

Calcolo delle Probabilità

Appunti del corso 2025/26

DAVID BARBATO
`BARBATO@MATH.UNIPD.IT`

DIPARTIMENTO DI MATEMATICA
UNIVERSITÀ DEGLI STUDI DI PADOVA
VIA TRIESTE 63, 35121 PADOVA, ITALY

Versione 1.0

Ultima modifica: 23 ottobre 2025.

Segnalazioni di errori, osservazioni, suggerimenti e critiche sono molto graditi.

Indice

1	Gli spazi misurabili	7
1.1	σ -algebra e spazi misurabili	7
1.1.1	σ -algebra ed informazione	12
1.1.2	d-system e lemma di Dynkin	14
1.1.3	Teorema della classe monotona $\dagger$	16
1.1.4	Prodotto di σ -algre	16
2	Gli spazi di misura	19
2.1	Introduzione agli spazi di misura	19
2.1.1	Funzioni additive e σ -additive	19
2.1.2	Definizione di misura	19
2.1.3	Quasi sempre, quasi ovunque, quasi certamente, quasi... 20	
2.2	Un teorema di unicit� della misura	21
2.3	Un teorema di esistenza della misura	22
2.4	Propriet� degli spazi di probabilit�	22
2.4.1	Legge di una variabile aleatoria	26
2.4.2	Funzione di ripartizione	26
3	Integrazione	29
3.1	Valore atteso di una variabile aleatoria	29
3.2	Teoremi limiti	31
3.3	Variabili aleatorie discrete e assolutamente continue	33
3.4	Disuguaglianza di Markov	35
3.5	Disuguaglianza di Jensen	36
3.6	Spazi L^p	38
3.6.1	Monotonia degli spazi L^p	39
3.6.2	Disuguaglianze di Holder e Minkowski	39
3.6.3	Completezza	41
4	Indipendenza	43
4.1	Probabilit� condizionata	43
4.2	Indipendenza tra eventi	44
4.2.1	Indipendenza di due eventi	44

4.2.2	Indipendenza di una famiglia generica di eventi	45
4.3	Indipendenza di σ -algebre	46
4.4	Indipendenza di variabili aleatorie	46
4.5	Indipendenza e p-system	47
4.6	Indipendenza a blocchi	49
4.7	Varianza, covarianza ed indipendenza	49
4.8	Secondo Lemma di Borel-Cantelli	51
4.8.1	Secondo lemma di Borel-Cantelli per variabili non correlate	52
4.9	σ -algebra coda e legge 0-1 di Kolmogorov	53
4.10	Unicit� della distribuzione congiunta di v.a. indipendenti . .	54
4.11	Cifre decimali di un numero a caso in $(0, 1)$	55
A	Esercizi	57
A.1	Capitolo 1	57
A.1.1	Esercizi teorici	57
A.1.2	Esercizi di comprensione	59
A.2	Capitolo 2	61
A.2.1	Esercizi teorici	61
A.2.2	Esercizi di comprensione	61
A.3	Capitolo 3	64
A.3.1	Esercizi teorici	64
A.4	Capitolo 4	65
A.4.1	Esercizi teorici	65
A.4.2	Esercizi di comprensione	65
B	Esistenza della misura di Lebesgue	67

Introduzione

Capitolo 1

Gli spazi misurabili

In questo capitolo introdurremo le σ -algebre e gli spazi misurabili. Tale introduzione è necessaria anche per chi ha già incontrato gli spazi misurabili in un corso di teoria della misura. Uno dei punti chiave del capitolo sarà la comprensione del legame che c'è tra σ -algebra e informazione. Lo strumento più adatto per comprendere tale legame è il criterio di misurabilità, teorema 1.8. Tale teorema di fatto afferma che una variabile Y è $\sigma(X)$ misurabile se e solo se può essere determinata a partire dal valore assunto dalla variabile X , ovvero la variabile aleatoria Y è $\sigma(X)$ misurabile se e solo se il suo valore può essere determinato dall'informazione di X .

1.1 σ -algebra e spazi misurabili

In questo paragrafo introdurremo la definizione di σ -algebra e di spazio misurabile. Per farci un'idea iniziale su cosa sia una σ -algebra possiamo pensare ad essa come alla famiglia dei sottoinsiemi di un certo insieme Ω ai quali vogliamo assegnare una probabilità. A differenza di quanto visto nel corso di probabilità del primo anno, assegneremo una probabilità solo ad una famiglia di sottoinsiemi di Ω e non a tutti. Ci sono due valide ragioni per fare ciò: la prima è che non sempre è possibile farlo (vedi insieme di Vitali per la misura di Lebesgue); la seconda è che non sempre ci conviene farlo (questo concetto sarà evidente alla fine del corso quando studieremo filtrazioni e martingale).

Affinché una σ -algebra sia compatibile con le proprietà della probabilità (o in generale di una misura) è necessario che soddisfi alcune condizioni, da ciò segue la definizione:

Definizione 1.1 (σ -algebra). Dato un insieme Ω , una σ -algebra $\mathcal{H}$ in Ω è una collezione di sottoinsiemi di Ω che soddisfa le seguenti proprietà.

- $\Omega \in \mathcal{H}$ e $\emptyset \in \mathcal{H}$
- Se $A \in \mathcal{H}$ allora $A^c := \Omega \setminus A \in \mathcal{H}$
- Per ogni $\forall A, B \in \mathcal{H}$ si ha $A \cup B \in \mathcal{H}$ e $A \cap B \in \mathcal{H}$
- $\forall \{A_n\}_{n \in \mathbb{N}}$ se $A_n \in \mathcal{H} \ \forall n \in \mathbb{N}$ allora $\bigcup_{n=1}^{\infty} A_n \in \mathcal{H}$ e $\bigcap_{n=1}^{\infty} A_n \in \mathcal{H}$

gli elementi di $\mathcal{H}$ sono detti insiemi misurabili.

La seconda condizione è detta chiusura rispetto al complementare, la terza invece chiusura rispetto all'unione ed intersezione finita, e la quarta chiusura rispetto all'unione numerabile e all'intersezione numerabile. Osserviamo anche che le ipotesi sono ridondanti per esempio la chiusura rispetto al complementare più la chiusura rispetto all'unione dà la chiusura rispetto all'intersezione. Una collezione di sottoinsiemi di Ω che soddisfa le prime 3 condizioni si dirà un'algebra. La proposizione che segue indica delle condizioni sufficienti affinché una collezione di sottoinsiemi sia una σ -algebra.

Proposizione 1.1. *Dato un insieme Ω e una collezione $\mathcal{H}$ di sottoinsiemi di Ω consideriamo le seguenti condizioni:*

- 1 $\Omega \in \mathcal{H}$
- 1' $\emptyset \in \mathcal{H}$
- 2 Se $A \in \mathcal{H}$ allora $A^c := \Omega \setminus A \in \mathcal{H}$
- 3 $\forall \{A_n\}_{n \in \mathbb{N}}$ se $A_n \in \mathcal{H} \ \forall n \in \mathbb{N}$ allora $\bigcup_{n=1}^{\infty} A_n \in \mathcal{H}$
- 3' $\forall \{A_n\}_{n \in \mathbb{N}}$ se $A_n \in \mathcal{H} \ \forall n \in \mathbb{N}$ allora $\bigcap_{n=1}^{\infty} A_n \in \mathcal{H}$

Se vale la condizione 2, una condizione tra la 1 e la 1' e una condizione tra la 3 e la 3' allora $\mathcal{H}$ è una σ -algebra.

Definizione 1.2. La coppia $(\Omega, \mathcal{H})$ con Ω insieme e $\mathcal{H}$ sua σ -algebra si dirà spazio misurabile.

Proprietà 1.2 (σ -algebra). *Dimostrare che se A e B sono elementi di una σ -algebra $\mathcal{H}$ allora anche $A \setminus B$ e $A \Delta B$ lo sono.*

Esempio 1.1 (σ -algre). *Dato un insieme non vuoto Ω esso ha almeno due σ -algre: $\mathcal{H}_1 := \{\Omega, \emptyset\}$ e $\mathcal{H}_2 := \mathcal{P}(\Omega)$. $\mathcal{H}_1$ e $\mathcal{H}_2$ sono rispettivamente la σ -algebra meno fine e la più fine di Ω .*

Esempio 1.2 (σ -algebre). Dato un insieme non vuoto Ω e una sua partizione $\{A_i\}_{i \in I}$. È possibile costruire una σ -algebra $\mathcal{H}$ in modo tale che sia la più fine tra le σ -algebre che non spezzano gli insiemi della partizione.

$$\mathcal{H} = \cup_{J \subseteq I} \{ \cup_{i \in J} A_i \}$$

Esercizio 1.1. Dimostrare che se una σ -algebra è finita allora la sua cardinalità è 2^n per qualche $n \in \mathbb{N}$.

Proprietà 1.3 (σ -algebra). Dimostrare le seguenti affermazioni.

- 1 Se $\mathcal{H}_1$ e $\mathcal{H}_2$ sono σ -algebre di Ω allora anche $\mathcal{H}_1 \cap \mathcal{H}_2$ lo è.
- 1' Se $\{\mathcal{H}_i\}_{i \in I}$ è una famiglia di σ -algebre di Ω allora anche $\cap_{i \in I} \mathcal{H}_i$ è una σ -algebra.
- 2 Non è vero che per ogni insieme Ω e per ogni $\mathcal{H}_1$ e $\mathcal{H}_2$, σ -algebre di Ω vale $\mathcal{H}_1 \cup \mathcal{H}_2$ è una σ -algebra di Ω .
- 2' Date $\mathcal{H}_1$ e $\mathcal{H}_2$, σ -algebre di Ω allora $\mathcal{H}_1 \cup \mathcal{H}_2$ è una σ -algebra se e solo se ($\mathcal{H}_1 \subseteq \mathcal{H}_2$ oppure $\mathcal{H}_2 \subseteq \mathcal{H}_1$)
- 3 Dato $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ collezione di sottoinsiemi di Ω esiste la più piccola σ -algebra di Ω contenente $\mathcal{A}$. Questa σ -algebra sarà detta σ -algebra generata da $\mathcal{A}$ e sarà indicata con $\sigma(\mathcal{A})$.
- 3' In particolare se $\mathcal{H}_1$ e $\mathcal{H}_2$ sono σ -algebre di Ω allora esiste la più piccola σ -algebra contenente $\mathcal{H}_1$ e $\mathcal{H}_2$.

La proposizione che segue mostra un approccio standard per dimostrare che gli elementi di una σ -algebra soddisfano una assegnata proprietà.

Proposizione 1.4. Dato Ω insieme non vuoto. Sia $\mathcal{A}$ una collezione di sottoinsiemi di Ω e sia P una proprietà dei sottoinsiemi di Ω , (cioè $P(E) \in \{\text{"vero"}, \text{"falso"}\}$ per ogni $E \subseteq \Omega$). Se valgono le seguenti condizioni:

1. $P(\Omega) = \text{"vero"}$ oppure $P(\emptyset) = \text{"vero"}$.
2. $P(E) = \text{"vero"}$ per ogni $E \in \mathcal{A}$.
3. Per ogni $E \in \sigma(\mathcal{A})$, $P(E) = \text{"vero"}$ se e solo se $P(E^c) = \text{"vero"}$
4. Per ogni $\{E_n\}_{n \in \mathbb{N}}$ collezione di elementi di $\sigma(\mathcal{A})$ con $P(E_n) = \text{"vero"}$ per ogni $n \in \mathbb{N}$ si ha $P(\cap_{n \in \mathbb{N}} E_n) = \text{"vero"}$.

allora per ogni $E \in \sigma(\mathcal{A})$ si ha $P(E) = \text{"vero"}$.

Detto con parole semplici, la proposizione precedente ci dice che se una proprietà è vera per gli elementi di $\mathcal{A}$, è vera su Ω ed è chiusa per complementare e intersezione numerabile allora è vera su tutti gli elementi della σ -algebra generata da $\mathcal{A}$.

Osservazione 1.5. La proposizione precedente resta vera se nel punto (4) sostituiamo l'intersezione con l'unione.

Dimostrazione. Proposizione 1.4

Consideriamo la collezione $\mathcal{H}$ di elementi di $\sigma(\mathcal{A})$ per i quali la proprietà è vera

$$\mathcal{H} := \{E \in \sigma(\mathcal{A}) \mid P(E) = \text{"vero"}\}.$$

Le condizioni (1), (3) e (4) ci dicono che $\mathcal{H}$ soddisfa le ipotesi della proposizione 1.1 quindi $\mathcal{H}$ è una σ -algebra. La condizione (2) ci dice che $\mathcal{H}$ contiene $\mathcal{A}$. Poiché $\sigma(\mathcal{A})$ è la più piccola σ -algebra che contiene $\mathcal{A}$, e $\mathcal{H}$ contiene $\mathcal{A}$ allora si deve avere $\mathcal{H} \supseteq \sigma(\mathcal{A})$ e quindi $P(E) = \text{"vero"}$ per ogni $E \in \sigma(\mathcal{A})$. $\square$

Definizione 1.3. Dati due spazi misurabili $(\Omega, \mathcal{H})$ e $(F, \mathcal{F})$ un'applicazione $X : \Omega \rightarrow F$ si dirà misurabile da $(\Omega, \mathcal{H})$ a $(F, \mathcal{F})$ se per ogni $A \in \mathcal{F}$ si ha $X^{-1}(A) \in \mathcal{H}$.

La nozione di applicazione misurabile ha molte similitudini con quella di applicazione continua; per le applicazioni misurabili si richiede che la controimmagine di un misurabile sia misurabile mentre per le applicazioni continue si richiede che la controimmagine di un aperto sia aperto. Vale la regola di composizione, se si compongono due applicazioni misurabili si ottiene ancora un'applicazione misurabile.

Esercizio 1.2. Sia $(F, \mathcal{F})$ uno spazio misurabile. Sia $X : \Omega \rightarrow F$ un'applicazione. Dimostrare che la collezione $\mathcal{H}$ di sottoinsiemi di Ω data da

$$\mathcal{H} := X^{-1}(\mathcal{F}) = \{A \subseteq \Omega \mid \exists B \in \mathcal{F} \text{ con } X^{-1}(B) = A\}$$

è una σ -algebra. $\mathcal{H}$ è la più piccola σ -algebra di Ω che rende misurabile X .

Enunceremo ora una proposizione che utilizzeremo spesso per dimostrare che una applicazione è misurabile.

Proposizione 1.6. Siano $(\Omega, \mathcal{H})$ e $(F, \mathcal{F})$ due spazi misurabili. Sia $X : \Omega \rightarrow F$ un'applicazione. Sia inoltre $\mathcal{A} \subseteq \mathcal{P}(F)$ tale che $\sigma(\mathcal{A}) = \mathcal{F}$. Se per ogni $A \in \mathcal{A}$ vale $X^{-1}(A) \in \mathcal{H}$ allora $X : (\Omega, \mathcal{H}) \rightarrow (F, \mathcal{F})$ è misurabile.

La proposizione precedente ci dice che per mostrare che una applicazione X è misurabile è sufficiente verificare che gli elementi di un insieme di generatori di $\mathcal{F}$ abbia controimmagine in $\mathcal{H}$.

Dimostrazione. Per ogni $A \in \mathcal{F}$ consideriamo la proprietà P definita da:

$$P(A) = \begin{cases} \text{"vero"} & \text{se } X^{-1}(A) \in \mathcal{H} \\ \text{"falso"} & \text{altrimenti} \end{cases}$$

La proprietà P soddisfa le condizioni della proposizione 1.4 quindi per ogni $A \in \sigma(\mathcal{A}) = \mathcal{F}$ si ha $X^{-1}(A) \in \mathcal{H}$ e dunque $X : (\Omega, \mathcal{H}) \rightarrow (F, \mathcal{F})$ è misurabile. $\square$

Definizione 1.4. Dato uno spazio topologico Ω si chiama σ -algebra dei boreliani, $\mathcal{B}(\Omega)$, la più piccola σ -algebra che contiene gli aperti.

Esercizio 1.3. Dimostrare che se $X : \mathbb{R}^n \rightarrow \mathbb{R}^m$ è un'applicazione continua allora è anche misurabile come applicazione da $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ in $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$.

Definizione 1.5. Dato un insieme Ω , un sottoinsieme delle parti $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ si dice un p-system se è chiuso per intersezione cioè

$$\forall A, B \in \mathcal{A} \text{ si ha } A \cap B \in \mathcal{A}$$

Importanti esempi di p-system in $\mathbb{R}$ sono i seguenti:

$$\begin{cases} \mathcal{A}_1 := \{ (a, +\infty) \mid a \in \mathbb{R} \} \\ \mathcal{A}_2 := \{ (a, b] \mid a, b \in \mathbb{R} \} \\ \mathcal{A}_3 := \{ (a_1, b_1] \cup \dots \cup (a_n, b_n] \mid n \in \mathbb{N}, a_1, b_1, \dots, a_n, b_n \in \overline{\mathbb{R}} \} \\ \mathcal{A}_4 := \{ (-\infty, a] \mid a \in \mathbb{R} \} \end{cases}$$

$\mathcal{A}_3$ in realtà è anche un'algebra.

Esercizio 1.4. Dimostrare che:

$$\sigma(\mathcal{A}_1) = \sigma(\mathcal{A}_2) = \sigma(\mathcal{A}_3) = \sigma(\mathcal{A}_4) = \mathcal{B}(\mathbb{R})$$

Gli esempi $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}_3$ e $\mathcal{A}_4$ restano p-system anche se nella definizione sostituiamo $\mathbb{R}$ con $\mathbb{Q}$ in questo ultimo caso sono numerabili e vale ancora $\sigma(\mathcal{A}_k) = \mathcal{B}(\mathbb{R})$.

Enunceremo adesso per completezza alcune proprietà di misurabilità delle applicazioni a valori in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Tali proprietà dovrebbero essere già note agli studenti dai corsi di analisi.

Proposizione 1.7. *Siano X, Y e $\{X_n\}_{n \in \mathbb{N}}$ applicazioni misurabili da uno spazio misurabile $(\Omega, \mathcal{H})$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Sono applicazioni misurabili:*

1. X^+ e X^-
2. $X + Y$
3. $X \cdot Y$
4. X/Y nell'ipotesi $Y(\omega) \neq 0$ per ogni $\omega \in \Omega$
5. $\sup_n X_n$
6. $\inf_n X_n$
7. $\limsup_n X_n$
8. $\liminf_n X_n$

dove $\limsup_n X_n := \inf_n \{\sup_{m>n} X_m\} = \lim_n \{\sup_{m>n} X_m\}$
e $\liminf_n X_n := \sup_n \{\inf_{m>n} X_m\} = \lim_n \{\inf_{m>n} X_m\}$

Dimostrazione.

1) Sia $\mathcal{A}_1$ il p -system definito a pagina 11. Mostriamo che per ogni $A \in \mathcal{A}_1$ vale $X_+^{-1}(A) \in \mathcal{H}$, cosicché grazie all'esercizio 1.4 e alla proposizione 1.6 si avrà X_+ misurabile. Per ipotesi per ogni $a \in \mathbb{R}$ si ha $X^{-1}(a, \infty) \in \mathcal{H}$, per X_+ vale $(X_+)^{-1}(a, \infty) = X^{-1}(a, \infty)$ se $a \geq 0$ e $X_+^{-1}(a, \infty) = \Omega$ se $a < 0$ in entrambi i casi $X_+^{-1}(a, \infty) \in \mathcal{H}$. La dimostrazione per X_- è analoga.

2) procediamo come nel caso precedente e mostriamo che $(X+Y)^{-1}((a, \infty)) \in \mathcal{H}$. Si verifica facilmente che vale la seguente identità

$$(X + Y)^{-1}((a, \infty)) = \bigcup_{q \in \mathbb{Q}} (X^{-1}((q, \infty)) \cap Y^{-1}((q - a, \infty)))$$

dunque $(X + Y)^{-1}((a, \infty))$ è misurabile. Questo punto potrà essere risolto in maniera più elegante utilizzando l'esercizio 1.6.

6) Sia $X := \sup_{n \in \mathbb{N}} X_n$ procediamo ancora come nel caso 1) e mostriamo che $X^{-1}((a, \infty)) \in \mathcal{H}$. Per $\omega \in \Omega$ vale $X(\omega) > a$ se e solo se esiste $n \in \mathbb{N}$ tale che $X_n(\omega) > a$ dunque

$$X^{-1}((a, \infty)) = \bigcup_{n \in \mathbb{N}} X_n^{-1}((a, \infty))$$

per ipotesi $X_n^{-1}((a, \infty))$ appartiene ad $\mathcal{H}$ quindi anche $X^{-1}((a, \infty))$ appartiene ad $\mathcal{H}$

I restanti casi sono lasciati per esercizio. □

1.1.1 σ -algebra ed informazione

Così come negli spazi topologici la topologia è associata ad un concetto intuitivo di forma, in modo simile, negli spazi di probabilità la σ -algebra è

associata al concetto di informazione. Capire in che senso la σ -algebra contenga l'informazione del sistema è cruciale per poter sviluppare la corretta intuizione sugli argomenti che saranno svolti nella seconda parte del corso. Purtroppo però quando ci si avvicina per la prima volta alla probabilità è molto difficile capire in che senso una σ -algebra è associata all'informazione di un modello. Il prossimo teorema ci spiega in che senso la σ -algebra associata ad una variabile aleatoria X contenga la sua informazione.

