12

sabato 1 aprile 2023 17:18

Probabilità sulla retta reale

Def: Dati $\Omega \neq \emptyset$, $\mathcal{C} \subseteq \mathcal{P}(\Omega)$, σ-algebra generata da $\mathcal{C}$

$\sigma(\mathcal{C})$ = la più piccola σ-algebra contenente $\mathcal{C}$ = $\bigcap_{\substack{\mathcal{F}\ \sigma\text{-algebra su } \Omega \\ \mathcal{F} \supseteq \mathcal{C}}} \mathcal{F}$

Oss: Intersezione di una famiglia di σ-algebre è σ-algebra (esercizio)

in particolare $\sigma(\mathcal{C}) = \bigcap_{\mathcal{F}\ \sigma\text{-alg} \supseteq \mathcal{C}} \mathcal{F}$ è una σ-algebra (la più piccola contenente $\mathcal{C}$)

Def. Dato $\mathcal{X} \neq \emptyset$ spazio metrico separabile (spesso $\mathcal{X} \subseteq \mathbb{R}^d$)

σ-algebra dei boreliani su $\mathcal{X}$: σ-algebra generata dagli aperti di $\mathcal{X}$

$\mathcal{B}(\mathcal{X}) = \sigma\{A \subseteq \mathcal{X} \mid A \text{ aperto}\}$

Proprietà dei boreliani:

- $\mathcal{B}(\mathcal{X})$ contiene tutti gli insiemi aperti, chiusi ed è generata dai chiusi
  tutti gli insiemi finiti o numerabili (unione numerabile di $\{x\}$, chiusi)
- $\mathcal{B}(\mathbb{R})$ contiene tutti gli intervalli (aperti, chiusi, semiaperti, semichiusi) e tutte le semirette
- $\mathcal{B}(\mathbb{R}) = \sigma\{(a,b] \mid a,b \in \mathbb{R}, a<b\}$
  $= \sigma\{(-\infty, x] \mid x \in \mathbb{R}\}$
- $\mathcal{B}(\mathbb{R}^d) = \sigma\{A_1 \times \dots \times A_d \mid A_1, \dots A_d \in \mathcal{B}(\mathbb{R})\}$
  $= \sigma\{(-\infty, x_1] \times \dots \times (-\infty, x_d] \mid x_1, \dots x_d \in \mathbb{R}\}$
- $\mathcal{B}(\mathbb{R}^d) \neq \mathcal{P}(\mathbb{R}^d)$
- Se $\mathcal{X}$ è metrico separabile e $\mathcal{Y} \subseteq \mathcal{X}$, $\mathcal{Y} \neq \emptyset$, ha la metrica indotta da $\mathcal{X}$,
  $\mathcal{B}(\mathcal{Y}) = \{A \cap \mathcal{Y} \mid A \in \mathcal{B}(\mathcal{X})\}$

(in pratica, $\mathcal{B}(\mathbb{R}^d)$ contiene tutti gli insiemi interessanti, ma non tutti i sottoinsiemi di $\mathbb{R}^d$)

Def.: Dato $(\Omega, \mathcal{F})$ spazio misurabile, una misura μ su $(\Omega, \mathcal{F})$ è una funzione

$\mu: \mathcal{F} \to [0, +\infty]$ con $\mu(\emptyset) = 0$ e μ σ-additiva

Oss.: P è prob su $(\Omega, \mathcal{F})$ $\Leftrightarrow$ P è misura su $(\Omega, \mathcal{F})$ e $P(\Omega) = 1$

Def: $N \in \mathcal{F}$ si dice μ-trascurabile: se $\mu(N) = 0$

Una proprietà q vale μ-q.o.: se $\exists N \in \mathcal{F}$ μ-trascurabile t.c. q vale su N^c.

Prop (misura di Lebesgue): Esiste un'unica misura m su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ che soddisfi $m([a,b]) = b-a$ $\forall a,b \in \mathbb{R}, a<b$. Essa è detta misura di Lebesgue (lunghezza)

Oss: $m|_{[0,1]} : \mathcal{B}([0,1]) \to [0,1]$ è misura di probabilità (probabilità uniforme su $[0,1]$)

Si può dimostrare, assumendo l'assioma della scelta, che $m|_{[0,1]}$ non si estende a $\mathcal{P}([0,1])$

Sia $(\Omega, \mathcal{F}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ e sia P probabilità su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

Def: Funzione di ripartizione (FdR) di P

$F: \mathbb{R} \to [0,1]$ definita da $F(x) = P((-\infty, x])$, $x \in \mathbb{R}$

Prop (proprietà di F):

a) F è non decrescente

b) F è continua a destra (cioè $\forall x \in \mathbb{R}$, $\lim_{y \downarrow x} F(y) = F(x)$)

c) $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to +\infty} F(x) = 1$

Dim:

a) $\forall x \leq y$, $F(x) = P((-\infty, x]) \overset{\text{monotonia}}{\leq} P((-\infty, y]) = F(y)$

b) Sia $x_n \downarrow x$. $(-\infty, x_n] \downarrow (-\infty, x]$ (cioè $(-\infty, x_n] \supseteq (-\infty, x_{n+1}]$, $\bigcap_n (-\infty, x_n] = (-\infty, x]$)

quindi per continuità di P per succ. decrescenti, $F(x_n) = P((-\infty, x_n]) \downarrow P((-\infty, x]) = F(x)$

c) Sia $x_n \downarrow -\infty$. $(-\infty, x_n] \downarrow \emptyset$, quindi per cont. per succ. decrescenti $F(x_n) \downarrow P(\emptyset) = 0$

Sia $x_n \uparrow +\infty$. $(-\infty, x_n] \uparrow \mathbb{R}$, quindi per cont. per succ. crescenti, $F(x_n) \uparrow P(\mathbb{R}) = 1$

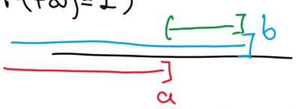
Prop (prob $\overset{1-1}{\longleftrightarrow}$ FdR):

Data una funzione $F: \mathbb{R} \to \mathbb{R}$ che soddisfa (a), (b), (c) della prop. precedente, esiste un'unica probabilità P su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ avente F per funzione di ripartizione.

In particolare, F determina univocamente P.

[La dimostrazione, che non vediamo, si basa sul fatto che le semirette $(-\infty, x]$, $x \in \mathbb{R}$, generano $\mathcal{B}(\mathbb{R})$ e sono chiuse per intersezione finita (π-sistema).

Calcolo di prob di intervalli con FdR: $(F(-\infty) = 0,\ F(+\infty) = 1)$



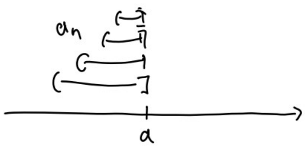
- $P((a,b]) = F(b) - F(a)$ $\forall -\infty \leq a < b \leq +\infty$

  infatti $(a,b] = (-\infty, b] \setminus (-\infty, a]$ e quindi $P((a,b]) = P((-\infty, b]) - P((-\infty, a])$

- $P\{a\} = F(a) - F(a-)$ $\forall a \in \mathbb{R}$, dove $F(a-) = \lim_{x \nearrow a} F(x)$

  infatti, se $a_n \nearrow a$, allora $(a_n, a] \downarrow \{a\}$ e quindi

  $P\{a\} = \lim_n P((a_n, a]) = \lim_n F(a) - F(a_n) = F(a) - F(a-)$



- $P([a,b]) = F(b) - F(a-)$ $\forall -\infty \leq a < b \leq +\infty$

  poiché $P([a,b]) = P((a,b]) + P\{a\}$

$P((a,b)) = F(b-) - F(a)$

$P([a,b)) = F(b-) - F(a-)$ } (esercizio)

Def: P si dice continua se F è continua, cioè se
$P\{a\}=0 \quad \forall a \in \mathbb{R}$

Esistono due grandi classi (non esaustive) di prob su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$:

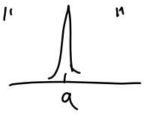
1) Prob. discrete: $\exists\, \Omega_0 \subseteq \mathbb{R}$ al più numerabile t.c. $P(\Omega_0)=1$ $\quad (\Omega_0=\{x_1, x_2, \dots\})$
senza perdita di generalità, possiamo prendere $\Omega_0 = R_P$ il range di P

Esempi: uniforme su $\{x_1, \dots x_N\}$, $B(p)$, $B(n,p)$, $G(p)$, $H(N, N_1, n)$, $P(\lambda)$, ..

Come abbiamo visto, assegnare P discreta equivale ad assegnare Ω_0 e $p: \Omega_0 \to \mathbb{R}$ densità discreta

Oss: Dato $a \in \mathbb{R}$, definiamo la misura di prob. δ_a (Delta di Dirac in a) su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$:

$$\delta_a(A) = \begin{cases} 1 & \text{se } a \in A \\ 0 & \text{se } a \notin A \end{cases}$$

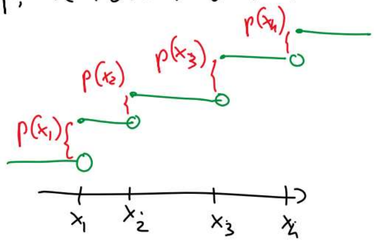
"  " massa concentrata in a

Allora P discreta si scrive come $P = \sum_i p(x_i)\,\delta_{x_i}$

Data P discreta, con range $R_P = \{x_1, x_2, \dots\}$ e densità discreta p, la FdR F di P è

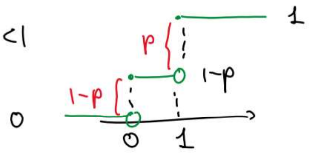
$$F(x) = \sum_{i,\, x_i \le x} p(x_i)$$

infatti $F(x) = P((-\infty, x]) = \sum_{x_i \in (-\infty, x]} p(x_i)$



Se R_P non ha punti di accumulazione, F è costante a tratti, con salti negli x_i e ampiezza dei salti $p(x_i)$

Esempio: $B(p)$, $0<p<1$



Esempio: dati $(p_r)_{r \in \mathbb{Q}}$ $\quad p_r > 0 \ \forall r$, $\sum_r p_r = 1$

$P = \sum_{r \in \mathbb{Q}} p_r \delta_r$ è prob discreta con FdR "non costante a tratti"

2) Probabilità assolutamente continue (rispetto alla misura di Lebesgue):

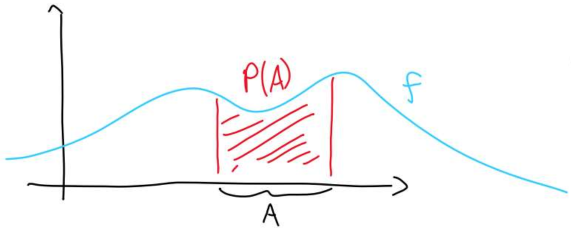
Def: P prob. su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ è assolutamente continua: se
$\exists\, f: \mathbb{R} \to \mathbb{R}$ boreliana (cioè t.c. $f^{-1}(A) \in \mathcal{B}(\mathbb{R})$ $\forall A \in \mathcal{B}(\mathbb{R})$) t.c.

$$P(A) = \int_A f(x)\,dx \quad \forall A \in \mathcal{B}(\mathbb{R})$$

Una tale funzione f si dice densità di P (rispetto alla misura di Lebesgue).

[Si assume implicitamente che f soddisfi $\int_{\mathbb{R}} |f(x)|\,dx < \infty$]

Oss: L'integrale $\int_A f(x)\,dx$ è inteso nel senso di Lebesgue, ma in molti esempi si riduce all'integrale di Riemann (eventualmente improprio)



Oss: La prob $P(A)$ è l'area sottesa da f su A.

Prop (caratterizzazione e unicità della densità)

a) Se P prob su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ha densità f, allora
g è densità per $P \iff f = g$ Lebesgue q.o.

b) Se P prob su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ha densità f, allora f soddisfa

(i) $f \geq 0$ Lebesgue q.o.

(ii) $\int_{\mathbb{R}} f(x)\,dx = 1$

c) Viceversa, data $f: \mathbb{R} \to \mathbb{R}$ boreliana che soddisfa (i) e (ii), allora esiste un'unica prob P su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ avente f per densità, ed è data da

$$P(A) = \int_A f(x)\,dx, \quad A \in \mathcal{B}(\mathbb{R})$$

Dim: [q.o. = Lebesgue q.o.]

Ricordiamo: • Se $h \geq 0$ q.o. allora • $\int h\,dx \geq 0$

• $\int h\,dx = 0 \iff h = 0$ q.o.

• teor conv. monotona: $h_n \geq 0$ q.o., $h_n \uparrow h \implies \int h_n\,dx \uparrow \int h\,dx$

a) Se f è densità e $g = f$ q.o., allora

$$\left|P(A) - \int_A g(x)\,dx\right| = \left|\int_A f\,dx - \int_A g\,dx\right| \leq \int_{\mathbb{R}} |f-g|\,dx = 0$$

Viceversa, se f e g sono densità per P

$$\int 1\!\!1_{\{f>g\}}(f-g)\,dx = \int_{\{f>g\}} f\,dx - \int_{\{f>g\}} g\,dx = P\{f>g\} - P\{f>g\} = 0,$$

Poiché $1\!\!1_{\{f>g\}}(f-g) \geq 0$, deve essere $1\!\!1_{\{f>g\}}(f-g) = 0$ q.o., cioè $f \leq g$ q.o.

