Integrale di Lebesgue

Misure e loro proprietà

Dato $\Omega \neq \emptyset$, diciamo σ-algebra ogni famiglia $\mathcal{R}$ di sottoinsiemi di Ω $\begin{cases} \Omega \in \mathcal{R} \\ A \in \mathcal{R} \to A^c \in \mathcal{R} \\ (A_n)_{n\geq 1} \in \mathcal{R} \to \bigcup_{n\geq 1} A_n \in \mathcal{R} \end{cases}$

Analoga definizione è $\emptyset, \Omega \in \mathcal{R}$
$A_1 \setminus A_2 \in \mathcal{R}$
$\bigcup_{i\in I} A_i, \bigcap_{i \in I} A_i \in \mathcal{R}$ con I finito

Esempi di σ-algebre:

- $\{\emptyset, \Omega\}$ è la più piccola σ-algebra
- 2^{Ω} è la più grande σ-algebra
- $\sigma(\mathcal{G}) = \bigcap \{\mathcal{R} \mid \mathcal{G} \subset \mathcal{R}, \mathcal{R}$ è σ-algebra$\}$ σ-algebra generata dalla famiglia di insiemi $\mathcal{G} \subseteq 2^{\Omega}$

  è l'insieme di tutti gli elementi di tutte le σ-algebre includenti $\mathcal{G}$
  $\neq \emptyset$ perché almeno $\emptyset$ c'è

  dim
  $\Omega \in \mathcal{R}$ quindi $\Omega \in \sigma(\mathcal{G})$
  $A \in \mathcal{R}$ quindi $A^c \in \sigma(\mathcal{G})$
  $(A_n)_{n\geq 1} \in \mathcal{R}$ quindi $\bigcup_{n\geq 1} A_n \in \sigma(\mathcal{G})$.

  $\sigma(\mathcal{G}) = \mathcal{G}$ se $\mathcal{G}$ è σ-algebra
  $\sigma(\mathcal{G}) \subseteq \sigma(\mathcal{G}')$ se $\mathcal{G} \subseteq \mathcal{G}'$

- $\mathcal{R} \cap S = \{A \cap S \mid A \in \mathcal{R}\}$ con $S \subset \Omega$. traccia della σ-algebra su S

  dim
  perché il più grande elemento di $\mathcal{R} \cap S$ è $\Omega \cap S$, allora $S \cap \Omega \cap S \in \mathcal{R} \cap S$
  $(A_n)_{n\geq 1} \cap S \in \mathcal{R}$, $\forall n \in \mathbb{N}$ $\exists A_n \in \mathcal{R} \mid A_n \cap S \in \mathcal{R} \cap S$, allora
  $\bigcup_{n \geq 1} (A_n \cap S) = \bigcup_{n\geq 1} A_n \cap S \in \mathcal{R}$

- $T^{-1}(\mathcal{R}) = \{T^{-1}(A) : A \in \mathcal{R}\}$ σ-algebra indotta da una applicazione T, con $T: \Omega_0 \neq \emptyset \to \Omega$

  dim
  $T(\Omega_0) = \Omega$ cioè $T^{-1}(\Omega) = \Omega_0$
  $A_0 \in \Omega_0$, $A_0 = T^{-1}(A)$ $A \in \mathcal{R}$, $T^{-1}(A) \in T^{-1}(\mathcal{R})$, allora $\underbrace{(T^{-1}(A))^c}_{\in T^{-1}(\mathcal{R})} = T^{-1}(A^c)$ anche $T^{-1}(A^c) \in T^{-1}(\mathcal{R})$
  $T^{-1}(A_n)_{n\geq 1} \in T^{-1}(\mathcal{R})$, allora $\underbrace{T^{-1}\left(\bigcup_{n\geq 1} A_n\right)}_{\bigcup_{n\geq 1} T^{-1}(A_n)} \in T^{-1}(\mathcal{R})$

Retta reale ampliata

$\mathbb{R}^* = \mathbb{R} \cup \{-\infty, +\infty\}$; l'ordinamento completo su $\mathbb{R}^*$ ottenuto prolungando l'usuale ordinamento per grandezza dei numeri reali con $-\infty$ più piccolo di tutti i reali e $+\infty$ più grande di tutti i reali. Sull'aritmetica di calcolo la novità sono: $0(-\infty) = 0(+\infty) = 0$ e $\frac{a}{-\infty} = \frac{a}{+\infty} = 0$ con $a \in \mathbb{R}$. Non si accettano rapporti di infiniti, rapporti con denominatori nulli o somme di infiniti di segno opposto o differenze di infiniti di uguale segno.

- $\mathcal{B}$ σ-algebra di Borel su $\mathbb{R}$, $\mathcal{B} = \sigma(\{]a,b] : a<b; a,b \in \mathbb{R}\})$
  è generata da famiglia degli intervalli limitati inferiormente aperti e superiormente chiusi
  è generata dalla famiglia delle semirette inferiori $]-\infty, a]$
  è generata dalla famiglia degli insiemi aperti

- $\mathcal{B}^*$ σ-algebra di Borel su $\mathbb{R}^*$, $\mathcal{B}^* = \sigma\{(]a,b] : a<b,\ a,b \in \mathbb{R}^*\}$ 2
  è generata dalla famiglia degli intervalli di $\mathbb{R}^*$ inferiormente aperti e superiormente chiusi
  è generata dalla famiglia degli intorni di base della retta reale ampliata
  è generata dalla famiglia $\{[-\infty, a],\ a \in \mathbb{R}\}$ delle semirette inferiori della retta reale ampliata di origine un n° reale

$\mathcal{B} \subset \mathcal{B}^*$, gli elementi di $\mathcal{B}$ o $\mathcal{B}^*$ sono detti boreliani di $\mathbb{R}$ o $\mathbb{R}^*$

Misura su $\mathcal{A}$

data $\mathcal{A}$ σ-algebra qualsiasi, $m: \mathcal{A} \to [0,+\infty]$ è una MISURA su $\mathcal{A}$ se
- $m(\emptyset) = 0$
- $m\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n\geq 1} m(A_n)$ con $(A_n)_{n\geq 1}$ disgiunta

NUMERABILE ADDITIVITÀ

Misura di Lebesgue (unidimensionale)

estende la nozione di lunghezza ai boreliani

$\mathcal{M} = \{S \subseteq \mathbb{R} : \lambda_*(S) = \lambda^*(S)\}$ famiglia insiemi dove misura esterna e interna coincidono

$\lambda : \mathcal{M} \to [0,+\infty]$ misura di Lebesgue def come $\lambda(S) = \lambda_*(S) = \lambda^*(S)$

$\mathcal{M}$ è σ-algebra | $\mathcal{B} \subset \mathcal{M} \subset 2^{\mathbb{R}}$; $\lambda(\mathbb{Q}) = 0$

è una estensione del concetto di lunghezza

Misura di conteggio

$\# J$ n° elementi di un gls insieme finito

fix $S \subseteq \Omega$

$\delta_S(A) : 2^{\Omega} \to [0,+\infty]$

$$\delta_S(A) = \begin{cases} \# A\cap S & A \cap S \text{ finito} \\ +\infty & \text{altrimenti} \end{cases}$$

n° finito elementi

conta n° elementi comuni ad A ad S ; è detta misura di conteggio indotta da S su Ω

dim
$\delta_S(\emptyset) = \# \emptyset \cap S = 0$
$(A_n)_{n\geq 1}$ succ disgiunta in 2^{Ω}, voglio dim numerabile additività; pongo $A = \bigcup_{n\geq 1} A_n$
- sia $A\cap S$ finito, allora $\exists\, m \in \mathbb{N} : \forall n > m \quad A_n \cap S = \emptyset$, ne segue $A\cap S = \bigcup_{i=1}^{m} (A_i \cap S)$ e $\forall n > m$ $\delta_S(A_n) = 0$. Risulta quindi $\delta_S(A) = \# A\cap S = \sum_{i=1}^{m} \# A_i \cap S = \sum_{i=1}^{m} \delta_S(A_i) = \sum_{n\geq 1} \delta_S(A_n)$
- sia $A \cap S$ infinito allora $\delta_S(A) = +\infty$. Ho due casi:
  - $A_m \cap S$ infinito per qualche m, si ha $\sum_{n\geq 1} \delta_S(A_n) \geq \delta_S(A_m) = +\infty$
  - $A_n \cap S$ finito $\forall n \geq 1$, $\exists$ una sottosuccessione $(A_{k_n})_{n\geq 1} \mid A_{k_n} \cap S \neq \emptyset\ \forall n$. Da cui
  $\sum_{n\geq 1} \delta_S(A_n) \underset{\text{localmente}}{\geq} \sum_{n\geq 1} \delta_S(A_{k_n}) \geq \underbrace{1 + 1 + \ldots}_{\text{almeno 1 elemento } \in S} = +\infty$

In ogni caso $\sum_{n\geq 1} \delta_S(A_n) = +\infty$

Proprietà della misura

Sia m una misura su $\mathcal{A}$. Allora:

1) ADDITIVITÀ $m\left(\bigcup_{i=1}^{m} A_i\right) = \sum_{i=1}^{m} m(A_i)$ se $(A_i)_{i\leq m}$ disgiunta

dim $A_1, \ldots, A_m$ disgiunti allora $A_1, \ldots, A_m, \emptyset, \ldots, \emptyset$ disgiunti $\forall n > m$ $A_n = \emptyset$. Allora $m\left(\bigcup_{i=1}^{m} A_i\right) = m\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n\geq 1} m(A_n) = \sum_{i=1}^{m} m(A_i)$

2) $m(A_1 \cup A_2) + m(A_1 \cap A_2) = m(A_1) + m(A_2)$

dim, $A_1 \cup A_2 = (A_1 \setminus A_2) \cup A_2$, $m(A_1 \cup A_2) = m(A_1 \setminus A_2) + m(A_2)$, $m(A_1 \cup A_2) + \underline{m(A_1 \cap A_2)} = m(A_1 \setminus A_2) + m(A_2) + \underbrace{m(A_1 \cap A_2)}$ $= m(A_1) + m(A_2)$

3) $m(A_2 \setminus A_1) = m(A_2) - m(A_1)$ se $A_1 \subseteq A_2$ e $m(A_1) < +\infty$

dim, $m(A_1 \cup A_2) + m(A_1 \cap A_2) = m(A_1) + m(A_2)$, da cui segue $m(A_2 \setminus A_1) + m(A_1) = m(A_2)$ da cui $m(A_2 \setminus A_1) = m(A_2) - m(A_1)$

4) MONOTONIA $m(A_1) \le m(A_2)$ se $A_1 \subseteq A_2$

dim, $\underbrace{m(A_2 \setminus A_1)}_{\ge 0} = m(A_2) - m(A_1)$ da cui $m(A_2) \ge m(A_1)$

5) SUBADDITIVITÀ $m\left(\bigcup_{i \in I} A_i\right) \le \sum_{i \in I} m(A_i)$ I discreto

dim, I discreto + numerabile: $A_1, \dots, A_n, \dots \subseteq \mathbb{R}$, $\bigcup_{i \in I} A_i = A_1 \cup A_2 / A_1 \cup A_3 / A_2 \cup A_1 \cup \dots \cup A_n / A_1 \cup \dots \cup A_{n-1} \cup \dots$

$m\left(\bigcup_{i \in I} A_i\right) = m(A_1) + \underbrace{m(A_2 \setminus A_1)}_{\le m(A_2)} + \underbrace{m(A_3 \setminus A_2 \cup A_1)}_{\le m(A_3)} + \dots + \underbrace{m(A_n \setminus A_1 \cup \dots \cup A_{n-1})}_{\le m(A_n)} + \dots$

$\le \sum_{i \in I} m(A_i)$

6) CONTINUITÀ DAL BASSO $m(A_n) \uparrow m\left(\bigcup_{n \ge 1} A_n\right)$ se $(A_n)_{n \ge 1}$ succ non decrescente

ossia $\lim_{n \to +\infty} m(A_n) = m\left(\bigcup_{n \ge 1} A_n\right)$

dim, $(A_n)_{n \ge 1}$ non decrescente, dunque $m(A_n)$ non decrescente. Ho due casi:

- se $\exists K \mid m(A_K) = +\infty$ allora $\forall n \ge K$ $m(A_n) = +\infty$, allora dalla monotonia $m\left(\bigcup_{n \ge 1} A_n\right) = +\infty$ da cui $\lim_{n \to +\infty} m(A_n) = m\left(\bigcup_{n \ge 1} A_n\right)$
- $\forall n$ $m(A_n)$ finito, allora $m\left(\bigcup_{n \ge 1} A_n\right) = m(A_1) + m(A_2 \setminus A_1) + \dots + m(A_n \setminus A_{n-1}) + \dots$ $= m(A_1) + \sum_{n \ge 1} m(A_{n+1} \setminus A_n)$ da df serie $= \lim_{n \to +\infty} m(A_1) + \sum_{i=1}^{n} m(A_{i+1} \setminus A_i) =$ $= \lim_{n \to +\infty} m\left(A_1 \cup \bigcup_{i \ge 1}^{n} A_{i+1} \setminus A_i\right) = \lim_{n \to +\infty} m(A_{n+1}) = \lim_{n \to +\infty} m(A_n)$.

7) CONTINUITÀ DALL'ALTO $m(A_n) \downarrow m\left(\bigcap_{n \ge 1} A_n\right)$ se $(A_n)_{n \ge 1}$ non crescente e $m(A_{\bar{n}}) < +\infty$ per qualche n

dim, ossia $\lim_{n \to +\infty} m(A_n) = m\left(\bigcap_{n \ge 1} A_n\right)$

$\forall n \ge \bar{n}$ $m(A_n) \le +\infty$ perché $(A_n)_{n \ge 1}$ non crescente. La succ $(A_{\bar{n}} \setminus A_n)_{n \ge 1}$ non decrescente ← non compl di quit decrescente

$m(A_{\bar{n}}) - m\left(\bigcap_{n \ge 1} A_n\right) = m\left(A_{\bar{n}} \setminus \bigcap_{n \ge 1} A_n\right) = m\left(A_{\bar{n}} \cap \left(\bigcup_{n \ge 1} A_n^c\right)\right) = m\left(\bigcup_{n \ge 1} (A_{\bar{n}} \cap A_n^c)\right) = m\left(\bigcup_{n \ge 1} A_{\bar{n}} \setminus A_n\right)$

$\overset{6)}{=} \lim_{n \to +\infty} m(A_{\bar{n}} \setminus A_n) = \lim_{n \to +\infty} m(A_{\bar{n}}) - m(A_n) = m(A_{\bar{n}}) - \lim_{n \to +\infty} m(A_n)$

togliendo $m(A_{\bar{n}})$ perché finito ho $m\left(\bigcap_{n \ge 1} A_n\right) = \lim_{n \to +\infty} m(A_n)$

8) FORMULA DI INCLUSIONE ESCLUSIONE $m\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{\emptyset \ne J \subseteq \{1, \dots, n\}} (-1)^{\#J - 1} m\left(\bigcap_{j \in J} A_j\right)$, con $m(A_i) < +\infty$ $i = 1, \dots, n$

9) Sia $m(A_i) = 0$ $\forall i \in I$, I discreto. Allora $m\left(\bigcup_{i \in I} A_i\right) = 0$

dim, $\sum_{i \in I} m(A_i) = 0$, allora anche $m\left(\bigcup_{i \in I} A_i\right) \le 0$ $m\left(\bigcup_{i \in I} A_i\right) \ge 0$

10) $m(\Omega) < +\infty$, $m(A_i) = m(\Omega)$ $\forall i \in I$, I finito allora $m(\bigcap_{i \in I} A_i) = m(\Omega)$

dim
$m(\Omega \setminus A_i) = m(A_i^c) \overset{3)}{=} m(\Omega) - m(A_i) = 0$ $\forall i \in I$. Posto $A = \bigcap_{i \in I} A_i$ da 3) $m(A^c) = m(\bigcup_{i \in I} A_i^c) = 0$, dalla 3) $m(\Omega \cap (\bigcup_{i \in I} A_i^c)^c) = m(\Omega \setminus A) = m(\Omega) - m(A) = m(\Omega)$

Applicazioni misurabili

Dati due spazi di misura $(\Omega, \mathcal{A})$, $(\Omega', \mathcal{A}')$, una applicazione $f: \Omega \to \Omega'$ è $(\mathcal{A}, \mathcal{A}')$-misurabile se $f^{-1}(A') \in \mathcal{A}$ $\forall A' \in \mathcal{A}'$

ogni applicazione costante di Ω in Ω' è $(\mathcal{A}, \mathcal{A}')$-misurabile;
se $\mathcal{A} = 2^{\Omega}$ o $\mathcal{A}' = \{\emptyset, \Omega\}$ tutte le applicazioni $f: \Omega \to \Omega'$ sono $(\mathcal{A}, \mathcal{A}')$-misurabili

Teor. f applicazione $(\mathcal{A}, \mathcal{A}')$-misurabile
$\emptyset \neq S \subset \Omega$
data $\mathcal{A} \cap S \Rightarrow f|_S$ è applicazione $(\mathcal{A} \cap S, \mathcal{A}')$-misurabile

dim tesi: $\forall A' \in \mathcal{A}'$ $(f|_S)^{-1}(A') \in \mathcal{A} \cap S$
$\forall A' \in \mathcal{A}'$ $(f|_S)^{-1} = \{\omega \in S \mid f|_S(\omega) \in A'\} = \{\omega \in \Omega \mid \omega \in S \wedge f(\omega) \in A'\} = \{\omega \in \Omega \mid \omega \in S\} \cap \{\omega \in \Omega \mid f(\omega) \in A'\}$
$= S \cap f^{-1}(A') \in S \cap \mathcal{A}$ dalla $(\mathcal{A}, \mathcal{B})$-mis di f

Teor. Dato un terzo spazio di misura $(\Omega'', \mathcal{A}'')$, $f: \Omega \to \Omega'$ $(\mathcal{A}, \mathcal{A}')$-misurabile, $g: \Omega' \to \Omega''$ $(\mathcal{A}', \mathcal{A}'')$-misurabile, allora la composta $g \circ f$ è $(\mathcal{A}, \mathcal{A}'')$-misurabile

Criterio standard di misurabilità: se $\sigma(\mathcal{F}') = \mathcal{A}'$ $\Rightarrow f^{-1}(F') \in \mathcal{A}$ $\forall F' \in \mathcal{F}'$ $\to$ f è $(\mathcal{A}, \mathcal{A}')$-misurabile

dim posto $\mathcal{F}'' = \{A' \mid f^{-1}(A') \in \mathcal{A}\}$ si ha:
- $\Omega' \in \mathcal{F}''$ perché $f^{-1}(\Omega') \in \mathcal{A}$ perché $f^{-1}(\Omega') = \{\omega \mid f(\omega) \in \Omega'\} = \Omega \in \mathcal{A}$
- $F'' \in \mathcal{F}'' \Rightarrow F''^c \in \mathcal{F}''$ perché posso $F'' \mid f^{-1}(F'') \in \mathcal{A}$ segue che $[f^{-1}(F'')]^c \in \mathcal{A}$ che è $f^{-1}((F'')^c) \in \mathcal{A}$ ossia $(F'')^c \in \mathcal{A}$
- $(F''_n)_{n \geq 1} \in \mathcal{F}'' \Rightarrow (\bigcup_{n \geq 1} F''_n) \in \mathcal{F}''$ perché $\{(F''_n)_{n \geq 1} \mid f^{-1}((F''_n)_{n \geq 1}) \in \mathcal{A}\}$ e anche $\bigcup_{n \geq 1} [f^{-1}(F''_n)] \in \mathcal{A}$ che è $f^{-1}(\bigcup_{n \geq 1} F''_n) \in \mathcal{A}$ ossia $(\bigcup_{n \geq 1} F''_n) \in \mathcal{F}''$

Quindi $\mathcal{F}''$ è σ-algebra. Da $\mathcal{F}' \subset \mathcal{F}''$ ottengo $\mathcal{A}' = \sigma(\mathcal{F}') \subseteq \mathcal{F}'' \subseteq \mathcal{A}'$, $\mathcal{F}'' = \mathcal{A}'$ quindi gli elementi $A' \in \mathcal{A}'$ sono portati in $\mathcal{A}$ da f^{-1}, quindi f è $(\mathcal{A}, \mathcal{A}')$-misurabile

Siano $\mathbb{P}$ una partizione dell'evento certo e $\mathcal{F}$ una σ-algebra di eventi logicamente dipendenti da $\mathbb{P}$. Allora una variabile aleatoria X è un'applicazione di $\mathbb{P}$ in $\mathbb{R}$ tale che $X^{-1}(]-\infty, x]) = \{X \leq x\} \in \mathcal{F}$ $\forall x$ reale. Ricordato che la famiglia $\{]-\infty, x] : x \in \mathbb{R}\}$ è un sistema di generatori dei boreliani, si ha che X è $(\mathcal{F}, \mathcal{B})$-misurabile
Le variabili aleatorie sono quindi un caso particolare di funzioni misurabili

Teor.
$\forall A' \in \mathcal{A}'$ $f^{-1}(A') \in \mathcal{A} \Leftrightarrow \forall F' \in \mathcal{F}'$ $f^{-1}(F') \in \mathcal{A}$

Chiamiamo $\mathcal{A}$-Borel misurabile (in breve Borel misurabile) ogni funzione $(\mathcal{A}, \mathcal{B})$-misurabile o $(\mathcal{F}, \mathcal{B}^*)$-misurabile.

Se $f: \Omega \to \mathbb{R} \subset \mathbb{R}^*$ è $(\mathcal{A}, \mathcal{B})$-misurabile $\Rightarrow (\mathcal{A}, \mathcal{B}^*)$-misurabile

Teor.
Sia $\emptyset \neq S \in \mathcal{B}^*$ allora le funzioni $S \to \mathbb{R}^*$ aventi insieme di pt di discontinuità discreto sono $(\mathcal{B}^* \cap S)$-Borel misurabili

dim (solo per funzioni continue)

TH: f è $(\mathcal{R},\mathcal{B})$-misurabile, cioè devo dire che
$\forall B \in \mathcal{B} \Rightarrow f^{-1}(B) \in \mathcal{R}$ cioè per criterio standard, $\forall a \in \mathbb{R} \Rightarrow f^{-1}(]a,+\infty[) \in \mathcal{R}$.
posto $]a,+\infty[= W$: · se $W = \emptyset \in \mathcal{R}$
· se $W \neq \emptyset$, $\forall \omega \in W$, $f(\omega) \in]a,+\infty[$ cioè $f(\omega) > a$. Dalla apertura di W e f continua: $\exists U_\omega \in \mathbb{R} \mid f(U_\omega) \subseteq]a,+\infty[$, allora $W = \bigcup_{\omega \in W} \omega$ aperto $\Rightarrow W \in \mathcal{B}$. $f^{-1}(]a,+\infty[) \in \mathcal{R}$ $\forall a \in \mathbb{R}$ ha dà la misurabilità di f per criterio standard di $\mathcal{B}$

Teor Siano $f: \Omega \to \mathbb{R}^*$ $\mathcal{B}$-misurabile e $S \in \mathcal{B}^*$ t.c. $f(\Omega) \subseteq S$. Allora qualunque sia $g: S \to \mathbb{R}^*$ avente insieme di pt di discontinuità discreto, $g \circ f$ è $\mathcal{B}^*$-misurabile

dim. TH $\forall A'' \in \mathcal{B}''$, $(g \circ f)^{-1}(A'') \in \mathcal{R}$ $\quad \Omega \underset{f^{-1}}{\overset{f}{\rightleftarrows}} \cdot \underset{g^{-1}}{\overset{g}{\rightleftarrows}} \mathcal{B}''$, $f^{-1}(g^{-1}(A'')) \in \mathcal{R}$

Teor Siano f, g, f_n $\mathcal{B}$-misurabili. Allora $\{f=g\}, \{f<g\}, \{f>g\}, \{f \leq g\}, \{f \geq g\} \in \mathcal{R}$

dim. $\{f<g\} \in \mathcal{R}$
$\forall \omega \in \Omega$ $f(\omega) < g(\omega)$; $\mathbb{Q}$ numerabile $q_1 < q_2 < \dots < q_n < \dots$ $\exists q_n \mid f(\omega) < q_n < g(\omega)$ da cui
$\{f<g\} = \{\omega \mid f(\omega) < g(\omega)\} = \bigcup_{n \geq 1} \{\omega \mid f(\omega) < q_n < g(\omega)\} = \bigcup_{n} \{f(\omega) < q_n\} \cap \bigcup_{n \geq 1} \{q_n < g(\omega)\} =$
$= \underbrace{f^{-1}(\underbrace{]-\infty, q_n]}_{\in \mathcal{B}})}_{\in \mathcal{R}} \cap \underbrace{g^{-1}(\underbrace{]q_n, +\infty[}_{\in \mathcal{B}})}_{\in \mathcal{R}} \in \mathcal{R}$

$\{f \neq g\} \in \mathcal{R}$
$\{f \neq g\} = \underbrace{\{f<g\}}_{\in \mathcal{R}} \cup \underbrace{\{f>g\}}_{\in \mathcal{R}}$ quindi $\in (\{f \neq g\})^c$
etc

Teor Sono $(\mathcal{R},\mathcal{B})$-misurabili le funzioni $\inf_{n \geq 1} f_n$, $\sup_{n \geq 1} f_n$, $\liminf_{n \to \infty} f_n$, $\limsup_{n \to \infty} f_n$, $\lim_{n \to \infty} f_n$

Teor Appartengono ad $\mathcal{R}$ gli insiemi di convergenza
$A_{lim} = \{\omega \in \Omega : \exists \lim_{n \to \infty} f_n(\omega)\}$, $A^{\mathbb{R}}_{lim} = \{\omega \in A_{lim} : \lim_{n \to \infty} f_n(\omega) \in \mathbb{R}\}$
e $\{f_n \to f\} = \{\omega \in \Omega : f_n(\omega) \to f(\omega)\}$

Teor Sono Borel-misurabili le funzioni $f \cdot g$ e se definite ovunque $f+g$ e f/g

Teor Sia f derivabile. Allora f' è Borel-misurabile

dim. f derivabile in x_0, $\exists$ finito $f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \overset{\text{unicità}}{=} \lim_{n \to \infty} \frac{f(x_0 + 1/n) - f(x_0)}{1/n} = f'(x_0)^+$
non è misurabile x composta misurabile, denominatore è misurabile perché costante, idem numeratore e limite, quindi derivata è misurabile

Teor Sia $f_n \geq 0$, allora la serie $\sum f_n$ è funzione Borel-misurabile

$f: \Omega \to \mathbb{R}^*$ è una funzione $\mathcal{R}$-semplice (in breve semplice) se è $\mathcal{R}$-Borel misurabile e l'insieme-immagine $f(\Omega)$ è finito (n° finito valori)

Teor $I_S(\omega)$ è funzione indicatrice di $S \subset \Omega$: $I_S(\omega) = \begin{cases} 1 & \omega \in S \\ 0 & \omega \notin S \end{cases}$
è una funzione semplice se $S \in \mathcal{R}$ $\Rightarrow$ $I_S(\omega) = 1$, $I_S(\omega) = 0$

Teor $I_S(\omega)$ è una funzione $(\mathcal{R},\mathcal{B})$-misurabile $\Leftrightarrow S \in \mathcal{R}$

(1/0 prodotto di un insieme è log due $\log \mathbb{P}$)

dim/ preso qls $B \in \mathcal{B}$:

$\Rightarrow$)

$$I_S^{-1}(B) = \begin{cases} \emptyset & \text{se } 0,1 \notin B \\ S^c & \text{se } 0 \in B, 1 \notin B \\ S & \text{se } 0 \notin B, 1 \in B \\ \Omega & \text{se } 0,1 \in B \end{cases}$$

perché $\nexists \omega$ che finisca altrove
perché $I_S^{-1}(0) \in \mathcal{A}$ «x misurab» $\to \{\omega \mid \omega \notin S\} \in \mathcal{A}$
perché $I_S^{-1}(1) \in \mathcal{A}$ «x misurab» $\to \{\omega \mid \omega \in S\} \in \mathcal{A}$
perché $\forall \omega \in \Omega$ finisce in 0 o 1

$\Leftarrow$) idem: gli $\omega \mid I_S(\omega)=1$ stanno in $\mathcal{A} \to$ è misur.
$\omega \mid I_S(\omega)=0$ " " $\to$ "

Proprietà:

$I_{S^c} = 1 - I_S$ dim $I_{S^c}(\omega) = \begin{cases} 0 & \omega \notin S^c \\ 1 & \omega \in S^c \end{cases}$ e $I_S(\omega) = \begin{cases} 1 & \omega \in S \\ 0 & \omega \notin S \end{cases}$