Teorema 1.8 (Criterio di misurabilità). *Sia Ω un insieme, sia $(E, \mathcal{E})$ uno spazio misurabile, sia X un'applicazione da Ω in E e sia $\sigma(X) := X^{-1}(\mathcal{E})$ la più piccola σ -algebra che rende misurabile X . Sia infine Y un'applicazione da $(\Omega, \sigma(X))$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ allora Y è misurabile se e solo se esiste un'applicazione f da $(E, \mathcal{E})$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ misurabile tale che $Y = f \circ X$.*

$$\begin{array}{ccc} (\Omega, \sigma(X)) & \xrightarrow{X} & (E, \mathcal{E}) \\ & \searrow Y & \downarrow f \\ & & (\mathbb{R}, \mathcal{B}(\mathbb{R})) \end{array}$$

Dimostrazione. La dimostrazione del *se* è ovvia: Y è misurabile perché è composizione di due applicazioni misurabili. Per dimostrare il *solo se* occorre costruire una funzione f misurabile che fa commutare il diagramma. (La dimostrazione è non banale perché l'applicazione X potrebbe non essere suriettiva inoltre l'insieme degli elementi di E che non appartengono all'immagine di X potrebbe non essere misurabile.) Osserviamo per cominciare che si può supporre $Y \geq 0$ senza perdere in generalità. Infatti se così non fosse potremmo porre $Y = Y^+ - Y^-$, risolvendo il problema per Y^+ e Y^- otterrei $Y^+ = f^+ \circ X$ e $Y^- = f^- \circ X$ e quindi posto $f := f^+ - f^-$ avrei $Y = f \circ X$. Consideriamo quindi il caso $Y \geq 0$.

Per $q \geq 0$ razionale consideriamo l'applicazione $Y_q := q \cdot \mathbb{1}_{Y \geq q}$,

$$Y_q = \begin{cases} q & \text{se } Y \geq q \\ 0 & \text{altrimenti} \end{cases}$$

tali funzioni soddisfano l'uguaglianza:

$$Y = \sup_{q \in \mathbb{Q}_+} Y_q$$

Facciamo vedere ora che esiste f_q misurabile tale che $Y_q = f_q \circ X$. Poiché Y è $\sigma(X)$ misurabile allora per costruzione lo è anche Y_q , sia $A_q := Y_q^{-1}(\{q\})$, poiché A_q appartiene a $\sigma(X)$ allora esiste B_q in $\mathcal{E}$ tale che $A_q = X^{-1}(B_q)$,

$$\text{ponendo ora } f_q(e) := \begin{cases} q & \text{se } e \in B_q \\ 0 & \text{altrimenti} \end{cases} \quad \text{si ha} \quad Y_q = f_q \circ X$$

sappiamo che il sup di applicazioni misurabili è misurabile quindi:

$$\text{data } f := \begin{cases} \sup_{q \in \mathbb{Q}_+} f_q & \text{se il sup è finito} \\ 0 & \text{altrimenti} \end{cases} \quad \text{si ha } Y = f \circ X$$

□

Due osservazioni sulla dimostrazione precedente:

- 1) se l'applicazione X non è suriettiva, la scelta degli insiemi B_q potrebbe non essere unica e quindi anche l'applicazione f potrebbe non essere unica.
- 2) Il $\sup_{q \in \mathbb{Q}_+} f_q(e)$ può effettivamente non essere finito, ma tale evenienza può verificarsi solo nel caso in cui e non appartenga all'immagine di X .

Osservazione 1.9 (Informazione $\sigma(X)$ della v.a. X). *Dal teorema 1.8, Y è una applicazione $\sigma(X)$ misurabile se e solo se può essere calcolata a partire da X e vale $Y = f \circ X$, cioè se è noto $X(\omega) = a$ posso determinare $Y(\omega) = f(a)$ senza conoscere il valore esatto di ω .*

1.1.2 d-system e lemma di Dynkin

In questo paragrafo parleremo dei d-system e del lemma di Dynkin, (lemma 1.11). Il lemma di Dynkin sarà utile nel prossimo capitolo quando studieremo gli spazi di misura. Utilizzeremo tale lemma per dimostrare due importanti teoremi, uno sull'unicità delle misure e l'altro sull'indipendenza a partire da un p-system. Va sottolineato che mentre questi due teoremi avranno un sacco di applicazioni, il lemma di Dynkin non sarà quasi più utilizzato.

Definizione 1.6. Dato un insieme Ω , un sottoinsieme delle parti $\mathcal{D} \in \mathcal{P}(\Omega)$ si dice un d-system se soddisfa le seguenti proprietà:

- $\Omega \in \mathcal{D}$
- se $A, B \in \mathcal{D}$ e $A \subseteq B$ allora $B \setminus A \in \mathcal{D}$
- se per ogni $n \in \mathbb{N}$ si ha $A_n \in \mathcal{D}$ e $A_n \subseteq A_{n+1}$ allora $\bigcup A_n \in \mathcal{D}$.

Esercizio 1.5. Sia Ω un insieme e sia $\mathcal{G} \subseteq \mathcal{P}(\Omega)$ consideriamo le seguenti ipotesi su $\mathcal{G}$:

- (i) $\Omega \in \mathcal{G}$
- (ii) Se $A \in \mathcal{G}$ allora $A^c \in \mathcal{G}$
- (iii) Se $A_n \in \mathcal{G}$ e $A_n \cap A_m = \emptyset$ per ogni $n \neq m$ allora $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{G}$
- (a) $\Omega \in \mathcal{G}$
- (b) Se $A, B \in \mathcal{G}$ e $A \subseteq B$ allora $B/A \in \mathcal{G}$

(c) Se $A_n \in \mathcal{G}$ e $A_n \uparrow A$ allora $A \in \mathcal{G}$

Dimostrare che le ipotesi (i), (ii) e (iii) implicano le ipotesi (a), (b) e (c) e viceversa le ipotesi (a), (b) e (c) implicano le ipotesi (i), (ii) e (iii).

Grazie all'esercizio precedente possiamo dare una definizione equivalente di d -system:

Definizione 1.7. Dato un insieme Ω , un sottoinsieme delle parti $\mathcal{D} \subseteq \mathcal{P}(\Omega)$ si dice un d -system se soddisfa le seguenti proprietà:

- $\Omega \in \mathcal{D}$
- se $A \in \mathcal{D}$ allora $A^c \in \mathcal{D}$
- se per ogni $n, m \in \mathbb{N}$ con $n \neq m$ si ha $A_n \in \mathcal{D}$ e $A_n \cap A_m = \emptyset$ allora $\bigcup A_n \in \mathcal{D}$.

Le due definizioni 1.6 e 1.7 sono perfettamente equivalenti. La doppia definizione ci sarà molto utile nelle applicazioni; quando sappiamo che un insieme $\mathcal{D}$ è un d -system potremo utilizzare tutte e sei le condizioni dell'esercizio 1.5 mentre quando dovremo dimostrare che $\mathcal{D}$ è un d -system potremo dimostrare solo le tre più semplici.

Proposizione 1.10. Una collezione di insiemi $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ è una σ -algebra se e solo se è sia un p -system che un d -system.

Dimostrazione. Dimostrare per esercizio. □

Lemma 1.11 (Dynkin). Se una collezione di insiemi $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ è un p -system allora

$$\sigma(\mathcal{A}) = d(\mathcal{A})$$

Dove $d(\mathcal{A})$ è il più piccolo d -system che contiene $\mathcal{A}$.

Dimostrazione. Ovviamente vale $\sigma(\mathcal{A}) \supseteq d(\mathcal{A})$. Grazie alla proposizione precedente per concludere la dimostrazione basterà mostrare che l'insieme $\mathcal{G} := d(\mathcal{A})$ è chiuso per intersezione finita. Occorre mostrare che se A e B appartengono a $\mathcal{G}$ allora anche $A \cap B$ appartiene a $\mathcal{G}$, il caso in cui sia A che B appartengono anche ad $\mathcal{A}$ è ovvio perchè $\mathcal{A}$ è un p -system. Il trucco per completare la dimostrazione è di studiare prima il caso in cui uno A appartiene a $\mathcal{A}$ e B no, e poi il caso generale con A e B in $\mathcal{G}$.

Caso 1) Sia A in $\mathcal{A}$ fissato.

Sia $\mathcal{F}_A$ la collezione di insiemi B di $\mathcal{G}$ che soddisfano la condizione $A \cap B \in \mathcal{G}$:

$$\mathcal{F}_A := \{B \in \mathcal{G} \mid A \cap B \in \mathcal{G}\}$$

sappiamo per ipotesi che $\mathcal{G}$ soddisfa tutte le condizioni dell'esercizio 1.5, mostreremo che $\mathcal{F}_A$ soddisfa le condizioni della definizione 1.7 e contiene $\mathcal{A}$

dunque $\mathcal{F}_A = d(\mathcal{A}) = \mathcal{G}$.

Se $B \in \mathcal{A}$ allora poichè $\mathcal{A}$ è p -system si ha $A \cap B \in \mathcal{A} \subseteq \mathcal{G}$ quindi $\mathcal{A} \subseteq \mathcal{F}_A$

Se $B = \Omega$ allora $A \cap B = A$ e quindi $\mathcal{F}_A$ soddisfa la condizione (i) di 1.5.

Se $B \in \mathcal{F}_A$ allora $A \cap B \in \mathcal{G}$, $A \cap B^c = A \setminus (A \cap B) \in \mathcal{G}$ dunque $\mathcal{F}_A$ soddisfa la condizione (ii) di 1.5.

Se $B_n \in \mathcal{F}_A$ per ogni n e $B_n \cap B_m = \emptyset$ per $n \neq m$ allora per ogni n si ha $B_n \cap A \in \mathcal{G}$, inoltre $A \cap (\bigcup_n B_n) = \bigcup_n (A \cap B_n) \in \mathcal{G}$ quindi $\mathcal{F}_A$ soddisfa anche la condizione (iii) di 1.5.

Questo conclude la dimostrazione del caso 1) ovvero se $A \in \mathcal{A}$ e $B \in \mathcal{G}$ allora $A \cap B \in \mathcal{G}$.

Caso 2) Sia A in $\mathcal{G}$ fissato. La dimostrazione da qui in poi procede come nel caso 1) ed è lasciata per esercizio. $\square$

1.1.3 Teorema della classe monotona [†]

Il teorema della classe monotona, è un importante teorema di analisi. Il suo studio è facoltativo. Ne parlo qui solo perché è una conseguenza immediata del lemma di Dynkin.

Teorema 1.12 (Teorema della classe monotona). *Sia Ω un insieme, sia $\mathcal{A}$ un p -system e $\mathcal{F} := \sigma(\mathcal{A})$. Sia $\mathcal{K}$ una classe di funzioni limitate da Ω in $\mathbb{R}$, se $\mathcal{K}$ soddisfa le seguenti proprietà:*

1. $\mathcal{K}$ è uno spazio vettoriale su $\mathbb{R}$.
2. $\mathbb{1}_\Omega \in \mathcal{K}$
3. $\mathbb{1}_A \in \mathcal{K}$ per ogni $A \in \mathcal{A}$
4. per ogni successione $f_n \in \mathcal{K}$ con $f_n \geq 0$, $f_n \uparrow f$ con f limitata vale $f \in \mathcal{K}$

allora $\mathcal{K}$ contiene tutte le funzioni $\mathcal{F}$ misurabili e limitate.

Dimostrazione. Sia $\mathcal{G} := \{B \in \mathcal{F} | \mathbb{1}_B \in \mathcal{K}\}$. È facile verificare che $\mathcal{G}$ contiene $\mathcal{A}$ ed è un d -system quindi: $\mathcal{G} = \sigma(\mathcal{A}) = \mathcal{F}$. Poiché $\mathcal{K}$ contiene le funzioni indicatrici degli elementi di $\mathcal{F}$ ed è uno spazio vettoriale allora contiene le funzioni semplici. Se f è positiva, limitata e $\mathcal{F}$ misurabile posso approssimarla dal basso con funzioni semplici e dalla proprietà 4 deduco che appartiene a $\mathcal{K}$. Infine il caso generale si deduce da $f = f^+ - f^-$. $\square$

1.1.4 Prodotto di σ -algebre

Consideriamo una famiglia di spazi misurabili $\{(\Omega_i, \mathcal{F}_i)\}_{i \in I}$. È possibile definire il prodotto di tali spazi misurabili procedendo in maniera analogo

[†]Argomento facoltativo.

a quanto gli studenti hanno già visto in topologia con il prodotto di spazi topologici.

$$(\Omega, \mathcal{F}) = \left(\prod_{i \in I} \Omega_i, \bigotimes_{i \in I} \mathcal{F}_i \right)$$

dove $\times_{i \in I}$ indica il prodotto cartesiano degli insiemi Ω_i , e $\bigotimes_{i \in I} \mathcal{F}_i$ indica la più piccola σ -algebra di Ω che renda le proiezioni $\pi_i : \Omega \rightarrow \Omega_i$ misurabili.

Proposizione 1.13. *Siano $\{(\Omega_i, \mathcal{F}_i)\}_{i \in I}$ spazi misurabili, sia $(\Omega, \mathcal{F}) = (\times_{i \in I} \Omega_i, \bigotimes_{i \in I} \mathcal{F}_i)$ e siano $\{\mathcal{A}_i\}_{i \in I}$ collezioni di insiemi tali che per ogni $i \in I$ vale $\sigma(\mathcal{A}_i) = \mathcal{F}_i$ allora*

$$\mathcal{F} = \sigma \left(\left\{ A = \prod_{i \in I} A_i \mid \exists k \in I, A_k \in \mathcal{A}_k, A_i = \Omega_i \text{ per ogni } i \neq k \right\} \right) \quad (1.1)$$

$$\mathcal{F} = \sigma \left(\left\{ A = \prod_{i \in I} A_i \mid \exists J \subseteq I \text{ finito}, A_i \in \mathcal{A}_i \text{ per ogni } i \in J, A_i = \Omega_i \text{ per ogni } i \notin J \right\} \right) \quad (1.2)$$

$$\mathcal{F} = \sigma \left(\left\{ A = \prod_{i \in I} A_i \mid \exists J \subseteq I \text{ al più numerabile}, A_i \in \mathcal{A}_i \text{ per ogni } i \in J, A_i = \Omega_i \text{ per ogni } i \notin J \right\} \right) \quad (1.3)$$

Notare che l'uguaglianza (1.3) in generale non è più vera se si toglie l'ipotesi al più numerabile.

Dimostrazione. L'equivalenza tra le tre equazioni è immediata perchè i generatori della seconda e della terza sono le intersezioni finite o al più numerabili degli insiemi che generano la prima σ algebra. Basta che dimostriamo la prima. Per la proposizione 1.6 $\pi_k : \Omega \rightarrow \Omega_k$ è misurabile se e solo se sono misurabili le controimmagini degli elementi di $\mathcal{A}_k$. Se $A_k \in \mathcal{A}_k$ allora $\pi_k^{-1}(A_k) = \times_{i \in I} A_i$ con $A_i = \Omega_i$ per ogni $i \neq k$. Questi sono esattamente gli insiemi presenti in (1.1). $\square$

Proposizione 1.14. *Sia I un insieme più che numerabile. Siano $\{(\Omega_i, \mathcal{F}_i)\}_{i \in I}$ spazi misurabili e sia $(\Omega, \mathcal{F}) = (\times_{i \in I} \Omega_i, \bigotimes_{i \in I} \mathcal{F}_i)$ allora per ogni $A \in \mathcal{F}$ esiste $J \subseteq I$ al più numerabile tale che:*

$$A \in \sigma \{ \{ \pi_j \}_{j \in J} \}$$

Dimostrazione. Dimostrazione per esercizio. $\square$

Proposizione 1.15. *Siano $(\Omega, \mathcal{H})$ e $\{(E_i, \mathcal{E}_i)\}_{i \in I}$ spazi misurabili. Sia $(E, \mathcal{E}) := (\times_{i \in I} E_i, \otimes_{i \in I} \mathcal{E}_i)$ supponiamo che $\{X_i\}_{i \in I}$ sia una famiglia di applicazioni $X_i : \Omega \rightarrow E_i$, e sia $X := (X_i)_{i \in I}$ il blocco delle variabili X_i . $X : \Omega \rightarrow E$. Allora sono cose equivalenti:*

- $X : (\Omega, \mathcal{H}) \rightarrow (E, \mathcal{E})$ è misurabile
- $\forall i \in I \quad X_i : (\Omega, \mathcal{H}) \rightarrow (E_i, \mathcal{E}_i)$ è misurabile

Dimostrazione. Dimostrazione per esercizio. □

Esercizio 1.6. *Dimostrare che vale:*

$$\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$$

L'esercizio precedente permette di dimostrare in maniera immediata che la somma o il prodotto di applicazioni misurabili è un'applicazione misurabile. In generale se X e Y sono applicazioni misurabili reali e $\Phi : (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2)) \rightarrow (E, \mathcal{E})$ è un'applicazione misurabile allora anche $\Phi(X, Y)$ è misurabile.

Esercizio 1.7. (***) *Sia $\{\Omega_i\}_{i \in I}$ una famiglia di spazi metrici completi e separabili, sia $\{\mathcal{T}_i\}_{i \in I}$ la famiglia delle corrispondenti topologie e supponiamo infine che l'insieme degli indici I sia al più numerabile. Dimostrare che posto $\Omega = \times_{i \in I} \Omega_i$ e $\mathcal{T} = \otimes_{i \in I} \mathcal{T}_i$ allora vale*

$$\mathcal{B}(\Omega) = \bigotimes_{i \in I} \mathcal{B}(\Omega_i) \tag{1.4}$$

dove $\mathcal{B}(\Omega)$ e $\mathcal{B}(\Omega_i)$ indicano la σ -algebra dei boreliani ovvero $\mathcal{B}(\Omega) := \sigma(\mathcal{T})$ e $\mathcal{B}(\Omega_i) = \sigma(\mathcal{T}_i)$.

Capitolo 2

Gli spazi di misura

Le misure sono speciali applicazioni che associano agli elementi di una σ -algebra un elemento di $\overline{\mathbb{R}}$. Esistono molti tipi di misure: misure con segno, misure positive, misure σ -finite, misure finite e misure di probabilità. Il nostro corso chiaramente si concentrerà sulle misure di probabilità. Talvolta sarà necessario fare riferimento alla misura di Lebesgue su $\mathbb{R}$, (misura σ -finita). Le misure non σ -finite verranno utilizzate solo per i controesempi, mentre le misure con segno non verranno mai utilizzate.

2.1 Introduzione agli spazi di misura

2.1.1 Funzioni additive e σ -additive

Sia $\mathcal{A}$ un'algebra, sia μ_0 una funzione $\mu_0 : \mathcal{A} \rightarrow [0, +\infty]$. La funzione μ_0 si dice additiva se per ogni A e B in $\mathcal{A}$ vale

$$A \cap B = \emptyset \implies \mu_0(A \cup B) = \mu_0(A) + \mu_0(B)$$

Definizione 2.1. Data $\mathcal{A}$ algebra. Una funzione $\mu_0 : \mathcal{A} \rightarrow [0, +\infty]$ si dice σ -additiva se per ogni successione $\{A_n\}_{n \in \mathbb{N}}$ di elementi di $\mathcal{A}$ disgiunti ($\forall j \neq k$ vale $A_j \cap A_k = \emptyset$) tali che $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$ vale:

$$\mu_0\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu_0(A_n)$$

2.1.2 Definizione di misura

Dato uno spazio misurabile $(\Omega, \mathcal{H})$. Una applicazione $\mu : \mathcal{H} \rightarrow [0, +\infty]$ è detta misura se è σ -additiva. La tripletta $(\Omega, \mathcal{H}, \mu)$ è detta spazio di misura. Inoltre:

- Una misura μ è detta di probabilità se $\mu(\Omega) = 1$

- Una misura μ è detta finita se $\mu(\Omega) < \infty$
- Una misura μ è detta σ -finita se esiste una successione $\{A_n\}_{n \in \mathbb{N}}$ di elementi di $\mathcal{H}$ tali che per ogni n si ha $\mu(A_n) < \infty$ e inoltre $\bigcup_n (A_n) = \Omega$.

Nel caso di misure σ -finite poiché $\mathcal{H}$ è una σ -algebra allora esiste anche una successione di insiemi misurabili disgiunti (oppure crescenti) $\{B_n\}_{n \in \mathbb{N}}$ di misura finita tali che $\bigcup_n (B_n) = \Omega$.

2.1.3 Quasi sempre, quasi ovunque, quasi certamente, quasi...

Sia $(\Omega, \mathcal{H}, \mu)$ uno spazio di misura. una proprietà P degli elementi di Ω si dice vera quasi ovunque, quasi sempre o quasi certamente se:

$$\exists A \in \mathcal{H} \text{ con } \mu(A) = 0 \text{ tale che } \forall \omega \notin A \text{ si abbia } P(\omega) = \text{“vero”}$$

in questo caso si dirà che per quasi ogni $\omega \in \Omega$ vale $P(\omega)$.

Sugli spazi di probabilità si può passare al complementare e dire che P è quasi certamente vera se esiste $B \in \mathcal{H}$ con $\mu(B) = 1$ tale che per ogni $\omega \in B$ è vera $P(\omega)$.

Osservazione 2.1. Il fatto che una proprietà P sia vera quasi certamente non implica che l'insieme degli ω in cui P è vera sia un insieme misurabile.

Osservazione 2.2. A differenza degli operatori 'per ogni' che commutano tra di loro ($\forall x \forall y P(x, y)$ è equivalente a $\forall y \forall x P(x, y)$) gli operatori 'per quasi ogni' non commutano tra di loro né con gli operatori 'per ogni'. Se $(\Omega, \mathcal{H}, \mu)$ è uno spazio di misura allora in generale le seguenti sono due affermazioni diverse:

- per quasi ogni $x \in \Omega$ per quasi ogni $y \in \Omega$ vale $P(x, y)$
- per quasi ogni $y \in \Omega$ per quasi ogni $x \in \Omega$ vale $P(x, y)$

L'operatore 'per quasi ogni' in generale non commuta nemmeno con 'per ogni' ma vale la seguente proposizione:

Proposizione 2.3. Sia $(\Omega, \mathcal{H}, \mu)$ è uno spazio di misura sia E un insieme al più numerabile (di solito $E = \mathbb{N}$) allora sono cose equivalenti:

1. per quasi ogni $\omega \in \Omega$, per ogni $n \in E$ vale $P(\omega, n)$
2. per ogni $n \in E$, per quasi ogni $\omega \in \Omega$ vale $P(\omega, n)$

Esercizio 2.1.

