Appunti sintetici di calcolo delle probabilità (corso progredito)

June 2, 2018

[1]
[2]

[1] Appunti interamente tratti dal libro "Elementi di statistica delle decisioni" del prof. emerito Silvano Holzer disponibile su `https://www.openstarts.units.it/bitstream/10077/2743/1/testoHolzer.pdf`

[2] L'indice analitico potrebbe risultare eccessivamente specifico, ma ciò è stato fatto allo scopo di avere sottomano l'ossatura del corso, così da agevolare il ripasso finale. Alcuni nomi di teoremi sono stati inventati

Contents

1 Integrale di Lebesgue

Fissiamo un insieme ambiente e chiamiamo **σ-algebra** (su Ω) ogni famiglia $\mathcal{A}$ di sottoinsiemi di Ω tale che
$-\Omega \in \mathcal{A}$
$-A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$
$-A_n \in \mathcal{A}$ per tutti gli $n \geq 1 \Rightarrow \bigcup_{n\geq 1} A_n \in \mathcal{A}$

cioè tale che contenga l'insieme ambiente e sia chiusa per complementazione e unioni discrete. Ricorrendo alle formule di De Morgan, otteniamo la definizione equivalente (avendo denotato, come sempre faremo, con A_i elementi generici di $\mathcal{A}$):

$-A_1 \backslash A_2 \in \mathcal{A}$;
$-\emptyset, \Omega \in \mathcal{A}$;
$-\bigcup_{i\in I} A_i, \bigcap_{i\in I} A_i \in \mathcal{A}$, se I è un insieme discreto.

1.1 Sigma-algebra generata

Def 1.1. σ-algebra generata dalla famiglia $\mathcal{G} \subseteq 2^{\Omega}$.È la σ-algebra su Ω:
$\sigma(\mathcal{G}) = \cap\{\ \mathcal{A}$: $\mathcal{G} \subset \mathcal{A}$ e $\mathcal{A}$ σ-algebra su $\Omega\}$

In parole povere è la più piccola sigma algebra che contiene $\mathcal{G}$
Ovviamente, se $\mathcal{G} \subseteq \mathcal{G}', \sigma(\mathcal{G}) \subseteq \sigma(\mathcal{G}')$ e $\sigma(\mathcal{G}) = \mathcal{G}$, se $\mathcal{G}$ è una σ-algebra.

1.2 Traccia della sigma-algebra su S

Def 1.2. Traccia della σ-algebra $\mathcal{A}$ su $S \subset \Omega$ è la σ-algebra su S formata da da tutti gli insiemi che stanno in sia in A che in S:

$$\mathcal{A} \cap S = \{A \cap S : A \in \mathcal{A}\}$$

1.3 Sigma-algebra indotta da una applicazione

Def 1.3. σ-algebra indotta da una applicazione $\tau : \Omega_0 \neq \emptyset \to \Omega$ È la σ-algebra su Ω_0:

$$\tau^{-1}(\mathcal{A}) = \{\tau^{-1}(A) : A \in \mathcal{A}\}$$

formata dalle controimmagini degli elementi della σ-algebra $\mathcal{A}$.

1.4 Retta reale ampliata

Def 1.4. $\mathbb{R}^* = \mathbb{R} \cup \{-\infty,\ +\infty\}$ ottenuta aggiungendo all'insieme $\mathbb{R}$ dei numeri reali (retta reale) i simboli $-\infty$ e $+\infty$.

• Data questa aggiunta, dobbiamo escludere le espressioni aritmetiche che coinvolgono rapporti di infiniti o rapporti con denominatori nulli, oppure somme di infiniti di segno opposto o differenze di infiniti di ugual segno, poiché sono forme indeterminate.

1.5 sigma-algebra di Borel

- La σ-algebra di Borel (di $\mathbb{R}$) è la σ-algebra $\mathcal{B}$ su $\mathbb{R}$ generata dalla famiglia degli intervalli limitati inferiormente aperti e superiormente chiusi, cioè

  $$\mathcal{B} = \sigma(\{]a,\ b]\ :\ a < b,\ a,\ b \in \mathbb{R}\})$$

  La sua introduzione, fatta nel 1898 da Emile Borel, inaugurò una nuova era dell'analisi matematica fornendo il punto di partenza sia di una classificazione topologica degli insiemi di punti che della formulazione astratta della nozione d'integrale.

  $\mathcal{B}$ è anche la σ-algebra generata da una qualsiasi famiglia d'intervalli (limitati o no) del medesimo tipo.

  $\mathcal{B}$ è anche generata dalla famiglia $\mathcal{U}$ degli insiemi aperti di $\mathbb{R}$ e quindi anche dalla famiglia $\{W^c : W \in \mathcal{U}\}$ degli insiemi chiusi.

- La σ-algebra di Borel (di $\mathbb{R}^*$) è la σ-algebra B^* su $\mathbb{R}^*$ generata dalla famiglia degli intervalli di $\mathbb{R}^*$ inferiormente aperti e superiormente chiusi.

  Pertanto, B^* è pure generata dalla famiglia degli insiemi aperti della retta reale ampliata.

  B^* è anche generata dalla famiglia $\{[-\infty,\ a]\ :\ a \in \mathbb{R}\}$ delle semirette inferiori della retta reale ampliata di origine un numero reale.

Infine sappiamo che vale $B \subset B^*$

Gli elementi di B e di B^* vengono chiamati rispettivamente **boreliani** della retta reale e della retta reale ampliata.

Richiamo brevemente la nozione di probabilità così poi da estenderla e vederne le analogie.

Def 1.5. Una probabilità su una σ-algebra di eventi è una applicazione $\Pr : \mathcal{F} \to [0, 1]$ tale che:

- $\Pr(\Omega) = 1 \to \Pr(\emptyset) = 0$

- $\Pr(\bigvee_{n\geq 1} E_n) = \sum_{n\geq 1} \Pr(E_n)$ per ogni successione $(E_n)_{n\geq 1}$ di eventi a due a due incompatibili.

(ocio: probabilità definita su sigma algebra! Non su omega)

1.6 Misura

Def 1.6. Con riferimento ad una σ-algebra arbitraria $\mathcal{A}$, una **misura** su $\mathcal{A}$ è un'applicazione $\mathrm{m} : \mathcal{A} \to [0,\ +\infty]$ se:

- $\mathrm{m}(\emptyset) = 0$ la misura dell'insieme vuoto vale zero
- $\mathrm{m}\ (\bigcup_{n\geq 1} A_n) = \sum_{n\geq 1} \mathrm{m}(A_n)$ per ogni successione disgiunta $(A_n)_{n\geq 1}$ **numerabile additività**

Introduzione alla misura di Lebesgue Preso un intervallo di estremi a, b con $a \leq b$ la lunghezza dell'intervallo è $b - a$. Se prendo una sequenza finita di intervalli I_n a due a due disgiunti a_1, b_1; a_2, b_2; ...; a_n, b_n quanto misura la lunghezza della loro unione? La lunghezza della loro unione è pari alla somma delle lunghezze $\lg(\bigvee_{n\geq 1} I_n) = \sum_{i=1}^{n} (b_i - a_i)$. Se prendo una semiretta di estremi $a, +\infty$ (oppure $-\infty, a$) quanto sarà lunga la semiretta? Sarà lunga $+\infty$, ecco che interviene la condizione vista nell'altra pagina ossia che la misura può valere anche $+\infty$.

Teor. Lebesgue Si può dimostrare -risultato della tesi di dottorato di Lebesgue- che partendo da certi insiemi elementari esiste una ed una sola misura $\lambda(\cdot)$ sui boreliani tale che questa misura calcolata su un qualsiasi intervallo è proprio la lunghezza dell'intervallo:

$$\exists!\ \lambda(\cdot) \text{ su } \mathcal{B} \text{ t.c. } \lambda(\text{intervallo}) = \text{lunghezza(intervallo)}$$

Questa misura è la generalizzazione del concetto di lunghezza.
Ma siamo sicuri che l'insieme dei boreliani coincida con l'insieme delle parti? Cioè che la $\mathcal{B}$ raggiunga tutti i sottoinsiemi? Se fosse vero potremmo misurare la lunghezza di ogni insieme (intervalli, rette, semirette, ...). La risposta è no. L'italiano Vitali ha dimostrato che esistono insiemi non boreliani per i quali non si può parlare di lunghezza, che sono cioè non misurabili; sono insiemi patologici. Perciò $\mathcal{B} \subsetneqq 2^{\mathbb{R}}$. Si può parlare di lunghezza solo sui borealiani, non su qualsiasi insieme di reali. Ma poco male perché gli insiemi di interesse per le applicazioni sono tutti misurabili.

1.7 Misura di Lebesgue (unidimensionale)

Considerata la famiglia $\Im = \{\emptyset, \mathbb{R},\ ]a, b],\ ]-\infty, b],]a, +\infty[\}$

-un insieme elementare E è una unione finita di elementi di $\Im$ a due a due disgiunti

-per lunghezza dell'insieme elementare $E =]\ \alpha_1, \beta_1] \cup \cdots \cup]\ \alpha_n, \beta_n]$ si intende l'elemento della retta reale ampliata:

$$\lg(E) = \sum_{i=1}^{n} (\beta_i - \alpha_i) \ .$$

Una misura di Lebesgue è una misura sui boreliani t.c. su unioni finite disgiunte di intervalli, semirette,.. essa è la usuale lunghezza. Quindi, solamente l'insieme vuoto ha lunghezza nulla. Mentre gli insiemi elementari illimitati sono gli unici di lunghezza infinita.

Considerata allora la famiglia: $\mathcal{M} = \{S \subset \mathbb{R} | \lambda_*(S) = \lambda^*(S)\}$ degli **insiemi misurabili secondo Lebesgue**, la funzione $\lambda : \mathcal{M} \to [0, +\infty]$ così definita

$$\lambda(S) = \lambda_*(S) = \lambda^*(S)\}$$

è una estensione ad insiemi piu complessi della nozione elementare di lunghezza, essendo $\lambda(E)$=lg(E).
(Date le misure interna ed esterna di un insieme, è logico pensare che, quando queste coincidono, abbiamo a che fare esattamente con la misura di quell'insieme, più o meno come l'area delle partizioni inferiori e superiori per Riemann.)

Risulta che $\mathcal{M}$ è una σ-algebra tale che $\mathcal{B} \subset \mathcal{M} \subset 2^{\mathbb{R}}$ e che $\lambda(\cdot)$ è una misura, detta **misura di Lebesgue (unidimensionale)**.

1.8 Misura di conteggio

Indicato con $\#J$ il numero di elementi di un qualsiasi insieme finito J, fissiamo $S \subseteq \Omega$. Allora, la funzione d'insieme $\gamma_S : 2^{\Omega} \mapsto [0,\ +\infty]$ così definita:

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

che conta il numero di elementi comuni ad A e S, è una misura, detta **misura di conteggio indotta** da S su Ω.

Thm 1.1 (dim)**.** *Proprietà della misura*

- ADDITIVITÀ*:* $\mathrm{m}(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathrm{m}(A_i)$ *, se* $(A_i)_{i \leq n}$ *è disgiunta;*
- $\mathrm{m}(A_1 \cap A_2) + \mathrm{m}(A_1 \cup A_2) = \mathrm{m}(A_1) + \mathrm{m}(A_2)$ *;*

⋆ $\mathrm{m}(A_2 \backslash A_1) = \mathrm{m}(A_2) - \mathrm{m}(A_1)$ *, se* $A_1 \subseteq A_2$ e $\mathrm{m}(A_1) < +\infty$*;*

- MONOTONIA *:* $\mathrm{m}(A_1) \leq \mathrm{m}(A_2)$ *, se* $A_1 \subseteq A_2$*;*
- SUBADDITIVITÀ*:* $\mathrm{m}(\bigcup_{i \in I} A_i) \leq \sum_{i \in I} \mathrm{m}(A_i)$ *, se* I *è discreto;*
- CONTINUITÀ DAL BASSO*:* $\mathrm{m}(A_n) \uparrow \mathrm{m}(\bigcup_{n \geq 1} A_n)$ *, se* $(A_n)_{n \geq 1}$ *è una successione non decrescente;*
- CONTINUITÀ DALL'ALTO*:* $\mathrm{m}(A_n) \downarrow \mathrm{m}(\bigcap_{n \geq 1} A_n)$ *, se* $(A_n)_{n \geq 1}$ *è una successione non crescente tale che* $\mathrm{m}(A_m) < +\infty$ *per qualche* m*;*
- FORMULA D'INCLUSIONE-ESCLUSIONE*: sia* $\mathrm{m}(A_i) < +\infty$ $(i = 1, \ldots, n)$*. Allora,* $\mathrm{m}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{\emptyset \subset J \subseteq \{1,\ldots,n\}} (-1)^{\# J - 1} \mathrm{m}\left(\bigcap_{j \in J} A_j\right)$;

  Ossia una cosa del tipo: $\bigcup_{i=1}^{n} A_i| = \sum_{i=1}^{n} |A_i| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j| + \sum_{1 \leq i < j < k \leq n} |A_i \cap A_j \cap A_k| - \cdots (-1)^{n-1} |A_1 \cap \cdots \cap A_n|$ *senza valori assoluti e con la funzione misura*
- *Sia* $\mathrm{m}(A_i) = 0$ *per ogni* $i \in I$ *con* I *discreto. Allora,* $\mathrm{m}(\bigcup_{i \in I} A_i) = 0$*;*
- *Sia* $\mathrm{m}(\Omega) < +\infty$ *e* $\mathrm{m}(A_i) = \mathrm{m}(\Omega)$ *per ogni* $i \in I$ *con* I *discreto. Allora,* $\mathrm{m}(\bigcap_{i \in I} A_i) = \mathrm{m}(\Omega)$.