$I_{S \cap T} = I_S \cdot I_T$ dim $I_{S\cap T}(\omega) = \begin{cases} 1 & \omega \in (S \cap T) \\ 0 & \omega \notin (S\cap T) \end{cases}$
$\to \omega \in S,\ I_S = 1$ e $\omega \in T,\ I_T = 1$
$\to \omega \notin S, \omega \in T \to I_S \cdot I_T = 0$
$\omega \in S, \omega \notin T \to I_S \cdot I_T = 0$
$\omega \notin S, \omega \notin T \to I_S \cdot I_T = 0$

$I_{S \cup T} = I_S + I_T - I_{S \cap T}$ dim $I_{S\cup T}(\omega) = \begin{cases} 1 & \omega \in S \cup T \\ 0 & \omega \notin S\cup T \end{cases}$
$\to I_S + I_T - I_{S\cap T} = 1+1-1 = 1$
$\to I_S + I_T - I_{S \cap T} = 0+0-0 = 0$

Data f semplice è posto $\{f = y\} = f^{-1}(\{y\})$ cioè $\{\omega \mid f(\omega) = y\}$ $\forall y \in \mathbb{R}^*$:

- $(\{f=y\})_{y \in f(\Omega)}$ è P finita di Ω costituita da elementi di $\mathcal{A}$ $\{f = y_1\} = f^{-1}(\{y_1\})$
  $y_1 \dots y_n$ determinazioni
- $f = \sum_{y \in f(\Omega)} y \cdot I_{\{f=y\}}$ cioè $f(\omega) = I_{\{f=y_1\}}(\omega) y_1 + \dots + I_{\{f=y_n\}}(\omega) y_n$ che comprende $y_1 \dots y_n$
  solo un indicatore per ω è verificato 1

* **Lemma fondamentale**

Sussistono le seguenti proposizioni:

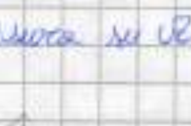
- Somme quozienti (se definiti ovunque) prodotti di funzioni semplici sono ancora funzioni semplici
- * Sia f una funzione Borel misurabile non negativa $\exists$ allora una successione $(f_n)_{n \geq 1}$ di funzioni semplici non negative a valori finiti t.c. $f_n \uparrow f$
- Sia f una funzione Borel misurabile $\exists$ allora una successione $(f_n)_{n \geq 1}$ di funzioni semplici a valori finiti t.c. $f_n \to f$ e $|f_n| \leq |f|$ $\forall n$

Integrale di Lebesgue

$f: \Omega \to \mathbb{R}^*$ funzione $\mathcal{A}$-Borel misurabile e m una misura su $\mathcal{A}$
definiamo l'integrale di Lebesgue per passi

1) funzioni semplici non negative

$f \geq 0$ semplice $\Rightarrow \int_\Omega f\, dm = \sum_{y \in f(\Omega)} y\, m(\{f = y\}) \geq 0$



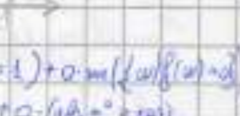
in particolare $m(A) = \int_\Omega I_A\, dm$ $\forall A \in \mathcal{A}$

se $f = I_A(\omega)$, $\forall A \in \mathcal{A}$, allora $\int_\Omega f\, dm = 1 \cdot m(\{\omega \mid f(\omega) = 1\}) + 0 \cdot m(\{\omega \mid f(\omega) = 0\})$
$\to A \in \mathcal{A}$ (all'inizio)
$= m(\{I_A(\omega) = 1\}) + 0 \cdot (\text{può} = \infty + \infty)$
$= m(\{\omega \in A\}) + 0 = m(A)$

2) funzioni Borel misurabili non negative

$f \geq 0 \Rightarrow \int_\Omega f\, dm = \sup \left\{ \int_\Omega g\, dm : 0 \leq g \leq f,\ e\ g \text{ funzione semplice} \right\} \geq 0$

prendo la + grande funzione f_n con area $\leq$ area f



dove la costante nulla c'è $\to \int_\Omega 0 = 0 \cdot$ qls cosa $= 0$

Il prossimo risultato consente il passaggio del limite sotto il segno di integrale nel caso di successioni non decrescenti di funzioni $\mathcal{B}$-misurabili, non negative

* **Teor convergenza monotona** Sia $(f_n)_{n \geq 1}$ funzioni $\mathcal{B}$-misurabili non negative e $f_n \uparrow f$,
allora $\int_\Omega f_n\, dm \uparrow \int_\Omega f\, dm$ cioè $\lim_{n \to +\infty} \int_\Omega f_n\, dm = \int_\Omega \underbrace{\lim_{n \to +\infty} f_n}_{f}\, dm$

Conseguentemente per il lemma fondamentale l'integrale di una funzione Borel-misurabile non negativa è il limite di una stazionaria successione non decrescente di integrali di funzioni semplici non negative a valori finiti

lemma: f B-mis, non negativa $\exists f_n \geq 0$ semplici : $f_n \uparrow f$

theor conv monotona: $(g_n)_{n \geq 1}$ B-mis, ≥ 0, $g_n \uparrow g$, $\int_\Omega g_n \uparrow \int_\Omega g$

Poniamo per definizione 3° passo

3) funzioni Borel-misurabili qualsiasi

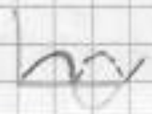
$f^+ = \max(0, f)$ parte positiva $\quad f^- = \max(0, -f)$ parte negativa

f^+, f^- sono B-misurabili se f B-misurabile

dim
- $g(x) = \max(0, x)$ è f continua, quindi $f^+ = (g \circ f)(x)$ perché composta di Borel-mis e B-misurabili
- $g(x) = \max(0, -x)$ è f continua, quindi $f^- = (g \circ f)(x)$, perché composta di B-mis. tot

$f = f^+ - f^-$, $\quad |f| = f^+ + f^-$

$|f|$



$$(fg)^+ = \begin{cases} f \cdot g^+ & \text{se } f \geq 0 \\ f \cdot g^- & \text{se } f \leq 0 \end{cases} \qquad (fg)^- = \begin{cases} f \cdot g^- & \text{se } f \geq 0 \\ -fg & \text{se } f \leq 0 \end{cases}$$

dim

$f \geq 0 \to f \cdot g^+ = \max(0, g) \cdot f = \max(0, f \cdot g) = (fg)^+$

$f \leq 0 \to f \cdot g^- = \max(0, -g) \cdot f = \max(0, f \cdot g) = (fg)^+$

dim

$f \geq 0 \to f \cdot g^- = \max(0, -g) f = \max(-gf, 0) = (fg)^-$

$f \leq 0 \to f \cdot g^+ = -f \max(0, g) = \max(0, -fg) = (fg)^-$

Sia f qls (B-mis)

supposto $\int_\Omega f^+ dm - \int_\Omega f^- dm$ non siano entrambi infiniti, diciamo **m-sommabile** (in breve **sommabile**) la funzione f e poniamo:

$$\int_\Omega f\, dm = \int_\Omega f^+ dm - \int_\Omega f^- dm \quad \text{che sarà} \begin{cases} +\infty & \text{se } \int_\Omega f^+ dm = +\infty \\ -\infty & \text{se } \int_\Omega f^- dm = +\infty \\ \text{finito} & \text{se } \int_\Omega f^+ dm, \int_\Omega f^- dm \text{ finiti} \end{cases}$$

Nell'ultimo caso diremo che f è **m-integrabile** (o in breve **integrabile**)

La dimostrazione standard è:
- constatare che la proprietà sussiste nell'ambito delle funzioni semplici non negative
- verificare che rimane valida anche nell'ambito delle funzioni Borel-misurabili non negative tramite il passo precedente, il lemma fondamentale e il teor convergenza monotona
- provare che rimane valida anche nell'ambito delle funzioni sommabili ricorrendo al passo precedente e alle parti positiva e negativa

Sia $A \in \mathcal{A}$, allora $f \cdot I_A$ è una funzione Borel-misurabile

Inoltre tenuto conto delle disuguaglianze $(f \cdot I_A)^+ = f^+ \cdot I_A \leq f^+$, $(f \cdot I_A)^- = f^- I_A \leq f^-$ risulta $f \cdot I_A$ sommabile (integrabile) se f è sommabile (integrabile). Conseguentemente nel caso di sommabilità della funzione f poniamo:

$$\int_A f\, dm = \int_\Omega f \cdot I_A\, dm \quad \text{e}$$

diciamo $\int_A f\, dm$ l'integrale di Lebesgue di f su A.

Giustificazione della notazione:

se $m = \lambda$, $\Omega = \mathbb{R}$ $\quad \int_\mathbb{R} f\, d\lambda = \int_\mathbb{R} f\, dx$, $\int_{[a,b]} f\, d\lambda = \int_a^b f(x)\, dx$

se f limitata e integrabile secondo Riemann su $[a,b]$ allora integrale L e R coincidono

Proprietà elementari dell'integrale di Lebesgue

Sussistono le seguenti proposizioni

1) LINEARITÀ RISPETTO ALLA FUNZIONE: Siano f, g funzioni sommabili e α, β numeri reali t.c. $\alpha f + \beta g$ risulta definita ovunque allora $\int_\Omega (\alpha f + \beta g)\, dm = \alpha \int_\Omega f + \beta \int_\Omega g\, dm$ ogni volta che non sono infiniti di segno opposto gli addendi che compaiono al secondo membro dell'uguaglianza

2) LINEARITÀ RISPETTO ALLA MISURA: Siano m' una misura su $\mathbb{R}$ e f una funzione m-sommabile e m'-sommabile, allora qualunque siano i numeri reali non negativi α e α' si ha $\int_\Omega f\, d(\alpha m + \alpha' m') = \alpha \int_\Omega f\, dm + \alpha' \int_\Omega f\, dm'$ ogni volta che non sono infiniti di segno opposto gli integrali che compaiono al secondo membro dell'uguaglianza

dim

— Siano m' una misura su $\mathbb{R}$, α e $\alpha' \geq 0$ reali, $m'' = \alpha m + \alpha' m'$ con integrali a 2° membro non di segno opposto

- $f \geq 0$ semplice, $\int_\Omega f\, dm'' = \sum_{y \in f(\Omega)} y \cdot m''(\{y \in f(\Omega)\}) = \sum_{y \in f(\Omega)} y \left[\alpha \cdot m(\{f = y\}) + \alpha' m'(\{f = y\})\right] =$
$= \alpha \sum_{y \in f(\Omega)} y\, m(\{f = y\}) + \alpha' \sum_{y \in f(\Omega)} y\, m'(\{f = y\}) = \alpha \int_\Omega f\, dm + \alpha' \int_\Omega f\, dm'$

- $f \geq 0$ B-mis., $\exists (f_n)_{n \geq 1}$ di f semplici non negative $f_n \uparrow f$, $\int_\Omega f\, dm'' =$
$= \lim_{n \to \infty} \int_\Omega f_n\, dm'' = \lim_{n \to \infty} \left(\alpha \int_\Omega f_n\, dm + \alpha' \int_\Omega f_n\, dm'\right) = \alpha \lim_{n \to \infty} \int_\Omega f_n\, dm + \alpha' \lim_{n \to \infty} \int_\Omega f_n\, dm' =$
$= \alpha \int_\Omega f\, dm + \alpha' \int_\Omega f\, dm'$

- f qls B-mis e sommabile risp m e m'
$\int_\Omega f^+ dm'' = \alpha \int_\Omega f^+ dm + \alpha' \int_\Omega f^+ dm'$
$\int_\Omega f^- dm'' = \alpha \int_\Omega f^- dm + \alpha' \int_\Omega f^- dm'$
$\Rightarrow \int_\Omega f^+ dm'' - \int_\Omega f^- dm''$ non infiniti segno opposto (i casi non accettati sono esclusi dalla m, m' sommabilità)

Inoltre dalle due relazioni discende $\int_\Omega f\, dm'' = \left(\alpha \int_\Omega f^+ dm + \alpha' \int_\Omega f^+ dm'\right) - \left(\alpha \int_\Omega f^- dm + \alpha' \int_\Omega f^- dm'\right)$
$= \alpha \int_\Omega f\, dm + \alpha' \int_\Omega f\, dm'$

3) MONOTONIA: Siano f, g sommabili t.c. $f \leq g$ allora $\int_\Omega f\, dm \leq \int_\Omega g\, dm$

4) ADDITIVITÀ: Sia f una funzione sommabile, allora $\int_{A_1 \cup A_2} f\, dm = \int_{A_1} f\, dm + \int_{A_2} f\, dm$, $A_1 \cap A_2 = \emptyset$

5) ANNULLAMENTO SUGLI INSIEMI DI MISURA NULLA: Sia f una funzione sommabile e $m(A) = 0$ allora $\int_A f\, dm = 0$

dim

- $f \geq 0$ semplice. Posto $f = \sum_{y \in f(\Omega)} y\, I_{\{f = y\}}$, dalla monotonia della misura $0 \leq \int_\Omega f \cdot I_A\, dm =$
$= \int_\Omega \left[\sum_{y \in f(\Omega)} y\, I_{\{f = y\}}\right] I_A\, dm = \int_\Omega \left[\sum_{y \in f(\Omega)} y\, I_{\{f = y\} \cap A}\right] dm = \sum_{y \in f(\Omega)} y\, m(\underbrace{\{f = y\} \cap A}_{\subseteq A}) \leq \sum_{y \in f(\Omega)} y \cdot m(A) = 0$

- $f \geq 0$ B-misurabile. Dal lemma fondamentale $0 \leq f_n \uparrow f$, $0 \leq f_n I_A \uparrow f \cdot I_A$ con $f \cdot I_A$ ancora semplice, B-misur. non negativa. Dal teorema conv. monotona ho che
$0 = \int_\Omega f_n I_A\, dm \to \int_\Omega f I_A\, dm = \int_A f\, dm$

- f arbitraria: $\int_A f\, dm = \int_\Omega f I_A\, dm = \int_\Omega (f I_A)^+ dm - \int_\Omega (f \cdot I_A)^- dm = \underbrace{\int_A f^+ dm}_{f^+ \geq 0,\ m(A) = 0} - \underbrace{\int_A f^- dm}_{f^- \geq 0,\ m(A) = 0} = 0$

$(f \cdot I_A)^{\pm} = f^{\pm} I_A$ — passo precedente

Teor

f, g funzioni B-misurabili, allora:

- f sommabile se: $g \leq f$ con g sommabile t.c. $-\infty < \int_\Omega g\, dm$
(1) $f \leq g$ con g sommabile t.c. $\int_\Omega g\, dm < +\infty$

dim

(1) se $f \leq g$, g sommabile, $\int_\Omega g\, dm < +\infty$, $f^+ = \max(0, f) \leq \max(0, g) = g^+$ da cui $\int_\Omega f^+ dm \leq \int_\Omega g^+ dm < +\infty$ perché se solo se $\int_\Omega g^+ dm = +\infty$ allora $\int_\Omega g\, dm = +\infty$, già escluso dalla $\int_\Omega f^+ dm < +\infty$ deriva $\int_\Omega f < +\infty$ (ossia) perché $\int_\Omega f\, dm = +\infty$ sse $\int_\Omega f^+ = +\infty$

- f integrabile $\iff$ $|f|$ integrabile

dim

$\int_\Omega |f|\, dm = \int_\Omega f^+ dm + \int_\Omega f^- dm \Rightarrow$ perché entrambi finiti $\to \int_\Omega f\, dm$ integrabile

- f integrabile se $|f| \leq g$ con g integrabile

dim: da $0 \leq |f| \leq g$ segue $\int_\Omega |f|\, dm \leq \int_\Omega g\, dm < +\infty$, cioè $\int_\Omega f^+ dm + \int_\Omega f^- dm < +\infty$, entrambe finite

- $f \cdot g$ integrabile se f^2, g^2 integrabili

dim

$\int_\Omega f^2 + g^2 \, dm = \int f^2 + \int g^2 < +\infty$; $0 \leq (|f| - |g|)^2 = f^2 - 2|f| \cdot |g| + g^2$, $2|f||g| \leq f^2 + g^2$, $|f| \cdot |g| \leq \frac{f^2 + g^2}{2} \leq f^2 + g^2$ allora $0 \leq \int_\Omega |f| \cdot |g| \leq \int_\Omega f^2 + g^2 < +\infty$ da cui $0 \leq \int_\Omega |f||g| < +\infty$ allora $|f| \cdot |g|$ integrabile e $|f \cdot g|$ anche integrabile

- $m(\Omega) < +\infty$ allora f è integrabile se f^2 è integrabile. Inoltre (allora) ogni funzione limitata è integrabile

dim

- $1 = \int_\Omega (I_\Omega)^2 dm = \int_\Omega I_\Omega dm = m(\Omega) (< +\infty)$, allora $(I_\Omega)^2$ è integrabile. Dalla precedente: f^2 integrabile e $(I_\Omega)^2$ integrabile ho $(f \cdot I_\Omega)$ integrabile cioè $\int_\Omega I_\Omega f \, dm = \int_\Omega 1 \cdot f \, dm = \int_\Omega f \, dm$ integrabile.
- se f limitata, allora $\exists\, n \mid |f| \leq n = n \cdot I_\Omega$ e $\int_\Omega n \cdot I_\Omega \, dm = n \cdot m(\Omega) < +\infty$ dalla 3) (nota bile) $|f| \leq n \cdot I_\Omega$ integrabile ho che $|f|$ integrabile e quindi f integrabile

Chiamiamo **m-trascurabile** (in breve **trascurabile**) ogni sottoinsieme S di Ω per il quale esiste un insieme $A \supseteq S$ tale che $m(A) = 0$.
Sono trascurabili, oltre all'insieme vuoto, i sottoinsiemi di un insieme trascurabile e l'unione di una famiglia discreta di insiemi trascurabili. Inoltre per la monotonia della misura, un elemento di $\mathcal{A}$ è trascurabile se e solo se è di misura nulla.

Diremo che una proprietà P o una relazione binaria $\prec$ riguardanti applicazioni di dominio Ω sussiste **m-quasi ovunque** (in breve **quasi ovunque**) se il suo campo di validità include il complementare di un insieme trascurabile. Per indicare tale situazione useremo rispettivamente le notazioni P (m-q.o.) e $\prec$ (m-q.o.).

f finita (m-q.o.) se l'insieme $\{|f| = +\infty\}$ è trascurabile
$f \leq g$ (m-q.o.) se l'insieme $\{f > g\}$ è trascurabile
$f \geq g$ (m-q.o.) se l'insieme $\{f < g\}$ è trascurabile
$f = g$ (m-q.o.) se l'insieme $\{f \neq g\}$ è trascurabile
$\exists \lim_n f_n$ (m-q.o.) se l'insieme complementare è trascurabile
$f_n \xrightarrow{n \to \infty} f$ (m-q.o.) se l'insieme $\{f_n \to f\}^c$ è trascurabile
$f_n \uparrow f$ (m-q.o.) se l'insieme $\{\omega \in \Omega \mid \neg(f_n(\omega) \uparrow f(\omega))\}$ è trascurabile

Teor

Sia f una funzione sommabile allora

1) $\int_\Omega g \, dm \leq \int_\Omega f \, dm$ se $g \leq f$ (m-q.o.) e g sommabile ✓ (basato) MONOTONIA

dim

$A_1 = \{g \leq f\}$, $g \cdot I_{A_1} \leq f \cdot I_{A_1}$ e $m(A_1^c) = 0$. Poiché $(g\, I_{A_1})^+ \leq (f \cdot I_{A_1})^+$ e $(g\, I_{A_1})^- \geq (f \cdot I_{A_1})^-$ allora

$$\int_\Omega g \, dm = \int_{A_1} g \, dm = \int_\Omega (g\, I_{A_1})^+ dm - \int_\Omega (g\, I_{A_1})^- dm \leq \int_\Omega (f \cdot I_{A_1})^+ dm - \int_\Omega (f \cdot I_{A_1})^- dm$$
$$= \int_\Omega f\, I_{A_1} dm = \int_\Omega f \, dm$$

2) $\int_\Omega g \, dm = \int_\Omega f \, dm$ se $f = g$ (m-q.o.) e g $(\mathcal{E}, \mathcal{B})$-misurabile (non arrotondato)

dim

g $(\mathcal{E}, \mathcal{B})$-misurabile e $m(\{f \neq g\}) = 0$, allora $\{f \leq g\}$ m-q.o. e $\{f \geq g\}$ m-q.o.
dalla 1) segue $\int g \leq \int f$ e $\int f \leq \int g$ se g sommabile. Provo allora la sommabilità di g: f sommabile $\to \int_\Omega f^+ dm < +\infty \to f^+ = g^+$ (m-q.o.) $\to \int_\Omega f^+ dm = \int_\Omega g^+ dm < +\infty$
$\to \int_\Omega f^- dm < +\infty$, $f^- = g^-$ (m-q.o.) $\to \int_\Omega f^- dm = \int_\Omega g^- dm < +\infty$
cioè dalla sommabilità di f gli integrali delle relative parti positiva e negativa non sono entrambi infiniti. Ad es. quello della parte positiva lo supponiamo finito, e provo che $\int_\Omega g^+ \leq \int_\Omega f^+$; a tal fine dalla 1) basta constatare che $f^+ = g^+$ m-q.o. Dalla $\{f^+ \neq g^+\} = \{\max(0, f) \neq \max(0, g)\} \subseteq \{f \neq g\}$ dalla monotonia misura, $m(\{f^+ \neq g^+\}) \leq m(\{f \neq g\}) = 0$.

3) $\left| \int_\Omega f \, dm \right| \leq \int_\Omega |f| \, dm$

dim $-|f| \leq f \leq |f|$ per monotonia (m-q.o) e linearità, $-\int_\Omega |f| dm \leq \int_\Omega f dm \leq \int_\Omega |f| dm$ " $-A \leq B \leq A \Leftrightarrow |B| \leq A$ "
allora $|\int_\Omega f dm| \leq \int_\Omega |f| dm$

4) $m(A) \inf f(A) \leq \int_A f dm \leq m(A) \sup f(A)$ $\forall A$ TEOR MEDIA INTEGRALE

dim

$I_A(\omega) \cdot \inf_{\omega \in A} f(\omega) \leq f(\omega) I_A(\omega) \leq I_A(\omega) \cdot \sup_{\omega \in A} f(\omega)$

$\int_\Omega \inf f(A) I_A dm \leq \int_\Omega I_A f dm \leq \int_\Omega \sup f(A) I_A dm$, perché cost $\inf f(A) m(A) \leq \int_A f dm \leq \sup f(A) m(A)$

5) se f misurabile e $f \geq 0$ e $\int_\Omega f dm = 0$, allora $f = 0$ (m-q.o.)

dim

$A \stackrel{\circ}{=} \{\omega \mid f(\omega) > 0\}$, (x assurdo) $m(A) > 0$ $(A_n)_{n \geq 1} = \{f > 1/n\}$ è succ non decrescente al $\uparrow n \Rightarrow \{f > 1/n\}$ si allarga
$(A_n)_{n \geq 1} = \{f > \frac{1}{n}\}$ è succ non decrescente,
dalla continuità del basso $m(A_n) \uparrow m(\bigcup_{n \geq 1} A_n)$, allora $\exists k \mid m(A_k) > 0$. Data $g = I_{A_k}/k$
allora $g \leq f$ perché $\begin{cases} \omega \notin A_k, \text{ allora } g = 0, f \geq 0 \\ \omega \in A_k \quad g = 1/k, f \geq 0 \end{cases}$, dalla monotonia m-q.o.
$\int_\Omega g dm \leq \int_\Omega f dm = 0$ allora $\int g dm = 0 = \frac{1}{k} \int_\Omega I_{A_k} dm = \frac{1}{k} m(A_k)$ contraddizione

6) se $m(A) > 0$ e $f(\omega) > 0$ $\forall \omega \in A$, allora $\int_A f dm > 0$

dim

sia $\int_\Omega f I_A dm = \int_A f dm = 0$ (x assurdo). Poiché $f I_A \geq 0$ si ha $\stackrel{5)}{}$ $f I_A = 0$ (m-q.o)
Osservato che $A = \{f I_A \neq 0\}$ ottengo la contraddizione $(0 < m(A) \underset{(\text{x5})}{=} m(\{f I_A \neq 0\}) = 0$ perché
$f I_A = 0$ (m-q.o)

7) Se f è integrabile, allora f è finita quasi ovunque

dim

sia f integrabile, allora $|f|$ è integrabile. Posto $A = \{|f| = +\infty\}$, e sia (x assurdo)
$m(A) > 0$. Considerata la funzione semplice $g = +\infty \cdot I_A$, otteniamo $g \leq |f|$ e
quindi la contraddizione
$+\infty = +\infty\, m(A) = \int_\Omega g dm \leq \int_\Omega |f| dm < +\infty$, ($|f|$ integrabile)

Ulteriori proprietà di convergenza

Teor della convergenza dominata $(f_n)_{n \geq 1}$ $(\mathcal{E}, \mathcal{B})$-misurabili, $f_n \to f$. Inoltre sia
g integrabile, $|f_n| \leq g$ (m-q.o) $\forall n$. Allora f, f_n $(n \geq 1)$ sono integrabili
e si ha $\int_\Omega f_n dm \to \int_\Omega f dm$

(diciamo uguale alla conv. convergenza di Analisi I)

Teor d'integrazione per serie Siano $(f_n)_{n \geq 1}$ $\mathcal{B}$-misurabili non negative
e $(a_n)_{n \geq 1}$ reali non negativi. Allora $\int_\Omega (\sum_{n \geq 1} a_n f_n) dm = \sum_{n \geq 1} a_n \int_\Omega f_n dm$

Integrale di Lebesgue e serie numeriche Sia S insieme discreto avente
$N \leq +\infty$ elementi e sia $f : \Omega \to \mathbb{R}^*$ una funzione δ_S-misurabile. Considerata
allora una numerazione $s_1, s_2, \dots$ di S si ha
$$\int_\Omega f \, d\delta_S = \sum_{k=1}^{N} f^+(s_k) - \sum_{k=1}^{N} f^-(s_k) = \sum_{k=1}^{N} f^+(s_k) + \sum_{k=1}^{N} (-f(s_k))$$

dim

- $f \geq 0$, $g = \sum_{n \geq 1} f(s_n) I_{\{s_n\}}$ allora $g = f$ (δ_S-q.o.) e quindi per teor integr per serie $\int_\Omega f d\delta_S = \int_\Omega g \, d\delta_S = \int_\Omega \sum_{k=1}^{N} f(s_k) I_{\{s_k\}} d\delta_S = \sum_{k=1}^{N} f(s_k) \int_\Omega I_{\{s_k\}} d\delta_S =$
$= \sum_{k=1}^{N} f(s_k) \delta_S(\{s_k\}) = \sum_{k=1}^{N} f(s_k)$
- f qls Osservo che $\int_\Omega f^+ d\delta_S = \sum_{k=1}^{N} f^+(s_k)$, $\int_\Omega f^- d\delta_S = \sum_{k=1}^{N} f^-(s_k)$ da cui, dalla def di S, ho la tesi

Se $N = +\infty$ l'integrale può essere calcolato sommando le serie dei valori, rispettivamente,
positivi e negativi della restrizione $f|_S$. Sussiste l'uguaglianza ogni volta che
la serie numerica $\sum_{n \geq 1} f(s_n)$ è permutabile, come avviene quando f è di
segno costante (serie a termini positivi) oppure è δ_S-integrabile (serie assolutamente convergente) l'uguaglianza è $\int_\Omega f d\delta_S = \sum_{n \geq 1} f(s_n)$

$\sum^{+\infty} a_n$ conv assolut $\Leftrightarrow \sum^{+\infty} |a_n| < +\infty$

Probabilità 1

Eventi variabili aleatorie, enti aleatori

$\Omega \neq \emptyset$ evento certo

$\mathcal{A}$ σ-algebra su Ω degli eventi di interesse

Chiamiamo
- caso elementare di Ω ogni elemento di Ω
- evento di Ω ogni elemento di $\mathcal{A}$
- probabilità sugli eventi di Ω ogni misura P su $\mathcal{A}$ t.c. $P(\Omega)=1$
- variabile aleatoria su Ω ogni funzione $(\mathcal{A},\mathcal{B})$-misurabile a valori nella retta reale
- variabile aleatoria estesa su Ω ogni funzione $(\mathcal{A},\mathcal{B})$-misurabile a valori nella retta reale ampliata

Considerato inoltre uno spazio di misura $(\mathcal{X}, \mathcal{N})$, chiamiamo
- ente aleatorio su Ω a valori in $\mathcal{X}$ ogni applicazione di Ω in $\mathcal{X}$ tale che sia $(\mathcal{A},\mathcal{N})$-misurabile

ente aleatorio: $\underset{\mathcal{A}}{\Omega} \to \underset{\mathcal{N}}{\mathcal{X}}$ $(\mathcal{A},\mathcal{N})$-misurabile

variabile aleatoria: $\underset{\mathcal{A}}{\Omega} \to \underset{\mathcal{B}}{\mathbb{R}}$ $(\mathcal{A},\mathcal{B})$-misurabile