- Dimostrare che l'implicazione $(1) \implies (2)$ della proposizione precedente è vera anche se l'insieme E non è numerabile.
- Mostrare con un esempio che l'implicazione $(1) \implies (2)$ della proposizione precedente può non essere vera se l'insieme E non è numerabile.

2.2 Un teorema di unicità della misura

Il prossimo teorema è una delle applicazioni principali del lemma di Dynkin.

Teorema 2.4. *Sia $(\Omega, \mathcal{H})$ uno spazio misurabile sia $\mathcal{A}$ un p -system tale che $\sigma(\mathcal{A}) = \mathcal{H}$ e $\Omega \in \mathcal{A}$. Se μ_1 e μ_2 sono due misure finite su $(\Omega, \mathcal{H})$ che coincidono su $\mathcal{A}$ allora μ_1 e μ_2 sono uguali su tutto $\mathcal{H}$.*

Una delle applicazioni più comuni di questo teorema sarà il seguente corollario.

Corollario 2.5. *Due variabili aleatorie reali X e Y hanno la stessa distribuzione se e solo se per ogni intervallo (a, b) vale $P(X \in (a, b)) = P(Y \in (a, b))$.*

Il corollario 2.5 resta valido se si considerano gli intervalli $[a, b]$ chiusi oppure se ci si limita a considerare solo gli intervalli con estremi a e b in $\mathbb{Q}$.

Dimostrazione del teorema 2.4. Cominciamo dalla tesi. Devo mostrare che per ogni $A \in \mathcal{H}$ vale $\mu_1(A) = \mu_2(A)$; sia $\mathcal{G}$ la collezione degli insiemi che soddisfano tale proprietà.

$$\mathcal{G} := \{A \in \mathcal{H} \mid \mu_1(A) = \mu_2(A)\}$$

Per costruzione $\mathcal{G}$ contiene il p -system $\mathcal{A}$, se mostriamo che $\mathcal{G}$ è un d -system allora con il lemma di Dynkin possiamo concludere $\mathcal{H} \supseteq \mathcal{G} \supseteq d(\mathcal{A}) = \sigma(\mathcal{A}) = \mathcal{H}$. Resta solo da mostrare che $\mathcal{G}$ è un d -system.

- (i) $\Omega \in \mathcal{A} \subseteq \mathcal{G}$
- (ii) Se $A \in \mathcal{G}$ allora $\mu_1(A) = \mu_2(A)$
quindi $\mu_1(A^c) - \mu_2(A^c) = \mu_1(\Omega) - \mu_1(A) - (\mu_2(\Omega) - \mu_2(A)) = 0$
- (iii) Sia $\{A_n\}_{n \in \mathbb{N}}$ una successione di elementi di $\mathcal{G}$ disgiunti, per ogni n vale $\mu_1(A_n) = \mu_2(A_n)$ e quindi

$$\mu_1\left(\bigcup_n A_n\right) = \sum_n \mu_1(A_n) = \sum_n \mu_2(A_n) = \mu_2\left(\bigcup_n A_n\right)$$

□

Esercizio 2.2. *Mostrare con un controesempio che nel teorema 2.4 l'ipotesi $\Omega \in \mathcal{A}$ non può essere rimossa.*

Il teorema 2.4 può essere utilizzato anche per dimostrare che la misura di Lebesgue su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ è unica. La misura di Lebesgue non è finita quindi il teorema 2.4 non può essere applicato in maniera diretta. Affinché la dimostrazione funzioni, è sufficiente spezzare la retta $\mathbb{R}$ in unione di intervalli $[n, n+1)$ e mostrare l'unicità su ciascun intervallo.

2.3 Un teorema di esistenza della misura

Il prossimo teorema è il teorema di estensione di Carathéodory. Si tratta di un teorema che mostra l'esistenza di misure con certe proprietà. L'uso principale per noi è quello di mostrare l'esistenza della misura di Lebesgue. La sua trattazione dettagliata avviene nel corso di Analisi Reale. La sua dimostrazione classica fa uso del metodo di Carathéodory, chi non frequenta il corso di Analisi reale può trovare tale metodo esposto nell'appendice del Williams. In tutti i casi la dimostrazione del teorema di estensione è solo per i più curiosi, per questo corso è sufficiente conoscerne l'enunciato.

Teorema 2.6 (Estensione di Carathéodory). *Sia $(\Omega, \mathcal{H})$ uno spazio misurabile, $\mathcal{A}$ un'algebra tale che $\sigma(\mathcal{A}) = \mathcal{H}$. Se $\mu_0 : \mathcal{A} \rightarrow [0, +\infty]$ è una funzione σ -additiva allora esiste una misura μ su $(\Omega, \mathcal{H})$ che coincide con μ_0 su $\mathcal{A}$. Inoltre se $\mu_0(\Omega) < \infty$ allora tale misura è unica.*

Dal teorema precedente si può dedurre il seguente teorema di esistenza della misura di Lebesgue su $[0, 1]$.

Teorema 2.7. *Esiste una misura μ su $([0, 1], \mathcal{B}([0, 1]))$ tale che per ogni $0 \leq a \leq b \leq 1$ vale $\mu(a, b) = b - a$.*

Anche per questo teorema vale quanto detto per quello precedente, la dimostrazione di questo teorema non è richiesta per l'esame, in appendice per i più curiosi può essere trovata una dimostrazione che in parte completa e in parte è alternativa a quella presente sul Williams.

2.4 Proprietà degli spazi di probabilità

La seguente proposizione riassume alcune proprietà delle misure di probabilità che dovrebbero essere note.

Proposizione 2.8. *Sia $(\omega, \mathcal{H}, \mathbb{P})$ uno spazio di probabilità. Siano A, B e $\{A_n\}_{n \in \mathbb{N}}$ elementi di $\mathcal{H}$. Valgono le seguenti affermazioni:*

$$(a) \quad \mathbb{P}(A) + \mathbb{P}(B) = \mathbb{P}(A \cup B) + \mathbb{P}(A \cap B)$$

$$(b) \quad \mathbb{P}(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

$$(c) \quad \text{Se } A_n \uparrow A \text{ allora } \mathbb{P}(A_n) \uparrow \mathbb{P}(A)$$

$$(d) \quad \text{Se } A_n \downarrow A \text{ allora } \mathbb{P}(A_n) \downarrow \mathbb{P}(A)$$

Ad eccezione della proprietà (d) le precedenti proprietà valgono anche nel caso di misure infinite.

Dimostrazione. La proprietà (a) si deduce facilmente dalla additività di $\mathbb{P}$ sugli insiemi disgiunti.

Dimostriamo (b) (proprietà subadditiva della misura). Sia $\{B_n\}_{n \in \mathbb{N}}$ la famiglia di eventi definita da: $B_1 := A_1$, $B_2 = A_2 \setminus A_1$, $B_n := A_n \setminus (A_1 \cup \dots \cup A_{n-1})$. Gli eventi B_n soddisfano le seguenti proprietà:

1. $\{B_n\}_{n \in \mathbb{N}}$ sono una famiglia di eventi disgiunti
2. $B_n \subseteq A_n$ per ogni n
3. $\cup_{n \in \mathbb{N}} B_n = \cup_{n \in \mathbb{N}} A_n$

dunque

$$P(\cup_{n \in \mathbb{N}} A_n) = P(\cup_{n \in \mathbb{N}} B_n) = \sum_{n \in \mathbb{N}} P(B_n) \leq \sum_{n \in \mathbb{N}} P(A_n)$$

dove le prime due uguaglianze sono dedotte dalle proprietà 3 e 1 mentre la disuguaglianza segue dalla proprietà 2.

Dimostriamo ora (c).

Sia $B_1 = A_1$ e per $n > 1$ sia $B_n = A_n \setminus A_{n-1}$ allora gli insiemi $\{B_n\}_{n \in \mathbb{N}}$ sono disgiunti, $\cup_{k=1}^n B_k = A_n$ e $\cup_n B_n = A$. Quindi

$$\mathbb{P}(A) = \sum_n \mathbb{P}(B_n) = \lim_n \sum_{k=1}^n \mathbb{P}(B_k) = \lim_n \mathbb{P}(A_n)$$

Infine la dimostrazione di (d) si ottiene passando al complementare ed utilizzando (c). $\square$

Indichiamo la funzione indicatrice di un insieme con la notazione: $\mathbb{1}_A$. Valgono le seguenti definizioni per i limiti sugli insiemi:

$$\begin{aligned} \limsup_{n \rightarrow \infty} A_n = A & \iff \limsup_{n \rightarrow \infty} \mathbb{1}_{A_n} = \mathbb{1}_A \\ \liminf_{n \rightarrow \infty} A_n = A & \iff \liminf_{n \rightarrow \infty} \mathbb{1}_{A_n} = \mathbb{1}_A \\ \lim_{n \rightarrow \infty} A_n = A & \iff \lim_{n \rightarrow \infty} \mathbb{1}_{A_n} = \mathbb{1}_A \end{aligned}$$

È necessario però aggiungere qualche spiegazione per capire meglio l'intuizione che c'è sotto le definizioni di $\limsup$ e $\liminf$ di una famiglia di eventi. Se non si comprende tale intuizione sarà difficile poi capire come e quando possono essere utilizzati. Sia $(\Omega, \mathcal{H}, \mathbb{P})$ uno spazio di probabilità, sia $\{A_n\}_{n \in \mathbb{N}}$ una famiglia di eventi. Allora:

$$\begin{aligned} A = \limsup_n A_n &= \bigcap_n \bigcup_{m \geq n} A_m \\ &= \{\omega \in \Omega \mid \forall m \exists n > m \text{ tale che } \omega \in A_n\} \\ &= \{\omega \in \Omega \mid \omega \in A_n \text{ per infiniti } n\} \\ &= \{\omega \in \Omega \mid \#\{n \in \mathbb{N} \mid \omega \in A_n\} = \infty\} \end{aligned}$$

A parole l'evento A è l'evento “si verificano infiniti A_n ”.

Sia ancora $\{A_n\}_{n \in \mathbb{N}}$ una famiglia di eventi. Allora:

$$\begin{aligned} A &= \liminf_n A_n = \bigcup_n \bigcap_{m \geq n} A_m \\ &= \{\omega \in \Omega \mid \exists m \text{ tale che } \forall n \geq m \text{ si ha } \omega \in A_n\} \\ &= \{\omega \in \Omega \mid \#\{n \in \mathbb{N} \mid \omega \notin A_n\} < \infty\} \end{aligned}$$

A parole l'evento A è l'evento “da un certo punto in poi A_n si verifica sempre”, o anche “ A_n si verifica sempre a meno di un insieme finito di indici”. Valgono i seguenti lemmi:

Lemma 2.9 (lemma di Fatou per eventi). *Sia $(\Omega, \mathcal{H}, \mathbb{P})$ spazio di probabilità, per ogni famiglia di eventi $\{A_n\}_{n \in \mathbb{N}}$ vale*

$$\mathbb{P}\left(\liminf_n A_n\right) \leq \liminf_n \mathbb{P}(A_n)$$

Dimostrazione. Sia $A := \liminf_n A_n$.

$$\mathbb{P}(A) = \mathbb{P}\left(\liminf_n A_n\right) = \mathbb{P}\left(\bigcup_n \bigcap_{m \geq n} A_m\right)$$

Sia $B_n := \bigcap_{m \geq n} A_m$ allora si ha $B_n \subseteq A_n$ e $B_n \uparrow A$ quindi

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_n B_n\right) = \lim_n \mathbb{P}(B_n) \leq \liminf_n \mathbb{P}(A_n)$$

□

Lemma 2.10 (lemma di Fatou inverso per eventi). *Sia $(\Omega, \mathcal{H}, \mathbb{P})$ spazio di probabilità, per ogni famiglia di eventi $\{A_n\}_{n \in \mathbb{N}}$ vale*

$$\mathbb{P}\left(\limsup_n A_n\right) \geq \limsup_n \mathbb{P}(A_n)$$

Dimostrazione. È sufficiente passare al complementare e applicare il lemma precedente. □

Esercizio 2.3. *Trovare un esempio in cui le disuguaglianze dei due precedenti lemmi di Fatou siano disuguaglianze strette.*

Lemma 2.11 (Primo lemma di Borel-Cantelli). *Sia $(\Omega, \mathcal{H}, \mathbb{P})$ spazio di probabilità, sia $\{A_n\}_{n \in \mathbb{N}}$ una famiglia di eventi. Se vale $\sum_n \mathbb{P}(A_n) < \infty$ allora*

$$\mathbb{P}\left(\limsup_n (A_n)\right) = 0$$

Dimostrazione. Sia $A := \limsup_n A_n$.

$$\mathbb{P}(A) = \mathbb{P}\left(\limsup_n A_n\right) = \mathbb{P}\left(\bigcap_n \bigcup_{m \geq n} A_m\right)$$

Sia $B_n := \bigcup_{m \geq n} A_m$ allora si ha $B_n \supseteq A_n$ e $B_n \downarrow A$ quindi

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcap_n B_n\right) = \lim_n \mathbb{P}(B_n) \leq \mathbb{P}(B_m) \quad \forall m$$

per la subattività si ha:

$$\mathbb{P}(A) \leq \mathbb{P}(B_m) \leq \sum_{n=m}^{\infty} \mathbb{P}(A_n) \quad \forall m$$

poichè la serie $\sum_n \mathbb{P}(A_n)$ converge allora si ha

$$\lim_{m \rightarrow \infty} \sum_{n=m}^{\infty} \mathbb{P}(A_n) = 0$$

□

Il primo lemma di Borel-Cantelli verrà utilizzato molto spesso durante il corso quindi è importante comprendere bene il significato della sua tesi $P(\limsup_n (A_n)) = 0$.

$$\begin{aligned} A &:= \limsup A_n = \{\omega \in \Omega \mid \#\{n \mid \omega \in A_n\} = \infty\} \\ A^c &= \{\omega \in \Omega \mid \#\{n \mid \omega \in A_n\} < \infty\} \end{aligned}$$

Quindi

$$P(\limsup_n (A_n)) = 0 \quad \Longleftrightarrow \quad \text{per quasi ogni } \omega \in \Omega \quad \#\{n \mid \omega \in A_n\} < \infty \quad (2.1)$$

Esercizio 2.4. Consideriamo una successione di variabili aleatorie $(X_n)_{n \in \mathbb{N}}$ con distribuzione $P(X_n = 1) = \frac{1}{n^2}$ e $P(X_n = 0) = 1 - \frac{1}{n^2}$. Sia S infine la somma $S := \sum_n X_n$. Calcolare la probabilità $P(S = +\infty)$.

Svolgimento. Dobbiamo calcolare la probabilità che S sia uguale a $+\infty$. Di fatto la variabile S conta quante volte si verifica l'evento $X_n = 1$ cioè:

$$S = \#\{n \in \mathbb{N} \mid X_n = 1\}$$

L'idea è di utilizzare il primo lemma di Borel-Cantelli per dimostrare che la probabilità che l'evento $X_n = 1$ si verifichi infinite volte è zero.

Per $n \in \mathbb{N}$ sia $A_n := \{X_n = 1\}$. Vale $S(\omega) = +\infty$ se e solo se ω appartiene ad infiniti A_n ovvero

$$\{S = +\infty\} = \limsup_n A_n$$

Verifichiamo ora le ipotesi del primo lemma di Borel-Cantelli

$$\sum_n P(A_n) = \sum_n \frac{1}{n^2} < \infty$$

dunque

$$P(\{S = +\infty\}) = P(\limsup_n A_n) = 0$$

□

2.4.1 Legge di una variabile aleatoria

Sia $(\Omega, \mathcal{H}, P)$ uno spazio di probabilità, sia $(E, \mathcal{E})$ uno spazio misurabile e sia $X : (\Omega, \mathcal{H}, P) \rightarrow (E, \mathcal{E})$ misurabile. X induce una misura μ su $(E, \mathcal{E})$ che di solito si denota con $\mu = P(X^{-1}) = X(P)$ ed è definita da:

$$\mu(A) = P(X^{-1}(A)) \quad \forall A \in \mathcal{E}$$

μ è detta legge di X oppure distribuzione di X .

Sia ora $(F, \mathcal{F})$ un ulteriore spazio misurabile e sia $Y : (E, \mathcal{E}) \rightarrow (F, \mathcal{F})$ misurabile. Allora posto $Z = Y \circ X$ e $\mu_2 = \mu(Y^{-1})$ vale:

$$\mu_2 = P(Z^{-1}) \quad \text{ovvero} \quad (Y \circ X)(P) = Y(X(P))$$

2.4.2 Funzione di ripartizione

Sia X una variabile aleatoria da $(\Omega, \mathcal{H}, P)$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ si definisce funzione di ripartizione la seguente funzione $F_X : \mathbb{R} \rightarrow \mathbb{R}$:

$$F_X(t) = P(X \leq t)$$

La funzione di ripartizione gode delle seguenti quattro proprietà caratterizzanti:

- (i) F_X è non decrescente
- (ii) F_X è continua a destra
- (iii) $\lim_{t \rightarrow -\infty} F_X(t) = 0$
- (iv) $\lim_{t \rightarrow +\infty} F_X(t) = 1$

Verificare che la funzione di ripartizione di una v.a. reale soddisfa le condizioni precedenti è immediato. Dimosteremo che vale anche il viceversa se F è una funzione che soddisfa le quattro proprietà enunciate allora esiste una v.a. che ha F come funzione di ripartizione.

Proposizione 2.12. *Sia $F : \mathbb{R} \rightarrow \mathbb{R}$ una funzione, se F soddisfa le proprietà (i),..., (iv) allora esiste una variabile aleatoria X tale che la sua funzione di ripartizione è uguale ad F .*

Prima di tutto spieghiamo l'idea nel caso speciale in cui F sia strettamente crescente e continua. In questo caso la funzione $F : \mathbb{R} \rightarrow (0, 1)$ è invertibile, sia $H : (0, 1) \rightarrow \mathbb{R}$ la sua inversa. La funzione H è a sua volta continua e strettamente crescente. Sia $(\Omega, \mathcal{H}, P) = ((0, 1), \mathcal{B}(0, 1), \mathcal{L})$ dove $\mathcal{L}$ è la misura di Lebesgue. Sia $X := H$, X è non decrescente quindi misurabile inoltre

$$F_X(t) = P(X \leq t) = P(\{\omega \in \Omega | X(\omega) \leq t\}) = \quad (2.2)$$

$$= P(\{\omega \in \Omega | \omega \leq F(t)\}) = \quad (2.3)$$

$$= \mathcal{L}((-\infty, F(t)) \cap (0, 1)) \quad (2.4)$$

$$= \mathcal{L}(0, F(t)) = F(t) \quad (2.5)$$

In generale però una funzione non crescente e continua a destra non è inver-

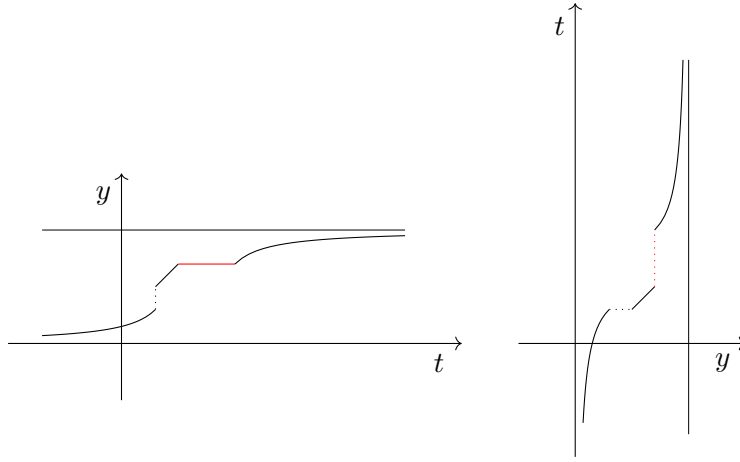


Figura 2.1: Sulla sinistra la funzione F sulla destra la sua inversa.

tibile. Cerchiamo allora di costruire una funzione che assomigli più possibile all'inversa di F e che sia monotona. Osserviamo la figura 2.1. Nella parte sinistra della figura è rappresentata il grafico di una funzione di distribuzione F che non è continua e non è strettamente crescente. A sinistra il possibile grafico dell'inversa di F , il grafico dell'inversa si ottiene scambiando le variabili t e y cioè scambiando gli assi e ribaltando il grafico di sinistra. Si nota subito che ci sono due difficoltà nel costruire l'inversa: la prima difficoltà è quella rappresentata dalla linea tratteggiata in nero dove non c'è un'inversa; la seconda difficoltà è costituita dalla linea tratteggiata in rosso dove l'inversa non è unica. La prima difficoltà si supera imponendo la monotonia. La seconda difficoltà invece resta ci sono più funzioni monotone che coindono con l'inversa dove questa è unica. Tra queste ce ne sono due privilegiate la funzione q detta quantile che nei casi come quello rappresentato dai puntini in rosso seleziona l'estremo superiore e la funzione H detta pseudoinversa

che nei casi in cui l'inversa non sia unica seleziona il minimo.

$$\begin{aligned} q(y) &:= \inf\{t \in \mathbb{R} \mid F(t) > y\} = \sup\{t \in \mathbb{R} \mid F(t) \leq y\} \\ H(y) &:= \inf\{t \in \mathbb{R} \mid F(t) \geq y\} = \sup\{t \in \mathbb{R} \mid F(t) < y\} \end{aligned}$$

si verifica facilmente che le funzioni q e H sono non decrescenti, la funzione q è continua a destra mentre la funzione H è continua a sinistra, vale $q \geq H$. Le funzioni q ed H sono diverse solo nei punti di discontinuità dove la prima è continua a destra mentre la seconda è continua a sinistra. Inoltre poiché i punti di discontinuità di una funzione monotona sono al più numerabili allora anche l'insieme dei punti in cui q e H sono diverse è al più numerabile. Un po' più difficile è verificare che vale la seguente relazione:

$$H(y) \leq \bar{t} \quad \Longleftrightarrow \quad y \leq F(\bar{t}) \quad (2.6)$$

Dimostriamo la (2.6).