1.9 Applicazioni misurabili

1.9.1 Misurabilità

Def 1.7. Dati due **spazi di misura** $(\Omega, \mathcal{A})$ e $(\Omega', \mathcal{A}')$ una applicazione $f : \Omega \mapsto \Omega'$ è $(\mathcal{A}, \mathcal{A}')$-**misurabile** se $f^{-1}(A') \in \mathcal{A}$ per ogni $A' \in \mathcal{A}'$

Vediamo ora due teoremi che affermano in breve che restringendo il dominio di una applicazione misurabile il risultato è ancora una applicazione misurabile; e che la composta fra due applicazioni misurabili è ancora una applicazione misurabile.

Thm 1.2 (dim)**.** *Sussistono le seguenti proposizioni:*

- *Siano* f *un'applicazione* $(\mathcal{A}, \mathcal{A}')$*-misurabile* e S *un sottoinsieme non vuoto di* Ω*. Allora, considerata* $\mathcal{A} \cap S$*, traccia della* σ*-algrebra* $\mathcal{A}$ *su* S*, la restrizione* $f|_S$ *è un'applicazione* $(\mathcal{A} \cap S, \mathcal{A}')$*-misurabile;*

- *Dati uno spazio di misura* $(\Omega, \mathcal{A})$*, uno spazio di misura* $(\Omega', \mathcal{A}')$ *ed un ultimo spazio di misura* $(\Omega'', \mathcal{A}'')$*, siano* $f : \Omega \to \Omega'$ $(\mathcal{A}, \mathcal{A}')$*- misurabile* e

$g : \Omega' \to \Omega''$ $(\mathcal{A}',\ \mathcal{A}'')$-misurabile. Allora, l'applicazione composta $g \circ f$ è $(\mathcal{A},\ \mathcal{A}'')$-misurabilie

Thm 1.3 ($\star$ dim)**.** ***Criterio standard di misurabilità.*** *Dati due spazi di misura $(\Omega, \mathcal{A})$ e $(\Omega', \mathcal{A}')$ sia $\sigma(\mathcal{F}') = \mathcal{A}'$, sia cioè $\mathcal{F}'$ un insieme generatore di $\mathcal{A}'$. Allora $f : \Omega \mapsto \Omega'$ è $(\mathcal{A}, \mathcal{A}')$-misurabile se $f^{-1}(F') \in \mathcal{A}$ per ogni $F' \in \mathcal{F}'$.*

Dice che non occorre verificare la misurabilità su ogni insieme di $\mathcal{A}$ (per fortuna! perché potrebbero essere tanti), ma basta verificarla sulla famiglia di i-fnsiemi che forma il generatore di $\mathcal{A}'$.

Le variabili aleatorie sono un caso particolare di funzioni misurabili.

Thm 1.4. $\forall A' \in \mathcal{A}', f^{-1} \in \mathcal{A} \Leftrightarrow \forall F' \in \mathcal{F}', f^{-1}(F') \in \mathcal{A}$

Cioè per dimostrare la misurabilità della funzione f basterà dimostrare la condizione definita su un sistema di generatori, e non su ogni elemento della $\mathcal{A}'$

1.9.2 Borel-misurabilità

Def 1.8. Chiamiamo **$\mathcal{A}$-Borel misurabile** (in breve **Borel misurabile**) ogni funzione $(\mathcal{A}, \mathcal{B})$-misurabile o $(\mathcal{A}, \mathcal{B}^*)$-misurabile

1.9.3 Teor su funzioni Borel-misurabili

Thm 1.5 (dim solo per funzioni continue, dim parziale)**.** *Sussistono le seguenti proposizioni:*

- *Dati gli spazi misurabili $(\mathbb{R}^*, \mathcal{B}^*)'$ e $(\mathbb{R}^*, \mathcal{B}^*)''$, sia $\emptyset \neq S \in B^{*\prime}$. Allora le funzioni di S in $\mathbb{R}^{*\prime\prime}$ che hanno un numero discreto di punti di discontinuità sono $(B^{*\prime} \cap S)$-Borel misurabili ;*

Un esempio sono le funzioni continue oppure le funzioni monotone.

- *Dati gli spazi misurabili $(\Omega, \mathcal{A})$, $(\mathbb{R}^*, \mathcal{B}^*)'$ e $(\mathbb{R}^*, \mathcal{B}^*)''$ siano $f : \Omega \mapsto \mathbb{R}^*$ Borel misurabile e $S \in \mathcal{B}^{*\prime}$ tale che $nf(\Omega) \subseteq S$. Allora, qualunque sia la funzione $g : S \mapsto \mathbb{R}^{*\prime\prime}$ avente l'insieme dei punti di discontinuità discreto, $g \circ f$ è Borel misurabile.*

Nota che non serve richiedere che g sia borel-misurabile, perché scatta il teorema precedente.

Thm 1.6. *Siano f, g, f_n funzioni Borel misurabili. Allora, dopo aver denotato compattamente con "$\{f = g\}$" $= \{\omega | f(\omega) = g(\omega)\}$:*

- *$\{f = g\}, \{f < g\}, \{f > g\}, \{f \leq g\} e \{f \geq g\} \in \mathcal{A}$;*

- *Sono Borel misurabili le funzioni*

$$\inf_{n\geq 1} f_n \quad \sup_{n\geq 1} f_n \quad \lim_{n\to+\infty} \inf f_n \quad \lim_{n\to+\infty} \sup f_n \quad \lim_{n\to+\infty} f_n$$

- *Sono Borel misurabili le funzioni f e g, se definite ovunque, $f+g$ e $\frac{f}{g}$;*
- *Sia f derivabile. Allora f' è Borel misurabile;*
- *Sia $f_n \geq 0$. Allora, la serie $\sum_{n\geq 1} f_n$ è una funzione Borel misurabile.*

1.9.4 Funzione semplice

Def 1.9. $f : \Omega \to \mathbb{R}^*$ è una $\mathcal{A}$-**funzione semplice** (in breve **funzione semplice**) se è $\mathcal{A}$-Borel misurabile e l'insieme-immagine $f(\Omega)$ è finito

1.9.5 Funzione indicatrice

Thm 1.7 (dim rapida)**.** *La funzione indicatrice di $S \subset \Omega$ $I_s(\omega)$ è una funzione $(\mathcal{A}, \mathcal{B})$- misurabile se e solo se $S \in \mathcal{A}$*

Thm 1.8. *Se $S \in \mathcal{A}$, la funzione indicatrice di $S \subset \Omega$: $I_S(\omega) = \begin{cases} 1 & \text{se } \omega \in S \\ 0 & \text{se } \omega \notin S \end{cases}$ è una funzione semplice*

1.9.6 Proprietà della funzione indicatrice

Thm 1.9 (dim)**.** *Valgono le seguenti proprietà:*

- $I_{S^c} = 1 - I_S$
- $I_{S\cap T} = I_S \cdot I_T$
- $I_{S\cup T} = I_S + I_T - I_{S\cap T}$

Data f semplice (qui $(\mathcal{A}, \mathcal{B}^*)$-misurabile) e posto per convezione "$\{f = y\}$" $= f^{-1}(\{y\})$ per ogni $y \in \mathbb{R}^*$:

- $(\{f = y\})_{y\in f(\Omega)}$ *è una partizione finita di Ω formata da elementi di $\mathcal{A}$*
- $f = \sum_{y\in f(\Omega)} yI_{\{f=y\}}$

La prima: in pratica raggruppo gli ω in base alla loro immagine
La seconda: è una combinazione lineare con pesi i valori $y_1, ..., y_n$. Esplicitamente è: $y_1 \cdot I_{\{f=y_1\}}(\omega) + ... + y_n \cdot I_{\{f=y_n\}}(\omega)$, fino n perché funzione semplice

1.9.7 Lemma fondamentale (funzioni semplici)

Thm 1.10. ***Lemma fondamentale*** *Sussistono le seguenti proposizioni:*

• *Somme, quozienti (se definiti ovunque) e prodotti di funzioni semplici sono ancora funzioni semplici;*

⋆⋆ *Sia f una funzione Borel misurabile non negativa. Esiste allora una successione $(f_n)_{n\geq 1}$ di funzioni semplici, non negative, e a valori finiti tale che $f_n \uparrow f$;*

• *Sia f una funzione Borel misurabile. Esiste allora una successione $(f_n)_{n\geq 1}$ di funzioni semplici a valori finiti tale che $f_n \to f$ e $|f_n| \leq |f|$ per ogni n.*

1.10 Integrale di Lebesgue

$f : \Omega \to \mathbb{R}^*$ funzione $\mathcal{A}$-Borel misurabile e sia m una misura su $\mathcal{A}$. Definiamo l'integrale di Lebesgue per passi, di seguito i primi due mentre il terzo e ultimo passo avverrà in seguito:

Integrale di funzioni semplici non negative 1) **Integrale di funzioni semplici non negative**

Sia $f \geq 0$ semplice. Allora:

$$\int_\Omega f d\mathrm{m} = \sum_{y\in f(\Omega)} y \cdot \mathrm{m}(\{f = y\}) \geq 0$$

In particolare, $\mathrm{m}(A) = \displaystyle\int_\Omega I_A d\mathrm{m}$ per ogni $A \in \mathcal{A}$.

1.10.1 Integrale di funzioni Borel misurabili non negative

2) **Integrale di funzioni Borel misurabili non negative**

Sia $f \geq 0$. Allora:

$$\int_\Omega f d\mathrm{m} = \sup\{\int_\Omega g\, d\mathrm{m} : 0 \leq g \leq f \text{ e } g \text{ funzione semplice } \} \geq 0$$

Il prossimo risultato consente il passaggio del limite sotto il segno d'integrale nel caso di successioni non decrescenti di funzioni Borel misurabili non negative.

1.10.2 Teor della convergenza monotona

Thm 1.11. ***Teorema della convergenza monotona*** *Sia $(f_n)_{n\geq 1}$ una successione di funzioni Borel misurabili non negative tale che $f_n \uparrow f$. Allora,*

$$\int_\Omega f_n d\mathrm{m} \uparrow \int_\Omega f d\mathrm{m}$$

(ricordo che f_n è una successione non decrescente)

In sostanza la tesi del secondo punto del Lemma fondamentale diventa l'ipotesi del Teor qui sopra:
successione di funzioni semplici, non negative : $f_n \uparrow f$ con f B-mis $\Rightarrow$
successione di funzioni B-mis, non negative : $g_n \uparrow g$ con g (≥ 0, B-mis)

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

1.10.3 Funzioni Borel misurabili qualsiasi

1.10.4 Parte positiva e negativa di una funzione

Le funzioni:

parte positiva: $f^+ = \max(0,\ f)$

parte negativa: $f^- = \max(0,\ -f)$

sono Borel misurabili in quanto trasformate continue di f.

Thm 1.12. *Vale inoltre che*

$$f = f^+ - f^-$$

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

$$|f| = f^+ + f^-$$

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

3) **Integrale di una funzione f qualsiasi**
Sia f qualsiasi.

1.10.5 Funzione sommabile

Considerati gli integrali $\int_\Omega f^+ d\text{m}$, $\int_\Omega f^- d\text{m}$ e supposto che non siano entrambi infiniti (ocio), chiamiamo m-**sommabile** (in breve **sommabile**) la funzione f e poniamo:

$$\int_\Omega f d\text{m} = \int_\Omega f^+ d\text{m} - \int_\Omega f^- d\text{m}$$

Il valore dell'integrale di una funzione sommabile f è quindi:

- $+\infty$, se $\int_\Omega f^+ d\mathrm{m} = +\infty$
- $-\infty$, se $\int_\Omega f^- d\mathrm{m} = +\infty$
- finito, se $\int_\Omega f^+ d\mathrm{m}$ e $\int_\Omega f^- d\mathrm{m}$ sono entrambi finiti.

1.10.6 Funzione integrabile

In quest'ultimo caso diremo che f è **m-integrabile** (in breve **integrabile**). Ovvero che $\int_\Omega f d\mathrm{m}$ è finito.