Se $\mathcal{X} = \mathbb{R}$ e $\mathcal{N} = \mathcal{B}$ ho una variabile aleatoria, caso particolare di ente aleatorio

Dato un ente aleatorio qualsiasi X, useremo le notazioni
$P(X\in A)$ per $P(\{X\in A\})$
$P(X\le x)$ per $P(\{X\le x\})$
$P(X-Y\ge 0)$ per $P(\{X-Y\ge 0\})$
Inoltre abbrevieremo con il solito "v.a." il nome "variabile aleatoria" intesa sia al singolare che al plurale

Legge di un ente aleatorio

Dato X ente aleatorio a valori in $\mathcal{X}$, quindi $(\mathcal{A},\mathcal{N})$-misurabile, consideriamo l'applicazione $P_X : \mathcal{N} \to [0,1]$:

$$P_X(A) = P(\{X\in A\}) \qquad \underset{\mathcal{A}}{\Omega} \xrightarrow{X} \mathcal{X}, \quad \mathcal{N} \xrightarrow{P_X} [0,1]$$

è una probabilità su $\mathcal{N}$

dim
- $P_X(\mathcal{X}) = P(X\in\mathcal{X}) = P(\Omega) = 1$
- $P_X(\bigcup_{n\ge1} A_n) = P(\{X \in \bigcup_{n\ge1} A_n\}) = P(\{X^{-1}(\bigcup_{n\ge1} A_n)\}) = P(\bigcup_{n\ge1}\{X^{-1}(A_n)\}) =$ (disgiunti) $= \sum_{n\ge1} P(\{X^{-1}(A_n)\}) = \sum_{n\ge1} P_X(A_n)$ notato che $(\{X\in A_n\})_{n\ge1}$ è una successione disgiunta

La probabilità P_X si chiama legge (o distribuzione) di X. È la probabilità immagine P_X di P mediante X. P_X associa $\forall A\in\mathcal{N}$ la probabilità che X assuma come determinazione un elemento dell'insieme A.

$$(P\circ X^{-1})(A) = P(X\in A) = P(\underbrace{X^{-1}(A)}_{\in\mathcal{A}}) = P_X(A)$$

$\Omega \xrightarrow{X} \mathcal{X}$, $\mathcal{A} \circlearrowleft P$, $\mathcal{N} \circlearrowleft P_X$ — l'ente aleatorio trasporta la probabilità definita su $\mathcal{A}$, nella probabilità definita su $\mathcal{N}$

Se X è v.a. posso considerare la funzione di ripartizione di X, cioè la funzione $F_X : \mathbb{R} \to [0,1]$, $F_X(x) = P(X\le x) = P_X(]-\infty,x])$ legge della variabile aleatoria calcolata sulle semirette

$$\omega \to \underbrace{X(\omega)}_{\in\mathbb{R}} \to \underbrace{F(X(\omega))}_{[0,1]}$$

Teor Proprietà $S_x =]-\infty, x]$

- $F_X(x) \le F_X(x')$ se $x \le x'$ MONOTONIA

  dim. da monotonia della misura

- $F(-\infty) = \lim_{x \to -\infty} F_X(x) = 0$

  dim. $S_x =]-\infty, x]$ $\forall x \in \mathbb{R}$, $(S_{-n})_{n \ge 1}$ non crescente, $\Rightarrow \bigcap_{n \ge 1} (S_{-n}) = \emptyset$, da continuità dall'alto della misura $F_X(\bigcap_{n \ge 1} (S_{-n})) = \lim_{n \to +\infty} F_X(S_{-n}) = F_X(\emptyset) = 0$

- $F_X(+\infty) = \lim_{x \to +\infty} F_X(x) = 1$

  dim. $S_x =]-\infty, x]$ $\forall x \in \mathbb{R}$, $(S_n)_{n \ge 1}$ non decrescente, da continuità dal basso $\bigcup_{n \ge 1} S_n = \mathbb{R}$, $F_X(\bigcup_{n \ge 1} S_n) = \lim_{n \to +\infty} F_X(S_n) = F_X(\mathbb{R}) = 1$

- $F_X(x^+) = \lim_{t \to x^+} F_X(t) = F_X(x)$

  dim. $S_x \stackrel{t \to x^+}{=}]-\infty, x]$ $\forall x \in \mathbb{R}$, $(S_{x+1/n})_{n \ge 1}$ non crescente, da continuità dall'alto $\Rightarrow \bigcap_{n \ge 1} (S_{x+1/n}) = S_x$ $F_X(\bigcap_{n \ge 1} S_{x+1/n}) = F_X(S_x) = P_X(]-\infty, x]) = F_X(x)$

- $F_X(x^-) = \lim_{t \to x^-} F_X(t) = P(X < x) = P_X(]-\infty, x[)$

  dim. $S_x \stackrel{t \to x^-}{=}]-\infty, x]$ $\forall x \in \mathbb{R}$, $S_{x-1/n}$ non decrescente e da continuità dal basso allora $F_X(\bigcup_{n \ge 1} S_{x-1/n}) = F_X(S_x) = P(X \in]-\infty, x[) = P(X < x) = P_X(]-\infty, x[)$

  crescente per $t \to x$

- $F_X(x^+) - F_X(x^-) = P(X = x)$

  dim. segue dalle due precedenti

Inoltre se m' è una misura su $\mathcal{B}$ tale che $m'(]-\infty, x]) = F_X(x)$ per ogni numero reale x, si ha $m' = P_X$

la legge è l'unica misura sui Boreliani t.c. sulle semirette inferiori coincide con FdR.

*Teor fondamentale calcolo delle probabilità

Sia $g: \mathcal{X} \to \mathbb{R}^*$ funzione sommabile rispetto P_X, allora $\int_{\{X \in A\}} g(X)\,dP = \int_A g\,dP_X$ $\forall A \in \mathcal{N}$

dim

- $g = I_{A'}$ con $A' \in \mathcal{N}$. Allora $g(X(\omega)) = \begin{cases} 1 & \text{se } X(\omega) \in A' \\ 0 & \text{se } X(\omega) \notin A' \end{cases} = I_{\{X \in A'\}}(\omega)$ $\forall \omega \in \Omega$

  allora $\int_A g\,dP_X = \int I_{A'} \cdot I_A\, P_X = \int_{A' \cap A} dP_X = P(X \in (A' \cap A)) = P(\{X \in A'\} \cap \{X \in A\}) = \int_\Omega I_{\{X \in A'\}} \cdot I_{\{X \in A\}}\,dP = \int_{\{X \in A\}} I_{\{X \in A'\}}\,dP = \int_{\{X \in A\}} g(X)\,dP = \int_{\{X \in A\}} g(X)\,dP$.

- $g \ge 0$ semplice, allora $g = \sum_{y \in g(\mathcal{X})} y \cdot I_{\{g = y\}}$, da cui $\int_A g\,dP_X = \int_A (\sum_{y \in g(\mathcal{X})} y \cdot I_{\{g=y\}})\,dP_X = \sum_{y \in g(\mathcal{X})} y \int_A I_{\{g=y\}}\,dP_X \stackrel{\text{PASSO PRECEDENTE}}{=} \sum_{y \in g(\mathcal{X})} y \int_{\{X \in A\}} I_{\{g=y\}}(X)\,dP = \int_{\{X \in A\}} (\sum_{y \in g(\mathcal{X})} y\, I_{\{g=y\}}(X))\,dP = \int_{\{X \in A\}} g(X)\,dP$.

  $g_n \uparrow g$, $g_n(x) \uparrow g(x)$ $\forall x$ (*)
  pongo $x = X(\omega)$
  $g_n(X(\omega)) \uparrow g(X(\omega))$

- $g \ge 0$ $(\mathcal{N}, \mathcal{B})$-misurabile, dal lemma fondamentale $0 \le$ [illegible] [illegible] lici allora $0 \le g_n \cdot I_A \uparrow g \cdot I_A$ e $g_n(X) \uparrow g(X)$ (*) cioè $g_n(X(\omega)) \uparrow g(X(\omega))$ $\forall \omega \in \Omega$. Essendo $(g_n(X))_{n \ge 1}$ funz. $\mathcal{N}, \mathcal{B}$-misurabili: $\int_A g\,dP_X \stackrel{\text{dal teor conv monotona}}{=} \lim_{n \to +\infty} \int_A g_n\,dP_X \stackrel{\text{PASSO PRECEDENTE}}{=} \lim_{n \to +\infty} \int_{\{X \in A\}} g_n(X)\,dP = \int_{\{X \in A\}} \lim_{n \to +\infty} g_n(X)\,dP = \int_{\{X \in A\}} g(X)\,dP$ PASSO PREC

- g qls $(\mathcal{N}, \mathcal{B})$-misurabile $(g(X))^+ = g^+(X)$, $(g(X))^- = g^-(X)$ allora $\int_A g^+\,dP_X \stackrel{!}{=} \int_{\{X \in A\}} g^+(X)\,dP = \int_{\{X \in A\}} g(X)^+\,dP$ (idem per g^-). Dalla arbitrarietà di A la P_X-sommabilità di g equivale alla sommabilità di $g(X)$
  $\int g^+ - \int g^- = \int g\,dP_X$ e $\int_{\{X \in A\}} g(X)^+ - \int_{\{X \in A\}} g(X)^- = \int_{\{X \in A\}} g(X)\,dP$. Inoltre
  $\int_A g\,dP_X = \int_{\{X \in A\}} g(X)\,dP$

*Il risultato dice che il meccanismo di sorteggio con cui estraggo X non conta, bensì conta la legge d'arrivo. Se modelli di estrazione diversi portano alla stessa legge allora nulla cambia.

Funzione di densità di un ente aleatorio

Esempio introduttivo

- v.a. discreta con rango $B = \{x_1, x_2, \dots\}$, $f(x) = P(X = x) \ge 0$

  data ν_B su $\mathcal{N} = 2^{\mathbb{R}}$, allora $\forall A \subset \mathbb{R}$ $P(X \in A) = P_X(A) = P(\bigcup_{n \in \{n \mid x_n \in A\}} \{X = x_n\}) \stackrel{\text{disgiunti}}{=} \sum_{n \in \{n \mid x_n \in A\}} f(x_n) = \sum_{n \ge 1} f(x_n) \cdot I_A(x_n) = \int_A f\,d\nu_B$ $\forall A \in \mathcal{N}$

posso ricostruire la legge tramite integrale e misura contegg.

• v.a. assolutamente continua, con funzione di densità $f \geq 0$ con $F_X(x) = \int_{-\infty}^{x} f(t)dt \quad \forall x \in \mathbb{R}$

allora $P_X(B) = \int_B f(t)dt \quad \forall B \in \mathcal{B}$

Nei due casi diversi considerati la legge P_X può essere determinata via integrazione di una funzione $f \geq 0$ rispetto ad una misura di riferimento (nel primo caso $\#_Q$; nel secondo λ). Posso quindi unificare i due casi

* Ritornando al caso generale, supponiamo che sulla σ-algebra $\mathcal{N}$ sia definita una misura di riferimento (arbitraria) μ.
Chiamiamo μ-densità di X e.g. ogni funzione $f: \mathcal{X} \to [0,+\infty[$ $(\mathcal{N},\mathcal{B})$-misurabile t.c. $P_X(A) = \int_A f d\mu \quad \forall A \in \mathcal{N}$

$$\Omega \to \underset{\mathcal{N}}{\mathcal{X}} \xrightarrow{f} \underset{\mathcal{B}}{\mathbb{R}^+}$$

Una μ-densità è finita (μ-q.o.)
Oss $\int_A f d\mu$ è finito $\Rightarrow P_X(A)$ finita allora l'insieme dove $\int_A f d\mu = +\infty$ è trascurabile

Teor
f μ-densità di X e $\tilde{f}$ μ-densità $\Rightarrow \tilde{f} = f$ (μ-q.o.) cioè $\mu(f \neq \tilde{f}) = 0$

dim
Supposto per assurdo che $\mu(\{f \neq \tilde{f}\}) = \mu(\{f < \tilde{f}\}) + \mu(\{f > \tilde{f}\}) > 0$, assumiamo ad esempio $\mu(\{f < \tilde{f}\}) > 0$; osserviamo che $A = \{f < \tilde{f}\} \in \mathcal{N}$ risulta che $\int_A f d\mu = P(X \in A) = \int_A \tilde{f} d\mu$, da cui per linearità $0 = \int_A \tilde{f} d\mu - \int_A f d\mu = \int_A (\tilde{f} - f) d\mu$
Da $\tilde{f} - f > 0$ su A segue $0 = \int_A (\tilde{f} - f) d\mu > 0$ (th. est. sugli integrali).
contraddizione
Quindi la μ-densità è definita a meno di insiemi μ-trascurabili

Teor
cfr. Teor fondamentale
dim del th. fondam.
Sia f μ-densità, $g: \mathcal{X} \to \mathbb{R}^+$ sommabile rispetto alla legge di X, allora $\int_A g \, dP_X = \int_A g f d\mu \quad \forall A \in \mathcal{N}$

dim
• g semplice e ≥ 0, $g = \sum_{y \in g(\mathcal{X})} y I_{\{g=y\}}$, $\int_A g \, dP_X = \int_X g I_A dP_X \overset{lin}{=} \sum_{y \in g(\mathcal{X})} y \int_X I_{\{g=y\}} I_A dP_X$
$= \sum_{y \in g(\mathcal{X})} y \left(\int_X I_{\{g=y\} \cap A} dP_X\right) = \sum_{y \in g(\mathcal{X})} y P_X(\{g=y\} \cap A) = \sum_{y \in g(\mathcal{X})} y \int_{\{g=y\} \cap A} f d\mu =$
hanno additività numerabile
$= \sum_{y \in g(\mathcal{X})} y \int_A I_{\{g=y\}} f d\mu = \int_A \left[\sum_{y \in g(\mathcal{X})} y I_{\{g=y\}}\right] f d\mu = \int_A [g] f d\mu$

• $g \geq 0$, $(\mathcal{N},\mathcal{B})$-misurabile, $\exists\ 0 \leq g_n \uparrow g$, g_n semplice $\Rightarrow 0 \leq g_n f \uparrow g f$ con $(g_n f)_{n \geq 1}$ successione di funzioni $\mathcal{B}$-misurabili non negative, allora
lemma fond.
$\int_A g \, dP_X \underset{\text{conv. monot.}}{=} \lim_{n \to +\infty} \int_A g_n dP_X \underset{\text{passo precedente}}{=} \lim_{n \to +\infty} \int_A g_n f d\mu \underset{\text{conv. monot.}}{=} \int_A \left(\lim_{n \to +\infty} g_n\right) f d\mu = \int_A g f d\mu$

• g qls $\mathcal{B}$-misurabile
passo prec. $\quad g^+ f = (gf)^+$
$\int_A g^+ dP_X = \int_A g^+ f d\mu = \int_A (gf)^+ d\mu$
$\int_A g^- dP_X = \int_A g^- f d\mu = \int_A (gf)^- d\mu$
passo precedente $\quad (g^- f) = (gf)^-$

allora g è P_X-sommabile se $g \cdot f$ è μ-sommabile
Inoltre
$\int_A g \, dP_X = \int_A g^+ dP_X - \int_A g^- dP_X = \int_A (gf)^+ d\mu - \int_A (gf)^- d\mu = \int_A g f d\mu$ nel caso di sommabilità

Teor fondamentale del calcolo delle probabilità

Sia $g: \mathcal{X} \to \mathbb{R}^+$ una funzione sommabile rispetto alla legge di X. Quindi allora

$$\int_{\{X \in A\}} g(X) dP = \begin{cases} \int_A g \, dP_X \\ \int_A g f d\mu & \text{se } f \text{ è } \mu\text{-densità di } X \end{cases} \qquad \forall A \in \mathcal{N}$$

sx Th. ???

Due v.a. aleatori X, Y a valori in $\mathcal{X}$ si dicono equidistribuiti se $P_X = P_Y$ cioè se hanno la medesima legge

Conseguentemente, X e Y sono equidistribuite se sono uguali quasi certamente (cioè P-q.c.). Infatti, posto $C = \{X = Y\}$ si ha $P(C) = 1$ e quindi $P_X(A) = P(X \in A) = P(\{X \in A\} \cap C) + P(\{X \in A\} \cap C^c) = P(\{X \in A\} \cap C) + 0 = P((Y \in A) \cap C) = P((Y \in A) \cap C) + 0 = P(Y \in A) = P_Y(A) \quad \forall A \in \mathcal{B}$.

Speranza matematica

$\mathbb{P}$ realizzazione dell'evento certo
$\mathcal{F}$ σ-algebra di eventi logicamente dipendenti da $\mathbb{P}$
P_r probabilità su $\mathcal{F}$
$X: \mathbb{P} \to \mathbb{R}$ variabile aleatoria (è $(\mathcal{F}, \mathcal{B})$-misurabile)
$E(X)$ speranza matematica

3° P. commerciale

- X con rango $B = \{x_1, \dots, x_n\}$. Risulta $E(X) = \sum_{i=1}^{n} x_i P(X = x_i) = \int_{\mathbb{P}} X \, dP_r$

- X assolutamente continua, ma f funzione di densità, dal Teor. fondamentale con $g(x) = x$ abbiamo $E(X) = \int_{\mathbb{R}} x f(x) dx = \int_{\mathbb{R}} x f(x) \lambda(dx) = \int_{\mathbb{P}} X \, dP_r$

così possiamo dire che la speranza matematica di X è un integrale.

La speranza matematica (o valor medio) di una v.a. estesa P-sommabile X è l'elemento della retta reale ampliata $E(X) = \int_{\Omega} X \, dP$
nato dal bisogno di calcolo del valore per v.a. non discrete e non continue.
Nelle ipotesi del Teor. fondamentale abbiamo

$$E(g(X)) = \int_{\Omega} g(X) dP = \begin{cases} \int_{\mathbb{R}} g \, dP_X \\ \int_{\mathbb{R}} g f \, d\mu \text{ se } f \text{ è } \mu\text{-densità di } X \end{cases}$$

Se g è l'identità $E(X) = \begin{cases} \int_{\mathbb{R}} x P_X(dx) \\ \int_{\mathbb{R}} x f(x) \mu(dx) \text{ se } f \text{ è } \mu\text{-densità di } X \end{cases}$

Teor

Risulta $E(I_A) = P(A)$ e $E(\alpha I_\Omega) = \alpha \quad \forall A \in \mathcal{A}, \alpha \in \mathbb{R}^*$. Inoltre, supposto che X, Y siano var. aleatorie estese con speranza matematica, risulta:
- LINEARITÀ: se $\alpha, \beta \in \mathbb{R}^*$ sono t.c. $\alpha X + \beta Y$ è definita ovunque, allora $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$ se $\alpha E(X)$ e $\beta E(Y)$ non sono infiniti di segno opposto
- $E(X) = E(Y)$ se $X = Y$ (P-q.c.)
- MONOTONIA: $E(X) \leq E(Y)$ se $X \leq Y$ (P-q.c.)
- $|E(X)| \leq E(|X|)$
- INTERNALITÀ: $\inf X \leq E(X) \leq \sup X$
- INTERNALITÀ STRETTA: $\alpha < E(X) < \beta$, se $\alpha < X(\omega) < \beta$ per ogni ω;

dim. $\alpha < E(X)$ e $E(X) < \beta$ si dim. allo stesso modo. Proviamo la prima:
$E(X) - \alpha = E(X) - E(\alpha I_\Omega) = E(X - I_\Omega) = \int_{\Omega} (X - \alpha I_\Omega) dP > 0$, perché $(X - \alpha I_\Omega) > 0$ e quindi anche l'integrale. Quindi $E(X) - \alpha > 0$.

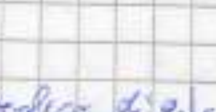
- DISUGUAGLIANZA DI MARKOV: $X \geq 0$, $a > 0$ allora $P(X \geq a) \leq E(X)/a$

dim. Ho $a \cdot I_{\{X \geq a\}} \leq X$ perché $\begin{cases} \forall X(\omega) \geq a & a \leq X \\ \forall X(\omega) < a & 0 \leq X \end{cases}$, per monotonia ho che
$a \cdot E(I_{\{X \geq a\}}) \leq E(X)$ da cui $P(\{X \geq a\}) \leq E(X)/a$

- Se Z una variabile aleatoria, essa ammette speranza matematica se $Z \geq X$ e $E(X) > -\infty$ o $Z \leq X$ e $E(X) < +\infty$

Introduzione alla disuguaglianza di Jensen
- J intervallo aperto (limitato o no) della retta reale
- $g: J \to \mathbb{R}$ convessa se $a, b \in J$ due pt arbitrari $| a < b$, il grafico di $g|_{[a,b]}$ sta sotto la retta passante per i pt $(a, g(a))$, $(b, g(b))$. Formalmente

$$g(x) \leq \frac{g(b) - g(a)}{b - a} x + \frac{g(a) b - g(b) a}{b - a} \quad \forall x \in [a,b]$$



Osservato che, dato $x \in [a, b]$ risulta $x = \frac{b - x}{b - a} a + \frac{x - a}{b - a} b$ con $\begin{cases} \frac{b - x}{b - a} + \frac{x - a}{b - a} = 1 \\ \frac{b - x}{b - a}, \frac{x - a}{b - a} \geq 0 \end{cases}$
possiamo concludere che ogni pt dell'intervallo $[a, b]$ è ottenibile come miscela degli estremi dell'intervallo

Quindi se una funzione g è convessa se $\forall a,b \in J$ e $\forall \alpha \in [0,1]$ risulta
$g(\alpha a+(1-\alpha)b) \le \alpha g(a)+(1-\alpha)g(b)$ immagine della miscela ≤ miscela delle immagini
ne segue per induzione
$g(\alpha_1 x_1+\dots+\alpha_n x_n) \le \alpha_1 g(x_1)+\dots+\alpha_n g(x_n) \quad \forall x_1,\dots,x_n \in J,\ \alpha_1,\dots,\alpha_n \ge 0 \mid \alpha_1+\dots+\alpha_n=1$

Lemma

$g: J \to \mathbb{R}$ convessa. Allora · Per ogni punto a di J $\exists$ un reale $b \mid g(x) \ge b(x-a)+g(a)\ \forall x \in J$ — retta passante per x con $b=\partial$

· g è continua $\Rightarrow$ $\mathcal{B}\cap J$-Borel misurabile

Teor **Disuguaglianza di Jensen** Dato un intervallo aperto J (limitato o no) della retta reale. Siano $g: J \to \mathbb{R}$ una funzione convessa e $X: \Omega \to J$ una v.a. con speranza matematica. Allora $E(X) \in J$. Inoltre, la v.a. $g(X)$ ammette speranza matematica e risulta $g(E(X)) \le E(g(X))$ cioè $g\left[\int_\Omega X\,dP\right] \le \int_\Omega g(X)\,dP$

dim
- dalla $\inf J < X(\omega) < \sup J\ \forall \omega$ ho che dalla integrabilità stretta, $E(X) \in J$ osservando
- proviamo ora che $g(X)$ è una v.a. con sper. matematica: dal Lemma so che $g(X)$ è una v.a. con speranza matematica
- per quanto riguarda la sommabilità esiste, per il Lemma un numero reale b t.c. $g(x) \ge b\cdot(x-E(X))+g(E(X))$ per ogni $x \in J$, considerata allora la v.a. $Y=b[X-E(X)]+g(E(X))$ si ha $g(X) \ge Y$ e $E(Y)=b[E(X)-E(X)]+$ $+E(g(E(X))I_\Omega)=g(E(X)) \in \mathbb{R}$, dunque $g(X)$ è sommabile
- da $g(X) \ge Y$ per monotonia $E(g(X)) \ge g(E(X))$

$$\begin{array}{ccccc} \Omega & \xrightarrow{X} & J & \longrightarrow & \mathbb{R} \\ \mathcal{A} & & \mathcal{B}\cap J & & \mathcal{B} \\ P & & \underset{g(X)}{\rightsquigarrow} & & \end{array}$$

Varianza

Data una v.a. X con speranza matematica finita la varianza di X è il momento centrale secondo: $Var(X)=E[(X-E(X))^2]$

Teor

Sia X una v.a. con speranza matematica finita. Allora:
- $Var(X)=E(X^2)-[E(X)]^2$

dim
Per linearità della sp. matematica $Var(X)=E(X^2-2XE(X)+E(X)^2)=$
$=E(X^2)-2E(X)E(X)+E(X)^2=E(X^2)-E(X)^2$

- $Var(\alpha X+\beta)=\alpha^2 Var(X) \quad \forall \alpha,\beta$ reali

dim
Per linearità della sp. matematica $Var(\alpha X+\beta)=E\left(\left([\alpha X+\beta]-E(\alpha X+\beta)\right)^2\right)$
$=E[((\alpha X+\beta)-(\alpha E(X)+\beta))^2]=E(\alpha^2(X-E(X))^2)=\alpha^2 E((X-E(X))^2)=$
$=\alpha^2 Var(X)$

- $Var(X)=0$ se e solo se $X=E(X)$ (P-q.c.)
- Se $\varepsilon>0$, allora $P(|X-E(X)| \ge \varepsilon) \le Var(X)/\varepsilon^2$ DISUGUAGLIANZA DI CEBICEV

dim
Dalla disuguaglianza di Markov risulta $P(|X-E(X)| \ge \varepsilon)=$
$=P((X-E(X))^2 \ge \varepsilon^2) \le E((X-E(X))^2)/\varepsilon^2=Var(X)/\varepsilon^2$

$\left[P(X \ge a) \le E(X)/a \quad a>0,\ X \ge 0\right]$

Covarianza

Considerate due var. aleatorie X, Y con speranza matematica finite e t.c. sia finita la speranza matematica $E(X\cdot Y)$ del loro prodotto, la covarianza di X e Y è la differenza: $Cov(X,Y)=E(X\cdot Y)-E(X)\cdot E(Y)$

Teor definizione equivalente

$E[(X-E(X))(Y-E(Y))]=E(XY-XE(Y)-YE(X)+E(X)E(Y))=$
$=E(XY)-E(X)E(Y)-E(Y)E(X)+E(X)E(Y)=E(XY)-E(X)E(Y)$

Inoltre vale $Cov(X,Y) = Cov(Y,X)$ e $Cov(X,X) = Var(X)$.

Teor X, Y, Z variabili aleatorie con speranza matematica finita e t.c. $E(XY)$ e $E(YZ)$ finite allora

- $Cov(\alpha X, Y) = \alpha Cov(X,Y) \quad \forall \alpha \in \mathbb{R}$

dim

$$Cov(\alpha X, Y) = E(\alpha X Y) - E(\alpha X)E(Y) \overset{\text{linearità}}{=} \alpha E(XY) - \alpha E(X)E(Y) = \alpha Cov(X,Y)$$

- $Cov(X+Y, Z) = Cov(X,Z) + Cov(Y,Z)$

dim

$$Cov(X+Y,Z) = E[(X+Y)Z] - E(X+Y)E(Z) \overset{\text{linear}}{=} E(X \cdot Z) + E(Y \cdot Z) - E(X)E(Z) - E(Y)E(Z) = Cov(X,Z) + Cov(Y,Z)$$

- $[Cov(X,Y)]^2 \leq Var(X) Var(Y)$ DISUGUAGLIANZA DI CAUCHY SCHWARTZ

dim

Pongo $U = X - \mu$ e $V = Y - \nu$ con $\mu = E(X)$ e $\nu = E(Y)$, la tesi è che $E(U \cdot V) \leq E(U^2)E(V^2)$

- se uno fra U, V è zero la disuguagl è vera: infatti se $E(U^2) = 0 \Rightarrow X = E(X)$ P-q.c. (dn Lebesgue) $\Rightarrow U = 0$ P-q.c. Da cui $U \cdot V = 0$ P-q.c. $(+\infty \cdot 0 = 0) \Rightarrow E(U \cdot V) = 0 = 0 \cdot E(V^2)$
- se entrambi $E(U^2), E(V^2) > 0$
  - se uno dei 2: $E(U^2) = +\infty \Rightarrow E(U \cdot V)^2 \leq +\infty$ la diseg è soddisfatta
  - se $E(U^2), E(V^2)$ finite allora dato un arbitrario $t \in \mathbb{R}$ consideriamo la v.a. $tU+V$, dalla linearità della sp. matem. risulta:

$(0 \leq) E((tU+V)^2) = E(t^2U^2 + V^2 + 2tUV) = t^2E(U^2) + 2tE(U \cdot V) + E(V^2)$ diseq di secondo grado a coefficienti reali $E(U^2)t^2 + 2E(U \cdot V)t + E(V^2) \geq 0$ che ha per soluzione data l'arbitrarietà di t ogni t reale, in quanto $(tU+V)^2 \geq 0$ e quindi $E(tU+V)^2 \geq 0$ $\forall t$. Solo se $E(U^2)t^2 + 2E(UV)t + E(V^2) = 0$ ha una soluzione t^* ciò significa che il $\Delta \leq 0$ (non $>$) ossia $E(U \cdot V)^2 - E(U^2) \cdot E(V^2) \leq 0$ che è proprio la tesi

- Inoltre se $X_1, \dots, X_n$ sono v.a. con speranza matematica finita t.c. $E(X_i \cdot X_j)$ finita $\forall i \neq j$, si ha $Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i=1}^{n}\sum_{j>i} Cov(X_i, X_j)$

dim

procedo per induzione

- $n=2$ $\quad Var(X+Y) = E([(X+Y) - E(X+Y)]^2) = E([(X - E(X)) + (Y - E(Y))]^2) =$
$= E((X-E(X))^2) + E((Y-E(Y))^2) + 2E[(X-E(X))(Y-E(Y))] =$
$= Var(X) + Var(Y) + 2Cov(X,Y)$
- supponiamo l'uguaglianza verificata per $n \geq 2$, e la dimostro per $n = m+1$

$Var\left(\sum_{i=1}^{m} X_i\right) = Var\left(\sum_{i=1}^{m-1} (X_i) + X_m\right) = Var\left(\sum_{i=1}^{m-1} X_i\right) + Var(X_m) + 2Cov\left(\sum_{i=1}^{m-1} X_i, X_m\right) =$
$= \sum_{i=1}^{m-1} Var(X_i) + 2\sum_{i=1}^{m-1}\sum_{j>i}^{m-1} Cov(X_i, X_j) + Var(X_m) + 2\, Cov\left(\sum X_i, X_m\right) =$
$= \sum^{m} Var(X_i) + Var(X_m) + 2\sum_{i=1}^{m}\sum_{j>i} Cov(X_i, X_j) + 2\sum_{i=1}^{m} Cov(X_i, X_m)$
$= \sum^{m} Var(X_i) + 2 \cdot \sum^{m}_{i}\sum_{j>i} Cov(X_i, X_j)$.