Per simmetria, deve essere anche $f \geq g$ q.o. e quindi $f = g$ q.o.

b) Se f è densità allora

$$0 \geq \int_{\{f<0\}} f(x)\,dx = P\{f<0\} \geq 0, \text{ quindi } \int 1\!\!1_{\{f<0\}} f\,dx = 0 \text{ quindi } 1\!\!1_{\{f<0\}} f = 0 \text{ q.o. } \checkmark$$

cioè $f \geq 0$ q.o.

$$\int_{\mathbb{R}} f(x)\,dx = 1$$

c) Unicità: Se P ha densità f, deve essere $P(A) = \int_A f(x)\,dx \quad \forall A \in \mathcal{B}(\mathbb{R})$

Esistenza: mostriamo che $P(A) = \int_A f(x)\,dx$ verifica la def. di prob.

• $P(A) = \int_A f(x)\,dx \underset{f \geq 0 \text{ q.o.}}{\geq} 0 \qquad \forall A \in \mathcal{B}(\mathbb{R})$

- $P(\mathbb{R}) = \int_{\mathbb{R}} f(x)\,dx = 1$
- dati A_n a due a due disgiunti, $n \in \mathbb{N}^+$, abbiamo

$$\mathbb{1}_{\bigcup_n A_n} = \sum_n \mathbb{1}_{A_n} \quad \left(\text{infatti } \mathbb{1}_{\bigcup_n A_n}(\omega) = 1 \iff \omega \in \bigcup_n A_n \iff \omega \in A_{n_0} \text{ per uno e un solo } n_0 \iff \sum_n \mathbb{1}_{A_n}(\omega) = 1 \right)$$

quindi

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \int \mathbb{1}_{\bigcup_n A_n} f(x)\,dx = \int \sum_{n=1}^{\infty} \mathbb{1}_{A_n} f(x)\,dx$$

$$\xrightarrow{\text{teor. conv. monotona}} = \lim_{N\to\infty} \int \sum_{n=1}^{N} \mathbb{1}_{A_n} f(x)\,dx$$

$$\xrightarrow{\text{additività}} = \lim_{N\to\infty} \sum_{n=1}^{N} \int \mathbb{1}_{A_n} f(x)\,dx = \sum_{n=1}^{\infty} P(A_n)$$

FdR di prob. assolutamente continua: P prob su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, con FdR F

- Se P è assolutamente continua, con densità f, allora

$$F(x) = P((-\infty, x]) = \int_{-\infty}^{x} f(y)\,dy \quad \forall x \in \mathbb{R}$$

[Si può dimostrare anche il viceversa]

- Se P è assolutamente continua, allora è continua (poiché F è continua)

Il viceversa è falso

esempio: P prob associata a F scala di Cantor

- Se F è continua su $\mathbb{R}$ e C^1 a tratti (cioè F è C^1 eccetto che in un insieme di pt isolati), allora P è assolutamente continua, con densità

$$f(x) = F'(x) \quad (\text{dove esiste } F')$$

Dim:

Prendiamo $f = F'$ dove esiste, estendendo $f = 0$ dove F' non esiste. Notiamo che $f \geq 0$ poiché F è non-decrescente e, per il teor fondamentale del calcolo,

$$F(x) = \int_{-\infty}^{x} f(y)\,dy \quad \forall x \in \mathbb{R}$$

In particolare, $\int_{-\infty}^{+\infty} f(x)\,dx = \lim_{x\to+\infty} F(x) = 1$ e quindi f è una densità. Prendiamo Q prob su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ avente f per densità. Allora P e Q hanno la stessa FdR e quindi $P = Q$ e P ha densità $f = F'$.

Oss: Se P è assol. continua con densità f, allora $P\{f=0\} = \int_{\{f=0\}} f\,dx = 0$

Def: Data P prob su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, supporto (topologico) di P: $\operatorname{supp} P \subseteq \mathbb{R}$ chiuso t.c.

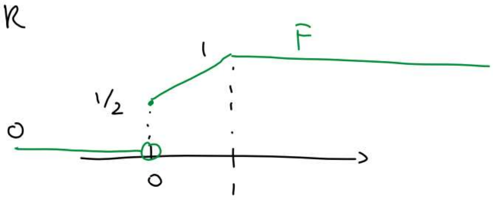
$$(\operatorname{supp} P)^c = \bigcup_{A \text{ aperto} \subseteq \mathbb{R},\ P(A)=0} A$$

Lemma: $P((\operatorname{supp} P)^c) = 0$

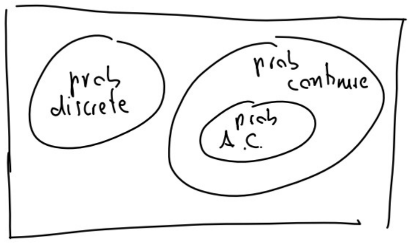
Esistono prob P su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ che non sono né discrete né continue

Es: P prob associata alla FdR

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2}(1+x) & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$



Prob reali

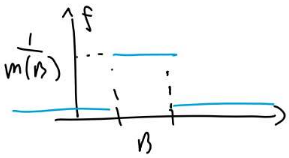


Esempi notevoli di probabilità assolutamente continue

P prob su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ con densità f, m misura di Lebesgue

- Dato $B \in \mathcal{B}(\mathbb{R})$ con $0 < m(B) < \infty$, distribuzione uniforme su B $(\mathcal{U}(B))$

$$f(x) = \frac{1}{m(B)} \mathbb{1}_B(x) = \begin{cases} \frac{1}{m(B)} & x \in B \\ 0 & x \notin B \end{cases}$$

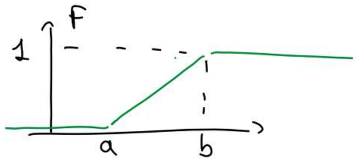


$\forall A \in \mathcal{B}(\mathbb{R})$, $P(A) = \int_A \frac{1}{m(B)} \mathbb{1}_B(x)\,dx = \frac{m(A \cap B)}{m(B)}$

In particolare, se $B = (a,b)$ intervallo, $a<b$,

$$f(x) = \frac{1}{b-a} \mathbb{1}_{(a,b)}(x)$$

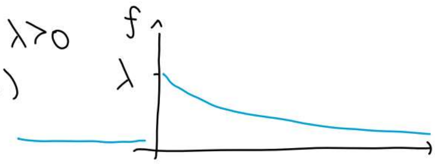
FdR: $F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x < b \\ 1 & x \ge b \end{cases}$



Significato: "scegliamo un pt a caso su (a,b), senza preferenze"

- Distribuzione esponenziale di parametro $\lambda > 0$ $(\mathcal{E}xp(\lambda))$

$$f(x) = \lambda e^{-\lambda x} \mathbb{1}_{(0,+\infty)}(x)$$



FdR $F(x) = (1 - e^{-\lambda x}) \mathbb{1}_{x \ge 0}$

Significato: "tempi di attesa senza memoria"

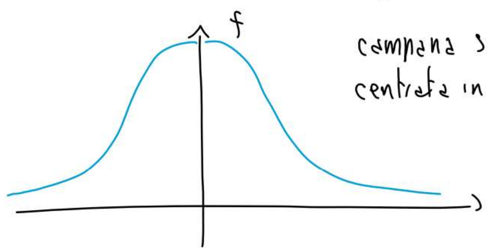
- Distribuzione Gamma di parametri $r > 0$ e $\lambda > 0$ $(\Gamma(r,\lambda))$

$$f(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} \mathbb{1}_{(0,+\infty)}(x)$$

dove $\Gamma(r) = \int_0^\infty x^{r-1} e^{-x}\,dx$ $\quad (\Gamma(r) = (r-1)!\ \forall r \in \mathbb{N}^+)$

- Distribuzione gaussiana, o normale, standard: $\mathcal{N}(0,1)$:

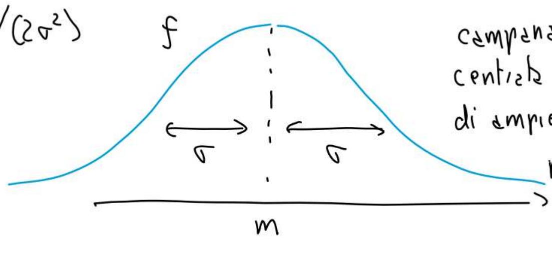
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$



campana simmetrica
centrata in $x=0$

- Distribuzione gaussiana, o normale, di media $m \in \mathbb{R}$ e varianza $\sigma^2 > 0$: $\mathcal{N}(m, \sigma^2)$:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-m)^2/(2\sigma^2)}$$



campana simmetrica
centrata in $x=m$
di ampiezza proporzionale a σ
pt di flesso in $m \pm \sigma$

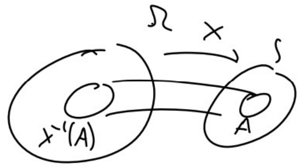
14

lunedì 3 aprile 2023 21:58

Variabili aleatorie generali

Def: Dati $(\Omega, \mathcal{F})$, $(S, \mathcal{S})$ spazi misurabili, una variabile aleatoria (v.a) da Ω in S è una funzione $X: \Omega \to S$ misurabile da $\mathcal{F}$ a $\mathcal{S}$, cioè t.c.

$X^{-1}(A) \in \mathcal{F} \quad \forall A \in \mathcal{S}$



Tipicamente: $S = \mathbb{R}$ e in questo caso si prende $\mathcal{S} = \mathcal{B}(\mathbb{R})$: v.a. reale

$S = \mathbb{R}^d$ " " " " " " $\mathcal{S} = \mathcal{B}(\mathbb{R}^d)$: vettore aleatorio

Ricordiamo: dato $A \subseteq S$, $X^{-1}(A) := \{X \in A\} := \{\omega \in \Omega \mid X(\omega) \in A\}$

X^{-1} commuta con le operazioni insiemistiche $\cup, \cap, {}^c$

Oss: Se $\mathcal{F} = \mathcal{P}(\Omega)$, ogni funzione $X: \Omega \to S$ è misurabile

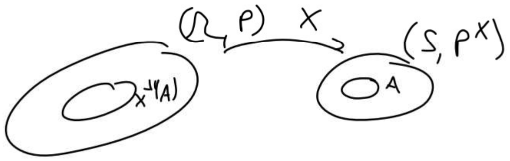
Fatto: Se $\mathcal{C} \subseteq \mathcal{P}(S)$ genera $\mathcal{S}$ (cioè $\mathcal{S} = \sigma(\mathcal{C})$) e $X^{-1}(C) \in \mathcal{F} \ \forall C \in \mathcal{C}$, allora X è misurabile

In particolare, se $(S, \mathcal{S}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, X è misurabile se, ad es, $(-\infty, x] \in \mathcal{F} \ \forall x \in \mathbb{R}$

Def. Dati $(\Omega, \mathcal{F}, P)$ spazio di probabilità, $(S, \mathcal{S})$ spazio misurabile, $X: \Omega \to S$ v.a. legge (o distribuzione) di X su $(S, \mathcal{S})$ (o misura immagine di P tramite X)

P^X (o $X_\# P$ o $P \circ X^{-1}$ o $X(P)$): misura di probabilità su $(S, \mathcal{S})$ definita da

$$P^X(A) = P\{X \in A\} = P(X^{-1}(A)), \quad A \in \mathcal{S}$$



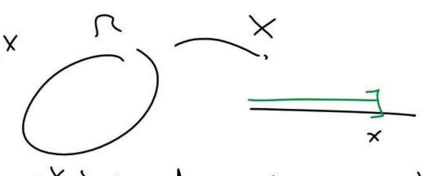
Lemma: P^X è una probabilità su $(S, \mathcal{S})$

Dim come nel caso discreto.

Per v.a. X reale,

- chiamiamo funzione di ripartizione F^X di X la FdR di P^X

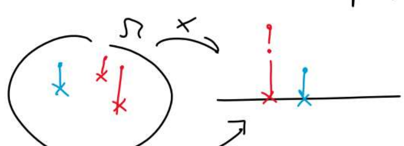
$F^X(x) = P^X((-\infty, x]) = P\{X \le x\} \quad \forall x \in \mathbb{R}$



- diciamo che X è discreta, con densità discreta p^X, se P^X è discreta con dens. discr. p^X:

$p^X(x) = P^X\{x\} = P\{X = x\}$,

$P\{X \in A\} = P^X(A) = \sum_{x \in A} p^X(x) \left(= \sum_{x \in A \cap R_{p^X}} p^X(x)\right) \quad \forall A \in \mathcal{B}(\mathbb{R})$



Oss: non è detto che Ω sia discreta

- diciamo che X è continua se P^X è continua, cioè $P\{X = x\} = P^X\{x\} = 0 \quad \forall x \in \mathbb{R}$

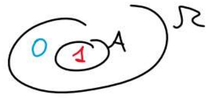
· diciamo che X è assolutamente continua con densità f, se P^X è assol. cont. con dens. f

$$P\{X\in A\} = P^X(A) = \int_A f(x)\,dx \qquad \forall A\in\mathcal{B}(\mathbb{R})$$

Analogamente per X v.a. a valori in $\mathbb{R}^d$

Oss: Dato $A\subseteq\Omega$, $\mathbb{1}_A:\Omega\to\mathbb{R}$, $\mathbb{1}_A(\omega)=\begin{cases}1 & \text{se } \omega\in A\\ 0 & \text{se } \omega\notin A\end{cases}$,

$A\in\mathcal{F} \iff \mathbb{1}_A$ è misurabile (da $\mathcal{F}$ a $\mathcal{B}(\mathbb{R})$) (esercizio)



Costruzione canonica: dato $(S,\mathcal{S})$ sp misurabile e Q probabilità su $(S,\mathcal{S})$, prendiamo $(\Omega,\mathcal{F})=(S,\mathcal{S})$, $P=Q$ prob su $(\Omega,\mathcal{F})$ e $X:\Omega\to S$, $X=id$ allora X è v.a. di legge $P^X=Q$

Composizione di v.a.: dati $(\Omega,\mathcal{F})$, $(S,\mathcal{S})$, $(S',\mathcal{S}')$ sp. misurabili, $X:\Omega\to S$ v.a., $\varphi:S\to S'$ misurabile, allora $\varphi\circ X:\Omega\to S'$ è v.a. (cioè misurabile):

infatti $\forall A\in\mathcal{S}'$, $\{\varphi(X)\in A\} = \{X\in \underbrace{\varphi^{-1}(A)}_{\in\mathcal{S} \text{ per mis. di } \varphi}\}\in\mathcal{F}$ per mis. di X

Oss: Se X è discreta, con range R_{P_X}, allora $\varphi(X)$ è discreta, con range $\varphi(R_{P_X})$

infatti · $P\{\varphi(X)\in\varphi(R_{P_X})\}\geq P\{X\in\varphi(R_{P_X})\}=1$

· $\#\varphi(R_{P_X})\leq \# R_{P_X}\leq \#\mathbb{N}$

· $\forall y=\varphi(x)$, $x\in R_{P_X}$, $P\{\varphi(X)=y\}\geq P\{X=x\}>0$

Invece se $X:\Omega\to\mathbb{R}$ è continua, non è detto che $\varphi(X)$ sia continua.

es: $\varphi\equiv 0$.

