Capitolo 1

Elementi della Teoria della Misura

Per fissare il quadro nel quale opereremo richiamiamo il concetto di tribú.

1.1 Tribú

Definizione 1.1.1. Dato un insieme non vuoto Ω, si chiama *algebra di sottoinsiemi* di Ω, ogni famiglia $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, non vuota, che sia stabile per la complementazione, per l'unione finita e tale che l'insieme vuoto appartenga ad $\mathcal{A}$; cioè:

(a) $\Omega \in \mathcal{A}$;

(b) $A \in \mathcal{A} \Longrightarrow A^c \in \mathcal{A}$;

(c) $A, B \in \mathcal{A} \Longrightarrow A \bigcup B \in \mathcal{A}$. ◇

Naturalmente, se $\mathcal{A}$ è un'algebra (di sottoinsiemi di Ω) e se A e B sono in $\mathcal{A}$, allora vi appartengono anche $A \cap B$, $A \setminus B$, e $A \Delta B$; infatti

$$A \bigcap B = \left(A^c \bigcup B^c\right)^c, \qquad A \setminus B = A \cap B^c,$$
$$A \Delta B = (A \setminus B) \bigcup (B \setminus A).$$

Definizione 1.1.2. Si chiama *tribú*, o σ*–algebra*, una famiglia $\mathcal{F}$ di sottoinsiemi di Ω, $\mathcal{F} \subseteq \mathcal{P}(\Omega)$, che goda delle seguenti proprietà:

(a) $\Omega \in \mathcal{F}$;

(b) $A \in \mathcal{F} \Longrightarrow A^c \in \mathcal{F}$;

(c) $\forall n \in \mathbb{N} \quad A_n \in \mathcal{F} \Longrightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$. ◇

Una tribú è dunque stabile rispetto all'operazione di unione numerabile. Usando le leggi di de Morgan è immediato dimostrare il seguente

Teorema 1.1.1. *Sia $\mathcal{F}$ una tribú di sottoinsiemi di Ω. Allora*

(a) *$\mathcal{F}$ è stabile per le unioni finite:*

$$A_i \in \mathcal{F}(i = 1, \dots, n) \implies \bigcup_{i=1}^{n} A_i \in \mathcal{F}\,;$$

(b) *$\mathcal{F}$ è stabile rispetto alle intersezioni numerabili:*

$$\forall\, n \in \mathbb{N} \quad A_n \in \mathcal{F} \implies \bigcap_{n\in\mathbb{N}} A_n \in \mathcal{F}\,;$$

(c) *$\mathcal{F}$ è stabile rispetto alle intersezioni finite:*

$$A_i \in \mathcal{F} \quad (i = 1, 2, \dots, n) \implies \bigcap_{i=1}^{n} A_i \in \mathcal{F}\,.$$

Si osservi che una tribú è anche un'algebra. Però non tutte le algebre sono anche tribú; si veda a tal fine l'Esercizio 1.2.

La classe delle tribú di sottoinsiemi di un insieme non vuoto Ω è ordinata, parzialmente, rispetto all'inclusione e contiene una piú piccola tribú, la *tribú banale*, $\mathcal{N} := \{\emptyset, \Omega\}$ ed una piú grande tribú, che è la famiglia delle parti $\mathcal{P}(\Omega)$, sicché, per ogni tribú $\mathcal{F}$, si ha $\mathcal{N} \subseteq \mathcal{F} \subseteq \mathcal{P}(\Omega)$.

Sia A un sottoinsieme proprio e non vuoto di Ω, cioè $A \neq \emptyset$ e $A \neq \Omega$; la famiglia $\mathcal{F}(A) := \{\emptyset, A, A^c, \Omega\}$ è un'algebra; è anzi, una tribú, poiché ogni algebra finita è anche una tribú, dato che ogni successione è necessariamente composta da un numero finito di insiemi distinti, sicché ogni unione numerabile è, di fatto, un'unione finita; essa è la piú piccola tribú che contenga A (e si dice *generata da* A). Infatti se $\mathcal{G}$ è una tribú che contiene A, risulta, per definizione,

$$A \in \mathcal{G}, \quad A^c \in \mathcal{G}, \quad \emptyset \in \mathcal{G}, \quad \Omega \in \mathcal{G}\,,$$

onde $\mathcal{F}(A) \subseteq \mathcal{G}$. Il teorema seguente è di dimostrazione banale.

Teorema 1.1.2. *Se $\{\mathcal{F}_\iota : \iota \in I\}$ è un'arbitraria famiglia di tribú di sottoinsiemi di Ω, è una tribú anche $\cap_{\iota\in I}\mathcal{F}_\iota$.*

Quest'ultimo risultato consente di risolvere il problema dell'esistenza della piú piccola tribú che contenga un'assegnata famiglia $\mathcal{C}$ di sottoinsiemi di Ω. Tale tribú si indica con $\mathcal{F}(\mathcal{C})$ e si dice *generata da* $\mathcal{C}$; essa è eguale all'intersezione di tutte le tribú in $\mathcal{P}(\Omega)$ che contengano $\mathcal{C}$. Si noti che la famiglia della quale si considera l'intersezione non è vuota perché vi appartiene almeno $\mathcal{P}(\Omega)$. Un esempio fondamentale di tale situazione è fornito dalla famiglia $\mathcal{I}$ degli intervalli aperti della retta reale $\mathbb{R}$, $\mathcal{I} := \{]a, b[\colon a, b \in \mathbb{R}, a \leq b\}$; si osservi che la condizione $a \leq b$, in luogo di quella piú naturale $a < b$, fa sí che l'insieme vuoto $\emptyset$ sia considerato come un particolare intervallo, ciò che è comodo. $\mathcal{I}$ non è un'algebra (e pertanto neanche una tribú), poiché se, ad esempio, $a < b < c < d$, l'unione $]a, b[\cup]c, d[$ non è un intervallo. La tribú generata da $\mathcal{I}$ si chiama *tribú di Borel* e la si denota con $\mathcal{B}(\mathbb{R})$ o, se non sorgono ambiguità, semplicemente con $\mathcal{B}$; i suoi insiemi si chiamano *boreliani*. Vale il seguente utile

Teorema 1.1.3. *La tribú di Borel $\mathcal{B}(\mathbb{R})$ è generata da una qualsiasi delle seguenti famiglie:*

(a) *le semirette del tipo* $]-\infty, x] \quad (x \in \mathbb{R})$*;*

(b) *gli insiemi aperti di* $\mathbb{R}$*;*

(c) *gli insiemi chiusi di* $\mathbb{R}$*.*

Dimostrazione. Sia $\mathcal{B} = \mathcal{B}(\mathbb{R})$ e si indichi con $\mathcal{B}_1$ la tribú generata dalla famiglia indicata in (a). Si osservi che anche gli intervalli aperti a sinistra e chiusi a destra, cioè del tipo $]a, b]$, con a e b in $\mathbb{R}$, appartengono a $\mathcal{B}$. Infatti

$$]a, b] = \bigcap_{n \in \mathbb{N}} \left]a, b + \frac{1}{n}\right[.$$

Ora $]x, +\infty[= \cup_{n\in\mathbb{N}}]x, x + n]$ che appartiene a $\mathcal{B}$ onde

$$]-\infty, x] =]x, +\infty[^c \in \mathcal{B}\,,$$

e perciò $\mathcal{B}_1 \subseteq \mathcal{B}$. D'altra parte, $]x, +\infty[=]-\infty, x]^c \in \mathcal{B}_1$ onde, se $x < y$, $]x, y] =]-\infty, y] \cap]x, +\infty[\in \mathcal{B}_1$. Infine $]x, y[= \cup_{n\in\mathbb{N}}]x, y - 1/n] \in \mathcal{B}_1$. Dunque $\mathcal{B} \subseteq \mathcal{B}_1$ e quindi $\mathcal{B}_1 = \mathcal{B}$.

Si indichi con $\mathcal{B}_2$ la tribú generata dagli aperti. Poiché l'intervallo $]x, y[$ è esso stesso un aperto, si ha $\mathcal{B} \subseteq \mathcal{B}_2$. Se $A \subseteq \mathbb{R}$ è aperto, esiste, com'è noto, una successione $(]x_n, y_n[)$ di intervalli aperti tale che si possa rappresentare A nella forma $A = \cup_{n\in\mathbb{N}}]x_n, y_n[$ onde $A \in \mathcal{B}$ e quindi $\mathcal{B}_2 \subseteq \mathcal{B}$. Che $\mathcal{B}$ sia generata anche dagli insiemi chiusi è ora immediato. $\square$

Se $\mathcal{F}$ è la tribú generata dalla famiglia $\mathcal{C}$, non si può, in generale, dare una descrizione costruttiva degli elementi di $\mathcal{F}$ partendo dagli elementi di $\mathcal{C}$.

Definizione 1.1.3. Si diranno *misurabili* (o, ove vi sia possibilità di confusione, *$\mathcal{F}$–misurabili*) gli insiemi appartenenti ad una prefissata tribú $\mathcal{F}$. Si dirà *spazio misurabile* la coppia $(\Omega, \mathcal{F})$ costituita da un insieme non vuoto Ω e da una tribú $\mathcal{F}$ di suoi sottoinsiemi.

1.2 Richiami di probabilità elementare

Richiamo qui concetti e definizioni già noti al lettore del corso introduttivo di probabilità. Alcuni dei concetti richiamati saranno in seguito estei o generalizati. La definizione di probabilità è formalizzata nella sezione successiva, Definizione 1.3.1. Dato uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, una funzione misurabile $X : \Omega \to \mathbb{R}$ si dice *variabile aleatoria*. Tale concetto sarà poi esteso al caso di funzioni che assumano anche i valori $\pm\infty$.Si dice che una variabile aleatoria X è *discreta* quando assume un numero finito di valori.

In uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ due insiemi A e B in $\mathcal{F}$ si dicono *indipendenti* se si ha

$$\mathbb{P}\left(A \bigcap B\right) = \mathbb{P}(A)\,\mathbb{P}(B)\,.$$

Due variabili aleatorie X e Y definite nello stesso spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ si dicono *indipendenti* se, per ogni scelta di due boreliani A e B, si ha

$$\mathbb{P}\left(X^{-1}(A) \cap Y^{-1}(B)\right) = \mathbb{P}\left(X^{-1}(A)\right) \mathbb{P}\left(Y^{-1}(B)\right) .$$

Questi concetti si estendono immediatamente a piú di due insiemi o di due variabili aleatorie.

Ricordiamo che l'indipendenza di insiemi è legata all'indipendenza di tribú.

Date n tribú $\mathcal{F}_1, \dots, \mathcal{F}_n$ contenute in $\mathcal{F}$, queste si dicono indipendenti se per ogni scelta di n insiemi $A_1 \in \mathcal{F}_1, \dots, A_n \in \mathcal{F}_n$ risulta

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mathbb{P}(A_i) .$$

Due insiemi A e B sono indipendenti se, e solo se, sono indipendenti le tribuú $\mathcal{F}(A)$ e $\mathcal{F}(B)$ che essi generano. Per verificare che gli insiemi misurabili $A_1, \dots, A_n$ siano indipendenti occorre allora controllare che risulti

$$\mathbb{P}\left(\bigcap_{i=1}^{n} \overline{A_i}\right) = \prod_{i=1}^{n} \mathbb{P}(\overline{A_i}) ,$$

ove, per ogni indice i, $\overline{A_i}$ può essere uno degli insiemi A_i, A_i^c, Ω o $A_i = \emptyset$.

Gli insiemi misurabili $(A_k)_{k\in\mathbb{N}}$ di una successione si dicono indipendenti se per ogni scelta di $n \geq 2$ sono indipendenti gli insiemi $A_1, \dots, A_n$.

Siano date n variabili aleatorie $X_1, \dots, X_n$ nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ queste si dicono *indipendenti* se risulta

$$\mathbb{P}\left(\bigcap_{k=1}^{n} X^{-1}(B_k)\right) = \prod_{k=1}^{n} \mathbb{P}\left(X^{-1}(B_k)\right)$$

per ogni scelta di n boreliani $B_1, \dots, B_n$.

Infine le variabili aleatorie $(X_k)_{k\in\mathbb{N}}$ di una successione si dicono indiependenti se sono indipendenti $X_1, \dots, X_n$ per ogni scelta di $n \geq 2$.

Le variabili aleatorie discrete piú semplici sono quelle di Bernoulli: Y è una variabile aleatoria di Bernoulli se assume due soli valori, convenzionalmente chiamati *successo* e *fallimento* e indicati con s oppure 1 e f oppure 0; cosí $\mathbb{P}(Y = 1) = p$ e $\mathbb{P}(Y = 0) = q = 1 - p$. Il processo di Bernoulli, già incontrato nel primo corso di probabilità, è costituito da una successione (X_n) di variabili aleatorie indipendenti tutte con la stessa legge di Bernoulli, $\mathbb{P}(X_n = 1) = 1$ e $\mathbb{P}(X_n = 0) = q = 1 - p$.

Dato un insieme $A \subset \Omega$ la sua *funzione indicatrice* $\mathbf{1}_A$ è cosí definita

$$\mathbf{1}_A(\omega) := \begin{cases} 1 & \omega \in A , \\ 0, & \omega \notin A . \end{cases}$$

Nel seguito si indicherà il limite a sinistra in $x_0 \in \mathbb{R}$ di una funzione $f : \mathbb{R} \to \mathbb{R}$, se esiste, mediante la notazione

$$\ell^- f(x_0) := \lim_{\substack{x\to x_0 \\ x<x_0}} f(x) .$$

1.3 Probabilità e misura

Ricordiamo la definizione di probabilità già nota dal corso introduttivo.

Definizione 1.3.1. Dato uno spazio misurabile $(\Omega, \mathcal{F})$ — vale a dire un insieme non vuoto Ω ed una tribú $\mathcal{F}$ di suoi sottoinsiemi — si dice *(misura di) probabilità* su $(\Omega, \mathcal{F})$ ogni funzione $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ che soddisfaccia alle seguenti condizioni:

(P.1) $\mathbb{P}(A) \geq 0$ per ogni insieme $A \in \mathcal{F}$;

(P.2) $\mathbb{P}(\Omega) = 1$;

(P.3) per ogni successione $(A_n)_{n\in\mathbb{N}}$ di insiemi misurabili disgiunti ($A_n \in \mathcal{F}$, per ogni $n \in \mathbb{N}$, con $A_j \cap A_k = \emptyset$ $(j \neq k)$), vale la proprietà:

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

detta di *additività numerabile* o σ*–additività.* ◇

Anche nel caso discreto, a ben guardare, nasce la necessità di studiare situazioni piú complesse. Infatti, nello studio del processo di Bernoulli, viene naturale considerare spazî dei risultati Ω che non sono finiti o numerabili: cosí una generica "storia" può essere rappresentata mediante una successione

$$(x_1, \dots, x_n, \dots)$$

ove, per ogni $n \in \mathbb{N}$, $x_n = 0$ oppure $x_n = 1$; ma è noto che tali successioni possono essere poste in corrispondenza biunivoca con i numeri dell'intervallo $[0, 1]$, mediante

$$x = \sum_{n\in\mathbb{N}} \frac{x_n}{2^n}\,,$$

se si adotta la convenzione che i numeri diadici siano rappresentati dalla successione che ha infiniti zeri. È inoltre noto che l'intervallo $[0, 1]$ non è numerabile.

Quando, nello studio della passeggiata aleatoria di Bernoulli, abbiamo considerato lo spazio di tutte le possibili storie, $\Omega := \{0, 1\}^{\mathbb{N}}$, abbiamo lasciato in sospeso la questione di quale sia la tribú da adottare per definire le probabilità atte a rispondere alle domande che ci ponemmo allora. Si consideri, per ogni $n \in \mathbb{N}$ il "segmento iniziale" $\Omega_n := \{0, 1\}^n$, ed, in questo si consideri la famiglia delle parti, $\mathcal{F}'_n := \mathcal{P}(\Omega_n)$. Si consideri ora la famiglia di sottoinsiemi di $\{0, 1\}^{\mathbb{N}}$ data da

$$\mathcal{F}_n := \mathcal{F}'_n \times \prod_{k=n}^{\infty} \{0, 1\}\,;$$

evidentemente, la famiglia $\mathcal{F}_n$ è adeguata per affrontare tutti i problemi che riguardano questioni sul processo di Bernoulli sino al tempo $t = n$. Si può qundi considerare la famiglia di sottoinsiemi di Ω definita da

$$\bigcup_{n\in\mathbb{N}} \mathcal{F}_n\,.$$

Questa, tuttavia non è una tribú; sarà, perciò naturale porci nella tribú generata da tale famiglia

$$\bigvee_{n\in\mathbb{N}} \mathcal{F}_n := \mathcal{F}\left(\bigcup_{n\in\mathbb{N}} \mathcal{F}_n\right).$$

In generale, si pone dunque il problema di definire le probabilità senza porre restrizioni sulla cardinalità di Ω. Storicamente, si sono presentate due strade che, entrambe, delineeremo, seppur succintamente.

La prima strada consiste nel partire, dato un insieme non vuoto Ω, da un'algebra $\mathcal{A}$ di suoi sottoinsiemi; è naturale considerare un'algebra poiché, come si è visto, questa è stabile rispetto alle operazioni sugli insiemi, purché eseguite in numero finito. Una funzione $\mu : \mathcal{A} \to [0, +\infty]$ si dice *finitamente additiva* se

$$\mu\left(A \bigcup B\right) = \mu(A) + \mu(B)\,, \tag{1.3.1}$$

per ogni coppia di insiemi disgiunti A e B di $\mathcal{A}$, $A \cap B = \emptyset$.

In particolare, una probabilità $\mathbb{P} : \mathcal{A} \to [0, 1]$ è una funzione è finitamente additiva se obbedisce alla (1.3.1), che prende il nome di additività semplice e che si scrive

$$\mathbb{P}\left(A \bigcup B\right) = \mathbb{P}(A) + \mathbb{P}(B) \qquad (A, B \in \mathcal{A},\ A \bigcap B = \emptyset)\,.$$

Vale il seguente

Teorema 1.3.1. (Tarski). *Ogni misura finitamente additiva e finita μ definita su un'algebra $\mathcal{A}$ di sottoinsiemi di Ω può essere estesa a $\mathcal{P}(\Omega)$; vale a dire che esiste una misura finitamente additiva e finita $\nu : \mathcal{P}(\Omega) \to [0, 1]$ tale che risulti $\nu(A) = \mu(A)$ per ogni insieme A di $\mathcal{A}$.*

La dimostrazione si può dare agevolmente (ma non è la dimostrazione originale di Tarski) mediante il teorema di Hahn–Banach; esso presenta però l'inconveniente che l'estensione ν di μ non è, solitamente, unica. Di contro, nel caso delle probabilità, non si pongono restrizioni ai sottoinsiemi di Ω che possono essere riguardati come eventi, ai quali, cioè, si può attribuire una probabilità.

Vi è un secondo modo di procedere, che consiste nel richiedere che la funzione $\mathbb{P}$ soddisfaccia, anziché all'assioma di additività semplice, a quello piú forte di *additività numerabile* (o σ–*additività*), che abbiamo già incontrato.

Naturalmente, occorre che la famiglia di sottoinsiemi sulla quale è definita la probabilità sia tale che, se vi sono contenuti tutti gli insiemi di una successione di insiemi, vi appartenga anche la loro unione. Perciò è naturale definire la probabilità, in questa seconda accezione, in una tribu $\mathcal{F}$ di sottoinsiemi di Ω.

Poiché adotteremo, come abbiamo già fatto, e come è tradizione, questo secondo punto di vista, la proprietà di additività numerabile rende opportuna una digressione attraverso le misure, dato che le probabilità sono particolari misure. Tuttavia, mettiamo in guardia il lettore dal pensare che la probabilità sia un mero esempio di misura con la restrizione che sia eguale a 1 la misura di tutto lo spazio, o come spesso si dice, sia una misura normalizzata. Bastano i soli concetti di indipendenza e di condizionamento per mostrare la ricchezza della Probabilità.

1.4 Successioni di insiemi

Le definizioni che seguono, e che non dipendono da una topologia, sono giustificate, *a posteriori*, dal legame con i concetti dello stesso nome per le funzioni a valori reali (si vedano gli esercizî). Per aiutare l'intuizione, si tenga presente che le definizioni ricalcano quelle per le successioni di numeri reali, se si fa corrispondere all'estremo superiore e all'estremo inferiore tra numeri reali l'unione e l'intersezione, rispettivamente, tra insiemi. Naturalmente, quest'analogia serve solo ad intuire i risultati, ma non a dimostrarli.