Oss

Nel caso particolare in cui $Cov(X,Y)^2 = Var(X) \cdot Var(Y)$ (C-S), la t^* è unica soluzione di $E(U^2)t^2 + 2E(U \cdot V)t + E(V^2) = 0$ ed è $t^* = -\frac{E(U \cdot V)}{E(U^2)}$

Considero $t^*U + V$: che $E((t^*U+V)^2) = 0$ (da sopra), da cui

(solita) per teor "$X \geq 0$ e $E(X) = 0 \Rightarrow P(X=0) = 1$" ho che $P(t^*U + V = 0) = 1$,
$P(Y - \nu = -t^*(X - \mu)) = 1$, $P(Y = -t^*X + (t^*\mu + \nu)) = 1$ cioè Y trasformata affine di X, crescente se $E(U \cdot V) > 0$, decrescente se $E(U \cdot V) < 0$
$P(Y = -t^*X + b) = 1$ a meno di un insieme di prob zero.

$$t^* = \frac{E((X-E(X))(Y-E(Y)))}{E((X-E(X))^2)} = \frac{Cov(X,Y)}{Var(X)} \Rightarrow Y = \frac{Cov(X,Y)}{Var(X)} X + b \text{ P-q.c.}$$

OSS

da C-S $\frac{Cov(X,Y)^2}{Var(X)Var(Y)} \leq 1$ cioè $\left(\frac{Cov(X,Y)}{\sqrt{Var(X)} \cdot \sqrt{Var(Y)}}\right)^2 \leq 1$ chiamo $\rho_{X,Y} = \left(\frac{Cov}{\sigma(X)\sigma(Y)}\right)^2$

INDICE DI CORRELAZIONE DI BRAVAIS:

- se $\rho = -1$ $X(o\, Y)$ è trasf affine decrescente dell'altra P-q.c.
- se $\rho = 1$ $X(o\, Y)$ è " " crescente " P-q.c.

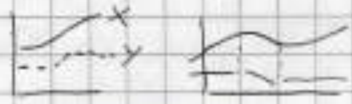
$\rho = 0$ $\quad X, Y$ sono non correlate

$-1 \leq \rho < 0$ $\quad X, Y$ sono correlate negativamente

$0 < \rho < 1$ $\quad$ " $\quad$ " $\quad$ positivamente

Due variabili aleatorie X, Y si dicono comonotone se:
$[X(\omega) - X(\omega')][Y(\omega) - Y(\omega')] \geq 0 \quad \forall \omega, \omega'$

se $X \uparrow$, allora $Y \uparrow$ oppure $Y \rightarrow$; assicura che non possono essere correlate negativamente
$\quad \downarrow \qquad\qquad \downarrow$ oppure $\rightarrow$



Teor X, Y v.a. comonotone con $E(X), E(Y), E(X \cdot Y)$ finite, allora $Cov(X,Y) \geq 0$

dim
fix ω_0 arbitrario, allora dalla comonotonicità $(X - X(\omega_0))(Y - Y(\omega_0)) \geq 0$
$\Rightarrow XY - Y(\omega_0)X - X(\omega_0)Y + X(\omega_0)Y(\omega_0) \geq 0$ sono numeri
$XY + X(\omega_0)Y(\omega_0) \geq Y(\omega_0)X + X(\omega_0)Y \quad / E(\cdot)$
$E(XY) + X(\omega_0)Y(\omega_0) \geq Y(\omega_0)E(X) + X(\omega_0)E(Y)$, dalla arbitrarietà di ω_0
$E(XY) + XY \geq YE(X) + XE(Y) \quad / E(\cdot)$
$2E(XY) \geq 2E(X)E(Y)$
$Cov(X,Y) = E(XY) - E(X)E(Y) \geq 0$

Os $Cov(X,Y) \geq 0 \not\Rightarrow X, Y$ comonotone

Probabilità 2

Misura prodotto

Introduzione: Dato un insieme $A \subset \Omega = \Omega_1 \times \Omega_2$ vogliamo misurare A sfruttando l'analogia col principio di Cavalieri (affette salame)

$\forall x$ calcolo la sezione di A corrispondente

Data $m(\cdot)$ misura, $(A_n)_{n \geq 1}$ una P discreta di Ω e $m(A_n) < +\infty$ $\forall n$, allora m è detta misura σ-finita.
Sono esempi di misure σ-finite le probabilità, la misura di Lebesgue n-dimensionale e le misure di conteggio indotte da insiemi discreti.

Dato $\Omega_i \neq \emptyset$, $\mathcal{R}_i$ una σ-algebra su Ω_i, m_i misura σ-finita su $\mathcal{R}_i$ $(i = 1,2)$
Posto $\Omega = \Omega_1 \times \Omega_2$ consideriamo la:
- famiglia dei rettangoli misurabili $\mathcal{B} = \{A_1 \times A_2 : A_i \in \mathcal{R}_i,\ i = 1,2\}$
- σ-algebra prodotto: $\mathcal{R}_1 \otimes \mathcal{R}_2 = \sigma(\mathcal{B})$ di $\mathcal{R}_1$ e $\mathcal{R}_2$

Consideriamo infine, $\forall S \subset \Omega$ la:
- sezione S relativa a $\omega_1 \in \Omega_1$: $S(\omega_1) = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in S\}$
- sezione S relativa a $\omega_2 \in \Omega_2$: $S(\omega_2) = \{\omega_1 \in \Omega_1 : (\omega_1, \omega_2) \in S\}$

introduciamo una misura sulla σ-algebra $\mathcal{R} = \mathcal{R}_1 \otimes \mathcal{R}_2$ tale che verifichi le due seguenti proprietà, suggerite dalla nozione di area delle figure piane e del relativo principio di Cavalieri:

(a) la misura di un rettangolo misurabile $A_1 \times A_2$ di base A_1 e altezza A_2 è pari al prodotto $m_1(A_1) \cdot m_2(A_2)$

(b) due insiemi $A_1, A_2 \in \mathcal{R}$ hanno misura uguale se, per qualche $i = 1,2$ danno luogo a sezioni di misura uguale in corrispondenza ad ogni scelta dell'elemento che le individua in Ω_i, cioè se
$m_j(A_1(\omega_i)) = m_j(A_2(\omega_i))$ con $j \neq i$ $\forall \omega_i \in \Omega_i$

Lemma
Sia $A \in \mathcal{R}_1 \otimes \mathcal{R}_2$ allora:
- $A(\omega_i) \in \mathcal{R}_j$ $\forall \omega_i$ $i = 1,2$
- $\omega_i \to m_j(A(\omega_i))$ è $\mathcal{R}_i$-Borel misurabile $i = 1,2;\ j \neq i$

Introduciamo sulla σ-algebra prodotto la **misura prodotto** di m_1 e m_2:
$$m_1 \times m_2(A) = \int_{\Omega_1} m_2(A(\omega_1))\, m_1(d\omega_1)$$

$m_1 \times m_2(A)$ è una misura

dim/
- $m_1 \times m_2(\emptyset) = \emptyset$ perché se $A = \emptyset$ allora $m_2(A) = 0 \Rightarrow \int_{\Omega_1} 0 \cdot m_1(d\omega_1) = 0$
- $m_1 \times m_2\left(\bigcup_{n \geq 1} A_n\right) = \int_{\Omega_1} m_2\left(\bigcup_{n \geq 1} A_n(\omega_1)\right) m_1(d\omega_1) = \int_{\Omega_1} \sum_{n \geq 1} m_2(A_n(\omega_1))\, m_1(d\omega_1) \overset{\text{teor. integr. per serie}}{=}$
$= \sum_{n \geq 1} \int_{\Omega_1} m_2(A_n(\omega_1))\, m_1(d\omega_1) = \sum_{n \geq 1} m_1 \times m_2(A_n)$ $\forall (A_n)_{n \geq 1}$ disgiunti

Il seguente teorema assicura che la misura prodotto è l'unica misura sulla σ-algebra prodotto che verifica le proprietà (a), (b)

Teor. Fubini:
- $m_1 \times m_2(A_1 \times A_2) = m_1(A_1)\, m_2(A_2)$ $\forall A_1 \times A_2$
- $m_1 \times m_2(A) = \int_{\Omega_2} m_1(A(\omega_2))\, m_2(d\omega_2)$ $\forall A \in \mathcal{R}_1 \otimes \mathcal{R}_2$

Inoltre qualsiasi sia la misura m' sulla σ-algebra prodotto t.c. $m'|_{\mathcal{B}} = m_1 \times m_2|_{\mathcal{B}}$ risulta $m' = m_1 \times m_2$

Oss $\int_{\Omega_1} m_2(A(\omega_1)) \cdot m_1(d\omega_1)$ $\exists$ se $m_2(\cdot)$ integrabile $\Rightarrow m_2(A(\cdot))$ è $(\mathcal{R}, \mathcal{B})$-misurabile (e idem per $m_1(A(\cdot))$ se integro su Ω_2), ma non basta occorre che m_1 sia σ-finita (σ_2 idem).
Inoltre l'integrale ha senso se $A(\omega_1) \in \mathcal{R}_2$, $A(\omega_2) \in \mathcal{R}_1$, quindi se $\mathcal{B} = A_1 \times A_2) \in \mathcal{R}_1 \otimes \mathcal{R}_2$ dal $\sigma(\mathcal{B}) = \mathcal{R}_1 \otimes \mathcal{R}_2$.

Sintesi del nuovo costrutto

$\Omega_1 \quad \Omega_2 \qquad \Omega = \Omega_1 \times \Omega_2$

$\mathcal{A}_1 \quad \mathcal{A}_2 \quad \to \mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\{A_1 \times A_2 : A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\})$

$m_1 \quad m_2 \qquad m = m_1 \times m_2$

σ-finite

$$m(A) = \int_{\Omega_2} m_1(A(\omega_2)) \cdot m_2(d\omega_2) = \int_{\Omega_1} m_2(A(\omega_1)) \cdot m_1(d\omega_1)$$

$$m(A_1 \times A_2) = m_1(A_1) \cdot m_2(A_2)$$

m' misura $m(\cdot)$ su $\mathcal{A}$ t.c. $m'(A_1 \times A_2) = m_1(A_1) \cdot m_2(A_2)$

Ora mi pongo il problema di integrare su 2 variabili

Come per integrale Riemann anche qui ci sono formule di riduzione. I due seguenti teoremi sono due pilastri fondamentali della teoria dell'integrazione in quanto consentono di ridurre l'integrale doppio $\iint_\Omega f \, dm_1 \times m_2 = \iint_{\Omega} f(\omega_1, \omega_2) \, m_1 \times m_2 (d\omega_1 \times \omega_2)$ ad un integrale iterato, come pure invertire l'ordine delle integrazioni successive

Teor. di Tonelli

Sia $f \geq 0$ $\mathcal{A}_1 \otimes \mathcal{A}_2$ - Borel misurabile, allora $\forall A_1, A_2$, le funzioni:

$\omega_1 \to \int_{A_2} f(\omega_1, \cdot) \, dm_2 \quad$ e $\quad \omega_2 \mapsto \int_{A_1} f(\cdot, \omega_2) \, dm_1$

sono rispettivamente $(\mathcal{A}_1, \mathcal{B})$-misurabile e $(\mathcal{A}_2, \mathcal{B})$-misurabile. Inoltre risulta:

$$\iint_{A_1 \times A_2} f \, dm_1 \times m_2 = \int_{A_1} \left(\int_{A_2} f(\omega_1, \cdot) \, dm_2 \right) m_1(d\omega_1) = \int_{A_2} \left(\int_{A_1} f(\cdot, \omega_2) \, dm_1 \right) m_2(d\omega_2)$$

indipendente da ordine di integraz.

Teor. di Fubini

Sia f $m_1 \times m_2$ - sommabile. Risulta: $A_1^{(1)} \subseteq \{\omega_1 : f(\omega_1, \cdot) \text{ è } m_2\text{-sommabile}\} \in \mathcal{A}_1$

$A_2^{(2)} \subseteq \{\omega_2 : f(\cdot, \omega_2) \text{ è } m_1\text{-sommabile}\} \in \mathcal{A}_2$

e $m_i((A_i^{(i)})^c) = 0 \quad (i = 1, 2)$. Inoltre $\forall A_i \ (i = 1, 2)$, le funzioni

$$g_1 : \omega_1 \mapsto \begin{cases} \int_{A_2} f(\omega_1, \cdot) \, dm_2 & \text{se } \omega_1 \in A_1^{(1)} \\ 0 & \text{se } \omega_1 \notin A_1^{(1)} \end{cases} \qquad g_2 : \omega_2 \mapsto \begin{cases} \int_{A_1} f(\cdot, \omega_2) \, dm_1 & \text{se } \omega_2 \in A_2^{(2)} \\ 0 & \text{se } \omega_2 \notin A_2^{(2)} \end{cases}$$

sono rispettivamente $(\mathcal{A}_1, \mathcal{B})$-misurabile e $(\mathcal{A}_2, \mathcal{B})$-misurabile e tali che

$$\iint_{A_1 \times A_2} f \, dm_1 \times m_2 = \int_{A_1} g_1 \, dm_1 = \int_{A_2} g_2 \, dm_2$$

~~formula~~ di riduzione

Estensione a n° finito di variabili

La nozione di misura prodotto ed i relativi teoremi di Tonelli e Fubini possono essere estesi al caso di più di due fattori

Siano $\mathcal{A}_k$ una σ-algebra su un insieme $\Omega_k \neq \emptyset$ e m_k una misura σ-finita su $\mathcal{A}_k$ $(k = 1, \dots, m)$

Posto $\Omega = \Omega_1 \times \dots \times \Omega_m$ consideriamo la:

- famiglia dei rettangoli misurabili $\mathcal{R}^{(m)} = \{A_1 \times \dots \times A_m : A_k \in \mathcal{A}_k \ (k = 1, \dots, m)\}$
- σ-algebra prodotto $\mathcal{A}_1 \otimes \dots \otimes \mathcal{A}_m = \sigma(\mathcal{R}^{(m)})$

Teor. Esiste una sola misura $m = m_1 \times \dots \times m_m$ sulla σ-algebra prodotto - detta misura prodotto di $m_1, \dots, m_m$, tale che:

$$m_1 \times \dots \times m_m (A_1 \times \dots \times A_m) = \prod_{k=1}^{m} m_k(A_k) \text{ per ogni rettangolo misurabile } A_1 \times \dots \times A_m$$

Sussiste la seguente generalizzazione dei teoremi di Fubini e Tonelli che permette di ridurre l'integrale multiplo:

$\int_{\Omega} f \, d m_1 \times \dots \times m_m = \int_{\Omega} f(\omega_1, \dots, \omega_m) \, m_1 \times \dots \times m_m (d\omega_1 \times \dots \times d\omega_m)$ ad un integrale iterato come pure di invertire l'ordine delle integrazioni successive

Teor Sia f m-sommabile. Data una permutazione $k_1, \dots, k_m$ di $\{1, \dots, m\}$ si ha $\int_{\Omega} f \, dm = \int_{\Omega_{k_1}} \left(\dots \left(\int_{\Omega_{k_m}} f(\omega_1, \dots, \omega_m) \, m_{k_m}(d\omega_{k_m}) \right) \dots \right) m_{k_1}(d\omega_{k_1})$, dove nel caso che non sia $f \geq 0$ ogni integrale

$\int_{\Omega_{k}} \left(\dots \left(\int_{\Omega_{k_m}} f(\omega_1, \dots, \omega_m) \, m_{k_m}(d\omega_{k_m}) \right) \dots \right) m_{k}(d\omega_{k})$ $(k>1)$

esiste $\forall (\omega_{k_1}, \dots, \omega_{k_{k-1}})$ appartenente ad un insieme $m_{k_1} \times \dots$-trascurabile ovunque altrove per effettuare l'integrazione successiva relativa alla misura $m_{k_{k-1}}$ si fa tale insieme tutti valori zero.

esempio $m=3$

$$\Omega = \Omega_1 \times \Omega_2 \times \Omega_3$$
$$\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2 \otimes \mathcal{A}_3 = \sigma(A_1 \times A_2 \times A_3 \mid A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2, A_3 \in \mathcal{A}_3)$$
$$m = m_1 \times m_2 \times m_3$$

$$\iiint_{A_1 \times A_2 \times A_3} f(\omega_1, \omega_2, \omega_3) \, m(d\omega_1 \times \omega_2 \times \omega_3) = \int_{A_1} \left(\iint_{A_2 \times A_3} f(\omega_1, \omega_2, \omega_3) \, m_2 \times m_3 (dm_2 \times m_3) \right) m_1(d\omega_1)$$
$$= \int_{A_2} \left(\iint_{A_1 \times A_3} f(\omega_1, \omega_2, \omega_3) \, m_1 \times m_3 (dm_1 \times m_3) \right) m_2(d\omega_2)$$

Misura di Lebesgue multidimensionale

considerato lo spazio numerico delle m-ple reali $(m \geq 2)$, la σ-algebra di Borel (di $\mathbb{R}^m$) è la σ-algebra prodotto $\mathcal{B}^{(m)} = \mathcal{B} \otimes \dots \otimes \mathcal{B}$, e la misura di Lebesgue m-dimensionale è la misura prodotto $\lambda^{(m)} = \lambda \times \dots \times \lambda$. Inoltre i boreliani di $\mathbb{R}^m$ sono gli elementi di $\mathcal{B}^{(m)}$. Analogamente al caso unidimensionale, $\mathcal{B}^{(m)}$ può essere vista come la σ-algebra generata dalla famiglia: - degli insiemi aperti di m-ple
- degli insiemi $S_{x_1, \dots, x_m} =]-\infty, x_1] \times \dots \times]-\infty, x_m]$ $(x_1, \dots, x_m \in \mathbb{R})$

La σ-algebra di Borel di $(\mathbb{R}^*)^m$ è per analogia la σ-algebra prodotto $(\mathcal{B}^*)^{(m)} = \mathcal{B}^* \otimes \dots \otimes \mathcal{B}^* \supset \mathcal{B}^{(m)}$ costituito dai boreliani di $(\mathbb{R}^*)^m$

Leggi congiunte

$\Omega \to X_1$ ente aleatorio $: \Omega \to \mathcal{X}_1$ $\mathcal{N}_1$
$\mathcal{A} \to X_2$ ente aleatorio $: \Omega \to \mathcal{X}_2$ $\mathcal{N}_2$
P

Pongo $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$
$\mathcal{N} = \mathcal{N}_1 \otimes \mathcal{N}_2 = \sigma(A_1 \times A_2 \mid A_1 \in \mathcal{N}_1, A_2 \in \mathcal{N}_2)$
$X = (X_1, X_2) : \omega \to (X_1(\omega), X_2(\omega))$ coppia aleatoria = ogni ente aleatorio su Ω a valori in $\mathcal{X}$

Teor

X è coppia aleatoria se e solo se X_i è un ente aleatorio a valori in $\mathcal{X}_i$ $(i=1,2)$

dim

Poiché $\mathcal{N} = \sigma(\mathcal{R})$ per criterio standard di misurabilità basta osservare che $X_1^{-1}(A_1) = X^{-1}(A_1 \times \mathcal{X}_2)$, $X_2^{-1}(A_2) = X^{-1}(\mathcal{X}_1 \times A_2)$ e $X^{-1}(A_1 \times A_2) = X^{-1}(A_1) \cap X^{-1}(A_2)$ $\forall A_1 \in \mathcal{N}_1$ e $\forall A_2 \in \mathcal{N}_2$

in quanto $\mathcal{R}$ ha come sist di generatori $\{A_1 \times A_2 \mid A_i \in \mathcal{R}_i\}$, perciò basta dim la $(\mathcal{A}, \mathcal{N})$-misurabilità sui sist di generatori.

$X^{-1}(A_1 \times A_2) = \{\omega \mid X(\omega) \in A_1 \times A_2\} = \{\omega \mid X_1(\omega), X_2(\omega) \in A_1 \times A_2\} = \{\omega \mid X_1(\omega) \in A_1\} \cap \{\omega \mid X_2(\omega) \in A_2\} = \{\omega \mid \omega \in X_1^{-1}(A_1) \cap X_2^{-1}(A_2)\}$

Chiamiamo **Legge (o distribuzione) congiunta** di X_1, X_2 la legge P_X di X in coppia aleatoria. $P_X = P_{X_1} \cdot P_{X_2}$ Legge congiunta della coppia aleatoria

Supponendo ci sia una misura di riferimento μ_i su N_i $i=1,2$ chiamiamo **(μ_1,μ_2)-densità congiunta** di X_1, X_2 ogni funzione (N,B)-misurabile $f:$ $\mathcal{X} \to [0,+\infty]$: $P_X(A) = P(X\in A) = \int_A f\, d(\mu_1\times\mu_2) = \int_A f(x_1,x_2)(dx_1\times dx_2)$

Chiamiamo **legge marginale** di X_i la legge P_{X_i};
Chiamiamo **μ_i-densità marginale** di X_i una μ_i-densità di X_i $i=1,2$
Chiamiamo **funzione di ripartizione congiunta** delle variabili (o v.) aleatorie X_1, X_2 la funzione $F_X(x_1,x_2) = P_X(]-\infty,x_1]\times]-\infty,x_2]) = P((X_1,X_2)\le(x_1,x_2))$

Schema della legge dell'ente aleatorio

$(\Omega,\mathcal{A},P)$

- v.a. X_1: $\mathcal{X}_1$ $(\mathcal{A},N_1)$-misurab; N_1; P_{X_1}; μ_1-mis rif $\Rightarrow \begin{cases} f_1 \ge 0 \\ P_{X_1}(A) = \int_A f_1\, d\mu_1 \end{cases}$
- v.a. X_2: $\mathcal{X}_2$ $(\mathcal{A},N_2)$-misurab; N_2; P_{X_2}; μ_2-mis rif $\Rightarrow$ f_2, $P_{X_2}(A) = \int_A f_2\, d\mu_2$

$\Rightarrow$ legge congiunta coppia aleatoria
$P_X = P_{X_1}\cdot P_{X_2}$

$\Rightarrow$ e dopo aggiunta delle densità marginali sorgerà la densità congiunta
$\mu = \mu_1\times\mu_2$, $\mu(A) = \int_{\mathcal{X}} \mu_2(A(w_1))\, \mu_1(dw_1)$
da cui
$f: \mathcal{X} \to \mathbb{R}^+$ $\quad f \ge 0$
$P_X(A) = \iint_A f\, d\mu_1\times\mu_2$

Bastano ∞ μ-densità congiunte (caso univariato)

Teorema fondamentale del calcolo delle probabilità (caso bivariato)
$g: \mathcal{X}\to\mathbb{R}^+$ una funzione sommabile rispetto alla legge congiunta di X_1 e X_2 allora posto $\mu=\mu_1\times\mu_2$, si ha:

$$\int_{\{X\in A\}} g(X_1,X_2)\, dP = \begin{cases} \int_A g\, dP_X \\ \int_A g f\, d\mu \quad \text{se } f \text{ è una } \mu\text{-densità di } X_1, X_2 \end{cases} \qquad \forall A\in N$$

Data FdR congiunta come trovo le marginali? Una porta vuota $\lim_{x\to\infty}$ quello sono

Teor **Risultato**: $P_{X_1}(A_1) = P_X(A_1\times\mathcal{X}_2)$, $P_{X_2}(A_2) = P_X(\mathcal{X}_1\times A_2)$ $\forall A_i\in N_i$
Se f è una $(\mu_1\cdot\mu_2)$ densità congiunta di X_1, X_2 la funzione

$$f_{X_i}(x_i) = \begin{cases} \int_{\mathcal{X}_2} f(x_1,x_2)\, \mu_2(dx_2) & \text{se } i=1 \\ \int_{\mathcal{X}_1} f(x_1,x_2)\, \mu_1(dx_1) & \text{se } i=2 \end{cases}$$

distinguo l'altra variabile con integrazione

$\forall x_i\in\mathcal{X}_i$, la funzione è una μ_i-densità marginale di X_i $(i=1,2)$

def. marg. $\quad$ $\mathcal{X}_1 \times \mathcal{X}_2 \ \ (\Lambda_2 = \mathcal{X}_2)$ $\quad$ def. congiunta

Dim. • Dato $\Lambda_1 \in \mathcal{A}_1$, risulta $P_{X_1}(\Lambda_1) \overset{\cdot}{=} P(X_1 \in \Lambda_1) \overset{\cdot}{=} P(\underline{X} \in \Lambda_1 \times \mathcal{X}_2) = P_{\underline{X}}(\Lambda_1 \times \mathcal{X}_2)$

•• Se f una (μ_1, μ_2)-densità congiunta di X_1, X_2. Dal teor. di Tonelli otteniamo

$$P_{X_1}(\Lambda_1) = P_{\underline{X}}(\Lambda_1 \times \mathcal{X}_2) = \underset{\text{def densità congiunta}}{\iint_{\Lambda_1 \times \mathcal{X}_2}} f \, d\mu_1 \times \mu_2 \underset{\text{Tonelli}}{=} \int_{\Lambda_1} \left(\int_{\mathcal{X}_2} f(x_1, \cdot) \, d\mu_2 \right) \mu_1(dx_1) \quad \forall \Lambda_1 \in \mathcal{A}_1$$

passo prec.

La legge che associa a Λ_1 l'integrale rispetto a μ_1 di una funzione non negativa $\Rightarrow$ è una μ_1-densità

Conseguentemente f_{X_1} è una μ_1-densità marginale di X_1.

Indipendenza

Introduzione:

1ºm. ENS. nel caso di var. aleat. X_1 indip. X_2: $P(X_1 \le x_1 \mid X_2 \le x_2) = P(X_1 \le x_1)$

cioè $P(X_1 \in]-\infty, x_1] \mid X_2 \in]-\infty, x_2]) = F_{X_1}(x_1)$

2ºm. ENS. • $P(X_1 \le x_1, X_2 \le x_2) = P(X_1 \le x_1) \cdot P(X_2 \le x_2)$ cioè

$$P(\underline{X} \in]-\infty, x_1] \times]-\infty, x_2]) = P(\underbrace{X_1 \in]-\infty, x_1]}_{\text{rett. mis. di } \mathcal{A}_1}) \cdot P(\underbrace{X_2 \in]-\infty, x_2]}_{\text{rett. mis. in } \mathcal{A}_2})$$

allora vorremo che rett. × rett. $\in \mathcal{A}$

$$\overset{\Delta}{=} P(\underline{X} \in \underset{\mathcal{A}_1}{\Lambda_1} \times \underset{\mathcal{A}_2}{\Lambda_2}) = P(X_1 \in \Lambda_1) \cdot P(X_2 \in \Lambda_2)$$

*Due var. aleatorie X_1, X_2 sono indipendenti se $P_{\underline{X}}(\Lambda_1 \times \Lambda_2) = P_{X_1}(\Lambda_1) \cdot P_{X_2}(\Lambda_2)$ $\forall$ rettangolo misurabile $\Lambda_1 \times \Lambda_2$

area del rettangolo = prodotto lunghezze dei lati; già so che la misura prodotto è l'unica che ha sto requisito

Teor.