Sia ora $A \in \mathcal{A}$. Allora fI_A è una funzione Borel misurabile; inoltre, tenuto conto delle disuguaglianze

$$(fI_A)^+ = f^+ I_A \leq f^+$$

e

$$(fI_A)^- = f^- I_A \leq f^-$$

fI_A risulta sommabile (integrabile) se f è sommabile (integrabile). Conseguentemente, nel caso di sommabilità della funzione f, poniamo:

$$\int_A f d\mathrm{m} = \int_\Omega f I_A d\mathrm{m}$$

e chiamiamo "$\int_A f d\mathrm{m}$" **l'integrale di Lebesgue di f su A.**

1.11 Proprietà dell'integrale di Lebesgue

Thm 1.13 (Proprietà dell'integrale di Lebesgue)**.** *Sussistono le seguenti proposizioni:*

1.11.1 Bilinearità

- LINEARITÀ RISPETTO ALLA FUNZIONE*: Siano f, g funzioni sommabili e α, β numeri reali tali che $\alpha f + \beta g$ risulti definita ovunque. Allora,*

$$\int_\Omega (\alpha f + \beta g) d\mathrm{m} = \alpha \int_\Omega f d\mathrm{m} + \beta \int_\Omega g d\mathrm{m}$$

ogni qual volta non siano infiniti di segno opposto gli addendi che compaiono al secondo membro dell'uguaglianza;

- *(dim)* LINEARITÀ RISPETTO ALLA MISURA*: Siano* m *una misura su* $\mathcal{A}$ *e* f *una funzione* m*-sommabile e* m′*-sommabile. Allora, qualunque siano* i *numeri reali non negativi* α *e* α'*, si ha*

$$\int_\Omega f d(\alpha \mathrm{m} + \alpha' \mathrm{m}') = \alpha \int_\Omega f d\mathrm{m} + \alpha' \int_\Omega f d\mathrm{m}'$$

ogni qual volta non siano infiniti di segno opposto gli integrali che compaiono al secondo membro dell'uguaglianza;

1.11.2 Monotonia

- MONOTONIA *Siano* f, g *funzioni sommabili tali che* $f \leq g$*. Allora*

$$\int_\Omega f \, d\mathrm{m} \leq \int_\Omega g \, d\mathrm{m}$$

1.11.3 Additività

- ADDITIVITÀ*: Sia* f *una funzione sommabile. Allora presi due insiemi disgiunti qualsiasi* A_1 e A_2*, risulta*

$$\int_{A_1 \cup A_2} f d\mathrm{m} = \int_{A_1} f d\mathrm{m} + \int_{A_2} f d\mathrm{m}$$

- MONOTONIA *Siano* f, g *funzioni sommabili tali che* $f \leq g$*. Allora*

$$\int_\Omega f \, d\mathrm{m} \leq \int_\Omega g \, d\mathrm{m}$$

1.11.4 Annullamento sugli insiemi di misura nulla

⋆ *(dim)* ANNULLAMENTO SUGLI INSIEMI DI MISURA NULLA *: Sia* f *una funzione sommabile e* $\mathrm{m}(A) = 0$*. Allora*

$$\int_A f \, d\mathrm{m} = 0$$

1.12 Teor su funzioni sommabili e integrabili

Thm 1.14. *Siano* f, g *due funzioni Borel misurabili. Allora:*

- *(dim)* f *è sommabile se:*
  ⋆ $f \geq g$ *con* g *funzione sommabile t.c.* $\int_\Omega g d\mathrm{m} > -\infty$
  - $f \leq g$ *con* g *funzione sommabile t.c.* $\int_\Omega g d\mathrm{m} < +\infty$
- *(dim)* f *è integrabile se e solo se lo è il suo valore assoluto;*
- *(dim)* f *è integrabile se* $|f| \leq g$ *con* g *funzione integrabile;*

⋆ *(dim)* $f \cdot g$ *è integrabile se* f^2, g^2 *sono integrabili;*

1.13 Insieme trascurabile

Def 1.10. Chiamiamo **m-trascurabile** (in breve **trascurabile**) ogni sottoinsieme S di Ω per il quale esista un insieme $A \supseteq S$ tale che $\mathrm{m}(A) = 0$.

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.

1.14 Proprietà m-quasi ovunque

Def 1.11. Diremo che una proprietà P o una relazione binaria $\asymp$ 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 $\asymp$ (m-q.o.)

Oss. Non particolarmente utile:
$-\{f\}$ finita (m-q.o.), se l'insieme $\{|f| = +\infty\}$ è trascurabile
$-\{f \leq g\}$ (m-q.o.), se l'insieme $\{f > g\}$ è trascurabile
$-\{f < g\}$ (m-q.o.), se l'insieme $\{f \geq g\}$ è trascurabile
$-\{f = g\}$ (m-q.o.), se l'insieme $\{f \neq g\}$ è trascurabile
$-\{\lim\limits_{n\to+\infty} f_n\}$ (m-q.o.), se l'insieme $A^c_{\lim}$ è trascurabile;
$-\{f_n \to f\}$ (m-q.o.), se l'insieme $\{f_n \to f\}^c$ è trascurabile;
$-\{f_n \uparrow f\}$ (m-q.o.), se l'insieme $\{\omega \in \Omega \ : \ \neg(f_n(\omega) \uparrow f(\omega))\}$ è trascurabile.

In pratica se una funzione è definita su un insieme di ω trascurabile, allora è definita m-quasi ovunque sul complementare di quell'insieme di ω.

1.15 Teor su integrali di Lebesgue

Thm 1.15 (dim)**.** *Sia f una funzione sommabile. Allora:*

- $\displaystyle\int_\Omega g\,d\mathrm{m} \leq \int_\Omega f d\mathrm{m}$, *se $g \leq f$ (m-q.o.) e g è sommabile*
- $\displaystyle\int_\Omega g\,d\mathrm{m} = \int_\Omega f d\mathrm{m}$, *se $f = g$ (m-q.o.) e g è Borel-misurabile (ocio)*
- $\displaystyle\left|\int_\Omega f d\mathrm{m}\right| \leq \int_\Omega |f|\,d\mathrm{m}$

- $\mathrm{m}(A)\inf f(A) \leq \int_A f\mathrm{dm} \leq \mathrm{m}(A)\sup f(A)$ *per ogni* A
  *(come teor media integrale: b*h=area)*

$\star$ *Se* $f \geq 0$ e $\int_\Omega f\mathrm{dm} = 0$, *allora* $f = 0$ *(m-q.o.)*

$\star\star$ *Se* $\mathrm{m}(A) > 0$ e $f(\omega) > 0$ *per ogni* $\omega \in A$, *allora* $\int_A f\mathrm{dm} > 0$

$\star\star\star$ *Se* f *è integrabile, allora* f *è finita quasi ovunque.*

1.16 Ulteriori proprietà di convergenza

1.16.1 Teor della convergenza dominata

Thm 1.16. ***Teorema della convergenza dominata*** *Sia* $(f_n)_{n\geq 1}$ *una successione di funzioni Borel misurabili tale che* $f_n \to f$. *Inoltre, sia* g *una funzione integrabile tale che* $|f_n| \leq g$ *(m-q.o.) per ogni* n. *Allora* f *ed anche* $(f_n)_{n\geq 1}$ *sono integrabili* e *si ha*

$$\int_\Omega f_n \, \mathrm{dm} \to \int_\Omega f \, \mathrm{dm}$$

1.16.2 Teor d'integrazione per serie

Thm 1.17. ***Teorema d'integrazione per serie*** *Siano* $(f_n)_{n\geq 1}$ *una successione di funzioni Borel misurabili non negative* e $(\alpha_n)_{n\geq 1}$ *una successione di numeri reali non negativi. Allora,*

$$\int_\Omega \left(\sum_{n\geq 1} \alpha_n f_n \right) \mathrm{dm} = \sum_{n\geq 1} \alpha_n \left(\int_\Omega f_n \mathrm{dm} \right)$$

Questo teorema ha riunificato i casi di v.a. a valori finiti, infinitamente numerabili e assolutamente continue

1.16.3 Legame fra integrale di Lebesgue e serie numeriche

Osservazione 1.18. ***Integrale di Lebesgue e serie numeriche*** *Sia* $S \subseteq \Omega$ *un insieme discreto avente* $N \leq +\infty$ *elementi* e *sia* $f : \Omega \mapsto \mathbb{R}^*$ una *funzione* γ_S*-sommabile. Considerata allora una numerazione* $s_1, s_2, \ldots$ *di* S, *riesce*

$$\int_\Omega f d\gamma_S = \sum_{i=1}^{N} f^+(s_i) - \sum_{i=1}^{N} f^-(s_i) = \sum_{i=1}^{N} f^+(s_i) + \sum_{i=1}^{N} (-f^-(s_i))$$

quindi, se $N = +\infty$*, l'integrale può essere calcolato sommando le serie dei valori, rispettivamente, positivi* e *negativi della restrizione* $f|_S$.

Sussiste l'uguaglianza $\int_\Omega f d\gamma_S = \sum_{n\geq 1} f(S)$ *ogniqualvolta 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* γ_S*-integrabile (serie assolutamente convergente).*

2 PROBABILITA' 1

2.1 Eventi, variabili aleatorie, probabilità, enti aleatori

Dati

- $\Omega \neq \emptyset$ (partizione dell'evento certo)
- $\mathcal{A}$ σ-algebra su Ω (degli eventi di interesse)

Def 2.1. 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 A tale che $\mathrm{P}(\Omega) = 1$;
- **variabile aleatoria** (su Ω) ogni funzione $\mathcal{A}$-Borel misurabile a valori nella retta reale;
- **variabile aleatoria estesa** (su Ω) ogni funzione $\mathcal{A}$-Borel misurabile a valori nella retta reale ampliata;

2.1.1 Ente aleatorio

Considerato lo spazio di misura$(\Omega, \mathcal{A})$ ed inoltre un altro spazio di misura $(\mathfrak{X}, \aleph)$, chiamiamo **ente aleatorio (su Ω)** a valori in $\mathfrak{X}$ ogni applicazione di Ω in $\mathfrak{X}$ che sia $(\mathcal{A}, \aleph)$-misurabile.

2.1.2 Legge di un ente aleatorio

Dato un ente aleatorio X a valori in $\mathfrak{X}$ e $A \in \aleph$, consideriamo l'applicazione $\mathrm{P}_X : \aleph \to [0, 1]$:

$$\mathrm{P}_X(\mathrm{A}) = \mathrm{P}(\{X \in \mathrm{A}\}) \ = \mathrm{P}(\{\omega | X(\omega) \in \mathrm{A}\}).$$

che è una probabilità su $\aleph$.

Def 2.2. La probabilità P_X si chiama **legge** (o **distribuzione**) dell'ente aleatorio X.

Essa è la 'probabilità immagine' di P mediante X.

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

La legge di X associa ad ogni $\mathrm{A} \in \aleph$ la probabilità $\mathrm{P}(X \in \mathrm{A})$ che l'ente aleatorio X assuma, come determinazione, un elemento dell'insieme A.

$$\begin{array}{lll} \Omega \xrightarrow{X} & \mathfrak{X} & \\ \mathcal{A} & \aleph_A & \text{Ocio che } \{X \in \mathcal{A}\} \in \mathcal{A}, \text{ mentre } A \in \aleph \\ P & P_X & \end{array}$$

Thm 2.1 (dim x casa)**.** *Risulta:*

$\bullet F_X(x) \leq F_X(x')$, *se* $x \leq x'$*;*
$\bullet F_X(-\infty) = \lim_{x \to -\infty} F_X(x) = 0$*;*
$\bullet F_X(+\infty) = \lim_{x \to +\infty} F_X(x) = 1$*;*
$\bullet F_X(x^+) = \lim_{t \to x^+} F_X(t) = F_X(x)$ *;*
$\bullet F_X(x^-) = \lim_{t \to x^-} F_X(t) = \mathrm{P}(X < x) = \mathrm{P}_X(]-\infty,\ x[)$ *;*
$\bullet F_X(x^+) - F_X(x^-) = \mathrm{P}(X = x)$.
$\star$ *Inoltre, se* m' *è una misura su* $\mathcal{B}$ *tale che* $\mathrm{m}]-\infty, x]) = F_X(x)$ *per ogni numero reale* x*, si ha* $\mathrm{m}' = \mathrm{P}_X$.

2.1.3 Ente aleatorio trasformato

Thm 2.2 (dim)**.** *Sia* $g : \mathrm{X} \mapsto \mathbb{R}^*$ *una funzione sommabile rispetto alla legge di X. Riesce allora*

$$\int_{\{X \in \mathrm{A}\}} g(X) d\mathrm{P} = \int_{\mathrm{A}} g \, d\mathrm{P}_X$$

per ogni $\mathrm{A} \in \aleph$.

2.1.4 Funzione di densità di un ente aleatorio

Posso unificare i 2 seguenti casi poiché la Legge è ricostruibile via integrazione.

Def 2.3. VARIABILI ALEATORIE DISCRETE V.a. X con rango $B = \{x_1,\ x_2,\ \ldots\}$
$f(x) = \mathrm{P}(X = x) \geq 0$ (funzione di probabilità)
Considerata la misura di conteggio γ_B su $\aleph = 2^{\mathbb{R}}$, otteniamo

$$\mathrm{P}_X(\mathrm{A}) \ = \mathrm{P}\Big(\bigcup_{n \in \{n : x_n \in \mathrm{A}\}} \{X = x_n\}\Big) = \sum_{n \in \{n : x_n \in \mathrm{A}\}} f(x_n)$$

$$= \sum_{n \geq 1} f(x_n) I_{\mathrm{A}}(x_n) = \int_{\mathrm{A}} f d\gamma_B$$

per ogni $\mathrm{A} \in \aleph$.

Def 2.4. VARIABILI ALEATORIE ASSOLUTAMENTE CONTINUE V.a. assolutamente continua X con f funzione di densità, cioè $f \geq 0$ e

$$F_X(x) = \int_{-\infty}^{x} f(t)dt$$

per ogni numero reale x. Allora

$$\mathrm{P}_X(B) = \int_B f(t)dt$$

per ogni boreliano B.

Nei due casi classici 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 γ_B; nel secondo λ).

Ritornando al caso generale, supponiamo che sulla σ-algebra $\aleph$ sia definita una **misura di riferimento** μ.

2.1.5 Mu-densità di un ente aleatorio

Def 2.5. Chiamiamo **μ-densità di** X ogni funzione $f : \mathbb{R} \mapsto [0,\ +\infty]$ $\aleph$-Borel misurabile tale che

$$\mathrm{P}_X(\mathrm{A}) = \int_{\mathrm{A}} f d\mu$$

per ogni $\mathrm{A} \in \aleph$

2.1.6 Non unicità della densità di un ente

Thm 2.3 (dim)**.** *Siano $f, \tilde{f}$ due μ-densità di X. Allora,*

$$\tilde{f} = f \ (\mu - q.o.)$$

Una μ-densità è definita ($\mu-$q.o.), **a meno di insiemi μ-trascurabili** quindi non è unica.

2.1.7 Collegamento fra densità e legge di un ente

Thm 2.4 (dim.)**.** *Sia f una μ-densità di X. Se $g : \mathfrak{X} \to \mathbb{R}^*$ è sommabile rispetto alla legge di X, allora*

$$\int_{\mathrm{A}'} g \, d\mathrm{P}_X = \int_{\mathrm{A}'} g f d\mu$$

per ogni $A' \in \aleph$.