Definizione 1.4.1. Data una successione $(A_n)_{n\in\mathbb{N}}$ di sottoinsiemi di Ω l'insieme

$$\limsup_{n\to\infty} A_n := \bigcap_{n\in\mathbb{N}} \bigcup_{j\geq n} A_j$$

si chiama *limite superiore* o *massimo limite* di (A_n) mentre l'insieme

$$\liminf_{n\to\infty} A_n := \bigcup_{n\in\mathbb{N}} \bigcap_{j\geq n} A_j$$

si dice *limite inferiore* o *minimo limite* di (A_n). ◇

Gli insiemi $\limsup_{n\to\infty} A_n$ e $\liminf_{n\to\infty} A_n$ che si ottengono mediante operazioni numerabili sono misurabili se tale è ogni insieme della successione (A_n).

L'insieme $\limsup_{n\to\infty} A_n$ è costituito da tutti, e soli, i punti di Ω che appartengono ad infiniti insiemi della successione (A_n); dire, infatti, che il punto ω appartiene a $\limsup_{n\to\infty} A_n = \cap_{n\in\mathbb{N}} \cup_{j\geq n} A_j$ equivale a dire che, per ogni numero naturale n, esiste un altro naturale $j \geq n$ tale che $\omega \in A_j$.

Analogamente, l'insieme $\liminf_{n\to\infty} A_n$ è costituito da tutti, e soli, i punti di Ω che appartengono definitivamente alla successione, tali cioè da appartenere a tutti gli insiemi della succesione $\{A_n\}$ tranne, al piú, ad un numero finito di essi. Infatti, dire che ω è nell'insieme $\liminf_{n\to\infty} A_n = \cup_{n\in\mathbb{N}} \cap_{j\geq n} A_j$ equivale a dire che esiste un naturale n tale che $\omega \in A_j$ per ogni $j \geq n$.

Valgono le seguenti inclusioni

$$\bigcap_{n\in\mathbb{N}} A_n \subseteq \liminf_{n\to+\infty} A_n \subseteq \limsup_{n\to+\infty} A_n \subseteq \bigcup_{n\in\mathbb{N}} A_n \,. \tag{1.4.1}$$

L'inclusione centrale segue da quanto precede. Per stabilire la (1.4.1), basta perciò mostrare una delle rimanenti inclusioni, per esempio la prima, l'altra dimostrandosi per passaggio al complementare.

Definizione 1.4.2. Si dirà che la successione (A_n) di sottoinsiemi di Ω *converge* se si ha

$$\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = A \,.$$

In tal caso, si scrive $\lim_{n\to+\infty} A_n = A$ (oppure $A_n \to A$). Una successione (A_n) si dirà *crescente* se è $A_n \subseteq A_{n+1}$ per ogni $n \in \mathbb{N}$, *decrescente* se è $A_n \supset A_{n+1}$ per ogni $n \in \mathbb{N}$, *monotona* se è o crescente o decrescente. ◇

Il teorema che segue è l'analogo del teorema sulle successioni reali monotone.

Teorema 1.4.1. *Ogni successione (A_n) monotona converge e risulta:*

(a) $\lim_{n\to+\infty} A_n = \bigcap_n A_n$ *se* (A_n) *è decrescente;*

(b) $\lim_{n\to+\infty} A_n = \bigcup_n A_n$ *se* (A_n) *è crescente.*

Dimostrazione. Basta dimostrare una sola delle due eguaglianze, l'altra ottenendosi per complementazione. Si supponga, dunque, che (A_n) sia crescente; allora

$$\limsup_{n\to+\infty} A_n = \bigcap_{n\in\mathbb{N}} \bigcup_{j\geq n} A_j = \bigcap_{n\in\mathbb{N}} \bigcup_{j\in\mathbb{N}} A_j = \bigcup_{j\in\mathbb{N}} A_j \,,$$
$$\liminf_{n\to+\infty} A_n = \bigcup_{n\in\mathbb{N}} \bigcap_{j\geq n} A_j = \bigcup_{n\in\mathbb{N}} A_n \,,$$

che dimostra l'asserto. □

1.5 Misure

Diamo qui una piú ampia definizione di misura che consentirà per esempio lo studio delle estensioni di una misura.

Definizione 1.5.1. Sia $\mathcal{C}$ una famiglia di sottoinsiemi di un insieme non vuoto Ω. Si dice *misura* su $\mathcal{C}$ una funzione $\mu : \mathcal{C} \to \overline{\mathbb{R}}_+$ tale che:

(a) non sia identicamente eguale a $+\infty$ (esiste cioè un insieme $A \in \mathcal{C}$ con $\mu(A) < +\infty$);

(b) μ sia numerabilmente additiva, nel senso che, se (A_n) è una successione di insiemi disgiunti di $\mathcal{C}$, l'unione dei quali è ancora in $\mathcal{C}$,

$$\forall\, n \in \mathbb{N} \quad A_n \in \mathcal{C} \quad A_i \bigcap A_j = \emptyset \quad (i \neq j) \quad \text{e} \quad \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{C} \,,$$

si ha

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n) \,. \tag{1.5.1}$$

Di solito, $\mathcal{C}$ è un'algebra o una tribú; in quest'ultimo caso non occorre postulare che l'unione $\cup_{n\in\mathbb{N}} A_n$ appartenga a $\mathcal{C}$. ◇

Nel parlare di misure è piú comodo e naturale porsi, anziché in $\mathbb{R}$ o in $\mathbb{R}_+$, in $\overline{\mathbb{R}} := [-\infty, +\infty]$. Si noti che $\overline{\mathbb{R}}$ *non* è un gruppo giacché, per ogni reale x, è $x + \infty = +\infty$ e $x - \infty = -\infty$; non è definita l'operazione $\infty - \infty$.

Valgono, inoltre, le convenzioni $0 \cdot (\pm\infty) = 0$ e

$$\begin{gathered}
\forall\, x \in \mathbb{R} \qquad -\infty < x < +\infty \,, \\
\forall\, x > 0 \qquad x \cdot (+\infty) = +\infty \,, \quad \text{e} \quad x \cdot (-\infty) = -\infty \,, \\
\forall\, x < 0 \qquad x \cdot (+\infty) = -\infty \,, \quad \text{e} \quad x \cdot (-\infty) = +\infty \,, \\
(\pm\infty) \cdot (\pm\infty) = +\infty \,, \qquad (\mp\infty) \cdot (\pm\infty) = -\infty \,.
\end{gathered}$$

Se l'insieme vuoto $\emptyset$ appartiene a $\mathcal{C}$, si ha $\mu(\emptyset) = 0$. Infatti sia $A \in \mathcal{C}$ tale che $\mu(A) < +\infty$ e si consideri la successione (A_n) di elementi di $\mathcal{C}$ cosí definita: $A_1 := A$, $A_n = \emptyset$ per ogni $n \geq 2$; allora la (1.5.1) dà

$$\mu(A) = \mu(A) + \mu(\emptyset) + \dots + \mu(\emptyset) + \dots$$

cioè $\mu(\emptyset) = 0$. Similmente, si prova che μ è finitamente additiva. Se gli insiemi A, B e $A \cup B$ appartengono a $\mathcal{C}$, basta considerare la successione (A_n) definita da $A_1 := A$, $A_2 := B$, $A_n := \emptyset$ per $n \geq 3$. Dall'additività semplice scende anche che una misura è isotona: se $A \subseteq B$ e se gli insiemi A, B e $B \setminus A$ sono in $\mathcal{C}$, allora $\mu(A) \leq \mu(B)$.

Definizione 1.5.2. Dato un insieme non vuoto Ω, si dice *anello* una famiglia $\mathcal{R}$ di sottoinsiemi di Ω che sia stabile rispetto all'intersezione ed alla differenza finita

$$A \bigcap B \in \mathcal{R} \quad \text{e} \quad A \Delta B \in \mathcal{R},$$

per ogni coppia A e B di insiemi di $\mathcal{R}$. $\diamond$

Questa definizione assicura che il risultato di tutte le operazioni eseguite sopra un numero finito di insiemi dell'anello $\mathcal{R}$ appartiene ancora ad $\mathcal{R}$ (si vedano, a tal fine, gli esercizî). Si deve eccettuare la complementazione, a meno che Ω non appartenga a $\mathcal{R}$.

Prima di enunciare il prossimo risultato, che dà un criterio utile, spesso utilizzato, per stabilire per la σ–additività, conviene introdurre la definizione di funzione finitamente additiva.

La proprietà espressa dalla (1.3.1) vale, in particolare, per le probabilità. Dalla (1.3.1) scende, come immediata conseguenza, l'isotonia della funzione d'insieme μ, vale a dire che, se A e B sono due insiemi di $\mathcal{R}$ con $A \subseteq B$, allora $\mu(A) \leq \mu(B)$.

Teorema 1.5.1. *Sia $\mathcal{R}$ un anello di sottoinsiemi di Ω e sia $\mu : \mathcal{R} \to \overline{\mathbb{R}}_+$ una funzione d'insieme finitamente additiva; allora*

(a) *se μ è σ–additiva, essa passa al limite lungo tutte le successioni crescenti di insiemi di $\mathcal{R}$; in altre parole, se (A_n) è una successione di insiemi di $\mathcal{R}$ con*

$$A_n \subseteq A_{n+1} \quad e \quad \bigcup_{n \in \mathbb{N}} A_n = A \in \mathcal{R},$$

allora $\lim_{n \to +\infty} \mu(A_n) = \mu(A)$*;*

(b) *se μ è σ–additiva, essa passa al limite lungo tutte le successioni decrescenti di insiemi di $\mathcal{R}$ per le quali esista un indice k tale che $\mu(A_k) < +\infty$; in altre parole, se (A_n) è una successione di insiemi di $\mathcal{R}$ con*

$$A_n \supset A_{n+1}, \quad \bigcap_{n \in \mathbb{N}} A_n = A \in \mathcal{R}$$

ed esiste almeno un indice $k \in \mathbb{N}$ per il quale $\mu(A_k) < +\infty$, allora

$$\lim_{n \to +\infty} \mu(A_n) = \mu(A);$$

(c) *se μ passa al limite lungo tutte le successioni crescenti di insiemi di $\mathcal{R}$, cioè, se per ogni successione (A_n) di insiemi di $\mathcal{R}$ con $A_n \subseteq A_{n+1}$ per ogni $n \in \mathbb{N}$, e $\cup_{n\in\mathbb{N}} A_n = A \in \mathcal{R}$, è $\lim_{n\to+\infty} \mu(A_n) = \mu(A)$, allora μ è σ–additiva;*

(d) *se μ passa al limite lungo tutte le successioni decrescenti di insiemi di $\mathcal{R}$ che tendono all'insieme vuoto $\emptyset$, cioè, se per ogni successione (A_n) di insiemi di $\mathcal{R}$ con $A_n \supset A_{n+1}$ e $\cap_{n\in\mathbb{N}} A_n = \emptyset$, è $\lim_{n\to+\infty} \mu(A_n) = 0$, allora μ è σ–additiva.*

Dimostrazione. (a) Si supponga che $A_n \uparrow A = \cup_{n\in\mathbb{N}} A_n$. Se esiste un indice k tale che $\mu(A_k) = +\infty$, si ha $\mu(A_n) \geq \mu(A_k) = +\infty$ per ogi $n \geq k$, sicché $\mu(A) = +\infty$ e $\mu(A_n) \to \mu(A)$. Possiamo quindi supporre che sia $\mu(A_n) < +\infty$ per ogni $n \in \mathbb{N}$. Si scriva A come unione disgiunta,

$$A = A_1 \bigcup (A_2 \setminus A_1) \bigcup (A_3 \setminus A_2) \bigcup \dots = A_1 \bigcup \left(\bigcup_{n\in\mathbb{N}} (A_{n+1} \setminus A_n) \right) ,$$

onde, grazie alla σ–additività,

$$\begin{aligned} \mu(A) &= \mu(A_1) + \sum_{n\in\mathbb{N}} \mu\left(A_{n+1} \setminus A_n\right) \\ &= \mu(A_1) + \lim_{k\to+\infty} \sum_{n=1}^{k} \mu\left(A_{n+1} \setminus A_n\right) \\ &= \lim_{k\to+\infty} \left(\mu(A_1) + \sum_{n=1}^{k} \{\mu(A_{n+1}) - \mu(A_n)\} \right) = \lim_{k\to+\infty} \mu(A_{k+1}) . \end{aligned}$$

(b) Si supponga che esista un indice k tale che $\mu(A_k) < +\infty$ e si ponga, per ogni $n \geq k$, $C_n := A_k \setminus A_n$. La successione $(C_n)_{n\geq k}$ è tutta costituita da insiemi di $\mathcal{R}$, è crescente ed ammette come limite l'insieme $A_k \setminus A$; per quanto dimostrato in (a), si ha $\mu(C_n) \to \mu(A_k \setminus A) = \mu(A_k) - \mu(A)$. Poiché $\mu(C_n) = \mu(A_k) - \mu(A_n)$, segue che $\mu(A_n) \to \mu(A)$.

(c) Sia (A_n) una successione di insiemi disgiunti di $\mathcal{R}$ e si supponga che anche l'unione $A = \cup_{n\in\mathbb{N}} A_n$ appartenga a $\mathcal{R}$. Posto $E_n := \cup_{j=1}^{n} A_j$, si ha che gli insiemi E_n appartengono a $\mathcal{R}$ per ogni $n \in \mathbb{N}$ e che la successione (E_n) è crescente e tende a A, sicché

$$\sum_{j=1}^{n} \mu(A_j) = \mu(E_n) \xrightarrow[n\to+\infty]{} \mu(A) \quad \text{cioè} \quad \sum_{n\in\mathbb{N}} \mu(A_n) = \mu(A) .$$

(d) Con la stessa notazione di (c), si ponga $F_n := A \setminus E_n = \cup_{j=n+1}^{\infty} A_j$; allora F_n appartiene a $\mathcal{R}$ per ogni $n \in \mathbb{N}$, la successione (F_n) è decrescente e $\cap_{n\in\mathbb{N}} F_n = \emptyset$. Scende dall'ipotesi fatta che $\mu(F_n) \downarrow 0$, e, poiché $\mu(A) = \sum_{j=1}^{n} \mu(A_j) + \mu(F_n)$, si ha

$$\mu(A) = \sum_{j\in\mathbb{N}} \mu(A_j) ,$$

che conclude la dimostrazione. □

È immediato il corollario

Corollario 1.5.1. *Per una funzione $\mu : \mathcal{R} \to \mathbb{R}_+$, finita e finitamente additiva, sono equivalenti le condizioni*

(a) μ *è* σ*–additiva;*

(b) $\forall\, n \in \mathbb{N} \quad A_n \in \mathcal{R}, A_n \uparrow A \in \mathcal{R} \Longrightarrow \mu(A_n) \uparrow \mu(A)$*;*

(c) $\forall\, n \in \mathbb{N} \quad A_n \in \mathcal{R}, A_n \downarrow A \in \mathcal{R} \Longrightarrow \mu(A_n) \downarrow \mu(A)$.

Nel seguito chiameremo *spazio mensurale* ogni terna $(\Omega, \mathcal{F}, \mu)$ formata da un insieme non vuoto Ω, da una tribú $\mathcal{F}$ di sottoinsiemi di Ω e da una misura (σ–additiva) μ definita in $\mathcal{F}$. Se poi risulta $\mu(\Omega) = 1$, μ si dirà *(misura di) probabilità* e la terna $(\Omega, \mathcal{F}, \mu)$ si dirà *spazio di probabilità*; in tal caso, si usa solitamente $\mathbb{P}$ in luogo di μ per indicare la probabilità.

È bene sottolineare che una probabilità $\mathbb{P}$ è definita in una tribú $\mathcal{F}$ che, in genere, differisce da $\mathcal{P}(\Omega)$ e che a $\mathbb{P}$ si richiede di soddisfare all'assioma (P.3) di additività numerabile. Ci si può domandare se, come per le misure finitamente additive, sia possibile estendere una probabilità $\mathbb{P}$ dalla tribú $\mathcal{F}$ a tutta la famiglia delle parti $\mathcal{P}(\Omega)$; tale possibilità è esclusa, in generale, dal seguente teorema del quale non daremo la dimostrazione.

Teorema 1.5.2. (Ulam). *Se Ω è un insieme non numerabile* card$(\Omega) > \aleph_0$, *la sola misura μ definita su $\mathcal{P}(\Omega)$ e che si annulli su tutti i punti di Ω, vale a dire, tale che, per ogni $\omega \in \Omega$, $\mu(\{\omega\}) = 0$, è quella identicamente nulla, $\mu(A) = 0$ per ogni sottoinsieme A di Ω.*

A chi voglia trattare di probabilità si presenta dunque un'alternativa: considerare come eventi tutti i possibili sottoinsiemi di un insieme Ω ed usare probabilità finitamente additive, oppure, rinunciare a considerare come eventi tutti i sottoinsiemi di Ω, restringendo *a priori* la classe degli eventi ad una tribú $\mathcal{F}$, ma usare probabilità numerabilmente additive. La strada che tradizionalmente si segue è la seconda, soprattutto in virtú della maggiore ricchezza di risultati che deriva da un assioma piú forte (additività numerabile invece che semplice). Esistono però fenomeni nei quali solo l'uso di probabilità finitamente additive consente di trarre previsioni in accordo con le osservazioni; un esempio importante è costituito dalla cosiddetta legge di Benford, ossia il fatto che nelle serie di numeri la prima cifra significativa non sia distribuita uniformemente (si vedano a tal fine i lavori (Scozzafava, 1981) e (Fuchs & Letta, 1984). Tuttavia, come già detto, in queste lezioni adotteremo sistematicamente gli assiomi di Kolmogorov.

Il seguente esempio non contraddice al Teorema 1.5.2.

Esempio 1.5.1. Sia ω_0 un punto dell'insieme non vuoto Ω. Si dice *misura di Dirac concentrata in* ω_0 la probabilità $\delta_{\omega_0} : \mathcal{P}(\Omega) \to [0, 1]$ definita, per ogni sottoinsieme A di Ω, da $\delta_{\omega_0}(A) := \mathbf{1}_A(\omega_0)$. Si osservi che la condizione del Teorema 1.5.2 è violata nell'unico punto ω_0. ■

Osserviamo infine che le parti (b) e (c) del corollario 1.5.1 possono essere rafforzate: per una probabilità $\mathbb{P}$ e per ogni successione $(A_n) \subseteq \mathcal{F}$, non necessariamente monotona, convergente a un insieme $A \in \mathcal{F}$, si ha $\mathbb{P}(A_n) \xrightarrow[n \to +\infty]{} \mathbb{P}(A)$. (Si vedano gli esercizî).

1.6 Estensione di misure

Ci proponiamo di affrontare in questa sezione il problema, che si pone spesso, di estendere una data misura da una famiglia $\mathcal{C}$ ad una famiglia piú ampia che la contiene.

Si è già incontrata la definizione di anello e quella di algebra di una famiglia di sottoinsiemi di Ω. Introdurremo nel seguito altre famiglie di sottoinsiemi la cui considerazione è utile trattando di probabilità.

Definizione 1.6.1. Una famiglia $\mathcal{C}$ di sottinsiemi di Ω che sia stabile rispetto all'intersezione finita,

$$\forall A, B \in \mathcal{C} \quad A \bigcap B \in \mathcal{C}$$

si dice π*–classe* o π*–sistema.* ◇

Definizione 1.6.2. Si dice *semi–anello* una famiglia $\mathcal{S}$ di sottoinsiemi di Ω che contenga l'insieme vuoto, che sia una π–classe e tale che per ogni coppia di insiemi A e B in $\mathcal{S}$ esista una partizione finita di $A \setminus B$ in insiemi di $\mathcal{S}$:

(a) $\emptyset \in \mathcal{S}$;

(b) $A, B \in \mathcal{S} \quad \Longrightarrow \quad A \bigcap B \in \mathcal{S}$;

(c) per ogni coppia A e B di insiemi in $\mathcal{S}$, si può esprimere la loro differenza $A \setminus B$ come unione disgiunta di una famiglia finita di insiemi di $\mathcal{S}$. ◇

Un esempio notevole di semi–anello è dato, in $\mathbb{R}$, dalla famiglia $\mathcal{I}$ degli intervalli $]a, b]$ aperti a sinistra e chiusi a destra; in $\mathbb{R}^n$ è un semi–anello la famiglia $\mathcal{I}^n$ dei *rettangoli*

$$R =]a_1, b_1] \times]a_2, b_2] \times \cdots \times]a_n, b_n] \ .$$

Definizione 1.6.3. Una famiglia non vuota di sottoinsiemi di Ω si dice *classe monotona* se è stabile rispetto ai limiti di successioni monotone. ◇

È evidente che una tribú è una classe monotona; viceversa, ogni algebra $\mathcal{A}$ che sia anche una classe monotona è una tribú; infatti, se $A_n \in \mathcal{A}$ per ogni $n \in \mathbb{N}$, allora, posto $B_n := \cup_{j=1}^n A_j \in \mathcal{A}$, si ha $B_n \uparrow \cup_{n \in \mathbb{N}} A_n$. Un anello che sia anche una classe monotona non è necessariamente una tribú perché potrebbe non appartenervi l'insieme Ω.