Sono equivalenti le proposizioni: $\quad$ / a fine corso vedremo il perché di questo nome, come prodotto marginali

• X_1 e X_2 sono "indipendenti"

•• $P_{\underline{X}} = P_{X_1} \cdot P_{X_2}$ a meno di insiemi trascurabili

••• Se $f_{\underline{X}}$ è una (μ_1, μ_2)-densità congiunta di X_1, X_2 la • è equivalente a $f_{\underline{X}}(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$ a meno di un insieme $\mu_1 \times \mu_2$-trascurabile

Teor. $\quad$ funzione costante

Siano $\mathcal{G}_i$ σ-algebra su Ω_i e $g_i : \mathcal{X}_i \to \Omega_i$ un'applicazione $(\mathcal{A}_i, \mathcal{G}_i)$-misurabile $(i = 1, 2)$. Siano inoltre X_1, X_2 indipendenti. Allora sono pure indipendenti gli enti aleatori trasformati $g_1(X_1)$ e $g_2(X_2)$

Dim.

$$P_{\underline{Y}}(G_1 \times G_2) \overset{\text{def}}{=} P((g_1(X_1), g_2(X_2)) \in G_1 \times G_2) = P(\{g_1(X_1) \in G_1\} \cap \{g_2(X_2) \in G_2\}) =$$

$$= P(\{X_1 \in g_1^{-1}(G_1)\} \cap \{X_2 \in g_2^{-1}(G_2)\}) = P((X_1, X_2) \in g_1^{-1}(G_1) \times g_2^{-1}(G_2)) =$$

$$\overset{\text{def legge cong.}}{=} P_{\underline{X}}(g_1^{-1}(G_1) \times g_2^{-1}(G_2)) \overset{\text{indip.}}{=} P_{X_1}(g_1^{-1}(G_1)) \cdot P_{X_2}(g_2^{-1}(G_2)) =$$

$(g \circ f)^{-1}(\cdot) = f^{-1}(g^{-1}(\cdot))$

$$= P(X_1^{-1}(g_1^{-1}(G_1))) \cdot P(X_2^{-1}(g_2^{-1}(G_2))) = P((g_1 \circ X_1)^{-1}(G_1)) \cdot P((g_2 \circ X_2)^{-1}(G_2)) =$$

$$= P_{g_1(X_1)}(G_1) \cdot P_{g_2(X_2)}(G_2) \quad \forall \text{ rettangolo misurabile } G_1 \times G_2 \in \mathcal{G}_1 \otimes \mathcal{G}_2$$

Teor. $\quad$ ovvero $X : \mathbb{R} \to [0, +\infty[$

Siano X_1, X_2 v.a. indipendenti. Allora se $X_i \ge 0$ $(i = 1, 2)$ o $E(X_i)$ finita $i = 1, 2$ esiste la speranza matematica del prodotto e si ha $E(X_1 \cdot X_2) = E(X_1) E(X_2)$

Dim. Risulta $P_{\underline{X}} = P_{X_1} \cdot P_{X_2}$;

• $X_i \ge 0$ $(i = 1, 2)$, allora $P_{X_i}([0, +\infty[) = P_{\underline{X}}([0, +\infty[\times [0, +\infty[) = 1 \quad (i = 1, 2)$

Sono ben tramite teor. di Tonelli otteniamo

$$E(X_1 \cdot X_2) = \int_{\mathbb{R}^2} x_1 x_2 \, P_X(dx_1 \times dx_2) \overset{\text{Tonelli}}{=} \int_{[0,+\infty[\times[0,+\infty[} x_1 x_2 \, P_{X_1} \times P_{X_2}(dx_1 \times dx_2) =$$

$$= \int_{[0,+\infty[} \left(\int_{[0,+\infty[} x_2 \, P_{X_2}(dx_2) \right) x_1 \, P_{X_1}(dx_1) = \left(\int_{[0,+\infty[} x_1 \, P_{X_1}(dx_1) \right) \left(\int_{[0,+\infty[} x_2 \, P_{X_2}(dx_2) \right) =$$

$$= \left(\int_{\mathbb{R}} x_1 \, P_{X_1}(dx_1) \right) \left(\int_{\mathbb{R}} x_2 \, P_{X_2}(dx_2) \right) = E(X_1) \, E(X_2)$$

- $E(X_i)$ finita $(i = 1,2)$. Allora $|X_1|, |X_2|$ (non negativi) sono indipendenti e quindi poiché per ... ci sono indip. sono indip. anche
$$E(|X_1 \cdot X_2|) = E(|X_1| \cdot |X_2|) = E(|X_1|) \cdot E(|X_2|) < +\infty \quad \text{perché } X_i \text{ finita } i = 1,2.$$
Ne segue che $X_1 \cdot X_2$ ha speranza matematica finita$^{(*)}$. Quindi esiste finita.
Allora $f(x_1, x_2) = x_1 \cdot x_2$ è $P_{X_1} \times P_{X_2}$-integrabile (da Lebesgue)
$P_{X_1}(\{x_1 \in \mathbb{R} : f(x_1, \cdot) \text{ è } P_{X_2}\text{-sommabile}\}) = P_{X_1}(\mathbb{R}) = 1$. Conseguentemente, per il teor. di Fubini,
$$E(X_1 \cdot X_2) = \int_{\mathbb{R}^2} x_1 \cdot x_2 \, P_{X_1} \times P_{X_2}(dx_1 \times dx_2) \overset{\text{Tonelli}}{=} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} x_2 \, P_{X_2}(dx_2) \right) x_1 \, dP_{X_1} \quad \text{e quindi } E(X_1 \cdot X_2) = E(X_1) E(X_2)$$
(definendo ... $\int_{\{x \in N\}} g(x)\, dP = \dots$) (come prima)

(*) perché $X_1 X_2 \le |X_1 X_2|$ per monotonia $E(X_1 X_2) \le E(|X_1 X_2|) < +\infty$

Cor

$A_1, A_2 \in \mathcal{A}$ sono due eventi indipendenti se e solo se I_{A_1}, I_{A_2} v.a. sono indipendenti

dim

per quanto appena dimostrato: $X = (I_{A_1}, I_{A_2})$, perché entrambe ...
$$E(I_{A_1}) \cdot E(I_{A_2}) = E(I_{A_1} \cdot I_{A_2}) = E(I_{A_1 \cap A_2}) = P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$$
ritroviamo così la vecchia nozione di probabilità

Le definizioni date si estendono "pari pari" alle n-uple aleatorie ($n \ge 3$) e rimangono validi tutti i risultati ottenuti

Abbiamo visto che X_1, X_2 enti aleatori indip. $\Leftrightarrow P_X = P_{X_1} \cdot P_{X_2}$

$$X_1, \dots, X_n \text{ indipendenti} \Leftrightarrow P_X = P_{X_1} \times \dots \times P_{X_n} \quad \text{con } X = (X_1, \dots, X_n)$$

ma se ho una successione?

Intanto osserviamo che se gli enti aleatori $X_1, \dots, X_n$ sono indipendenti, allora lo sono pure quelli ottenuti facendo una selezione tra di essi. Considerato la k-upla aleatoria $Y = (X_1, \dots, X_k)$ si ha $Y \in (A_1 \times \dots \times A_k) \Leftrightarrow X \in (A_1 \times \dots \times A_k \times \mathcal{X}_{k+1} \times \dots \times \mathcal{X}_m)$
e quindi
$$P_Y(A_1 \times \dots \times A_k) = P_X(A_1 \times \dots \times A_k \times \mathcal{X}_{k+1} \times \dots \times \mathcal{X}_m) =$$
$$= P_{X_1}(A_1) \dots P_{X_k}(A_k) \cdot P_{X_{k+1}}(\mathcal{X}_{k+1}) \dots P_{X_m}(\mathcal{X}_m)$$
$$= P_{X_1}(A_1) \dots P_{X_k}(A_k)$$

Data una successione di enti aleatori $X_1, X_2, \dots$ si dicono indipendenti se sono indipendenti $X_{k_1}, \dots, X_{k_n}$ $\forall$ scelta di $k_1, \dots, k_n$ distinti e di n

ad es. $A_1, \dots, A_n, \dots \in \mathcal{A}$ eventi indipendenti $\Leftrightarrow I_{A_1}, \dots, I_{A_n}, \dots$ v.a. indipendenti
ed esse lo sono se $P(A_{k_1} \cap \dots \cap A_{k_m}) = P(A_{k_1}) \dots P(A_{k_m})$

Convergenze di variabili aleatorie

Come | fa | fatto
+ Σ pp / caso

Non vedremo successioni di var. aleatorie. Le successioni di variabili aleatorie servono per descrivere fenomeni come "tempo di attesa prima di ottenere testa nel lancio di una moneta", "tempo di attesa per uscita del 2 nel gioco del lotto".

$X_1, \ldots, X_n, \ldots$ $(\mathcal{A}, \mathcal{B})$-misurabili $\Omega \xrightarrow{X_n} \mathbb{R}\ (\mathbb{R}^*)$, $\mathcal{A}$, P, $\mathcal{B}\ (\mathcal{B}^*)$

succ. di variabili aleatorie

Oss. fix $\omega \in \Omega$ $\forall \omega \in \Omega$ $X_1(\omega), X_2(\omega), \ldots, X_n(\omega), \ldots$ è succ. numerica. Mi chiedo $\exists \lim_{n \to +\infty} X_n(\omega) = \underbrace{X(\omega)}_{n° \text{ certo}}$?
Posso raccogliere gli ω per cui limite $\exists$ finito $\{\omega \mid \lim_{n \to +\infty} X_n(\omega) = X(\omega)\}$ che è successione numerica

$(X_n)_{n \geq 1}$, $X : \Omega \to \mathbb{R}$, denoto $\{\omega \mid \lim_{n \to +\infty} X_n(\omega) = X(\omega)\}$ con $\{X_n \to X\}$ l'insieme di convergenza della succ. X_n convergente ad X
e l'insieme dove la succ. numerica converge ad un n° certo

Se $\{X_n \to X\} = \Omega$ allora si parla di convergenza certa. Questa è la convergenza puntuale dell'analisi: una succ. di funzioni converge puntualmente ad una funzione limite se in ogni punto la succ. converge al valore $X_n \to X$. Cioè $\forall \omega \in \Omega$ $X_n(\omega) \to X(\omega)$

Se non c'è convergenza su ogni $\omega \in \Omega$, ma l'insieme degli ω per cui non c'è convergenza ha misura nulla, allora diremo c'è convergenza quasi certa, cioè una succ. di v.a. converge q.c. a un valore numerico.

Convergenza quasi certa

Una successione di v.a. $(X_n)_{n \geq 1}$ converge quasi certamente alla v.a. X se $X_n \to X\ (P\text{-q.c.})$ cioè $P(\{X_n \to X\}) = 1$ ove $\{X_n \to X\} = \{\omega \mid X_n(\omega) \to X(\omega)\}$ è l'insieme di convergenza

Teor \ Oss.

$$\{X_n \to X\} = \{\omega \mid \lim_{n \to +\infty} X_n(\omega) = X(\omega)\}$$ si traduce insiemisticamente così:

$$= \{\omega \in \Omega : \forall n' \exists n \ \forall m \geq n \ (|X_m(\omega) - X(\omega)| \leq 1/n')\} =$$

fatto così con $\forall \varepsilon > 0\ \exists n\ \forall m \geq n\ |X_m(\omega) - X_n(\omega)| \leq \varepsilon$ con $\varepsilon = 1/n'$, perché $n' \to +\infty$, $\varepsilon \to 0$

$$= \bigcap_{n' \geq 1} \{\omega \in \Omega \mid \exists n\ \forall m \geq n\ (|X_m(\omega) - X(\omega)| \leq 1/n')\}$$

$$= \bigcap_{n' \geq 1} \bigcup_{n \geq 1} \{\omega \in \Omega \mid \forall m \geq n\ (|X_m(\omega) - X(\omega)| \leq 1/n')\}$$

$$= \bigcap_{n' \geq 1} \bigcup_{n \geq 1} \bigcap_{m \geq n} \{\omega \in \Omega \mid |X_m(\omega) - X(\omega)| \leq 1/n'\}$$

Mi chiedo: • fix n'
• $X_1, \ldots, X_n \mid X_m, \ldots$
• $|\cdot| < 1/n'$
$\Rightarrow$ collezione gli ω

1) Considero solo $\underline{\quad\quad}_{n' \geq 1}$

$$\left(\bigcup_{n \geq 1} \bigcap_{m \geq n} \{|X_m - X| \leq 1/n'\}\right)_{n' \geq 1} = f(n')$$, intuitivamente

all'aumentare di n' la banda si assottiglia, quindi l'insieme degli ω sarà non crescente. Dalla continuità dall'alto della misura ho che

$\Rightarrow P\left(\bigcup_{n\ge1}\bigcap_{m\ge n}\{|X_m - X| < 1/n'\}\right) \uparrow P(X_n \to X) \qquad n' \to +\infty$

2) Considero ora $n \geq 1$

$\left(\bigcap_{m\ge n}\{|X_m-X|<\frac{1}{n'}\}\right)_{n\ge1} = f(n)$, intuitivamente

più aumenta n, minore è il numero di insiemi su cui interseco, avrò un insieme più ampio via via. Quindi l'unione degli omega è non decrescente. Dalla continuità dal basso risulta

$\Rightarrow P\left(\bigcap_{m\ge n}\{|X_m-X|<1/n'\}\right) \uparrow P\left(\bigcup_{n\ge1}\bigcap_{m\ge n}|X_m-X|<1/n'\right) \qquad n\to+\infty$

Teor caratterizzazione convergenza quasi certa

Le seguenti proposizioni sono fra loro equivalenti:

i) $X_n \to X$ (P-q.c.)

ii) $\lim_{n\to+\infty} P\left(\bigcap_{m\ge n}\{|X_m-X|<\varepsilon\}\right) = 1 \quad \forall \varepsilon>0$ — non P-q.c. perché c'è un limite fra i piedi

iii) $\lim_{n\to+\infty} P\left(\bigcap_{m\ge n}\{|X_m-X|\le\varepsilon\}\right)=1 \quad \forall\varepsilon>0$

dim/

sia $\varepsilon>0$ e n' t.c. $\frac{1}{n'} < \varepsilon$:

- i) ⇒ ii) poiché $X_n\to X$ (P-q.c.) $\Rightarrow \{X_n\to X\} \subseteq \bigcup_{n\ge1}\bigcap_{m\ge n}\{|X_m-X|<1/n'\}$ perché $\{X_n\to X\}$ ha un "$\cap$" in più. Dalla monotonia:

$1 = P(\{X_n\to X\}) \overset{②}{\le} P\left(\bigcup_{n\ge1}\bigcap_{m\ge n}\{|X_m-X|<1/n'\}\right) \le P\left(\bigcup_{n\ge1}\bigcap_{m\ge n}\{|X_m-X|<\varepsilon\}\right) =$
$= \lim_{n\to+\infty} P\left(\bigcap_{m\ge n}\{|X_m-X|<\varepsilon\}\right)$

- ii) ⇒ iii) $\lim_{n\to+\infty} P\left(\bigcap_{m\ge n}\{|X_m-X|<\varepsilon\}\right)=1 \quad \forall\varepsilon>0$, poiché

$\bigcap_{m\ge n}\{|X_m-X|<\varepsilon\} \subseteq \bigcap_{m\ge n}\{|X_m-X|\le\varepsilon\}$, dalla monotonia $P(\dots) \le P(\dots)$ da cui

$1=\lim_{n\to+\infty} P\left(\bigcap_{m\ge n}\{|X_m-X|<\varepsilon\}\right) \le \lim_{n\to+\infty} P\left(\bigcap_{m\ge n}\{|X_m-X|\le\varepsilon\}\right)$

- iii) ⇒ i) $P(X_m\to X) \overset{①}{=} \lim_{n'\to+\infty} P\left(\bigcup_{n\ge1}\bigcap_{m\ge n}\{|X_m-X|\le 1/n'\}\right) =$

$\overset{②}{=} \lim_{n'\to+\infty}\left[\lim_{n\to+\infty} P\left(\bigcap_{m\ge n}\{|X_m-X|\le 1/n'\}\right)\right] = 1$

pongo $\varepsilon = \frac{1}{n'}$, e ottengo la 1)

Teor unicità quasi certa del limite e convergenza P-q.c. delle trasformate

Sia $X_n\to X$ P-q.c., allora:

a) se $X_n\to Y$ P-q.c., allora $X=Y$ (P-q.c.)

b) se $g:\mathbb{R}\to\mathbb{R}$ è una funzione continua, allora $g(X_n)\to g(X)$ (P-q.c.)

dim/

a) $X_n\to X$ (P-q.c.) o $P(\{X_n\to X\})=1 \overset{\text{Tesi}}{\Rightarrow} X=Y$ (P-q.c.) cioè $P(X=Y)=1$
$X_n\to Y$ (P-q.c.) o $P(\{X_n\to Y\})=1$

per notazione
$\Rightarrow X_n\to X = \{\omega \mid X_n(\omega)\to X(\omega)\}$ e
$X_n\to Y = \{\omega \mid X_n(\omega)\to Y(\omega)\}$

unicità limite in $\mathbb{R}$
$= \{\omega \mid (X_n\to X)\cap(X_n\to Y)\} \subseteq \{\omega \mid X(\omega)=Y(\omega)\}$, monotonia

$P(\{X_n \to X\} \cap \{X_n \to Y\}) \leq P(\{X = Y\})$ $\quad (P(A \cap B) = P(A) + P(B) - P(A \cup B))$

$\underbrace{P(\{X_n \to X\})}_{=1} + \underbrace{P(\{X_n \to Y\})}_{=1} - \underbrace{P(\{X_n \to X\} \cup \{X_n \to Y\})}_{\leq 1 \text{ per monotonia}} \leq P(\{X = Y\})$

$1 \leq P(\{X = Y\})$ cioè $X = Y$ (P-q.c.)

b) $\{X_n \to X\}$ (P-q.c.) $\quad \overset{\text{per ipotesi}}{\{\omega \mid X_n(\omega) \to X(\omega)\}} \subseteq \{\omega \mid g(X_n(\omega)) \to g(X(\omega))\}$ (leggere le successioni convergono alla sua limite)

per monotonia $1 \leq P\{g(X_n) \to g(X)\}$ cioè $g(X_n) \to g(X)$ (P-q.c.)

Introduzione alla convergenza in probabilità

Teor Bernoulli (o legge empirica del caso)

$E_1, \dots, E_n, \dots$ eventi indipendenti ed equiprobabili (p)

allora

$\varepsilon > 0 \quad S_n = \sum_{i=1}^{n} |E_i| \quad \lim_{n \to +\infty} P\left(\left|\frac{S_n}{n} - p\right| < \varepsilon\right) = 1$ freq. relativa si discosta dalla comune probabilità meno di ε

Generalizzo con $p \to X$ v.a. gli, $\frac{S_n}{n} \to X_n$ v.a. gli:

- $\varepsilon > 0 \quad \lim_{n \to +\infty} P(|X_n - X| < \varepsilon) = 1$ considero 1 v.a. e faccio il limite
- $\varepsilon > 0 \quad \lim_{n \to +\infty} P\left(\bigcap_{m \geq n} |X_m - X| < \varepsilon\right) = 1$ considero in blocco le X_m che seguono X → è conv. quasi certa.

Osservo che

$\bigcap_{m \geq n} |X_m - X| < \varepsilon \Rightarrow |X_n - X| < \varepsilon$ perché ad es. se $m = n$ vale l'implicazione

per monotonia $1 = P\left(\bigcap_{m \geq n} |X_m - X| < \varepsilon\right) \leq P(|X_n - X| < \varepsilon)$

⇒ Morale: se vale c.q.c. vale il teor di Bernoulli generalizzato. Arrivo così alla nozione di convergenza in probabilità.

Convergenza in probabilità

Una successione $(X_n)_{n \geq 1}$ di v.a. converge in probabilità alla v.a. X (in simboli $X_n \xrightarrow{P} X$) se $\forall \varepsilon > 0$ risulta $\lim_{n \to +\infty} P(|X_n - X| < \varepsilon) = 1$ oppure $\lim_{n \to +\infty} P(|X_n - X| \leq \varepsilon) = 1$

Teor La convergenza quasi certa implica quella in probabilità

$P\left(\bigcap_{m \geq n} \{|X_m - X| < \varepsilon\}\right) \subseteq \{|X_n - X| < \varepsilon\}$

dim/ Dato $\varepsilon > 0$, dalla $\bigcap_{m \geq n} \{|X_m - X| < \varepsilon\} \subseteq \{|X_n - X| < \varepsilon\}$ valida $\forall n$, dal teor caratt conv. quasi certa e monotonia probabilità:

$1 = \lim_{n \to +\infty} P\left(\bigcap_{m \geq n} \{|X_m - X| < \varepsilon\}\right) \leq \lim_{n \to +\infty} P(|X_n - X| < \varepsilon)$

Lemma Se $X_n \xrightarrow{P} X$, allora $\exists$ una sottosuccessione convergente (P-q.c.) a X (!)

dim/ so che $\lim_{n \to +\infty} P(|X_n - X| < \varepsilon) = 1$ dato $\varepsilon > 0$, negando limite

$\lim_{n \to +\infty} P(|X_n - X| \geq \varepsilon) = 0$. Quindi $\forall \varepsilon > 0 \ \forall n \ \exists K_n \mid P(|X_{K_n} - X| \geq \varepsilon) < \frac{1}{2^n}$ (fisso ε, per def. limite $\forall \varepsilon' > 0 \ \exists K'_n \mid |a'_n| < \varepsilon'$ e $\varepsilon' = \frac{1}{2^n}$)

$\Rightarrow P\left(\bigcap_{m \geq n} \{|X_{K_m} - X| < \varepsilon\}\right) = 1 - P\left(\bigcup_{m \geq n} \{|X_{K_m} - X| \geq \varepsilon\}\right) \geq$ subadditività

$\geq 1 - \sum_{m \geq n} P(|X_{K_m} - X| \geq \varepsilon) \geq 1 - \sum_{m \geq n} \frac{1}{2^m} = 1 - \frac{1}{2^{n-1}}$ (serie geom. di rag. $\frac{1}{2}$)

quindi finire

$P\left(\bigcap_{m\geq n}\{|X_{k_m}-X|<\varepsilon\}\right)\geq 1-\frac{1}{2^{n-1}}$ / $\lim_{n\to+\infty}$

$\lim_{n\to+\infty} P\left(\bigcap_{m\geq n}\{|X_{k_m}-X|<\varepsilon\}\right)\geq 1$

cioè $X_{k_m}\to X$ (P-q.c)

$\left[\ \sum_{m\geq n}\frac{1}{2^m}=\frac{1}{2^n}+\frac{1}{2^{n+1}}+\ldots = \frac{1}{2^n}\left(1+\frac{1}{2}+\ldots\right) = \frac{1}{2^n}\left(\frac{1}{1-\frac{1}{2}}\right)=\frac{1}{2^n}(2)=\frac{1}{2^{n-1}}\ \right.$

Teor. caratterizzazione convergenza in probabilità

Le seguenti proposizioni sono equivalenti

i) $X_n \xrightarrow{P} X$

ii) ogni sottosuccessione di $(X_n)_{n\geq 1}$ ammette una (sotto)-sottosuccessione convergente (P-q.c) a X

per $(X_n)_{n\geq 1}$

dim

i) → ii) data sottosucc. $(X_{k_n})_{n\geq 1}$ risulta $X_{k_n}\xrightarrow{P} X$; dal lemma so che $\exists$ sottosuccessione di X_{k_n} convergente P-q.c a X

ii) → i) (x ASSURDO)

$\neg$i) → $\neg$ii) $X_{k_n} \not\xrightarrow{P} X$ cioè $\not\exists \lim_{n\to+\infty} P(|X_{k_n}-X|<\varepsilon)=1\ \ \forall\varepsilon>0$

cioè $\exists\varepsilon>0\ \forall m\ \exists n\geq m : |a_n|\geq\varepsilon$ la quale non potrà essere sottosucc. convergente:

$m=1\quad n_1\quad |a_{n_1}|\geq\varepsilon$

$m=2\quad n_2\quad |a_{n_2}|\geq\varepsilon$

...

$m=m\quad n_m\quad |a_{n_m}|\geq\varepsilon$

cioè $a_{n_1}, a_{n_2}, \ldots, a_{n_m}, \ldots$ per cui $|X_{n_1}-X|\geq\varepsilon$, $|X_{n_2}-X|\geq\varepsilon$, ...

cioè nessuna sottosuccessione è convergente a zero, perché tutti hanno elementi $\geq\varepsilon>0$. Quindi $\not\exists X_{a_n}\to X$. Quindi nessuna (sottosucc.) converge P-q.c a X (negando ii)

Teor. unicità quasi certa del limite (1/CQC) e convergenza in prob. della trasformata (NO, ORO)

(1-sto) Sia $X_n\xrightarrow{P} X$ risulta:

i) se $X_n\xrightarrow{P} Y$, allora $X=Y$ (P-q.c)

ii) se $g:\mathbb{R}\to\mathbb{R}$ è una funzione continua, allora $g(X_n)\xrightarrow{P} g(X)$ (da finire lemma di caratterizzazione)

dim

1) $\lim_{n\to+\infty} P(|X_n-X|<\varepsilon)=1$

$\lim_{n\to+\infty} P(|X_n-Y|<\varepsilon)=1$

$P(X=Y)=1$

$E\wedge F\to G$ "→" densità C

Dato n', $\forall m$ $\left(|X_n-X|<\frac{1}{n'}\right)\wedge\left(|X_n-Y|<\frac{1}{n'}\right)\Rightarrow\left(|X-Y|<\frac{2}{n'}\right)$

al massimo della distanza

per monotonia probabilità

$P(|X-Y|<2/n')\geq P\left((|X_m-X|<1/n')\cap(|X_n-Y|<1/n')\right) =$ $P(A\cap B)=P(A)+P(B)-P(A\cup B)$

$= P(\ldots)+P(\ldots)-P\left(\{|X_n-X|<1/n'\}\cup\{|X_n-Y|<1/n'\}\right)\geq$

$\geq P(\ldots)+P(\ldots)-1$ e passando al limite

cost rispetto n

$\Leftrightarrow \lim_{n\to+\infty} P(|X-Y|<2/n')\geq 1+1-1=1$. Ora $\{|X-Y|<2/n'\}_{n'\geq 1}$ è una crescente,

dalla continuità dall'alto $1=\lim_{n'\to+\infty} P(|X-Y|<2/n')=P\left(\bigcap_{n'\geq 1}|X-Y|<2/n'\right)$

Il passo finale è dire che

$P(X=Y)=P\left(\bigcap_{n'\geq 1}|X-Y|<2/n'\right)$

$\subseteq$ $X(\omega)=Y(\omega)\Rightarrow X(\omega)-Y(\omega)=0\Rightarrow\forall n'\geq 1\ |X(\omega)-Y(\omega)|<2/n'\Rightarrow\omega\in\bigcap_{n'\geq 1}|X(\omega)-Y(\omega)|<2/n'$

(non fai la prob)

$\supseteq$ $\omega\in\bigcap_{n'\geq 1}\{\omega\,|\,|X(\omega)-Y(\omega)|<2/n'\}\Rightarrow\forall n'\geq 1\ |X(\omega)-Y(\omega)|<2/n'$

$\&\ \lim_{n'\to+\infty} 0\leq|X(\omega)-Y(\omega)|\leq\lim_{n'\to+\infty} 2/n'=0$, allora $X(\omega)=Y(\omega)$ / per forza per il teorema

ii) Data una sottosuccessione $(g(X_{k_n}))_{n\geq1}$, allora per il lemma la sottosuccessione $(X_{k_n})_{n\geq1}$ ammette una sottosuccessione $(X_{k_{h_n}})_{n\geq1}$ convergente P-quasi certamente a X. Allora $(g(X_{k_{h_n}}))_{n\geq1}$ converge P-quasi (=KFc) certamente a $g(X)$. Ne segue, per l'arbitrarietà della sottosuccessione $(g(X_{k_n}))_{n\geq1}$ la tesi

Sintesi:
- convergenza certa
- convergenza quasi certa
- convergenza in probabilità
- convergenza in distribuzione

} unicità q.c. del limite | convergenza via via più debole

Leggi dei grandi numeri

Le leggi dei grandi numeri forniscono una formulazione matematica alla idea empirica che:
- la media di un gran numero di realizzazioni di v.a. indipendenti ed equidistribuite è verosimilmente vicina alla comune speranza matematica;
- la frequenza relativa di successo (rapporto tra il numero di eventi che si verificano tra quelli considerati e la numerosità di questi ultimi) di un gran numero di prove indipendenti ed equiprobabili è verosimilmente vicina alla comune speranza matematica (legge empirica del caso) → 2° che convergenza?