2.1.8 Teor fondamentale del calcolo delle probabilità

Thm 2.5 (dim)**.** ***Teorema fondamentale del calcolo delle probabilità*** *Sia* $g : \mathfrak{X} \mapsto \mathbb{R}^*$ *una funzione sommabile rispetto alla legge di* X*. Riesce allora*

$$\int_{\{X \in \mathrm{A}\}} g(X) d\mathrm{P} = \begin{cases} \displaystyle\int_{\mathrm{A}} g \, d\mathrm{P}_X \\ \displaystyle\int_{\mathrm{A}} g \, f d\mu \quad \text{se } f \text{ é una } \mu\text{−densità di } X \end{cases} \qquad \textit{per ogni}$$

$\mathrm{A} \in \aleph$.

2.1.9 Enti aleatori equidistribuiti

Def 2.6. Due enti aleatori X, Y a valori in $\mathfrak{X}$ si dicono **equidistribuiti** se $\mathrm{P}_X = \mathrm{P}_Y$, cioè se hanno la medesima legge.

Thm 2.6 (dim)**.** *Conseguentemente,* X *e* Y *sono equidistribuiti se sono uguali* ***quasi certamente*** *(cioè* P*-quasi ovunque).*

3 Speranza matematica di una variabile aletaoria

Mentre prima..

- $\mathcal{P}$ partizione dell'evento certo
- $\mathcal{F}\sigma$-algebra di eventi logicamente dipendenti da $\mathcal{P}$
- Pr probabilità su $\mathcal{F}$
- $X : \mathcal{P} \to \mathbb{R}$ variabile aleatoria (cioè una funzione $\mathcal{F}$-Borel misurabile)
- E(X) speranza matematica è pari a:
  - se X ha rango $B = \{x_1, ..., \ x_n\}$ risulta $\mathrm{E}(X) = x_1 \mathrm{Pr}(X = x_1) + \cdots + x_n \mathrm{Pr}(X = x_n) = \int_{\mathcal{P}} X d\mathrm{Pr}$
  - se X è assolutamente continua: denotata con f una sua funzione di densità, dal Teorema fondamentale (ponendo $g(x) = x$) otteniamo $\mathrm{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx = \int_{\mathbb{R}} x f(x) \lambda(dx) = \int_{\mathcal{P}} X d\mathrm{Pr}$

Ora invece..

Def 3.1. La **speranza matematica** (o **valore medio**) di una v.a. estesa P-sommabile X è l'elemento della retta reale ampliata: $E(X) = \int_{\Omega} X d\mathrm{P}$

Nelle ipotesi del teorema fondamentale otteniamo

$$\mathrm{E}(g(X)) = \int_\Omega g(X) d\mathrm{P} = \begin{cases} \displaystyle\int_{\mathbb{R}} g\, d\mathrm{P}_X \\ \\ \displaystyle\int_{\mathbb{R}} g f d\mu \quad \text{se } f \text{ è una } \mu - \text{densità di } X \end{cases}$$

Nell'ulteriore caso particolare che g sia l'identità si ha

$$\mathrm{E}(X) = \begin{cases} \displaystyle\int_{\mathbb{R}} x\, \mathrm{P}_X\ (dx) \\ \\ \displaystyle\int_{\mathbb{R}} x f(x) \mu(dx) \quad \text{se } f \text{ è una } \mu - \text{densità di } X \end{cases}$$

3.1 Proprietà

Thm 3.1 (dim solo se c'è $\star$)**.** *Riesce* $\mathrm{E}(I_A) = \mathrm{P}(A)$ *e* $\mathrm{E}(\alpha I_\Omega) = \alpha$ *per ogni* $A \in \mathcal{A}$ *e* $\alpha \in \mathbb{R}^*$*. Inoltre, supposto che* X, Y *siano* v.a. *estese con speranza matematica, risulta:*

3.1.1 Linearità

- LINEARITÀ *: Se* $\alpha, \beta \in \mathbb{R}^*$ *sono tali che* $\alpha X + \beta Y$ *è definita ovunque, allora*
$$\mathrm{E}(\alpha X + \beta Y) = \alpha \mathrm{E}(X) + \beta \mathrm{E}(Y)$$
se $\alpha \mathrm{E}(X)$ *e* $\mathrm{E}(Y)$ *non sono infiniti di segno opposto;*

3.1.2 Uguaglianza

- *se* $X = Y$ $(\mathrm{P} - \text{q.c.})$ *allora* $\mathrm{E}(X) = \mathrm{E}(Y)$ *;*

3.1.3 Monotonia

- MONOTONIA*: se* $X \leq Y$ $(\mathrm{P} - \text{q.c.})$ *allora* $\mathrm{E}(X) \leq \mathrm{E}(Y)$

3.1.4 Valore assoluto

- $|\mathrm{E}(X)| \leq \mathrm{E}(|X|)$ *;*

3.1.5 Internalità

- INTERNALITÀ*:* $\inf X \leq \mathrm{E}(X) \leq \sup X$*;*

3.1.6 Internalità stretta

⋆ INTERNALITÀ STRETTA *se* $\alpha < X(\omega) < \beta$ $\forall\omega$ *allora* $\alpha < \mathrm{E}(X) < \beta$*;*

3.1.7 Disuguaglianza di Markov

⋆⋆ *Se* $X \geq 0$ *e* $a > 0$*, allora* $\Pr(X \geq a) \leq \dfrac{\mathrm{E}(X)}{a}$ *disuguaglianza di Markov;*

3.1.8 Ammissione

- *Infine, una* v.a. Z *ammette speranza matematica se* $Z \geq X$ *e* $\mathrm{E}(X) > -\infty$ *o* $Z \leq X$ *e* $\mathrm{E}(X) < +\infty$.

Def 3.2. Dato un intervallo aperto J (limitato o no) della retta reale diremo che $g : J \mapsto \mathbb{R}$ è **convessa** se, dati due punti arbitrari $a, b \in J$ tali che $a < b$, il grafico di $g_{[a,b]}$ sta sotto la retta passante per i punti $(a,\ g(a))$, $(b,\ g(b))$. Formalmente

$$g(x) \leq \frac{g(b) - g(a)}{b - a}x + \frac{g(a)b - g(b)a}{b - a}$$

per ogni $x \in [a,\ b]$.

Osservazione 3.2. *Osservato che, dato* $x \in [a,\ b]$*, risulta*

$$x = \frac{b - x}{b - a}a + \frac{x - a}{b - a}b$$

con

$$\frac{b - x}{b - a} + \frac{x - a}{b - a} = 1,\ \frac{b - x}{b - a},\ \frac{x - a}{b - a} \geq 0$$

possiamo concludere che ogni punto dell′*intervallo* $[a,\ b]$ *è ottenibile come* ***mistura*** *degli estremi dell'intervallo. Pertanto* g *è convessa se per ogni* $a, b \in J$ *e per ogni* $\alpha \in [0, 1]$ *risulta*

$$g(\alpha a + (1 - \alpha)b) \leq \alpha g(a) + (1 - \alpha)g(b)$$

Ne segue per induzione

$$g(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leq \alpha_1 g(x_1) + \cdots + \alpha_n g(x_n)$$

per ogni x_1*,...,* $x_n \in J$ *e* α_1*,... ,* $\alpha_n \geq 0$ *tali che* $\alpha_1 + \cdots + \alpha_n = 1$

Lemma 3.3. $g : J \mapsto \mathbb{R}$ *convessa. Allora:*

- *Per ogni punto* a *di* J *esiste un numero reale* b *tale che:*

$$g(x) \geq b(x - a) + g(a)\,\text{per ogni}\,x \in J$$

- g *è continua e quindi* $B \cap J$*-Borel misurabile*

3.1.9 Disuguaglianza di Jensen

Thm 3.4 (dim)**.** ***Disuguaglianza di Jensen*** *Dato un intervallo aperto J (limitato o no) $\subset \mathbb{R}$, siano $g : J \to \mathbb{R}$ una funzione convessa e $X : \Omega \mapsto J \subset \mathbb{R}$ una* v.a. *con speranza matematica. Allora* $\mathrm{E}(X) \in J$*. Inoltre, la* v.a. *$g(X)$ ammette speranza matematica* e *risulta* $g(\mathrm{E}(X)) \leq \mathrm{E}(g(X))$, *cioè*

$$g\left[\int_\Omega X d\mathrm{P}\right] \leq \int_\Omega g(X) d\mathrm{P}$$

3.2 Varianza di una variabile aletaoria

Def 3.3. Data una v.a. X con speranza matematica finita, la **varianza** di X è il *momento centrale secondo*:

$$Var(X) = \mathrm{E}((X - \mathrm{E}(X))^2)$$

3.3 Proprietà della varianza

Thm 3.5 (dim)**.** *Sia X una* v.a. *con speranza matematica finita. Allora:*

3.3.1 Def alternativa

⋆ $Var(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$*;*

3.3.2 Varianza di una trasformata affine lineare

⋆⋆ $\mathrm{Var}(\alpha X + \beta) = \alpha^2 \mathrm{Var}(X)$ *per ogni α, β reali;*

3.3.3 Varianza nulla

- *Var(X)* $= 0$ *se e solo se* $X = \mathrm{E}(X)$ *(P-q.c.);*

3.3.4 Disuguaglianza di Bienaymé - Čebičev

⋆⋆⋆ *Se $\epsilon > 0$, allora*

$$\mathrm{Pr}(|X - \mathrm{E}(X)| \geq \epsilon) \leq \frac{\mathrm{Var}(X)}{\epsilon^2} \quad \textit{disuguaglianza di Bienaymé} - \text{Čebičev}$$

3.4 Covarianza di una variabile aletaoria

Def 3.4. Considerate due v.a. X, Y con speranza matematica finita e tali che esista finita anche la speranza matematica del loro prodotto $\mathrm{E}(XY)$, la covarianza di X e Y è la differenza:

$$\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$$

3.5 Proprietà della covarianza

Thm 3.6 (dim)**.** *Riesce che:*

3.5.1 Commutativa

- $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$, *e in particolare,* $\mathrm{Var}(X) = \mathrm{Cov}(X, X)$

3.5.2 Def alternativa

- $\mathrm{Cov}(X, Y) = \mathrm{E}((X - \mathrm{E}(X))(Y - \mathrm{E}(Y)))$

Thm 3.7 (dim solo se c'è $\star$)**.** *Siano* X, Y, Z v.a. *con speranza matematica finita e con* $\mathrm{E}(XY)$ *e* $\mathrm{E}(YZ)$ *finite. Allora:*

- $\star$ $\mathrm{Cov}(\alpha X,\ Y) = \alpha \mathrm{Cov}(X,\ Y)$ *per ogni numero reale* α*;*
- $\star$ $\mathrm{Cov}(X + Z, Y) = \mathrm{Cov}(X,\ Y) + \mathrm{Cov}(Y,\ Z)$ *;*

3.5.3 Disuguaglianza di Cauchy-Schwarz

- $\star$ $\mathrm{Cov}(X, Y)^2 \leq \mathrm{Var}(X)\mathrm{Var}(Y)$ *(Disuguaglianza di Cauchy-Schwarz).*

  Inoltre, se $X_1, \ldots, X_n$ *sono* v.a. *con speranza matematica finita* e *tali che* $\mathrm{E}(X_i X_j)$ *finita per ogni* $i \neq j$*, si ha*

$$\mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2 \sum_{i=1}^{n-1} \sum_{j>i} \mathrm{Cov}(X_i,\ X_j).$$

In tutto le disuguaglianze da sapere sono:

- Markov
- Jensen
- Cebicev
- Chauchy-Schwartz

3.6 Indice di correlazione di Bravais

Def 3.5. Chiamiamo **indice di correlazione di Bravais** il seguente rapporto

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Oss. Da Chauchy-Schwartz, purché entrambe le varianze siano ¿0 e finite segue che $\rho_{X,Y} \leq 1$ da cui:

- $\rho_{X,Y} = -1$ allora X (o Y) è trasformata affine decrescente dell'altra (P-q.c.)
- $\rho_{X,Y} = 1$ allora X (o Y) è trasformata affine crescente dell'altra (P-q.c.)
- $\rho_{X,Y} = 0$ allora X e Y sono *non correlate*
- $-1 < \rho_{X,Y} < 0$ allora sono correlate negativamente
- $0 < \rho_{X,Y} < 1$ allora sono correlate positivamente
- se vale l'uguale allora è *quasi certo* che una v.a. è trasformata affine dell'altra con coeff > 0 se Cov>0 (idem per < 0)

3.7 Comonotonia

Def 3.6. Due v.a. X, Y si dicono **comonotone** se

$$[X(\omega) - X(\omega')][Y(\omega) - Y(\omega')] \geq 0$$

per ogni ω, ω'.

Sono comonotone se hanno lo circa stesso andamento rispetto alla crescenza/decrescenza dell'altra: ad esempio se una v.a. cresce l'altra non decresce.

3.7.1 Teor su v.a. comonotone

Thm 3.8 (dim)**.** *Siano* X, Y *v.a. comonotone con* $E(X), E(Y), E(XY)$ *finite. Allora* $\mathrm{Cov}(X, Y) \geq 0$

La comonotonia assicura che non possono essere correlate negativamente

4 PROBABILITA' 2

4.1 Misura prodotto

4.1.1 Misura sigma-finita

Def 4.1. Una misura m è **σ-finita** se esiste una partizione discreta $(A_n)_{n\geq 1}$ di Ω con $\mathrm{m}(A_n) < +\infty$ per ogni n.

Misure σ-finite sono le misure finite (in particolare la probabilità), la misura di Lebesgue unidimensionale e le misure di conteggio indotte da insiemi discreti.