Si dimostra facilmente che, data un'arbitraria famiglia non vuota $\mathcal{C}$ di sottoinsiemi di Ω, esistono un piú piccolo anello ed una piú piccola classe monotona che contengono $\mathcal{C}$; li si denotano rispettivamente con $\mathcal{R}(\mathcal{C})$ e con $m(\mathcal{C})$ e si chiamano rispettivamente *anello* e *classe monotona generati da* $\mathcal{C}$.

Teorema 1.6.1. (della classe monotona). *La classe monotona $m(\mathcal{A})$ e la tribú $\mathcal{F}(\mathcal{A})$ generate da un'algebra $\mathcal{A}$ coincidono ($m(\mathcal{A}) = \mathcal{F}(\mathcal{A})$).*

Dimostrazione. L'inclusione $m(\mathcal{A}) \subseteq \mathcal{F}(\mathcal{A})$ è ovvia. Per dimostrare l'inclusione inversa, basta far vedere che $m(\mathcal{A})$ è una tribú e questo, a sua volta, scenderà dallo stabilire che $m(\mathcal{A})$ è un'algebra. A tale scopo, per ogni $B \subseteq \Omega$, si ponga

$$\mathcal{D}(B) := \left\{ A \subseteq \Omega : A \setminus B, B \setminus A, A \bigcup B \in m(\mathcal{A}) \right\} .$$

Se (A_n) è una successione decrescente di elementi di $\mathcal{D}(B)$ e $A = \cap_{n\in\mathbb{N}} A_n$, allora $(A_n \setminus B)$, $(B \setminus A_n)$, $(A_n \cup B)$ sono successioni monotone in $m(\mathcal{A})$ sicché i limiti $A \setminus B$, $B \setminus A$, $A \cup B$ appartengono a $m(\mathcal{A})$ onde A appartiene a $\mathcal{D}(B)$. Similmente si procede per una successione crescente. Perciò, $\mathcal{D}(B)$ è una classe monotona per ogni $B \subseteq \Omega$. Inoltre, se B è in $\mathcal{A}$, si ha $\mathcal{D}(B) \supset \mathcal{A}$. Poiché A appartiene a $\mathcal{D}(B)$ se, e solo se, B appartiene a $\mathcal{D}(A)$, necessariamente si ha $\mathcal{D}(B) \supset m(\mathcal{A})$ se B è in $m(\mathcal{A})$. Ciò significa che, se A e B sono in $m(\mathcal{A})$ vi sono anche $A \setminus B$, $B \setminus A$, $A \cup B$, sicché $m(\mathcal{A})$ è un anello; è, anzi, un'algebra perché Ω appartiene a $\mathcal{A}$ e quindi a $m(\mathcal{A})$. □

Il seguente risultato è talvolta utile perché caratterizza gli insiemi di un anello generato da un semi–anello.

Teorema 1.6.2. *L'anello $\mathcal{R}(\mathcal{S})$ generato da un semi–anello $\mathcal{S}$ è costituito da tutti, e soli, gli insiemi che sono un'unione disgiunta e finita di insiemi di $\mathcal{S}$:*

$$\mathcal{R}(\mathcal{S}) = \bigcup_{n\in\mathbb{N}} \left\{ E = \bigcup_{j=1}^{n} A_j : A_j \in \mathcal{S}, A_i \cap A_j = \emptyset \ (i \neq j) \right\} .$$

Dimostrazione. è evidente che $\mathcal{R}(\mathcal{S})$ contiene gli insiemi della forma detta perché un anello è stabile rispetto alle unioni finite. Viceversa, sia

$$\mathcal{V} := \bigcup_{n\in\mathbb{N}} \left\{ E = \bigcup_{j=1}^{n} A_j : A_j \in \mathcal{S}, \ A_i \bigcap A_j = \emptyset \ (i \neq j) \right\} .$$

Poiché è ovvio che $\mathcal{V} \supset \mathcal{S}$, basta mostrare che $\mathcal{V}$ è un anello. Siano E e F in $\mathcal{V}$ con $E = \cup_{i=1}^{n} A_i$ e $F = \cup_{j=1}^{m} B_j$ e si ponga $C_{ij} := A_i \cap B_j$, ottenendo una famiglia disgiunta di insiemi di $\mathcal{S}$. Ora, $E \cap F = \cup_{i=1}^{n} \cup_{j=1}^{m} C_{ij}$, sicché $\mathcal{V}$ è una π–classe. Per induzione, scende dalla definizione di semi–anello che

$$A_i = \left(\bigcup_{j=1}^{n} C_{ij} \right) \bigcup \left(\bigcup_{k=1}^{r(i)} D_{ik} \right) \qquad (i = 1, 2, \ldots, n) ,$$

$$B_j = \left(\bigcup_{i=1}^{n} C_{ij} \right) \bigcup \left(\bigcup_{k=1}^{s(j)} G_{jk} \right) \qquad (j = 1, 2, \ldots, m) ,$$

ove $\{D_{ik} : 1 \leq i \leq n, 1 \leq k \leq r(i)\}$ e $\{G_{jk} : 1 \leq j \leq m, 1 \leq k \leq s(j)\}$ sono famiglie finite di insiemi disgiunti di $\mathcal{S}$. Perciò

$$E \Delta F = \left(\bigcup_{i=1}^{n} \bigcup_{k=1}^{r(i)} D_{ik} \right) \bigcup \left(\bigcup_{j=1}^{n} \bigcup_{k=1}^{s(j)} G_{jk} \right)$$

che è in $\mathcal{V}$; $\mathcal{V}$ è stabile anche rispetto alla differenza simmetrica ed è quindi un anello. □

Siano date due classi $\mathcal{C}$ e $\mathcal{D}$ di sottoinsiemi di Ω; se $\mu : \mathcal{C} \to \overline{\mathbb{R}}_+$ e $\mathcal{C} \subseteq \mathcal{D}$, si dice che la funzione $\nu : \mathcal{D} \to \overline{\mathbb{R}}_+$ *estende* μ se accade che $\nu(E) = \mu(E)$ per ogni insieme $E \in \mathcal{C}$.

Teorema 1.6.3. *Se $\mathcal{S}$ è un semi–anello e $\mu : \mathcal{S} \to \mathbb{R}_+$ è finitamente additiva, esiste una sola estensione finitamente additiva ν di μ all'anello $\mathcal{R}(\mathcal{S})$ generato da $\mathcal{S}$. Se μ è una misura anche ν è una misura.*

Dimostrazione. Per il Teorema 1.6.2, ogni insieme E di $\mathcal{R} = \mathcal{R}(\mathcal{S})$ è della forma $E = \cup_{i=1}^n A_i$ ove $\{A_i\}$ è una partizione di E in $\mathcal{S}$. Si definisca

$$\nu(E) := \sum_{i=1}^n \mu(A_i) . \tag{1.6.1}$$

La (1.6.1) è una buona definizione. Infatti, sia $E = \cup_{j=1}^m B_j$ un'altra decomposizione di E in insiemi disgunti di $\mathcal{S}$. Posto $C_{ij} := A_i \cap B_j$, gli insiemi C_{ij} appartengono a $\mathcal{S}$ e sono disgiunti; inoltre $A_i = \cup_{j=1}^m C_{ij}$ $(i = 1, 2, \dots, n)$ e $B_j = \cup_{i=1}^n C_{ij}$ $(j = 1, 2, \dots, m)$. Poiché μ è additiva in $\mathcal{S}$, si ha

$$\sum_{i=1}^n \mu(A_i) = \sum_{i=1}^n \sum_{j=1}^m \mu(C_{ij}) = \sum_{j=1}^m \mu(B_j) ,$$

sicché la (1.6.1) è una buona definizione.

Siano ora E_1 e E_2 due insiemi disgiunti di $\mathcal{R}$ con decomposizioni $E_1 = \cup_{i=1}^n A_i$ e $E_2 = \cup_{j=1}^m B_j$. Una possibile decomposizione di $E_1 \cup E_2$ in insiemi disgiunti di $\mathcal{S}$ è data da

$$E_1 \bigcup E_2 = \left(\bigcup_{i=1}^n A_i \right) \bigcup \left(\bigcup_{j=1}^m B_j \right) ,$$

sicché

$$\nu\left(E_1 \bigcup E_2\right) = \sum_{i=1}^n \mu(A_i) + \sum_{j=1}^m \nu(B_j) = \nu(E_1) + \nu(E_2) ,$$

vale a dire che ν è finitamente additiva.

Per dimostrare che ν è la sola estensione additiva di μ a $\mathcal{R}$, si supponga che ne esista un'altra $\widetilde{\nu}$; se, come sopra, E è in $\mathcal{R}$, si ha

$$\begin{aligned}\widetilde{\nu}(E) = \widetilde{\nu}\left(\bigcup_{i=1}^n A_i \right) &= \sum_{i=1}^n \widetilde{\nu}(A_i) \quad (\text{perché } \widetilde{\nu} \text{ è additiva}) \\ &= \sum_{i=1}^n \mu(A_i) \quad (\text{perché } \widetilde{\nu} \text{ estende } \mu) \\ &= \nu(E) ,\end{aligned}$$

sicché $\widetilde{\nu} = \nu$.

Si supponga ora che μ sia una misura su $\mathcal{S}$ e sia (E_n) una successione disgiunta di insiemi di $\mathcal{R}$ tale che appartenga a $\mathcal{R}$ la sua unione $E = \cup_{n \in \mathbb{N}} E_n$. Si introducano le unioni disgiunte di insiemi di $\mathcal{R}$

$$E = \bigcup_{r=1}^k A_r, \qquad E_n = \bigcup_{i=1}^{k(n)} B_{ni} \quad (n \in \mathbb{N}) .$$

Posto $C_{rni} := A_r \cap B_{ni}$ $(n \in \mathbb{N}; r = 1, 2, \dots, k; i = 1, 2, \dots, k(n))$, la famiglia $\{C_{rni}\}$ è composta di insiemi disgiunti di $\mathcal{S}$; inoltre,

$$A_r = \bigcup_{n\in\mathbb{N}} \bigcup_{i=1}^{k(n)} C_{rni} \qquad \text{e} \qquad B_{ni} = \bigcup_{r=1}^{k} C_{rni}$$

sono decomposizioni in insiemi disgiunti di $\mathcal{S}$. Poiché μ è σ–additiva, si ha

$$\mu(A_r) = \sum_{n\in\mathbb{N}} \sum_{i=1}^{k(n)} \mu(C_{rni}) \qquad \text{e} \qquad \mu(B_{ni}) = \sum_{r=1}^{k} \mu(C_{rni})\,.$$

L'ordine di somma nelle serie a termini positivi può essere cambiato a piacimento, sicché

$$\begin{aligned}\nu(E) &= \sum_{r=1}^{k} \mu(A_r) = \sum_{r=1}^{k} \sum_{n\in\mathbb{N}} \sum_{i=1}^{k(n)} \mu(C_{rni}) \\ &= \sum_{n\in\mathbb{N}} \sum_{i=1}^{k(n)} \sum_{r=1}^{k} \mu(C_{rni}) = \sum_{n\in\mathbb{N}} \sum_{i=1}^{k(n)} \mu(B_{ni}) = \sum_{n\in\mathbb{N}} \nu(E_n)\,.\end{aligned}$$

Perciò, anche l'estensione ν di μ è σ–additiva. □

Prima di poter enunciare il fondamentale teorema di estensione di Carathéodory occorrono alcune premesse.

Definizione 1.6.4. Dato un insieme non vuoto Ω, ogni funzione $\mu^* : \mathcal{P}(\Omega) \to \mathbb{R}_+$ tale che:

(a) $\mu^*(\emptyset) = 0$;

(b) μ^* sia isotona: se $E \subseteq F$ allora $\mu^*(E) \leq \mu^*(F)$;

(c) μ^* sia numerabilmente subadditiva:

$$E \subseteq \bigcup_{n\in\mathbb{N}} E_n \implies \mu^*(E) \leq \sum_{n\in\mathbb{N}} \mu^*(E_n)$$

si dice *misura esterna* su Ω. ◇

Teorema 1.6.4. *Sia μ^* una misura esterna su Ω e sia $\mathcal{M}$ la famiglia dei sottoinsiemi di Ω definita da*

$$\mathcal{M} := \left\{ E \subseteq \Omega : \forall\, A \subseteq \Omega \quad \mu^*(A) = \mu^*\left(A \bigcap E\right) + \mu^*\left(A \setminus E\right) \right\}\,;$$

allora

(a) *$\mathcal{M}$ è una tribú;*

(b) *la restrizione μ di μ^* a $\mathcal{M}$ è una misura.*

Dimostrazione. (a) Dimostriamo dapprima che $\mathcal{M}$ è stabile rispetto all'unione finita; a tal fine, basta far vedere che anche $E_1 \cup E_2$ appartiene a $\mathcal{M}$, se vi appartengono E_1 e E_2. Poiché E_1 è in $\mathcal{M}$, si ha per ogni $A \subseteq \Omega$,

$$\mu^*(A) = \mu^*\left(A \bigcap E_1\right) + \mu^*\left(A \setminus E_1\right) ; \tag{1.6.2}$$

ma anche E_2 è in $\mathcal{M}$ e dunque

$$\begin{aligned} \mu^*\left(A \setminus E_1\right) &= \mu^*\left[(A \setminus E_1) \bigcap E_2\right] + \mu^*\left[(A \setminus E_1) \setminus E_2\right] \\ &= \mu^*\left(A \bigcap E_1^c \bigcap E_2\right) + \mu^*\left[A \setminus \left(E_1 \bigcup E_2\right)\right] . \end{aligned} \tag{1.6.3}$$

Ora $(A \cap E_1^c \cap E_2) \cup (A \cap E_1) = A \cap (E_1 \cup E_2)$; perciò, sostituendo la (1.6.3) nella (1.6.2) si ottiene, ricorrendo alla subadditività di μ^*,

$$\begin{aligned} \mu^*(A) &= \mu^*\left(A \bigcap E_1\right) + \mu^*\left(A \bigcap E_1^c \bigcap E_2\right) + \mu^*\left[A \setminus \left(E_1 \bigcup E_2\right)\right] \\ &\geq \mu^*\left[A \bigcap (E_1 \bigcup E_2)\right] + \mu^*\left[A \setminus \left(E_1 \bigcup E_2\right)\right] . \end{aligned}$$

Poiché la diseguaglianza nell'altro verso vale in virtú della subadditività di μ^*, si ha che $E_1 \cup E_2$ appartiene a $\mathcal{M}$.

Si vede subito che la stessa relazione che garantisce che E appartiene a $\mathcal{M}$ assicura anche che vi appartenga il suo complementare E^c. È, poi, ovvio che appartiene a $\mathcal{M}$ l'insieme vuoto $\emptyset$. Dunque $\mathcal{M}$ è un'algebra. Per mostrare che essa è una tribú occorre far vedere che è in $\mathcal{M}$ l'unione di ogni successione (E_n) di insiemi di $\mathcal{M}$. Basta, anzi, supporre che gli insiemi della successione siano disgiunti, perché avendo già mostrato che $\mathcal{M}$ è un'algebra, si può sostituire ogni unione numerabile con un'unione numerabile disgiunta. Pertanto si supporrà che sia $E_j \cap E_k = \emptyset$ se $j \neq k$. Posto $F_n := \cup_{j=1}^n E_j$, si ha, per ogni $n \in \mathbb{N}$ e per ogni $A \subseteq \Omega$,

$$\mu^*\left(A \bigcap F_n\right) = \sum_{j=1}^n \mu^*\left(A \bigcap E_j\right) . \tag{1.6.4}$$

Ciò è ovviamente vero per $n = 1$. Si supponga ora che la (1.6.4) valga per un naturale n. Poiché F_n appartiene a $\mathcal{M}$, prendendo $A \cap F_{n+1}$ in luogo di A nella definizione di $\mathcal{M}$, si ha, per $n+1$,

$$\mu^*\left(A \bigcap F_{n+1}\right) = \mu^*\left(A \bigcap F_n\right) + \mu^*\left(A \bigcap E_{n+1}\right) = \sum_{j=1}^{n+1} \mu^*\left(A \bigcap E_j\right) ,$$

sicché la (1.6.4) vale per ogni $n \in \mathbb{N}$. Se $E = \cup_{n\in\mathbb{N}} E_n$, scende dall'isotonia di μ^* che

$$\mu^*\left(A \bigcap E\right) \geq \mu^*\left(A \bigcap F_n\right) = \sum_{j=1}^n \mu^*\left(A \bigcap E_j\right) ,$$

onde

$$\mu^*\left(A \bigcap E\right) \geq \sum_{n\in\mathbb{N}} \mu^*\left(A \bigcap E_n\right) .$$

La subadditività di μ^* dà ora

$$\mu^*\left(A \bigcap E\right) = \sum_{n\in\mathbb{N}} \mu^*\left(A \bigcap E_n\right) . \tag{1.6.5}$$

Cosí, per ogni sottoinsieme A di Ω e per ogni $n \in \mathbb{N}$, si ha

$$\mu^*(A) = \mu^*\left(A \bigcap F_n\right) + \mu^*\left(A \setminus F_n\right) \geq \sum_{j=1}^{n} \mu^*\left(A \bigcap E_j\right) + \mu^*\left(A \setminus E\right) ,$$

e, di qui,

$$\mu^*(A) \geq \mu^*\left(A \bigcap E\right) + \mu^*\left(A \setminus E\right) ,$$

sicché

$$\mu^*(A) = \mu^*(\left(A \bigcap E\right) + \mu^*(A \setminus E) ;$$

perciò E appartiene a $\mathcal{M}$.

Ponendo $A = \Omega$ nella (1.6.5) si vede che μ^* è σ–additiva in $\mathcal{M}$. □

Subito dopo aver introdotto una nuova definizione, saremo in grado di enunciare il teorema d'estensione di Carathéodory.

Definizione 1.6.5. Una misura μ su $(\Omega, \mathcal{F})$ si dice *σ–finita* se si può esprimere Ω come unione numerabile di insiemi misurabili ciascuno dei quali abbia misura finita, $\Omega = \cup_{n\in\mathbb{N}} E_n$, ove per ogni $n \in \mathbb{N}$, è $E_n \in \mathcal{F}$ e $\mu(E_n) < +\infty$. ◇

Teorema 1.6.5. (Carathéodory). *Sia $\mathcal{R}$ un anello di sottoinsiemi di Ω. Si supponga che Ω si possa esprimere come unione di insiemi di $\mathcal{R}$, $\Omega = \cup_{n\in\mathbb{N}} E_n$, con $E_n \in \mathcal{R}$ per ogni $n \in \mathbb{N}$ e sia $\mu : \mathcal{R} \to \overline{\mathbb{R}}_+$ una misura su $\mathcal{R}$. Esiste allora una misura ν sulla tribú $\mathcal{F}(\mathcal{R})$ generata da $\mathcal{R}$ che estende μ. Se μ è σ–finita, l'estensione ν è unica ed è anch'essa σ–finita.*

Dimostrazione. Ogni sottoinsieme E di Ω può essere ricoperto dall'unione di una successione di insiemi di $\mathcal{R}$; si può quindi definire

$$\mu^*(E) := \inf \left\{ \sum_{n\in\mathbb{N}} \mu(F_n) \right\} ,$$

eseguendo l'estremo inferiore su tutte le successioni (F_n) di insiemi di $\mathcal{R}$ tali che $E \subseteq \cup_{n\in\mathbb{N}} F_n$. È evidente che la funzione μ^* cosídefinita su $\mathcal{P}(\Omega)$ è positiva, isotona, e tale che $\mu^*(\emptyset) = 0$. Sia ora $E \subseteq \cup_{n\in\mathbb{N}} E_n$. Se esiste almeno un indice n per il quale $\mu^*(E_n) = +\infty$, si ha banalmente,

$$\mu^*(E) \leq \sum_{n\in\mathbb{N}} \mu^*(E_n) .$$

Se, invece, risulta $\mu^*(E_n) < +\infty$ per ogni $n \in \mathbb{N}$, si scelgano in $\mathcal{R}$, per ogni $\varepsilon > 0$, insiemi F_{nk} con $k \in \mathbb{N}$, in modo che sia, per ogni $n \in \mathbb{N}$,

$$E_n \subseteq \bigcup_{k\in\mathbb{N}} F_{nk} \qquad \text{e} \qquad \sum_{k\in\mathbb{N}} \mu^*(F_{nk}) < \mu^*(E_n) + \frac{\varepsilon}{2^n} .$$

Segue di qui che μ^* è una misura esterna. Si definisca ora $\mathcal{M}$ come nel Teorema 1.6.4. Dimostriamo, in primo luogo, che $\mathcal{M}$ contiene $\mathcal{R}$. Se $E \in \mathcal{R}$ e $\mu^*(A) < +\infty$, si scelga una successione (E_n) di insiemi di $\mathcal{R}$, tale che A sia contenuto nell'unione $\cup_{n\in\mathbb{N}} E_n$ e che

$$\mu^*(A) + \varepsilon \geq \sum_{n\in\mathbb{N}} \mu(E_n)\,;$$

ora

$$\begin{aligned}\mu^*(A) + \varepsilon &\geq \sum_{n\in\mathbb{N}} \left[\mu\left(E_n \bigcap E\right) + \mu\left(E_n \setminus E\right)\right] \\ &\geq \mu^*\left(A \bigcap E\right) + \mu^*\left(A \setminus E\right)\,,\end{aligned}$$

per la subadditività di μ^*. Se $\mu^*(A) = +\infty$, l'ultima diseguglianza è ovvia. Siccome ε è arbitrario, E appartiene a $\mathcal{M}$.