Date le variabili aleatorie $X_1, \dots, X_n$ poniamo $S_n = X_1 + \dots + X_n$. Se $E(X_1), \dots, E(X_n)$ sono finite, allora $E\left(\frac{S_n}{n}\right) = \frac{E(X_1) + \dots + E(X_n)}{n}$

Dati gli eventi $A_1, \dots, A_n$ poniamo $S_n = I_{A_1} + \dots + I_{A_n}$, allora $E\left(\frac{S_n}{n}\right) = \frac{P(A_1) + \dots + P(A_n)}{n}$ cioè la media degli indicatori è interpretabile come freq. di successi

La legge empirica del caso è formalizzata in modi diversi dalla legge debole/forte dei grandi numeri: la convergenza nel primo caso è in probabilità, nel secondo è quasi certa. $E\left(\frac{S_n}{n}\right)$, $\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right|$

Legge debole dei grandi numeri

Una successione $(X_n)_{n\geq1}$ di v.a. verifica la legge debole dei grandi numeri se $\frac{S_n}{n} - E\left(\frac{S_n}{n}\right) \xrightarrow{Pr} 0$ $\quad \left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| \xrightarrow{Pr} 0$

Teor. di Markov:
Sia la successione di v.a. $(X_n)_{n\geq1}$ t.c. $\lim_{n\to\infty} \frac{Var(S_n)}{n^2} = 0$, allora la successione verifica la legge debole dei grandi numeri.

dim. Posto $Y_n = \frac{S_n}{n} - E\left(\frac{S_n}{n}\right)$ $\forall n \geq 1$, dalla disuguaglianza di Čebyšev ho che
$$P(|Y_n - \underbrace{E(Y_n)}_{0}| \geq \varepsilon) \leq Var(Y_n)/\varepsilon^2, \quad P(|Y_n| \geq \varepsilon) \leq \frac{Var\left(\frac{S_n}{n} + \text{cost}\right)}{\varepsilon},$$
$$P(|Y_n| \geq \varepsilon) \leq \frac{Var(S_n)}{n^2 \cdot \varepsilon}$$
∴ $\lim_{n\to\infty}$ e per ipotesi ho che
$$\lim_{n\to\infty} P(|Y_n| \geq \varepsilon) = 0 \quad \text{cioè } \frac{S_n}{n} - E\left(\frac{S_n}{n}\right) \xrightarrow{Pr} 0$$

Teor. di Čebyšev
Sia la successione di v.a. $(X_n)_{n\geq1}$ t.c. $Cov(X_i, X_j) \leq 0$ $(i \neq j)$ ← a due v.a. comunque e $\lim_{n\to\infty} \frac{Var(X_1) + \dots + Var(X_n)}{n^2} = 0$ allora $(X_n)_{n\geq1}$ verifica la legge debole grandi numeri

dim/ Risulta $0 \leq \frac{Var(S_n)}{n^2} = \frac{\sum_{i=1}^{n} Var(X_i) + 2\sum_{i=1}^{n}\sum_{j>i} Cov(X_i, X_j)}{n^2} \leq \frac{\sum_{i=1}^{n} Var(X_i)}{n^2}$, al limite:

$0 \leq \lim_{n\to+\infty} \frac{Var(S_n)}{n^2} = \frac{\sum Var(X_i) + 2\sum\sum Cov}{n^2} \leq \lim_{n\to+\infty} \frac{\sum_{i}^{n} Var(X_i)}{n^2} \underset{\times ip}{=} 0$

Ottengo

$\lim_{n\to+\infty} \frac{Var(S_n)}{n^2} = 0$, da cui per Markov $X_1, \dots, X_n, \dots$ è convergente in probabilità e quindi verifica la legge debole dei grandi numeri

Os: Come caso particolare di questo teorema (con ipotesi + forti di queste) è che le succ di v.a. sono "indipendenti" ed hanno medesima distribuzione. Anche il teor di Bernoulli è un caso particolare / sottoprodotto di qst teorema, perché richiedeva l'indipendenza e l'equidistribuzione - basterebbe come condizioni per teor di Bernoulli la non correlazione e la medesima varianza $\forall X_n$

A sua volta l'ipotesi del teor precedente $\lim_{n\to+\infty} \frac{Var(X_1)+\dots+Var(X_n)}{n^2} = 0$ è verificata da · $\lim_{n\to+\infty} \frac{Var(X_n)}{n} = 0$

e da · $\sum_{n\geq1} \frac{Var(X_n)}{n^2} < +\infty$ (convergente)

Legge forte dei grandi numeri

Una successione $(X_n)_{n\geq1}$ di v. aleatorie verifica la legge forte dei grandi numeri se $\frac{S_n}{n} - E\left(\frac{S_n}{n}\right) \to 0$ (P-q.c) — si avrebbe parlando che lo scarto valga zero

** Teor di Rajchmann: Sia $Cov(X_i, X_j) \leq 0$ $(i \neq j)$ e le varianze $Var(X_n)$ $(n\geq1)$ equilimitate. Allora $(X_n)_{n\geq1}$ verifica la legge forte dei grandi numeri

Corollario: Risulta

i) Siano $(X_n)_{n\geq1}$ indipendenti, equidistribuite e con varianza (comune) finita. Allora $\frac{S_n}{n} \to E(X_1)$ (P-q.c) — $\frac{S_n}{n} - E(X_1) \to 0 \doteq \frac{S_n}{n} \to E(X_1)$

ii) (Borel) Siano gli eventi $(A_n)_{n\geq1}$ indipendenti ed equiprobabili allora $\frac{S_n}{n} \to P(A_1)$ (P-q.c) — la freq rel di successo converge alla probabilità comune

dim/ · $Var(A_n) = Var(I_{A_n})$ — · altresì ritroviamo la legge empirica del caso (caso P-q.c)

via integrali

$= E(I_{A_n}^2) - [E(I_{A_n})]^2 = P(A_n) - [P(A_n)]^2 = P(A_n) \cdot [1 - P(A_n)] \leq 1 \ \forall n$

· $E(I_{A_n}) = P(A_n) \ \forall n$ (?)

× Rajchmann

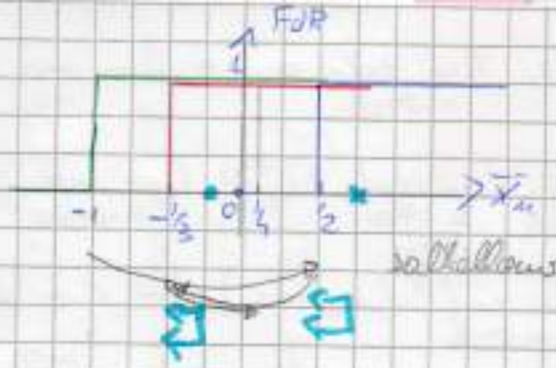
$\lim_{n\to+\infty} P\left(\bigcap_{m\geq n} \left|\frac{S_m}{m} - P(A_1)\right| \leq \varepsilon\right) = 1$ vuol dire $\frac{S_n}{n} \to P(A_1)$ P-q.c.

Convergenza in distribuzione

Introduzione: la convergenza anche di $(X_n)_{n\geq1}$ di v.a. non si traduce in una corrispondente convergenza (dell'analisi) sulle F di R. Per convincersi pensiamo a $X_n = (-1)^n/n$, $n\geq1$ e $X \equiv 0$

$F_{X_n}(x) = \begin{cases} 0 & \text{se } x < (-1)^n/n \\ 1 & \text{se } x \geq (-1)^n/n \end{cases}$ $\quad F_X(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$

$n=1$ $\quad X_1 = -1$
$n=2$ $\quad X_2 = 1/2$
$n=3$ $\quad X_3 = -1/3$
$n=4$ $\quad X_4 = 1/4$



- a sx di 0 le FdR convergono a 1
- a dx di 0 le FdR convergono a 0
  (fix • x da un certo pt in poi per n grande sono qui sopra)

$F_{X_n} \to 0 \quad \forall x \neq 0$, $F_{X_n}(0)$ non ammette limite ($\forall x$ e $\exists x$): $F_{X_{2n}}(0) = 0$, $F_{X_{2n+1}}(0) = 1$ e in 0 la X vale 1. La convergenza puntuale non implica la convergenza ($n \to \infty$) in distribuz.

↳ morale: c'è convergenza solo in ogni pt di continuità della funzione limite quindi nei pti di discontinuità delle FdR delle $(X_n)_{n \geq 1}$ nulla si può dire.

Uso del problema fra convergenza nell'analisi delle v.a. e convergenza delle FdR

Una successione di v.a. $(X_n)_{n \geq 1}$ converge in distribuzione alla v.a. X ($X_n \xrightarrow{d} X$) se $F_n(x) \to F(x)$ in ogni pt di continuità x di F, avendo indicato rispettivamente con F e F_n le funzioni di ripartizione delle v.a. X e X_n ($n \geq 1$)

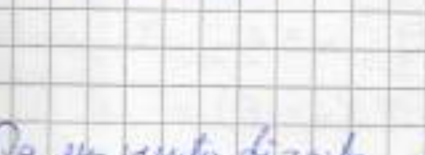
Oss: riguardo la conv. delle v.a. uso delle FdR, che è una succ. numerica

Da un punto di vista intuitivo la differenza sostanziale fra la convergenza quasi certa quella in probabilità e la convergenza in distribuzione può essere espressa dicendo che nelle prime due X_n tende ad essere uguale a X con alta probabilità, mentre nella convergenza in distribuzione X_n tende ad avere la medesima legge di X.

Lim sup e Lim inf

$a_1, \dots, a_n, \dots$ numeri reali (in disordine)

$\inf\{a_1, \dots, a_n, \dots\} = \alpha_1 \qquad \sup\{a_1, \dots, a_n, \dots\} = \beta_1$
$\inf\{\cancel{a_1}, a_2, \dots, a_n, \dots\} = \alpha_2 \qquad \sup\{\cancel{a_1}, a_2, \dots, a_n, \dots\} = \beta_2$
$\vdots$
$\inf\{\cancel{\ }, \dots, \cancel{\ }, a_n, \dots\} = \alpha_n \qquad \sup\{\cancel{\ }, \dots, \cancel{\ }, a_n, \dots\} = \beta_n$



↳ estremo inferiore o sale o rimane dov'è => successione non decrescente

↳ estremo superiore o rimane dov'è o cala => successione non crescente

Vale che $\alpha_n \leq a_n \leq \beta_n$, α_n e β_n succ monotone $\Rightarrow \exists$ limite

$\liminf_{n \to +\infty} (a_n) = \lim_{n \to +\infty} \alpha_n$ limite inferiore della succ

$\limsup_{n \to +\infty} (a_n) = \lim_{n \to +\infty} \beta_n$ limite superiore della succ.

$\liminf_{n \to +\infty} a_n = \limsup_{n \to +\infty} a_n = \lambda \iff \exists \lim_{n \to +\infty} a_n = \lambda$

Caratterizzazione della nozione di limite di una successione

$\liminf_{n\to+\infty} a_n \leq \limsup_{n\to+\infty} a_n$ (sempre)

Se $\liminf_{n\to+\infty} a_n \geq \limsup_{n\to+\infty} a_n \Rightarrow \limsup_{n\to+\infty} a_n = \liminf_{n\to+\infty} a_n$ (li uguali)

$\liminf_{n\to+\infty} (a_n + c) = c + \liminf_{n\to+\infty} (a_n)$

$\limsup_{n\to+\infty} (a_n + c) = c + \limsup_{n\to+\infty} (a_n)$

$|X_n - X| < \varepsilon$
$-\varepsilon < X_n - X < \varepsilon$
$-\varepsilon - X_n < -X < -X_n + \varepsilon$
$\varepsilon + X_n > X > X_n - \varepsilon$
$X_n - \varepsilon < X < X_n + \varepsilon$
$\Rightarrow X < X_n + \varepsilon$

Teor La convergenza in probabilità implica quella in distribuzione (non andare oltre inf)

dim/ Supposto $X_n \xrightarrow{Pr} X$, sia $\varepsilon > 0$. Allora $\lim_{n\to+\infty} P(|X_n - X| \geq \varepsilon) = 0$.

Dato x reale, risulta $|X_n - X| < \varepsilon \Rightarrow X < X_n + \varepsilon$, e quindi $\{X_n \leq x \wedge |X_n - X| < \varepsilon\} \subseteq \{X \leq x + \varepsilon\}$. Dalla additività e monotonia della probabilità risulta allora:

- $F_n(x) = P\big((X_n \leq x) \wedge |X_n - X| < \varepsilon) \vee (X_n \leq x) \wedge |X_n - X| \geq \varepsilon)\big) \overset{\text{disgiunti}}{=}$
$= P(X_n \leq x \wedge |X_n - X| < \varepsilon) + P(X_n \leq x \wedge |X_n - X| \geq \varepsilon) \leq$ monotonia
$\leq P(X \leq x + \varepsilon) + P(|X_n - X| \geq \varepsilon)$ da cui otteniamo
$\limsup_{n\to+\infty} F_n(x) \leq P(X \leq x + \varepsilon) + \limsup_{n\to+\infty} P(|X_n - X| \geq \varepsilon) =$ so $\exists \lim_{n\to+\infty} P(\ )$ ripetiz.
$= P(X \leq x + \varepsilon) + \lim_{n\to+\infty} P(|X_n - X| \geq \varepsilon) = P(X \leq x + \varepsilon) + 0$
$\hookrightarrow \limsup_{n\to+\infty} F_n(x) \leq F(x + \varepsilon)$

- Inoltre $\{X \leq x - \varepsilon \wedge |X_n - X| < \varepsilon\} \subseteq \{X_n \leq x\}$. Ne segue
$F(x - \varepsilon) = P(X \leq x - \varepsilon \wedge |X_n - X| < \varepsilon) + P(X \leq x - \varepsilon \wedge |X_n - X| \geq \varepsilon) \leq$
$\leq P(X_n \leq x) + P(|X_n - X| \geq \varepsilon)$ da cui otteniamo
$F(x - \varepsilon) \leq \liminf_{n\to+\infty} P(X_n \leq x) + \liminf_{n\to+\infty} P(|X_n - X| \geq \varepsilon)$
$= \liminf_{n\to+\infty} P(X_n \leq x) + \lim_{n\to+\infty} P(|X_n - X| \geq \varepsilon) = \liminf_{n\to+\infty} P(X_n \leq x)$
$\hookrightarrow F(x - \varepsilon) \leq \liminf_{n\to+\infty} F_n(x)$

Sia ora x di continuità per F

$\Rightarrow \limsup_{n\to+\infty} F_n(x) \leq \lim_{\varepsilon\to0^+} F(x + \varepsilon) = F(x) = \lim_{\varepsilon\to0^+} F(x - \varepsilon) \leq \liminf_{n\to+\infty} F_n(x)$

$x-\varepsilon \quad x \quad x+\varepsilon$

se una funz. è > di un valore anche il suo limite lo sarà — prop. FdR (FC dx) — x è PC

da cui
$F_n(x) \to F(x)$ con $n \to +\infty$.

Il motivo per cui è importante la convergenza in distribuzione si evince dal seguente teorema.

Teor di caratterizzazione della convergenza in distribuzione (Portmanteau theorem, 1947, Alexandrof)

di valigia/pillola

Le seguenti proposizioni sono equivalenti:

- $X_n \xrightarrow{d} X$
- $\lim_{n\to+\infty} E(f(X_n)) = E(f(X))$, per ogni funzione $f: \mathbb{R} \to \mathbb{R}$ continua e limitata

se f è l'identità non vale xK non è limitata!

Se $(X_n)_{n\geq1}$ sono equilimitate, cioè assumono valori in un insieme limitato $[-K, +K]$, allora se $g(x) = x^m$ (limitata), FC, il Portmanteau assicura la convergenza dei momenti: i momenti centrali di qls ordine convergeranno al momento centrale del medesimo ordine della variabile aleatoria limite

C'è convergenza in distribuzione sse c'è convergenza delle sper. matematiche delle trasformate continue e limitate.
Esso collega la convergenza delle FdR nei pt continuità della FdR della v.a. limite con convergenza di speranze matematiche (e in particolare converg. momenti di qls nel caso di v.a. equilimitate)

15:00

Teor. unicità in legge (unic. limite) o in distribuzione del limite e convergenza in distribuzione della trasformata

Sia $X_n \xrightarrow{d} X$, sussistono allora le seguenti proposizioni:
i) se $X_n \xrightarrow{d} Y$, allora X, Y sono equidistribuite (uso $P_X = P_Y$) — unicità in equidistribuzione
ii) se $g: \mathbb{R} \to \mathbb{R}$ è una funzione continua, allora $g(X_n) \xrightarrow{d} g(X)$ — non serve limitata

dim/
i) TH $\forall x \in \mathbb{R}$ $F_X(x) = F_Y(x)$; allora fix x: tolgo da $\mathbb{R}$ tutti i pt di discontinuità per le F_X e F_Y, al massimo un numerabile (somma di 2 numerabili). $\mathbb{R}$ resta un continuo di pt.
Poiché $(x_n)_{n\geq1}$ tende ad un certo x reale [illegible], considero la successione decrescente $(x_n)_{n\geq1}$ convergente ad x e costituita da pt di continuità sia di F_X che di F_Y. Allora da $(x_n) \downarrow x$ e x_k pt di continuità per F_X e F_Y, ho che $\lim_{n\to+\infty} F_n(x_k) = F_X(x_k) = F_Y(x_k)$, allora
$$F_X(x) \underset{\text{FC a dx}}{=} \lim_{n\to+\infty} F_X(x_n) \underset{\text{sopra}}{=} \lim_{n\to+\infty} F_Y(x_n) \underset{\text{FC a dx}}{=} F_Y(x)$$

ii) Sia f una funzione continua e limitata. Allora, la funzione composta $f \circ g$ è pure continua e limitata. Ne segue per il Portmanteau theorem
$$E(f(g(X_n))) = E((f\circ g)(X_n)) \to E((f\circ g)(X)) = E(f(g(X)))$$
Data l'arbitrarietà della funzione f, abbiamo $g(X_n) \xrightarrow{d} g(X)$ (sempre per il medesimo teor)

La convergenza in distribuzione serve:

1) nell'ambito delle variabili aleatorie limitate la conv. in distrib. assicura la convergenza di tutti i possibili momenti

2) (motivo) a capire il teor limite centrale, la cui tesi è una convergenza in distribuzione, ed ha enorme importanza dal pdv calcolistico: "centrale" nel senso di importanza, permette di calcolare le proprietà delle v.a. S_n o $\frac{S_n}{n}$

Prima del T.L.C. vediamo un teorema su FdR (limite) continua

Teor. di Polya

Se la successione di f.d.r. $(F_n)_{n\geq1}$ converge ad una f.r. F continua allora la convergenza è uniforme

CONVERGENZA UNIFORME
$(f_n)_{n\geq1}: X \to \mathbb{R}$, $f: X \to \mathbb{R}$
$f_n \to f$ unif se
$\forall \varepsilon > 0$ $\exists n_\varepsilon \in \mathbb{N}$ $|f_n(x) - f(x)| < \varepsilon$ $\forall x \in X$ e $\forall n \in \mathbb{N}$
[illegible]
[illegible]
BASE f_n

Date le v.a. $X_1, \dots, X_n$ con speranza matematica α e varianza $\sigma^2 > 0$ finita, se chiamo somma ridotta di $X_1, \dots, X_n$ la seguente trasformata della v.a. S_n: $S_n^* = \frac{S_n - \alpha n}{\sqrt{\sigma^2 n}}$
Chiaramente $E(S_n^*) = 0$. Inoltre, se $X_1, \dots, X_n$ indipendenti, si ha $Var(S_n^*) = 1$

Idea TLC: standardizzo mia v.a., tolgo Σ Cov dividendo (successione standardizzata di S_n^*) allora esse convergono in distrib. a una qls v.a. $\sim N(0,1)$, ma la FdR di $N(0,1)$ è continua (è integrale) allora basta Polya

se $f(x) =$
i momenti
medesimo

C'è convergen
della trasfo
esso collega
con conver
di qls nel ca

Teor. unicit
convergenza

Sia X_n
i) se X
ii) se g:

dim.)
i) TH $\forall x \in \mathbb{R}$
al massimo
Poiché $(X$
$(X_n)_{n\ge1}$ conv
da $(x_n) \forall x$
$F_X(x) =$

ii) Sia f
E(f(
Dato
medesim

La convergen
1) nell'ambito
convergenza
2) (media) e cap
zione
importan

Prima del 1. Lr.

$F_{X_n}(y) \to F_X(y)$ y cont. per F_X

$F_{X_n}(z) \to F_Y(z)$ z cont. per F_Y

x qualsiasi $\overset{\text{Tesi}}{\Rightarrow} F_X(x) = F_Y(x)$

$\exists\, x_m \downarrow x$ $\quad x_m$ p. cont. di F_X, F_Y

$$\left.\begin{array}{l} F_{X_m}(x_m) \to F_X(x_m) \\ F_{X_m}(x_m) \to F_Y(x_m) \end{array}\right] \begin{array}{l} F_X(x_m) = F_Y(x_m) \\ \forall m \end{array}$$

$$F_X(x) = \lim_{m\to+\infty} F_X(x_m) = \lim_{m\to+\infty} F_Y(x_m) = F_Y(x)$$

pt 1) nuova dimostr.

Teor. di Polya

Se la successione di fdr $(F_n)_{n\ge1}$ converge ad una f.r. F continua allora la convergenza è uniforme

CONVERGENZA UNIFORME
$(f_n)_{n\ge1}: X \to \mathbb{R}$, $f: X \to \mathbb{R}$
$f_n \to f$ unif. se
$\forall \varepsilon > 0\ \exists n_\varepsilon \in \mathbb{N}\ |f_n(x) - f(x)| < \varepsilon$
$\forall x \in X$ e $\forall n \ge n_\varepsilon$

Date le v.a. $X_1, \dots, X_n$ con speranza matematica α e varianza $\sigma^2 > 0$ finita, si chiama somma ridotta di $X_1 \dots X_n$ la seguente trasformata della v.a. S_n: $S_n^* = \frac{S_n - \alpha n}{\sqrt{\sigma^2 n}}$

Chiaramente $E(S_n^*) = 0$. Inoltre, se $X_1 \dots X_n$ indipendenti, si ha $Var(S_n^*) = 1$

Idea TCL: standardizzo succ. v.a., Lindeberg-Lévy (...) allora esse convergono in distrib. a una qls v.a. $\sim N(0,1)$, ma la FdR di $N(0,1)$ è continua (è un integrale) allora scatta Polya

Teor limite centrale: Siano le v.a. $X_1, X_2, \dots$ indipendenti, equidistribuite e con varianza finita non nulla. Allora, la successione $(F_{S_n^*})_{n\geq1}$ delle f.d.r. delle somme ridotte converge uniformemente alla f.d.r. della distribuzione normale $N(0,1)$

In altri termini la successione $(S^*)_{n\geq1}$ converge in distribuzione a una qualsiasi v.a. distribuita secondo la distribuzione normale $N(0,1)$.

uniforme significa $\forall \varepsilon>0 \; \exists m \; \forall n>m \; |F_{S_n^*} - \Phi(x)| < \varepsilon \; \forall x$

Teor: Nelle ipotesi del teorema del limite centrale assumendo la finitezza e positività del momento centrale assoluto terzo $\beta = E(|X_1 - E(X_1)|^3)$, risulta:

$$\underbrace{\sup_{x\in\mathbb{R}} \left| F_{S_n^*}(x) - \Phi(x) \right|}_{} \leq \frac{K \cdot \beta}{\sqrt{(\sigma^2)^3}} \cdot \frac{1}{\sqrt{n}}$$

disuguaglianza di Berry-Essen con ($0{,}403 \leq K \leq 0{,}478$ allo stato attuale delle conosc.)

• il massimo errore commesso sostituendo FdR con FdR ridotta

↳ Per ottenere livello di precisione di K alla seconda cifra decimale, $n = 10'000$ osservazioni E^5 ...

Ma sperimentalmente si è visto che per le usuali distribuzioni statistiche si è visto che una buona approx si ottiene con $30 < n < 50$. (notare) Se sono asimmetriche il n° n va aumentato

Se consideriamo una success. di eventi indipendenti di medesima probabilità p, una buona approx la si ottiene prendendo $p(1-p)\cdot n \geq 5$

Inoltre nel caso di v.a. a valori interi, si ottiene in generale una migliore approssimazione di $P(S_n \leq x)$ sostituendo in $\frac{x - \alpha n}{\sqrt{\sigma^2 n}}$ il valore x con $x + 1/2$.

Il teorema del limite centrale giustifica l'approssimazione, che viene usualmente fatta nelle applicazioni pratiche, della legge di S_n con quella della distribuzione normale $N(0,1)$ nel caso di indipendenza ed equidistribuzione. Approssimazione che, indicata con Φ la funzione di ripartizione della Normale $N(0,1)$, si basa sulla seguente relazione $P(S_n \leq x) =$

$$= P\left(S_n^* \leq \frac{x - \alpha n}{\sqrt{\sigma^2 n}}\right) \underset{\text{se } n \text{ grande}}{\approx} \Phi\left(\frac{x - \alpha n}{\sqrt{\sigma^2 n}}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\alpha n}{\sqrt{\sigma^2 n}}} e^{-t^2/2}\, dt$$

Inoltre
$$\underbrace{P\left(\left|\frac{S_n}{n} - \alpha\right| \geq \varepsilon\right)}_{\text{come Cebicev}} \underset{\text{limite}}{\approx} 2\Phi\left(-\varepsilon\sqrt{\frac{n}{\sigma^2}}\right) = \sqrt{\frac{2}{\pi}} \int_{\varepsilon\sqrt{\frac{n}{\sigma^2}}}^{+\infty} e^{-t^2/2}\, dt$$
che fornisce un valore più preciso - per n grande - della disuguaglianza di Cebicev, => si usa questa proprietà.

Questo tipo di problemi va sotto "l'approssimazione normale", che consiste nel calcolare probabilità relative a S_n o $\frac{S_n}{n}$ tramite la Normale $N(0,1)$

1 Convergenza quasi certa → permettono di calcolare la f.r. media comune di una succ.
2 Convergenza in probab.
3 Convergenza in distribuz. - ha natura più calcolistica

Questi 3 risultati non ci permettono di calcolare la FdR comune delle $(X_n)_{n\geq1}$ a partire dalle FdR osservate, perché 1e2 permettono di individuare la speranza matematica, TCL (3) permette di trovare FdR_{S_n} o $FdR_{S_n^*}$. Lo farà il prox teorema.

Teor. di Glivenko-Cantelli

Le leggi dei grandi numeri e del teorema del limite centrale considerano il comportamento asintotico di, rispettivamente, media aritmetica e della somma ridotta.
Il prossimo teorema fornisce invece una formulazione matematica precisa dell'intuizione empirica che la funzione di ripartizione che distribuisce la probabilità in modo uguale su un gran n° di realizzazioni di v.a. indipendenti ed equidistribuite, è verosimilmente vicina/prossima alla comune funzione di ripartizione.

Date le var. aleatorie $X_1, \dots, X_n$ la funzione di dominio $\mathbb{R} \times \Omega$:

$$F^{(X_1,\dots,X_n)}(x,\omega) = \left(I_{\{X_1 \le x\}}(\omega) + \dots + I_{\{X_n \le x\}}(\omega)\right) \cdot \frac{1}{n} =$$
$$= \frac{I_{]-\infty,x]}(X_1(\omega)) + \dots + I_{]-\infty,x]}(X_n(\omega))}{n}$$

è detta funzione di ripartizione empirica (relativa a $X_1, \dots, X_n$)

semi (o campione) o empirica

$F^{(X_1,\dots,X_n)}(\cdot,\omega)$ è una funzione di ripartizione per ogni $\omega \in \mathbb{R}$, e quindi dipendendo dai casi elementari, è una funzione di ripartizione aleatoria); risulta infatti:

$$F^{(X_1,\dots,X_n)}(x,\omega) \le F^{(X_1,\dots,X_n)}(x',\omega), \text{ se } x < x' \quad \left(\text{perché } I_{]-\infty,x]} \le I_{]-\infty,x']}\right)$$
$$F^{(X_1,\dots,X_n)}(x,\omega) = 0 \text{ se } x < \min(X_1(\omega),\dots,X_n(\omega))$$
$$F^{(X_1,\dots,X_n)}(x,\omega) = 1 \text{ se } x \ge \max(X_1(\omega),\dots,X_n(\omega))$$
$$F^{(X_1,\dots,X_n)}(x^+,\omega) = F^{(X_1,\dots,X_n)}(x,\omega) \quad \left(\text{in quanto } \lim_{t \to x^+} I_{\{X_n \le t\}} = I_{\{X_n \le x\}}\right)$$

Oss. Se $x_1, \dots, x_n$ tutte uguali la FdR distrib. uniforme:

se però almeno 2 si coincidono allora non è più uniforme

Teor. di Glivenko-Cantelli

Siano le v.a. $X_1, X_2, \dots, X_n, \dots$ indipendenti con comune funzione di ripartizione F. Allora $\sup_{x \in \mathbb{R}} \left| F^{(X_1,\dots,X_n)}(x,\cdot) - F(x) \right| \to 0$ (P-q.c.)

l'errore commesso sost. FdR empirica con F

v.a. (ω non fix)

cioè

$$P\left(\left\{\omega \in \Omega \mid F^{(X_1,\dots,X_n)}(\cdot,\omega) \text{ converge uniformemente a } F\right\}\right) = 1$$

scelta fatta la ipotesi

con pratica certezza, la convergenza così ottenuta è di tipo uniforme risp. F.