Dato $\Omega_i \neq \emptyset$, siano $\mathcal{A}_i$ una σ algebra su Ω_i, m_i una misura σ-finita su $\mathcal{A}_i$ $(i = 1, 2)$ e posto $\Omega = \Omega_1 \times \Omega_2$, consideriamo la:

4.2 Famiglia dei rettangoli misurabili

Def 4.2. famiglia dei **rettangoli misurabili**

$$\mathcal{R} = \{\mathrm{A}_1 \times \mathrm{A}_2 \ : \ \mathrm{A}_i \in \mathcal{A}_i (i = 1, 2)\}$$

4.3 Sigma-algebra prodotto

Def 4.3. **σ-algebra prodotto**:

$$\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{R})(\text{di } \mathcal{A}_1 \,\mathrm{e}\, \mathcal{A}_2)$$

Considerate infine, per ogni $S \subset \Omega$ la:

Def 4.4. **sezione di S relativa a** $\omega_1 \in \Omega_1$:

$$S(\omega_1) = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in S\}$$

Def 4.5. **sezione di S relativa a** $\omega_2 \in \Omega_2$:

$$S(\omega_2) = \{\omega_1 \in \Omega_1 : (\omega_1, \omega_2) \in S\},$$

Ora verrà introdotta una misura sulla σ-algebra $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$ tale che verifica le seguenti due proprietà (suggerite dalla nozione di area delle figure piane e del relativo Principio di Cavalieri):

(a) la misura di un rettangolo misurabile $\mathrm{A}_1 \times \mathrm{A}_2$ di "base A_1 e \altezza" A_2 è uguale al prodotto $\mathrm{m}_1(\mathrm{A}_1)\mathrm{m}_2(\mathrm{A}_2)$;

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

4.4 Unicità della misura prodotto

Lemma 4.1 (dim x casa)**.** *Sia* $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$*. Allora:*

- $A(\omega_i) \in \mathcal{A}_j$ *per ogni* $\omega_i (i = 1, 2)$ *;*
- $\omega_i \mapsto \mathrm{m}_j(A(\omega_i))$ *è* $\mathcal{A}_i$*-Borel misurabile* $(i = 1, 2; j \neq i)$

Def 4.6. Definiamo sulla σ-algebra prodotto la misura prodotto di m_1 e m_2:

$$\mathrm{m}_1 \times \mathrm{m}_2\ (A) = \int_{\Omega_1} \mathrm{m}_2(A(\omega_1))\mathrm{m}_1(d\omega_1)$$

Thm 4.2 (dim)**.** *Si dimostra che* $\mathrm{m}_1 \times \mathrm{m}_2(A)$ *è una misura.*

Il prossimo teorema assicura che la misura prodotto è l'unica misura sulla σ-algebra prodotto che verifica le proprietà (a) e (b) dell'altra pagina.

Thm 4.3. *Risulta:*

- $\mathrm{m}_1 \times \mathrm{m}_2(\mathrm{A}_1 \times \mathrm{A}_2) = \mathrm{m}_1(\mathrm{A}_1)\mathrm{m}_2(\mathrm{A}_2)$ *per ogni* $\mathrm{A}_1 \times \mathrm{A}_2$*;*
- $\mathrm{m}_1 \times \mathrm{m}_2(A) = \displaystyle\int_{\Omega_2} \mathrm{m}_1(A(\omega_2))\mathrm{m}_2(d\omega_2)$ *per ogni* $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$.

Thm 4.4. *Risulta inoltre che qualunque sia la misura* m' *sulla* σ*-algebra prodotto tale che* $\mathrm{m}'|_{\mathcal{R}} = \mathrm{m}_1 \times \mathrm{m}_2|_{\mathcal{R}}$*, riesce* $\mathrm{m}' = \mathrm{m}_1 \times \mathrm{m}_2$.

I seguenti teoremi sono due "pietre miliari" della teoria dell'integrazione in quanto consentono di **ridurre** l' **integrale doppio**

$$\int_{\Omega} f d\mathrm{m}_1 \times \mathrm{m}_2 = \int_{\Omega} f(\omega_1,\ \omega_2)\mathrm{m}_1 \times \mathrm{m}_2(d\omega_1 \times d\omega_2)$$

ad un **integrale iterato** come pure di invertire l'ordine delle integrazioni successive.

4.5 Teor di Tonelli

Il seguente teorema di Tonelli dice che l'integrale di una funzione non negativa sul prodotto di due spazi sigma finiti rispetto alla misura prodotto, coincide con l'integrale iterato rispetto alle due misure.

Thm 4.5. ***Teorema di Tonelli*** *Sia* $f \geq 0$*,* $\mathcal{A}_1 \otimes \mathcal{A}_2$*-Borel misurabile. Allora, per ogni* A_1 *e* A_2*, le funzioni:*

$$\omega_1 \to \int_{\mathrm{A}_2} f(\omega_1, \cdot\ )d\mathrm{m}_2$$

$$\omega_2 \to \int_{\mathrm{A}_1} f(\ \cdot, \omega_2)d\mathrm{m}_1$$

sono, rispettivamente, $\mathcal{A}_1$-Borel misurabile e $\mathcal{A}_2$-Borel misurabile. Risulta inoltre che:

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

4.6 Teor di Fubini

Il seguente teorema di Fubini dice che se l'integrale iterato ha valore finito, allora il valore dell'integrale è indipendente dall'ordine di integrazione. Esso dice che se l'integrale esiste a meno di una costante (q.o.) allora vale la formula di riduzione.

Thm 4.6. ***Teorema di Fubini*** *Sia $f m_1\times m_2$-sommabile. Risulta:*

$$A_1^{(f)} \subseteq \{\omega_1 : f(\omega_1,\ \cdot) \text{ è } m_2-\text{sommabile}\} \in \mathcal{A}_1$$

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

e $m_i\left((A_i^{(f)})^c\right) = 0\ (i=1,2)$.
Inoltre, per ogni $A_i (i=1,2)$, le funzioni:

$$g_1 : \omega_1 \mapsto \begin{cases} \displaystyle\int_{A_2} f(\omega_1,\ \cdot) dm_2 & \text{se } \omega_1 \in A_1^{(f)}\\ 0 & \text{se } \omega_1 \notin A_1^{(f)}\end{cases}$$

$$g_2 : \omega_2 \mapsto \begin{cases} \displaystyle\int_{A_1} f(\cdot\ ,\omega_2) dm_1 & \text{se } \omega_2 \in A_2^{(f)}\\ 0 & \text{se } \omega_2 \notin A_2^{(f)}\end{cases}$$

La conclusione dei due teoremi è la stessa ma le ipotesi sono molto diverse.

4.7 Estensione dei risultati a 3 o più variabili

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

Def 4.7. Siano

- $\mathcal{A}_h$ una σ-algebra su un insieme $\Omega_h \neq \emptyset$
- m_h una misura σ-finita su $\mathcal{A}_h (h=1,\ \ldots,\ m)$
- $\Omega = \Omega_1\times\cdots\times\Omega_m$

4.7.1 Rettangoli misurabili

Def 4.8. Definiamo la famiglia dei **rettangoli misurabili**:

$$\mathcal{R}^{(m)} = \{\mathrm{A}_1 \times \cdots \times \mathrm{A}_m \ : \ \mathrm{A}_h \in \mathcal{A}_h \ (h = 1, \ \ldots, \ m)\}$$

4.7.2 Sigma-algebra prodotto m-dimensionale

Def 4.9. Definiamo la σ-algebra prodotto $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_m = \sigma(\mathcal{R}^{(m)})$

4.7.3 Misura prodotto

Thm 4.7. *Esiste una sola misura* $\mathrm{m} = \mathrm{m}_1 \times \cdots \times \mathrm{m}_m$ *sulla* σ*-algebra prodotto, detta **misura prodotto di*** $(\mathrm{m}_1, \ \ldots, \ \mathrm{m}_m)$*, tale* che*:*

$$\mathrm{m}_1 \times \cdots \times \mathrm{m}_m(\mathrm{A}_1 \times \cdots \times \mathrm{A}_m) = \prod_{h=1}^{m} \mathrm{m}_h(\mathrm{A}_h)$$

per ogni rettangolo misurabile $\mathrm{A}_1 \times \cdots \times \mathrm{A}_m$

4.7.4 Teor di Fubini e Tonelli m-dimensionali

Thm 4.8. *Sussiste la generalizzazione dei teoremi di Tonelli e Fubini che permette di ridurre l'**integrale multiplo** ad un **integrale iterato**:*

$$\int_\Omega f \ d(\mathrm{m}_1 \times \cdots \times \mathrm{m}_m) = \int_\Omega f(\omega_1, \ \ldots, \ \omega_m) \mathrm{m}_1 \times \cdots \times \mathrm{m}_m(d\omega_1 \times \ \cdots \ \times d\omega_m).$$

Il seguente teorema permette di invertire l'ordine delle integrazioni successive

Thm 4.9. *Sia f* m*-sommabile. Data una permutazione $h_1, \ldots, h_m$ di $\{1, \ldots, m\}$ si ha*

$$\int_\Omega f d\mathrm{m} = \int_{\Omega_{h_1}} \left(\cdots \left(\int_{\Omega_{h_m}} f(\omega_1, \ \ldots, \ \omega_m) \mathrm{m}_{h_m}(d\omega_{h_m}) \right) \cdots \right) \mathrm{m}_{h_1}(d\omega_{h_1}),$$

dove, nel caso che non sia $f \geq 0$, ogni integrale

$$\int_{\Omega_{h_k}} \left(\cdots \left(\int_{\Omega_{h_m}} f(\omega_1, \ \ldots, \ \omega_m) \mathrm{m}_{\mathrm{h_m}}(d\omega_{h_m}) \right) \cdots \right) \mathrm{m}_{h_k}(d\omega_{h_k}) \quad (k > 1)$$

esiste per ogni $(\omega_{h_1}, \ \ldots, \ \omega_{h_{k-1}})$ non appartenente ad un insieme $\mathrm{m}_{h_{k-1}}$ -trascurabile e *viene esteso (per effettuare l'integrazione successiva relativa alla misura $\mathrm{m}_{h_{k-1}}$) su tale insieme dandogli valore zero.*

4.7.5 Sigma-algebra di Borel m-dimensionale

Def 4.10. Definiamo la **σ-algebra di Borel** di $(\mathbb{R}^*)^m$ come la σ-algebra prodotto $\mathcal{B}^{(m)} = \overbrace{\mathcal{B} \otimes \cdots \otimes \mathcal{B}}^{\text{m volte}}$ che:
[-] come nel caso unidimensionale $(\mathbb{R}^*)^m$ può essere vista come la σ-algebra generata dalla famiglia degli insiemi aperti di m-uple, oppure dagli insiemi $S_{x_1,...,x_m} =]-\infty, x_1] \times ... \times]-\infty, x_m]$ $(x_1, ..., x_m \in \mathbb{R})$
[-] è costituita dai **boreliani** di $(\mathbb{R}^*)^m$

4.7.6 Misura di Lebesgue m-dimensionale

Def 4.11. Definiamo la **misura di Lebesgue** m-**dimensionale** come la misura prodotto

$$\lambda^{(m)} = \overbrace{\lambda \times ... \times \lambda}^{\text{m volte}}$$

4.8 Coppia aleatoria

Def 4.12. Denotiamo :
con $\aleph_i$ la σ-algebra su $\mathfrak{X}_i \neq \emptyset$ $(i = 1, 2)$
con $X_i : \Omega \mapsto \mathrm{X}_i (i = 1, 2)$
con $\mathfrak{X} = \mathfrak{X}_1 \times \mathfrak{X}_2$
con $\aleph = \aleph_1 \otimes \aleph_2$

Def 4.13. $\mathbf{X} = (\mathrm{X}_1, \mathrm{X}_2) : \omega \mapsto (X_1(\omega), X_2(\omega))$ **coppia aleatoria**. Cioé ogni ente aleatorio su Ω a valori in $\mathfrak{X}$

Thm 4.10 (dim x casa)**.** *$\boldsymbol{X}$ è una coppia aleatoria se e solo se X_i è un ente aleatorio a valori in $\mathfrak{X}_i$ $(i = 1, 2)$.*

4.9 Legge congiunta

Def 4.14. Legge (o **distribuzione**) **congiunta** di X_1, X_2 è la legge $\mathrm{P_X}$ di $\mathbf{X}$

4.10 Densità congiunta

Def 4.15. Sia μ_i misura σ-finita di riferimento su $\aleph_i (i = 1, 2)$, allora chiamo (μ_1, μ_2)-**densità congiunta** di X_1, X_2 ogni funzione $\aleph$-Borel misurabile $f : \mathfrak{X} \mapsto [0, +\infty]$ tale che
$\mathrm{P_X(A) = P(X \in A)} = \int_{\mathrm{A}} f d(\mu_1 \times \mu_2) = \int_{\mathrm{A}} f(x_1, x_2) \mu_1 \times \mu_2 (dx_1 \times dx_2)$

4.11 Legge marginale e densità marginale di X_i

Def 4.16. **Legge marginale** di X_i è la legge P_{X_i}; mentre la **densità marginale** di X_i è una sua μ_i-densità $(i = 1, 2)$

Breve riepilogo su cosa è stato definito: $\Omega \xrightarrow{X_1} \mathfrak{X}_1$

$\mathcal{A}$ $\quad \aleph_1$

$P \quad P_{X_1} \; / \; \mu_1 = \begin{cases} f \geq 0 \\ P_{X_1}(A \in \aleph) = \int_A f d\mu \end{cases}$

$\Omega \xrightarrow{X_2} \mathfrak{X}_2$

$\mathcal{A}$ $\quad \aleph_2$

$P \quad P_{X_2} \; / \; \mu_2 = \begin{cases} f \geq 0 \\ P_{X_2}(A \in \aleph) = \int_A f d\mu \end{cases}$

4.12 Funzione di ripartizione congiunta

Def 4.17. Definiamo la **Funzione di ripartizione congiunta** delle v.a. X_1, X_2, la funzione:

$$F_{\mathrm{X}}(x_1,\ x_2) = \mathrm{P}_{\mathrm{X}}\left(]-\infty,\ x_1] \times]-\infty,\ x_2]\right) = \mathrm{P}((X_1,\ X_2) \leq (x_1,\ x_2))$$

4.13 Teor fondamentale del calcolo delle probabilità (bivariato)

Thm 4.11. ***Teorema fondamentale del calcolo delle probabilità (caso bivariato*** *Sia* $g : \mathrm{X} \mapsto \mathbb{R}^*$ *una funzione sommabile rispetto alla legge congiunta di* $X_1 \,\mathrm{e}\, X_2$. *Allora, posto* $\mu = \mu_1 \times \mu_2$, *si ha*

$$\int_{\{\mathrm{X} \in \mathrm{A}\}} g(X_1, X_2) d\mathrm{P} = \begin{cases} \displaystyle\int_{\mathrm{A}} g d\mathrm{P}_{\mathrm{X}} \\ \\ \displaystyle\int_{\mathrm{A}} g f d\mu \end{cases}$$

per ogni $\mathrm{A} \in \aleph$

Thm 4.12 (dim). *Risulta:*

- $\mathrm{P}_{X_1}(\mathrm{A}_1) = \mathrm{P}_{\boldsymbol{X}}(\mathrm{A}_1 \times \mathfrak{X}_2)$, $\mathrm{P}_{X_2}(\mathrm{A}_2) = \mathrm{P}_{\boldsymbol{X}}(\mathfrak{X}_1 \times \mathrm{A}_2)$ *per ogni* $\mathrm{A}_i \in \aleph_i$;
- •• *Se* f *è una* (μ_1, μ_2)*-densità congiunta di* X_1, X_2, *la funzione:*

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

per ogni $x_i \in \mathrm{X}_i$, *è una* μ_i*-densità marginale di* $X_i (i = 1, 2)$.