Per il Teorema 1.6.4, $\mathcal{M}$ è una tribú, sicché essa include necessariamente la tribú $\mathcal{F}(\mathcal{R})$ generata da $\mathcal{R}$. Poiché la restrizione di μ^* a $\mathcal{M}$ è una misura, l'ulteriore restrizione ν di μ^* a $\mathcal{F}(\mathcal{R})$ è ancora una misura. Per tutti gli insiemi E di $\mathcal{R}$ si ha $\nu(E) = \mu^*(E) = \mu(E)$, sicché ν è effettivamente un'estensione di μ da $\mathcal{R}$ a $\mathcal{F}(\mathcal{R})$.

Si supponga ora che μ (e quindi μ^*) sia σ–finita su $\mathcal{R}$; allora si può scrivere Ω nella forma $\Omega = \cup_{n\in\mathbb{N}} E_n$ con (E_n) successione crescente di insiemi di $\mathcal{R}$ e $\mu(E_n) < \infty$ per ogni $n \in \mathbb{N}$. Per dimostrare l'unicità dell'estensione, sia λ una qualsiasi estensione di μ a $\mathcal{F}(\mathcal{R})$ e si definisca, per $n \in \mathbb{N}$, $\mathcal{E}_n := \{E \in \mathcal{F}(\mathcal{R}) : E \subseteq E_n, \lambda(E) = \mu(E)\}$. è immediato verificare che $\mathcal{E}_n$ è un'algebra; di piú, poiché le misure finite passano al limite lungo le successioni monotone, $\mathcal{E}_n$ è una classe monotona, e quindi, in virtú del Teorema 1.6.1, $\mathcal{E}_n \supset \mathcal{F}(\mathcal{R}_n)$ ove $\mathcal{R}_n := \mathcal{R} \cap E_n = \{E \in \mathcal{R} : E \subseteq E_n\}$. Dunque, per ogni $n \in \mathbb{N}$, l'estensione di μ da $\mathcal{R}_n$ a $\mathcal{F}(\mathcal{R}_n)$ è unica. Ma per ogni $E \in \mathcal{F}(\mathcal{R})$ si ha $(E \cap E_n) \uparrow E$, sicché applicando il Teorema 1.5.1 si ha l'unicità dell'estensione. □

Nella dimostrazione dell'unicità nel teorema di Carathéodory avremmo potuto usare anche il successivo Teorema 1.6.7 che introduciamo perché vi faremo ricorso, sia pure tacitamente, nel seguito.

Definizione 1.6.6. Una famiglia $\mathcal{L}$ di sottoinsiemi di Ω si dice essere un λ–*sistema* se sono verificate le condizioni:

(a) Ω appartiene a $\mathcal{L}$;

(b) $A, B \in \mathcal{L},\ A \subseteq B \implies B \setminus A \in \mathcal{L}$;

(c) $A_n \in \mathcal{L}\ (n \in \mathbb{N}),\ A_n \uparrow A \implies A \in \mathcal{L}$. ◇

Teorema 1.6.6. *Se $\mathcal{P}$ è un π–sistema di sottoinsiemi di Ω e se $\mathcal{L}$ è un λ–sistema che contiene $\mathcal{P}$, allora $\mathcal{L}$ contiene anche la tribú $\mathcal{F}(\mathcal{P})$ generata da $\mathcal{P}$.*

Dimostrazione. Si indichi con $\mathcal{L}(\mathcal{P})$ il piú piccolo λ–sistema che contiene $\mathcal{P}$. Basta dimostrare che $\mathcal{L}(\mathcal{P})$ è una tribú.

Si osservi che, se un λ–sistema $\mathcal{L}$ è stabile rispetto alle intersezioni finite, è una tribú. Infatti, $\mathcal{L}$ è stabile rispetto alla complementazione perché, se A appartiene a $\mathcal{L}$, si ha $A^c = \Omega \setminus A \in \mathcal{L}$,; inoltre $\mathcal{L}$ è stabile rispetto alle unioni finite perché

$A \cup B = (A^c \cap B^c)^c$ ed infine è stabile rispetto alle unioni numerabili perché $\cup_{j=1}^n A_j \uparrow \cup_{n\in\mathbb{N}} A_n$. Basta, perciò, dimostrare che $\mathcal{L}(\mathcal{P})$ è stabile rispetto alle intersezioni finite.

Si prenda A in $\mathcal{P}$ e si definisca

$$\mathcal{G}(A) := \left\{B \subseteq \Omega : A \bigcap B \in \mathcal{L}(\mathcal{P})\right\} .$$

Allora $\mathcal{G}(A)$ è un λ–sistema. Infatti Ω appartiene a $\mathcal{G}(A)$ perché $A \cap \emptyset = \emptyset$ è in $\mathcal{P}$ e dunque anche in $\mathcal{L}(\mathcal{P})$. Se B e C appartengono a $\mathcal{G}(A)$ e $C \subseteq B$, si ha $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$ che appartiene a $\mathcal{L}(\mathcal{P})$; perciò $B \setminus C$ appartiene a $\mathcal{G}(A)$. Infine, se (B_n) è una successione di insiemi di $\mathcal{G}(A)$ che tende crescendo a B, si ha $B_n \cap A \uparrow B \cap A$ e dunque anche B appartiene a $\mathcal{G}(A)$. Poiché $\mathcal{P}$ è contenuto in $\mathcal{G}(A)$, si ha $\mathcal{L}(\mathcal{P}) \subseteq \mathcal{G}(A)$. Perciò, se A appartiene a $\mathcal{P}$ e B a $\mathcal{L}(\mathcal{P})$, si ha che $A \cap B$ appartiene a $\mathcal{L}(\mathcal{P})$ e, di qui, che $\mathcal{G}(B)$ contiene $\mathcal{P}$ se B è in $\mathcal{L}(\mathcal{P})$; ma $\mathcal{G}(B)$ è un λ–sistema e dunque $\mathcal{G}(B)$ contiene $\mathcal{L}(\mathcal{P})$. Questo implica che, se B e C appartengono a $\mathcal{L}(\mathcal{P})$ vi appartiene anche $B \cap C$. □

Siamo ora in grado di dimostrare il seguente

Teorema 1.6.7. *Dato uno spazio misurabile $(\Omega, \mathcal{F})$, se due misure μ_1 e μ_2 coincidono sugli insiemi di un π–sistema $\mathcal{P}$ di sottoinsiemi di $\mathcal{F}$, $\mu_1(A) = \mu_2(A)$ per ogni $A \in \mathcal{P}$, allora coincidono sulla tribú $\mathcal{F}(\mathcal{P})$ generata da $\mathcal{P}$, $\mu_1(A) = \mu_2(A)$ per ogni $A \in \mathcal{F}(\mathcal{P})$.*

Dimostrazione. Sia A un insieme di $\mathcal{P}$ tale che $\mu_1(A) = \mu_2(A) < +\infty$. Si ponga

$$\mathcal{L} := \left\{B \in \mathcal{F} : \mu_1\left(A \bigcap B\right) = \mu_2\left(A \bigcap B\right)\right\} .$$

Banalmente, si ha che Ω appartiene a $\mathcal{L}$. Siano ora B e C due insiemi di $\mathcal{P}$ con $B \supset C$; allora

$$\begin{aligned}\mu_1(A \cap (B \setminus C)) &= \mu_1\left(A \bigcap B\right) - \mu_1\left(A \bigcap C\right) \\ &= \mu_2\left(A \bigcap B\right) - \mu_2\left(A \bigcap C\right) = \mu_2(A \cap (B \setminus C)) ,\end{aligned}$$

sicché $B \setminus C$ appartiene a $\mathcal{L}$. Infine sia (B_n) una successione crescente di insiemi di $\mathcal{L}$; se $B = \cup_{n\in\mathbb{N}} B_n$, si ha

$$\begin{aligned}\mu_1\left(A \bigcap B\right) &= \lim_{n\to+\infty} \mu_1\left(A \bigcap B_n\right) = \lim_{n\to+\infty} \mu_2\left(A \bigcap B_n\right) \\ &= \mu_2\left(A \bigcap B\right) .\end{aligned}$$

Dunque $\mathcal{L}$ è un λ–sistema. Allora il teorema precedente assicura che $\mathcal{L}$ include la tribú $\mathcal{F}(\mathcal{P})$ generata da $\mathcal{P}$. Perciò se $A \in \mathcal{P}$ è tale che $\mu_1(A) = \mu_2(A) < +\infty$ e se B è in $\mathcal{F}(\mathcal{P})$, si ha $\mu_1(A \cap B) = \mu_2(A \cap B)$. Consideriamo una successione crescente (A_n) di insiemi di $\mathcal{P}$ tali che $\cup_{n\in\mathbb{N}} A_n = \Omega$ e che $\mu_1(A_n) = \mu_2(A_n) < +\infty$ per ogni $n \in \mathbb{N}$. Allora, per ogni $B \in \mathcal{F}(\mathcal{P})$, si ha

$$\mu_1(A \bigcap B) = \lim_{n\to+\infty} \mu_1(A \bigcap B_n) = \lim_{n\to+\infty} \mu_2(A \bigcap B_n) = \mu_2(A \bigcap B) ,$$

che fornisce l'asserto. □

Le π–classi ed i λ–sistemi si usano anche per lo studio dell'indipendenza.

Teorema 1.6.8. *Nello spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$, *si considerino le* π*–classi* $\mathcal{C}_1$, $\dots$, $\mathcal{C}_n$ *di insiemi di* $\mathcal{F}$*; se* $\mathcal{C}_1$, $\dots$, $\mathcal{C}_n$ *sono indipendenti, vale a dire se*

$$\mathbb{P}\left(\bigcap_{j=1}^{n} A_j\right) = \Pi_{j=1}^{n} \mathbb{P}(A_j)\,, \tag{1.6.6}$$

per ogni scelta di A_j *in* $\mathcal{C}_j$ $(j = 1, \dots, n)$*, allora sono indipendenti anche le tribú* $\mathcal{F}(\mathcal{C}_1)$, $\dots$, $\mathcal{F}(\mathcal{C}_n)$ *generate da* $\mathcal{C}_1$, $\dots$, $\mathcal{C}_n$ *rispettivamente.*

Dimostrazione. Se $\mathcal{C}_j$ non contiene già Ω si consideri la famiglia $\mathcal{C}'_j$ ottenuta aggiungendo Ω a $\mathcal{C}_j$; ogni nuova famiglia $\mathcal{C}'_j$ è anch'essa una π–classe e la (1.6.6) vale per ogni scelta di A_j in $\mathcal{C}'_j$ $(j = 1, 2, \dots, n)$. Si fissino ora $A_2, A_3, \dots, A_n$ in $\mathcal{C}'_2, \mathcal{C}'_3, \dots, \mathcal{C}'_n$ rispettivamente e sia $\mathcal{L}_1$ la famiglia degli insiemi di $\mathcal{F}$ per i quali vale la (1.6.6). Si controlla immediatamente che $\mathcal{L}_1$ è un λ–sistema; inoltre, esso contiene $\mathcal{C}'_1$, sicché, per il Teorema 1.6.6 contiene anche la tribú $\mathcal{F}(\mathcal{C}'_1) = \mathcal{F}(\mathcal{C}_1)$ generata da $\mathcal{C}'_1$ o da $\mathcal{C}_1$. Pertanto, la (1.6.6) vale se A_1 è in $\mathcal{F}(\mathcal{C}_1)$ e A_j in $\mathcal{C}_j$ per $j = 2, 3, \dots, n$. Si ripeta lo stesso ragionamento per gli insiemi di $\mathcal{F}(\mathcal{C}_1), \mathcal{C}_2, \dots, \mathcal{C}_n$ per concludere che la (1.6.6) vale quando gli insiemi $A_1, A_2, \dots, A_n$ siano scelti in $\mathcal{F}(\mathcal{C}_1), \mathcal{F}(\mathcal{C}_2), \mathcal{C}_3, \dots, \mathcal{C}_n$ rispettivamente; si proceda poi nello stesso modo per i rimanenti indici. $\square$

1.7 Le misure di Borel–Stieltjes

Un esempio importante di spazio misurabile è dato da $(\mathbb{R}, \mathcal{B})$. Si dirà che una misura μ su $\mathcal{B}$ è *localmente finita* se accade che, per ogni intervallo limitato I, sia $\mu(I) < +\infty$. Ovviamente, ogni misura finita su $(\mathbb{R}, \mathcal{B})$ è anche localmente finita.

Teorema 1.7.1. *Se* $\mu : \mathcal{B} \to \overline{\mathbb{R}}_+$ *è una misura localmente finita, allora la funzione* $F : \mathbb{R} \to \mathbb{R}$ *definita, a meno di una costante, da*

$$F(b) - F(a) := \mu(]a, b]) \quad (a < b)$$

è crescente e continua a destra.

Dimostrazione. Si osservi che

$$F(x) = \begin{cases} F(0) + \mu(]0, x])\,, & \text{se } x \geq 0\,, \\ F(0) - \mu(]x, 0])\,, & \text{se } x < 0; \end{cases}$$

basta perciò fissare $F(0)$ arbitrariamente. Se $x' < x''$, si ha $F(x'') - F(x') = \mu(]x', x'']) \geq 0$, sicché F è crescente. Sia ora (x_n) un'arbitraria successione reale che decresce a x $\quad (x_n \downarrow x)$. Allora $F(x_n) - F(x) = \mu(]x, x_n]) \xrightarrow[n \to +\infty]{} 0$. $\square$

In effetti, in probabilità, si segue una via diversa per definire univocamente la funzione F che sarà detta *funzione di ripartizione*. Anziché fissare il valore $F(0)$ di F nell'origine, si pone eguale a zero il limite di F a $-\infty$.

Vale il seguente importante reciproco dell'ultimo teorema.

Teorema 1.7.2. *Se $F : \mathbb{R} \to \mathbb{R}$ è crescente e continua a destra, esiste un'unica misura $\mu : \mathcal{B} \to \mathbb{R}_+$ tale che $\mu(]a,b]) = F(b) - F(a)$ per ogni intervallo $]a,b]$ con $a < b$. Tale misura è localmente finita e si dice* misura di Borel–Stieltjes *associata a F; la si indica spesso con μ_F.*

Dimostrazione. Sia $\mathcal{R}$ l'anello delle unioni finite di intervalli disgiunti, aperti a sinistra e chiusi a destra, $\mathcal{R} := \cup_{n\in\mathbb{N}}\mathcal{R}_n$ ove

$$\mathcal{R}_n := \left\{ A : A = \bigcup_{i=1}^{n} \left]a_i, b_i\right] , a_1 \le b_1 \le a_2 \le b_2 \le \dots \le a_n \le b_n \right\} .$$

In $\mathcal{R}$ si definisca $\nu : \mathcal{R} \to \mathbb{R}_+$ mediante

$$\nu(A) := \sum_{i=1}^{n} \{F(b_i) - F(a_i)\} ;$$

ν, cosí definita, è finitamente additiva; di piú, essa è una misura su $\mathcal{R}$ come si dimostrerà di seguito.

Se A è in $\mathcal{R}$, esiste, per ogni $\varepsilon > 0$, un insieme B di $\mathcal{R}$ tale che la sua chiusura $\overline{B}$ sia contenuta in A, vale a dire $B \subseteq \overline{B} \subseteq A$, e che $\nu(A) - \nu(B) < \varepsilon$. Infatti, se $A =]a,b]$, B sarà del tipo $]a+h,b]$ con $h > 0$ sufficientemente piccolo (per la continuità a destra di F si ha

$$\lim_{h\downarrow 0} \nu(]a+h,b]) = \lim_{h\downarrow 0}[F(b) - F(a+h)] = F(b) - F(a) = \nu(A) .$$

Si procede in maniera analoga se A è un'unione finita di intervalli. Sia ora (A_n) una successione decrescente di insiemi di $\mathcal{R}$ con $\cap_{n\in\mathbb{N}} A_n = \emptyset$. Fissato $\varepsilon > 0$, si scelga, per ogni $n \in \mathbb{N}$, $B_n \in \mathcal{R}$ in modo che sia $B_n \subseteq \overline{B}_n \subseteq A_n$ e che $\nu(A_n) - \nu(\overline{B}_n) < \varepsilon\, 2^{-n}$. Allora, si ha

$$\bigcap_{n\in\mathbb{N}} \overline{B}_n = \emptyset .$$

Esiste $r \in \mathbb{N}$ tale che

$$\bigcap_{i=1}^{r} \overline{B}_i = \emptyset .$$

Per rendersene conto, si osservi che gli insiemi $\overline{\mathbb{R}} \setminus \overline{B}_n$ formano un ricoprimento aperto del compatto $\overline{\mathbb{R}} = [-\infty, +\infty]$; vi è, pertanto, un ricoprimento aperto finito di $\overline{\mathbb{R}}$:

$$\overline{\mathbb{R}} = \bigcup_{i=1}^{r} (\overline{\mathbb{R}} \setminus \overline{B}_i) = \bigcup_{i=1}^{r} \left(\overline{\mathbb{R}} \bigcap \overline{B}_i^c\right) = \overline{\mathbb{R}} \bigcap \left(\bigcup_{i=1}^{r} \overline{B}_i^c\right) ,$$

sicché $\cap_{i=1}^{r} \overline{B}_i = \emptyset$ e dunque, a maggior ragione, $\cap_{i=1}^{r} B_i = \emptyset$ e

$$A_r = A_r \bigcap \left(\bigcap_{i=1}^{r} B_i\right)^c = \bigcup_{i=1}^{r} (A_r \setminus B_i) \subseteq \bigcup_{i=1}^{r} (A_i \setminus B_i) ,$$

onde

$$\begin{aligned} \nu(A_r) &\le \nu \left[\bigcup_{i=1}^{r} (A_i \setminus B_i)\right] \\ &\le \sum_{i=1}^{r} \nu\,(A_i \setminus B_i) < \varepsilon \sum_{i=1}^{r} 2^{-i} = \varepsilon(1 - 2^{-r}) < \varepsilon . \end{aligned}$$

Perciò $\lim_{n\to+\infty} \nu(A_n) = 0$ e, in virtú del Teorema 1.5.1, ν è σ–additiva. L'anello $\mathcal{R}$ soddisfà alle ipotesi del teorema di Carathéodory perché $\mathbb{R}$ è l'unione numerabile degli intervalli $]-n, n]$ e perché ν è σ–finita, in quanto $\nu(]-n, n]) = F(n)-F(-n) < +\infty$. Esiste perciò un'unica misura μ su $\mathcal{B}$ tale che $\mu(A) = \nu(A)$ per ogni insieme $A \in \mathcal{R}$. In particolare, $\mu(]a, b]) = \nu(]a, b]) = F(b) - F(a)$. □

La funzione identità soddisfà alle ipotesi del Teorema 1.7.2: $F(x) = x$ per ogni $x \in \mathbb{R}$; la misura associata si dice di Borel–Lebesgue; essa fa corrispondere ad ogni intervallo $[a, b]$ la sua lunghezza $b - a$. Questa costruzione della misura di Lebesgue non è l'unica possibile.