Se $y \leq F(\bar{t})$ allora $\bar{t} \in \{t \in \mathbb{R} \mid F(t) > y\}$ quindi $H(y) \leq \bar{t}$.

Se invece $y > F(\bar{t})$ allora poiché F è continua a destra esiste $\varepsilon > 0$ tale che $y > F(\bar{t} + \varepsilon)$ allora $\bar{t} + \varepsilon \notin \{t \in \mathbb{R} \mid F(t) > y\}$ quindi $H(y) \geq \bar{t} + \varepsilon > \bar{t}$. $\square$

A questo punto grazie alla (2.6) posto $X := H$ vale quanto detto in (2.2).

Funzione di ripartizione per variabili aleatorie in $\overline{\mathbb{R}}$

Talvolta per questioni di compattezza può essere utile lavorare con le variabili aleatorie a valori in $\overline{\mathbb{R}}$. La funzione di ripartizione si definisce ancora come $F_X(t) = P(X \leq t)$ ed è caratterizzata dalle seguenti proprietà:

- (i) F_X è non decrescente
- (ii) F_X è continua a destra
- (iii) $F_X(-\infty) \geq 0$
- (iv) $F_X(+\infty) = 1$

Capitolo 3

Integrazione

3.1 Valore atteso di una variabile aleatoria

La definizione di integrale di un'applicazione misurabile di solito è sviluppata durante i corsi di analisi. In queste dispense presenterò per completezza le nozioni di base senza dimostrazioni. Mi limiterò ad enunciare i teoremi e le definizioni che verranno usate durante il corso; per i dettagli delle dimostrazioni rimando gli studenti ad un buon corso di analisi (o a un buon libro).

Sia $(\Omega, \mathcal{H}, \mu)$ uno spazio di misura. Sia $\mathcal{S}$ lo spazio delle funzioni semplici da $(\Omega, \mathcal{H})$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, (funzioni misurabili che assumono solo un numero finito di valori). L'integrale di una funzione semplice si definisce nel seguente modo:

$$\forall Y \in \mathcal{S} \quad \int_{\Omega} Y(\omega) d\mu(\omega) = \sum_{x \in \mathbb{R}} x \cdot \mu(\{Y = x\}) \quad (3.1)$$

dove come al solito con $\{Y = x\}$ si intende l'insieme: $\{\omega \in \Omega \mid Y(\omega) = x\}$. Si dimostra che l'integrale definito in 3.1 è lineare sulle funzioni semplici ovvero $\int_{\Omega} \alpha X + Y d\mu = \alpha \int_{\Omega} X d\mu + \int_{\Omega} Y d\mu$, è monotono (se $X \leq Y$ allora $\int X d\mu \leq \int Y d\mu$) e se $X = \mathbb{1}_A$ allora $\int X d\mu = \mu(A)$.

In generale se X è un'applicazione misurabile non negativa si definisce l'integrale:

$$\int_{\Omega} X(\omega) d\mu(\omega) := \sup_{\substack{Y \in \mathcal{S} \\ 0 \leq Y \leq X}} \int_{\Omega} Y(\omega) d\mu(\omega) \quad (3.2)$$

Si dimostra che l'integrale definito in (3.2) per le funzioni non negative è lineare, monotono e coincide con quello definito in (3.1) quando X è semplice. In pratica l'integrale di X si ottiene approssimando X dal basso con funzioni semplici. Si dimostra inoltre che non è necessario calcolare il *sup* della (3.2) su tutte le Y ma è sufficiente farlo su una qualunque successione che tende ad X dal basso vale la seguente proposizione:

Proposizione 3.1. *Sia X da $(\Omega, \mathcal{H}, \mu)$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ misurabile e non negativa, siano Y_n semplici tali che $0 \leq Y_n \leq X$ e $Y_n \uparrow X$ allora*

$$\int_{\Omega} X(\omega) d\mu(\omega) := \sup_n \int_{\Omega} Y_n(\omega) d\mu(\omega) \quad (3.3)$$

Data X non negativa è sempre possibile costruire una successione $\{Y_n\}_{n \in \mathbb{N}}$ che soddisfa le condizioni della proposizione precedente, se per esempio $\{q_n\}_{n \in \mathbb{N}}$ è una successione dei numeri razionali positivi con $q_0 = 0$ allora si può porre

$$Y_n := \sup\{q_k | 0 \leq k \leq n, q_k \leq X\} \quad (3.4)$$

La definizione generale dell'integrale di una funzione non necessariamente positiva è data da:

$$\int_{\Omega} X(\omega) d\mu(\omega) = \int_{\Omega} X^+(\omega) d\mu(\omega) - \int_{\Omega} X^-(\omega) d\mu(\omega) \quad (3.5)$$

con la solita convenzione sulla somma degli infiniti. A partire dalla definizione (3.5) e (3.2) è possibile dimostrare varie proprietà: monotonia:

$$X \leq Y \quad \implies \quad \int_{\Omega} X(\omega) d\mu(\omega) \leq \int_{\Omega} Y(\omega) d\mu(\omega)$$

linearità (nell'ipotesi che X e Y ammettano integrali finiti):

$$\int_{\Omega} \lambda_1 X(\omega) + \lambda_2 Y(\omega) d\mu(\omega) = \lambda_1 \int_{\Omega} X(\omega) d\mu(\omega) + \lambda_2 \int_{\Omega} Y(\omega) d\mu(\omega)$$

Se $\mu = P$ è una misura di probabilità allora scriveremo:

$$\mathbb{E}[X] = \int_{\Omega} X(\omega) dP(\omega)$$

per indicare $\mathbb{E}[X]$ a parole ci sono molti modi diversi, i più comuni sono: *speranza matematica, valore atteso, valore medio e media.*

Teorema 3.2 (Regola di composizione). *Sia $(\Omega, \mathcal{H}, \mu)$ uno spazio di misura. Sia ϕ misurabile da $(\Omega, \mathcal{H})$ a $(E, \mathcal{E})$. Sia $\nu := \phi(\mu)$. Sia infine $X : (E, \mathcal{E}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ misurabile e $Y = X(\phi)$. Se X è integrabile vale:*

$$\int_E X d\nu = \int_{\Omega} Y d\mu \quad (3.6)$$

Dimostrazione. La dimostrazione procede con un metodo standard, si dimostra prima la tesi per il caso X funzione indicatrice poi si studia il caso X funzione semplice e infine si passa al caso generale.

Consideriamo per cominciare il caso $X = \mathbb{1}_A$ con $A \in \mathcal{E}$.

$$\int_E X d\nu = \nu(A) = \mu(\phi^{-1}(A)) = \int_{\Omega} \mathbb{1}_{\phi^{-1}(A)} d\mu = \int_{\Omega} Y d\mu$$

Caso X funzione semplice. Se X è una funzione semplice allora è combinazione lineare di funzioni indicatrici. Allora dalla linearità degli integrali della (3.6) si deduce la tesi.

Caso X funzione misurabile positiva. Se X_n sono funzioni semplici positive con $X_n \uparrow X$ allora posto $Y_n := X_n(\phi)$ si ha $Y_n \uparrow Y$. La tesi segue dalla proposizione 3.1. $\square$

3.2 Teoremi limiti

Ora enunceremo tre importanti teoremi limiti di analisi e per completezza inserirò anche la dimostrazione.

Teorema 3.3 (Convergenza monotona). *Sia $(\Omega, \mathcal{H}, \mu)$ uno spazio di misura. Siano X_n e X variabili aleatorie quasi certamente non negative. Se X_n è non decrescente in n e tende ad X quasi certamente allora*

$$\int_{\Omega} X_n d\mu \uparrow \int_{\Omega} X d\mu$$

se μ è una misura di probabilità il teorema precedente diventa

$$\begin{cases} 0 \leq X_n \leq X \\ X_n \uparrow X \text{ q.c.} \end{cases} \implies \mathbb{E}[X_n] \uparrow \mathbb{E}[X]$$

Dimostreremo il teorema di convergenza monotona, la sua dimostrazione sarà molto semplice perchè gli aspetti più noiosi sono stati nascosti all'interno della dimostrazione della proposizione 3.1.

Dimostrazione Teorema 3.3.

Per la monotonia dell'integrale si ha $\int_{\Omega} X_n d\mu$ crescente in n e minore di $\int_{\Omega} X d\mu$. Occorre solo mostrare che $\int_{\Omega} X d\mu$ non è strettamente maggiore del limite degli integrali di X_n . Sia $\varepsilon > 0$ per definizione di integrale esiste una funzione semplice Y tale che $0 \leq Y \leq X$ e $\int_{\Omega} Y d\mu \geq \int_{\Omega} X d\mu - \varepsilon$. Poniamo $Y_n = \min\{Y, X_n\}$. Quindi $X_n \geq Y_n$. Dalla monotonia e dalla proposizione 3.1 otteniamo:

$$\lim_{n \rightarrow \infty} \int_{\Omega} X_n d\mu \geq \lim_{n \rightarrow \infty} \int_{\Omega} Y_n d\mu = \int_{\Omega} Y d\mu \geq \int_{\Omega} X d\mu - \varepsilon$$

$\square$

Osservazione 3.4. *L'ipotesi $X_n \geq 0$ del teorema di convergenza monotona può essere eventualmente indebolita in $X_n \geq Y$ per ogni n e Y v.a. che ammette media finita.*

Esercizio 3.1. *Dimostrare l'osservazione precedente. Sia Y v.a. tale che $\mathbb{E}[Y^-] < \infty$ allora vale:*

$$\begin{cases} Y \leq X_n \leq X \\ X_n \uparrow X \text{ q.c.} \end{cases} \implies \mathbb{E}[X_n] \uparrow \mathbb{E}[X]$$

Teorema 3.5 (Lemma di Fatou).

Sia X_n una successione di funzioni misurabili quasi certamente non negative. Vale:

$$\int_{\Omega} \liminf_{n \rightarrow \infty} X_n d\mu \leq \liminf_{n \rightarrow \infty} \int_{\Omega} X_n d\mu$$

Dimostrazione.

$$\begin{aligned} \int_{\Omega} \liminf_{n \rightarrow \infty} X_n d\mu &= \int_{\Omega} \sup_n \inf_{k \geq n} X_k d\mu = \sup_n \int_{\Omega} \inf_{k \geq n} X_k d\mu \\ &\leq \sup_n \inf_{k \geq n} \int_{\Omega} X_k d\mu = \liminf_n \int_{\Omega} X_k d\mu \end{aligned}$$

dove il primo e l'ultimo uguale sono la definizione di $\liminf$, il secondo uguale segue dal teorema di convergenza monotona mentre la disuguaglianza è dovuta alla proprietà monotona per cui per ogni $k \geq n$ vale $\int_{\Omega} X_k d\mu \geq \int_{\Omega} \inf_{k \geq n} X_k d\mu$. $\square$

L'ipotesi $X_n \geq 0$ quasi certamente può in realtà essere indebolita.

Teorema 3.6 (Lemma di Fatou bis).

Sia Y funzione misurabile tale che $\int_{\Omega} Y^- d\mu < \infty$. Sia $\{X_n\}_{n \in \mathbb{N}}$ una successione di applicazioni misurabili. Se per ogni n vale $X_n \geq Y$ quasi certamente allora

$$\int_{\Omega} \liminf_{n \rightarrow \infty} X_n d\mu \leq \liminf_{n \rightarrow \infty} \int_{\Omega} X_n d\mu$$

La dimostrazione del precedente teorema è immediata, basta definire le variabili $\widetilde{X}_n := X_n - Y$ ed applicare ad esse il teorema 3.5.

Teorema 3.7 (Lemma di Fatou inverso).

Sia Y funzione misurabile tale che $\int_{\Omega} Y^+ d\mu < \infty$. Sia $\{X_n\}_{n \in \mathbb{N}}$ una successione di applicazioni misurabili. Se per ogni n vale $X_n \leq Y$ quasi certamente allora

$$\int_{\Omega} \limsup_{n \rightarrow \infty} X_n d\mu \geq \limsup_{n \rightarrow \infty} \int_{\Omega} X_n d\mu$$

Anche in questo caso la dimostrazione è immediata basta definire le applicazioni $\widetilde{X}_n := -X_n$ e $\widetilde{Y} := -Y$ ed applicare il teorema 3.6.

Teorema 3.8 (Convergenza dominata).

Sia Y funzione misurabile che ammetta integrale finito (ovvero vale $\int_{\Omega} |Y| d\mu < \infty$). Siano X e $\{X_n\}_{n \in \mathbb{N}}$ applicazioni misurabili. Se per ogni n vale $|X_n| \leq Y$ quasi certamente allora

$$X_n \xrightarrow[n \rightarrow \infty]{} X \quad \text{q.c.} \quad \implies \quad \int_{\Omega} X_n d\mu \xrightarrow[n \rightarrow \infty]{} \int_{\Omega} X d\mu$$

3.3. VARIABILI ALEATORIE DISCRETE E ASSOLUTAMENTE CONTINUE 33

Dimostrazione. È sufficiente applicare i due lemmi precedenti 3.7 e 3.6

$$\limsup_{n \rightarrow \infty} \int_{\Omega} X_n d\mu \leq \int_{\Omega} X d\mu \leq \liminf_{n \rightarrow \infty} \int_{\Omega} X_n d\mu$$

□

Esercizio 3.2. Siano $X, Y, \{X_n\}_{n \in \mathbb{N}}$ applicazioni misurabili tali che per ogni n vale $Y \leq X_n \leq X$. Se $\int_{\Omega} Y^- d\mu < \infty$ allora vale

$$X_n \xrightarrow[n \rightarrow \infty]{q.c.} X \quad \implies \quad \int_{\Omega} X_n d\mu \xrightarrow[n \rightarrow \infty]{} \int_{\Omega} X d\mu$$

Esercizio 3.3. Costruire un esempio di variabili aleatorie $\{X_n\}_{n \in \mathbb{N}}$ e X quasi certamente non-negative, tali che:

- $X_n, X \in L^1$
- $X_n \xrightarrow[n \rightarrow \infty]{q.c.} X$
- $\lim_n \mathbb{E}[X_n] \neq \mathbb{E}[X]$

Esercizio 3.4. Siano $\{X_n\}_{n \in \mathbb{N}}$ e X variabili aleatorie in L^1 quasi certamente non-negative tali che $X_n \xrightarrow[n \rightarrow \infty]{q.c.} X$. Dimostrare che

$$\liminf_n \mathbb{E}[X_n] \geq \mathbb{E}[X]$$

3.3 Variabili aleatorie discrete e assolutamente continue

Questa sezione estende agli spazi di misura $(E, \mathcal{E}, \mu)$ alcune nozioni di probabilità viste nel corso di probabilità del primo anno e che in quel caso riguardavano lo spazio $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{L})$.

Spazi di probabilità discreta e variabili aleatorie discrete

Uno spazio di probabilità $(E, \mathcal{E}, \mu)$ si dice discreto se esiste un insieme $A \in \mathcal{E}$ al più numerabile tale che $\mu(A) = 1$. La funzione da E in $[0, 1]$ definita da $p_x := \mu(\{x\})$ è detta densità discreta della misura μ .

Una variabile aleatoria X da uno spazio di probabilità $(\Omega, \mathcal{H}, P)$ in $(E, \mathcal{E})$ si dice discreta se la sua legge è discreta ovvero se esiste un insieme $A \in \mathcal{E}$ al più numerabile tale che $P(X \in A) = 1$. In questo caso si ha:

$$\sum_{e \in E} P(X = e) = 1$$

Data $g : (E, \mathcal{E}) \rightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ vale

$$\mathbb{E}[g(X)] = \sum_{e \in E} P(X = e) \cdot g(e) = \sum_{t \in \overline{\mathbb{R}}} P(g(X) = t) \cdot t$$

con le solite regole sulla somma degli infiniti.

Variabili aleatorie assolutamente continue e misure di probabilità assolutamente continue

Siano $(\Omega, \mathcal{H}, P)$ uno spazio di probabilità e $(E, \mathcal{E}, \mu)$ uno spazio di misura. Introduciamo ora la definizione di variabile aleatoria e misura assolutamente continua a valori in $(E, \mathcal{E}, \mu)$ (In quasi tutte le nostre applicazioni sarà $(E, \mathcal{E}, \mu) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathcal{L}^n)$ e la maggior parte delle volte sarà $n = 1$.)

Sia f da $(E, \mathcal{E})$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ misurabile, non negativa e tale che $\int_E f \, d\mu = 1$. Si può facilmente verificare che la funzione ν da $\mathcal{E}$ in $[0, 1]$ data da:

$$\nu(A) := \int_A f(x) \, d\mu(x) := \int_E \mathbb{1}_A(x) \cdot f(x) \, d\mu(x)$$

è una misura di probabilità. In questo caso diremo che ν è assolutamente continua rispetto a μ e f sarà detta densità di ν rispetto a μ . (Talvolta soprattutto in analisi incontrerete la notazione $f = \frac{d\nu}{d\mu}$ anche detta derivata di Radon Nikodyn)

Una variabile aleatoria X da $(\Omega, \mathcal{H})$ in $(E, \mathcal{E})$ si dice assolutamente continua (rispetto a μ) se la sua legge è assolutamente continua ovvero se esiste una funzione misurabile $f_X : (E, \mathcal{E}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ detta densità di X tale che per ogni $A \in \mathcal{E}$ vale

$$\int_A f_X(x) d\mu = P(X \in A)$$

Vale la seguente caratterizzazione delle funzioni di densità

$$f \geq 0 \quad q.c. \quad \int_E f(x) d\mu = 1 \quad (3.7)$$

La necessità di (3.7) è immediata: la seconda condizione segue da $\int_E f(x) d\mu = P(X \in E) = 1$ mentre la prima segue dal seguente esercizio.

Esercizio 3.5. *Sia f è una funzione misurabile quasi certamente maggiore o uguale a zero (o quasi certamente minore o uguale). Allora vale:*

$$\int_E f(x) d\mu = 0 \quad \Longleftrightarrow \quad f = 0 \quad q.c.$$

La versione probabilistica dell'esercizio precedente è:

Esercizio 3.6. *Sia X è una variabile aleatoria quasi certamente maggiore o uguale a zero (o quasi certamente minore o uguale a zero). Allora vale:*

$$\mathbb{E}[X] = 0 \quad \Longleftrightarrow \quad X = 0 \quad q.c.$$

Svolgimento Esercizio 3.5. Consideriamo il caso $f \geq 0$. L'implicazione $\Leftarrow$ è ovvia. Vediamo ora la seconda implicazione. Sia $\varepsilon > 0$ allora $f \geq \varepsilon \mathbb{1}_{f > \varepsilon}$ quindi

$$\forall \varepsilon > 0 \quad 0 = \int_{\mathbb{R}^n} f(x) dx \geq \int_{\mathbb{R}^n} \varepsilon \mathbb{1}_{f > \varepsilon} dx = \varepsilon \cdot \mathcal{L}^n(\{f > \varepsilon\}) \geq 0$$

Quindi per ogni $\varepsilon > 0$ vale $\mathcal{L}^n(\{f > \varepsilon\}) = 0$.

Dimostriamo ora la caratterizzazione (3.7). Dobbiamo dimostrare che se f soddisfa (3.7) allora esiste uno spazio di probabilità $(\Omega, \mathcal{H}, P)$ ed una variabile aleatoria X che ammette f come funzione di densità. Si procede con un trucco si pone $\Omega := E$, $\mathcal{H} = \mathcal{E}$ e $P(A) := \int_A f(x) d\mu$ si verifica facilmente che $(\Omega, \mathcal{H}, P)$ è uno spazio di probabilità e $X(\omega) := \omega$ ammette banalmente f come densità.

Proposizione 3.9. *Sia X una variabile aleatoria da $(\Omega, \mathcal{H}, P)$ in $(E, \mathcal{E}, \mu)$ assolutamente continua con densità f_X . Sia g misurabile da $(\mathbb{E}, \mathcal{E})$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Se $g(X)$ ammette media allora vale:*

$$\mathbb{E}[g(X)] = \int_E f_X(x) g(x) d\mu(x)$$

Dimostrazione. Sia $Y := g(X)$, la tesi diventa:

$$\int_{\Omega} Y(\omega) dP(\omega) = \int_E f_X(x) g(x) d\mu(x)$$

Sia $P_X := P(X^{-1})$ la legge di X , per la regola di composizione si ha:

$$\int_{\Omega} Y(\omega) dP(\omega) = \int_E g(x) dP_X(x)$$

la tesi diventa:

$$\int_E g(x) dP_X(x) = \int_E f_X(x) g(x) d\mu(x) \quad (3.8)$$

Per dimostrare la (3.8) si procede applicando il metodo standard su g : prima si considera il caso g funzione indicatrice, poi g semplice infine g positiva.

CASO g funzione indicatrice: sia $g := \mathbb{1}_A$ con $A \in \mathcal{E}$ allora si ha

$$\int_E g(x) dP_X(x) = P(X \in A) = \int_E f_X(x) \mathbb{1}_A d\mu(x) = \int_E f_X(x) g(x) d\mu(x)$$

dove la seconda uguaglianza segue dalla definizione di funzione di densità.

CASO g semplice: segue immediatamente dal caso precedente e dal fatto che i due integrali della (3.8) sono lineari in g .