4.14 Indipendenza di enti aleatori

Due enti aleatori X_1 e X_2 sono **indipendenti** se

$$\mathrm{P_X}(\mathrm{A}_1 \times \mathrm{A}_2) = \mathrm{P}_{X_1}(\mathrm{A}_1)\mathrm{P}_{X_2}(\mathrm{A}_2)$$

per ogni rettangolo misurabile $\mathrm{A}_1 \times \mathrm{A}_2$.

4.14.1 Teor su equivalenza indipendenza fra e.a.

Thm 4.13. *Sono equivalenti le proposizioni:*

•X_1 *e* X_2 *sono indipendenti;*

•• $\mathrm{P_X} = \mathrm{P}_{X_1} \times \mathrm{P}_{X_2}$.

••• *Se* f_X *è una* $(\mu_1,\ \mu_2)$*-densità congiunta di* X_1 *e* X_2*, la proposizione • è equivalente alla:* $f_\mathrm{X}(x_1,\ x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$*, a meno di un insieme* $\mu_1 \times \mu_2$*-trascurabile*

4.14.2 Teor su indipendenza di e.a. trasformati

Thm 4.14 (dim)**.** *Siano* $\mathcal{G}_i$ σ*-algebra su* $\mathfrak{Y}_i$ *e* $g_i : \mathfrak{X}_i \mapsto \mathfrak{Y}_i$ *un'applicazione* $(\aleph_i, \mathcal{G}_i)-$ *misurabile* $(i = 1, 2)$*. Siano inoltre* X_1, X_2 *indipendenti. Allora, sono pure indipendenti gli enti aleatori trasformati* $g_1(X_1)$ *e* $g_2(X_2)$*.*

4.14.3 Teor su conseguenza indipendenza v.a. su E()

Thm 4.15 (dim)**.** *Siano* X_1, X_2 *v.a. indipendenti. Allora, se* $X_i \geq 0$ $(i = 1, 2)$ *o* $\mathrm{E(X)}$ *finita* $(i = 1, 2)$*, esiste la speranza matematica del prodotto* e *si ha*

$$\mathrm{E}(X_1X_2) = \mathrm{E}(X_1)\mathrm{E}(\mathrm{X}_2)$$

Oss. Estensioni ad m-uple
Le definizioni date si estendono "pari pari" alle m-ple aleatorie ($m \geq 3$) e rimangono validi tutti i risultati ottenuti.

4.14.4 Teor su sottoinsieme di e.a. indipendenti

Thm 4.16. *Se gli enti aleatori* $X_1, \ldots, X_m$ *sono* ***indipendenti****, allora lo sono pure quelli ottenuti operando una selezione tra di essi.*

4.14.5 Indipendenza successione di eventi

Def 4.18. I termini della successione $X_1, X_2, \ldots$ di enti aleatori si dicono **indipendenti** se sono indipendenti $X_{k_1}, \ldots, X_{k_n}$ (ocio) per ogni scelta di $k_1, \ldots, k_n$ distinti e di n.

5 VARIABILI ALEATORIE

I vari tipi di convergenza di seguito, che vedremo, sono via via più deboli, (la convergenza puntuale dell'analisi matematica non rientra nell'elenco perché è più forte di tutte)

1 (convergenza certa)

2 convergenza quasi certa: ha unicità q.c. del limite

3 convergenza in probabilità: ha unicità q.c. del limite

4 convergenza in distribuzione

5.1 Convergenza quasi certa

Def 5.1. Una successione $(X_n)_{n\geq 1}$ di v.a. converge **quasi certamente** alla v.a. X se: $X_n \to X$;

$$(\mathrm{P}-\text{q.c.}), \text{cioè } \mathrm{P}(X_n \to X) = 1,$$

ove $\{X_n \to X\} = \{\omega \ : \ X_n(\omega) \to X(\omega)\}$ è l'**insieme di convergenza**

5.1.1 Teor caratterizzazione cqc

Thm 5.1 (dim)**.** *Le seguenti proposizioni sono equivalenti:*

i) $X_n \to X \ (\mathrm{P}-\text{q.c.})$ *;*

ii) $\lim_{n\to+\infty} \mathrm{P}(\bigcap_{m\geq n} \{|X_m - X| < \epsilon\}) = 1$ *qualunque sia* $\epsilon > 0$*;*

iii) $\lim_{n\to+\infty} \mathrm{P}(\bigcap_{m\geq n} \{|X_m - X| \leq \epsilon\}) = 1$ *qualunque sia* $\epsilon > 0$.

5.1.2 Teor unicità q.c. del limite e convergenza in probabilità della trasformata continua

Thm 5.2 (dim)**.** *Sia* $X_n \to X(\mathrm{P}-\text{q.c.})$ *. Risulta:*

- *Se* $X_n \to Y$ *(P-q.c.), allora* $X = Y$ *(P-q.c.);*
- *Se* $g : \mathbb{R} \to \mathbb{R}$ *è una funzione continua, allora* $g(X_n) \to g(X)$ *(P-q.c.).*

5.2 Convergenza in probabilità

Def 5.2. Una successione $(X_n)_{n\geq 1}$ di v.a. converge in probabilità alla v.a. X (in simboli $X_n \to^{\text{pr}} X$) se per ogni $\epsilon > 0$ risulta

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

oppure

$$\lim_{n\to+\infty} \mathrm{P}(|X_n - X| \leq \epsilon) = 1$$

5.2.1 Teor cqc implica cip

Thm 5.3 (dim)**.** *La convergenza quasi certa implica quella in probabilità.*

5.2.2 Teor di caratterizzazione cip

Thm 5.4 (dim)**.** *Le seguenti proposizioni sono equivalenti:*

i) $X_n \to^{\text{pr}} X$*;*

ii) Ogni sottosuccessione di $(X_n)_{n\geq 1}$ *ammette una sottosuccessione convergente* P*-quasi certamente a* X*.*

5.2.3 Teor unicità q.c. del limite e convergenza in probabilità della trasformata continua

Thm 5.5 (dim)**.** *Sia* $X_n \overset{Pr}{\to} X$*. Risulta:*

(i) Se $X_n \overset{Pr}{\to} Y$*, allora* $X = Y$ *(P-q.c.);*

(ii) Se $g : \mathbb{R} \to \mathbb{R}$ *è una funzione continua, allora* $g(X_n) \overset{Pr}{\to} g(X)$ *.*

5.3 Leggi dei grandi numeri

Le leggi dei grandi numeri forniscono una formulazione matematica precisa all'intuizione empirica che:

- la media aritmetica 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 probabilità.

5.3.1 Somma di v.a.

Def 5.3. Date le v.a.$X_1, \dots, X_n$, poniamo

$$S_n = X_1 + \cdots + X_n$$

Se $\mathrm{E}(X_1), \dots, \mathrm{E}(X_n)$ sono finite, allora

$$\mathrm{E}\left(\frac{S_n}{n}\right) = \frac{\mathrm{E}(X_1) + \cdots + \mathrm{E}(X_n)}{n}$$

5.3.2 Somma di v.a. indicatrici

Def 5.4. Dati gli eventi $A_1, \dots, A_n$, poniamo

$$S_n = I_{A_1} + \cdots + I_{A_n}$$

Allora

$$\mathrm{E}\left(\frac{S_n}{n}\right) = \frac{\mathrm{P}(A_1) + \cdots + \mathrm{P}(A_n)}{n}$$

5.4 Legge debole dei grandi numeri

Def 5.5. Una successione $(X_n)_{n\geq 1}$ di v.a. verifica la **legge debole dei grandi numeri** se

$$\frac{S_n}{n} - \mathrm{E}\left(\frac{S_n}{n}\right) \to 0$$

5.4.1 Teor di Markov (cond. suff.)

Thm 5.6. ***Teorema di Markov*** *(dim). Sia la successione di* v.a. $(X_n)_{n\geq 1}$ *tale che:*

$$\lim_{n\to+\infty} \frac{\mathrm{Var}(S_n)}{n^2} = 0 \quad (condizione\ \ di\ \ Markov).$$

Allora, la successione verifica la legge debole dei grandi numeri.

5.4.2 Teor di Cebicev (altra condiz suff)

Thm 5.7 (dim)**.** *Sia la successione di* v.a.$(X_n)_{n\geq 1}$ *tale che* $\mathrm{Cov}(X_i, X_j) \leq 0$ $(i \neq j)$ *e*

$$\lim_{n\to+\infty} \frac{\mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n)}{n^2} = 0. \qquad (3)$$

Allora $(X_n)_{n\geq 1}$ *verifica la legge debole dei grandi numeri.*

5.4.3 Teor su condizione suff per altra condiz suff

Osservazione 5.8. *L'ipotesi (3) sussiste se è verificata una delle due condizioni:*

- $\lim_{n\to+\infty} \frac{\mathrm{Var}(X_n)}{n} = 0$;
- $\sum_{n\geq 1} \frac{\mathrm{Var}(X_n)}{n^2} < +\infty$

Dunque sussiste la legge debole dei grandi numeri, per successioni di v.a. non positivamente correlate, se le varianze $\sigma_n^2 = \mathrm{Var}(X_n)(n \geq 1)$ sono equilimitate oppure non divergono troppo velocemente.

5.5 Legge forte dei grandi numeri

Def 5.6. Una successione $(X_x)_{n\geq 1}$ di v.a. verifica la **legge forte dei grandi numeri** se

$$\frac{S_n}{n} - \mathrm{E}\left(\frac{S_n}{n}\right) \to 0 \; (P - q.c.)$$

5.5.1 Teorema di Rajchman

Thm 5.9. ***Teorema di Rajchman****. Sia* $\mathrm{Cov}(X_i, X_j) \leq 0 \, (i \neq j)$ *e le varianze* $\mathrm{Var}(\mathrm{X}_n)(n \geq 1)$ *equilimitate. Allora,* $(X_n)_{n\geq 1}$ *verifica la legge forte dei grandi numeri*

5.5.2 Corollario teor di Rajchman

Corollario. *Risulta:*

i) [dim] Siano $X_1, X_2, \ldots$ *indipendenti, equidistribuite* e *con varianza finita. Allora* $\frac{S_n}{n} \to \mathrm{E(X)}$ *(P-q.c.)*

ii) (Borel) Siano gli eventi $A_1, A_2, \ldots$ *indipendenti ed equiprobabili. Allora* $\frac{S_n}{n} \to \mathrm{P(A)}$ *(P-q.c.)*

5.6 Convergenza in distribuzione

La convergenza puntuale di una successione di v.a. ad una v.a. non assicura la convergenza delle rispettive funzioni di ripartizione alla funzione di ripartizione della v.a. limite.

Def 5.7. Una successione $(X_n)_{n\geq 1}$ di v.a. **converge in distribuzione** alla v.a. X (in simboli $X_n \overset{di}{\to} X$) se $F_n(x) \to F(x)$ in ogni punto di continuità x di F avendo indicato, rispettivamente, con F ed F_n le funzioni di ripartizione delle v.a. X e X_n $(n \geq 1)$.

Da un punto di vista intuitivo, la differenza sostanziale tra 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.