I concetti e le proprietà di:

· uguaglianza q.c. per v.a. ($X=Y$ q.c. se $P\{X=Y\}=1$)

· uguaglianza in legge per v.a. ($X\overset{(d)}{=}Y$ se $P^X=P^Y$)

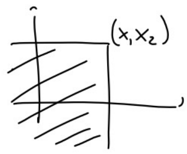
si estendono senza difficoltà al caso generale.

Distribuzioni congiunte e indipendenza per v.a. generali

Probabilità su $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$:

- Prop: Esiste un'unica misura m su $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ t.c.
  $m([a_1,b_1]\times\ldots\times[a_d,b_d]) = (b_1-a_1)\cdot\ldots\cdot(b_d-a_d) \quad \forall a_1,b_1,\ldots a_d,b_d$, con $a_i<b_i \ \forall i$
  Essa è detta misura di Lebesgue d-dimensionale (volume)

- Def: Data P prob su $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, funzione di ripartizione di P:
  $F: \mathbb{R}^d \to \mathbb{R}$ definita da $F(x_1,\ldots x_d) = P((-\infty, x_1]\times\ldots\times(-\infty, x_d])$



Prop:
- F è non decrescente rispetto a ogni variabile
  $F(x_1,\ldots x_{i-1}, x, x_{i+1},\ldots x_d) \le F(x_1,\ldots x_{i-1}, y, x_{i+1},\ldots x_d) \quad \forall x\le y, \ \forall i$
- F è continua a destra rispetto a ogni variabile
  $\lim_{y\downarrow x} F(x_1,\ldots x_{i-1}, y, x_{i+1},\ldots x_d) = F(x_1,\ldots x_{i-1}, x, x_{i+1},\ldots x_d) \quad \forall x, \ \forall i$
- $\forall i, \ \lim_{x\to-\infty} F(x_1,\ldots x_{i-1}, x, x_{i+1},\ldots x_d) = 0$
  $\lim_{x_1\to+\infty}\ldots\lim_{x_d\to+\infty} F(x_1,\ldots x_d) = 1$

Prop: F determina univocamente P (cioè se P e Q hanno la stessa FdR, allora $P=Q$)

Dim:
Se P e Q hanno la stessa FdR, allora P e Q coincidono su
$\mathcal{C} = \{(-\infty,x_1]\times\ldots\times(-\infty,x_d] \mid (x_1,\ldots x_d)\in\mathbb{R}^d\}$. Poiché $\mathcal{C}$ è chiusa per inters. finita e $\sigma(\mathcal{C}) = \mathcal{B}(\mathbb{R}^d)$,
per il lemma di Dynkin, abbiamo $P=Q$ su $\mathcal{B}(\mathbb{R}^d)$.

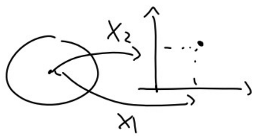
- P prob su $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ è detta discreta: se $\exists\, \Omega_0 \subseteq \mathbb{R}^d$ al più numerabile t.c. $P(\Omega_0)=1$
- P prob su $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ è detta assolutamente continua: se $\exists f: \mathbb{R}^d\to\mathbb{R}$ boreliana t.c.
  $P(A) = \int_A f(x)\,dx \quad \forall A\in\mathcal{B}(\mathbb{R}^d)$ (integrale di Lebesgue)
  f si dice densità di P.

Prop: a) Se P ha densità f, allora P ha densità $g \Leftrightarrow f=g$ Lebesgue q.o.
b) Se P ha densità f, allora
  i) $f\ge 0$ Lebesgue-q.o.
  ii) $\int_{\mathbb{R}^d} f(x)\,dx = 1$
c) Viceversa, se $f:\mathbb{R}^d\to\mathbb{R}$ boreliana soddisfa (i) e (ii), allora esiste un'unica prob P su $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ che ha densità f

FdR: Se P è assolutamente continua con densità f, allora P ha FdR
$F(x_1,\ldots x_d) = \int_{-\infty}^{x_1}\ldots\int_{-\infty}^{x_d} f(y_1,\ldots y_d)\,dy_1\ldots dy_d \quad \forall (x_1,\ldots x_d)\in\mathbb{R}^d \quad (\Delta)$

Viceversa, se P ha F che soddisfa (Δ) per qualche densità f (cioè $f: \mathbb{R}^d \to \mathbb{R}$ boreliana che soddisfa (i) e (ii)), allora P ha densità f.
infatti, detta Q la prob con densità f, Q e P hanno le stesse FdR e quindi $P=Q$.

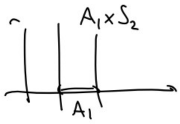
"Se F è sufficientemente regolare, allora $f(x_1 \dots x_d) = \partial_{x_1} \dots \partial_{x_d} F(x_1 \dots x_d)$"
In generale, si può usare la formula $f = \partial_{x_1} \dots \partial_{x_d} F$ per indovinare f e poi verificare che f è densità e che $F(x_1 \dots x_d) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_d} f(y_1 \dots y_d)\, dy_1 \dots dy_d$ (e quindi P ha densità f)



Date $X_1: \Omega \to S_1, \dots X_n: \Omega \to S_n$, $A_1 \subseteq S_1, \dots A_n \subseteq S_n$,
chiamiamo $(X_1, \dots X_n): \Omega \to S_1 \times \dots \times S_n$, $(X_1, \dots X_n)(\omega) = (X_1(\omega), \dots X_n(\omega))$
vettore aleatorio

Ricordiamo

$\{(X_1, \dots X_n) \in A_1 \times \dots \times A_n\} = \{X_1 \in A_1 \dots X_n \in A_n\} = \bigcap_{i=1}^n \{X_i \in A_i\}$
$\{X_i \in A_i\} = \{(X_1 \dots X_n) \in S_1 \times \dots \times S_{i-1} \times A \times S_{i+1} \times \dots \times S_n\}$



Def: Dati $(S_1, \mathcal{S}_1), \dots (S_n, \mathcal{S}_n)$ spazi misurabili, si definisce σ-algebra prodotto $\mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n$ su $S_1 \times \dots \times S_n$:
$$\mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n = \sigma\{\overbrace{A_1 \times \dots \times A_n}^{\text{insiemi cilindrici}} \mid A_1 \in \mathcal{S}_1, \dots A_n \in \mathcal{S}_n\}$$

Oss: Dato $(\Omega, \mathcal{F})$ sp. mis., $X_i: \Omega \to S_i$, $i = 1, \dots n$,
X_i è misurabile da $\mathcal{F}$ a $\mathcal{S}_i$, $\forall i = 1 \dots n$ $\iff$ $(X_1, \dots X_n): \Omega \to S_1 \times \dots \times S_n$ è misurabile da $\mathcal{F}$ a $\mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n$

In particolare, $(X_1, \dots X_n): \Omega \to \mathbb{R}^n$ è misurabile (in $\mathcal{B}(\mathbb{R}^n)$) $\iff$ X_i è mis. (in $\mathcal{B}(\mathbb{R})$) $\forall i$.

Dim: $\Rightarrow$) $\forall A_1 \in \mathcal{S}_1, \dots A_n \in \mathcal{S}_n$,
$\{(X_1, \dots X_n) \in A_1 \times \dots \times A_n\} = \underbrace{\{X_1 \in A_1\}}_{\in \mathcal{F}} \cap \dots \cap \underbrace{\{X_n \in A_n\}}_{\in \mathcal{F}} \in \mathcal{F}$ poiché X_i sono mis. da $\mathcal{F}$ a $\mathcal{S}_i$
poiché gli insiemi cilindrici generano $\mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n$, $(X_1 \dots X_n)$ è mis da $\mathcal{F}$ a $\mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n$
$\Leftarrow$): $\forall i = 1, \dots n$, $\forall A_i \in \mathcal{S}_i$, $\{X_i \in A_i\} = \{(X_1 \dots X_n) \in S_1 \times \dots \times S_{i-1} \times A_i \times S_{i+1} \times \dots \times S_n\} \in \mathcal{F}$
poiché $(X_1 \dots X_n)$ è mis da $\mathcal{F}$ a $\mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n$

Def: Data $X = (X_1, \dots X_n): \Omega \to S_1 \times \dots \times S_n$ v.a (misurabile da $\mathcal{F}$ a $\mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n$)
legge congiunta di $(X_1, \dots X_n)$: la legge P^X di X su $(S_1 \times \dots \times S_n, \mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n)$
leggi marginali: le leggi $P^{X_1}, \dots P^{X_n}$ rispettivamente di $X_1, \dots X_n$ su $(S_1, \mathcal{S}_1), \dots (S_n, \mathcal{S}_n)$
e più in generale la legge di $(X_j)_{j \in J}$ su $(\times_{j \in J} S_j, \otimes_{j \in J} \mathcal{S}_j)$, per $J \subsetneq I = \{1, \dots n\}$

Oss: $\forall A_i \in \mathcal{S}_i$, $P\{X_1 \in A_1, \dots X_n \in A_n\} = P\{(X_1 \dots X_n) \in A_1 \times \dots \times A_n\} = P_{(X_1 \dots X_n)}(A_1 \times \dots \times A_n)$

Per $X=(X_1, \dots X_n)$ v.a a valori in $\mathbb{R}^n$,

- chiamiamo FdR di X la FdR di P_X: $F_X(x_1, \dots x_n) = P\{X_1 \le x_1, \dots X_n \le x_n\}$
- diciamo che X è discreta con densità discreta p_X: se P_X è discreta con dens. discreta p_X
- diciamo che X è assolutamente continua con densità $f: \mathbb{R}^n \to \mathbb{R}$: se P_X è assol. cont. con densità f:

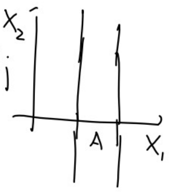
$$P\{X \in A\} = \int_A f(x_1, \dots x_n)\, dx_1 \dots dx_n \quad \forall A \in \mathcal{B}(\mathbb{R}^n)$$

Prop: La legge congiunta determina le leggi marginali: precisamente, $\forall i = 1, \dots n$

$$P_{X_i}(A) = P\{X_i \in A\} = P\{(X_1, \dots X_n) \in S_1 \times \dots \times S_{i-1} \times A_i \times S_{i+1} \times \dots \times S_n\} \quad \forall A \in \mathcal{J}_i$$

e più in generale, $\forall J \subsetneq I$,

$$P_{(X_j)_{j\in J}}(A) = P\{(X_j)_{j\in J} \in A, (X_i)_{i \in I \setminus J} \in \times_{i\in I\setminus J} S_i\} \quad \forall A \in \otimes_{j\in J} \mathcal{J}_j$$



Dim: (come nel caso discreto)

La prima formula segue dall'osservazione

$$\{X_i \in A\} = \{(X_1, \dots X_n) \in S_1 \times \dots \times S_{i-1} \times A_i \times S_{i+1} \times \dots \times S_n\}$$

e analogamente la seconda da $\{(X_j)_{j\in J} \in A\} = \{(X_j)_{j\in J} \in A, (X_i)_{i\in I\setminus J} \in \times_{i\in I\setminus J} S_i\}$

Caso discreto:

$X = (X_1, \dots X_n): \Omega \to \mathbb{R}^n$ è discreto $\iff$ $X_i: \Omega \to S_i$ è discreta $\forall i = 1, \dots n$

Infatti: • se X è discreto con range R_{P_X}, allora detta π_i la proiezione canonica, $X_i = \pi_i(X)$ è discreto con range $\pi_i(R_{P_X})$

• se X_i sono discreti con range $R_{P_{X_i}}$, allora $X \in R_{P_{X_1}} \times \dots \times R_{P_{X_n}}$ q.c. e $\# R_{P_{X_1}} \times \dots \times R_{P_{X_n}} \le \# \mathbb{N}$, quindi $R_{P_{(X_1 \dots X_n)}} \subseteq R_{P_{X_1}} \times \dots \times R_{P_{X_n}}$