CASO g positiva: poiché entrambi gli argomenti degli integrali dell'equazione (3.8) sono monotoni in g possiamo applicare il teorema di convergenza monotona a destra e sinistra ad una successione monotona di funzioni semplici che approssimano g . $\square$

3.4 Disuguaglianza di Markov

Sia X una variabile aleatoria non negativa che ammette media finita. Siamo interessati a stimare la probabilità che X sia maggiore di una certa costante $c > 0$, conoscendo il valore medio $\mathbb{E}[X]$

$$\mathbb{E}[X] \geq \mathbb{E}[X \mathbf{1}_{X \geq c}] \geq \mathbb{E}[c \mathbf{1}_{X \geq c}] = cP(X \geq c) \quad (3.9)$$

quindi si ottiene la seguente disuguaglianza di Markov:

Proprietà 3.10. *Per ogni v.a. X quasi certamente non-negativa e per ogni $c > 0$ vale:*

$$P(X > c) \leq \frac{\mathbb{E}[X]}{c} \quad (3.10)$$

Se g è una funzione positiva crescente con $\mathbb{E}[g(X)] < \infty$ allora l'argomento utilizzato in (3.9) può essere ripetuto su $g(X)$:

$$\mathbb{E}[g(X)] \geq \mathbb{E}[g(X)\mathbf{1}_{g(X) \geq g(c)}] \geq \mathbb{E}[g(c)\mathbf{1}_{g(X) \geq g(c)}] = g(c)P(X \geq c)$$

da cui si ottiene:

Proprietà 3.11. *Per ogni v.a. X , per ogni $c > 0$ e per ogni funzione g non-negativa, debolmente crescente e con $g(c) > 0$ vale:*

$$P(X > c) \leq \frac{\mathbb{E}[g(X)]}{g(c)} \quad (3.11)$$

la formula (3.11) ha il vantaggio di essere valida anche nei casi in cui la v.a. X non sia positiva o non ammetta media finita.

3.5 Disuguaglianza di Jensen

La disuguaglianza di Jensen è una disuguaglianza che riguarda i valori attesi di funzioni convesse.

Definizione 3.1. Un insieme $C \subseteq \mathbb{R}^n$ si dice convesso se per ogni $x, y \in C$ e per ogni $\mu, \lambda \in (0, 1)$ con $\mu + \lambda = 1$ si ha $z := \mu x + \lambda y \in C$.

L'intersezione di insiemi convessi è ancora un insieme convesso. Dato un insieme $A \in \mathbb{R}^n$ il più piccolo insieme convesso che contiene A si dice l'involuppo convesso di A .

Definizione 3.2. Dato C un sottoinsieme convesso di $\mathbb{R}^n$. Una funzione $\Phi : C \rightarrow \mathbb{R}$ si dice convessa se per ogni $x, y \in C$ e per ogni $\mu, \lambda \in (0, 1)$ con $\mu + \lambda = 1$ si ha:

$$\Phi(\mu x + \lambda y) \leq \mu \Phi(x) + \lambda \Phi(y)$$

Esercizio 3.7. *Una funzione convessa Φ su un insieme aperto C è continua.*

Valore atteso in $\mathbb{R}^n$

Sia $X = (X_1, X_2, \dots, X_n)$ una variabile aleatoria a valori in $\mathbb{R}^n$ tale che $\mathbb{E}[|X|] < \infty$. Il valore atteso di X si definisce nel seguente modo.

$$\mathbb{E}[X] := (\mathbb{E}[X_1], \mathbb{E}[X_2], \dots, \mathbb{E}[X_n])$$

Il valore atteso definito in questo modo è lineare, se A è una matrice possiamo definire il prodotto matrice per vettore nel modo usuale e si verifica facilmente che vale:

$$\mathbb{E}[AX^t] = A\mathbb{E}[X^t]$$

questo ci assicura che il valore atteso è tensoriale ovvero se O è la matrice associata ad un cambio di base allora vale

$$O^{-1}\mathbb{E}[OX^t] = \mathbb{E}[X^t]$$

Valore atteso su un insieme convesso di $\mathbb{R}^n$

Sia $C \subseteq \mathbb{R}^n$ convesso. Allora valgono i seguenti teoremi di separazione con iperpiani (in inglese *Hyperplane separation theorem*):

Teorema 3.12 (versione 1). *Sia C un chiuso convesso di $\mathbb{R}^n$, sia P un punto di $\mathbb{R}^n$ che non appartiene a C . Esiste un iperpiano che separa in senso stretto P e C ovvero esiste un'applicazione affine ψ con $\psi(x) = a \cdot x + b$, $a \in \mathbb{R}^n$ e $b \in \mathbb{R}$ tale che:*

$$\psi(P) < 0 \quad e \quad \psi(x) > 0 \quad \forall x \in C \quad (3.12)$$

Teorema 3.13 (versione 2). *Sia C un convesso di $\mathbb{R}^n$, sia P un punto di $\mathbb{R}^n$ che non appartiene a C . Esiste un iperpiano che separa in senso debole P e C ovvero esiste un'applicazione affine ψ con $\psi(x) = a \cdot x + b$, $b \in \mathbb{R}$, $a \in \mathbb{R}^n$ $a \neq 0$ tale che:*

$$\psi(P) \leq 0 \quad e \quad \psi(x) \geq 0 \quad \forall x \in C \quad (3.13)$$

Nella prima versione l'ipotesi C insieme chiuso permette di escludere il caso P appartiene alla frontiera di C e quindi da luogo alle disuguaglianze strette. Il lemma 3.13 vale per C non chiuso, in questo caso se P appartiene alla frontiera di C le disuguaglianze non possono essere strette.

Teorema 3.14. *Sia $C \subseteq \mathbb{R}^n$ un insieme convesso. Sia X una variabile aleatoria a valori in C di media finita. Allora vale $\mathbb{E}[X] \in C$.*

Dimostrazione. Procederemo per induzione su n .

CASO $n = 1$. Se $n = 1$ allora gli insiemi convessi sono intervalli (per esempio (a, b) oppure $[a, \infty)$, la tesi segue banalmente dalla monotonia dell'integrale: $X > a$ implica $\mathbb{E}[X] > a$, $X \geq a$ implica $\mathbb{E}[X] \geq a$ e così via

PASSO INDUTTIVO: vale per ogni $k < n$ allora vale per n .

Sia k la dimensione del più piccolo spazio affine W tale che $P(X \in W) = 1$. se $k < n$ allora la tesi segue dall'ipotesi induttiva poiché $C \cap W$ è ancora un insieme convesso.

Resta il caso $k = n$ cioè l'immagine di X non è contenuta (quasi certamente)

in nessun iperpiano.

Procediamo per assurdo sia $P := \mathbb{E}[X]$ e sia ψ la funzione affine definita dal teorema 3.13. Vale $\psi[\mathbb{E}[X]] \leq 0$ e $\psi(X) \geq 0$ inoltre poiché l'immagine di X non è contenuto (quasi certamente) nell'iperpiano $\psi = 0$ allora $P(\psi(X) > 0) > 0$. Dall'esercizio 3.6 segue $\mathbb{E}[\psi(X)] > 0$ mentre dalla linearità di ψ si ottiene $\psi(\mathbb{E}[X]) = \mathbb{E}[\psi(X)] \leq 0$ $\square$

Teorema 3.15 (Disuguaglianza di Jensen). *Sia X una v.a. a valori in $C \subseteq \mathbb{R}^n$, sia Φ un'applicazione convessa su C . Se X e $\Phi(X)$ ammettono media finita allora vale:*

$$\Phi(\mathbb{E}[X]) \leq \mathbb{E}[\Phi(X)]$$

Dimostrazione. Sia $Y := (X, \Phi(X))$ variabile aleatoria a valori in $\mathbb{R}^{n+1}$ e sia $D \subseteq \mathbb{R}^{n+1}$ l'insieme seguente:

$$D := \{(x_1, \dots, x_n, x_{n+1}) \in \mathbb{R}^{n+1} | x_{n+1} \geq \Phi((x_1, \dots, x_n))\}$$

Dalla convessità di Φ segue la convessità dell'insieme D . La tesi segue immediatamente dal teorema 3.14:

$$\mathbb{E}[Y] \in D \quad \implies \quad \mathbb{E}[\Phi(X)] \geq \Phi(\mathbb{E}[X])$$

$\square$

3.6 Spazi L^p .

In questo paragrafo vedremo due applicazioni della disuguaglianza di Jensen: la monotonia degli spazi L^p (3.14) e il teorema di Holder.

Definizione 3.3. Sia $(\Omega, \mathcal{H}, \mu)$ uno spazio di misura. Sia X misurabile da $(\Omega, \mathcal{H})$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ e $p \geq 1$ diremo che X è in $\mathcal{L}^p$ se:

$$\int_{\Omega} |X|^p d\mu < \infty$$

Indicheremo inoltre con $L^p := \mathcal{L}^p / \mathcal{R}$ lo spazio che si ottiene quozientando per la relazione di equivalenza:

$$X \sim_{\mathcal{R}} Y \quad \iff \quad X = Y \text{ q.c.}$$

In maniera simile diremo che X è in $\mathcal{L}^\infty$ se

$$\|X\|_\infty := \inf \{x \geq 0 | \mu(|X| > x) = 0\} < \infty$$

Poiché $x + y \leq 2 \max(x, y)$ allora $|X + Y|^p \leq 2^p \max(|X|^p, |Y|^p) \leq 2^p (|X|^p + |Y|^p)$ da cui si deduce che L^p è uno spazio vettoriale. Definiamo ora la

seguito funzione che (in seguito grazie alla disuguaglianza di Minkowski) sarà una norma per gli spazi L^p :

$$\|X\|_p := \mathbb{E}[|X|^p]^{\frac{1}{p}}$$

3.6.1 Monotonia degli spazi L^p

Proposizione 3.16. *Sia $(\Omega, \mathcal{H}, P)$ uno spazio di probabilità. Se $1 < p < q$ allora*

$$L^p(\Omega, \mathcal{H}, P) \supseteq L^q(\Omega, \mathcal{H}, P) \quad (3.14)$$

$$\|X\|_p \leq \|X\|_q \quad (3.15)$$

Si può dimostrare che l'inclusione (3.14) vale anche se lo spazio di misura non è uno spazio di probabilità ma solo uno spazio di misura finita. Se μ è una misura finita e $1 \leq p \leq q$ allora vale:

$$\|X\|_{L^p(\Omega, \mathcal{H}, \mu)} \leq C \|X\|_{L^q(\Omega, \mathcal{H}, \mu)}$$

con $C = (\mu(\Omega))^{\frac{1}{p} - \frac{1}{q}}$

Ora dimostreremo la proposizione precedente nel caso $\|X\|_p, \|X\|_q < \infty$ lasciando per esercizio il caso $\|X\|_p = \infty$.

Dimostrazione. Sia X misurabile da $(\Omega, \mathcal{H}, P)$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. La tesi è:

$$\begin{aligned} \mathbb{E}[|X|^p]^{\frac{1}{p}} &\leq \mathbb{E}[|X|^q]^{\frac{1}{q}} \\ &\Downarrow \\ \mathbb{E}[|X|^p]^{\frac{1}{p}} &\leq \mathbb{E}[(|X|^p)^{\frac{q}{p}}]^{\frac{1}{q}} \end{aligned}$$

Applicando ora la disuguaglianza di Jensen alla funzione convessa $\Phi(x) = |x|^{\frac{q}{p}}$ si ottiene $\Phi(\mathbb{E}[|X|^p]) \leq \mathbb{E}[\Phi(|X|^p)]$ quindi

$$\begin{aligned} \mathbb{E}[|X|^p]^{\frac{q}{p}} &\leq \mathbb{E}[(|X|^p)^{\frac{q}{p}}] \\ &\Downarrow \\ \mathbb{E}[|X|^p]^{\frac{1}{p}} &\leq \mathbb{E}[|X|^q]^{\frac{1}{q}} \end{aligned}$$

□

3.6.2 Disuguaglianze di Holder e Minkowski

Enunciamo ora le disuguaglianze di Holder e Minkowski. Per completezza ci sono le dimostrazioni ma non saranno richieste per l'esame.

Teorema 3.17. *Siano X e Y variabili aleatorie. Siano $p \geq 1$ e $q \geq 1$ tali che $\frac{1}{p} + \frac{1}{q} = 1$. Se $X \in L^p$ e $Y \in L^q$ allora:*

$$\mathbb{E}[XY] \leq \mathbb{E}[|XY|] \leq \|X\|_p \|Y\|_q$$

Dimostrazione. Senza perdere in generalità possiamo supporre X e Y non negative. Questo semplificherà le notazioni. Sia $c := \mathbb{E}[X^p]$,

$$\mathbb{E}[XY] = c \mathbb{E} \left[\frac{X^p}{c} \cdot \frac{Y}{X^{p-1}} \right] = c \mathbb{E} \left[\frac{X^p}{c} \cdot \left(\frac{Y^q}{X^p} \right)^{\frac{1}{q}} \right] \quad (3.16)$$

Se poniamo $f := \frac{X^p}{c}$ allora f è non negativa e vale $\mathbb{E}[f] = 1$ quindi f è la funzione di densità di una misura di probabilità Q , sia $\mathbb{E}_Q$ il valore atteso associato a Q (ovvero $\mathbb{E}_Q[Z] = \mathbb{E}[f \cdot Z]$). Applicando la disuguaglianza di Jensen alla funzione concava $\Phi(x) = x^{\frac{1}{q}}$ la (3.16) diventa

$$\mathbb{E}[XY] = c \mathbb{E}_Q \left[\left(\frac{Y^q}{X^p} \right)^{\frac{1}{q}} \right] \leq c \mathbb{E}_Q \left[\left(\frac{Y^q}{X^p} \right) \right]^{\frac{1}{q}} = c \mathbb{E} \left[\frac{X^p}{c} \cdot \left(\frac{Y^q}{X^p} \right) \right]^{\frac{1}{q}} = \|X\|_p \|Y\|_q$$

□

Teorema 3.18. *Siano X e Y variabili aleatorie. Sia $p \geq 1$. Se $X, Y \in L^p$ allora:*

$$\|X\|_p + \|Y\|_p \geq \|X + Y\|_p \quad (3.17)$$

Dimostrazione. Se $p = 1$ la tesi è immediata poiché vale $|x + y| \leq |x| + |y|$. Sia $p > 1$, e q tale che $\frac{1}{p} + \frac{1}{q} = 1$. Senza perdere in generalità supponiamo X e Y non negative.

$$\begin{aligned} \mathbb{E}[|X + Y|^p] &= \mathbb{E}[X|X + Y|^{p-1}] + \mathbb{E}[Y|X + Y|^{p-1}] \\ &\leq \|X\|_p \|(X + Y)^{p-1}\|_q + \|Y\|_p \|(X + Y)^{p-1}\|_q \\ &\leq (\|X\|_p + \|Y\|_p) \mathbb{E}[(X + Y)^p]^{\frac{1}{q}} \end{aligned}$$

$$\mathbb{E}[|X + Y|^p]^{\frac{1}{p}} \leq \|X\|_p + \|Y\|_p$$

□

La disuguaglianza (3.17) ci dice che $\|\cdot\|_p$ soddisfa la disuguaglianza triangolare ed è facile ora verificare che è una norma per lo spazio L^p .

3.6.3 Completezza

Una successione di variabili aleatorie $\{X_n\}_{n \in \mathbb{N}}$ si dice di Cauchy rispetto alla norma $\|\cdot\|_{L^p}$ se per ogni $\epsilon > 0$ esiste $\bar{n}$ tale che per ogni $n, m \geq \bar{n}$ vale $\|X_n - X_m\|_p < \epsilon$.

Proposizione 3.19. *Le successioni di Cauchy in L^p convergono. Ovvero se $\{X_n\}_{n \in \mathbb{N}}$ è una successione di Cauchy in L^p allora esiste $X \in L^p$ tale che $\|X_n - X\|_p \xrightarrow{n \rightarrow \infty} 0$.*

Dimostrazione. La dimostrazione consta di due parti prima occorre determinare la v.a. X e poi mostrare che $\|X_n - X\|_p \xrightarrow{n \rightarrow \infty} 0$.

Poiché la successione $\{X_n\}_{n \in \mathbb{N}}$ è una successione di Cauchy allora per ogni $k \in \mathbb{N}$ esiste n_k tale che

$$\forall n, m \geq n_k \quad \mathbb{E}[|X_n - X_m|^p] < \frac{1}{2^{2pk}}$$

Mostreremo ora che $\{X_{n_k}\}_{k \in \mathbb{N}}$ converge quasi certamente e indicheremo con X tale limite. Sia

$$A_k := \left\{ \omega \in \Omega : |X_{n_k}(\omega) - X_{n_{k-1}}(\omega)| > \frac{1}{2^k} \right\}$$

applicando la disuguaglianza di Markov si ottiene

$$P(A_k) = P\left(|X_{n_k} - X_{n_{k-1}}|^p > \frac{1}{2^{pk}}\right) \leq \frac{\mathbb{E}[|X_{n_k} - X_{n_{k-1}}|^p]}{\frac{1}{2^{pk}}} \leq \frac{\frac{1}{2^{2pk}}}{\frac{1}{2^{pk}}} = \frac{1}{2^{pk}}$$

La somma $\sum_k P(A_k)$ è finita. Possiamo applicare il primo lemma di Borel-Cantelli che ci dà: per quasi ogni $\omega \in \Omega$ esiste $\bar{k}$ tale che per ogni $k \geq \bar{k}$ vale $\omega \notin A_k$. Poiché

$$X_{n_k}(\omega) = X_0(\omega) + \sum_{i=1}^k (X_{n_i}(\omega) - X_{n_{i-1}}(\omega))$$

e la serie è assolutamente convergente ovvero: $\sum_{i=1}^k |X_{n_i}(\omega) - X_{n_{i-1}}(\omega)| \leq \sum_{i=1}^k \frac{1}{2^i} < 1$ allora $\{X_{n_k}\}_{k \in \mathbb{N}}$ ammette limite quasi certo, sia X tale limite. Dobbiamo ora mostrare che $\mathbb{E}[|X_n - X|^p] \xrightarrow{n \rightarrow \infty} 0$.

$$\mathbb{E}[|X_n - X|^p] = \mathbb{E}[\liminf_k |X_n - X_{n_k}|^p] \leq \liminf_k \mathbb{E}[|X_n - X_{n_k}|^p]$$

per $n > n_{\bar{k}}$ e $k > \bar{k}$ vale $\mathbb{E}[|X_n - X_{n_k}|^p] < \frac{1}{2^{2pk}}$ quindi

$$\forall n > n_{\bar{k}} \quad \mathbb{E}[|X_n - X|^p] < \frac{1}{2^{2p\bar{k}}}$$

□

Conclusione spazi L^p

In conclusione la disuguaglianza di Minkowski e la completezza ci permettono di dire che gli spazi L^p sono spazi di Banach, inoltre si verifica facilmente che lo spazio L^2 è uno spazio di Hilbert con prodotto scalare $\langle X, Y \rangle := E[XY]$

Capitolo 4

Indipendenza

In questo capitolo tratteremo l'indipendenza. A meno che non sia specificato diversamente, indicheremo con $(\Omega, \mathcal{H}, P)$ lo spazio di probabilità su cui si trovano gli eventi o sono definite le variabili aleatorie.

4.1 Probabilità condizionata

La probabilità condizionata dovrebbe essere già nota dal corso del primo anno. Potete testare le vostre competenze con l'esercizio A.4.7, esercizio che richiede la conoscenza dei concetti di base di probabilità condizionata e indipendenza.

In questo paragrafo richiamerò solo le definizioni e le proprietà principali, affinché possiate leggerle ed assicurarvi di ricordare tutto.

Definizione 4.1. Dati due eventi A e B su $(\Omega, \mathcal{H}, P)$ con $P(A) > 0$ si dice probabilità condizionata di B dato A (o sapendo A) la quantità:

$$P(B|A) := \frac{P(B \cap A)}{P(A)}$$

Proposizione 4.1. Dato un evento $A \in \mathcal{H}$ con $P(A) > 0$ la funzione Q da $\mathcal{H}$ in $[0, 1]$ data da:

$$Q(B) = P_A(B) := P(B|A) \quad \forall B \in \mathcal{H}$$

è una misura di probabilità ovvero la terna $(\Omega, \mathcal{H}, Q)$ è a sua volta uno spazio di probabilità.

Nelle ipotesi della proposizione precedente indicheremo con $\mathbb{E}_{P_A}$ la speranza rispetto alla misura di probabilità P_A (cioé $\mathbb{E}_{P_A}[X] = \int X \, dP_A$)

Proposizione 4.2 (regola della catena). *Siano $(A_1 \dots A_n)$ eventi con $P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0$ vale:*

$$P(A_1 \cap \dots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \dots P(A_n|A_1 \cap A_2 \dots \cap A_{n-1})$$

Di solito la proposizione precedente viene utilizzata con $n = 2$ e diventa:

$$P(A \cap B) = P(A) \cdot P(B|A)$$

Proposizione 4.3. *Sia $(A_i)_{i \in I}$ una partizione di Ω costituita da eventi di misura positiva con I insieme al più numerabile vale:*

$$P(B) = \sum_{i \in I} P(B|A_i)P(A_i)$$

La proposizione precedente viene spesso utilizzata con la partizione $\{A, A^c\}$ per la quale si ha:

$$P(B) = P(B|A)P(A) + P(B|A^c)P(A^c)$$

Teorema 4.4 (formula di Bayes). *Sia $(A_i)_{i \in I}$ una partizione di Ω costituita da eventi di misura positiva con I insieme al più numerabile, sia B un ulteriore evento con probabilità positiva, per ogni $j \in I$ vale:*

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i \in I} P(B|A_i)P(A_i)}$$

4.2 Indipendenza tra eventi

Introdurremo dapprima l'indipendenza di due eventi e poi quella di una famiglia generica di eventi.

4.2.1 Indipendenza di due eventi

Intuitivamente due eventi E ed F sono indipendenti se il verificarsi di uno non modifica il nostro grado di fiducia nell'altro. Questo vuol dire che se $P(F) > 0$ allora E ed F sono indipendenti se

$$P(E|F) = P(E) \tag{4.1}$$

ovvero il fatto che si sia verificato F non ha modificato la nostra fiducia nell'evento E . Per $P(F) > 0$ le equazioni (4.1) e (4.2) sono equivalenti tuttavia per la definizione di indipendenza è preferibile la seconda perché mette in evidenza la simmetria tra gli eventi E ed F e soprattutto perché ha senso anche quando $P(F) = 0$.

Definizione 4.2. Diciamo che due eventi E ed F sono indipendenti se e solo se vale

$$P(E \cap F) = P(E)P(F) . \quad (4.2)$$

Una conseguenza di (4.2) è che gli eventi di misura zero o uno sono indipendenti da qualunque altro evento. Notare che gli eventi di misura zero e uno sono tutti e soli gli eventi indipendenti da tutti gli altri. Si può verificare che E ed F sono indipendenti se e solo se lo sono E ed F^c .