1.8 Funzioni semplici e funzioni misurabili

Definizione 1.8.1. Dati due spazî misurabili $(\Omega, \mathcal{F})$ e $(\Omega', \mathcal{F}')$ si dice che la funzione $f : \Omega \to \Omega'$ è *misurabile* (rispetto alle tribú $\mathcal{F}$ e $\mathcal{F}'$) se, per ogni insieme B di $\mathcal{F}'$, l'immagine inversa di B mediante f appartiene alla tribú $\mathcal{F}$, cioè

$$\forall\, B \in \mathcal{F}' \quad f^{-1}(B) \in \mathcal{F}\,.$$

Spesso si scrive $f : (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$ per indicare una funzione misurabile nel senso di questa definizione. Se poi f è una funzione da Ω in $\mathbb{R}$ o in $\overline{\mathbb{R}}$ spesso si sottintende la tribú di Borel $\mathcal{B}$ in $\mathbb{R}$ o in $\overline{\mathbb{R}}$. In questo caso si scrive $f \in \mathcal{L}^0$, indicando lo spazio delle funzioni misurabili da Ω in $\mathbb{R}$ o in $\overline{\mathbb{R}}$ con $\mathcal{L}^0$ o, se si vuole indicare esplicitamente la tribú considerata, con $\mathcal{L}^0(\Omega, \mathcal{F})$, oppure ancora con $\mathcal{L}^0(\mathcal{F})$. ◇

Il seguente teorema è di dimostrazione immediata

Teorema 1.8.1. *Siano $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{F}')$ e $(\Omega'', \mathcal{F}'')$ tre spazî misurabili e siano $f : \Omega \to \Omega'$ e $g : \Omega' \to \Omega''$ due funzioni misurabili, rispettivamente rispetto alle coppie di tribú $\mathcal{F}$, $\mathcal{F}'$ e $\mathcal{F}'$, $\mathcal{F}''$. Allora è misurabile rispetto a $\mathcal{F}$ e a $\mathcal{F}''$ la funzione composta $g \circ f : \Omega \to \Omega''$.*

Per verificare se una funzione sia misurabile si ricorre spesso al seguente criterio.

Teorema 1.8.2. *Siano $(\Omega_1, \mathcal{F}_1)$ e $(\Omega_2, \mathcal{F}_2)$ spazî misurabili e sia $\mathcal{A} \subseteq \mathcal{F}_2$ una famiglia di sottoinsiemi di Ω_2 che genera la tribú $\mathcal{F}_2$, cioè $\mathcal{F}(\mathcal{A}) = \mathcal{F}_2$. Sono allora equivalenti, per una funzione $f : \Omega_1 \to \Omega_2$, le proprietà:*

(a) *f è misurabile;*

(b) *$\forall\, A \in \mathcal{A} \quad f^{-1}(A) \in \mathcal{F}_1$.*

Dimostrazione. L'implicazione (a) $\Longrightarrow$ (b) è banale.

(b) $\Longrightarrow$ (a) Sia $\mathcal{G}$ la famiglia dei sottoinsiemi di Ω_2 la cui immagine inversa appartiene a $\mathcal{F}_1$: $\mathcal{G} := \{A \subseteq \Omega_2 : f^{-1}(A) \in \mathcal{F}_1\}$. Per ipotesi $\mathcal{A} \subseteq \mathcal{G}$. $\mathcal{G}$ è una tribú: $\Omega_2 \in \mathcal{G}$ perché $f^{-1}(\Omega_2) = \Omega_1 \in \mathcal{F}$; se $A \in \mathcal{G}$, è $f^{-1}(A^c) = [f^{-1}(A)]^c \in \mathcal{F}$, cioè $A^c \in \mathcal{G}$. Infine, se A_n appartiene a $\mathcal{G}$ per ogni $n \in \mathbb{N}$ si ha $f^{-1}(\cup_{n\in\mathbb{N}} A_n) = \cup_{n\in\mathbb{N}} f^{-1}(A_n) \in \mathcal{F}_1$, cioè $\cup_{n\in\mathbb{N}} A_n \in \mathcal{G}$. Ma allora, dalla definizione di tribú generata, scende che $\mathcal{G} \supset \mathcal{F}_2$, sicché f risulta misurabile. □

Corollario 1.8.1. *Sia $(\Omega, \mathcal{F})$ uno spazio misurabile; affinché la funzione $f : \Omega \to \overline{\mathbb{R}}$ sia misurabile, rispetto a $\mathcal{F}$ e a $\mathcal{B}(\overline{\mathbb{R}})$, ognuna delle seguenti condizioni equivalenti è necessaria e sufficiente:*

(a) $\{f \leq c\} := f^{-1}([-\infty, c]) = \{\omega \in \Omega : f(\omega) \leq c\}$ *appartiene a $\mathcal{F}$ per ogni $c \in \mathbb{R}$;*

(b) $\{f > c\} := f^{-1}(]c, +\infty]) = \{\omega \in \Omega : f(\omega) > c\}$ *appartiene a $\mathcal{F}$ per ogni $c \in \mathbb{R}$;*

(c) $\{f \geq c\} := f^{-1}([c, +\infty]) = \{\omega \in \Omega : f(\omega) \geq c\}$ *appartiene a $\mathcal{F}$ per ogni $c \in \mathbb{R}$;*

(d) $\{f < c\} := f^{-1}([-\infty, c[) = \{\omega \in \Omega : f(\omega) < c\}$ *appartiene a $\mathcal{F}$ per ogni $c \in \mathbb{R}$.*

Dimostrazione. Ognuna delle famiglie indicate genera $\mathcal{B}(\overline{\mathbb{R}})$. □

Definizione 1.8.2. Una funzione $f : \mathbb{R} \to \mathbb{R}$ (oppure da $\overline{\mathbb{R}}$ in $\overline{\mathbb{R}}$) che sia misurabile rispetto alle tribú di Borel si dice *boreliana*. ◇

Si osservi la notazione usata per la prima volta nell'enunciato del Corollario 1.8.1, $\{f \leq c\} := f^{-1}([-\infty, c]) = \{\omega \in \Omega : f(\omega) \leq c\}$ e simili; nel seguito essa sarà usata sistematicamente. Cosí, ad esempio, si scriverà $\{a \leq f \leq b\}$ o $\{f \in [a, b]\}$ in luogo delle notazioni, equivalenti, ma piú pesanti, $f^{-1}([a, b])$ oppure

$$\{\omega \in \Omega : a \leq f(\omega) \leq b\}.$$

Un'ampia classe di funzioni $f : \mathbb{R} \to \mathbb{R}$ misurabili rispetto alla tribú di Borel $\mathcal{B}$ è data dalle funzioni continue (si vedano gli esercizî). Una funzione indicatrice $\mathbf{1}_A$ è misurabile se, e solo se, l'insieme A appartiene a $\mathcal{F}$. La classe delle funzioni misurabili è stabile rispetto a numerose operazioni eseguite in numero finito o numerabile, come dimostrano i due teoremi che seguono.

Teorema 1.8.3. *Siano $f, g : \Omega \to \overline{\mathbb{R}}$ funzioni misurabili e sia α un numero reale. Sono allora misurabili, se sono definite, anche le funzioni $\alpha \cdot f$, $f + g$, f^2, $f^+ := f \vee 0$, $f^- := -(f \wedge 0)$, $|f|$, $f \cdot g$, $f \vee g$, $f \wedge g$, f^n per ogni $n \in \mathbb{N}$. Inoltre, se f è misurabile e positiva sono misurabili anche le funzioni $f^{1/n}$, per ogni $n \in \mathbb{N}$, e f^r per ogni numero razionale positivo r, $r \in \mathbb{Q}_+$; se $f \leq 0$ è misurabile e n è dispari, anche $f^{1/n}$ è misurabile.*

Dimostrazione. La dimostrazione che $\alpha \cdot f$ è misurabile è una banale applicazione del Corollario 1.8.1.

Per $\omega \in \Omega$ è ovvio che si ha $f(\omega) + g(\omega) > c$ se, e solo se, esiste un razionale q tale che

$$f(\omega) > q > c - g(\omega).$$

Si numerino i razionali

$$\mathbb{Q} = \bigcup_{n \in \mathbb{N}} \{q_n\}.$$

Allora

$$\{f + g > c\} = \bigcup_{n \in \mathbb{N}} \Big(\{f > q_n\} \bigcap \{g > c - q_n\} \Big) ;$$

per il Corollario 1.8.1, ciascuno degli insiemi dei quali si fa l'unione appartiene a $\mathcal{F}$, sicché vi appartiene anche la loro unione numerabile. Dunque, $f+g$ è misurabile.

$$\{f^2 \le c\} = \begin{cases} \emptyset\,, & \text{se } c<0,\\ \{f=0\}, & \text{se } c=0,\\ \{-\sqrt{c} \le f \le \sqrt{c}\}\,, & \text{se } c>0; \end{cases}$$

poiché ognuno di questi insiemi è misurabile, f^2 è misurabile.

$$\{f^+ \ge c\} = \begin{cases} \Omega\,, & \text{se } c \le 0,\\ \{f \ge c\}\,, & \text{se } c>0, \end{cases}$$

sicché $\{f^+ \ge c\}$ è in $\mathcal{F}$ per ogni c reale; f^+ è dunque misurabile. In maniera analoga si stabilisce la misurabilità di f^-. Poiché $|f| = f^+ + f^-$, anche $|f|$ risulta misurabile. La misurabilità delle altre funzioni scende ora da quanto è già stato dimostrato e dalle identità

$$f \cdot g = \frac{1}{2}\left\{(f+g) - f^2 - g^2\right\}$$

$$f \vee g = \frac{1}{2}\{f+g+|f-g|\}\,, \qquad f \wedge g = f+g-(f \vee g)\,.$$

Si è visto sopra che f^2 è misurabile. Per induzione si supponga che sia misurabile f^n con $n \in \mathbb{N}$. Allora, dalla misurabilità del prodotto di due funzioni misurabili scende la misurabilità di $f^n \cdot f = f^{n+1}$, sicché f^n è misurabile per ogni $n \in \mathbb{N}$.

Vogliamo, infine, studiare la misurabilità di $f^{1/n}$ quando f è positiva; si ha

$$\left\{f^{1/n} \ge c\right\} = \begin{cases} \Omega, & \text{se } c \le 0,\\ \{f \ge c^n\}\,, & \text{se } c>0. \end{cases}$$

Se f è negativa e n è dispari, $n = 2k-1$ con $k \in \mathbb{N}$, si ha

$$\left\{f^{1/(2k-1)} \ge c\right\} = \left\{f \ge c^{2k-1}\right\}\,.$$

In ogni caso appartiene a $\mathcal{F}$. Dunque, $f^{1/n}$ è misurabile.

Infine, se r è un numero razionale positivo, si può scrivere $r = p/q$, con p e q naturali. L'asserto è ora conseguenza di quanto già visto. $\square$

Teorema 1.8.4. *Sia (f_n) una successione di funzioni misurabili da Ω in $\overline{\mathbb{R}}$; sono allora misurabili anche le seguenti funzioni:*

(a) $\vee_{n\in\mathbb{N}} f_n := \sup_{n\in\mathbb{N}} f_n$;

(b) $\wedge_{n\in\mathbb{N}} f_n := \inf_{n\in\mathbb{N}} f_n$,

(c) $\limsup_{n\to+\infty} f_n$;

(d) $\liminf_{n\to+\infty} f_n$;

(e) $\lim_{n\to+\infty} f_n$ *se la successione (f_n) converge.*

Dimostrazione. (a) $\{\vee_{n\in\mathbb{N}} f_n \leq c\} = \cap_{n\in\mathbb{N}}\{f_n \leq c\}$ e l'asserto segue ora da 1.8.1.

(b) segue da (a) perché $\wedge_{n\in\mathbb{N}} f_n = - \vee_{n\in\mathbb{N}} (-f_n)$.

(c) $\limsup_{n\to+\infty} f_n = \inf_{n\to+\infty} g_n$ con $g_n := \sup_{k\geq n} f_k$ e g_n è misurabile per ogni $n \in \mathbb{N}$ in virtú di (a); l'asserto segue ora da (b). In maniera analoga si dimostra (d). Quanto a (e), basta osservare che, se la successione (f_n) converge, si ha

$$\lim_{n\to+\infty} f_n = \limsup_{n\to+\infty} f_n = \liminf_{n\to+\infty} f_n .$$

Il risultato è cosí dimostrato. □

Il seguente corollario è un'immediata conseguenza dei due teoremi precedenti.

Corollario 1.8.2. *Sia f una funzione misurabile positiva, $f \geq 0$ e x un qualsiasi reale positivo; è allora misurabile la funzione f^x.*

Dimostrazione. Basta considerare un'arbitraria successione $(r_n)_{n\in\mathbb{N}}$ di razionali positivi convergenti a x ed osservare che $\lim_{n\to+\infty} f^{r_n} = f^x$. □

Definizione 1.8.3. Si dice $(\mathcal{F})$–*semplice* ogni funzione $f : \Omega \to \mathbb{R}$ che si possa scrivere nella forma

$$f = \sum_{j=1}^{n} c_j \, \mathbf{1}_{A_j} ,$$

ove $A_j \in \mathcal{F}$ per $j = 1, \dots, n$, $A_j \cap A_k = \emptyset$ se $j \neq k$, $\cup_{j=1}^n A_j = \Omega$ e $c_j \in \mathbb{R}$ $(j = 1, \dots, n)$. ◇

Si può fare a meno della condizione $\cup_{j=1}^n A_j = \Omega$; basterebbe, infatti, porre $A_{n+1} := \Omega \setminus \cup_{j=1}^n A_j$ se fosse $\cup_{j=1}^n A_j \neq \Omega$. Allora, si può scrivere $f = \sum_{j=1}^{n+1} c_j \, \mathbf{1}_{A_j}$, ponendo $c_{n+1} := 0$. Ogni funzione semplice assume un numero finito di valori. La rappresentazione di una funzione semplice non è unica, fatto che si sfrutta, per esempio nella dimostrazione del successivo Teorema 1.8.5. Se è ovvio dal contesto quale sia la tribú alla quale si fa riferimento si dirà che una funzione è semplice, anziché $\mathcal{F}$–semplice.

Teorema 1.8.5. *Siano f e g due funzioni semplici; tali sono allora anche le funzioni αf per ogni $\alpha \in \mathbb{R}$, $f \pm g$, $f \cdot g$, $f \vee g := \max\{f, g\}$, $f \wedge g := \min\{f, g\}$.*

Dimostrazione. Siano $c_1, \dots, c_m$ e $d_1, \dots, d_n$ numeri reali, $A_1, \dots, A_m$ e $B_1, \dots, B_n$ partizioni misurabili di Ω, vale a dire $\cup_{j=1}^m A_j = \cup_{k=1}^n B_k = \Omega$, con A_j e B_k in $\mathcal{F}$ per tutti gli indci j e k; infine $A_i \cap A_j = B_i \cap B_j = \emptyset$ per $i \neq j$. Si consideri la partizione piú fine di entrambe $(A_j \cap B_k)$ $(j = 1, \dots, m; k = 1, \dots, n)$. È ora evidente che si possono rappresentare f e g ricorrendo alla stessa partizione $(A_j \cap B_k)$

$$f = \sum_{k=1}^{n} \sum_{j=1}^{m} c_j \, \mathbf{1}_{A_j \cap B_k} ,$$

$$g = \sum_{k=1}^{n} \sum_{j=1}^{m} d_k \, \mathbf{1}_{A_j \cap B_k} ,$$

sicché

$$f + g = \sum_{k=1}^{n} \sum_{j=1}^{m} (c_j + d_k)\, \mathbf{1}_{A_j \cap B_k} ,$$

che rende evidente l'asserto per la somma. Si procede in maniera analoga negli altri casi. □

Teorema 1.8.6. *Ogni funzione $\mathcal{F}$–semplice è $\mathcal{F}$–misurabile.*

Dimostrazione. Sia $f = \sum_{j=1}^{n} c_j\, \mathbf{1}_{A_j}$; allora $\{f \le c\} = \cup \{A_j : c_j \le c\}$, che, come unione finita di insiemi di $\mathcal{F}$, appartiene a $\mathcal{F}$. □

L'importanza delle funzioni semplici è data dal seguente

Teorema 1.8.7. *Ogni funzione $f : \Omega \to \overline{\mathbb{R}}_+$ misurabile e positiva è il limite di almeno una successione crescente di funzioni semplici positive.*

Dimostrazione. Per ogni intero positivo s si ponga

$$A_{k,n} := \begin{cases} \left\{ \dfrac{(k-1)}{2^n} \le f < \dfrac{k}{2^n} \right\}, & \text{per } k = 1, \dots, 2^{2n}, \\ \{f \ge 2^n\}, & \text{per } k = 0. \end{cases} \tag{1.8.1}$$

Poiché f è misurabile, tutti gli insiemi $A_{k,n}$ appartengono a $\mathcal{F}$. Inoltre, se $k \ne k'$ si ha

$$A_{k,n} \bigcap A_{k',n} = \emptyset \qquad \text{e} \qquad \Omega = \bigcup_{k=0}^{2^{2n}} A_{k,n} .$$

La funzione $s_n : \Omega \to \mathbb{R}_+$ definita da

$$s_n := \sum_{s=1}^{2^{2n}} \frac{k-1}{2^n}\, \mathbf{1}_{A_{k,n}} + 2^n\, \mathbf{1}_{A_{0,n}} \tag{1.8.2}$$

è semplice e soddisfà alla diseguaglianza $0 \le s_n \le f$.

Se $k > 0$, si ha $A_{k,n} = A_{2k-1,n+1} \cup A_{2k,n+1}$ che è un'unione disgiunta. Se ω è in $A_{2k-1,n+1}$, si ha $s_n(\omega) = s_{n+1}(\omega) = (k-1)/2^n$, mentre, se ω è in $A_{2k,n+1}$, è

$$s_{n+1}(\omega) = \frac{2k-1}{2^{n+1}} = \frac{2k-2}{2^{n+1}} + \frac{1}{2^{n+1}} = s_n(\omega) + \frac{1}{2^{n+1}} ,$$

sicché, in ogni caso, $s_{n+1}(\omega) \ge s_n(\omega)$.. Se, poi, è $\omega \in A_{0,n}$, si osservi che

$$A_{0,n} = A_{0,n+1} \bigcup \left(\bigcup_{r=2^{2n+1}+1}^{2^{2n+2}} A_{r,n+1} \right) ,$$

e perciò $s_n(\omega) = 2^n \le s_{n+1}(\omega)$. Dunque, $s_n(\omega) \le s_{n+1}(\omega)$ per ogni $\omega \in \Omega$. Sia ora $\omega \in \Omega$ tale che $f(\omega) < +\infty$; esiste allora un numero naturale n tale $2^n > s_n(\omega)$; un tale punto appartiene ad un insieme $A_{k,n}$ con $k > 0$. Perciò $0 \le f(\omega) - s_n(\omega) \le 2^{-n}$, sicché $\lim_{n\to+\infty} s_n(\omega) = f(\omega)$. Se invece $f(\omega) = +\infty$, si ha $s_n(\omega) = 2^n$ per ogni $n \in \mathbb{N}$, sicché $\lim_{s\to+\infty} s_n(\omega) = +\infty$. □

Prima di chiudere questa sezione diamo il seguente teorema che in letteratura è pure noto come "Teorema della classe monotona" e che costiuisce una versione piú sofisticata del Teorema 1.6.1.

Teorema 1.8.8. *Sia $\mathcal{H}$ una classe di funzioni limitate da Ω in $\mathbb{R}$ s si supponga che $\mathcal{H}$ soddisfaccia alle proprietà:*

(a) *$\mathcal{H}$ è uno spazio vettoriale;*

(b) *la funzione identicamente eguale a 1 appartiene a $\mathcal{H}$;*

(c) *appartiene a $\mathcal{H}$ ogni funzione limitata su Ω che sia l'estremo superiore di una successione (f_n) crescente di funzioni positive di $\mathcal{H}$.*

Se, inoltre, sono in $\mathcal{H}$ le funzioni indicatrici di ogni insieme in un dato π–sistema $\mathcal{I}$, allora appartiene a $\mathcal{H}$ ogni funzione limitata definita in Ω che sia misurabile rispetto a $\mathcal{F}(\mathcal{I})$.