Dal caso discreto abbiamo:

Prop: Detta $p_{(X_1 \dots X_n)}$ la densità discreta congiunta di $(X_1, \dots X_n)$ e p_{X_i} le densità discrete marginali, vale

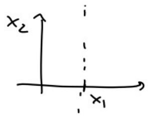
$$p_{X_i}(x) = \sum_{x_j \in \mathbb{R},\, j \ne i} p_{(X_1 \dots X_n)}(x_1, \dots x_{i-1}, x, x_{i+1}, \dots x_n) \qquad x \in \mathbb{R}$$

dove $\displaystyle\sum_{x_j \in \mathbb{R},\, j\ne i} = \sum_{(x_1, \dots x_{i-1}, x, x_{i+1}, \dots x_n) \in R_{P_{(X_1 \dots X_n)}}}$

Caso assolutamente continuo:

Prop: Se $X = (X_1, \dots X_n): \Omega \to \mathbb{R}^n$ è assolutamente continuo con densità congiunta $f: \mathbb{R}^n \to \mathbb{R}$, allora, $\forall i = 1, \dots n$, X_i è assolutamente continua con densità marginale $f_i: \mathbb{R} \to \mathbb{R}$ data da

$$f_i(x) = \int_{\mathbb{R}^{n-1}} f(x_1, \dots x_{i-1}, x, x_{i+1}, \dots x_n)\, dx_1 \dots dx_{i-1}\, dx_{i+1} \dots dx_n, \quad x \in \mathbb{R}$$



15/4

Dim: $i=1$ per semplicità di notazione

$$P\{X_1 \in A\} = P\{X_1 \in A, X_2 \in \mathbb{R}, \dots X_n \in \mathbb{R}\}$$
$$= \int_{A \times \mathbb{R}^{n-1}} f(x_1, x_2, \dots x_n)\, dx_1 \dots dx_n$$

Fubini -Tonelli
$$= \int_A \left(\int_{\mathbb{R}^{n-1}} f(x_1, x_2, \dots x_n)\, dx_2 \dots dx_n \right) dx_1$$
$$= \int_A f_1(x_1)\, dx_1 \qquad \forall A \in \mathcal{B}(\mathbb{R})$$

quindi f_1 è densità per X_1

Fubini -Tonelli:

Se $f: \mathbb{R}^2 \to \mathbb{R}$ è boreliana ≥ 0 o integrabile, allora

- $f(\cdot, x_2)$, $x_2 \mapsto \int f(x_1, x_2)\, dx_1$ sono boreliane, e analogamente scambiando x_1 e x_2
- $\iint f(x_1,x_2)\, dx_1 dx_2 = \int \left(\int f(x_1,x_2)\, dx_1 \right) dx_2 = \int \left(\int f(x_1,x_2)\, dx_2 \right) dx_1$

Analogamente per $f: \mathbb{R}^n \to \mathbb{R}$.

Inoltre $1\!\!1_{A_1 \times A_2}(x_1,x_2) = 1\!\!1_{A_1}(x_1) 1\!\!1_{A_2}(x_2)$ e quindi, se $f_1: \mathbb{R} \to \mathbb{R}$, $f_2: \mathbb{R} \to \mathbb{R}$ sono boreliane integrabili o ≥ 0

$$\iint_{A_1 \times A_2} f_1(x_1) f_2(x_2)\, dx_1 dx_2 = \iint 1\!\!1_{A_1 \times A_2}(x_1,x_2) f_1(x_1) f_2(x_2)\, dx_1 dx_2 =$$
$$= \iint 1\!\!1_{A_1}(x_1) 1\!\!1_{A_2}(x_2) f_1(x_1) f_2(x_2)\, dx_1 dx_2 = \int_{A_1} f_1(x_1)\, dx_1 \cdot \int_{A_2} f_2(x_2)\, dx_2$$

Analogamente per $A_1 \times \dots \times A_n$, $f_1, \dots f_n$.

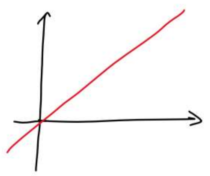
Oss: Se X_1, X_2 v.a reali sono assolutamente continue, non è detto che $(X_1, X_2): \Omega \to \mathbb{R}^2$ sia assolutamente continuo

Esempio: X_1 assolutamente continua, $X_2 = X_1$, (X_1, X_1) non è assolutamente continuo

idea: (X_1, X_1) è concentrato su $\{(x,x) | x \in \mathbb{R}\}$, che ha misura di Lebesgue 2D nulla

per assurdo: (X_1, X_1) ha densità f, allora

$$0 = \iint_{\{x_1 = x_2\}} f(x_1,x_2)\, dx_1 dx_2 = P\{X_1 = X_2\} = 1 \quad \lightning$$
$$\{x_1 = x_2\} = \{(x_1,x_1) | x_1 \in \mathbb{R}\}$$



Indipendenza di v.a: ($(\Omega, \mathcal{F}, P)$ sp di prob., $(S_i, \mathcal{J}_i)$ sp misurabili, $i=1, \dots n$)

Def: Date $X_1: \Omega \to S_1, \dots X_n: \Omega \to S_n$ v.a., $(X_1, \dots X_n)$ si dice famiglia di v.a. indipendenti: se

$$P\{X_1 \in A_1, \dots X_n \in A_n\} = P\{X_1 \in A_1\} \cdot \dots \cdot P\{X_n \in A_n\} \quad \forall A_1 \in \mathcal{J}_1, \dots A_n \in \mathcal{J}_n$$

equivalentemente $P_{(X_1, \dots X_n)}(A_1 \times \dots \times A_n) = P_{X_1}(A_1) \cdot \dots \cdot P_{X_n}(A_n) \quad \forall A_1 \in \mathcal{J}_1, \dots A_n \in \mathcal{J}_n$

Oss: Come nel caso discreto, $(X_i)_{i=1}^n$ sono indip $\iff \{X_i \in A_i\}_{i=1}^n$ sono indip. $\forall A_i \in \mathcal{J}_i, i=1, \dots n$

Oss: Come nel caso discreto, l'indipendenza è una proprietà della legge congiunta di $(X_1, \dots X_n)$

Def: Sia $(S_i, \mathcal{S}_i)$, $i \in I$, una famiglia di spazi misurabili (con I possibilmente infinito) sia $X_i : \Omega \to S_i$, $i \in I$, una famiglia di v.a.. $(X_i)_{i \in I}$ si dice famiglia di v.a. indipendenti: se, $\forall J \subseteq I$ finito, $(X_j)_{j \in J}$ è famiglia di v.a. indipendenti.

Oss: Se $X_1, \dots X_n$ sono indipendenti, allora le leggi marginali determinano la legge congiunta: dette $Y_1, \dots Y_n$ indipendenti con le stesse leggi marginali di $X_1, \dots X_n$ ($P_{X_i} = P_{Y_i}$ $\forall i$), allora

$$P_{(X_1 \dots X_n)}(A_1 \times \dots \times A_n) \underset{\text{indip}}{=} P_{X_1}(A_1) \cdot \ldots \cdot P_{X_n}(A_n) \underset{\text{stesse marginali}}{=} P_{Y_1}(A_1) \cdot \ldots \cdot P_{Y_n}(A_n) \underset{\text{indip.}}{=} P_{(Y_1 \dots Y_n)}(A_1 \times \dots \times A_n) \quad \forall A_i \in \mathcal{S}_i, \ i = 1, \dots n$$

Poiché gli insiemi cilindrici $A_1 \times \dots \times A_n$ sono chiusi per intersezione e generano $\mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n$, allora $P_{(X_1 \dots X_n)} = P_{(Y_1 \dots Y_n)}$ su $\mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n$ per lemma di Dynkin

Prop: Siano $X_i : \Omega \to \mathbb{R}$, $i = 1, \dots n$, v.a. reali, siano F la FdR di $(X_1 \dots X_n)$ e, per $i = 1, \dots n$, F_i la FdR di X_i. Allora

$(X_1 \dots X_n)$ è famiglia di v.a. indipendenti se e solo se

$$F(x_1 \dots x_n) = F_1(x_1) \cdot \ldots \cdot F_n(x_n) \quad \forall (x_1 \dots x_n) \in \mathbb{R}^n$$

Dim:

$\Rightarrow$): Se $X_1 \dots X_n$ sono indipendenti, allora

$$F(x_1 \dots x_n) = P\{(X_1, \dots X_n) \in (-\infty, x_1] \times \dots \times (-\infty, x_n]\} = P\{X_1 \le x_1, \dots X_n \le x_n\}$$

$$\text{indip} \to = P\{X_1 \le x_1\} \cdot \ldots \cdot P\{X_n \le x_n\} = F_1(x_1) \cdot \ldots \cdot F(x_n)$$

$\Leftarrow$): Siano $Y_1 \dots Y_n$ v.a. reali con $Y_i \overset{(d)}{=} X_i$ $\forall i = 1, \dots n$ e Y_i indipendenti

[Si può dimostrare che esistono tali Y_i su un opportuno spazio, si veda teor successivo]

Poiché $X_i \overset{(d)}{=} Y_i$, Y_i ha F_i come FdR, e per indipendenza la FdR di $(Y_1 \dots Y_n)$ è

$$F_{(Y_1 \dots Y_n)}(x_1 \dots x_n) = F_1(x_1) \cdot \ldots \cdot F(x_n) = F(x_1, \dots x_n) \ \forall (x_1 \dots x_n) \in \mathbb{R}^n$$

Quindi $(X_1 \dots X_n)$ e $(Y_1 \dots Y_n)$ hanno la stessa FdR, quindi hanno la stessa legge.

In particolare, $X_1 \dots X_n$ sono indipendenti

Prop: Sia $(X_1 \dots X_n) : \Omega \to \mathbb{R}^n$ un vettore aleatorio discreto ($\Leftrightarrow X_i$ sono v.a. discrete).

$(X_1, \dots X_n)$ è famiglia di v.a. indipendenti se e solo se

$$p_{(X_1 \dots X_n)}(x_1 \dots x_n) = p_{X_1}(x_1) \cdot \ldots \cdot p_{X_n}(x_n) \quad \forall (x_1 \dots x_n) \in \mathbb{R}^n$$

(equivalentemente $\forall (x_1 \dots x_n) \in R_{p_{(X_1 \dots X_n)}}$)

(come visto nel caso discreto)

Prop: Sia $(X_1, \dots X_n) : \Omega \to \mathbb{R}^n$ un vettore aleatorio.

a) Se $(X_1 \dots X_n)$ è assolutamente continuo e la sua densità congiunta f soddisfa

$$f(x_1 \dots x_n) = f_1(x_1) \cdot \ldots \cdot f_n(x_n) \quad \forall (x_1 \dots x_n) \in \mathbb{R}^n \qquad (*)$$

con $f_1, \dots f_n$ densità marginali, allora $(X_1, \dots X_n)$ è famiglia di v.a. indipendenti.

b) Viceversa, se $X_1, \dots X_n$ sono v.a. reali assolutamente continue e indipendenti, allora $(X_1 \dots X_n)$ è assolutamente continuo e la sua densità congiunta f soddisfa $(*)$.

Dim:

a) $\forall A_1, \dots A_n \in \mathcal{B}(\mathbb{R}^n)$

$$P\{X_1 \in A_1, \dots X_n \in A_n\} = \int_{A_1 \times \dots \times A_n} f(x_1 \dots x_n)\, dx_1 \dots dx_n$$

$$\overset{\text{ipotesi}}{=} \int_{A_1 \times \dots \times A_n} f(x_1) \dots f(x_n)\, dx_1 \dots dx_n$$

$$\overset{\text{Fubini Tonelli}}{=} \int_{A_1} \int_{A_2} \dots \int_{A_n} f(x_1) f(x_2) \dots f(x_n)\, dx_n \dots dx_2\, dx_1$$

$$= \int_{A_1} f(x_1)\, dx_1 \cdot \dots \cdot \int_{A_n} f(x_n)\, dx_n = P\{X_1 \in A_1\} \cdot \dots \cdot P\{X_n \in A_n\}$$

quindi $X_1 \dots X_n$ sono indipendenti

b) Dimostrazione usando criterio indipendenza per FdR

$$F(x_1, \dots x_n) \underset{\text{indip.}}{=} F(x_1) \cdot \dots \cdot F(x_n) =$$

$$= \int_{-\infty}^{x_1} f(y_1)\, dy_1 \cdot \dots \cdot \int_{-\infty}^{x_n} f(y_n)\, dy_n$$

$$\overset{\text{Fubini Tonelli}}{=} \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(y_1) \dots f(y_n)\, dy_1 \dots dy_n \qquad \forall (x_1 \dots x_n) \in \mathbb{R}^n$$

Esercizio: $f(x_1) \dots f(x_n)$ è una densità su $\mathbb{R}^n$

Quindi $P_{(X_1 \dots X_n)}$ ha densità $f(x_1) \dots f(x_n)$.