Esercizio 4.1. *Dimostrare che un evento A è indipendente da tutti gli altri eventi se e solo se vale $P(A) \in \{0, 1\}$.*

4.2.2 Indipendenza di una famiglia generica di eventi

Definizione 4.3. Una famiglia di eventi $\{E_i\}_{i \in I}$ si dice una famiglia di eventi indipendenti se e solo se per ogni $J \subseteq I$ con $|J| < \infty$ vale

$$P\left(\bigcap_{j \in J} E_j\right) = \prod_{j \in J} P(E_j) . \quad (4.3)$$

Dunque una famiglia di eventi $\{E_i\}_{i \in I}$ si dice una famiglia di eventi indipendenti se e solo se per tutti i sottoinsiemi finiti vale la formula (4.3). Vediamo ora una proprietà che sarà utile nei prossimi paragrafi.

Proprietà 4.5. *Siano $E_1, E_2, \dots, E_n$ eventi indipendenti. Se $F_1, F_2, \dots, F_n$ sono eventi tali che per ogni i o $F_i = E_i$ oppure $F_i = E_i^c$ allora anche $F_1, F_2, \dots, F_n$ sono eventi indipendenti.*

Dimostrazione. Vedi esercizio A.4.1 □

Con il seguente esercizio invece mostreremo che indipendenza a due a due non implica indipendenza.

Esercizio 4.2. *Consideriamo il lancio di due dadi. Consideriamo i seguenti eventi:*

$A :=$ “il risultato del primo lancio è pari”

$B :=$ “il risultato del secondo lancio è pari”

$C :=$ “il risultato della somma è pari”

Dimostrare che gli eventi $\{A, B\}$ sono indipendenti, gli eventi $\{A, C\}$ sono indipendenti, gli eventi $\{B, C\}$ sono indipendenti, ma gli eventi $\{A, B, C\}$ non sono indipendenti

Svolgimento. Calcolando le probabilità degli eventi A , B e C e delle loro intersezioni si ottiene:

$P(A) = 1/2$, $P(B) = 1/2$, $P(C) = 1/2$, $P(A \cap B) = 1/4$, $P(A \cap C) = 1/4$, $P(B \cap C) = 1/4$, $P(A \cap B \cap C) = 1/4$ dunque

$$P(A \cap B) = P(A)P(B) \quad P(A \cap C) = P(A)P(C)$$

$$P(B \cap C) = P(B)P(C)$$

$$P(A \cap B \cap C) \neq P(A)P(B)P(C)$$

□

4.3 Indipendenza di σ -algebre

Definizione 4.4. Diciamo che una famiglia di σ -algebre $\{\mathcal{F}_i\}_{i \in I}$ si dice una famiglia di σ -algebre indipendenti se e solo se per ogni $J \subseteq I$ con $|J| < \infty$ vale

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j) \quad \forall (A_j)_{j \in J} \in (\mathcal{F}_j)_{j \in J} \quad (4.4)$$

Dunque una famiglia di σ -algebre $\{\mathcal{F}_i\}_{i \in I}$ si dice una famiglia di σ -algebre indipendenti se e solo se per tutti i sottoinsiemi finiti J vale la formula (4.4).

Proposizione 4.6. Sia $\{E_i\}_{i \in I}$ una famiglia di eventi e siano $\mathcal{F}_i := \sigma(E_i)$ le σ -algebre generate. Sono cose equivalenti:

- (a) $\{E_i\}_{i \in I}$ è una famiglia di eventi indipendenti
- (b) $\{\mathcal{F}_i\}_{i \in I}$ è una famiglia di σ -algebre indipendenti

Dimostrazione. Osserviamo che $\mathcal{F}_i = \{\emptyset, E_i, E_i^c, \Omega\}$. Allora un'implicazione è ovvia e l'altra è una conseguenza dell'esercizio 4.1 e della proprietà 4.5. □

4.4 Indipendenza di variabili aleatorie

Definizione 4.5. Sia $\{X_i\}_{i \in I}$ una famiglia di variabili, e per ogni $i \in I$ sia $\mathcal{F}_i := \sigma(X_i)$. Diciamo che una famiglia di variabili aleatorie $\{X_i\}_{i \in I}$ è una famiglia di variabili aleatorie indipendenti se $\{\mathcal{F}_i\}_{i \in I}$ è una famiglia di σ -algebre indipendenti.

Proposizione 4.7. Per ogni $i \in I$ sia X_i variabile aleatoria a valori in $(E_i, \mathcal{E}_i)$. Le variabili aleatorie $\{X_i\}_{i \in I}$ sono indipendenti se e solo se per ogni $J \subseteq I$ finito e per ogni $(A_j)_{j \in J} \in (\mathcal{E}_j)_{j \in J}$ vale:

$$P\left(\bigcap_{j \in J} \{X_j \in A_j\}\right) = \prod_{j \in J} P(X_j \in A_j)$$

Proposizione 4.8. Sia $\{X_i\}_{i \in I}$ una famiglia di v.a. a valori in $(E_i, \mathcal{E}_i)_{i \in I}$. Siano ϕ_i applicazioni misurabili da $(E_i, \mathcal{E}_i)$ in $(F_i, \mathcal{F}_i)$. Sia infine $Y_i := \phi_i(X_i)$ per ogni $i \in I$. Se $\{X_i\}_{i \in I}$ è una famiglia di v.a. indipendenti allora anche le v.a. $\{Y_i\}_{i \in I}$ costituiscono una famiglia di v.a. indipendenti

Dimostrazione. La dimostrazione segue immediatamente da $\sigma(Y_i) \subseteq \sigma(X_i)$. □

Esercizio 4.3. Con le notazioni della proposizione precedente ($Y_i = \phi_i(X_i)$). Costruire un esempio in cui le v.a. $\{Y_i\}_{i \in I}$ siano indipendenti mentre le v.a. $\{X_i\}_{i \in I}$ non siano indipendenti.

4.5 Indipendenza e p-system

Procederemo per gradi, analizzeremo prima l'indipendenza tra un evento ed una σ -algebra generata da un p -system, poi studieremo l'indipendenza tra due σ -algebre ed infine studieremo il problema per una famiglia qualsiasi di σ -algebre.

Proposizione 4.9. Sia A un evento, $\mathcal{A}$ un p -system e $\mathcal{F} := \sigma(\mathcal{A})$. Se A è indipendente da tutti gli eventi di $\mathcal{A}$ allora A è indipendente anche da tutti gli eventi di $\mathcal{F}$.

Dimostrazione. Per cominciare poiché Ω è indipendente da tutti gli eventi possiamo assumere $\Omega \in \mathcal{A}$. Per il lemma di Dynkin vale:

$$\mathcal{F} := \sigma(\mathcal{A}) = d(\mathcal{A})$$

procederemo ora in maniera standard, dimostrando che gli eventi che soddisfano la proprietà richiesta sono un d-system e contengono $\mathcal{A}$. Sia

$$\mathcal{G} := \{B \in \mathcal{F} | P(A \cap B) = P(A) \cdot P(B)\}$$

Poiché A è indipendente dagli eventi di $\mathcal{A}$ allora $\mathcal{G} \supseteq \mathcal{A}$. Resta da dimostrare che $\mathcal{G}$ è un d-system.

(i) $\Omega \in \mathcal{A} \supseteq \mathcal{G}$

(ii) $B \in \mathcal{G}$ implica $P(A \cap B) = P(A) \cdot P(B)$.

$$P(A \cap B^c) = P(A) - P(A \cap B) = P(A)(1 - P(B)) = P(A) \cdot P(B^c)$$

quindi anche B^c appartiene a $\mathcal{G}$

(iii) Siano $\{B_n\}_{n \in \mathbb{N}}$ disgiunti e in $\mathcal{G}$, quindi $P(A \cap B_n) = P(A) \cdot P(B_n)$

$$\begin{aligned} P\left(A \cap \left(\bigcup_n B_n\right)\right) &= P\left(\bigcup_n (A \cap B_n)\right) = \sum_n P(A \cap B_n) \\ &= \sum_n P(A)P(B_n) = P(A) \sum_n P(B_n) = \\ &= P(A)P\left(\bigcup_n B_n\right) \end{aligned}$$

dunque anche $\bigcup_n B_n$ appartiene a $\mathcal{G}$. □

Proposizione 4.10. Siano $\mathcal{F}_1$ e $\mathcal{F}_2$ due σ -algebre e siano $\mathcal{A}_1$ e $\mathcal{A}_2$ due p -system tali che $\mathcal{F}_i := \sigma(\mathcal{A}_i)$. Se gli eventi di $\mathcal{A}_1$ sono indipendenti dagli eventi di $\mathcal{A}_2$ allora le σ -algebre $\mathcal{F}_1$ e $\mathcal{F}_2$ sono indipendenti.

Dimostrazione. Applichiamo dapprima la proposizione 4.9 ad $A \in \mathcal{A}_1$ e $\mathcal{A} = \mathcal{A}_2$ cosicché si ottiene che gli elementi di $\mathcal{A}_1$ sono indipendenti da quelli di $\mathcal{F}_2$. Applicchiamo ora la proposizione 4.9 ad $A \in \mathcal{F}_2$ e $\mathcal{A} = \mathcal{A}_1$ cosicché si ottiene che gli elementi di $\mathcal{F}_1$ sono indipendenti da quelli di $\mathcal{F}_2$. $\square$

Proposizione 4.11. *Sia $\{\mathcal{F}_i\}_{i \in I}$ una famiglia di σ -algebre, sia $\{\mathcal{A}_i\}_{i \in I}$ una famiglia di p-system tali che $\mathcal{F}_i := \sigma(\mathcal{A}_i)$. Se per ogni $J \subseteq I$ finito e per ogni $(A_j)_{j \in J} \in \times_{j \in J} \mathcal{A}_j$ vale*

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j)$$

allora $\{\mathcal{F}_i\}_{i \in I}$ è una famiglia di σ -algebre indipendenti.

Dimostrazione. Innanzitutto l'insieme Ω è indipendente da qualunque altro evento, quindi possiamo assumere senza perdere in generalità che Ω appartenga ai p-system $\mathcal{A}_i$. La definizione di famiglia di σ -algebre indipendenti si basa sulla verifica della condizione (4.4) su insiemi finiti di indici quindi senza perdere in generalità possiamo assumere $I = \{1, 2, \dots, n\}$ finito. Inoltre per la proposizione precedente il caso $n = 2$ è vero e per induzione possiamo assumere che sia vero per $n - 1$.

Sia dunque $n \geq 3$.

Ipotesi $\forall (A_i)_{i \in I} \in \times_{i \in I} (\mathcal{A}_i)$ vale $P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot \dots \cdot P(A_n)$

Tesi $\forall (B_i)_{i \in I} \in \times_{i \in I} (\mathcal{F}_i)$ vale $P(B_1 \cap \dots \cap B_n) = P(B_1) \cdot \dots \cdot P(B_n)$

Per induzione sappiamo che $P(B_1 \cap \dots \cap B_{n-1}) = P(B_1) \cdot \dots \cdot P(B_{n-1})$ quindi la tesi diventa:

$$P((B_1 \cap \dots \cap B_{n-1}) \cap B_n) = P(B_1 \cap \dots \cap B_{n-1}) \cdot P(B_n) \quad (4.5)$$

A questo punto l'idea è di dimostrare l'uguaglianza precedente utilizzando la proposizione 4.10. Definiremo ora le σ -algebre $\mathcal{G}_1$ e $\mathcal{G}_2$ e i p-system $\mathcal{B}_1$ e $\mathcal{B}_2$ a cui applicare la proposizione 4.10.

$$\mathcal{G}_1 := \sigma(\mathcal{F}_1, \dots, \mathcal{F}_{n-1}) \quad \mathcal{G}_2 := \mathcal{F}_n$$

$$\mathcal{B}_1 := \{A_1 \cap \dots \cap A_{n-1} \mid \forall i \ A_i \in \mathcal{A}_i\} \quad \mathcal{B}_2 = \mathcal{A}_n$$

Chiaramente $\mathcal{B}_1$ e $\mathcal{B}_2$ sono p-system e $\sigma(\mathcal{B}_2) = \mathcal{G}_2$.

Mentre $\sigma(\mathcal{B}_1) = \sigma(\mathcal{A}_1, \dots, \mathcal{A}_{n-1}) = \sigma(\mathcal{F}_1, \dots, \mathcal{F}_{n-1}) = \mathcal{G}_1$ Quindi per la proposizione 4.10 vale la (4.5). $\square$

4.6 Indipendenza a blocchi

Teorema 4.12. Sia $\{\mathcal{H}_i\}_{i \in I}$ una famiglia di σ -algebre indipendenti. Sia $\{I_j\}_{j \in J}$ una partizione di I . Se poniamo per ogni $j \in J$:

$$\mathcal{G}_j := \sigma(\{\mathcal{H}_i | i \in I_j\})$$

allora $\{\mathcal{G}_j\}_{j \in J}$ è una famiglia di σ -algebre indipendenti.

Dimostrazione. Per ogni $j \in J$ sia:

$$\mathcal{A}_j := \left\{ \bigcap_{i \in \Gamma} A_i \mid \Gamma \subseteq I_j \text{ finito e } \forall i \in \Gamma \quad A_i \in \mathcal{H}_i \right\}$$

Chiaramente $\mathcal{A}_j$ è un p -system e vale $\mathcal{G}_j = \sigma(\mathcal{A}_j)$. Se mostriamo che $\{\mathcal{A}_j\}_{j \in J}$ è una famiglia di p -system indipendenti allora per la proposizione 4.11 abbiamo finito. Mostriamo ora che $\{\mathcal{A}_j\}_{j \in J}$ è una famiglia di p -system indipendenti.

Siano $j_1, \dots, j_n \in J$ distinti e consideriamo gli eventi $B_k \in \mathcal{A}_{j_k}$ con $k \in \{1, \dots, n\}$. Dobbiamo mostrare che gli eventi $\{B_k\}_{1 \leq k \leq n}$ sono indipendenti. Poiché $B_k \in \mathcal{A}_{j_k}$ allora esistono interi m_k ed eventi $A_{i_{k,h}} \in \mathcal{H}_{i_{k,h}}$ tali che:

$$B_k = A_{i_{k,1}} \cap \dots \cap A_{i_{k,m_k}}$$

dove gli indici $i_{h,k}$ sono tutti distinti al variare di k e h .

$$\begin{aligned} P(B_1 \cap \dots \cap B_n) &= P\left(\bigcap_{\substack{k \in \{1, \dots, n\} \\ h \in \{1, \dots, m_k\}}} A_{i_{k,h}}\right) = \prod_{\substack{k \in \{1, \dots, n\} \\ h \in \{1, \dots, m_k\}}} P(A_{i_{k,h}}) \\ &= \prod_{k \in \{1, \dots, n\}} \left(\prod_{h \in \{1, \dots, m_k\}} P(A_{i_{k,h}}) \right) = P(B_1) \cdot \dots \cdot P(B_n) \end{aligned}$$

□

4.7 Varianza, covarianza ed indipendenza

Se X è una variabile aleatoria in L^2 si definisce varianza di X la grandezza:

$$VAR(X) := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

La varianza di una variabile aleatoria indica la sua *variabilità* ed è nulla se e solo se la variabile aleatoria è quasi certamente costante.

Dalla disuguaglianza di Markov (3.10) è possibile ricavare la disuguaglianza di Chebyshev per variabili aleatorie in L^2 .

Proposizione 4.13. *Sia X una v.a. in L^2 sia μ_X la sua media e sia c una costante positiva, vale:*

$$P(|X - \mu_X| \geq c) \leq \frac{VAR[X]}{c^2} \quad (4.6)$$

Se X e Y sono due variabili in L^2 si definisce covarianza la grandezza:

$$COVAR(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

La covarianza è bilineare e generalizza la varianza ($VAR(X) = COV(X, X)$), se vale $COVAR(X, Y) = 0$ allora si dice che le variabili aleatorie sono non correlate. Se X e Y sono due variabili aleatorie indipendenti allora la loro covarianza è zero, non vale il viceversa cioè può verificarsi che la covarianza sia zero senza che le variabili siano indipendenti. Una delle applicazioni più frequenti della proprietà di non correlazione è la seguente:

Proposizione 4.14. *Siano $X_1, \dots, X_n$ variabili aleatorie in L^2 a due a due non correlate, vale*

$$VAR[X_1 + \dots + X_n] = VAR[X_1] + \dots + VAR[X_n]$$

L'utilizzo più comune della proposizione precedente è quello per variabili aleatorie a due a due indipendenti (quindi non correlate!). Dalla disuguaglianza di Holder (o anche dalla disuguaglianza di Cauchy-Schwarz) si può dedurre che

$$COV(X, Y) \leq \sqrt{VAR(X) \cdot VAR(Y)}$$

Se le variabili aleatorie X e Y non sono costanti allora si può definire la seguente grandezza:

$$CORR(X, Y) := \frac{COV(X, Y)}{\sqrt{VAR(X) \cdot VAR(Y)}}$$

La grandezza $CORR(X, Y)$ è detta correlazione, è sempre compresa nell'intervallo $[-1, 1]$. La correlazione è molto utilizzata nelle applicazioni perché è invariante per cambiamenti di scala ovvero se α e β sono costanti positive allora vale:

$$CORR(\alpha X, \beta Y) = CORR(X, Y)$$

Per finire due eventi A e B si dicono non correlati se le loro funzioni indicatrici hanno covarianza zero, in questo caso vale: A e B sono non correlati se e solo se sono indipendenti. Quindi dire "eventi a due a due non correlati" è equivalente a dire "eventi a due a due indipendenti".

4.8 Secondo Lemma di Borel-Cantelli

Lemma 4.15. *Sia $\{A_n\}_{n \in \mathbb{N}}$ una successione di eventi indipendenti. Se vale $\sum_n P(A_n) = \infty$ allora*

$$P\left(\limsup_n A_n\right) = 1$$

Dimostrazione.

$$P\left(\limsup_n A_n\right) = P\left(\bigcap_n \bigcup_{k \geq n} A_k\right) = P\left(\bigcap_n B_n\right) =$$

Dove abbiamo posto $B_n := \bigcup_{k \geq n} A_k$, gli eventi B_n sono noncrescenti in n quindi

$$P\left(\limsup_n A_n\right) = \lim_n P(B_n)$$

$P(B_n)$ è non crescente quindi se vogliamo che il limite al secondo membro sia uguale ad 1 allora l'unica possibilità è dimostrare che $P(B_n) = 1$.

$$\begin{aligned} P(B_n) &= P\left(\bigcup_{k \geq n} A_k\right) = 1 - P\left(\left(\bigcup_{k \geq n} A_k\right)^c\right) = 1 - P\left(\bigcap_{k \geq n} A_k^c\right) \\ &= 1 - \prod_{k \geq n} P(A_k^c) = 1 - \prod_{k \geq n} (1 - P(A_k)) = 1 \end{aligned}$$

dove l'ultima uguaglianza è dovuta al lemma 4.16. $\square$

Lemma 4.16. *Data una successione $\{a_n\}_{n \in \mathbb{N}}$ con $a_n \in (0, 1)$ per ogni n . Vale:*

$$\prod_{n \in \mathbb{N}} (1 - a_n) = 0 \quad \Longleftrightarrow \quad \sum_{n \in \mathbb{N}} a_n = \infty$$

Dimostrazione. Consideriamo dapprima il caso semplice a_n non tende a zero allora esiste $\varepsilon > 0$ e una sottosuccessione n_k tale che per ogni k si ha $a_{n_k} > \varepsilon$. Quindi

$$\prod_{n \in \mathbb{N}} (1 - a_n) \leq \prod_{k \in \mathbb{N}} (1 - a_{n_k}) = 0$$

$$\sum_{n \in \mathbb{N}} a_n \geq \sum_{k \in \mathbb{N}} a_{n_k} = \infty$$

Consideriamo ora il caso a_n tende a zero.

$$\prod_{n \in \mathbb{N}} (1 - a_n) = \prod_{n \in \mathbb{N}} e^{\log(1 - a_n)}$$

Poiché gli esiti della produttoria e della sommatoria non dipendono da un numero finito di a_n allora possiamo considerare a_n piccolo cosicché

$$\begin{aligned} -2a_n &\leq \log(1 - a_n) \leq -a_n \\ \prod_{n \in \mathbb{N}} e^{-2a_n} &\leq \prod_{n \in \mathbb{N}} (1 - a_n) \leq \prod_{n \in \mathbb{N}} e^{-a_n} \\ e^{-2 \sum_n a_n} &\leq \prod_{n \in \mathbb{N}} (1 - a_n) \leq e^{-\sum_n a_n} \end{aligned}$$

$\square$

4.8.1 Secondo lemma di Borel-Cantelli per variabili non correlate

È possibile dimostrare il secondo lemma di Borel-Cantelli sotto ipotesi leggermente più deboli ovvero richiedendo che gli eventi siano solo a due a due indipendenti.

Lemma 4.17. *Sia $\{A_n\}_{n \in \mathbb{N}}$ una successione di eventi a due a due non correlati. Se vale $\sum_n P(A_n) = \infty$ allora*

$$P\left(\limsup_n A_n\right) = 1$$

Dimostrazione.

Prima di tutto cerchiamo di semplificare la tesi in qualcosa di equivalente ma più semplice.