Dimostrazione. Sia f una funzione postiva misurabile rispetto a $\mathcal{F}(\mathcal{I})$ e si supponga che sia $0 \leq f(\omega) \leq K$ per ogni $\omega \in \Omega$. Come nell'eq. (1.8.2) si definisca per $s \in \mathbb{N}$

$$f_s := \sum_{s=1}^{K \wedge 2^{2s}} \frac{k-1}{2^s} \mathbf{1}_{A_{k,s}} ,$$

ove gli insiemi $A_{k,s}$ sono stati introdotti in (1.8.1). Poiché f è misurabile rispetto a $\mathcal{F}(\mathcal{I})$, ogni insieme $A_{k,s}$ è in $\mathcal{F}(\mathcal{I})$, sicché $\mathbf{1}_{A_{k,s}}$ appartiene a $\mathcal{H}$; e, poiché $\mathcal{H}$ è uno spazio vettoriale, f_s appartiene a $\mathcal{H}$ per ogni $s \in \mathbb{N}$. Ma (f_s) tende crescendo a f sicché f è in $\mathcal{H}$.

Il caso generale scende dall'usuale decomposizione $f = f^+ - f^-$. □

Esempio 1.8.1. Si considerino lo spazio di probabilità $([0,1[\,, \mathcal{B}([0,1[), \lambda)$ e la funzione $R_n(x) := 2\,x_n - 1$ ove x_n è la n–esima cifra nello sviluppo binario di $x \in [0,1]$:

$$x = \sum_{n \in \mathbb{N}} \frac{x_n}{2^n} \quad x_n \in \{0,1\} ,$$

ove per evitare ambiguità non si considerano gli sviluppi con solo un numero finito di zeri. Evidentemente

$$R_1(x) = \begin{cases} -1\,, & x \in [0,1/2[\;, \\ 1\,, & x \in [1/2,1[\;, \end{cases} \quad R_2(x) = \begin{cases} -1\,, & x \in [0,1/4[\cup [1/4,3/4[\;, \\ 1\,, & x \in [1/4,1/2[\cup [3/4,1]\;, \end{cases}$$

$$R_3(x) = \begin{cases} -1\,, & x \in [0,1/8[\cup [1/4,3/8[\cup [1/2,5/8[\cup [3/4,7/8[\;, \\ 1\,, & x \in [1/8,1/4[\cup [3/8,1/2[\cup [5/8,3/4[\cup [7/8,1]\;, \end{cases}$$

$$\cdots\cdots\cdots$$

$$R_n(x) = \begin{cases} -1\,, & x \in \bigcup_{k=0}^{2^{n-1}-1} \left[\frac{2k}{2^n}, \frac{2k+1}{2^n}\right[\;, \\ \\ 1\,, & x \in \bigcup_{k=0}^{2^{n-1}-1} \left[\frac{2k+1}{2^n}, \frac{2k+2}{2^n}\right]\;. \end{cases}$$

Le funzioni R_n sono misurabili e si chiamano *funzioni di Rademacher*; esse si usano in alcuni esempî. ■

1.9 La definizione dell'integrale

La definizione dell'integrale sarà date a tappe; prima per le funzioni semplici positive, quindi per le funzioni misurabili positive, ed infine sarà estesa alle funzioni che si diranno *integrabili*.

L'integrale per le funzioni semplici positive. Sia $(\Omega, \mathcal{F}, \mu)$ uno spazio mensurale, sia

$$f = \sum_{j=1}^{n} c_j \, \mathbf{1}_{A_j}$$

una funzione semplice positiva, tale, cioè, che, per ogni indice j, si abbia $c_j \geq 0$ e gli insiemi A_j appartengano tutti a $\mathcal{F}$, siano disgiunti e costituiscano una partizione di Ω. Si definisca allora come integrale di f rispetto alla misura μ la somma

$$\int f \, \mathrm{d}\mu := \sum_{j=1}^{n} c_j \, \mu(A_j) \,. \tag{1.9.1}$$

Occorre mostrare che la definizione appena data (1.9.1) non dipende dalla particolare rappresentazione della funzione f. Si considerino due diverse rappresentazioni di f:

$$f = \sum_{j=1}^{n} c_j \, \mathbf{1}_{A_j} = \sum_{k=1}^{r} d_k \, \mathbf{1}_{B_k} \,,$$

con $A_i \cap A_j = B_i \cap B_j = \emptyset$ se $i \neq j$, $\cup_{j=1}^{n} A_j = \cup_{k=1}^{r} B_k = \Omega$, $c_j, d_k \geq 0$, $A_i, B_k \in \mathcal{F}$ per ogni $i \leq n$ e per ogni $k \leq r$. Perciò,

$$\mu(A_i) = \sum_{j=1}^{r} \mu\left(A_i \bigcap B_j\right) \qquad \text{e} \qquad \mu(B_j) = \sum_{i=1}^{n} \mu\left(A_i \bigcap B_j\right) .$$

Se $A_i \cap B_j \neq \emptyset$, accade che $f(\omega) = c_i = d_j$, per ogni $\omega \in A_i \cap B_j$; perciò,

$$\begin{aligned}\sum_{i=1}^{n} c_i \, \mu(A_i) &= \sum_{i=1}^{n} \sum_{j=1}^{r} c_i \, \mu\left(A_i \bigcap B_j\right) \\ &= \sum_{j=1}^{r} \sum_{i=1}^{n} d_j \, \mu\left(A_i \bigcap B_j\right) = \sum_{j=1}^{r} d_j \, \mu(B_j) \,,\end{aligned}$$

di modo che la (1.9.1) è una buona definizione.

Siano ora f e g due funzioni semplici positive

$$f = \sum_{i=1}^{n} c_i \, \mathbf{1}_{A_i} \quad \text{e} \quad g = \sum_{j=1}^{r} d_j \, \mathbf{1}_{B_j} \,;$$

ricorrendo alle rappresentazioni

$$f = \sum_{i=1}^{n} \sum_{j=1}^{r} c_i \, \mathbf{1}_{A_i \cap B_j} \quad \text{e} \quad g = \sum_{i=1}^{n} \sum_{j=1}^{r} d_j \, \mathbf{1}_{A_i \cap B_j}$$

si può scrivere $f + g = \sum_{i=1}^{n} \sum_{j=1}^{r} (c_i + d_j)\, \mathbf{1}_{A_i \cap B_j}$, onde

$$\begin{aligned}\int (f+g)\,\mathrm{d}\mu &= \sum_{i=1}^{n}\sum_{j=1}^{r}(c_i+d_j)\,\mu\left(A_i \bigcap B_j\right)\\ &= \sum_{i=1}^{n} c_i \sum_{j=1}^{r}\mu\left(A_i \bigcap B_j\right) + \sum_{j=1}^{r} d_j \sum_{i=1}^{n}\mu\left(A_i \bigcap B_j\right)\\ &= \sum_{i=1}^{n} c_i\,\mu(A_i) + \sum_{j=1}^{r} d_j\,\mu(B_j) = \int f\,\mathrm{d}\mu + \int g\,\mathrm{d}\mu\,.\end{aligned}$$

Nella stessa maniera si dimostra che, se α e β sono reali positivi, si ha

$$\int (\alpha\, f + \beta\, g)\,\mathrm{d}\mu = \alpha \int f\,\mathrm{d}\mu + \beta \int g\,\mathrm{d}\mu\,,$$

e che, se $0 \le f \le g$, è

$$\int f\,\mathrm{d}\mu \le \int g\,\mathrm{d}\mu\,.$$

Funzioni misurabili positive. Se $f : \Omega \to \overline{\mathbb{R}}_+$ è misurabile, esiste una successione di funzioni semplici positive (s_n) con $s_n \uparrow f$ (Teorema 1.8.7). Per ogni $n \in \mathbb{N}$, è definito l'integrale $\int s_n\,\mathrm{d}\mu$; inoltre, si è appena visto che la successione degli integrali

$$\left(\int s_n\,\mathrm{d}\mu\right)_{n\in\mathbb{N}}$$

è crescente; essa ammette perciò limite. Si definisca

$$\int f\,\mathrm{d}\mu := \lim_{n\to+\infty} \int s_n\,\mathrm{d}\mu\,. \tag{1.9.2}$$

Occorre dimostrare che la (1.9.2) è una buona definizione, vale a dire che il valore dell'integrale non dipende dalla particolare successione (s_n) di funzioni semplici scelta per approssimare f. La dimostrazione fa uso del seguente

Lemma 1.9.1. *Sia (s_n) una successione crescente di funzioni semplici positive. Si supponga che esista una funzione semplice positiva $g : \Omega \to \mathbb{R}_+$ tale che*

$$\lim_{n\to+\infty} s_n \ge g\,.$$

Allora

$$\lim_{n\to+\infty} \int s_n\,\mathrm{d}\mu \ge \int g\,\mathrm{d}\mu\,. \tag{1.9.3}$$

Dimostrazione. Sia $g = \sum_{i=1}^{k} c_i\, \mathbf{1}_{A_i}$ e si supponga dapprima che sia $\int g\,\mathrm{d}\mu = +\infty$. Esiste perciò un indice i tale che $c_i > 0$ e $\mu(A_i) = +\infty$. Preso ε in $]0, c_i[$, si ponga

$$E_n := \{s_n + \varepsilon \ge g\} \quad (n \in \mathbb{N})\,;$$

la successione $(A_i \cap E_n)_{n\in\mathbb{N}}$ è crescente e tende a A_i, onde

$$\lim_{n\to+\infty} \mu\left(A_i \bigcap E_n\right) = +\infty\,.$$

Ora,

$$\int s_n \, \mathrm{d}\mu \geq \int s_n \, \mathbf{1}_{A_i \cap E_n} \, \mathrm{d}\mu \geq (c_i - \varepsilon) \, \mu \left(A_i \bigcap E_n \right) \xrightarrow[n \to \infty]{} +\infty \,,$$

ciò che stabilisce la (1.9.3), in queste ipotesi. Se poi è $\int g \, \mathrm{d}\mu < +\infty$, si ponga

$$E := \{g > 0\} = \bigcup \{A_i : c_i > 0\} \in \mathcal{F} \,.$$

Poiché g è semplice, si ha $c := \min\{c_i : c_i > 0\} > 0$ e $\mu(E) < +\infty$. Fissato $\varepsilon \in \,]0, c[$, si ha

$$\begin{aligned}
\int s_n \, \mathrm{d}\mu &\geq \int s_n \, \mathbf{1}_{E \cap E_n} \, \mathrm{d}\mu \geq \int (g - \varepsilon) \, \mathbf{1}_{E \cap E_n} \, \mathrm{d}\mu \\
&= \int g \, \mathbf{1}_{E_n \cap E} \, \mathrm{d}\mu - \varepsilon \mu \left(E_n \bigcap E \right) \geq \sum_{c_i > 0} \int g \, \mathbf{1}_{E_n \cap A_i} \, \mathrm{d}\mu - \varepsilon \mu(E) \\
&= \sum_{c_i > 0} c_i \, \mu \left(E_n \bigcap A_i \right) - \varepsilon \, \mu(E) \,,
\end{aligned}$$

e, poiché $\mu(E_n \cap A_i) \to \mu(A_i)$,

$$\lim_{n \to +\infty} \int s_n \, \mathrm{d}\mu \geq \sum_{c_i > 0} c_i \, \mu(A_i) - \varepsilon \, \mu(E) = \int g \, \mathrm{d}\mu - \varepsilon \, \mu(E) \,,$$

e, di qui, la (1.9.3) in virtú dell'arbitrarietà di ε. □

Siano ora (s_n) e (t_r) due successioni crescenti di funzioni semplici positive entrambe tendenti a f. Fissato $r \in \mathbb{N}$, si ha $\lim_{n \to +\infty} s_n = f \geq t_r$ onde, per il Lemma 1.9.1,

$$\lim_{n \to +\infty} \int s_n \, \mathrm{d}\mu \geq \int t_r \, \mathrm{d}\mu \,,$$

e, di qui,

$$\lim_{n \to +\infty} \int s_n \, \mathrm{d}\mu \geq \lim_{r \to +\infty} \int t_r \, \mathrm{d}\mu \,.$$

Scambiando i ruoli delle successioni (s_n) e (t_r) si ottiene la diseguaglianza nell'altro verso, sicché si può concludere che

$$\lim_{n \to +\infty} \int s_n \, \mathrm{d}\mu = \lim_{r \to +\infty} \int t_r \, \mathrm{d}\mu \,,$$

ciò che dimostra che la (1.9.2) è una buona definizione.

Dalla stessa proprietà valida per le funzioni semplici positive si deduce che, se f e g sono funzioni misurabili e positive e se α e β sono numeri reali positivi, vale

$$\int (\alpha \, f + \beta \, g) \, \mathrm{d}\mu = \alpha \int f \, \mathrm{d}\mu + \beta \int g \, \mathrm{d}\mu \,.$$

Funzioni misurabili integrabili. Sia $f : \Omega \to \overline{\mathbb{R}}$ misurabile; tali sono anche le funzioni positive f^+ e f^-. Se tanto f^+ quanto f^- hanno integrale finito, si ponga

$$\int f \, \mathrm{d}\mu := \int f^+ \, \mathrm{d}\mu - \int f^- \, \mathrm{d}\mu \,.$$

Definizione 1.9.1. Si dice che una funzione $f : \Omega \to \overline{\mathbb{R}}$ è *integrabile* se tanto $\int f^+ \,\mathrm{d}\mu$ quanto $\int f^- \,\mathrm{d}\mu$ sono finiti; si dirà che f è *semi-integrabile* se solo uno dei due integrali $\int f^+ \,\mathrm{d}\mu$ o $\int f^- \,\mathrm{d}\mu$ è finito, precisamente f si dice *semi-integrabile inferiormente* se $\int f^+ \,\mathrm{d}\mu = +\infty$ mentre $\int f^- \,\mathrm{d}\mu < +\infty$ nel qual caso si porrà $\int f \,\mathrm{d}\mu := +\infty$; si dice invece che f è *semi-integrabile superiormente* se $\int f^+ \,\mathrm{d}\mu < +\infty$ e $\int f^- \,\mathrm{d}\mu = +\infty$, nel qual caso si pone $\int f \,\mathrm{d}\mu := -\infty$.

Se accade che sia $\int f^+ \,\mathrm{d}\mu = \int f^- \,\mathrm{d}\mu = +\infty$, non si definisce l'integrale $\int f \,\mathrm{d}\mu$.

Se $A \in \mathcal{F}$ si definisce l'integrale esteso all'insieme A mediante

$$\int\limits_A f \,\mathrm{d}\mu := \int f \,\mathbf{1}_A \,\mathrm{d}\mu \,,$$

se esiste l'integrale a secondo membro. $\diamond$

Dato lo spazio mensurale $(\Omega, \mathcal{F}, \mu)$, l'insieme delle funzioni $f : \Omega \to \overline{\mathbb{R}}$ integrabili rispetto alla misura μ sarà denotato con $\mathcal{L}^1 = \mathcal{L}^1(\Omega, \mathcal{F}, \mu)$. Nel seguito si incontreranno le proprietà di tale spazio.

1.10 Proprietà dell'integrale

Dato uno spazio mensurale $(\Omega, \mathcal{F}, \mu)$ si dice che una proprietà vale *quasi ovunque* (abbreviato in q.o. in italiano, a.e.=*almost everywhere* in inglese, p.p.=*presque partout* in francese) se l'insieme dei punti nei quali la proprietà non vale è contenuto in un insieme trascurabile, cioè misurabile e di misura nulla.

Teorema 1.10.1. *Siano $f, g : \Omega \to \overline{\mathbb{R}}$ funzioni integrabili; allora*

(a) *esiste finito l'integrale* $\int\limits_A f \,\mathrm{d}\mu$*;*

(b) *se $A \cap B = \emptyset$ con $A, B \in \mathcal{F}$, è*

$$\int\limits_{A \cup B} f \,\mathrm{d}\mu = \int\limits_A f \,\mathrm{d}\mu + \int\limits_B f \,\mathrm{d}\mu \,;$$

(c) *f è finita q.o.* $(\mu(\{|f| = +\infty\}) = 0)$*;*

(d) *$f + g$ è integrabile e*

$$\int (f+g) \,\mathrm{d}\mu = \int f \,\mathrm{d}u + \int g \,\mathrm{d}u \,;$$

(e) $f \geq 0 \Longrightarrow \int f \,\mathrm{d}\mu \geq 0$ *e* $f \geq g \Longrightarrow \int f \,\mathrm{d}\mu \geq \int g \,\mathrm{d}u$ *(isotonía);*

(f) *$|f|$ è integrabile e*

$$\left| \int f \,\mathrm{d}\mu \right| \leq \int |f \,\mathrm{d}\mu \,;$$

(g) *per ogni c reale, cf è integrabile e*

$$\int cf \,\mathrm{d}\mu = c \int f \,\mathrm{d}\mu \,;$$

(h) *$f \geq 0$ e $\int f \, \mathrm{d}\mu = 0 \Longrightarrow f = 0$ q.o.;*

(i) *$f = g$ q.o. $\Longrightarrow \int f \, \mathrm{d}\mu = \int g \, \mathrm{d}\mu$;*

(j) *se $h : \Omega \to \overline{\mathbb{R}}$ è misurabile e $|h| \leq |f|$, h è integrabile.*

Dimostrazione. (a) L'asserto scende dalle ovvie diseguaglianze

$$0 \leq f^+ \mathbf{1}_A \leq f^+ \quad \text{e} \quad 0 \leq f^- \mathbf{1}_A \leq f^- ;$$

la (b) scende dalle relazioni

$$\mathbf{1}_{A \cup B} = \mathbf{1}_A + \mathbf{1}_B ,$$
$$f^+ \mathbf{1}_{A \cup B} = f^+ \mathbf{1}_A + f^+ \mathbf{1}_B , \quad f^- \mathbf{1}_{A \cup B} = f^- \mathbf{1}_A + f^- \mathbf{1}_B .$$

(c) Se f non fosse finita q.o., almeno uno dei due insiemi

$$E_1 := \{f = +\infty\} \quad \text{e} \quad E_2 := \{f = -\infty\}$$

avrebbe misura strettamente positiva; si supponga, per esempio, che sia $\mu(E_1) > 0$. Poiché la definizione di integrale implica che, se $f \geq 0$, sia anche $\int f \, \mathrm{d}\mu \geq 0$, dalla diseguaglianza $0 \leq f^+ \mathbf{1}_{E_1} \leq f^+$ scende

$$\int f^+ \, \mathrm{d}\mu \geq \int_{E_1} f^+ \, \mathrm{d}\mu = +\infty ,$$

sicché f non sarebbe integrabile.

(d) Per la (c), si può supporre che f e g siano finite in tutto Ω. Allora

$$f + g = (f+g)^+ - (f+g)^- \quad \text{e} \quad f + g = f^+ - f^- + g^+ - g^- ,$$

onde, dal confronto, $(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+$. Poiché

$$(f+g)^+ \leq f^+ + g^+ \quad \text{e} \quad (f+g)^- \leq f^- + g^- ,$$

si ha, intanto, che $(f+g)^+$ e $(f+g)^-$ sono integrabili. L'asserto è noto per le funzioni misurabili positive e perciò

$$\int (f+g)^+ \, \mathrm{d}\mu + \int f^- \, \mathrm{d}\mu + \int g^- \, \mathrm{d}\mu = \int (f+g)^- \, \mathrm{d}\mu + \int f^+ \, \mathrm{d}\mu + \int g^+ \, \mathrm{d}\mu ,$$

relazione dalla quale scende facilmente quanto si voleva dimostrare.