Dim con lemma di Dynkin:

Dobbiamo verificare che

$$P_{(X_1 \dots X_n)}(A) = \int_A f_1(x_1) \dots f_n(x_n)\, dx_1 \dots dx_n \qquad \forall A \in \mathcal{B}(\mathbb{R}^n)$$
$$= Q(A)$$

dove Q è la prob. di densità $f_1(x_1) \dots f_n(x_n)$ (esercizio: $f_1(x_1) \dots f_n(x_n)$ è una densità)

Poiché $\mathcal{C} = \{A_1 \times \dots \times A_n \mid A_1, \dots A_n \in \mathcal{B}(\mathbb{R})\}$ è chiusa per intersezione finita e genera $\mathcal{B}(\mathbb{R}^n)$, basta verificare $P_{(X_1 \dots X_n)}(A_1 \times \dots \times A_n) = Q(A_1 \times \dots \times A_n)$ $\forall A_i \in \mathcal{B}(\mathbb{R})$

$$P_{(X_1 \dots X_n)}(A_1 \times \dots \times A_n) = P\{X_1 \in A_1, \dots X_n \in A_n\}$$

$$\overset{\text{indip}}{=} P\{X_1 \in A_1\} \cdot \dots \cdot P\{X_n \in A_n\}$$

$$= \int_{A_1} f_1(x_1)\, dx_1 \cdot \dots \cdot \int_{A_n} f_n(x_n)\, dx_n$$

$$\overset{\text{Fubini Tonelli}}{=} \int_{A_1 \times \dots \times A_n} f(x_1) \dots f(x_n)\, dx_1 \dots dx_n$$

Lemma (stabilità per composizione): Siano $(\Omega, \mathcal{F}, P)$ sp di prob, $(S_i, \mathcal{S}_i), (S_i', \mathcal{S}_i')$ spazi misurabili. Se $X_i : \Omega \to S_i$, $i \in I$, sono v.a. indipendenti e $\varphi_i : S_i \to S_i'$ sono misurabili, allora $\varphi_i(X_i) : \Omega \to S_i$, $i \in I$, sono v.a. indipendenti

Dim come nel caso discreto

Lemma (gruppi disgiunti di v.a. indip sono indip.)

Se $X_i : \Omega \to S_i$, $i \in I$, sono v.a. indip. e $J_1, \dots J_m \in I$ sono disgiunti e finiti, allora $(X_j)_{j \in J_1} : \Omega \to \times_{j \in J_1} S_j, \dots (X_j)_{j \in J_2} : \Omega \to \times_{j \in J_m} S_j$ sono v.a. indipendenti

15/7

La dimostrazione, che non vediamo, usa il lemma di Dynkin.

Esercizio: dimostrare il lemma per I finito, X_i assolutamente continue

Oss. (costruzione canonica): Data Q legge su $(S_1 \times \dots \times S_n, \mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n)$, la costruzione canonica fornisce $(\Omega, \mathcal{F}, P)$, $X = (X_1, \dots, X_n)$ v.a. a valori in $S_1 \times \dots \times S_n$ di legge Q

Teor: Dati $(S_1, \mathcal{S}_1, Q_1), \dots (S_n, \mathcal{S}_n, Q_n)$ spazi di prob, esiste un'unica probabilità (detta prob prodotto) $Q (= Q_1 \otimes \dots \otimes Q_n)$ su $(S_1 \times \dots \times S_n, \mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_n)$ che sia la legge di $(X_1, \dots, X_n)$, con $X_1, \dots X_n$ indipendenti e di leggi marginali $Q_1, \dots Q_n$ rispettivamente.

16

martedì 11 aprile 2023 17:43

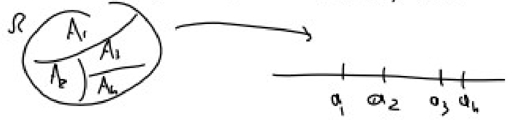
Valore atteso, momenti e varianza

$(\Omega, \mathcal{F}, P)$ spazio di prob..

Data $X: \Omega \to \mathbb{R}$ v.a reale, vogliamo definire il valore atteso $E[X]$ di X in modo t.c.

- per X discreta, $E[X]$ coincida con la definizione/caratterizzazione vista nel caso discreto
$$E[X] = \sum_i x_i p_X(x_i) = \sum_i x_i P\{X = x_i\}$$
- E abbia "buone proprietà" di passaggio al limite: "$X_n \to X$ in un senso opportuno $\Rightarrow E[X_n] \to E[X]$"

Def: $X: \Omega \to \mathbb{R}$ v.a. si dice semplice se assume un numero finito di valori, cioè
$$X = \sum_{i=1}^m a_i \mathbb{1}_{A_i}$$



con $a_1 \dots a_m \in \mathbb{R}$, $A_1, \dots A_m \in \mathcal{F}$, $m \in \mathbb{N}^+$.

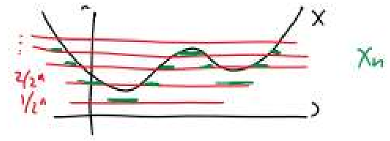
Data X semplice, $X = \sum_{i=1}^m a_i \mathbb{1}_{A_i}$, definiamo
$$E[X] = \int_\Omega X \, dP := \sum_{i=1}^m a_i P(A_i)$$

Oss (esercizio):
- $E[X]$ non dipende dalla scelta degli A_i e degli a_i
- In particolare, possiamo prendere gli a_i come tutti distinti, in questo caso $A_i = \{X = a_i\}$, e troviamo $E[X] = \sum_{i=1}^m a_i P\{X = a_i\}$, come visto nel caso discreto

Lemma: Data $X: \Omega \to \mathbb{R}$ v.a ≥ 0, esiste $(X_n)_n$ successione di v.a. semplici ≥ 0, nondecrescente (cioè $X_n(\omega) \leq X_{n+1}(\omega)$ $\forall n$, $\forall \omega \in \Omega$) e t.c. $X_n \nearrow X$ puntualmente (cioè $X_n(\omega) \nearrow X(\omega)$ $\forall \omega$).

Dim:

Basta prendere $X_n = \sum_{k=0}^{n2^n - 1} \frac{k}{2^n} \mathbb{1}_{[\frac{k}{2^n}, \frac{k+1}{2^n})}(X) + n \mathbb{1}_{[n, +\infty)}(X)$



Prop/def: Data $X: \Omega \to \mathbb{R}$ v.a ≥ 0, sia $(X_n)_n$ successione non decrescente di v.a. semplici, ≥ 0, con $X_n \nearrow X$ puntualmente. Definiamo valore atteso di X il numero
$$E[X] := \int_\Omega X \, dP := \lim_n E[X_n] \in [0, +\infty] \quad \left(\text{si scrive anche } \int_\Omega X(\omega) P(d\omega) = \int_\Omega X(\omega) dP(\omega)\right)$$
Il limite esiste e non dipende dalla successione $(X_n)_n$ scelta.

Def: Data $X: \Omega \to \mathbb{R}$ v.a., $X^+ = \max\{X, 0\}$, $X^- = \max\{-X, 0\}$, se $E[X^+] < \infty$ oppure $E[X^-] < \infty$ definiamo
$$E[X] := \int_\Omega X \, dP = E[X^+] - E[X^-] \in [-\infty, +\infty]$$
Diciamo che X è integrabile se $E[X^+] < \infty$ e $E[X^-] < \infty$, equivalentemente se $E[|X|] < \infty$

Oss: $|X| = X^+ + X^-$, $X = X^+ - X^-$
$$E[|X|] = E[X^+] + E[X^-]$$
se X è integrabile, $|E[X]| \leq E[|X|] < \infty$

Def: Data $X:\Omega\to\mathbb{R}$ v.a. ≥ 0 o integrabile, $A\in\mathcal{F}$,

$$\int_A X\,dP := \int_\Omega 1\!\!1_A X\,dP = E[1\!\!1_A \cdot X]$$

Oss: $E[1\!\!1_A] = P(A)$

Lemma (proprietà di $E[X]$):

a) $X = c$ q.c. $\Rightarrow E[X] = c$

b) $X \overset{(d)}{=} Y$, $X \geq 0$ o integrabile $\Rightarrow Y \geq 0$ q.c. o integrabile, $E[Y] = E[X]$

in particolare, $X = Y$ q.c. $\Rightarrow E[Y] = E[X]$

c) $X \geq 0$ q.c $\Rightarrow E[X] \geq 0$

d) $X \geq 0$ q.c., $E[X] = 0 \Rightarrow X = 0$ q.c.

e) $X \geq Y$ q.c., $E[Y^-] < \infty$ oppure $E[X^+] < \infty \Rightarrow E[X] \geq E[Y]$

f) linearità: X, Y entrambe integrabili, $a\in\mathbb{R} \Rightarrow aX+Y$ integrabile, $E[aX+Y] = aE[X]+E[Y]$

lo stesso se X, Y sono entrambe ≥ 0 q.c. e $a \geq 0$.

Lemma (Fatou): Siano X_n, $n\in\mathbb{N}^+$, v.a reali ≥ 0 q.c. Allora

$$E[\varliminf_n X_n] \leq \varliminf_n E[X_n]$$

Teor (di convergenza monotona): Siano X_n, $n\in\mathbb{N}$, X v.a. reali con $0 \leq X_n \nearrow X$ q.c. (cioè, per q.o. ω, $X_n(\omega) \geq 0$, $(X_n(\omega))_n$ non decrescente, $X_n(\omega) \nearrow X(\omega)$). Allora

$$E[X_n] \nearrow E[X]$$

Teor (di convergenza dominata): Siano X_n, $n\in\mathbb{N}$, X v.a reali, con $X_n \to X$ q.c (cioè $P\{X_n \to X\} = 1$)
Supponiamo che esista Y integrabile ≥ 0 con $|X_n| \leq Y$ q.c. $\forall n$. Allora X_n e X sono integrabili e $E[X_n] \to E[X]$.

Oss: Le stesse def. e proprietà si estendono all'integrale secondo Lebesgue

$$\int_\Omega \varphi\, d\mu$$

con $\varphi: \Omega \to \mathbb{R}$ boreliana, μ misura su $(\Omega, \mathcal{F})$. L'unica differenza è che nella def. di funzione semplice $\varphi = \sum_{i=1}^m a_i 1\!\!1_{A_i}$, si chiede anche che $\mu(A_i) < \infty$ $\forall i$.

In particolare, per $\mu = m$ misura di Lebesgue su $\mathbb{R}$ o su $\mathbb{R}^d$, si trova la def di

$$\int_{\mathbb{R}} \varphi\, dx \ , \ \int_{\mathbb{R}^d} \varphi\, dx \quad \text{(integrale rispetto alla misura di Lebesgue)}$$

Prop (valore atteso in funzione della legge): $(\Omega, \mathcal{F}, P)$ sp. di prob.

a) Sia $X: \Omega \to \mathbb{R}$ v.a.

- X è integrabile $\Leftrightarrow \int_{\mathbb{R}} |x|\, P^X(dx) < \infty$

- Se X è integrabile o ≥ 0 q.c., allora
  $$E[X] = \int_{\mathbb{R}} x \, P^X(dx)$$

b) Più in generale, date $(S, \mathcal{S})$ sp. mis, $X: \Omega \to S$ v.a., $\varphi: S \to \mathbb{R}$ boreliana,

- $\varphi(X)$ è integrabile $\iff \int_S |\varphi(x)| \, P^X(dx) < \infty$
- Se $\varphi(X)$ è integrabile o ≥ 0 q.c., allora
  $$E[\varphi(X)] = \int_S \varphi(x) \, P^X(dx)$$

Dim:

(a) Segue da (b) con $\varphi(x) = x$.

(b) Dimostriamo (b) per 1. $\varphi = \mathbb{1}_A$, 2. φ semplice, 3. $\varphi \geq 0$, 4. φ generica

1. Se $\varphi = \mathbb{1}_A$, allora $\varphi(X) = \mathbb{1}_A(X) = \mathbb{1}_{\{X \in A\}}$ e quindi
   $$E[\mathbb{1}_A(X)] = E[\mathbb{1}_{X \in A}] = P\{X \in A\} = P^X(A) = \int_S \mathbb{1}_A \, P^X(dx)$$
2. Per φ semplice (combinazione lin. di indicatrici) si usa (1.) e linearità di E
3. Per $\varphi \geq 0$, siano φ_n semplici, $0 \leq \varphi_n \uparrow \varphi$, quindi $0 \leq \varphi(X_n) \uparrow \varphi(X)$
   Per (2), $E[\varphi_n(X)] = \int_S \varphi_n(x) \, P^X(dx) \quad \forall n$.
   Per conv monotona, $E[\varphi(X_n)] \uparrow E[\varphi(X)]$
   $$\int_S \varphi_n(x) P^X(dx) \uparrow \int_S \varphi(x) P^X(dx)$$
   quindi $E[\varphi(X)] = \int_S \varphi(x) P^X(dx)$
4. Per φ generica (boreliana), $\varphi(X)$ è integrabile $\iff E[|\varphi(X)|] \overset{(3)}{=} \int |\varphi(x)| P^X(dx) < \infty$
   Scomponendo $\varphi = \varphi^+ - \varphi^-$, e applicando (3) a φ^+, φ^-, si ottiene $E[\varphi(X)] = \int \varphi(x) P^X(dx)$

Oss importante: $E[\varphi(X)]$ dipende solo dalla legge di X
(cioè se $X \overset{(d)}{=} Y$, allora $E[\varphi(X)] = E[\varphi(Y)]$)

Calcolo del valore atteso:

- Caso $X: \Omega \to S$ discreta, con densità discreta p_X: come abbiamo visto, per $\varphi: S \to \mathbb{R}$ boreliana,
  - $\varphi(X)$ integrabile $\iff E[|\varphi(X)|] = \sum_{x \in R_{p_X}} |\varphi(x)| \, p_X(x) < \infty$
  - se $\varphi(X)$ è integrabile o ≥ 0 q.c, $E[\varphi(X)] = \sum_{x \in R_{p_X}} \varphi(x) \, p_X(x)$
- Caso $X: \Omega \to \mathbb{R}$ assolutamente continua, con densità f: per $\varphi: \mathbb{R} \to \mathbb{R}$ boreliana
  - $\varphi(X)$ integrabile $\iff E[|\varphi(X)|] = \int_{\mathbb{R}} |\varphi(x)| f(x) \, dx < \infty$
  - se $\varphi(X)$ è integrabile o ≥ 0 q.c, $E[\varphi(X)] = \int_{\mathbb{R}} \varphi(x) f(x) \, dx$

Dim:

Secondo lo schema indicatrici, semplici, ≥ 0, generiche

1. $\varphi = \mathbb{1}_A$, $A \in \mathcal{B}(\mathbb{R})$: $E[\mathbb{1}_A(X)] = P\{X \in A\} = \int_A f(x)\,dx$
2. φ semplice: da (1.) usando linearità di E
3. $\varphi \geq 0$: prendiamo φ_n semplici, $0 \leq \varphi_n \uparrow \varphi$, allora per conv dominata e (2)

$$E[\varphi(X)] \underset{\text{conv. monotona}}{=} \lim_n E[\varphi_n(X)] \underset{(2)}{=} \lim_n \int \varphi_n(x) f(x)\,dx \underset{\text{conv monotona per } \varphi_n f}{=} \int \varphi(x) f(x)\,dx$$

4. φ generica: la condizione di integrabilità segue da (3) per $|\varphi|$
   l'uguaglianza $E[\varphi(X)] = \int \varphi(x) f(x)\,dx$ segue da (3) per φ_+, φ_-

Esempi notevoli:

- $X \sim \mathcal{U}((a,b))$, $a<b$ $\left(f(x) = \frac{1}{b-a} \mathbb{1}_{(a,b)}(x)\right)$

  X è limitata q.c., quindi $E[|X|] = \int |x| f(x)\,dx < \infty$

  $E[X] = \int x f(x)\,dx = \frac{1}{b-a}\int_a^b x\,dx = \frac{1}{b-a}\,\frac{1}{2}(b^2-a^2) = \frac{a+b}{2}$

- $X \sim \mathcal{E}xp(\lambda)$, $\lambda > 0$ $\left(f(x) = \lambda e^{-\lambda x} \mathbb{1}_{(0,+\infty)}(x)\right)$

  $X \geq 0$ q.c., quindi $\exists\, E[X] \in [0,+\infty]$

  $E[X] = \int_0^{+\infty} x\,\lambda e^{-\lambda x}\,dx \underset{\text{per parti}}{=} -x e^{-\lambda x}\Big|_{x=0}^{+\infty} + \int_0^{\infty} e^{-\lambda x}\,dx = 0 - \frac{1}{\lambda} e^{-\lambda x}\Big|_{x=0}^{+\infty} = \frac{1}{\lambda}$

- $X \sim \Gamma(r,\lambda)$, $r>0$, $\lambda>0$

  $E[X] = \frac{r}{\lambda}$

- $X \sim \mathcal{N}(0,1)$ $\left(f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\right)$

  $E[|X|] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |x| e^{-x^2/2}\,dx < \infty$

  $$E[X] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} x e^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^0 x e^{-x^2/2}\,dx + \int_0^{+\infty} x e^{-x^2/2}\,dx\right) = 0$$

  $$\int_{-\infty}^0 x e^{-x^2/2}\,dx \underset{y=-x}{=} \int_0^{+\infty} -y\, e^{-y^2/2}\,dy$$

- $X \sim \mathcal{N}(m,\sigma^2)$

  $E[X] = m$ (con cambio di variabile oppure standardizzazione, vedi dopo)

Def. Data $X: \Omega \to \mathbb{R}$ v.a. reale, mediana di X: ogni valore $m \in \mathbb{R}$ t.c. $P\{X \leq m\} \geq \frac{1}{2}$ e $P\{X \geq m\} \geq \frac{1}{2}$.

Esercizio: $m = \inf\{t \in \mathbb{R} \mid F_X(t) \geq \frac{1}{2}\}$ (con F_X FdR di X) è una mediana

Le definizioni e proprietà dei momenti si estendono senza difficoltà dal caso discreto al caso generale

in particolare ricordiamo disuguaglianza di Markov per $X \geq 0$: $\forall a > 0$

$$P\{X \geq a\} \leq \frac{1}{a} E[X]$$

e il suo corollario $P\{|X| > a\} \leq \frac{1}{a^p} E[|X|^p]$

Esempio notevole: momenti di una gaussiana standard $\mathcal{N}(0,1)$: per $X \sim \mathcal{N}(0,1)$, $p>0$

$$E[|X|^p] = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |x|^p e^{-x^2/2} dx < \infty$$

quindi X ammette momenti di ogni ordine $p>0$; per $p \in \mathbb{N}^+$:

$$E[X^p] = \int_{\mathbb{R}} x^p \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \begin{cases} 0 & \text{per } p \text{ dispari, per simmetria} \\ (p-1)!! & \text{per } p \text{ pari, usando induzione e integrazione per parti (esercizio)} \end{cases}$$

dove $(p-1)!! = (p-1)\cdot(p-3)\cdot \ldots \cdot 3 \cdot 1$

Le definizioni e proprietà di varianza e deviazione standard, per X con $E[|X|^2] < \infty$,

$$Var(X) = E[(X - E[X])^2], \quad \sigma(X) = \sqrt{Var(X)}$$

si estendono senza difficoltà dal caso discreto al caso generale, in particolare ricordiamo la dis. di Chebyshev

$$P\{|X - E[X]| > a\} \leq \frac{1}{a^2} Var(X)$$

Esempi notevoli:

- $X \sim \mathcal{U}((a,b))$ $\quad \left(f(x) = \frac{1}{b-a} \mathbb{1}_{(a,b)}(x)\right)$

$$E[X^2] = \int x^2 f(x) dx = \frac{1}{b-a} \int_a^b x^2 dx = \frac{1}{b-a} \frac{1}{3}(b^3 - a^3) = \frac{a^2+ab+b^2}{3}$$

$$Var(X) = E[X^2] - E[X]^2 = \frac{(b-a)^2}{12}$$

- $X \sim \mathcal{E}xp(\lambda)$, $\lambda > 0$ $\quad \left(f(x) = \lambda e^{-\lambda x} \mathbb{1}_{(0,+\infty)}(x)\right)$

$$E[X^2] = \int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx = -x^2 e^{-\lambda x}\Big|_{x=0}^{+\infty} + \int_0^{+\infty} 2x e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

$$Var(X) = E[X^2] - E[X^2] = \frac{1}{\lambda^2}$$

- $X \sim \Gamma(r, \lambda)$

$$Var(X) = \frac{r}{\lambda^2}$$

- $X \sim \mathcal{N}(0,1)$

$$Var(X^2) = E[X^2] = 1!! = 1$$

- $X \sim \mathcal{N}(m, \sigma^2)$

$$Var(X^2) = \sigma^2$$

17

sabato 15 aprile 2023 16:18

Indipendenza, covarianza e correlazione

$(\Omega, \mathcal{F}, P)$ sp. di prob.

Oss: Date $X: \Omega \to S_1$, $Y: \Omega \to S_2$ v.a. (con $(S_1, \mathcal{F}_1)$, $(S_2, \mathcal{F}_2)$ sp. misurabili), con legge congiunta $P_{(X,Y)}$ su $(S_1 \times S_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$, data $\varphi: S_1 \times S_2 \to \mathbb{R}$ boreliana con $\varphi(X,Y) \geq 0$ q.c. o $\varphi(X,Y)$ integrabile, vale

$$E[\varphi(X,Y)] = \int_{S_1 \times S_2} \varphi(x,y) \, P_{(X,Y)}(d(x,y))$$

Calcolo nel caso discreto: se (X,Y) è discreta ($\Leftrightarrow X, Y$ discrete), $E[\varphi(X,Y)] = \sum_{(x,y)} \varphi(x,y) \, p_{(X,Y)}(x,y)$ con $p_{(X,Y)}$ densità congiunta

Calcolo nel caso assolutamente continuo:

Se $(X,Y): \Omega \to \mathbb{R}^2$ è assolutamente continua con densità $f: \mathbb{R}^2 \to \mathbb{R}$, allora

- $\varphi(X,Y)$ è integrabile $\Leftrightarrow \int_{\mathbb{R}^2} |\varphi(x,y)| f(x,y) \, dx dy < \infty$
- se $\varphi(X,Y)$ è ≥ 0 q.c. o integrabile, allora
$$E[\varphi(X,Y)] = \int_{\mathbb{R}^2} \varphi(x,y) f(x,y) \, dx dy$$

Dim: come nel caso di una v.a. reale assolutamente continua (indicatrici, semplici, ≥ 0, generali)

Queste formule si generalizzano in modo naturale al caso $(X_1, \dots, X_n): \Omega \to S_1 \times \dots \times S_n$

- $E[\varphi(X_1, \dots, X_n)] = \int_{S_1 \times \dots \times S_n} \varphi(x_1, \dots, x_n) \, P_{(X_1, \dots, X_n)}(d(x_1, \dots, x_n))$
- caso discreto: $E[\varphi(X_1, \dots, X_n)] = \sum_{(x_1, \dots, x_n)} \varphi(x_1, \dots, x_n) \, p_{(X_1, \dots, X_n)}(x_1, \dots, x_n)$
- caso $(X_1, \dots, X_n): \Omega \to \mathbb{R}^n$
$$E[\varphi(X_1, \dots, X_n)] = \int_{\mathbb{R}^n} \varphi(x_1, \dots, x_n) f(x_1, \dots, x_n) \, dx_1 \dots dx_n$$

La dis di Schwarz si estende senza difficoltà dal caso discreto al caso generale

$$E[|XY|] \leq E[X^2]^{1/2} E[Y^2]^{1/2} \quad \text{per } X, Y \text{ v.a. reali}$$

Prop: Siano X, Y v.a. reali indipendenti. Se X, Y sono entrambe integrabili (o entrambe ≥ 0), allora XY è integrabile (o ≥ 0) e vale

$$E[XY] = E[X]E[Y]$$

Dim: Secondo lo schema indicatrici, semplici, ≥ 0, generali

1. $X = 1\!\!1_A$ $Y = 1\!\!1_B$, $A, B \in \mathcal{F}$

   $1\!\!1_A, 1\!\!1_B$ sono indip $\Leftrightarrow$ A e B sono indip.

   dim di $\Rightarrow$: $A = \{1\!\!1_A = 1\}$, $B = \{1\!\!1_B = 1\}$, quindi $P(A \cap B) = P\{1\!\!1_A = 1, 1\!\!1_B = 1\}$
   $= P\{1\!\!1_A = 1\} \cdot P\{1\!\!1_B = 1\} = P(A)P(B)$

   $$E[\underbrace{1\!\!1_A 1\!\!1_B}_{= 1\!\!1_{A\cap B}}] = P(A \cap B) \underset{\text{indip. di A e B}}{=} P(A)P(B) = E[1\!\!1_A]E[1\!\!1_B]$$

2. X, Y semplici: da (1) e linearità: $X = \sum_{i=1}^{n} a_i 1\!\!1_{A_i}$, $Y = \sum_{j=1}^{m} b_j 1\!\!1_{B_j}$

   $$E[XY] = E\Big[\sum_{i,j} a_i b_j 1\!\!1_{A_i} 1\!\!1_{B_j}\Big] = \sum_{i,j} a_i b_j E[1\!\!1_{A_i} 1\!\!1_{B_j}]$$
   $$\underset{(1)}{=} \sum_{i,j} a_i b_j E[1\!\!1_{A_i}]E[1\!\!1_{B_j}] = \sum_i a_i E[1\!\!1_{A_i}] \cdot \sum_j b_j E[1\!\!1_{B_j}]$$
   $$= E[X]E[Y]$$

3. $X, Y \geq 0$: da (2) e conv. monotona

   siano $X_n = h_n(X)$, $Y_n = h_n(Y)$, con $h_n : \mathbb{R} \to \mathbb{R}$, $h_n(x) = \sum_{k=0}^{2^n-1} \frac{k}{2^n} 1\!\!1_{[\frac{k}{2^n}, \frac{k+1}{2^n})}(x) + n 1\!\!1_{[n,+\infty)}(x)$

   X_n, Y_n sono semplici, $0 \leq X_n \uparrow X$, $0 \leq Y_n \uparrow Y$ e quindi $X_n Y_n$ sono semplici, $0 \leq X_n Y_n \uparrow XY$

   $\forall n$, X_n e Y_n sono indipendenti, perché funzioni di v.a. indipendenti

   $$E[XY] \underset{\text{conv. mon.}}{=} \lim_n E[X_n Y_n] \underset{(2)}{=} \lim_n E[X_n]E[Y_n] \underset{\text{conv mon.}}{=} E[X]E[Y]$$

4. X, Y integrabili: da (3)

   $X^+ = \max\{X, 0\}$, $Y^+ = \max\{Y, 0\}$ sono indip. poiché funz. di indip.,

   analogamente X^-, Y^+, sono indip., X^+, Y^- sono indip., X^-, Y^- sono indip.

   quindi si decompone $XY = X^+Y^+ - X^-Y^+ - X^+Y^- + X^-Y^-$ e si usa (3).