Tesi: $P(\limsup_n A_n) = 1$

Tesi: per quasi ogni $\omega \in \Omega$ vale $\#\{n | \omega \in A_n\} = \infty$

Tesi: per quasi ogni $\omega \in \Omega$, $\forall \bar{n} \in \mathbb{N}$ esiste $n \geq \bar{n}$ tale che $\omega \in A_n$

poiché $\mathbb{N}$ è numerabile posso scambiare l'ordine di "per quasi ogni" e "per ogni"

Tesi: $\forall \bar{n} \in \mathbb{N}$, per quasi ogni $\omega \in \Omega$, esiste $n \geq \bar{n}$ tale che $\omega \in A_n$

poiché $\sum_n P(A_n) = \infty$ implica $\sum_{n \geq \bar{n}} P(A_n) = \infty$ allora posso limitarmi a dimostrare il caso $\bar{n} = 1$ quindi la tesi diventa:

Tesi: per quasi ogni $\omega \in \Omega$ esiste $n \in \mathbb{N}$ tale che $\omega \in A_n$

L'idea ora è di calcolare media e varianza del numero di volte in cui ω appartiene ad A_k con $k \in \{1, \dots, n\}$ e poi utilizzare la disuguaglianza di Chebyshev.

$$S_n := \sum_{k=1}^n \mathbb{1}_{A_k}$$

$$\mathbb{E}[S_n] = \sum_{k=1}^n P(A_k)$$

$$VAR[S_n] = \sum_{k=1}^n P(A_k)(1 - P(A_k)) \leq \sum_{k=1}^n P(A_k)$$

Ora utilizziamo la disuguaglianza 4.6 di chebyshev con $c = \mathbb{E}[S_n]$

$$P(|S_n - \mathbb{E}[S_n]| \geq \mathbb{E}[S_n]) \leq \frac{VAR(S_n)}{(\mathbb{E}[S_n])^2}$$

da cui

$$P(S_n = 0) \leq \frac{\mathbb{E}[S_n]}{(\mathbb{E}[S_n])^2} = \frac{1}{\sum_{k=1}^n P(A_k)}$$

quindi $P(S_n = 0)$ tende a zero per n che tende all'infinito. $\square$

4.9 σ -algebra coda e legge 0-1 di Kolmogorov

Introdurremo ora la σ -algebra coda e la legge 0-1 di Kolmogorov.

Definizione 4.6. Data una successione di σ -algebre $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$, sia $\tau_n := \sigma(\{\mathcal{G}_k | k > n\})$, si dice σ -algebra coda la σ -algebra $\tau := \bigcap_n \tau_n$.

Intuitivamente gli eventi di τ sono quelli che possono essere determinati dalle informazioni presenti su una qualunque coda delle σ -algebre.

Teorema 4.18 (Legge 0-1 di Kolmogorov).

Sia $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$ una successione di σ -algebre, sia $\tau_n := \sigma\{\mathcal{G}_k | k > n\}$ e sia $\tau := \bigcap_n \tau_n$ la σ -algebra coda. Se $\{\mathcal{G}_n\}_{n \in \mathbb{N}}$ è una successione di σ -algebre indipendenti allora τ è una σ -algebra banale ovvero

$$\text{per ogni } A \in \tau \quad \text{vale } P(A) \in \{0, 1\}$$

Dimostrazione. Dimostreremo che la σ -algebra τ è indipendente da se stessa cosicché per ogni $A \in \tau$ si ha $P(A \cap A) = P(A) \cdot P(A)$ e quindi $P(A) \in \{0, 1\}$. Sia

$$\mathcal{F}_n := \sigma(\mathcal{G}_1, \dots, \mathcal{G}_n)$$

per il principio di indipendenza a blocchi $\mathcal{F}_n$ e τ_n sono indipendenti e poiché $\tau \subseteq \tau_n$ allora $\mathcal{F}_n$ e τ sono indipendenti. Sia

$$\mathcal{A} := \bigcup_n \mathcal{F}_n \quad \text{e} \quad \mathcal{G} := \sigma(\{\mathcal{G}_n\}_{n \in \mathbb{N}}) = \sigma(\mathcal{A})$$

chiaramente $\mathcal{A}$ è un p-system che genera $\mathcal{G}$ e per quanto dimostrato prima è indipendente da τ . Per il teorema 4.11 si ha τ è indipendente da $\mathcal{G}$. Infine poiché $\tau \subseteq \mathcal{G}$ allora τ è indipendente da τ . $\square$

La legge 0-1 di Kolmogorov ci dice che nelle ipotesi del teorema se un evento appartiene a τ allora è quasi certamente vero o quasi certamente falso senza vie di mezzo. Se una variabile aleatoria X reale è τ misurabile allora è quasi certamente costante ovvero esiste x tale che $P(X = x) = 1$.

Esempio 4.1. Siano $\{X_n\}_{n \in \mathbb{N}}$ indipendenti, sia $S_n := X_1 + \dots + X_n$. Sia τ la σ -algebra coda ovvero $\tau = \bigcap_{k > n} \tau_k$ con $\tau_k := \sigma(X_{k+1}, X_{k+2}, \dots)$. Per la legge 0-1 di Kolmogorov allora:

- $\exists C \in [-\infty, +\infty]$ tale che $\limsup_n X_n = C$ q.c.
- la probabilità che $\frac{S_n}{n}$ converga è zero oppure uno.
- se la probabilità che $\frac{S_n}{n}$ converga è maggiore stretto di zero allora esiste C tale che $\frac{S_n}{n}$ converge quasi certamente a C .

4.10 Unicità della distribuzione congiunta di v.a. indipendenti

Molto spesso gli esercizi iniziano con affermazioni del tipo: “siano X e Y variabili aleatorie indipendenti con distribuzione $X \sim Unif(a, b)$ e $Y \sim Esp(\lambda)$.”, si assume per ipotesi che esistano e si dà per scontato che la distribuzione congiunta sia unica. In questo paragrafo vedremo che l’unicità è vera. Mentre l’esistenza verrà mostrata quando si introdurranno le misure prodotto.

Teorema 4.19. *Siano $(\Omega_1, \mathcal{H}_1, P_1)$ e $(\Omega_2, \mathcal{H}_2, P_2)$ spazi di probabilità. Sia $\{(E_i, \mathcal{E}_i)\}_{i \in I}$ una famiglia di spazi misurabili. Siano*

$$\begin{aligned} X_i &: (\Omega_1, \mathcal{H}_1) \rightarrow (E_i, \mathcal{E}_i) \\ Y_i &: (\Omega_2, \mathcal{H}_2) \rightarrow (E_i, \mathcal{E}_i) \end{aligned}$$

misurabili con eguale distribuzione, $X_i(P_1) = Y_i(P_2)$ per ogni i . Sia $X := (X_i)_{i \in I}$ e $Y := (Y_i)_{i \in I}$. Se le v.a. $\{X_i\}_{i \in I}$ sono indipendenti e le v.a. $\{Y_i\}_{i \in I}$ sono indipendenti allora X e Y sono applicazioni misurabili a valori in $(E, \mathcal{E}) := (\times_{i \in I} E_i, \otimes_{i \in I} \mathcal{E}_i)$ con medesima legge ovvero $P_1(X^{-1}) = P_2(Y^{-1})$.

Dimostrazione. Per la proposizione 1.15 X e Y sono misurabili. Consideriamo il seguente p-system

$$\mathcal{A} = \left\{ A = \times_{i \in I} A_i \mid J \subseteq I \text{ finito}, A_i \in \mathcal{E}_i \text{ per ogni } i \in J, A_i = E_i \text{ per ogni } i \notin J \right\} \quad (4.7)$$

per la proposizione 1.13 $\mathcal{A}$ genera $\mathcal{E}$, mostreremo che le due misure $P_1(X^{-1})$ e $P_2(Y^{-1})$ coincidono sul p-system $\mathcal{A}$ cosicché per il teorema 2.4 sono uguali. Sia $A = \times_{i \in I} A_i \in \mathcal{A}$ come in (4.7),

$$\begin{aligned} P_1(X^{-1}(A)) &= P_1\left(\bigcap_{i \in J} X_i^{-1}(A_i)\right) = \prod_{i \in J} P_1(X_i^{-1}(A_i)) \\ &= \prod_{i \in J} P_1(Y_i^{-1}(A_i)) = P_1\left(\bigcap_{i \in J} Y_i^{-1}(A_i)\right) \\ &= P_1(Y^{-1}(A)) \end{aligned}$$

dove per la seconda e quarta uguaglianza è stata utilizzata l’indipendenza. $\square$

4.11 Cifre decimali di un numero a caso in $(0, 1)$

In questo paragrafo forniremo un metodo costruttivo per ottenere una famiglia numerabile di variabili aleatorie indipendenti. Consideriamo un numero U scelto a caso tra zero e uno ($U \sim Unif(0, 1)$). Siano $\{X_n\}_{n \in \mathbb{N}}$ le cifre di U dopo la virgola e sia $X := (X_n)_{n \in \mathbb{N}}$. Quindi X v.a. a valori in

$$(E, \mathcal{E}) := \left(\prod_{i \in \mathbb{N}} \{0, \dots, 9\}, \bigotimes_{i \in \mathbb{N}} \mathcal{P}(\{0, \dots, 9\}) \right)$$

Dimostreremo ora che le $\{X_n\}_{n \in \mathbb{N}}$ sono uniformi su $\{0, 1, \dots, 9\}$ e indipendenti.

DISTRIBUZIONE di X_n :

Facciamo un esempio:

$$P(X_1 = 5, X_2 = 7, X_3 = 2) = P(U \in [0.572, 0.573)) = \frac{1}{10^3}$$

in maniera analoga se $x_1, \dots, x_n$ sono cifre (elementi di $\{0, \dots, 9\}$) allora vale

$$P(X_1 = x_1, \dots, X_n = x_n) = \frac{1}{10^n}$$

quindi

$$P(X_n = x_n) = \sum_{x_1, x_2, \dots, x_{n-1}} P(X_1 = x_1, \dots, X_n = x_n) = \frac{10^{n-1}}{10^n} = \frac{1}{10}$$

INDIPENDENZA: per mostrare che le variabili $\{X_n\}_{n \in \mathbb{N}}$ sono indipendenti utilizzeremo la proposizione 4.11.

Per ogni n mi basterà dimostrare che le variabili $X_1, \dots, X_n$ sono indipendenti. Il p-system utilizzato per la proposizione 4.11 è quello dei singoletti $\mathcal{A}_n = \{\emptyset, \{X_n \in 0\}, \dots, \{X_n \in 9\}\}$. L'insieme vuoto è indipendente da tutto, per il resto vale:

$$P(X_1 = x_1, \dots, X_n = x_n) = 10^{-n} = P(X_1 = x_1) \cdot \dots \cdot P(X_n = x_n)$$

Abbiamo quindi mostrato che le $\{X_n\}_{n \in \mathbb{N}}$ sono i.i.d. Vedremo ora che vale anche il viceversa. Denotiamo con Φ l'applicazione descritta sopra che manda U in X : $X = \Phi(U)$. L'applicazione Φ è iniettiva, denotiamo che Φ^{-1} la sua inversa. Siano $\{Y_n\}_{n \in \mathbb{N}}$ variabili aleatorie indipendenti e uniformemente distribuite su $\{0, 1, \dots, 9\}$ e sia $Y = (Y_n)_{n \in \mathbb{N}}$ (per l'unicità della distribuzione di una famiglia di v.a. indipendenti) le v.a. X e Y hanno la stessa distribuzione e quindi anche $\Phi^{-1}(X)$ e $\Phi^{-1}(Y)$ hanno la stessa distribuzione. Questo vuol dire che data una variabile aleatoria $U \sim Unif(0, 1)$ è possibile costruire una successione di variabili aleatorie i.i.d. e uniformi su $\{0, \dots, 9\}$ e viceversa data una successione di variabili aleatorie i.i.d. con distribuzione uniforme su $\{0, \dots, 9\}$ è possibile costruire una variabile aleatoria uniforme su $(0, 1)$.

Supponiamo ora $U \sim Unif(0, 1)$ e $\{X_n\}_{n \in \mathbb{N}}$ siano le cifre di U allora possiamo porre $Y_n := X_{2n}$ e $Z_n := X_{2n+1}$ e ricavare due successioni di variabili indipendenti dalle quali posso ricavare due v.a. indipendenti con distribuzione uniforme su $(0, 1)$. Allo stesso modo se poniamo $Z_{k,n} := X_{2^k 3^n}$ otteniamo una successione $\{Z_{k,n}\}_{n \in \mathbb{N}}$ di successioni di variabili indipendenti, alle quali possiamo far corrispondere una successione di variabili aleatorie uniformi su $(0, 1)$ e indipendenti. Ovvero abbiamo costruito un'applicazione $\Psi : (0, 1) \rightarrow (0, 1)^{\mathbb{N}}$ tale che posto $T = (T_n)_{n \in \mathbb{N}} := \Psi(U)$ si abbia $\{T_n\}_{n \in \mathbb{N}}$ è una famiglia di variabili aleatorie indipendenti uniformi su $(0, 1)$. Non è possibile fare di meglio ovvero vale il seguente esercizio.

Esercizio 4.4. ***

Sia $U \sim Unif(0, 1)$ sia Ψ un'applicazione misurabile

$$\Psi : ((0, 1), \mathcal{B}(0, 1)) \rightarrow \left(\prod_{i \in I} (0, 1), \bigotimes_{i \in \mathbb{N}} \mathcal{B}((0, 1)) \right)$$

Sia $X := (X_i)_{i \in I} := \Psi(U)$. Se le variabili aleatorie $\{X_i\}_{i \in I}$ sono i.i.d. con distribuzione uniforme su $(0, 1)$ allora $\#(I) \leq \#(\mathbb{N})$.

Appendice A

Esercizi

Gli esercizi verranno divisi per capitoli. Per ciascun capitolo saranno divisi in esercizi teorici e di comprensione, gli esercizi teorici sono esercizi che è meglio conoscere almeno nell'enunciato; gli esercizi di comprensione sono esercizi per mettersi alla prova e migliorare la comprensione della teoria.

A.1 Capitolo 1

A.1.1 Esercizi teorici

Esercizio A.1.1. *Siano $\mathcal{A}_1$ e $\mathcal{A}_2$ due σ -algebre. Mostrare che $\mathcal{A}_1 \cup \mathcal{A}_2$ è una σ -algebra se e solo se $\mathcal{A}_1 \subseteq \mathcal{A}_2$ oppure $\mathcal{A}_2 \subseteq \mathcal{A}_1$.*

Esercizio A.1.2. *Mostrare che se $(A_n)_{n \in \mathbb{N}}$ è una successione crescente di eventi allora:*

$$\limsup A_n = \liminf A_n = \lim A_n = \cup A_n$$

Esercizio A.1.3. *Dimostrare che se $X : \mathbb{R}^n \rightarrow \mathbb{R}^m$ è un'applicazione continua allora è anche misurabile come applicazione da $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ in $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$.*

Esercizio A.1.4. *Consideriamo i seguenti insiemi:*

$$\begin{cases} \mathcal{A}_1 := \{ (-\infty, a] \mid a \in \mathbb{R} \} \\ \mathcal{A}_2 := \{ (a, b] \mid a, b \in \mathbb{R} \} \\ \mathcal{A}_3 := \{ (a_1, b_1] \cup \dots \cup (a_n, b_n] \mid n \in \mathbb{N}, a_1, b_1, \dots, a_n, b_n \in \overline{\mathbb{R}} \} \\ \mathcal{A}_4 := \{ (-a, b) \mid a, b > 0 \} \\ \mathcal{A}_5 := \{ [a, b] \mid a, b \in \mathbb{Q} \} \end{cases}$$

Dimostrare che:

$$\sigma(\mathcal{A}_1) = \sigma(\mathcal{A}_2) = \sigma(\mathcal{A}_3) = \sigma(\mathcal{A}_4) = \sigma(\mathcal{A}_5) = \mathcal{B}(\mathbb{R})$$

Esercizio A.1.5. Siano $\{X_n\}_{n \in \mathbb{N}}$ applicazioni misurabili da uno spazio misurabile $(\Omega, \mathcal{H})$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ dimostrare che sono misurabili le seguenti applicazioni:

1. $\sup_n X_n$
2. $\inf_n X_n$
3. $\limsup_n X_n$
4. $\liminf_n X_n$

dove $\limsup_n X_n := \inf_n \{\sup_{m > n} X_m\} = \lim_n \{\sup_{m > n} X_m\}$
e $\liminf_n X_n := \sup_n \{\inf_{m > n} X_m\} = \lim_n \{\inf_{m > n} X_m\}$

Esercizio A.1.6. Sia Ω un insieme e sia $\mathcal{G} \subseteq \mathcal{P}(\Omega)$ consideriamo le seguenti ipotesi su $\mathcal{G}$:

- (i) $\Omega \in \mathcal{G}$
- (ii) Se $A \in \mathcal{G}$ allora $A^c \in \mathcal{G}$
- (iii) Se $A_n \in \mathcal{G}$ e $A_n \cap A_m = \emptyset$ per ogni $n \neq m$ allora $\cup_{n \in \mathbb{N}} A_n \in \mathcal{G}$
- (a) $\Omega \in \mathcal{G}$
- (b) Se $A, B \in \mathcal{G}$ e $A \subseteq B$ allora $B/A \in \mathcal{G}$
- (c) Se $A_n \in \mathcal{G}$ e $A_n \uparrow A$ allora $A \in \mathcal{G}$

Dimostrare che le ipotesi (i), (ii) e (iii) implicano le ipotesi (a), (b) e (c) e viceversa le ipotesi (a), (b) e (c) implicano le ipotesi (i), (ii) e (iii).

Esercizio A.1.7. (*) Dimostrare che vale:

$$\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$$

Esercizio A.1.8. (***) Sia $\{\Omega_i\}_{i \in I}$ una famiglia di spazi metrici completi e separabili, sia $\{\mathcal{T}_i\}_{i \in I}$ la famiglia delle corrispondenti topologie e supponiamo infine che l'insieme degli indici I sia al più numerabile. Dimostrare che posto $\Omega = \times_{i \in I} \Omega_i$ e $\mathcal{T} = \otimes_{i \in I} \mathcal{T}_i$ allora vale

$$\mathcal{B}(\Omega) = \bigotimes_{i \in I} \mathcal{B}(\Omega_i) \tag{A.1}$$

dove $\mathcal{B}(\Omega)$ e $\mathcal{B}(\Omega_i)$ indicano la σ -algebra dei boreliani ovvero $\mathcal{B}(\Omega) := \sigma(\mathcal{T})$ e $\mathcal{B}(\Omega_i) = \sigma(\mathcal{T}_i)$.

A.1.2 Esercizi di comprensione

Esercizio A.1.9. Dimostrare che se una σ -algebra è finita allora la sua cardinalità è 2^n per qualche $n \in \mathbb{N}$.

Esercizio A.1.10. Sia Ω un insieme con cardinalità pari e maggiore di 2. Consideriamo la seguente collezione di sottoinsiemi di Ω :

$$\mathcal{A} := \{B \in \mathcal{P}(\Omega) \mid \#(B) \text{ pari} \}$$

dimostrare che $\mathcal{A}$ è un d -system ma non è una σ -algebra.

Esercizio A.1.11. Sia $\mathcal{C}$ la collezione di sottoinsiemi di $\mathbb{R}$ data da:

$$\mathcal{C} := \{(-a, a) : a > 0\}$$

e sia $\mathcal{A} = \sigma(\mathcal{C})$.

(a) Dimostrare che se A appartiene ad $\mathcal{A}$ allora A è un boreliano e vale $A = -A$ dove $-A := \{x : -x \in A\}$.

(b)** Dimostrare che se A è un boreliano e vale $A = -A$ allora A appartiene ad $\mathcal{A}$.

Esercizio A.1.12. Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ una funzione derivabile in ogni $x \in \mathbb{R}$. Mostrare che f' è misurabile. (Notare che f derivabile non vuol dire che ha derivata continua.)

Esercizio A.1.13. (*) Sia $(X_n)_{n \in \mathbb{N}}$ una successione di variabili aleatorie, sia $\mathcal{T}_n$ la sigma algebra generata da $(X_k)_{k \geq n}$ e sia $\mathcal{T} = \cap_n \mathcal{T}_n$ la sigma algebra coda. Dimostrare che:

(a) $\limsup_n X_n$ è $\mathcal{T}$ misurabile.

(b) $\liminf_n X_n$ è $\mathcal{T}$ misurabile.

(c) $\limsup_n \frac{X_n}{n}$ è $\mathcal{T}$ misurabile.

(d) $\limsup_n \frac{X_1 + \dots + X_n}{n}$ è $\mathcal{T}$ misurabile.

(e) Supponiamo ora $X_n(\omega) \geq 0$ per ogni $n \in \mathbb{N}$ e $\omega \in \Omega$ mostrare che

$\limsup_n \frac{\max\{X_1, \dots, X_n\}}{n}$ è $\mathcal{T}$ misurabile.

(f) Cosa si può dire di $\limsup_n \max\{X_1, \dots, X_n\}$? Costruire un esempio tale che il limite precedente non è $\mathcal{T}$ misurabile e vale $X_n(\omega) \geq 0$ per ogni $n \in \mathbb{N}$ e $\omega \in \Omega$.

Soluzioni

Exercise A.1.7

Questo risultato dovrebbe già essere noto dai precedenti corsi di analisi. Ci sono più modi per effettuare la dimostrazione. In queste note utilizzeremo i rettangoli di aperti:

$$\mathcal{A} := \{A_1 \times A_2 \mid A_i \text{ aperti di } \mathbb{R}\}$$

Chiaramente $\mathcal{A} \subset \mathcal{B}(\mathbb{R}^2)$ e $\mathcal{A} \subset \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$.

Basterà dimostrare: (i) $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}^2)$ e (ii) $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$.

(Dim.(i)) Sappiamo che un insieme aperto di $\mathbb{R}^2$ è unione numerabile di rettangoli aperti dunque $\sigma(\mathcal{A})$ contiene gli aperti di $\mathbb{R}^2$ e $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}^2)$.

(Dim.(ii)) $\sigma(\mathcal{A}) \supseteq \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ se le proiezioni naturali $\pi_i : \mathbb{R}^2 \rightarrow \mathbb{R}$ sono $\sigma(\mathcal{A})$ misurabili. Se A è un aperto di $\mathbb{R}$ allora $\pi_1^{-1}(A) = A \times \mathbb{R} \in \mathcal{A}$.

Esercizio A.1.8

La dimostrazione di questa uguaglianza è simile a quella dell'esercizio precedente. La difficoltà principale è di tipo topologico: occorre mostrare che gli aperti di $\mathcal{T}$ sono unione numerabile di rettangoli di aperti.