Medie condizionate

Condizionamento ad eventi non trascurabili

Supposto $P(H)>0$ cioè H evento non trascurabile, proviamo $P(A|H) = \frac{P(A\cap H)}{P(H)}$ $\forall A \in \mathcal{A}$ è una probabilità:

dim/ $P(\Omega|H) = \frac{P(\Omega\cap H)}{P(H)} = \frac{P(H)}{P(H)} = 1$

$$P\left(\bigcup_{n\geq1} A_n | H\right) = \frac{P\left(\left(\bigcup_{n\geq1} A_n\right)\cap H\right)}{P(H)} = \frac{P\left(\bigcup_{n\geq1}(A_n\cap H)\right)}{P(H)} = \frac{\sum_{n\geq1} P(A_n\cap H)}{P(H)} = \sum_{n\geq1}\left(\frac{P(A_n\cap H)}{P(H)}\right) =$$
$$= \sum_{n\geq1} P(A_n|H) \quad \text{con } (A_n)_{n\geq1} \text{ a 2 a 2 disgiunti}$$

Quindi $P(\cdot|H)$ è una probabilità, detta probabilità condizionata a H.

Come posso calcolare in modo agevole è $\int X\, dP(\cdot|H)$, con X v.a.?

Teor: X var aleatoria. Allora X è $P(\cdot|H)$ sommabile se e solo se $X\cdot I_H$ è P-sommabile. Inoltre nel caso di sommabilità (o $X\geq0$) vale che:
$$\int_\Omega X\, dP(\cdot|H) = \frac{\int_H X\, dP}{P(H)}$$

$P(\{X=y\}\cap H) = \int_\Omega I_{\{X=y\}} I_H\, dP$

dim • X semplice e non negativa: v.a. $X = \sum_{y\in X(\Omega)} y\cdot I_{\{X=y\}}$

$$\int_\Omega X\, dP(\cdot|H) \overset{\text{def integrale}}{=} \sum_{y\in X(\Omega)} y\, P(X=y|H) \overset{\text{def}}{=} \sum_{y\in X(\Omega)} y\cdot \frac{P(\{X=y\}\cap H)}{P(H)} = \frac{1}{P(H)}\sum_{y\in X(\Omega)} y\cdot\left(\int_\Omega I_{\{X=y\}} I_H\, dP\right)$$
prop integr. positivo
$$= \frac{1}{P(H)}\int_\Omega\left(\sum_{y\in X(\Omega)} y\, I_{\{X=y\}}\cdot I_H\right)dP = \frac{1}{P(H)}\int_H\left(\sum_{y\in X(\Omega)} y\cdot I_{\{X=y\}}\right)dP = \frac{\int_H X\, dP}{P(H)}.$$

• v.a. X $(\mathcal{A},\mathcal{B})$-misurabile, $X\geq0$. Allora $0\leq X_n \uparrow X$ con X_n funzioni semplici. Allora $0\leq (X_n\cdot I_H)\uparrow X\cdot I_H$, ed ho che
$$\int_\Omega X\, dP(\cdot|H) \overset{\text{teor. monot.}}{=} \lim_{n\to\infty}\int_\Omega X_n\, dP(\cdot|H) \overset{\text{passo prec}}{=} \lim_{n\to\infty}\frac{\int_H X_n\, dP}{P(H)} = \frac{1}{P(H)}\lim_{n\to\infty}\int_H X_n\, dP \overset{\text{teor monot}}{=} \frac{\int_H X\, dP}{P(H)}$$

• v.a. X qls. da
$$\int_\Omega X^+ dP(\cdot|H) \overset{\text{passo prec}}{=} \frac{\int_H X^+ dP}{P(H)} = \frac{\int_\Omega (X^+ I_H)\, dP}{P(H)} = \frac{\int_\Omega (X\cdot I_H)^+ dP}{P(H)}$$

Oss $X^+ \cdot I_H = \begin{cases} X\cdot I_H & X\geq0 \\ -X\cdot I_H & X<0\end{cases}$ $\quad (X\cdot I_H)^+ = \begin{cases} X\cdot I_H & X\cdot I_H\geq0 \\ -X I_H & X\cdot I_H<0\end{cases}$

su H: $\begin{cases} X & X\geq0 \\ -X & X<0\end{cases}$ $\quad$ su H: $\begin{cases} X & X\geq0 \\ -X & X<0\end{cases}$

su H^c: $\begin{cases} 0 & X\geq0 \\ 0 & X<0\end{cases} \Rightarrow 0$ sempre $\quad$ su H^c: $\begin{cases} 0 & 0\geq0 \\ 0 & \text{no senso}\end{cases} \Rightarrow 0$ sempre

da
$$\int_\Omega X^- dP(\cdot|H) \overset{\text{passo prec}}{=} \frac{\int_H X^- dP}{P(H)} = \frac{\int_\Omega (X^-\cdot I_H)\, dP}{P(H)} = \frac{\int_\Omega (X\cdot I_H)^- dP}{P(H)}$$

la $P(\cdot|H)$ sommabilità di X equivale alla P-sommabilità di $X\cdot I_H$.
Inoltre nel caso di sommabilità $\int_\Omega X\, dP(\cdot|H) = \int_\Omega X^+ dP(\cdot|H) - \int_\Omega X^- dP(\cdot|H) =$
$$\overset{\text{passo prec}}{=} \frac{\int_H X^+ dP}{P(H)} - \frac{\int_H X^- dP}{P(H)} = \frac{\int_H X\, dP}{P(H)}.$$

Data una v.a. X con speranza matematica, la speranza matematica condizionata di X ad H è identificata dall'elemento della retta reale ampliata $E(X|H) = \int X\, dP(\cdot|H)$, allora per teor precedente vale che:
$$E(X|H) \underset{\text{Teor prec}}{=} \frac{\int_H X\, dP}{P(H)} \underset{\text{def } E(\cdot)}{=} \frac{E(X\cdot I_H)}{P(H)}$$

Condizionamento a σ algebre

• **Esempio introduttivo**: evidenzia l'intuizione per formulare speranza matematica condizionata alle informazioni ottenibili. Lo scopo: arrivare a problemi di condizionamento anche su insiemi trascurabili

Estrazione a caso di un numero X (v.a.) dall'intervallo $[0,1]$

$\Omega = [0,1]$, $\mathcal{A} = \mathcal{B} \cap [0,1]$, $P = \lambda|_{[0,1]}$, $X(\omega) = \omega$, $E(x) = 1/2$, $X \sim U(0,1)$, $F_X(R_x) = y - x$

Immagino che l'estrazione sia fatta altrove e che la $\mathbb{P}$ di Ω sia:

$H_1 = \{X < 1/4\}$ $\quad H_3 = \{X > 3/4\}$ $\quad \mathbb{P} = \{H_1, H_2, H_3, H_4\}$

$H_2 = \{1/4 < X < 3/4\}$ $\quad H_4 = \{X = 1/4 \vee X = 3/4\}$

↳ eventi osservabili (o informazione) è la info disponibile

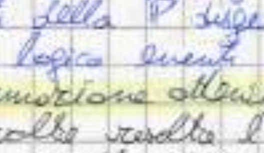
Ma io non prevedo il n° estratto

Qual è l'informazione che riesco ad ottenere da questi eventi? Una volta risolta l'incertezza su $\mathbb{P}$, di quali eventi potrò conoscere il valore di verità? $\mathcal{H}$, degli eventi di $\mathcal{H}$. $\mathcal{H} = \{ \bigcup_{j \in J} H_j : J \subseteq \{1,2,3,4\}\}$ insieme degli eventi della $\mathbb{P}$ disgiunti fra loro, con numerosità contabile. Poi tutt fondamentale della logica eventi $\mathcal{H} = \sigma(H_1, \dots, H_4)$ σ-algebra generata da $\mathbb{P}$ è detta informazione ottenibile ↳ se H_1 vero, allora $H_2 \vee H_3$ vero

Una volta risolta l'incertezza su $\mathbb{P}$ potrò dire il valore di verità solo degli eventi osservabili

X è una v.a. non osservabile

Lo scopo ora è calcolare $E(X \mid H_i)$ $i = 1,4$, già sapendo che $P(H_4) = 0$ e quindi non potrò calcolare $E(X \mid H_4)$. Perciò costruirò una intera struttura, condizionatamente alle $\mathcal{H}$ ind. di X, per risolvere questo problema



$P(H_1) = 1/4 \quad P(H_2) = 1/2 \quad P(H_3) = 1/4 \quad P(H_4) = 0$

$$E(X \mid H_1) = \frac{\int_0^{1/4} \omega\, d\omega}{P(H_1)} = \frac{\left(\frac{\omega^2}{2}\right)_0^{1/4}}{1/4} = \left(\frac{1}{16}\cdot\frac{1}{2} - 0\right)\cdot 4 = \frac{1}{8} \stackrel{\Delta}{=} y_1$$

$$E(X \mid H_2) = \frac{\int_{1/4}^{3/4} \omega\, d\omega}{P(H_2)} = \frac{\left(\frac{\omega^2}{2}\right)_{1/4}^{3/4}}{1/2} = \left(\frac{9}{16}\cdot\frac{1}{2} - \frac{1}{16}\cdot\frac{1}{2}\right)\cdot 2 = \frac{1}{2} \stackrel{\Delta}{=} y_2$$

$$E(X \mid H_3) = \frac{\int_{3/4}^{1} \omega\, d\omega}{P(H_3)} = \frac{\left(\frac{\omega^2}{2}\right)_{3/4}^{1}}{1/4} = \left(-\frac{9}{16}\cdot\frac{1}{2} + \frac{1}{2}\right)\cdot 4 = \frac{7\cdot 2}{16} = \frac{7}{8} = y_3$$

↳ A seconda dell'evento che si verifica, l'aggiornamento dell'informazione muta il valore atteso (solo 1 dei 3 accadrà)

Costruiamo una var. aleatoria osservabile, come mistura/posteriori dei y_i

$$Y = \sum_{i=1} y_i \cdot I_{H_i} = \frac{1}{8} I_{H_1} + \frac{1}{2} I_{H_2} + \frac{7}{8} I_{H_3} + \underbrace{y_4 \cdot I_{H_4}}_{\text{nulla}} \quad \text{arbitrario } y_4 \in \mathbb{R}; \quad Y: \Omega \to \mathbb{R}$$

Y associa ad H_i la sua speranza matematica condizionata

Y è $(\mathcal{H}, \mathcal{B})$-misurabile

dim/

verifico le controimmagini dei Boreliani $\in$ σ-algebra in quanto disgiunzioni degli eventi $H_1, \dots, H_4$: $\{Y \in]-\infty, x]\} = \{\emptyset\}$ se $x < 0$

H_1 se $x \leq 1/8$

$H_1 \cup H_2 \cup H_3$ se $x \leq 7/8$

Ω se $x \leq y_4$

} tutte contenute/stanno in $\mathcal{H} \Rightarrow (\mathcal{H}, \mathcal{B})$-mis.

Risolta l'incertezza su $\mathbb{P}$, conoscerò il valore di Y che si verificherà, cioè Y sarà osservabile

Scopriamo il legame fra X non oss. e Y oss.: (puntando a costruire $E(X \mid H) = \frac{E(X \cdot I_H)}{P(H)}$

mi concentro sul numeratore

a) $E(X \cdot I_H) = E(X \cdot I_{\bigcup_{j \in J} H_j}) \stackrel{\text{indicatori disgiunti}}{=} E(X \cdot \sum_{j \in J} I_{H_j}) = \sum_{j \in J \setminus \{4\}} E(X \cdot I_{H_j})$ in quanto

$H \in \mathcal{H}$ $\quad E(X \cdot I_{H_4}) = \int_{H_4} X\, dP = 0$ $\quad P(H_4) = 0$

$$\int_H X\, dP = \sum_j \int_{H_j} X\, dP = \sum_{j \in J \setminus 4} \left(\int_{H_j} X\, dP : P(H_j)\right) \cdot P(H_j) = \sum_{j \in J \setminus 4} E(X \mid H_j) P(H_j) = \sum_{j \in J \setminus 4} y_j P(H_j).$$

b) $Y \cdot I_H \overset{def}{=} \left(\sum_{i=1}^{4} y_i \cdot I_{H_i}\right) I_H = \sum_{i=1}^{4} y_i \cdot I_{H_i \cap H} \overset{(*)}{=} \sum_{j=1}^{4} y_j \cdot I_{H_j}$

$H_i \cap \bigcup_{j=1}^{4} H_j = \bigcup_j (H_i \cap H_j) = \bigcup_j H_j \Rightarrow I_{\bigcup H_j} = \sum_{j=1}^{4} I_{H_j} + 0$

$\int_H Y \, dP = \int_\Omega (Y \cdot I_H) \, dP \overset{(*)}{=} \sum_{j \in J} y_j P(H_j) = \sum_{j \in J \setminus \{4\}} y_j P(H_j)$

$\int_H X \, dP = \int_H Y \, dP$ ove $H = \bigcup_{j \in J} H_j$

(NB) se $J = \{4\}$ allora $H = H_4$ e $\int_H X \, dP = 0 = \int_H Y \, dP$

$\Rightarrow \int_H X \, dP = \int_H Y \, dP$ qualunque sia H osservabile, sia che H sia trascurabile sia non "

Posso quindi calcolare l'integrale della v.a. non osservabile X tramite l'integrale della v.a. osservabile Y

Se l'aggiornamento dell'informazione H è osservabile e se $P(H) > 0$ risulta $E(X|H) = \frac{\int_H X \, dP}{P(H)} = \frac{\int_H Y \, dP}{P(H)} = E(Y|H)$

la cui ipotesi in più!

La speranza matematica di X condizionata a gli evento non trascurabile posso calcolarla tramite v.a. Y

(NB) Nell'informazione disponibile se E è vero, $\bar{E}$ è falso $\Rightarrow$ vdv di 2 eventi!
Cioè un evento è osservabile anche la sua negazione è osservabile.
Se ho un numerabile di eventi osservabili $\Rightarrow$ l'unione osservabile è osservabile: se 1 evento $V \Rightarrow \bigcup$ è V, se tutti $F \Rightarrow \bigcup$ è F

* Conseguentemente la struttura degli eventi osservabili è una σ-algebra, quindi viene naturale identificare la informazione disponibile che si ottiene da un esperimento qualsiasi, come una σ-algebra di eventi

Morale:

Appare ragionevole ritenere che l'assunzione congiunta della $\mathcal{H}$-Borel misurabilità e della $\int_H X \, dP = \int_H Y \, dP$ per H osservabile (anche trasc.) possa essere intesa come caratterizzante la speranza matematica condizionata all'informazione ottenibile (rappresentata, nel caso particolare in esame, dalla σ-algebra generata $\mathcal{H}$).

• Formulazione generale (astrazione dall'esempio)

Adottando questo punto di vista e passando ad una formulazione generale considero, per descrivere formalmente l'informazione disponibile tramite un prefissato processo di osservazione, una σ-algebra $\mathcal{H} \subseteq \mathcal{A}$, i cui elementi chiameremo eventi osservabili, in quanto rappresentano gli eventi aleatori il cui valore di verità è acquistabile tramite il processo di osservazione

Assumiamo $\mathcal{H}$ sia σ-algebra: appare sensato che la famiglia degli eventi osservabili $\mathcal{H}$ sia σ-algebra in quanto:

dim Ω osservabile $\Rightarrow \Omega \in \mathcal{H}$
E osservabile $\Rightarrow \bar{E}$ osservabile $\Rightarrow E \in \mathcal{H} \Rightarrow \bar{E} \in \mathcal{H}$ (conosco entrambi vdv)
$E_1, E_n, \ldots$ osservabili $\Rightarrow \bigcup_{n \geq 1} E_n$ osservabile perché se $\exists E_i$ vero $\Rightarrow \bigcup_{n \geq 1} E_n$ vero
$\Rightarrow (E_n)_{n \geq 1} \in \mathcal{H} \Rightarrow \bigcup E_n \in \mathcal{H}$ · $\forall E_i$ falso $\Rightarrow \bigcup_{n \geq 1} E_n$ falso

* Data una var. aleatoria X con $\underline{E(X) \text{ finita}}$, chiamiamo versione (della speranza matematica condizionata di X a $\mathcal{H}$) (in simboli $E(X|\mathcal{H})$) ogni v.a. estesa $\boxed{Y}$ tale che:

$\left[X \text{ è } (\mathcal{A},\mathcal{B})\text{-misur} \right.$

- (a) sia $\mathcal{H}$-Borel misurabile
  (b) $\forall H \in \mathcal{H}$, risulta $\int_H X\,dP = \int_H Y\,dP$

s'intende come caratterizzante la speranza matematica condizionata alla informazione attuale

Il problema ora è arrivare ad $E(X|\mathcal{H})$ ← info che avrei / info disponibile

Ω casi elementari
$\mathcal{A}$ eventi di interesse
$\mathcal{H}$ insiemi di eventi osservabili o inform. disponibile
 - è σ-algebra
=> essendo una famiglia di osservabili posso risolvere l'incertezza su $\mathcal{H}$

$\mathcal{A}$ INFORMAZIONE MASSIMA $\mathcal{H}=\mathcal{A}$ => X è v.a. $\mathcal{B}$-mis (è misurabile) => X è una versione (!) della $E(X|\mathcal{H})$

$\mathcal{H}$ è sotto σ-algebra di $\mathcal{A}$

$\{\emptyset, \Omega\}$ INFORMAZ. NULLA (peggiore info disponibile) $\mathcal{H}=\{\emptyset,\Omega\}$ => Y v.a. costante, $\int_\Omega X\,dP = \int_\Omega Y\,dP = \int_\Omega k\,dP$ con $P(\Omega)=1 \Rightarrow k = E(X)$ $\Rightarrow E(X|\mathcal{H}) = E(X)$.

La (a) implica che la v.a. Y sia osservabile; cioè la $\mathcal{H}$-Borel misurabilità di Y => osservabilità di Y

dim/
fix $y \in \mathbb{R}$
$\{Y=y\} = \{Y \in \{y\}\} \in \mathcal{H}$ poiché $\{y\} \in \mathcal{B}$ e Y è $(\mathcal{H},\mathcal{B})$-misurabile
allora al variare di y, $\{Y \in \{y\}\}$ è P, quindi c'è un solo y_0 vero valore di Y | $\{Y = y_0\}$ è vero => ho trovato vero y_0.

Pertanto una volta risolta l'incertezza sugli osservabili, avrò individuato il "vero" valore di Y. Ciò non vale per X che è non osservabile

- Y è finita (P-q.c.), in quanto $E(Y) = E(X)$ => posso assumere che ogni versione sia sempre finita (P-q.c.)

dim/
$E(X) = \int_\Omega X\,dP$ finita per ipotesi -> $\int_H X\,dP$ finito => $\int_H Y\,dP$ finito, P-q.c. per le proprietà di $E()$

Se f è integrabile, allora f è finita (m-q.o.);

Quante sono le versioni Y di $E(X|\mathcal{H})$?
Teor. unicità P-q.c. della versione
 Y versione di $E(X|\mathcal{H})$
 Y' versione di $E(X|\mathcal{H})$
 $Y = Y'$ P-q.c. cioè la speranza matematica condizionata è definita a meno di insiemi di probabilità nulla

dim/
Posto $H \equiv \{Y \neq Y'\} = \{Y > Y'\} \cup \{Y < Y'\}$, allora $\{Y \neq Y'\} \in \mathcal{H}$ poiché Y, Y' sono $(\mathcal{H},\mathcal{B})$-misurabili. Sia per assurdo $P(H) > 0$. Supposto, ad esempio $P(Y' > Y) > 0$ allora applico b) due volte $\int_{\{Y'>Y\}} X\,dP = \int_{\{Y'>Y\}} Y\,dP = \int_{\{Y'>Y\}} Y'\,dP$ con $\{Y'>Y\} \in \mathcal{H}$. Da cui, essendo $E(X)$ finito:
↳ $0 = \int_{\{Y'>Y\}} Y'\,dP - \int_{\{Y'>Y\}} Y\,dP = \int_{\{Y'>Y\}} (Y'-Y)\,dP > 0$ ← propr. llb se $m(A)>0$

Quindi a rigore avrei una classe di funzioni Y' uguali fra loro a meno di eventi osservabili di probabilità nulla, cioè uguali P-q.c.
Per questo motivo, con un abuso di notazione, useremo il simbolo $E(X|\mathcal{H})$ per denotare anche una sua qualsiasi versione di $E(X|\mathcal{H})$ (in altre parole tralascio il "P-q.c.")

NB $E(X) \neq E(X|\mathcal{H})$
 numero   v.a. $(\mathcal{H},\mathcal{B})$-misurabile

Un evento osservabile H si chiama P-atomo (in breve atomo) se è ancora un evento di probabilità positiva non ripartibile in due eventi osservabili ancora di probabilità positiva

Ossia $\forall H_1 \subset H$, $P(H_1)=0$ oppure $P(H_1)=P(H)$

Teor.
Sia $(H_i)_{i\in I}$ una P divisibile di Ω costituita da eventi non trascurabili allora gli atomi della σ-algebra generata coincidono con gli elementi della partizione

** Teor.
Sia H un P-atomo di $\mathcal{H}$
$\Rightarrow \nexists\, H_1 \subset H : P(H_1) = P(H)$
$\Rightarrow E(X|\mathcal{H})(\omega) = E(X|H)\ \forall \omega \in H_1$

cioè non è vero che se prendo un H di $\mathcal{H}$ qls, la $E(X|\mathcal{H})(\omega) = E(X|H)$ con $\omega \in H$

Teor. Risultati

1) $E(E(X|\mathcal{H})) = E(X)$

dim $E(X) = \int_\Omega E(X|H)\,dP = E(E(X|H))$ con $\Omega = \mathcal{H}$

2) $E(X|\{\emptyset,\Omega\}) = E(X)$ $\quad E(X)$ condiz. alla info nulla produce informazione minima infatti conosco solo la media

dim l'unico evento di prob >0 è Ω, nella $\{\emptyset,\Omega\}$

3) $E(X|\mathcal{A}) = X$ (P-q.c.) $\quad$ l'informazione massima consiste nel sapere quale ω si verifica

dim, essendo X $(\mathcal{H},\mathcal{B})$-misurabile, se $\mathcal{H}=\mathcal{A} \Rightarrow$ conosco tutta X

4) $E(X|\mathcal{H}) = X$ (P-q.c.) se X è $(\mathcal{H},\mathcal{B})$-misurabile

dim X è $(\mathcal{H},\mathcal{B})$-misurabile ~~e~~ ~~osservabile~~, quindi la sua speranza matematica condizionata è lui.

Premessa a Levy

con $\mathcal{H}$ intermedia sfrutto Levy

** Teor Levy $\quad$ il risultato giustifica l'introduzione della speranza matematica condizionata alla σ-algebra

Sia $(\mathcal{H}_n)_{n\geq 1}$ succ. non decrescente di σ-algebre
$\mathcal{A} = \sigma(\bigcup_n \mathcal{H}_n)$ $\quad$ struttura informativa (filtrazione) $\quad$ succ. di stati di informazione, famiglie di osservabili sempre più grandi $\nearrow$
$\forall$ v.a. X, con $E(X)$ finito
$\Rightarrow E(X|\mathcal{H}_n) \to X$ (P-q.c.) $\quad$ la sper. mat. cond. ad ogni livello della inform. convergerà q.c. al mio osservabile X
ossia $P(\omega | E(X|\mathcal{H}_n)(\omega) \to X(\omega)) = 1$ conv q-certa

Oss La convergenza è solo puntuale purtroppo, non uniforme (!!G-Cantelli)

L'idea del teorema è che se le informazioni a disposizione sono non decrescenti ed esaustive, allora la corrispondente versione converge, al limite, alla v.a. non osservabile X

$\Rightarrow$ le sper. matematiche condizionate, ad ogni livello, sono approssimazioni di X, e sono via via meno rozze della info minima (che dà $E(X)$) ed arrivano alle info disponibile massima X.

NB Quando $E(X|\mathcal{H}) = X$ (P-q.c.)? Quando $\mathcal{H}=\mathcal{A}$ o X $(\mathcal{H},\mathcal{B})$-misurabile (visto sopra)

Teor X, Y, X_1, X_2
$E(X), E(Y), E(X_1), E(X_2)$
valgono le seguenti relazioni ma NB le $E(\cdot|\cdot)$ s'intendono vere P-q.c.

i) monotonia $E(X|\mathcal{H}) \leq E(Y|\mathcal{H})$ se $X \leq Y$ (P-q.c.)
ii) $E(X|\mathcal{H}) = E(Y|\mathcal{H})$ se $X = Y$ (P-q.c.)
iii) intermedità $a \leq E(X|\mathcal{H}) \leq b$ se $a \leq X \leq b$
iv) intermedità stretta $a < E(X|\mathcal{H}) < b$ se $a < X < b$ (P-q.c.)
v) linearità $E\left(\sum_{i=1}^{n} a_i X_i | \mathcal{H}\right) = \sum_{i=1}^{n} a_i E(X_i|\mathcal{H})$
vi) $|E(X|\mathcal{H})| \leq E(|X| \,|\mathcal{H})$
vii) convergenza monotona $E(X_n|\mathcal{H}) \uparrow E(X|\mathcal{H})$ se $X_n \uparrow X$ e $X_1 \geq 0$ (P-q.c.)
viii) $E\left(\sum_{n=1}^{\infty} a_n X_n | \mathcal{H}\right) = \sum a_n E(X_n|\mathcal{H})$ se $(a_n)_n \in \mathbb{R}^+$, $X_n \geq 0$ $\forall n$ e se $E(\sum_n a_n X_n|\mathcal{H})$ finita
ix) convergenza dominata $E(X_n|\mathcal{H}) \to E(X|\mathcal{H})$ se $|X_n| \leq Z$ (P-q.c.) $\forall n$, con $E(Z)$ finita e $X_n \to X$
x) $E(Z \cdot H|\mathcal{H}) = Z \cdot E(H|\mathcal{H})$ se Z è $(\mathcal{H}, \mathcal{B})$-misurabile e limitata
idea: Z è osservabile e limitata $\Rightarrow$ la tratto come un certo k $E(X) = E(X \cdot k)$
posto fuori ciò che è noto e misurabile
xi) Se K è una sotto-σ-algebra di $\mathcal{H}$, allora $E[E(X|\mathcal{H})|K] = E(X|K)$ e
$E[E(X|K)|\mathcal{H}] = E(X|K)$
se ho 2 informazioni differenti conta sempre la più piccola

Il seguente teorema dice come varia la $Var(Y)$ risp alla $Var(X)$ dopo l'arrivo di nuova informazione

Teor
X (non osservabile), $E(X)$, $Var(X)$ finite (per conto)
Y versione limitata di $E(X|\mathcal{H})$
$Var(X) = Var(Y) + E[(X-Y)^2] \geq Var(Y)$

(aspettazione v.a. non osservabile) (SSM fra ... e ...)