5.6.1 Teor link fra conv. in prob e conv. in distrib

Thm 5.10 (dim). *La convergenza in probabilità implica quella in distribuzione.*

5.6.2 Teor di caratterizzazione cid (Portmanteau)

Thm 5.11 (Portmanteau Theorem). *Le seguenti proposizioni sono equivalenti:*

- $X_n \to^{\mathrm{di}} X$*;*
- $\lim_{n\to+\infty} \mathrm{E}(f(X_n)) = \mathrm{E}(f(X))$*, per ogni funzione* $f : \mathbb{R} \to \mathbb{R}$ *continua* e *limitata.*

5.6.3 Teor unicità in distribuzione e convergenza in distribuzione della trasformata continua

(Ocio) Unicità non q.c., ma in legge

Thm 5.12 (dim). *Sia* $X_n \overset{di}{\to} X$*. Sussistono allora le seguenti proposizioni:*

(i) Se $X_n \overset{di}{\to} Y$*, allora* X, Y *sono equidistribuiti (cioè* $\mathrm{P}_X = \mathrm{P}_Y$*);*

(ii) Se $g : \mathbb{R} \to \mathbb{R}$ *è una funzione continua, allora* $g(X_n) \overset{di}{\to} g(X)$ *.*

Osservazione. La convergenza in distribuzione non assicura la convergenza del momento centrale m-simo delle v.a. della successione al momento centrale m-simo della v.a. limite. Ciò avviene solo se la v.a. è limitata

5.6.4 Teor di Polya

Thm 5.13. ***Teorema di Polya*** *Se la successione di funz. rip.* $(F_n)_{n\geq 1}$*, converge ad una funz. rip.* F *continua, allora la convergenza è uniforme.*

5.6.5 Somma ridotta

Date le v.a. $X_1, \ldots, X_n$ con speranza matematica α e varianza $\sigma^2 > 0$ finita, si chiama somma ridotta di $X_1, \ldots, X_n$ la seguente trasformata della v.a. S_n:

$$S_n^* = \frac{S_n - \alpha n}{\sqrt{\sigma^2 n}}.$$

Chiaramente, $\mathrm{E}(S_n^*) = 0$. Inoltre, se $X_1, \ldots, X_n$ sono indipendenti, si ha $\mathrm{Var}\,(S_n^*) = 1$.

5.7 Teor limite centrale

Thm 5.14 (Teorema limite centrale). *Siano le v.a.* $X_1, X_2, \ldots$ *indipendenti ed equidistribuite con varianza finita non nulla. Allora, la successione* $(F_{S_n^*})_{n\geq 1}$ *delle funz. rip. delle somme ridotte converge uniformemente alla funz. rip. della distribuzione normale* $\mathrm{N}(0,1)$.

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

5.7.1 Teor disuguaglianza di Berry-Essén

Nelle ipotesi del teorema limite centrale assumendo la finitezza e positività del momento centrale assoluto terzo $\beta = E(|X_1 - E(X_1)|^3)$, risulta che il massimo errore di approssimazione fra la funzione di ripartizione delle somme ridotte rispetto alla funzione di ripartizione della normale è dato da $\sup_{x\in\mathbb{R}} |F_{S_n^*}(x) - \Phi(x)| \leq \frac{K\beta}{\sqrt{(\sigma^2)^3}} \frac{1}{\sqrt{n}}$(**disuguaglianza di Berry-Essén**)
con$\frac{\sqrt{10}+3}{6\sqrt{2\pi}} \approx 0.409732 \leq K \leq 0.4785$ (allo stato attuale della conoscenza)

5.7.2 Approssimazione normale

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

$$\mathrm{P}(S_n \leq x) = \mathrm{P}(S_n^* \leq \frac{x-\alpha n}{\sqrt{\sigma^2 n}}) \approx \Phi(\frac{x-\alpha n}{\sqrt{\sigma^2 n}}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\alpha n}{\sqrt{\sigma^2 n}}} \mathrm{e}^{-\frac{t^2}{2}} dt.$$

dove $E(X_n) = \alpha$ e $Var(X_n) = \sigma^2 < +\infty \ \forall n$

L'esperienza empirica suggerisce che la soglia di applicabilità di questa "approssimazione normale"sia per n tra 30 e 50 per la maggior parte delle distribuzioni impiegate nella pratica che non siano troppo asimmetriche (valori che, è bene tenere presente, non sono nè suggeriti, nè giustificati da alcun risultato teorico). Nel caso di distribuzioni molto asimmetriche i valori 30, 50 devono essere aumentati. Ad esempio, se consideriamo una successione di eventi indipendenti di medesima probabilità p, una buona approssimazione la si ottiene prendendo $p(1-p)n \geq 5$. Osserviamo inoltre che, nel caso di v.a. a valori interi, si ottiene, in generale, una migliore approssimazione di $\mathrm{P}(S_n \leq x)$ sostituendo in $\frac{x-\alpha n}{\sqrt{\sigma^2 n}}$ il valore x con $x + \frac{1}{2}$.

Sempre in questo ambito, una ulteriore approssimazione normale di impiego usuale nella pratica è la seguente:

$$\mathrm{P}\left(\left|\frac{S_n}{n}-\alpha\right|\geq\epsilon\right)\approx 2\Phi\left(-\epsilon\sqrt{\frac{n}{\sigma^2}}\right)=\sqrt{\frac{2}{\pi}}\int_{\epsilon\sqrt{\frac{n}{\sigma^2}}}^{+\infty}\mathrm{e}^{-\frac{t^2}{2}}\,dt$$

che fornisce una approssimazione, per n sufficientemente grande, della probabilità che la media aritmetica delle prime n v.a. si discosti dalla comune speranza matematica α per almeno di $\epsilon > 0$

5.8 Teorema di Glivenko-Cantelli

Le leggi dei grandi numeri e il teorema limite centrale considerano il comportamento asintotico, rispettivamente, della media aritmetica e della somma ridotta.
Il prossimo teorema fornisce invece una formulazione matematica precisa all'intuizione empirica che la funzione di ripartizione che distribuisce la probabilità in modo uguale su un gran numero di realizzazioni di v.a. indipendenti ed equidistribuite è verosimilmente prossima alla comune funzione di ripartizione.
Date le v.a. $X_1,\ \ldots,\ X_n$, la funzione di dominio $\mathbb{R}\times\Omega$, con ω fissato e x che varia:

$$\begin{aligned}F^{(X_1,\ldots,X_n)}(x,\ \omega) &= \frac{I_{\{X_1\leq x\}}(\omega)+\cdots+I_{\{X_n\leq x\}}(\omega)}{n}\\ &= \frac{I_{]-\infty,x]}(X_1(\omega))+\cdots+I_{]-\infty,x]}(X_n(\omega))}{n},\end{aligned}$$

è detta **funzione di ripartizione empirica** (relativa a $X_1,\ldots,X_n$).
-$F^{(X_1,\ldots,X_n)}(\cdot,\omega)$ è una funzione di ripartizione per ogni $\omega\in\Omega$ e quindi, dipendendo dai casi elementari, è una funzione di ripartizione aleatoria; risulta infatti:

-$F^{(X_1,\ldots,X_n)}(x,\omega)\leq F^{(X_1,\ldots,X_n)}(x',\omega)$, se $x<x'$ (poichè $I_{]-\infty,x]}\leq I_{]-\infty,x']}$);

-$F^{(X_1,\ldots,X_n)}(x,\omega)\leq F^{(X_1,\ldots,X_n)}(x',\omega)$, se $x<x'$ (poichè $I_{]-\infty,x]}\leq I_{]-\infty,x']}$);

-$F^{(X_1,\ldots,X_n)}(x,\omega)=0$, se $x<\min(X_1(\omega),\ \ldots,\ X_n(\omega))$;

-$F^{(X_1,\ldots,X_n)}(x,\omega)=1$, se $x\geq\max(X_1(\omega),\ \ldots,\ X_n(\omega))$;

-$F^{(X_1,\ldots,X_n)}(x^+,\omega)=F^{(X_1,\ldots,X_n)}(x,\ \omega)$ (in quanto $\lim\limits_{t\to x^+} I_{\{X_i\leq t\}}=I_{\{X_i\leq x\}}$).

Thm 5.15. *Teorema di Glivenko-Cantelli*
Siano le v.a. $X_1, X_2, \ldots$ *indipendenti con comune funzione di ripartizione* F*. Allora*

$$\sup_{x \in \mathbb{R}} \left| F^{(X_1,\ldots,X_n)}(x,\ \cdot) - F(x) \right| \to 0 \ \ (P - q, \text{c.}),$$

cioé $\mathrm{P}(\{\omega \in \Omega$ *:* $F^{X_1,\ldots,X_n)}(\cdot, \omega)$ *converge uniformemente a* $F\}) = 1$

6 MEDIE CONDIZIONATE

6.1 Condizionamento a eventi non trascurabili

Def 6.1. Supposto $P(H) > 0$ poniamo

$$\mathrm{P}(A|H) = \frac{P(A \cap H)}{P(H)} \quad \forall A \in \mathcal{A}$$

Thm 6.1. *Si dimostra che essa è ancora una probabilità.*

Allora $\mathrm{P}(\cdot|H)$ è una probablità, detta **probabilità condizionata** a H.

Thm 6.2 (dim)**.** *Sia X una* v.a.. *Allora, X è $\mathrm{P}(\cdot|H)$-sommabile se* e *solo se XI_H è* P*-sommabile. Inoltre, nel caso di sommabilità,*

$$\int_\Omega X d\mathrm{P}(\cdot|H) = \frac{\displaystyle\int_H X dP}{\mathrm{P}(H)}$$

Def 6.2. Data una v.a. X con speranza matematica, la **speranza matematica condizionata di X a H** viene identificata con l'elemento della retta reale ampliata:

$$\mathrm{E}(X|H) = \int_\Omega X d\mathrm{P}(\cdot|H)$$

Allora,

$$\mathrm{E}(X|H) = \frac{\displaystyle\int_H X dP}{\mathrm{P}(H)} = \frac{\mathrm{E}(XI_H)}{\mathrm{P}(H)}$$

6.2 Condizionamento a sigma-algebre

Def 6.3. Definisco alcune quantità utili nel seguito:

- **Informazione** è una v.a. X non osservabile
- **Informazione ottenibile** $\mathcal{H}$ è la σ-algebra generata dagli **eventi osservabili** (es. $\sigma(\mathcal{H})$)
- una v.a. **osservabile** è costruibile attraverso indicatori ed elementi di $\mathcal{H}$ (es. $Y = \sum_{i=1}^4 y_i I_{H_i}$)

$\triangleright Y$ è $\mathcal{H}$-Borel misurabile
Dunque la v.a. Y consente di calcolare la speranza matematica condizionata di X a un qualsiasi evento osservabile non trascurabile.

Appare quindi ragionevole ritenere che l'assunzione congiunta della $\mathcal{H}$-Borel misurabilità e della $\displaystyle\int_H X dP = \int_H Y dP$ 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
Adottando questo punto di vista e passando a una formulazione generale, consideriamo, per descrivere formalmente *l'informazione ottenibile* 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 acquisibile tramite il processo di osservazione).

L'assunzione che la famiglia degli eventi osservabili $\mathcal{H}$ sia una σ-algebra appare del tutto naturale una volta osservato che l'evento certo è osservabile, che la negazione di un evento osservabile è osservabile e che la disgiunzione di eventi osservabili è osservabile.

6.2.1 Versione della speranza matematica condizionata di X

Def 6.4. Data una v.a. X con $\mathrm{E}(X)$ finito, chiamiamo **versione (della speranza matematica condizionata di X a $\mathcal{H}$)** (in simboli $\mathrm{E}(X|\mathcal{H})$) ogni v.a. estesa Y tale che:

(a) sia $\mathcal{H}$-Borel misurabile;

(b) per ogni $H \in \mathcal{H}$, risulti

$$\int_H X d\mathrm{P} = \int_H Y d\mathrm{P}.$$

- Y è una v.a. osservabile. Pertanto, una volta risolta l'incertezza sugli eventi osservabili viene individuato il "vero" valore di Y.

Thm 6.3. *Y è finita (P-q.c.) in quanto* $\mathrm{E}(Y) = \mathrm{E}(X)$.

Thm 6.4 (dim)**.** *Sia Y' un'altra versione. Allora $Y = Y'$ (P-q.c.).*

Dunque la speranza matematica condizionata è definita a meno di eventi osservabili di probabilità nulla.

Perciò con un **abuso di notazione**, useremo il simbolo $\mathrm{E}(X|\mathcal{H})$ per denotare anche una sua qualsiasi versione.

Def 6.5. Un evento osservabile H si chiama P-**atomo** (in breve **atomo**) se è un evento di probabilità positiva non ripartibile in due eventi osservabili ancora di probabilità positiva (in altri termini, per ogni $H_1 \subset H$ si ha $\mathrm{P}(H_1) = 0$ o $\mathrm{P}(H_1) = \mathrm{P}(H)$) .

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

Thm 6.5. *Sia H un* P*-atomo di $\mathcal{H}$. Esiste allora $H_1 \subseteq H$ tale che* $\mathrm{P}(H_1) = \mathrm{P}(H)$ e $\mathrm{E}(X|\mathcal{H})(\omega) = \mathrm{E}(X|H)$ *per ogni $\omega \in H_1$.*

• L'ipotesi di atomicità non è in generale rimovibile.