Esercizio A.1.11

(b) Sia $\Phi : \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(\mathbb{R})$ la seguente funzione:

$$\Phi(B) = (B \cap [0, \infty)) \cup (-(B \cap [0, \infty)))$$

L'insieme $A := \Phi(B)$ è simmetrico e coincide con B sulla semiretta positiva. Sia ora $\mathcal{F}$ la famiglia degli insiemi boreliani che sono mandati in $\mathcal{A}$ da Φ :

$$\mathcal{F} = \{B \in \mathcal{B}(\mathbb{R}) \mid \Phi(B) \in \mathcal{A}\}$$

$\mathcal{F}$ contiene l'insieme $\mathcal{A}_4$ dell'esercizio A.1.4 inoltre vale

$$\mathbb{R} \in \mathcal{F}$$

$$B \in \mathcal{F} \text{ implica } B^c \in \mathcal{F}$$

$$B_n \in \mathcal{F} \text{ per ogni } n \text{ implica } \bigcup_{n \in \mathbb{N}} B_n \in \mathcal{F}$$

Dunque

$$\mathcal{F} = \mathcal{B}(\mathbb{R}).$$

A questo punto possiamo concludere. Se A è un boreliano simmetrico allora $A \in \mathcal{F}$ e $\Phi(A) \in \mathcal{A}$, poiché $\Phi(A) = A$ allora A appartiene a $\mathcal{A}$.

A.2 Capitolo 2

A.2.1 Esercizi teorici

Esercizio A.2.1. *Mostrare che se $(A_n)_{n \in \mathbb{N}}$ è una successione di eventi tali che $\limsup A_n = \liminf A_n$ allora*

$$\lim P(A_n) = P(\lim A_n)$$

Esercizio A.2.2. *Sia $(\Omega, \mathcal{H}, P)$ uno spazio di probabilità. Per ogni A, B in $\mathcal{H}$ sia $d(A, B) := P(A \Delta B)$. Dimostrare che d è una pseudodistanza.*

Esercizio A.2.3. (*) *Sia $(\Omega, \mathcal{H}, P)$ uno spazio di probabilità e sia $\mathcal{A}$ un'algebra tale che $\sigma(\mathcal{A}) = \mathcal{H}$. Dimostrare che per ogni $H \in \mathcal{H}$ e $\epsilon > 0$ esiste $A \in \mathcal{A}$ tale che $P(H \Delta A) < \epsilon$.*

Esercizio A.2.4. *Sia $(\Omega, \mathcal{H}, \mu)$ è uno spazio di misura, e sia P una proposizione su $\Omega \times E$. Dimostrare che se E è un insieme numerabile (di solito $E = \mathbb{N}$) sono cose equivalenti:*

- per quasi ogni $\omega \in \Omega$, per ogni $n \in E$ vale $P(\omega, n)$
- per ogni $n \in E$, per quasi ogni $\omega \in \Omega$ vale $P(\omega, n)$

(*) *Trovare un controesempio nel caso in cui sia rimossa l'ipotesi E al più numerabile.*

Esercizio A.2.5. *Sia $(\Omega, \mathcal{F}, P)$ uno spazio di probabilità. Sia $\mathcal{G} \subset \mathcal{F}$ una sotto σ -algebra. Sia N l'insieme degli eventi trascurabili di $\mathcal{F}$ ($N := \{A \in \mathcal{F} | P(A) = 0\}$). Se*

$$\mathcal{H} := \sigma(N, \mathcal{G})$$

allora per ogni $A \in \mathcal{H}$ esiste $B \in \mathcal{G}$ tale che

$$P(A \Delta B) = 0$$

A.2.2 Esercizi di comprensione

Esercizio A.2.6. *Data una successione di eventi $(A_n)_{n \in \mathbb{N}}$*

(a) *Dimostrare che*

$$P(\liminf A_n) \leq \liminf P(A_n) \leq \limsup P(A_n) \leq P(\limsup(A_n)) \quad (\text{A.2})$$

(b) *Trovare un esempio in cui tutte e tre le disuguaglianze della disequazione (A.2) sono strette.*

Esercizio A.2.7. *Sia $(\Omega, \mathcal{H}, P)$ uno spazio di probabilità e sia $\mathcal{A}$ la collezione di eventi:*

$$\mathcal{A} := \{A \in \mathcal{H} | P(A) \in \{0, 1\}\}$$

dimostrare che $\mathcal{A}$ è una σ -algebra

Esercizio A.2.8. Sia Ω un insieme più che numerabile e sia $\mathcal{A}$ la collezione di sottoinsiemi di Ω data da:

$$\mathcal{A} := \{A \in \Omega \mid A \text{ oppure } A^c \text{ è finito o numerabile}\}$$

(a) Dimostrare che $\mathcal{A}$ è una σ -algebra e la seguente funzione P è una misura di probabilità su $(\Omega, \mathcal{A})$.

$$P(A) = \begin{cases} 0 & \text{se } A \text{ è finito o numerabile} \\ 1 & \text{se } A^c \text{ è finito o numerabile} \end{cases}$$

(b) A cosa serve l'ipotesi Ω è più che numerabile? Può essere rimossa?

SOLUZIONI

Esercizio A.2.2

L'applicazione $d : \mathcal{A} \times \mathcal{A} \rightarrow [0, \infty)$ che manda A e B in $P(A \triangle B)$ è una pseudodistanza se

- $d(A, A) = 0$
- $d(A, B) = d(B, A)$
- $d(A, B) + d(B, C) \geq d(A, C)$

Le prime due condizioni sono ovvie e la terza è una conseguenza dell'inclusione:

$$(A \triangle B) \cup (B \triangle C) \supseteq (A \triangle C)$$

Esercizio A.2.3

Occorre mostrare che tutti gli elementi di $\mathcal{H}$ si possono approssimare con elementi di $\mathcal{A}$. Indichiamo con $\mathcal{G}$ la collezione degli insiemi che soddisfano tale proprietà:

$$\mathcal{G} := \{B \in \mathcal{H} \mid \forall \varepsilon > 0 \exists A \in \mathcal{A} \text{ tale che } P(A \triangle B) < \varepsilon\}$$

è chiaro che $\mathcal{G}$ contiene $\mathcal{A}$ poiché $\mathcal{G} = \sigma(\mathcal{A})$ e $\mathcal{A}$ è un p-system allora per il lemma di Dynkin basta mostrare che $\mathcal{G}$ è un d-system.

Dimostrare che $\mathcal{G}$ contiene Ω ed è chiuso rispetto al complementare è banale. Resta da dimostrare che $\mathcal{G}$ è chiuso rispetto all'unione numerabile di insiemi disgiunti. Dimostreremo dapprima la chiusura finita e poi quella numerabile. Siano B_1, B_2 due elementi disgiunti di $\mathcal{G}$, quindi vale:

$$\forall \varepsilon > 0 \quad \exists A_1, A_2 \in \mathcal{A} \text{ tali che } P(A_1 \triangle B_1) < \varepsilon \text{ e } P(A_2 \triangle B_2) < \varepsilon$$

Sia ora $B = B_1 \cup B_2$ e $A = A_1 \cup A_2$ e dall'inclusione $((A_1 \cup A_2) \triangle (B_1 \cup B_2)) \subseteq (A_1 \triangle B_1) \cup (A_2 \triangle B_2)$ si ottiene:

$$P(A \triangle B) < 2\varepsilon$$

A questo punto utilizziamo la chiusura finita per dimostrare la chiusura numerabile. Siano $\{B_n\}_{n \in \mathbb{N}}$ elementi disgiunti di $\mathcal{G}$. Siano $B = \bigcup_n B_n$ e $C_n := B_1 \cup \dots \cup B_n$. Poiché $\mathcal{G}$ è chiuso rispetto all'unione finita gli elementi C_n appartengono a $\mathcal{G}$. Inoltre sappiamo che $C_n \uparrow B$. Per ogni $\varepsilon > 0$ esiste n e $A_n \in \mathcal{A}$ tali che

$$P(C_n \Delta B) < \varepsilon \quad P(A_n \Delta C_n) < \varepsilon$$

da cui si deduce:

$$P(A_n \Delta B) < 2\varepsilon$$

A.3 Capitolo 3

A.3.1 Esercizi teorici

Esercizio A.3.1. *Utilizzare la disuguaglianza di Jensen per dimostrare che se $1 \leq q \leq p$ allora $\|X\|_{L^p(\Omega, \mathcal{H}, P)} \geq \|X\|_{L^q(\Omega, \mathcal{H}, P)}$.*

Esercizio A.3.2. *(Disuguaglianza di Markov) Sia X una variabile aleatoria. Sia $g: \mathbb{R} \rightarrow [0, \infty)$ una applicazione non decrescente e sia $c \in \mathbb{R}$ tale che $g(c) > 0$ dimostrare che vale:*

$$P(X \geq c) \leq \frac{\mathbb{E}[g(X)]}{g(c)}$$

Esercizio A.3.3. *(*) Sia X una variabile aleatoria a valori in $\mathbb{N}$ dimostrare che*

$$\sum_{n=1}^{\infty} P(X \geq n) = \mathbb{E}[X] \quad (\text{A.3})$$

Esercizio A.3.4. *(*) Sia X una variabile aleatoria quasi certamente positiva mostrare che vale*

$$\sum_{n=1}^{\infty} P(X \geq n) \leq \mathbb{E}[X] \leq \sum_{n=0}^{\infty} P(X > n) \quad (\text{A.4})$$

Mostrare inoltre che per ogni $p \geq 1$ vale

$$\sum_{n=1}^{\infty} (n^p - (n-1)^p) P(X \geq n) \leq \mathbb{E}[X^p] \leq \sum_{n=0}^{\infty} ((n+1)^p - n^p) P(X > n)$$

Mostrare infine che se X è a valori in $\mathbb{N}$ allora le disuguaglianze precedenti diventano uguaglianze.

A.4 Capitolo 4

A.4.1 Esercizi teorici

Esercizio A.4.1. Siano $E_1, E_2, \dots, E_n$ eventi indipendenti. Se $F_1, F_2, \dots, F_n$ sono eventi tali che per ogni i o $F_i = E_i$ oppure $F_i = E_i^c$ allora anche $F_1, F_2, \dots, F_n$ sono eventi indipendenti.

Esercizio A.4.2. Dimostrare la seguente proposizione.

Sia $\{X_n\}_{n \in \mathbb{N}}$ una sequenza di variabili aleatorie. Se per ogni n la variabile X_n è indipendente da $(X_1, X_2, \dots, X_{n-1})$ allora le variabili $\{X_n\}_{n \in \mathbb{N}}$ sono indipendenti.

Esercizio A.4.3 (Indipendenza). Siano $\{A_n\}_{n \in I = \{1, 2, \dots, \bar{n}\}}$ eventi. Siano $\{\mathcal{G}_i\}_{i \in I}$ le σ -algebre generate dagli eventi: $\mathcal{G}_i := \sigma(A_i)$.

Dimostrare che gli eventi $\{A_n\}_{n \in I}$ sono indipendenti se e solo se lo sono anche le σ -algebras $\{\mathcal{G}_n\}_{n \in I}$.

Esercizio A.4.4. ** Sia X una variabile aleatoria con $\mathbb{E}[X] = M \neq 0$, $\text{VAR}(X) = \sigma^2$ dimostrare che:

$$(a) P(X = 0) \leq \frac{\sigma^2}{M^2}$$

$$(b) ** P(X = 0) \leq \frac{\sigma^2}{M^2 + \sigma^2}$$

Esercizio A.4.5. ***

Sia $U \sim \text{Unif}(0, 1)$ sia Ψ un'applicazione misurabile

$$\Psi : ((0, 1), \mathcal{B}(0, 1)) \rightarrow \left(\prod_{i \in I} (0, 1), \bigotimes_{i \in \mathbb{N}} \mathcal{B}((0, 1)) \right)$$

Sia $X := (X_i)_{i \in I} := \Psi(U)$. Se le variabili aleatorie $\{X_i\}_{i \in I}$ sono indipendenti e con distribuzione non banale (ovvero non quasi certamente costante) allora $\#(I) \leq \#(\mathbb{N})$.

A.4.2 Esercizi di comprensione

Esercizio A.4.6. Costruire un esempio di spazio di probabilità con tre eventi A , B e C tali che siano a due a due indipendenti ma non siano indipendenti in generale (ovvero $P(A \cap B \cap C) \neq P(A)P(B)P(C)$).

Esercizio A.4.7. L'emofilia è una malattia causata dal malfunzionamento del cromosoma X . Tale malfunzionamento impedisce ai geni del cromosoma di produrre un "fattore" cruciale per la coagulazione del sangue. Come è noto le donne hanno due cromosomi X mentre gli uomini ne hanno uno solo, questo fa sì che la presenza del cromosoma malfunzionante causi la malattia

quasi solo negli uomini in quanto nelle donne il secondo cromosoma X può sopperire alla carenza del primo.

Supponiamo che la regina abbia genitori sani ed un fratello ammalato di emofilia. (Quindi la madre della regina è una portatrice sana!) Supponiamo inoltre che la regina abbia 2 figli maschi entrambi sani e sia in attesa del terzo.

Qual è la probabilità che il terzo figlio sia sano?

SOLUZIONI

Esercizio A.4.1

È chiaro che se si riesce a fare il passaggio al complementare su un singolo evento, poi uno alla volta per induzione si può passare al complementare su qualunque sottoinsieme di eventi. La proprietà (4.3) va dimostrata su qualunque sottoinsieme di eventi, se tale sottoinsieme è costituito da m elementi con $m \in \{2, 3, \dots, n\}$ (i casi $m = 0$ e $m = 1$ sono sempre veri) allora a patto di permutare gli indici possiamo supporre di dover dimostrare che:

$$P(E_1 \cap \dots \cap E_{m-1} \cap E_m^c) = P(E_1) \dots P(E_{m-1})P(E_m^c) \quad (\text{A.5})$$

il punto chiave adesso sta nell'esprimere l'evento della prima probabilità della (A.5) in termini di sole intersezioni di eventi E_i per i quali possiamo utilizzare l'indipendenza, osserviamo che:

$$E_1 \cap \dots \cap E_{m-1} \cap E_m^c = (E_1 \cap \dots \cap E_{m-1}) \setminus (E_1 \cap \dots \cap E_{m-1} \cap E_m)$$

dunque

$$P(E_1 \cap \dots \cap E_{m-1} \cap E_m^c) = P(E_1 \cap \dots \cap E_{m-1}) - P(E_1 \cap \dots \cap E_{m-1} \cap E_m)$$

poiché stiamo supponendo gli E_i indipendenti si ha:

$$P(E_1 \cap \dots \cap E_{m-1} \cap E_m^c) = P(E_1) \dots P(E_{m-1}) - P(E_1) \dots P(E_{m-1})P(E_m)$$

da cui segue immediatamente la (A.5) e dunque la tesi.

Esercizio A.4.5 [Suggerimenti]

Considerare prima il caso di variabili bernoulliane di parametro $1/2$.

Utilizzare l'esercizio A.2.3 con un'algebra numerabile che genera i boreliani e poi l'esercizio A.2.2

Esercizio A.4.7

La probabilità che il terzo figlio maschio sia sano sapendo che i primi due sono sani è $\frac{9}{10}$.

Appendice B

Esistenza della misura di Lebesgue

Teorema B.1. *Esiste una misura μ su $([0, 1], \mathcal{B}([0, 1]))$ tale che per ogni $0 \leq a \leq b \leq 1$ vale $\mu(a, b) = b - a$.*

La dimostrazione che segue consta di due parti nella prima si dimostra che la definizione (B.1) è una buona definizione e μ_0 è una funzione additiva. Nella seconda parte si dimostra che μ_0 è σ -additiva. Sul libro di riferimento il Williams, la dimostrazione della prima parte è lasciata al lettore e la seconda è svolta in maniera diversa.

Dimostrazione. Per cominciare possiamo considerare la seguente algebra $\mathcal{A}$:

$$\mathcal{A} := \{[a_1, b_1) \cup [a_2, b_2) \cup \dots \cup [a_n, b_n) | n \in \mathbb{N}, 0 \leq a_1 < b_1 \leq a_2 < b_2 \leq \dots < b_n \leq 1\}$$

Abbiamo già visto che l'algebra $\mathcal{A}$ genera i boreliani. È chiaro anche che la misura richiesta dal teorema esiste se e solo se per ogni $a < b$ in $[0, 1]$ vale $\mu([a, b)) = b - a$. Inoltre per ogni a deve valere $\mu(\{a\}) = 0$ quindi $\mu(\{1\}) = 0$. Senza perdere in generalità possiamo costruire la misura su $([0, 1], \mathcal{B}([0, 1]))$. Consideriamo la seguente funzione $\mu_0 : \mathcal{A} \rightarrow [0, \infty]$:

$$\forall A \in \mathcal{A} \quad A = [a_1, b_1) \cup [a_2, b_2) \cup \dots \cup [a_n, b_n) \quad \mu_0(A) := \sum_{k=1}^n (b_k - a_k) \quad (\text{B.1})$$

Per completare la dimostrazione bisognerà dimostrare che la precedente definizione è una buona definizione e che la funzione μ_0 è σ -additiva. Dopo di che la tesi seguirà in modo immediato dal teorema di estensione di Carathéodory.

La dimostrazione della buona definizione di (B.1) è semplice ma piuttosto noiosa. Il punto chiave è che un insieme $A \in \mathcal{A}$ può avere più rappresentazioni della forma (B.1), per esempio se $a < b < c$ allora $[a, c) = [a, b) \cup [b, c)$,

bisogna verificare che qualunque sia la rappresentazione il risultato della somma non cambia. Nell'esempio precedente la verifica è immediata: $\mu_0([a, b] \cup [b, c]) = (b - a) + (c - b) = c - a = \mu_0([a, c])$. In generale un insieme $A \in \mathcal{A}$ ammette una sola rappresentazione con n minimo, tutte le altre rappresentazioni possono essere ricondotte a quella minima unendo una alla volta gli intervalli consecutivi.

Dalla buona definizione segue immediatamente che μ_0 è additiva.

Dimostrare che μ_0 è σ -additiva è più difficile. Sappiamo che μ_0 è additiva e questo ci tornerà molto utile. Per cominciare μ_0 additiva sugli insiemi disgiunti ci dà la subaddittività sugli insiemi non disgiunti e la monotonia rispetto all'inclusione. Sia ora $A \in \mathcal{A}$ e supponiamo $A = \bigcup_{n \in \mathbb{N}} A_n$ con gli $A_n \in \mathcal{A}$ disgiunti. Occorre mostrare che $\mu_0(A) = \sum_n \mu_0(A_n)$. Poiché $(\bigcup_{n \in \mathbb{N}} A_n) \cup A^c = [0, 1]$ e $\mu_0(A) + \mu_0(A^c) = \mu_0([0, 1]) = 1$ allora senza perdere in generalità possiamo supporre $A = [0, 1]$.

Indichiamo con $\{[a_n, b_n]\}_{n \in \mathbb{N}}$ la famiglia di tutti gli intervallini degli insiemi $\{A_n\}_{n \in \mathbb{N}}$. Abbiamo:

$$n \neq m \implies [a_n, b_n] \cap [a_m, b_m] = \emptyset \quad \text{e} \quad \bigcup_{n \in \mathbb{N}} [a_n, b_n] = [0, 1]$$

occorre mostrare che $\sum_n (b_n - a_n) = 1$. Una disuguaglianza è più semplice, poiché μ_0 è additiva allora

$$\sum_{k=1}^n (b_k - a_k) = \mu \left(\bigcup_{k=1}^n [a_k, b_k] \right) \leq 1$$

quindi vale

$$\sum_n (b_n - a_n) \leq 1$$

occorre ora mostrare che tale somma non può essere strettamente minore di 1.

La questione centrale è se sia possibile ricoprire $[0, 1]$ con una successione di intervallini $[a_n, b_n]$ di modo che la somma totale delle lunghezze sia strettamente minore di 1?

L'idea della dimostrazione che seguirà è di allargare un pochino gli intervalli $[a_n, b_n]$ e in modo da renderli aperti, a questo punto utilizzeremo il principio che permette di estrarre un ricoprimento finito da un ricoprimento di aperti, e sul ricoprimento finito potremo applicare l'addittività di μ_0 .

Sia $\varepsilon > 0$ piccolo. Per ogni n sia

$$c_n = a_n - \varepsilon(b_n - a_n) \quad \text{e} \quad d_n = b_n$$

cosicché per ogni n si ha

$$(c_n, d_n) \supset [a_n, b_n] \quad \text{e} \quad (d_n - c_n) = (1 + \varepsilon) \cdot (b_n - a_n)$$

$$\sum_n (d_n - c_n) = (1 + \varepsilon) \cdot \sum_n (b_n - a_n)$$

A questo punto poiché $\{(c_n, d_n)\}_{n \in \mathbb{N}}$ è un ricoprimento di aperti di $[0, 1 - \varepsilon]$ esiste $\bar{n}$ tale che

$$\bigcup_{n=1}^{\bar{n}} (c_n, d_n) \supseteq [0, 1 - \varepsilon]$$

per la subattività di μ_0

$$\sum_{n=1}^{\bar{n}} (b_n - a_n) = \frac{1}{1 + \varepsilon} \sum_{n=1}^{\bar{n}} (d_n - c_n) = \frac{1}{1 + \varepsilon} \sum_{n=1}^{\bar{n}} \mu_0([c_n, d_n)) \geq \frac{\mu_0([0, 1 - \varepsilon))}{1 + \varepsilon} \geq \frac{1 - \varepsilon}{1 + \varepsilon}$$

Per l'arbitrarietà di $\varepsilon > 0$ si ha:

$$\sum_n (b_n - a_n) \geq 1$$

□

Bibliografia

- [1] F. Caravenna, P. Dai Pra Probabilità, un introduzione attraverso modelli e applicazioni, *Springer*
- [2] David Williams Probability with Martingales *Cambridge mathematical textbooks*