Dim per (X, Y) v.a. assolutamente continua:

Se (X, Y) è assol cont con densità f, X e Y sono indip. $\Leftrightarrow f(x,y) = f_X(x) f_Y(y)$
con f_X e f_Y densità di X e Y risp.

$$E[|XY|] = \iint_{\mathbb{R}^2} |xy| f(x,y)\, dx\, dy \underset{\text{indip}}{=} \iint_{\mathbb{R}^2} |x| f_X(x) |y| f_Y(y)\, dx\, dy$$
$$\underset{\text{Fubini-Tonelli}}{=} \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} |x| f_X(x)\, dx \Big) |y| f_Y(y)\, dy = \int_{\mathbb{R}} |x| f_X(x)\, dx \int_{\mathbb{R}} |y| f_Y(y)\, dy$$
$$= E[|X|]\, E[|Y|] < \infty \text{ se } X, Y \text{ sono integrabili}$$

Per $E[XY]$, si ripetono i passaggi sopra senza i moduli e si ottiene $E[XY] = E[X]E[Y]$

Come nel caso discreto, si dimostra

Cor: Siano $X : \Omega \to S_1$, $Y : \Omega \to S_2$ v.a.. Allora

X, Y sono indip. $\Leftrightarrow E[g(X)h(Y)] = E[g(X)]E[h(Y)]$ $\forall g : S_1 \to \mathbb{R}$, $h : S_2 \to \mathbb{R}$ boreliane

con $g(X)$, $h(Y)$ entrambe ≥ 0 q.c.

o entrambe integrabili

Le def. e le proprietà di covarianza, coefficiente di correlazione, retta di regressione si estendono senza difficoltà dal caso discreto al caso generale, in particolare ricordiamo:

Prop: Date $X_1, .. X_n$ v.a. reali con $E[|X_i|^2] < \infty$ $\forall i$,

$$Var\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n Var(X_i) + \sum_{1 \leq i \neq j \leq n} Cov(X_i, X_j)$$

In particolare, se X_i sono indip o anche solo scorrelate a due a due

$$Var\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n Var(X_i)$$

18
sabato 15 aprile 2023 19:53

Trasformazioni di v.a.

Oss: Abbiamo visto che, se $X:\Omega\to\mathbb{R}$ è assolutamente continua e $\varphi:\mathbb{R}\to\mathbb{R}$ è boreliana, anche regolare, non è detto che $\varphi(X)$ sia assolutamente continua (es: $\varphi\equiv 0$)

Prop (Formula di cambio variabili): Sia $X:\Omega\to\mathbb{R}^d$ v.a. assolutamente continua, con densità f_X, con $f_X=0$ fuori da un aperto $O\subseteq\mathbb{R}^d$. Sia $\varphi:O\to O'$ un diffeomorfismo C^1 con $O'\subseteq\mathbb{R}^d$ aperto (cioè $\varphi:O\to O'$ è C^1, invertibile e φ^{-1} è C^1). Allora la v.a. $Y=\varphi(X):\Omega\to\mathbb{R}^d$ è assolutamente continua con densità

$$f_Y(y)=f_X(\varphi^{-1}(y))\,|\det D\varphi^{-1}(y)|\,\mathbb{1}_{O'}(y)\qquad y\in\mathbb{R}^d$$

[Idea: " $y=\varphi(x)$ $\quad x=\varphi^{-1}(y)\quad dx=|\det D\varphi^{-1}(y)|\,dy$ "]

Dim:

Per $A\in\mathcal{B}(\mathbb{R}^d)$, dobbiamo dim $P\{Y\in A\}=\int_A f_Y(y)\,dy$, f_Y come sopra

$$P\{\underset{\varphi(X)}{\overset{\shortparallel}{Y}}\in A\}=P\{X\in\varphi^{-1}(A)\}=\int \mathbb{1}_{\varphi^{-1}(A)}(x)f_X(x)\mathbb{1}_O(x)\,dx$$

cambio variabile: $x=\varphi^{-1}(y)$, $dx=\det D\varphi^{-1}(y)dy$

$$=\int \underbrace{\mathbb{1}_{\varphi^{-1}(A)}(\varphi^{-1}(y))}_{=\mathbb{1}_A(y)} f_X(\varphi^{-1}(y))\,|\det D\varphi^{-1}(y)|\,\overbrace{\mathbb{1}_O(\varphi^{-1}(y))}^{=\mathbb{1}_{O'}(y)}\,dy$$

$$=\int \mathbb{1}_A(y)\,f_Y(y)\,dy$$

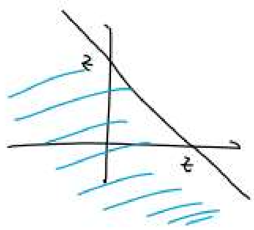
Nel caso generale, per $X:\Omega\to\mathbb{R}^d$ v.a., $\varphi:\mathbb{R}^d\to\mathbb{R}^m$ boreliana, si può calcolare la FdR di $Y=\varphi(X)$ per studiare le proprietà della legge di Y

$$F_Y(y_1,\dots,y_m)=P\{\varphi_1(X)\le y_1,\dots,\varphi_m(X)\le y_m\}$$

Prop (densità della somma):

Siano $X:\Omega\to\mathbb{R}$, $Y:\Omega\to\mathbb{R}$ v.a. reali con (X,Y) assolutamente continua di densità $f_{(X,Y)}$.
Allora $X+Y$ è assolutamente continua, con densità

$$f_{X+Y}(z)=\int_{\mathbb{R}} f_{(X,Y)}(x,z-x)\,dx=\int_{\mathbb{R}} f_{(X,Y)}(z-y,y)\,dy$$



Dim:

Sia F_{X+Y} la FdR di $X+Y$, vogliamo mostrare

$$F_{X+Y}(z)=\int_{-\infty}^{z}\dots\qquad\forall z\in\mathbb{R}$$

$$F_{X+Y}(z) = P\{X+Y\le z\} = P\{(X,Y)\in\{(x,y)|x+y\le z\}\}$$
$$= \iint 1\!\!1_{x+y\le z}\, f_{(X,Y)}(x,y)\,dx\,dy \quad \left[\text{dove } 1\!\!1_{x+y\le z} = 1\!\!1_{\{(x',y')|x'+y'\le z\}}(x,y)\right]$$
$$= \int \Big(\underbrace{\int 1\!\!1_{x+y\le z}\, f_{(X,Y)}(x,y)\,dy}\Big)\,dx$$

cambio var $y = z'-x$, $dy = dz'$

$$= \int 1\!\!1_{z'\le z}\, f_{(X,Y)}(x, z'-x)\,dz'$$
$$= \int\Big(\int 1\!\!1_{z'\le z}\, f_{(X,Y)}(x, z'-x)\,dx\Big)dy = \int_{-\infty}^{z} f_{X+Y}(z')\,dz'$$

L'altra formula si dimostra scambiando x e y.

Cor: (formula della convoluzione):

Se $X:\Omega\to\mathbb{R}$, $Y:\Omega\to\mathbb{R}$ sono v.a. reali, assolutamente continue con densità risp. f_X, f_Y, e indipendenti, allora $X+Y$ è assolutamente continua con densità

$$f_{X+Y}(z) = \int_{\mathbb{R}} f_X(x) f_Y(z-x)\,dx = \int_{\mathbb{R}} f_X(z-y) f_Y(y)\,dy =: f_X * f_Y(z)$$

(convoluzione di f_X e f_Y)

Dim: dalla prop precedente con $f_{(X,Y)}(x,y) = f_X(x) f_Y(y)$.

Prop Siano X, Y v.a. (discrete) a valori interi, con densità discreta congiunta $p_{(X,Y)}$. Allora $X+Y$ ha densità discreta $p_{X+Y}(k) = \sum_{j\in\mathbb{Z}} p_{(X,Y)}(j, k-j) = \sum_{j\in\mathbb{Z}} p_{(X,Y)}(j,k-j)$, $k\in\mathbb{Z}$.

In particolare, se X e Y sono indipendenti, di densità discrete risp. p_X, p_Y,

$$p_{X+Y}(k) = \sum_{j\in\mathbb{Z}} p_X(j)\, p_Y(k-j) = \sum_{j\in\mathbb{Z}} p_X(k-j)\, p_Y(j) \quad, k\in\mathbb{Z}$$

Applicazioni:

- $X\sim B(n,p)$, $Y\sim B(m,p)$ indipendenti $\Rightarrow X+Y\sim B(n+m,p)$ (riproducibilità delle binomiali)
- $X\sim \mathcal{P}(\lambda)$, $Y\sim \mathcal{P}(\mu)$ indipendenti $\Rightarrow X+Y\sim\mathcal{P}(\lambda+\mu)$ (" " Poisson)
- $X\sim \Gamma(r,\lambda)$, $Y\sim\Gamma(s,\lambda)$ indipendenti $\Rightarrow X+Y\sim\Gamma(r+s,\lambda)$ (" " Gamma)

in particolare $X_1,\dots X_n \sim \mathcal{E}(\lambda) = \Gamma(1,\lambda)$ indip. $\Rightarrow X_1+\dots+X_n\sim\Gamma(n,\lambda)$

Dim: $X\sim\Gamma(r,\lambda)$, $Y\sim\Gamma(s,\lambda)$ indip $\quad f_X(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} 1\!\!1_{(0,+\infty)}(x)$, $z>0$

$$f_{X+Y}(z) = \int_{\mathbb{R}} f_X(x) f_Y(z-x)\,dx = \int \frac{\lambda^r}{\Gamma(r)}\frac{\lambda^s}{\Gamma(s)} x^{r-1}(z-x)^{s-1} e^{-\lambda x} e^{-\lambda(z-x)} 1\!\!1_{x>0} 1\!\!1_{z-x>0}\,dx$$
$$= \frac{\lambda^{r+s}}{\Gamma(r)\Gamma(s)} \int_0^z x^{r-1}(z-x)^{s-1}\,dx \; e^{-\lambda z}\; 1\!\!1_{z>0}$$

$$\begin{array}{l} x = zu \\ dx = z\,du \end{array} = \frac{\lambda^{r+s}}{\Gamma(r)\Gamma(s)} \int_0^1 z^{r-1} u^{r-1} z^{s-1} (1-u)^{s-1} z\, du \quad e^{-\lambda z} \mathbb{1}_{z>0}$$

$$= \frac{\lambda^{r+s}}{\Gamma(r)\Gamma(s)} \underbrace{\int_0^1 u^{r-1}(1-u)^{s-1} du}_{B(r,s) = \frac{\Gamma(r)\Gamma(s)}{\Gamma(r+s)}} \; z^{r+s-1} e^{-\lambda z} \mathbb{1}_{z>0} = \frac{\lambda^{r+s}}{\Gamma(r+s)} z^{r+s-1} e^{-\lambda z} \mathbb{1}_{z>0}$$

V. a. gaussiane: standardizzazione

Prop: $X \sim N(m, \sigma^2)$, $a, b \in \mathbb{R}$ con $a \neq 0 \Rightarrow aX+b \sim N(am+b, a^2\sigma^2)$

Dim: $\left[f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-(x-m)^2/2\sigma^2\right) \right]$

Con formula di cambio variabile per $\varphi(x) = ax+b$, $|D\varphi^{-1}(y)| = \frac{1}{|a|}$: $Y = aX+b$ ha densità

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\left(\frac{y-b}{a} - m\right)^2 / 2\sigma^2\right) \frac{1}{|a|}$$

$$= \frac{1}{\sqrt{2\pi(a\sigma)^2}} \exp\left(-(y-(am+b))^2 / 2(a\sigma)^2\right)$$

Cor (standardizzazione): $X \sim N(m, \sigma^2) \iff \frac{X-m}{\sigma} \sim N(0,1)$

equivalentemente, $Z \sim N(0,1) \iff \sigma Z + m \sim N(m, \sigma^2)$

Calcoli con v.a. gaussiane: data $X \sim N(m, \sigma^2)$, dati $-\infty \leq a < b \leq +\infty$, $P\{a \leq X \leq b\} = ?$

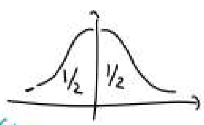
$P\{a \leq X \leq b\} = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx$ ma questo integrale non ha una forma esplicita (salvo per particolari valori di a e b)

- Se $X \sim N(0,1)$, allora

  $P\{a \leq X \leq b\} = \Phi(b) - \Phi(a)$ con Φ FdR di $N(0,1)$

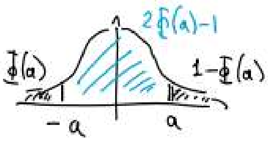
  e, per molti $x \in \mathbb{R}$, $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$ è noto con ottima approssimazione (ed è presente nei software più comuni)

  Proprietà di Φ:



  - $\Phi(0) = \frac{1}{2}$, cioè $P\{X \leq 0\} = P\{X \geq 0\} = \frac{1}{2}$
  - $\Phi(-x) = 1 - \Phi(x) \quad \forall x \in \mathbb{R}$

    in particolare $P\{-a \leq X \leq a\} = 2\Phi(a) - 1$



  - $\Phi(x) \approx 1$ per $x \geq 5$

    $\Phi(x) \approx 0$ per $x \leq -5$

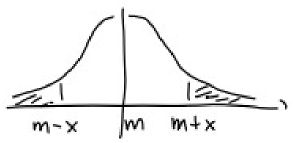
    cioè $P\{-5 \leq X \leq 5\}$ è quasi 1.
  - $P\{-3 \leq X \leq 3\} \approx 0.997$