6.2.2 Proprietà della speranza matematica condizionata

Thm 6.6. *Risulta:*

- $\mathrm{E}(\mathrm{E}(X|\mathcal{H})) = \mathrm{E}(X)$
- $\mathrm{E}(X|\{\emptyset,\ \Omega\}) = \mathrm{E}(X)$. *Qualora l'informazione ottenibile sia quella: minima (costituita solamente dagli eventi certo* e *impossibile) (**informazione nulla**), di X possiamo conoscere unicamente la speranza matematica;*
- $\mathrm{E}(X|\mathcal{A}) = X$ *(P-q.c.). Qualora l'informazione ottenibile sia quella: massima (consistente nel sapere quale caso elementare è vero), di X possiamo conoscere il suo "vero" valore.*
- $\mathrm{E}(X|\mathcal{H}) = X$ *(P-q.c.), se X è $\mathcal{H}$-Borel misurabile*

Thm 6.7. *Siano X, Y, X_1, X_2, etc.* v.a. *con speranza matematica finita. Valgono le proposizioni seguenti nelle quali le relazioni riguardanti speranze matematiche condizionate devono intendersi sussistere* P*-quasi certamente.*

- MONOTONIA : $\mathrm{E}(X|\mathcal{H}) \leq \mathrm{E}(Y|\mathcal{H})$, *se* $X \leq Y$ $(\mathrm{P}-\text{q.c.})$;
- $\mathrm{E}(X|\mathcal{H}) = \mathrm{E}(Y|\mathcal{H})$, *se $X = Y$ (P-q.c.);*
- $\mathrm{E}(X|\mathcal{H}) = \mathrm{E}(Y|\mathcal{H})$, *se $X = Y$ (P-q.c.);*
- INTERNALITÀ*: $a \leq \mathrm{E}(X|\mathcal{H}) \leq b$, se $a \leq X \leq b$ (P-q.c.);*
- INTERNALITÀ STRETTA*: $a < \mathrm{E}(X|\mathcal{H}) < b$, se $a < X < b$ (P-q.c.);*
- LINEARITÀ : $\mathrm{E}(\sum_{i=1}^{n} \alpha_i X_i|\mathcal{H}) = \sum_{i=1}^{n} \alpha_i \mathrm{E}(X_i|\mathcal{H})$ *qualunque siano* i *numeri reali $\alpha_1, \ldots, \alpha_n$;*
- $|\mathrm{E}(X|\mathcal{H})| \leq \mathrm{E}(|X||\mathcal{H})$;
- CONVERGENZA MONOTONA : $\mathrm{E}(X_n|\mathcal{H}) \uparrow \mathrm{E}(X|\mathcal{H})$, *se* $X_n \uparrow X$ e $X_1 \geq 0$ $(\mathrm{P}-\text{q.c.})$;
- $\mathrm{E}(\sum_{n\geq 1} \alpha_n X_n|\mathcal{H}) = \sum_{n\geq 1} \alpha_n \mathrm{E}(X_n|\mathcal{H})$, *se $(\alpha_n)_{n\geq 1}$ è una successione di numeri reali non negativi, $X_n \geq 0$ per ogni n* e *la speranza matematica della serie $\sum_{n\geq 1} \alpha_n X_n$ è finita;*

- CONVERGENZA DOMINATA : $\mathrm{E}(X_n|\mathcal{H}) \to \mathrm{E}(X|\mathcal{H})$, *se* $|X_n| \leq Z$ *(P-q.c.) per ogni* $n, \mathrm{E}(Z)$ *finita* e $X_n \to X$*;*
- $\mathrm{E}(ZX|\mathcal{H}) = Z\mathrm{E}(X|\mathcal{H})$, *se* Z *è* $\mathcal{H}$*-Borel misurabile* e *limitata; assicura che le* v.a. *osservabili* e *limitate possono essere "estratte" dall'operatore di speranza matematica condizionata.*
- *Se* $\mathcal{K}$ *è una sotto* σ*-algebra di* $\mathcal{H}$*, allora* $\mathrm{E}(\mathrm{E}(X|\mathcal{H})|\mathcal{K}) = \mathrm{E}(X|\mathcal{K})$ *e* $\mathrm{E}(\mathrm{E}(X|\mathcal{K})|\mathcal{H}) = \mathrm{E}(X|\mathcal{K})$.

6.2.3 Teorema di Lévy

Thm 6.8 (**Teorema di Lévy**)**.** *Sia* $(\mathcal{H}_n)_{n\geq 1}$ *una successione non decrescente di* σ*-algebre tali che*

$$\mathcal{A} = \sigma(\bigcup_{n\geq 1} \mathcal{H}_n)$$

*(**struttura informativa**). Allora, per ogni* v.a. X *con speranza matematica finita risulta*

$$\mathrm{E}(X|\mathcal{H}_n) \to X \ (\mathrm{P-q.c.})$$

Thm 6.9 (dim)**.** *Siano* X *una* v.a. *con varianza finita e* Y *una versione limitata di* $\mathrm{E}(X|\mathcal{H})$*. Allora*

$$\mathrm{Var}(X) = \mathrm{Var}(Y) + \mathrm{E}((X-Y)^2) \geq \mathrm{Var}(Y)$$

6.3 Funzione di regressione

In moltissime situazioni l'informazione ottenibile consiste nell'osservare il valore di un ente aleatorio $X : (\Omega,\ \mathcal{A}) \to (\mathfrak{X},\ \aleph)$, ove $\aleph$ contiene i singoletti. Poichè l'informazione deriva dall'osservazione di X, viene naturale identificare la σ-algebra degli eventi osservabili $\mathcal{H}$ con la σ-algebra indotta

$$X^{-1}(\aleph) = \{X^{-1}(\mathrm{A}) : \ \mathrm{A}\ \in \aleph\} \subseteq \mathcal{A}$$

da X su Ω (da un punto di vista interpretativo, la conoscenza dei valori di verità di tutti gli eventi di tale σ-algebra equivale alla conoscenza del vero valore di X in quanto $\{X = y\} = \{X \in \{y\}\} \in X^{-1}(\aleph)$ per ogni $y \in \mathfrak{X}$).

Lemma 6.10 (dim solo caso Z semplice)**.** *Sia* Z *una* v.a. $X^{-1}(\aleph)$*-Borel misurabile. Esiste allora una funzione* $\aleph$*-Borel misurabile* $g : \mathrm{X} \mapsto \mathbb{R}^*$ *tale che* $Z = g(X)$.

Data una v.a. Y con speranza matematica finita,

$$\mathrm{E}(Y|X) = \mathrm{E}(Y|X^{-1}(\aleph))$$

è la speranza matematica condizionata di Yall'osservazione di X.

6.3.1 Indipendenza fra v.a. Y ed e.a. X

Thm 6.11 *(dim $\star\star$)*. *Se la* v.a. Y e *l'ente aleatorio* X *sono* <u>*indipendenti,*</u> *allora*

$$\mathrm{E}(Y|X) = \mathrm{E}(Y) = \mathrm{E}(Y|\{\emptyset,\ \Omega\})$$

cioè l'informazione proveniente da X *equivale all'informazione nulla.*

Def 6.6. Chiamiamo **funzione di regressione di** Y **su** X ogni funzione

$$\mathrm{E}(Y|X=\cdot) : \mathfrak{X} \mapsto \mathbb{R}^*$$

tale che:

1. sia P_X-integrabile;
2. per ogni $\mathrm{A} \in \aleph$ risulti:

$$\int_{\{X\in \mathrm{A}\}} Y d\mathrm{P} = \int_{\mathrm{A}} \mathrm{E}(Y|X=x)\mathrm{P}_X(dx)$$

- è finita (P_X − q.c.)
- è definita a meno di insiemi P_X-trascurabili
- *(dim)* $\mathrm{E}(Y|X=X(\cdot))$ è una versione di $\mathrm{E}(Y|X)$.

Thm 6.12. *Sia* $x_0 \in \mathrm{X}$ *tale che* $\mathrm{P}(X=x_0) > 0$. *Allora,*

$$\mathrm{E}(Y|X=x_0) = \frac{\displaystyle\int_{\{X=x_0\}} Y d\mathrm{P}}{\mathrm{P}(X=x_0)} = \mathrm{E}(Y|\{X=x_0\}).$$

Si noti la la profonda ***differenza*** *concettuale delle due nozioni di* ***funzione di regressione*** e *di* ***speranza matematica condizionata all'informazione disponibile****: funzione dipendente dai valori osservabili la prima; dai casi elementari la seconda.*

6.3.2 Proprietà della funzione di regressione

Thm 6.13. *Siano* Y, Y_1, Y_2 *etc. v.a. con speranza matematica finita. Riesce allora:*

- $\mathrm{E}(Y) = \displaystyle\int_{\mathfrak{X}} \mathrm{E}(Y|X=x)\mathrm{P}_X(dx)$.
  Valgono inoltre le proposizioni seguenti nelle quali le relazioni riguardanti funzioni di regressione devono intendersi sussistere P_X*-quasi certamente.*

- *Se* $X = x_0 \in \mathfrak{X}$ *(P-q.c.), allora* $\mathrm{E}(Y|X=\cdot) = \mathrm{E}(Y)$*. Inoltre, se* X *è l'applicazione identica, allora* $\mathrm{E}(Y|X=\cdot) = Y$*;*
- $\mathrm{E}(\alpha I_\Omega|X=\cdot) = \alpha$*;*
- MONOTONIA*:* $\mathrm{E}(Y_1|X=\cdot) \leq \mathrm{E}(Y_2|X=\cdot)$ *, se* $Y_1 \leq Y_2$ *(P-q.c.);*
- $\mathrm{E}(Y_1|X=\cdot) = \mathrm{E}(Y_2|X=\cdot)$ *, se* $Y_1 = Y_2$ *(P-q.c.);*
- INTERNALITÀ*:* $a \leq \mathrm{E}(Y|X=\cdot) \leq b$*, se* $a \leq Y \leq b$ *(P-q.c.);*
- INTERNALITÀ STRETTA*:* $a < \mathrm{E}(Y|X=\cdot) < b$*, se* $a < Y < b$ *(P- q.c.);*
- LINEARITÀ*:* $\mathrm{E}(\sum_{i=1}^{n} \alpha_i Y_i|X=\cdot) = \sum_{i=1}^{n} \alpha_i \mathrm{E}(Y_i|X=\cdot)$ *qualunque siano* i *numeri reali* $\alpha_1, \ldots, \alpha_n$*;*
- $|\mathrm{E}(Y|X=\cdot)| \leq \mathrm{E}(|Y||X=\cdot)$*;*
- CONVERGENZA MONOTONA*:* $\mathrm{E}(Y_n|X=\cdot) \uparrow \mathrm{E}(Y|X=\cdot)$*, se* $Y_n \uparrow Y$ e $Y_1 \geq 0$ *(P-q.c.);*
- $\mathrm{E}(\sum_{n\geq 1} \alpha_n Y_n|X=\cdot) = \sum_{n\geq 1} \alpha_n \mathrm{E}(Y_n|X=\cdot)$ *, se* $(\alpha_n)_{n\geq 1}$ *è una successione di numeri reali non negativi,* $Y_n \geq 0$ *per ogni* n e *la speranza matematica della serie* $\sum_{n\geq 1} \alpha_n Y_n$ *è finita;*
- CONVERGENZA DOMINATA*:* $\mathrm{E}(Y_n|X=\cdot) \to \mathrm{E}(Y|X=\cdot)$*, se* $|Y_n| \leq Z$ *(P-q.c.) per ogni* n, $\mathrm{E}(Z)$ *finita* e $Y_n \to Y$*;*
- *Sia* $g : \mathrm{X} \mapsto \mathbb{R}^*$una *funzione* $\aleph$*-Borel misurabile tale che la v.a.* $g(X)Y$ *ammetta speranza matematica finita. Allora,*

  $$\mathrm{E}(g(X)Y|X=\cdot) = g(\cdot)\mathrm{E}(Y|X=\cdot) \ .$$

  In particolare, $\mathrm{E}(g(X)|X=\cdot) = g(\cdot)$ *;*

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. Notato che, data una funzione continua e crescente φ tale che $\varphi(0) = 0$ e $\lim_{x\to+\infty} \varphi(x) = +\infty, \mathrm{E}(\varphi(|Y - g(X)|))$ è una quantità che tende ad essere grande, se $g(X)$ assume valori distanti da Y, e piccola, se $g(X)$ assume valori vicini a Y, viene naturale usarla per misurare la bontà dell'approssimazione di Y con $g(X)$.

Tra le varie scelte possibili di φ, adottiamo quella relativa al **metodo dei minimi quadrati**, cioè poniamo $\varphi(x) = x^2$. Giungiamo così ad intendere la frase "errore più piccolo possibile" nel senso di "**errore quadratico medio più piccolo possibile**" e quindi a cercare una funzione g^* tale che

$$\mathrm{E}((Y - g^*(X))^2) \leq \mathrm{E}((Y - g(X))^2)$$

per ogni funzione g a quadrato P_X-integrabile.

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

Thm 6.14. *Sia Y una v.a. a quadrato integrabile. Allora,* $\mathrm{E}(Y|X = X(\cdot))$ *è a quadrato integrabile e risulta*

$$\begin{aligned}
\mathrm{E}((Y - \mathrm{E}(Y|X))^2) &= \mathrm{E}((Y - \mathrm{E}(Y|X = X(\cdot))^2) \\
&= \min\{\mathrm{E}((Y - g(X))^2) : \int_{\mathrm{X}} g^2 d\mathrm{P}_X < +\infty\}
\end{aligned}$$

6.4 Funzione di regressione lineare

Se, in particolare, $g(x) = ax + b$ $(a > 0)$ e X, Y v.a. a quadrato integrabile e Var(X) > 0 finita, si ha

$$\begin{aligned}
f(a, b) = \mathrm{E}((Y - (aX + b)^2) &= \mathrm{E}(Y^2 - 2(aX + b)Y + (aX + b)^2) \\
&= \mathrm{E}(Y^2) - 2\mathrm{E}(aXY + bY) + \mathrm{E}(a^2X^2 + 2abX + b^2) \\
&= \mathrm{E}(Y^2) - 2\mathrm{E}(aXY + bY) + \mathrm{E}(a^2X^2 + 2abX + b^2) \\
&= \mathrm{E}(Y^2) - 2a\mathrm{E}(XY) - 2b\mathrm{E}(Y) + a^2\mathrm{E}(X^2) + 2ab\mathrm{E}(X) + b^2
\end{aligned}$$

da cui otteniamo che il punto di minimo $(a^\star, b^\star)$ è soluzione del sistema

$$\begin{cases}
\dfrac{\partial}{\partial a} f(a,\ b) = -2\mathrm{E}(XY) + 2a\mathrm{E}(X^2) + 2b\mathrm{E}(X) = 0 \\[2ex]
\dfrac{\partial}{\partial b} f(a,\ b) = -2\mathrm{E}(Y) + 2a\mathrm{E}(X) + 2b = 0
\end{cases}$$

e quindi

$$a^\star = \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(X)} \qquad b^\star = \mathrm{E}(Y) - \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(X)}\mathrm{E}(X).$$

otteniamo così l'espressione della trasformata affine dell'osservabile X che meglio approssima in media quadratica il non osservabile Y

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

detta **retta di regressione lineare** di Y su X.