L'arrivo di informazione migliora la conoscenza di X

all'↑ dell'informazione la varianza della versione limitata cala

dim
$E(X^2)$, $E(Y)^2$ finite (=> sono 2 integrabili) in quanto $V(X) = E(X^2) - E(X)^2$ finita e $V(Y)$ finita perché Y limitata
allora posto $E(X) = \alpha \Rightarrow (X-Y)(Y-\alpha)$ v.a. è integrabile perché il prodotto è integrabile se $(X-Y)^2$ e $(Y-\alpha)^2$ è integrabile: X^2 e Y^2 integr. x ipotesi $\Rightarrow 2XY$ integrabile
e $\Rightarrow 2Y$ è integr. ($Y \cdot 1 \leq Y^2 1^2$ integrabili)

$$\Rightarrow Var(X) = E[(X-\alpha)^2] = E\left(((X-Y)+(Y-\alpha))^2\right) =$$
$$= E[(X-Y)^2] + E[(Y-\alpha)^2] + 2E[(X-Y)(Y-\alpha)]$$
$$= E[(X-Y)^2] + Var(Y) + 2\underbrace{E[(X-Y)(Y-\alpha)]}_{\text{se fosse } = 0 \text{ avrei finito}}$$

TH: $E[(X-Y)(Y-\alpha)] = 0$
da $(Y-\alpha)$ v.a. limitata e osservabile ($\mathcal{H}$-$\mathcal{B}$)-misurabile
allora
$$E[(X-Y)(Y-\alpha)|\mathcal{H}] \overset{x)}{=} (Y-\alpha)E[X-Y|\mathcal{H}] \quad (P\text{-q.c.})$$
$$= (Y-\alpha)\left[\underbrace{E(X|\mathcal{H})}_{\text{la posso sostituire con una sua versione}} - \underbrace{E[Y|\mathcal{H}]}_{\text{v.a. } Y \text{ } (\mathcal{H}\text{-}\mathcal{B})\text{-mis} \Rightarrow Y}\right] \quad (P\text{-q.c.})$$
$$= (Y-\alpha)(Y-Y) \quad (P\text{-q.c.})$$
$$= 0 \quad (P\text{-q.c.})$$

~~pongo $(X-Y)(Y-\alpha) \triangleq M$ allora finisco se dico $E(M|\mathcal{H}) = 0$ (P-q.c.)~~
~~$\Rightarrow$ da "$E[E(M|\mathcal{H})] = E(X)$" so che $E[E(M|\mathcal{H})] = E(M) = 0$, $\Rightarrow E(M) = 0$ allora~~
~~$E[(X-Y)(Y-\alpha)] = 0$~~ /

Oss Y versione limitata di $E(X|\mathcal{H}) \Rightarrow$ ogni versione Y' sarà ancora limitata (P-q.c.)

Funzione di regressione

X $(\mathcal{A},\mathcal{N})$-mis, Y non osserv.

In moltissime situazioni l'informazione ottenibile consiste nell'osservare il valore di un ente aleatorio $X: \frac{\Omega}{\mathcal{A}} \to \frac{\mathcal{X}}{\mathcal{N}}$, dove $\mathcal{N}$ contiene i singoletti. Perciò l'informazione dovuta dall'osservazione di X viene naturale identificarla con la σ-algebra di eventi osservabili $\mathcal{H}$, con la σ-algebra indotta $X^{-1}(\mathcal{N}) = \{X^{-1}(A): A \in \mathcal{N}\} \subseteq \mathcal{A}$ da X su Ω. È informazione disponibile.

Dà un fido interpretativo, la conoscenza di valori di verità di tutti gli eventi di tale σ-algebra equivale alla conoscenza del vero valore di X in quanto $\{X = y\} = \{X \in \{y\}\} \in X^{-1}(\mathcal{N})$ $\forall y \in \mathcal{X}$

$$\frac{\Omega}{\mathcal{A}}_{P} \xrightarrow{X \text{ e.a. osservabile}} \frac{\mathcal{X}}{\mathcal{N}} \qquad X\ (\mathcal{A},\mathcal{N})\text{-misurab} \qquad \mathcal{H} = X^{-1}(\mathcal{N}) = \{\{X \in B\}, B \in \mathcal{N}\}\ \sigma\text{-alg indotta da } X$$

$\mathcal{H} = X^{-1}(\mathcal{N})$ informazione disponibile
$= \{\{X \in B\}: B \in \mathcal{N}\}$ σ-algebra indotta da X

una volta osservato X conosciamo i valori degli eventi di $\mathcal{H}$

Sia Y v.a. non osservabile (∞), con $E(Y)$ finita:

$$E(Y|\mathcal{H}) = \underbrace{E(Y|X^{-1}(\mathcal{N}))}_{\substack{\text{vero significato} \\ \sigma\text{-algebra indotta}}} = \underbrace{E(Y|X)}_{\text{tecnografia}}$$

- è la pr. nat. di Y a fronte del fatto che posso conoscere il vero valore di X, con l'info = controimmagine
- è pr. sistem. condizionata di Y alla osservazione di X

Caso Y v.a.
X ente aleat. ($E(Y)$ finita)
X e Y indipendenti
$E(Y|X) = E(Y) = E(Y|\{\emptyset,\Omega\})$ — gia' visto

- l'informazione proveniente da X equivale alle informazioni nulla

- trasformate certe di e.a. indipendenti sono ancora indipendenti
- INDIPENDENTI: incontrata nel. legge congiunta: 2 e.a. indipendenti $\Leftrightarrow$ legge congiunta = prodotto marginali. Il motivo per cui è stato usato la parola indipendenze adesso è giustificato, mentre prima non aveva significato; dico che l'informazione disponibile tramite X equivale alla info nulla. Osservare X equivale all'osservare nulla.

dim. (non serve perché poi c'è la v.a. misurabile)

$A \in \mathcal{N}$, Y v.a.

$I_{\{X \in A\}} = I_A(X)$

Y e $I_{\{X\in A\}}$ sono indipendenti

$$I_A(X) = \begin{cases} 1 & X(\omega) \in A \Leftrightarrow \omega \in X^{-1}(A) \\ 0 & X(\omega) \notin A \end{cases} = \begin{cases} 1 & \{X \in A\} \\ 0 & \{X \notin A\} \end{cases} = I_{\{X \in A\}}$$

$$\Rightarrow \int_{\{X\in A\}} Y\, dP \overset{\text{def}}{=} E(Y \cdot I_{\{X \in A\}}) \overset{\text{v.a. indip}}{=} E(Y)P(X \in A) = \int_{\{X\in A\}} E(Y)\, dP \Rightarrow E(Y) = E(Y|X)$$

trasformata certa di X — trasformate di e.a. indip. sono ancora e.a. indip.

- $E(Y)$ è $\mathcal{H}$-$\mathcal{B}$-misur. $\Rightarrow$ regressione
- è v.a. (a), b)

Il seguente Lemma stabilisce che ogni funzione $(\mathcal{H},\mathcal{B})$-misurabile è trasformata certa dell'osservabile X

$$\begin{array}{ccc} \dfrac{\Omega}{\mathcal{A}} & \xrightarrow{X} & \dfrac{\mathcal{X}}{\mathcal{N}} \\ & \searrow^{Z} & \downarrow g \text{ v.a. } (\mathcal{N},\mathcal{B})\text{-misurab} \\ & & \dfrac{\mathbb{R}\ (\mathbb{R}^*)}{\mathcal{B}\ (\mathcal{B}^*)} \end{array}$$

$X^{-1}(\mathcal{N}) = \mathcal{H}$, Z $(\mathcal{H},\mathcal{B})$-mis

Come posso scrivere la v.a. Z? Come trasformata certa di X; l'idea è che risolta l'incertezza su X, l'ho risolta anche su $\mathcal{H}$, e allora l'ho risolta anche su Z. Il valore di X determina valore di Z

Lemma Z v.a. $(X^{-1}(\mathcal{N}), \mathcal{B})$-misurabile
$\exists\, g : \underset{\mathcal{N}}{\mathcal{X}} \to \underset{\mathcal{B}^*}{\mathbb{R}^*}$ $(\mathcal{N}, \mathcal{B}^*)$-misurabile t.c. $Z = g(X)$

- stabilisco che in presenza di una funzione $(\mathcal{A}, \mathcal{B}^*)$-misurabile, posso sempre trovare una trasformata certa dell'osservabile X che mi determina lei.
- siccome $E(Y|X)$ per def è $(\mathcal{A}, \mathcal{B})$-misurabile allora sostituisco a Z $E(Y|X)$ nel lemma, ottengo che $E(Y|X)$ è trasformata certa dell'osservabile X (perché dice che $\exists\, g$ t.c. $E(Y|X) = g(X)$) (cioè $E(Y|X^{-1}(\mathcal{N})) = g(X)$)

dim/ pongo $\underset{\mathcal{A}}{\mathcal{A}_1} = X^{-1}(\mathcal{N})$ TH: (solo caso Z v.a. semplice)

$$Z = \sum_{i=1} I_{H_i},\ H_i \in \mathcal{A}_1 \text{ e } \alpha_i \in \mathbb{R} \quad (i=1,\dots,n), \text{ allora } \exists\, h_1 \dots h_n \in \mathcal{N} \mid H_i = \{X \in h_i\}_{i=1,\dots,n}$$

$$\Rightarrow g = \sum_{i=1}^{n} \alpha_i \cdot I_{h_i} \quad (\mathcal{N}, \mathcal{B})\text{-misurabile}$$

$$g(X(\omega)) = \sum_{i=1}^{n} \alpha_i\, I_{h_i}(X(\omega)) = \sum_{i=1}^{n} \alpha_i \cdot I_{\{X \in h_i\}}(\omega) = \left(\sum_{i=1}^{n} \alpha_i \cdot I_{H_i}\right)(\omega) = Z(\omega)$$

sp. med. di Y cond. X $(\mathcal{A}_1, \mathcal{B})$-misurabile

Scelta una versione Z di $E(Y|X)$ $\exists$ per il lemma, una funzione g $(\mathcal{N}, \mathcal{B})$-misurabile : $Z = g(X)$ da cui

$$\int_{\{X \in A\}} Y\, dP \underset{\text{def sp. med. cond.}}{=} \int_{\{X \in A\}} Z\, dP = \int_{\{X \in A\}} g(X)\, dP \underset{\text{th. fond. calc. prob.}}{=} \int_A g\, P_X \quad \forall A \in \mathcal{N}$$

la conoscenza della legge e della g consente di ottenere ogni integrale di quel tipo con Y non osservabile, rispetto all'osservabile.

Ancora una volta per calcolare gli integrali relativi ad un non osservabile Y rispetto ad ogni osservabile, basta dare la legge dell'e.a. X e la funzione $g(\cdot)$. Ancora perché come per la legge dell'ente aleatorio possiamo sganciarci dal meccanismo di sorteggio
Come per le v.a. si era data solo la distrib. di prob. (2° lemma stesso 5) così ora sp. med. condiz. alle info date da e.a. osservabile, basta dare la funzione g x costruire l'integrale

$$\Omega \xrightarrow{X} \mathcal{X} \xrightarrow{g} \mathbb{R}\ (\mathbb{R}^*),\quad \mathcal{A}_1 \to \mathcal{N} \to \mathcal{B}\ (\mathcal{B}^*)$$

La g è la funzione di **regressione** come fra poco vedremo. Quindi tramite la regressione possiamo sganciarci dal meccanismo di sorteggio per valutare la speranza matematica condizionata alla informazione disponibile, data da un osservabile.

Chiamiamo **funzione di regressione** di Y su X ogni funzione $E(Y|X=\cdot) : \mathcal{X} \to \mathbb{R}^*$ "leggi: E di Y sapendo $=\cdot$"
tale che:
1) sia P_X-integrabile
2) $\forall A \in \mathcal{N}$ risulta: $\int_{\{X \in A\}} Y\, dP = \int_A E(Y|X=x)\, P_X(dx)$

tramite x e con fun. di regr. trovo valori della $E(Y|X=\cdot)$

La $E(Y|X=\cdot) : \mathcal{X} \to \mathbb{R}$ è finita (P_X-q.c.) perché l'integrale è finito
è definita a meno di insiemi P_X-trascurabili per def. dell'integrale: $Y = g(X)$, $Y' = g(X)$ sono P_X-q.c. uguali

$\underbrace{E(Y|X = X(\cdot))}$ è una versione di $E(Y|X)$

da completare

Segue che: $g(X) = E(Y|X)$, $g(X(\omega)) = E(Y|X)(\omega)$, $g(x) = E(Y|X=x)$ e
e $g(X(\dot\omega)) = E(Y|X = X(\dot\omega))$

Notazione ambigua
$g(\dot x) = E(Y|X = \dot x)$ — funz. regressione $\exists$ condizionamento! stenografia: "calcolo la f. di regr. in $\dot x$" $\Rightarrow$ def. sugli x
$\not\Leftarrow$ $E(Y|\{X = x\})$ — speranza matem. condiz. $\exists$ condiz. ad un evento: "calcolo la sp. mat. condizionata" $\Rightarrow$ def. sugli ω

f di regr → dipende da val osservabili

sp mat cond informaz dipende → dipende da casi trascurabili
↳ ha senso se $P(\{X=x\})=0$

Se $P(\{X=x\})>0$, cioè l'evento $\{X=x\}$ è non trascurabile, allora le 2 coincidono. Infatti

Teor. Sia $x_0 \in \mathcal{X} \mid P(\{X=x_0\})>0$, allora $E(Y|X=x_0)=\underbrace{\dfrac{\int_{\{X=x_0\}} Y\,dP}{P(X=x_0)}}_{\text{def}} = \underbrace{E(Y|\{X=x_0\})}_{\text{sp mat}}$

Ho omesso degli esempi, i quali servivano a far vedere come trovare la speranza mat cond alla info disponibile non sia così facile:
- già trovare la σ-algebra indotta è piuttosto difficile perché dipende da chi è X e che andamento ha
- poi c'è da trovare la f di regressione

Teor
Siano Y, Y_1, Y_2 etc v.a. con speranza matematica finita. Risulta allora:
$E(Y) = \int E(Y|X=x)\, P_X(dx)$
Valgono inoltre le proprietà seguenti nelle quali le relazioni riguardanti funzioni di regressione devono intendersi sussistere $P_X^{\mathcal{Q}}$-q.c.

- Se $X=x_0 \in \mathcal{X}$ P-q.c., allora $E(Y|X=\cdot)=E(Y)$. Inoltre se X è l'applicazione identica, allora $E(Y|X=\cdot)=Y$
- $E(a I_\Omega|X=\cdot)=a$
- MONOTONIA $E(Y_1|X=\cdot)\le E(Y_2|X=\cdot)$ se $Y_1 \le Y_2$ (P-q.c.)
- INTERVALLITÀ $a\le E(Y|X=\cdot)\le b$ se $a\le Y\le b$ (P-q.c.)
- INTERVALLITÀ STRETTA $a< E(Y|X=\cdot)<b$ se $a<Y<b$
- LINEARITÀ $E(\sum a_i Y_i|X=\cdot)=\sum a_i E(Y_i|X=\cdot)$ qualunque siano $a_1,\dots,a_n$ reali
- $E(\sum_{n\ge1} a_n Y_n|X=\cdot)=\sum_{n\ge1} a_n E(Y_n|X=\cdot)$ se $(a_n)_{n\ge1}\in\mathbb{R}^+$, $Y_n\ge0$, $E(\sum a_n Y_n)$ finita
- CONVERGENZA DOMINATA $E(Y_n|X=\cdot)\to E(Y|X=\cdot)$, se $|Y_n|\le Z$ (P-q.c.) $\forall n$, $E(Z)$ finita e $Y_n\to Y$
- $g: \mathcal{X}\to\mathbb{R}^+$ (N,B)-misurabile, $g(X)\cdot Y$ ammette $E(g(X)Y)$ finita allora
$E(g(X)\cdot Y|X=\cdot)=g(\cdot)\,E(Y|X=\cdot)$
In particolare $E(g(X)|X=\cdot)=g(\cdot)$.
- $|E(Y|X=\cdot)|\le E(|Y||X=\cdot)$;
- CONVERGENZA MONOTONA $E(Y_n|X=\cdot)\uparrow E(Y|X=\cdot)$ se $Y_n\uparrow Y$ e $Y_1\ge0$ (P-q.c.);

IMPORTANTI

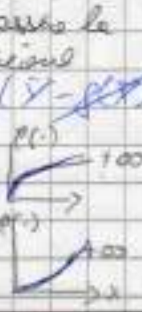
Retta di regressione lineare

Concludiamo mostrando l'importanza fondamentale della funzione di regressione nella soluzione del problema (centrale nella problematica statistica e di grande interesse applicativo) di stimare, a partire dall'osservabile X, il non osservabile Y commettendo un errore più piccolo possibile.

Consideriamo una stima $g(X)$ di Y
$\varphi(\cdot)$ funzione peso degli errori è t.c. $\varphi(0)=0$ crescente
↳ misura l'errore di sostituzione di Y continua $\lim_{x\to\infty}\varphi(x)=+\infty$
con una stima, ossia la bontà dell'approssimazione

$Y-g(X)$ $\quad |Y-g(X)|$ $\quad (Y-g(X))^2$
errore assoluto

$\varphi(|Y-g(X)|)$ può <



penalizza errori piccoli

" " grandi

$\Rightarrow$ allora $E[\varphi(|Y - g(X)|)]$ errore medio assoluto

Viene naturale pensare che l'errore medio assoluto assuma valori grandi se $g(X)$ assume valori distanti da Y (e piccoli se vicini)

Ordinano, tra le varie scelte di φ, quello relativo al metodo dei minimi quadrati ponendo $\varphi(x) = x^2$. Così l'errore più piccolo possibile sarà l'errore quadratico medio più piccolo possibile e quindi a cercare una funzione g^* t.c. $E((Y - g^*(X))^2) \leq E((Y - g(X))^2)$ $\forall g$ a quadrato P_X-integrabile

Il prossimo risultato collega la funzione di regressione con il MMQ assicurando che ogni funzione di regressione di Y su X fornisce un esempio di stima dei minimi quadrati di Y.

Teo: Se Y e $g(X)$ sono a quadrato integrabili allora $E(Y | X = X(\cdot))$ è a^2 integrabile e risulta:

$$E[(Y - E(Y|X))^2] = E[(Y - E(Y | X = X(\cdot)))^2] =$$

$$= \min_g \{ E[(Y - g(X))^2] : \int g^2 \, dP_X < +\infty \}$$

La speranza nel condiz. è la migliore approssimazione in media quadratica di non osservabile.

$\Rightarrow E(Y|X)$ è funzione che minimizza il danno quadratico medio.

La f di regressione che risolve il metodo dei MQ $\Rightarrow g^*(X) = E(Y | X = \cdot)$ è la funzione minima su tutte le funzioni a^2 integrabili.

Il teorema assicura che è possibile approssimare il non osservabile attraverso l'errore quadratico medio (il $|$ in media è più piccolo (come quadratico)) e che il risultato è la f di regressione.

Ma noi ricordando che la funzione di regressione è $E(Y|X)$, spe. nel cond. a X (osservabile) diventa la migliore approssimazione.

- Calcolare la f di regressione è difficile → sviluppo in serie di Taylor e semplifico troncando al $1°$ ordine: $g(X) = aX + b$ $(a>0)$
  $\min_g E[(Y - g(X))^2]$

Ponendo in particolare $g(x) = ax + b$ $(a>0)$ e X, Y v.a. a quadrato integrabili e $Var(X) > 0$ finita:

$$f(a,b) = E((Y - g(X))^2) = E[(Y - (aX + b))^2] =$$

$$= E(Y^2) + E(a^2X^2 + b^2 + 2abX) - 2E(Y(aX + b))$$

$$= E(Y^2) + [E(a^2X^2) + b^2 + 2abE(X)] - 2aE(XY) - 2bE(Y)$$

$$= E(Y^2) + a^2E(X^2) + b^2 + 2abE(X) - 2aE(X \cdot Y) - 2bE(Y)$$

da cui ottengo (a^*, b^*) pt minimo del sistema risolvendo

$$\begin{cases} \dfrac{\partial}{\partial a} f(a,b) = 2aE(X^2) + 2bE(X) - 2E(XY) \overset{!}{=} 0 \\ \dfrac{\partial}{\partial b} f(a,b) = 2b + 2aE(X) - 2E(Y) \overset{!}{=} 0 \end{cases}$$

$$\begin{cases} aE(X^2) + b(E(X)) = E(XY) \\ b + aE(X) = E(Y) \end{cases}$$

$b = E(Y) - aE(X)$

$a(E(X^2)) + (E(Y) - aE(X))E(X) = E(X \cdot Y)$

$$a\,E(\bar{X}^2) + E(X)E(\bar{Y}) - a\,E(\bar{X})^2 = E(\bar{X}\bar{Y})$$

$$a^* = \frac{Cov(\bar{X},\bar{Y})}{Var(\bar{X})}$$

$$b^* = E(\bar{Y}) - \frac{Cov(\bar{X},\bar{Y})}{Var(\bar{X})} \cdot E(\bar{X})$$

⇒ Abbiamo così l'espressione della trasformata affine dell'osservabile $\bar{X}$ che meglio approx, in media quadratica, il non osservabile $\bar{Y}$

$$y = \frac{Cov(\bar{X},\bar{Y})}{Var(\bar{X})}\,(x - E(\bar{X})) + E(\bar{Y})$$

detta retta di regressione lineare di $\bar{Y}$ su $\bar{X}$

Nelle applicazioni statistiche basta ciò per avere una lineare approx del non osservabile

Anche se ricorrendo a Taylor si troverebbero approx del 2° ordine migliori di questa

Invece qui cerchiamo il miglior approssimante lineare non tra le $f(x)$ a² integrabili, ma solo tra le lineari, perché vogliamo evitare problemi di calcolo ed determinare σ-algebre indotte

Γ La miglior approssimazione in media quadratica del non osservabile è la speranza matematica condizionata all'osservabile, del non osserv.

se g^* è Fd poss ho risolto il problema, ma come vedo fatto, se posto a x X ho la sper mat condiz (x def)

→ g^* risolve problema dei MQ

I DIMENSIONE

v.a. $X : \underset{\mathcal{A}}{\Omega} \to \underset{\mathcal{N}}{\mathcal{X}}$ $(\mathcal{A},\mathcal{N})$-misurabile

Legge $P_X : \mathcal{N} \to [0,1]$
$P_X(A) = P\{X\in A\}$ $\forall A\in\mathcal{N}$

Densità: sia μ misura $: \mathcal{N} \to [0,+\infty]$ $(\mathcal{N},\mathcal{B})$-misurabile
$\Rightarrow$ chiamo μ-densità ogni $f : \mathcal{X} \to [0,+\infty[$ $(\mathcal{N},\mathcal{B})$-misurabile
t.c. $P_X(A) = \int_A f\,d\mu$ $\forall A\in\mathcal{N}$

Teor: X v.a.
P_X legge
$g : \mathcal{X} \to [0,+\infty]$ misurabile risp P_X

$\int_A g\,dP_X = \int_A g f\,d\mu$ $\forall A\in\mathcal{N}$

Teor fondam. c.d.v.
X v.a.
P_X legge
$g : \mathcal{X} \to [0,+\infty]$ misurabile risp P_X

$$\int_{\{X\in A\}} g(X)\,dP = \int_A g\,dP_X = \int_A g f\,d\mu$$ con μ-densità

$$\begin{tikzcd} \underset{\mathcal{A}}{\Omega} \arrow[r, "X"] & \underset{\mathcal{N}}{\mathcal{X}} \arrow[r, "f"] \arrow[dr, "P_X"] \arrow[ddr, "\mu"'] & {[0,+\infty[} \\ & & {[0,1]} \\ & & {[0,+\infty]} \end{tikzcd}$$

II DIMENSIONI

v.a. $(X_1,X_2) = \underline{X} : \underset{\mathcal{A}}{\Omega} \to \underline{\mathcal{X}} = \mathcal{X}_1\times\mathcal{X}_2$, $\underline{\mathcal{N}} = (\mathcal{N}_1\otimes\mathcal{N}_2)$ $(\mathcal{A},\underline{\mathcal{N}})$-misurabile

Legge congiunta $P_{\underline{X}} : \underline{\mathcal{N}} \to [0,1]$
$P_{\underline{X}} = P_{X_1}\cdot P_{X_2}$

Densità congiunta: sia $\mu_1\times\mu_2$ misura $: \underline{\mathcal{N}} \to [0,+\infty]$ $(\underline{\mathcal{N}},\mathcal{B})$-misurabile
$\Rightarrow$ chiamo μ_1,μ_2 densità congiunta ogni
$f : \underline{\mathcal{X}} \to [0,+\infty[$ $(\underline{\mathcal{N}},\mathcal{B})$-misurabile
t.c. $P_{\underline{X}}(A) = \iint_A f\,d\mu_1\times\mu_2$ $\forall A\in\underline{\mathcal{N}}$

Teor: $\underline{X}$ v.a.
$P_{\underline{X}}$ legge congiunta
$g : \underline{\mathcal{X}} \to [0,+\infty]$ misurabile risp $P_{\underline{X}}$

$\iint_A g\,dP_{\underline{X}} = \iint_A g f\,d\mu_1\times\mu_2$ $\forall A\in\underline{\mathcal{N}}$

Teor fondam. c.d.v.
$\underline{X}$ v.a.
$P_{\underline{X}}$ legge congiunta
$g : \underline{\mathcal{X}} \to [0,+\infty]$ misurabile risp $P_{\underline{X}}$

$$\iint_{\{\underline{X}\in A\}} g(\underline{X}) = \iint_A g\,dP_{\underline{X}} = \iint_A g f\,d\mu_1\times\mu_2$$ con $\mu_1\times\mu_2$ densità congiunta

$$\begin{tikzcd} \underset{\mathcal{A}}{\Omega} \arrow[r, "\underline{X}"] & \underset{\underline{\mathcal{N}}}{\underline{\mathcal{X}}} \arrow[r, "f"] \arrow[dr, "P_{\underline{X}}"] \arrow[ddr, "\mu_1\times\mu_2"'] & {[0,+\infty[} \\ & & {[0,1]} \\ & & {[0,+\infty]} \end{tikzcd}$$

SPERANZA MATEMATICA CONDIZIONATA

$\mathcal{H} \subseteq \mathcal{A}$ informazione disponibile, σ-algebra
sia $X: \Omega \to \mathbb{R}$, $E(X)$ finita — X v.a. NON OSSERVABILE

$\Rightarrow$ SPERANZA MATEMATICA CONDIZIONATA di X a $\mathcal{H}$, [illegible] $E(X \mid \mathcal{H})$:
a) $(\mathcal{H}, \mathcal{B})$-misurabile $\Rightarrow$ osservabile [illegible]
b) $\forall H$ osservabile $\in \mathcal{H}$: $\int_H X \, dP = \int_H E(X \mid \mathcal{H}) \, dP$

$\Rightarrow E(X \mid \mathcal{H})$ è finita P-q.c.
[illegible]
"Y è una versione di $E(X \mid \mathcal{H})$" [illegible]

$\Rightarrow$ Y è [illegible] osservabile
$\Rightarrow E(Y)$ finita P-q.c. [illegible]

- $E(E(X \mid \mathcal{H})) = E(X)$
- $E(X \mid \mathcal{A}) = X$ se [illegible]
- $E(X \mid \{\emptyset, \Omega\}) = E(X)$, se [illegible]
- $E(X \mid \mathcal{H}) = X$, se X è $(\mathcal{H}, \mathcal{B})$-misurabile

- [illegible] $E(X \mid \mathcal{H}) \neq X$ (P-q.c.) [illegible]
- $X, V(X)$ finita [illegible]
$Var(X) \geq Var(Y) + E[(X-Y)^2] \geq Var(Y)$
- $(H_i)_{i \in I}$ [illegible] $P(H_i) > 0 \ \forall i \in I$
[illegible]
- [illegible] $E(X \mid H)$ [illegible]

Y è versione di $E(X \mid \mathcal{H})$ con X [illegible]

FUNZIONE DI REGRESSIONE

$\mathcal{H} \subseteq \mathcal{A}$ informazione disponibile, σ-algebra
sia $X: \Omega \to \mathcal{X}$, [illegible] $\mathcal{H} = X^{-1}(\mathcal{N})$ con [illegible]
X v.a. OSSERVABILE
$(X^{-1}(\mathcal{N}), \mathcal{N})$-misurabile
sia $Y: \Omega \to \mathbb{R}$ non osservabile, $E(Y)$ finita

$\Rightarrow E(Y \mid \mathcal{H}) = E(Y \mid X^{-1}(\mathcal{N})) = E(Y \mid X)$

Teorema [Y v.a. con $E(Y)$ finita]
Z versione di $E(Y \mid X)$ $\Rightarrow$ a) Z è $(\mathcal{H}, \mathcal{B})$-misurabile b) $\int_H Z \, dP = \int_H E(Y \mid X) \, dP \ \forall H \in \mathcal{H}$
$\exists g: \mathcal{X} \to \mathbb{R}$ t.c. $Z = g(X)$
$(\mathcal{N}, \mathcal{B})$-misur.
[illegible]

[illegible]

$\Rightarrow E(Y)$ finita [illegible]

FUNZIONE DI REGRESSIONE di Y su X
[illegible] $E(Y \mid X = \cdot)$ [illegible]

$\Rightarrow E(Y \mid X = \cdot)$ finita P-q.c.
[illegible]

$\Rightarrow E(Y \mid X = x(\cdot))$ è una versione di $E(Y \mid X)$
[illegible]

[illegible]

[illegible]

Z versione di $E(Y \mid X)$ [illegible]
$E(Y \mid X = x(\cdot))$ è versione di $E(Y \mid X)$