(e) La prima asserzione è ovvia; quanto alla seconda si scriva $f = g + (f - g)$, ove $f - g \geq 0$. Perciò

$$\int f \, \mathrm{d}\mu = \int g \, \mathrm{d}\mu + \int (f - g) \, \mathrm{d}\mu \geq \int g \, \mathrm{d}\mu .$$

(f) Si applichi la (d) a $|f| = f^+ + f^-$ che risulta quindi integrabile; inoltre

$$\left| \int f \, \mathrm{d}\mu \right| = \left| \int f^+ \, \mathrm{d}\mu - \int f^- \, \mathrm{d}\mu \right| \leq \int f^+ \, \mathrm{d}\mu + \int f^- \, \mathrm{d}\mu = \int |f| \, \mathrm{d}\mu .$$

(g) Se $c = 0$, è $cf = 0$ e $c\int f \,\mathrm{d}\mu = 0 = \int (cf)\,\mathrm{d}\mu$. Se $c > 0$, è $(cf)^+ = cf^+$ e $(cf)^- = cf^-$, sicché

$$\int cf \,\mathrm{d}\mu = \int cf^+ \,\mathrm{d}\mu - \int cf^- \,\mathrm{d}\mu = c\int f^+ \,\mathrm{d}\mu - c\int f^- \,\mathrm{d}\mu = c\int f \,\mathrm{d}\mu\,.$$

Se, invece, $c < 0$, si ha $(cf)^+ = -cf^-$ e $(cf)^- = -cf^+$, onde

$$\begin{aligned}\int cf \,\mathrm{d}\mu &= \int (cf)^+ \,\mathrm{d}\mu - \int (cf)^- \,\mathrm{d}\mu \\ &= \int (-c)f^- \,\mathrm{d}\mu + c\int f^+ \,\mathrm{d}\mu = c\int f \,\mathrm{d}\mu\,.\end{aligned}$$

(h) Si osservi che, definendo, per ogni $n \in \mathbb{N}$, $A_n := \{f \geq 1/n\}$, si ha $A_n \in \mathcal{F}$ per ogni $n \in \mathbb{N}$ e si può scrivere $\{f > 0\} = \cup_{n\in\mathbb{N}} A_n$; quindi $\mu(f > 0) \leq \sum_{n\in\mathbb{N}} \mu(A_n)$. D'altro canto, per ogni $n \in \mathbb{N}$, si ha

$$\frac{\mathbf{1}_{A_n}}{n} \leq f\,\mathbf{1}_{A_n} \leq f\,,$$

onde

$$\int f \,\mathrm{d}\mu \geq \frac{1}{n}\int \mathbf{1}_{A_n} \,\mathrm{d}\mu = \frac{\mu(A_n)}{n} \geq 0\,;$$

perciò $\mu(A_n) = 0$ per ogni $n \in \mathbb{N}$, e, di conseguenza $\mu(\{f > 0\}) = 0$.
(i) $f = g$ q.o. implica $f^+ = g^+$ q.o. e $f^- = g^-$ q.o.. Gli insiemi $A_{k,s}$, che nel Teorema 1.8.7 servivano per costruire l'approssimazione di una funzione misurabile positiva mediante funzioni semplici, hanno, per f^+ e per g^+, la stessa misura; perciò, se (s_n) e (s'_n) sono successioni crescenti di funzioni semplici tendono a f^+ e a g^+ rispettivamente, si ha $\int s_n \,\mathrm{d}\mu = \int s'_n \,\mathrm{d}\mu$ per ogni $n \in \mathbb{N}$ e, di qui, $\int f^+ \,\mathrm{d}\mu = \int g^+ \,\mathrm{d}\mu$. Similmente si mostra che $\int f^- \,\mathrm{d}\mu = \int g^- \,\mathrm{d}\mu$.
(j) Si ha $0 \leq h^+ \leq |f|$ e $0 \leq h^- \leq |f|$; perciò gli integrali delle funzioni semplici che approssimano h^+ e h^- sono crescenti e entrambe limitate superiormente. □

Segue, in particolare, dalle proprietà (d) e (g) del Teorema 1.10.1 che l'insieme $\mathcal{L}^1(\Omega, \mathcal{F}, \mu)$ delle funzioni $f : \Omega \to \overline{\mathbb{R}}$ integrabili rispetto alla misura μ è uno spazio lineare.

I teoremi che chiudono questa sezione sono tra quelli di uso piú frequente.

Teorema 1.10.2. (Teorema di Beppo Levi o di convergenza monotona). *Dato lo spazio mensurale $(\Omega, \mathcal{F}, \mu)$, se (f_n) è una successione crescente di funzioni misurabili positive che converge puntualmente alla funzione f (necessariamente misurabile e positiva), si può passare al limite sotto il segno d'integrale*

$$\lim_{n\to+\infty} \int f_n \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu\,,$$

nel senso che, se f è integrabile, si ha $\int f_n \,\mathrm{d}\mu \to \int f \,\mathrm{d}\mu$, mentre, se $\int f \,\mathrm{d}\mu = +\infty$, allora $\int f_n \,\mathrm{d}\mu \to +\infty$.

Dimostrazione. Per ogni $n \in \mathbb{N}$ sia $(s_{nk})_{k\in\mathbb{N}}$ una successione crescente di funzioni semplici positive tendente a $f_n := \lim_{k\to+\infty} s_{nk} = f_n$. Posto

$$g_k := \max\{s_{nk} : n \leq k\}\,,$$

la successione (g_k) è crescente; inoltre, la funzione g_k è semplice e positiva. Perciò $g := \lim_{k\to+\infty} g_k$ è una funzione misurabile positiva. Ora, si ha, per ogni $n \in \mathbb{N}$ e per $k \geq n$, $s_{nk} \leq g_k \leq f_k \leq f$, sicché, facendo tendere k a $+\infty$, si ha, per ogni $n \in \mathbb{N}$, $f_n \leq g \leq f$, diseguaglianza dalla quale scende, facendo tendere n a $+\infty$, che $g = f$. Siccome

$$\int s_{nk}\,\mathrm{d}\mu \leq \int g_k\,\mathrm{d}\mu \leq \int f_k\,\mathrm{d}\mu \qquad (n \leq k)\,,$$

e, per la definizione di integrale di una funzione misurabilie positiva,

$$\sup_{n\in\mathbb{N}} \int g_k\,\mathrm{d}\mu = \int g\,\mathrm{d}\mu\,,$$

segue facendo tendere k a $+\infty$, che

$$\int f_n\,\mathrm{d}\mu \leq \int g\,\mathrm{d}\mu \leq \lim_{k\to+\infty} \int f_k\,\mathrm{d}\mu\,;$$

si faccia ora tendere n a $+\infty$ per ottenere

$$\lim_{n\to+\infty} \int f_n\,\mathrm{d}\mu \leq \int g\,\mathrm{d}\mu \leq \lim_{k\to+\infty} \int f_k\,\mathrm{d}\mu\,,$$

cioè

$$\lim_{n\to+\infty} \int f_n\,\mathrm{d}\mu = \int g\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu\,,$$

vale a dire l'asserto. □

Naturalmente lo stesso risultato vale se, anziché avere la convergenza puntuale di (f_n) a f, la successione (f_n) converge a f quasi ovunque.

Teorema 1.10.3. (Fatou). *Se (f_n) è una successione di funzioni misurabili limitata inferiormente da una funzione integrabile g, $f_n \geq g \in \mathcal{L}^1$ per ogni $n \in \mathbb{N}$, allora*

$$\int \liminf_{n\to+\infty} f_n\,\mathrm{d}\mu \leq \liminf_{n\to+\infty} \int f_n\,\mathrm{d}\mu\,.$$

Dimostrazione. Si supponga, dapprima, $g = 0$. Posto $g_n := \inf_{k\geq n} f_k$ $(n \in \mathbb{N})$, (g_n) risulta essere una successione crescente di funzioni misurabili positive con

$$\lim_{n\to+\infty} g_n = \liminf_{n\to+\infty} f_n\,.$$

Poiché $f_n \geq g_n$ per ogni $n \in \mathbb{N}$, il teorema di Beppo Levi dà

$$\liminf_{n\to+\infty} \int f_n\,\mathrm{d}\mu \geq \lim_{n\to+\infty} \int g_n\,\mathrm{d}\mu = \int \lim_{n\to+\infty} g_n\,\mathrm{d}\mu = \int \liminf_{n\to+\infty} f_n\,\mathrm{d}\mu\,.$$

Se, invece, g non è identicamente nulla, basta applicare quanto appena dimostrato alla successione (h_n) con $h_n := f_n - g \geq 0$, tenendo presente che

$$\int h_n\,\mathrm{d}\mu = \int f_n\,\mathrm{d}\mu - \int g\,\mathrm{d}\mu \quad \text{e} \quad \liminf_{n\to+\infty} h_n = \liminf_{n\to+\infty} f_n - g\,,$$

che conclude la dimostrazione. □

Si applichi il lemma di Fatou alla successione $(-f_n)$.

Teorema 1.10.4. *Se (f_n) è una successione di funzioni misurabili limitata superiormente da una funzione integrabile g, $f_n \le g \in \mathcal{L}^1$ per ogni $n \in \mathbb{N}$, allora*

$$\int \limsup_{n\to+\infty} f_n \,\mathrm{d}\mu \ge \limsup_{n\to+\infty} \int f_n \,\mathrm{d}\mu \,.$$

Teorema 1.10.5. (Teorema di Lebesgue o di convergenza dominata). *Dato lo spazio mensurale $(\Omega, \mathcal{F}, \mu)$, sia $g : \Omega \to \overline{\mathbb{R}}_+$ una funzione integrabile e sia (f_n) una successione di funzioni misurabili $f_n : (\Omega, \mathcal{F}) \to (\overline{\mathbb{R}}, \mathcal{B})$ $(n \in \mathbb{N})$ che converge puntualmente (o quasi ovunque) alla funzione f; se $|f_n| \le g$ per ogni $n \in \mathbb{N}$, allora f è integrabile e si può passare al limite sotto il segno d'integrale:*

$$\lim_{n\to+\infty} \int f_n \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu \,.$$

Dimostrazione. Se $f_n \ge 0$ e $f_n \to 0$, i Teoremi 1.10.3 e 1.10.4 danno

$$\begin{aligned}\int \limsup_{n\to+\infty} f_n \,\mathrm{d}\mu = \int \liminf_{n\to+\infty} f_n \,\mathrm{d}\mu &\le \liminf_{n\to+\infty} \int f_n \,\mathrm{d}\mu \\ &\le \limsup_{n\to+\infty} \int f_n \,\mathrm{d}\mu \le \int \limsup_{n\to+\infty} f_n \,\mathrm{d}\mu \,;\end{aligned}$$

le diseguaglianze scritte sono, in effetti, tutte eguaglianze e perciò

$$\lim_{n\to+\infty} \int f_n \,\mathrm{d}\mu = 0 \,.$$

In generale, si consideri la successione (h_n) definita, per ogni $n \in \mathbb{N}$, da $h_n := |f_n - f|$. Per questa successione sono verificate le condizioni della prima parte, $h_n \ge 0$ e $h_n \to 0$. Inoltre, vale $h_n \le 2\,g$ per ogni $n \in \mathbb{N}$. Ora, per il Teorema 1.10.1 (f),

$$\left| \int f_n \,\mathrm{d}\mu - \int f \,\mathrm{d}\mu \right| \le \int h_n \,\mathrm{d}\mu \xrightarrow[n\to+\infty]{} 0 \,,$$

che dà l'asserto. □

Esempio 1.10.1. Sia Ω un insieme costituito da un'infinità numerabile o da un numero finito di punti ω_n. Assegnare sulla famiglia delle parti di Ω una misura μ equivale a dare numeri positivi $\mu_n = \mu(\{\omega_n\})$. Per ogni sottoinsieme A di Ω, si ha

$$\mu(A) = \sum_{\omega_n \in A} \mu_n \,.$$

Naturalmente, si avrà $\sum_n \mu_n = \mu(\Omega)$. Ogni funzione $f : \Omega \to \mathbb{R}$ è misurabile rispetto alla tribú $\mathcal{P}(\Omega)$; ha senso, perciò, considerare l'integrale

$$\int |f| \,\mathrm{d}\mu = \sum_n \int_{\{\omega_n\}} |f| \,\mathrm{d}\mu = \sum_n |f(\omega_n)| \, \mu_n \,.$$

La funzione f è dunque integrabile se, e solo se, $\sum_n |f(\omega_n)|\, \mu_n < +\infty$; in tal caso è

$$\int f \,\mathrm{d}\mu = \sum_n f(\omega_n)\, \mu_n \,. \tag{1.10.1}$$

In questo modo, ogni serie assolutamente convergente, oppure ogni somma finita di numeri reali, $\sum_n \alpha_n$ può essere riguardata come l'integrale di un'opportuna funzione rispetto alla *misura del contare*, vale a dire la misura ν definita sulla famiglia delle parti di $\mathbb{N}$ mediante $\nu(\{n\}) = 1$ per ogni $n \in \mathbb{N}$. Basta, infatti, definire una funzione f mediante $f(n) := \alpha_n$ per avere, secondo la (1.10.1),

$$\sum_n \alpha_n = \sum_n f(n) = \sum_n f(n)\, \nu(\{n\}) = \int f \,\mathrm{d}\nu \,.$$

La misura del contare non è finita ed attribuisce ad ogni sottoinsieme di $\mathbb{N}$ misura eguale a $+\infty$ se l'insieme è infinito, ed eguale alla sua cardinalità se è finito. ■

1.11 Misura immagine

Siano $(\Omega, \mathcal{F}, \mu)$ uno spazio mensurale, $(\Omega_1, \mathcal{F}_1)$ uno spazio misurabile e f una funzione misurabile tra $(\Omega, \mathcal{F})$ e $(\Omega_1, \mathcal{F}_1)$. Si può allora definire una funzione ν da $\mathcal{F}_1$ in $\overline{\mathbb{R}}_+$ mediante

$$\nu(E) := \mu\left[f^{-1}(E)\right] \qquad (E \in \mathcal{F}_1)\,.$$

è immediato controllare che ν cosí definita è una misura; essa si dice *misura immagine di μ mediante f* . Poiché $\nu(\Omega_1) = \mu(\Omega)$, ν è finita se, e solo se, tale è anche μ. Si è soliti indicare la misura immagine mediante $\mu\, f^{-1}$.

Vale il seguente importante teorema sul "cambio di variabile".

Teorema 1.11.1. *Siano $(\Omega, \mathcal{F})$ e $(\Omega_1, \mathcal{F}_1)$ due spazî misurabili e μ una misura su $(\Omega, \mathcal{F})$; se la funzione $g : \Omega_1 \to \overline{\mathbb{R}}$ è $\mathcal{F}_1$–misurabile, è*

$$\int_{\Omega_1} g \,\mathrm{d}(\mu\, f^{-1}) = \int_{\Omega} g \circ f \,\mathrm{d}\mu \,, \tag{1.11.1}$$

nel senso che, se esiste uno dei sue integrali, esiste anche l'altro e i due sono eguali.

Dimostrazione. Alla luce della definizione di integrale, ci si può limitare a considerare funzioni $g : \Omega_1 \to \overline{\mathbb{R}}$ positive. Sia, dapprima, $g = \mathbf{1}_E$ con $E \in \mathcal{F}_1$; allora

$$(g \circ f)(\omega) = \begin{cases} 1\,, & \text{se } \omega \in f^{-1}(E), \\ 0\,, & \text{se } \omega \in f^{-1}(E); \end{cases}$$

perciò, $(\mathbf{1}_E \circ f) = \mathbf{1}_{f^{-1}(E)}$ con $f^{-1}(E) \in \mathcal{F}$ e

$$\int_{\Omega} g \circ f \;\mathrm{d}\mu = \int_{\Omega} \mathbf{1}_{f^{-1}(E)} \;\mathrm{d}\mu = (\mu\, f^{-1})(E) = \nu(E) = \int_{\Omega_1} \mathbf{1}_E \;\mathrm{d}(\mu\, f^{-1})\,.$$

La (1.11.1) vale dunque per le funzioni indicatrici di insiemi di $\mathcal{F}_1$ e, come conseguenza della linearità degli integrali, per le funzioni $\mathcal{F}_1$–semplici positive. Sia g

una funzione $\mathcal{F}_1$–misurabile e positiva e sia (s_n) una successione crescente di funzioni semplici convergente a g; allora $(s_n \circ f)$ tende a $g \circ f$ crescendo e, di piú, ogni funzione $s_n \circ f$ è semplice. Il ricorso alla definizione di integrale di una funzione positiva misurabile completa la dimostrazione. □

La dimostrazione appena data è il prototipo di molte altre: si stabilisce il risultato per le funzioni semplici, quindi lo si estende alle funzioni semplici positive, quindi, ancora, alle funzioni misurabili positive ed infine alle funzioni integrabili, sicché, di fatto, la dimostrazione si riduce a controllare che il risultato sia vero per le sole funzioni indicatrici.

Sia μ una misura finita e $f : \Omega \to \overline{\mathbb{R}}$ una funzione misurabile; si definisca una funzione in $\mathbb{R}$ una funzione F mediante $F(x) := \mu(\{f \leq x\})$. Si dimostra facilmente che F è crescente e continua a destra e che

$$\lim_{x\to-\infty} F(x) = 0 \quad \text{e} \quad \lim_{x\to+\infty} F(x) = \mu(\Omega)\,.$$

Per la dimostrazione basta osservare che, per la misura immagine $\nu = \mu \circ f^{-1}$ su $(\mathbb{R}, \mathcal{B})$, si ha $F(x) = \nu(]-\infty, x])$ e tenere presente il Teorema 1.7.1. Se μ_F è la misura di Stieltjes associata a F (Teorema 1.7.2), vale il

Teorema 1.11.2. *Sia $(\Omega, \mathcal{F}, \mu)$ è uno spazio mensurale e $f : \Omega \to \overline{\mathbb{R}}$ una funzione misurabile; se la funzione $F : \mathbb{R} \to [0, \mu(\Omega)]$ è definita, per ogni $x \in \mathbb{R}$ da $F(x) = \mu(f \leq x)$ risulta $\mu f^{-1} = \mu_F$; inoltre se $g : \mathbb{R} \to \overline{\mathbb{R}}$ è misurabile vale l'eguaglianza*

$$\int_\Omega g \circ f \,\mathrm{d}\mu = \int_{\mathbb{R}} g \,\mathrm{d}\mu_F = \int_{\mathbb{R}} g \,\mathrm{d}(\mu f^{-1})\,, \tag{1.11.2}$$

nel senso che, se esiste uno dei due integrali, esiste anche l'altro e i due sono eguali.

Dimostrazione. Alla luce del Teorema 1.6.7, per dimostrare l'eguaglianza

$$\mu_F(B) = (\mu f^{-1})(B)$$

per ogni boreliano B basta mostrare che essa è vera per ogni intervallo $]a, b]$ con $a, b \in \mathbb{R}$ e $a < b$; in tal caso,

$$(\mu f^{-1})(]a, b]) = \mu(a < f \leq b) = F(b) - F(a) = \mu_F(]a, b])\,,$$

sicché risulta effettivamente $\mu f^{-1} = \mu_F$.

Per dimostrare la (1.11.2), si ricorra al teorema del cambio di variabile 1.11.1; poiché $\mu f^{-1} = \mu_F$ la (1.11.1) dà appunto la (1.11.2). □

Spesso l'integrale a secondo membro della (1.11.2) si scrive nella forma

$$\int_{\mathbb{R}} g \,\mathrm{d}\mu_F = \int_{\mathbb{R}} g(x) \,\mathrm{d}F(x) = \int_{\mathbb{R}} g \,\mathrm{d}F\,,$$

che si dice *integrale di Lebesgue–Stieltjes della funzione f rispetto alla f.r. F.*

Se nella (1.11.2) si prende come integrando $g : \mathbb{R} \to \mathbb{R}$ la funzione $g(x) := x^k$ con $k \in \mathbb{Z}_+$ si ottiene la formula che è spesso utile per il calcolo effettivo dei momenti

$$\int_\Omega f^k \,\mathrm{d}\mu = \int_{\mathbb{R}} x^k \,\mathrm{d}\mu_F(x) = \int_{\mathbb{R}} x^k \,\mathrm{d}F(x)\,.$$

1.12 Vocabolario

Diversi dei concetti che si conoscono dai corsi elementari hanno in probabilità un nome diverso da quello che hanno in un altro contesto; si rende perciò necessario un "vocabolario" che consenta di tradurre i nomi.

Definizione 1.12.1. Si dirà *spazio di probabilità* o *spazio probabilizzato* uno spazio mensurale $(\Omega, \mathcal{F}, \mathbb{P})$ nel quale la misura $\mathbb{P}$, detta, *probabilità* o *misura di probabilità* sia normalizzata, $\mathbb{P}(\Omega) = 1$. $\diamond$

Definizione 1.12.2. Dato uno spazio misurabile $(\Omega, \mathcal{F})$ si dice *variabile aleatoria (o casuale)*(spesso abbreviato in v.a.) una funzione $X : \Omega \to \overline{\mathbb{R}}$ che sia misurabile rispetto alle tribú $\mathcal{F}$ e $\mathcal{B}$. Qui, $\mathcal{B}$ è la tribú di Borel in $\overline{\mathbb{R}}$. $\diamond$

Benché nella definizione di v.a. non vi sia menzione di una probabilità definita su $(\Omega, \mathcal{F})$, si suole parlare di v.a. sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ perché i problemi tipici di calcolo delle probabilità sono solitamente riconducibili alla forma: *Qual è la probabilità che una v.a. X assuma valori in un boreliano A?*

Le v.a. sono, solitamente, indicate mediante le ultime lettere maiuscole dell'alfabeto, con suffissi, ove occorra.