- Se $X \sim \mathcal{N}(m, \sigma^2)$, ci si riconduce a $\mathcal{N}(0,1)$ tramite standardizzazione

$$P\{a \le X \le b\} = P\Big\{ \frac{a-m}{\sigma} \le \underbrace{\frac{X-m}{\sigma}}_{\sim \mathcal{N}(0,1)} \le \frac{b-m}{\sigma} \Big\} = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right)$$

In particolare

- $P\{X \le m\} = P\{X \ge m\} = \frac{1}{2}$
- $P\{X \le m-x\} = P\{X \ge m+x\} \quad \forall x \in \mathbb{R}$
- $P\{m - 5\sigma \le X \le m + 5\sigma\} \simeq 1$
- $P\{m - 3\sigma \le X \le m + 3\sigma\} \simeq 0.997$



Valore atteso, varianza e momenti di $X \sim \mathcal{N}(m, \sigma^2)$

$X = \sigma Z + m$, $Z \sim \mathcal{N}(0,1)$, quindi

- $E[X] = \sigma E[Z] + m = m$
- $Var(X) = \sigma^2 Var(Z) = \sigma^2$
- $E[|X-m|^p] = \sigma^p E[|Z|^p] = c_p Var(X)^{p/2} \quad \forall p > 0 \quad$ con $c_p = E[|Z|^p]$

(per una v.a. gaussiana, i momenti assoluti p-simi sono controllati dalla varianza)

Riproducibilità delle gaussiane:

Prop: $X \sim \mathcal{N}(m_1, \sigma_1^2)$, $Y \sim \mathcal{N}(m_2, \sigma_2^2)$ indipendenti $\Rightarrow X + Y \sim \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$

Dim:

1. $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0, \sigma^2)$ indip $\Rightarrow X + Y \sim \mathcal{N}(0, \sigma^2 + 1)$

$$f_{X+Y}(z) = \int f_X(x) f_Y(z-x)\, dx = \int \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(z-x)^2/2\sigma^2} dx$$

$$= \int \frac{1}{2\pi\sigma} \exp\left[-\frac{1}{2\sigma^2}\left((\sigma^2+1)x^2 - 2xz + z^2 \right)\right] dx$$

$$= \int \frac{1}{2\pi\sigma} \exp\left[-\frac{1}{2\sigma^2}\left(\sqrt{\sigma^2+1}\, x - \frac{1}{\sqrt{\sigma^2+1}} z \right)^2 - \frac{1}{2\sigma^2} \cdot \frac{\sigma^2}{\sigma^2+1} z^2 \right] dx$$

$$= \exp\left(-\frac{1}{2(\sigma^2+1)} z^2 \right) \cdot \frac{1}{2\pi\sigma} \int \exp\left(-\frac{1}{2\sigma^2}\left(\sqrt{\sigma^2+1}\, x - \frac{1}{\sqrt{\sigma^2+1}} z \right)^2 \right) dx \qquad \begin{array}{l} y = \sqrt{\sigma^2+1}\, x - \frac{1}{\sqrt{\sigma^2+1}} z \\ dy = \sqrt{\sigma^2+1}\, dx \end{array}$$

$$= \exp\left(-\frac{1}{2(\sigma^2+1)} z^2 \right) \cdot \frac{1}{2\pi\sigma} \cdot \frac{1}{\sqrt{\sigma^2+1}} \underbrace{\int \exp\left(-\frac{y^2}{2\sigma^2} \right) dy}_{= \sqrt{2\pi\sigma^2}}$$

$$= \exp\left(-\frac{1}{2(\sigma^2+1)} z^2 \right) \cdot \frac{1}{\sqrt{2\pi(\sigma^2+1)}} \qquad \text{densità } \mathcal{N}(0, \sigma^2+1)$$

2. Caso generale

$Z = \frac{X - m_1}{\sigma_1} \sim \mathcal{N}(0,1)$, $V = \frac{Y - m_2}{\sigma_1} \sim \mathcal{N}\left(0, \left(\frac{\sigma_2}{\sigma_1}\right)^2\right)$ indip., quindi

$$X + Y = m_1 + m_2 + \sigma_1 \underbrace{(Z + V)}_{\sim \mathcal{N}\left(0, \left(\frac{\sigma_2}{\sigma_1}\right)^2 + 1\right)} \sim \mathcal{N}\left(m_1 + m_2,\ \sigma_1^2 \left(\frac{\sigma_2^2}{\sigma_1^2} + 1 \right) = \sigma_1^2 + \sigma_2^2 \right)$$

19
martedì 25 aprile 2023 18:12

Teoremi limite (LGN, TCL)

$(\Omega, \mathcal{F}, P)$ spazio di prob.

Def: Data una sequenza (finita o infinita) di v.a. $X_1, \dots X_n, \dots$, con $X_i : \Omega \to S$, $(S, \mathcal{S})$ sp. misurabile, X_i si dicono indipendenti e identicamente distribuite (i.i.d): se sono indipendenti e hanno la stessa legge su $(S, \mathcal{S})$

La def di conv. in probabilità si estende senza difficoltà dal caso discreto al caso gen.

Teor (Legge (debole) dei grandi numeri, LGN):

Sia $X_1, \dots X_n$, una successione di v.a. reali i.i.d. dotate di momento secondo ($E[X_1^2]<\infty$) sia $m = E[X_1]$. Allora

$$\overline{X}_n := \frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{P} m \quad \text{per } n \to \infty$$

cioè $P\{|\overline{X}_n - m| > \varepsilon\} \to 0$ per $n \to \infty$, $\forall \varepsilon > 0$.

La dim. è la stessa del caso discreto

Oss: Come nel caso discreto, l'ipotesi di indip può essere sostituita con quella che le X_i siano a due a due scorrelate

Cor: Sia $X_1, \dots X_n, \dots$ una successione di v.a. i.i.d. a valori in S, con $(S, \mathcal{S})$ sp. mis.
Sia $\varphi : S \to \mathbb{R}$ boreliana con $E[\varphi(X_1)^2] < \infty$. Allora

$$\frac{1}{n} \sum_{i=1}^{n} \varphi(X_i) \xrightarrow{P} E[\varphi(X_1)] \quad \text{per } n \to \infty$$

Dim: Basta applicare la LGN alle v.a. reali i.i.d. $\varphi(X_i)$.

Oss: In particolare, per $\varphi = \mathbb{1}_A$, $A \in \mathcal{S}$, ($\mathbb{1}_A(X) = \mathbb{1}_{X \in A} \sim B(p)$ con $p = P\{X \in A\}$), troviamo
freq relativa di $\{X_i \in A\} = \frac{1}{n} \#\{i \mid X_i \in A\} \xrightarrow{P} P\{X \in A\}$

Applicazione: metodo Monte-Carlo per il calcolo approssimato degli integrali
Vogliamo calcolare $\int_a^b \varphi(x)dx$ per qualche $\varphi : (a,b) \to \mathbb{R}$ boreliana integrabile
assumiamo $\int_a^b |\varphi(x)|^2 dx < \infty$ (si potrebbe rimuovere questa ipotesi)

- notiamo che $\frac{1}{b-a} \int_a^b \varphi(x)dx = E[\varphi(U)]$ con $U \sim \mathcal{U}((a,b))$ (uniforme su (a,b))
- per la LGN, prese U_i i.i.d $\sim \mathcal{U}((a,b))$, $\frac{1}{n} \sum_{i=1}^{n} \varphi(U_i) \xrightarrow{P} E[\varphi(U)] = \frac{1}{b-a} \int_a^b \varphi(x)dx$

- quindi basta generare n v.a. U_i uniformi e calcolare
  $$\frac{1}{n}\sum_{i=1}^{n} \varphi(U_i)$$

Analogamente, $\int_{(0,1)^d} \varphi(x)dx$, con $\varphi : (0,1)^d \to \mathbb{R}$ boreliana t.c. $\int_{(0,1)^d} |\varphi(x)|^2 dx < \infty$,
può essere approssimato con $\frac{1}{n}\sum_{i=1}^{n} \varphi(U_i)$, con U_i i.i.d uniformi su $(0,1)^d$ (di densità $1\!\!1_{(0,1)^d}$)

Notare che
$$P\left\{ \left| \frac{1}{n}\sum_{i=1}^{n} \varphi(U_i) - \int_{(0,1)^d} \varphi(x)dx \right| > \varepsilon \right\} \leq \frac{1}{\varepsilon^2 n} Var(\varphi(U_1)) = O_{\varepsilon,\varphi}\left(\frac{1}{n}\right)$$ non dipende (direttamente) dalla dim. d

Teor (Teorema centrale del limite, TCL/TLC)

Sia $X_1, \ldots X_n, \ldots$ una successione di v.a. reali i.i.d. dotate di momento secondo ($E[X_1^2] < \infty$) e non costanti q.c., chiamiamo $m = E[X_1]$, $\sigma^2 = Var(X_1)$ ($0 < \sigma^2 < \infty$)

Allora
$$\frac{\sum_{i=1}^{n} X_i - nm}{\sqrt{n}\,\sigma} = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - m) \text{ converge in legge a } Z \sim \mathcal{N}(0,1), \text{ cioè}$$
detta F_n la FdR di $\frac{\sqrt{n}}{\sigma}(\bar{X}_n - m)$, Φ la FdR di $\mathcal{N}(0,1)$,
$$\lim_n F_n(x) = \Phi(x) \quad \forall x \in \mathbb{R}$$

Oss: Da $P\{a \leq Z \leq b\} = F(b) - F(a)$, si ricava
$$P\left\{a \leq \frac{\sqrt{n}}{\sigma}(\bar{X}_n - m) \leq b\right\} \to P\{a \leq Z \leq b\} \quad \forall a, b \in \mathbb{R},\ a < b$$
e analogamente per le disuguaglianze strette

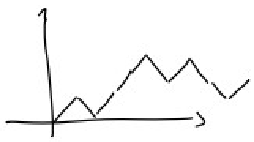
Oss Si può dimostrare che vale anche $E\left[\varphi\left(\frac{\sqrt{n}}{\sigma}(\bar{X}_n - m)\right)\right] \to E[\varphi(Z)] \quad \forall \varphi \in C_b(\mathbb{R})$

Oss: Per standardizzazione, $\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} X_i - nm\right) = \sqrt{n}(\bar{X}_n - m) \xrightarrow{\text{in legge}} Z_\sigma$, con $Z_\sigma \sim \mathcal{N}(0, \sigma^2)$
cioè $P\{a \leq \sqrt{n}(\bar{X}_n - m) \leq b\} \to P\{a \leq Z_\sigma \leq b\} \quad \forall a, b \in \mathbb{R},\ a < b$

Applicazione: convergenza delle leggi marginali di una passeggiata aleatoria

$(X_n)_{n \in \mathbb{N}^+}$ i.i.d. Rademacher, cioè $P\{X_i = 1\} = P\{X_i = -1\} = \frac{1}{2}$

$S_n = X_1 + \ldots + X_n$ passeggiata aleatoria simmetrica



consideriamo, per $t > 0$, $S_{[nt]}$
(comportamento a tempi grandi / passeggiata a step temporale $\frac{1}{n}$)

Per il TCL, per $n \to \infty$
$$\frac{1}{\sqrt{n}} S_{[nt]} = \sqrt{\frac{[nt]}{n}} \cdot \frac{1}{\sqrt{[nt]}} S_{[nt]} \xrightarrow{\text{in legge}} \sqrt{t}\, \underset{\sim \mathcal{N}(0,1)}{Z} \overset{(d)}{=} W_t \quad \text{con } W_t \sim \mathcal{N}(0, t)$$
questo è la convergenza delle leggi marginali: a t fissato, $\frac{1}{\sqrt{n}} S_{[nt]} \xrightarrow{\text{in legge}} W_t$

In realtà vale anche la convergenza come processo, cioè la convergenza delle leggi congiunte

$$\left(\frac{1}{\sqrt{n}} S_{[nt_1]}, \dots \frac{1}{\sqrt{n}} S_{[nt_n]}\right) \overset{\text{"in legge"}}{\longrightarrow} (W_{t_1}, \dots W_{t_n}) \qquad \forall\, t_1<t_2<\dots<t_n$$

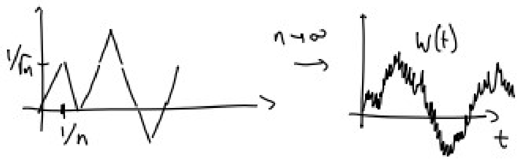
per un opportuno processo stocastico (= famiglia di v.a. indicizzate da t) W, noto come moto browniano o processo di Wiener

Notiamo

1. se lo step temporale scala come $\frac{1}{n}$, lo step spaziale scala come $\frac{1}{\sqrt{n}}$

   cioè $\delta W \approx \sqrt{\delta t}$

   ci aspettiamo allora che W sia $\approx \frac{1}{2}$-Hölder in t



   in effetti, è $\left(\frac{1}{2}-\varepsilon\right)$-Hölder $\forall \varepsilon>0$, ma non $\frac{1}{2}$-Hölder

2. la densità di W_t è $f_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)}$ $\forall t>0$, e soddisfa

$$\partial_t f(t,x) = \frac{1}{2}\partial_x^2 f(t,x) \qquad \forall (t,x)\in(0,+\infty)\times\mathbb{R}$$

   equazione della diffusione o equazione del calore