Alla nozione di proprietà valida quasi ovunque, si sostituisce, con lo stesso significato, quella di proprietà valida *quasi certamente* (abbreviato in q.c.). Altri cambiamenti di linguaggio si incontreranno tra breve.

Nel seguito si considereranno quasi esclusivamente v.a. X q.c. finite, tali, cioè, che $\mathbb{P}(|X| = +\infty) = 0$. Perciò, di fatto, le v.a. saranno funzioni definite in Ω a valori in $\mathbb{R}$ anziché in $\overline{\mathbb{R}}$. Per indicare che X è una v.a. scriveremo anche $X \in L^0$; e ricorreremo all'abuso di linguaggio che consiste nel considerare piuttosto le classi di equivalenza che non le singole funzioni. Un tale atteggiamento non può però essere adottato nel trattare di processi stocastici, che saranno però toccati solo marginalmente in queste lezioni.

Alla luce del Corollario 1.8.1, per verificare che una funzione $X : \Omega \to \mathbb{R}$ sia una v.a. basta, per esempio, mostrare che, per ogni $x \in \mathbb{R}$, appartiene alla tribú $\mathcal{F}$ l'insieme $\{X < x\} = X^{-1}([-\infty, x[)$.

Definizione 1.12.3. Data una v.a. X su $(\Omega, \mathcal{F}, \mathbb{P})$ si dice *legge* o *distribuzione* di X la misura immagine su $(\mathbb{R}, \mathcal{B})$ di $\mathbb{P}$ mediante X, cioè la misura $\mathbb{P}$ su $(\mathbb{R}, \mathcal{B})$ definita da $\mathbb{P}_X(B) := \mathbb{P}(X^{-1}(B)) = \mathbb{P}(X \in B)$ per ogni insieme $B \in \mathcal{B}$. $\diamond$

Ogni v.a. X induce quindi una probabilità $\mathbb{P}$ su $\mathcal{B}$. Si dice *funzione di ripartizione (o di distribuzione)*, di solito abbreviato in f.r., della v.a. X la funzione F_X da $\mathbb{R}$ in $[0, 1]$ definita da

$$F_X(x) := \mathbb{P}_X(]-\infty, x]) = \mathbb{P}(X \leq x) \quad (x \in \mathbb{R}).$$

Si scriverà F, in vece di F_X, per la f.r. di una v.a. X tutte le volte che ciò non generi confusione.

Definizione 1.12.4. Due v.a. X e X', non necessariamente definite sopra il medesimo spazio di probabilità, si dicono *isonome* o *identicamente distribuite*, o, ancora *somiglianti*, se hanno la stessa legge, cioè se $\mathbb{P}_X = \mathbb{P}_{X'}$. Due v.a. sono iisonome se, e solo se, esse hanno la stessa f.r.. $\diamond$

Analogamente, se una v.a. X definita in $(\Omega, \mathcal{F}, \mathbb{P})$ ammette integrale finito, tale integrale si indicherà mediante la notazione

$$\mathbb{E}(X) := \int X \, \mathrm{d}\mathbb{P} \,,$$

che si dirà *speranza (matematica)* di X. Si dice *varianza* di X, se esiste, finita, la quantità

$$\mathrm{V}(X) := \mathbb{E}\left[(X - \mathbb{E}(X)^2\right] \,.$$

Teorema 1.12.1. (Diseguaglianza di Markov). *Sia X una v.a. sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$; se X è positiva $(X \geq 0)$ si ha, quale che sia il numero reale $b > 0$,*

$$\mathbb{P}(X \geq b) \leq \frac{1}{b} \, \mathbb{E}(X) \,.$$

Dimostrazione. Poiché l'integrale che definisce la speranza è positivo, si ha $0 \leq \mathbb{E}(X) \leq +\infty$. Se $\mathbb{E}(X) = +\infty$, non vi è nulla da dimostrare. Si supponga, perciò, che sia $\mathbb{E}(X) < +\infty$, vale a dire, si supponga che X sia integrabile, $X \in L^1$, e si ponga $A(b) := \{X > b\}$. Poiché è evidente che $X \geq X \cdot \mathbf{1}_{A(b)}$, risulta, dalla definizione di speranza, che

$$\begin{aligned} \mathbb{E}(X) &\geq \mathbb{E}(X \cdot \mathbf{1}_{A(b)}) \geq \mathbb{E}(b \cdot \mathbf{1}_{A(b)}) \\ &= b \, \mathbb{E}(\mathbf{1}_{A(b)}) = b \, \mathbb{P}(A(b)) = b \, \mathbb{P}(X \geq b) \,, \end{aligned}$$

col che la diseguaglianza di Markov è provata. □

Il seguente è un corollario immediato della diseguaglianza di Markov.

Corollario 1.12.1. *Se c è un numero reale, $\varepsilon > 0$ e $n \in \mathbb{N}$, allora*

$$\mathbb{P}\left(|X - c| \geq \varepsilon\right) \leq \frac{\mathbb{E}\left(|X - c|^n\right)}{\varepsilon^n} \,. \tag{1.12.1}$$

Se, inoltre, la v.a. X ammette speranza e varianza finite si ha, per ogni numero reale $k > 0$,

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geq k\sqrt{V(X)}\right) \leq \frac{1}{k^2} \,. \tag{1.12.2}$$

Per la dimostrazione della (1.12.2) basta porre $c = \mathbb{E}(X)$ e $\varepsilon = V(X)$ nella (1.12.1). La (1.12.2) è nota con il nome di *diseguaglianza di Čebyšev*. Si noti che la diseguaglianza (1.12.1) ha senso anche quando non esiste finito il momento di ordine n: infatti se fosse $\mathbb{E}\left(|X - c|^n\right) = +\infty$ la (1.12.1) diverrebbe $\mathbb{P}\left(|X - c| \geq \varepsilon\right) < +\infty$, che è banalmente vera.

1.13 Gli spazî L^p

Gli spazî L^p introdotti da F. Riesz sono importanti in molti rami dell'Analisi, ma sono essenziali anche in Probabilità. Qui di seguito diamo la definizione e le proprietà che sono indispensabili in un primo approccio alla probabilità.

Definizione 1.13.1. Dato uno spazio mensurale $(\Omega, \mathcal{F}, \mu)$, si indica con

$$\mathcal{L}^p = \mathcal{L}^p(\Omega, \mathcal{F}, \mu) \qquad (p \in [1, +\infty[)$$

l'insieme delle funzioni misurabili f per le quali è integrabile $|f|^p$. $\diamond$

Dunque

$$\mathcal{L}^p := \left\{ f \in \mathcal{L}^0 : \int |f|^p \, \mathrm{d}\mu < +\infty \right\} .$$

Dato lo spazio mensurale $(\Omega, \mathcal{F}, \mu)$, due funzioni misurabili, $f, g : \Omega \to \overline{\mathbb{R}}$ si dicono *eguali (μ)–q.o.* se $\mu(f \neq g) = 0$; si scrive $f \simeq_\mu g$. È immediato controllare che l'eguaglianza q.o. è una relazione d'equivalenza. È pertanto naturale considerarne le classi d'equivalenza. Si ha l'abitudine di scrivere f per indicare tanto la funzione f quanto la sua classe d'equivalenza rispetto all'eguaglianza q.o., $\simeq_\mu$.

Si denota con $L^p = L^p(\Omega, \mathcal{F}, \mu)$ lo spazio quoziente di $\mathcal{L}^p$ rispetto alla relazione d'eguaglianza q.o.

$$L^p(\Omega, \mathcal{F}, \mu) := \mathcal{L}^p(\Omega, \mathcal{F}, \mu) / \simeq_\mu \ .$$

In appendice richiamiamo per comodità la definizione ed alcune proprietà delle funzioni convesse che saranno utili in queste lezioni.

In questa sezione supporremo che la misura μ sullo spazio misurabile $(\Omega, \mathcal{F})$ sia una probabilità. I risultati che seguono valgono, con opportune modifiche, anche se μ è una misura finita; se viene meno questa ipotesi non tutti continuano a valere.

Introduce una certa semplificazione nelle formule adottare la notazione, usuale in probabilità,

$$\mathbb{E}(f) := \int f \, \mathrm{d}\mu$$

quando si considera l'integrale di una funzione misurabile f (che sia integrabile).

Teorema 1.13.1. (Jensen). *Sia (a, b) un intervallo di $\mathbb{R}$ (con $-\infty \leq a < b \leq +\infty$), sia $\varphi : (a, b) \to \mathbb{R}$ convessa e sia f una funzione misurabile che assume valori in (a, b). Allora, se f è integrabile, vale la diseguaglianza*

$$\varphi(\mathbb{E}(f)) \leq \mathbb{E}(\varphi \circ f) \ . \tag{1.13.1}$$

Dimostrazione. Si osservi, innanzi tutto, che $\mathbb{E}(f)$ è in (a, b). Supponiamo, infatti che l'estremo a appartenga all'intervallo (a, b); in tal caso, non vi è nulla da provare perché, per ipotesi $f \geq a$ e, dunque, per l'isotonia dell'integrale, anche $\mathbb{E}(f) \geq a$. Supponiamo allora che l'intervallo (a, b) sia aperto in a, cioè che $f > a$. In tal caso, l'isotonia dà direttamente solo $\mathbb{E}(f) \geq a$, come prima, ma non necessariamente $\mathbb{E}(f) > a$, come asserito; se però fosse $\mathbb{E}(f) = a$, si avrebbe che è nullo l'integrale della funzione misurabile positiva $f - a$, e ciò accade solo se la funzione in questione è nulla q.o., vale a dire $f = a$ q.o., contrariamente all'ipotesi. Dunque, $\mathbb{E}(f) > a$. Analogamente si ragiona per l'altro estremo b.

Inoltre, φ è continua, sicché $\varphi \circ f$ è misurabile. Dalla (1.17.4) dell'Appendice segue che

$$(\varphi \circ f)(\omega) \geq \varphi(\mathbb{E}(f)) + \alpha \{f(\omega) - \mathbb{E}(f)\} ,$$

onde, integrando, si ottiene

$$\mathbb{E}(\varphi \circ f) \geq \varphi(\mathbb{E}(f)) + \alpha \{\mathbb{E}(f) - \mathbb{E}(f)\} ,$$

che conclude la dimostrazione. $\square$

Corollario 1.13.1. *Se* $r \geq 1$ *e* $s > 0$

$$\mathbb{E}^r(|f|) \leq \mathbb{E}(|f|^r) \,, \tag{1.13.2}$$
$$\mathbb{E}^r(|f|^s) \leq \mathbb{E}(|f|^{rs}) \,. \tag{1.13.3}$$

Dimostrazione. Si applichi la (1.13.1) alla funzione

$$x \mapsto \varphi(x) := x^r, \quad (x \geq 0, r \geq 1)$$

ottenendo cosí la (1.13.2). Applicando la diseguaglianza (1.13.2) cosí trovata alla funzione misurabile $g := |f|^s$, si ha la (1.13.3). □

Il risultato che segue è conseguenza immediata del precedente. Lo isoliamo per metterne in risalto l'importanza per ciò che seguirà.

Corollario 1.13.2. *Se* $0 < p < q$ *risulta*

$$\mathbb{E}^{1/p}(|f|^p) \leq \mathbb{E}^{1/q}(|f|^q) \,. \tag{1.13.4}$$

Dimostrazione. Nella (1.13.3) si ponga $r = q/p$ e $s = p$. □

La (1.13.4) asserisce che $\mathcal{L}^q(\Omega, \mathcal{F}, \mu)$ è incluso in $\mathcal{L}^p(\Omega, \mathcal{F}, \mu)$ se $0 < p < q$, vale a dire $\mathcal{L}^q \subseteq \mathcal{L}^p$. Analogamente, si ha $L^q \subseteq L^p$. Ponendo $\|\varphi\|_p := \mathbb{E}^{1/p}(|\varphi|^p)$ per una funzione misurabile φ di L^p con $p \geq 1$, la (1.13.4) si scrive $\|\varphi\|_p \leq \|\varphi\|_q \quad (p < q)$.

Lemma 1.13.1. *Siano* x, y, α, β *numeri strettamente positivi con* $\alpha + \beta = 1$. *Siano, inoltre,* $s, t > 0$, $p, q > 1$ *con* $\frac{1}{p} + \frac{1}{q} = 1$. *Allora,*

$$x^\alpha \, y^\beta \leq \alpha \, x + \beta \, y \,, \tag{1.13.5}$$
$$s \, t \leq \frac{s^p}{p} + \frac{t^q}{q} \,. \tag{1.13.6}$$

Dimostrazione. La (1.13.5) equivale alla diseguaglianza $\alpha \, \ln x + \beta \, \ln y \leq \ln(\alpha \, x + \beta \, y)$ che è ovvia perché la funzione logaritmo $\ln :]0, +\infty[\to \mathbb{R}$ è concava (ciò che vuol dire che $-\ln$ è convessa). Per ottenere la (1.13.6) basta poi porre

$$\alpha = \frac{1}{p}, \quad \beta = \frac{1}{q}, \quad x = s^p, \quad y = t^q \,,$$

concludendo cosíla dimostrazione. □

Si osservi che la (1.13.5) contiene, come caso particolare, la ben nota diseguaglianza tra la media geometrica e la media aritmetica di due numeri positivi x e y,

$$\sqrt{xy} \leq \frac{x + y}{2} \,.$$

Teorema 1.13.2. (Hölder). *Siano* $p, q \in]1, +\infty[$ *con*

$$\frac{1}{p} + \frac{1}{q} = 1 \,.$$

Se f *appartiene a* L^p, $f \in L^p$, *e* g *appartiene a* L^q, $g \in L^q$, *allora il prodotto* fg *appartiene a* L^1 *e*

$$\|fg\|_1 \leq \|f\|_p \, \|g\|_q \,, \tag{1.13.7}$$

cioè

$$\mathbb{E}\left(|f\,g|\right) \leq \mathbb{E}^{1/p}\left(|f|^p\right) \mathbb{E}^{1/q}\left(|f|^q\right) \,.$$

Dimostrazione. Non vi è nulla da dimostrare se $\mathbb{E}^{1/p}(|f|^p)\,\mathbb{E}^{1/q}(|g|^q) = 0$. Si supporrà perciò che sia $\mathbb{E}^{1/p}(|f|^p)\,\mathbb{E}^{1/q}(|g|^q) > 0$. Nella (1.13.6) si ponga

$$s = \frac{|f|}{\|f\|_p} \quad \text{e} \quad t = \frac{|g|}{\|g\|_q} \,,$$

per ottenere

$$\frac{|fg|}{\|f\|_p \, \|g\|_q} \leq \frac{1}{p} \, \frac{|f|^p}{\|f\|_p^p} + \frac{1}{q} \, \frac{|g|^q}{\|g\|_q^q} \,,$$

dalla quale si ricava la (1.13.7) integrando. □

Per $p = q = 2$, si ha la diseguaglianza di Schwarz, valida per le funzioni di L^2,

$$\|f\,g\|_1 \leq \|f\|_2 \, \|g\|_2 \qquad \Big(\mathbb{E}(|f\,g|) \leq \mathbb{E}^{1/2}(|f|)\,\mathbb{E}^{1/2}(|g|)\Big) \,.$$

Dalla diseguaglianza di Hölder segue quella di Minkowski.

Teorema 1.13.3. (Minkowski). *Se* f *e* g *appartengono a* L^p *con* $p \geq 1$, *allora*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \,. \tag{1.13.8}$$

Dimostrazione. Non vi è nulla da dimostrare se $p = 1$; si supponga, perciò, che sia $p > 1$. è, in primo luogo, evidente che $|f + g|^p$ è integrabile, o, equivalentemente, che $f + g$ appartiene a L^p, perché

$$|f + g|^p \leq 2^p \, \left(|f|^p \vee |g|^p\right) \,.$$

Dalla diseguaglianza di Hölder scende se, al solito, q è tale che $1/p + 1/q = 1$,

$$\begin{aligned}
\mathbb{E}\left(|f + g|^p\right) &\leq \mathbb{E}\left(|f| \; |f + g|^{p-1}\right) + \mathbb{E}\left(|g| \; |f + g|^{p-1}\right) \\
&\leq \|f\|_p \, \mathbb{E}^{1/q}\left(|f + g|^{pq-q}\right) + \|g\|_p \, \mathbb{E}^{1/q}\left(|f + g|^{pq-q}\right) \\
&= \left(\|f\|_p + \|g\|_p\right) \, \mathbb{E}^{1/q}\left(|f + g|^p\right) \,,
\end{aligned}$$

che equivale alla (1.13.8). □

La diseguaglianza di Minkowski assicura che L^p, con $p \geq 1$ è uno spazio vettoriale e che $\|\cdot\|_p : L^p \to \mathbb{R}_+$ è una norma su L^p (ma non su $\mathcal{L}^p$). Lo spazio L^p è uno spazio normato che risulta essere completo (Teorema di Fischer–Riesz). Ricordiamo che si chiama *spazio di Banach* uno spazio normato che sia completo rispetto alla topologia della norma. Dunque L^p per $p \geq 1$ è uno spazio di Banach, mentre L^2 è uno spazio di Hilbert con il prodotto interno definito da

$$\langle f, g \rangle := \mathbb{E}(f\,\bar{g}) = \int f\,\bar{g}\,\mathrm{d}P \,.$$

1.14 Misure definite da una densità

Siano μ una misura su $(\Omega, \mathcal{F})$ e $f : \Omega \to \overline{\mathbb{R}}_+$ una funzione misurabile positiva. Si definisca una funzione $\nu : \mathcal{F} \to \overline{\mathbb{R}}_+$ mediante

$$\nu(A) := \int_A f \, \mathrm{d}\mu \qquad (A \in \mathcal{F}) \,, \tag{1.14.1}$$

allora ν è una misura su $(\Omega, \mathcal{F})$. Infatti, se $A = \cup_{n \in \mathbb{N}} A_n$ con $A_n \in \mathcal{F}$ per ogni $n \in \mathbb{N}$ e $A_i \cap A_j = \emptyset$ se $i \neq j$, si ha, e si vedano gli esercizî per giustificare taluni dei passaggi,

$$\nu(A) = \int \mathbf{1}_A \, f \, \mathrm{d}\mu = \sum_{n \in \mathbb{N}} \int \mathbf{1}_{A_n} \, f \, \mathrm{d}\mu = \sum_{n \in \mathbb{N}} \int_{A_n} f \, \mathrm{d}\mu = \sum_{n \in \mathbb{N}} \nu(A_n) \,.$$

La misura ν è detta *di densità* f *e di base* μ; solitamente la si indica con $\nu = f \cdot \mu$.

Se la funzione f è integrabile, $f \in L^1(\mu)$, la misura ν è finita. Se poi, f è a valori in $\overline{\mathbb{R}}$ anziché in $\overline{\mathbb{R}}_+$, la (1.14.1) definisce in $\mathcal{F}$ una funzione ν che è ancora numerabilmente additiva, anche se non con valori positivi. Una funzione $\nu : \mathcal{F} \to \overline{\mathbb{R}}$ che sia σ–additiva si dice *misura reale* o *con segno*. In queste lezioni non ci occuperemo delle misure reali.

Teorema 1.14.1. *Se* $g : \Omega \to \overline{\mathbb{R}}$ *è integrabile rispetto alla misura* $\nu = f \cdot \mu$, *è*

$$\int g \, \mathrm{d}\nu = \int g \, f \, \mathrm{d}\mu \,. \tag{1.14.2}$$

Dimostrazione. La (1.14.2) è ovvia quando g è una funzione indicatrice, $g = \mathbf{1}_A$ con $A \in \mathcal{F}$; in tal caso, infatti, la (1.14.2) coincide con la (1.14.1). La (1.14.2) vale allora per le funzioni semplici a causa della linearità dell'integrale e, in virtú della definizione di integrale, per le funzioni positive misurabili. Ma è allora anche vera per una generica funzione integrabile. □

Lemma 1.14.1. *Per una funzione integrabile* $f : \Omega \to \overline{\mathbb{R}}$ *sono equivalenti le affermazioni:*

(a) $f = 0$ *μ–q.o.;*

(b) $\int_A f \, \mathrm{d}\mu = 0$ *per ogni* $A \in \mathcal{F}$.

Dimostrazione. Poiché l'implicazione (a) $\Longrightarrow$ (b) è ovvia, basterà dimostrare che vale l'implicazione inversa (b) $\Longrightarrow$ (a). Preso $A := \{f \geq 0\}$, si ha

$$\int f^+ \, \mathrm{d}\mu = \int \mathbf{1}_A \, f \, \mathrm{d}\mu = 0 \,,$$

sicché $f^+ = 0$ in virtú del Teorema 1.10.1(h). Similmente si procede per f^-. □

Scende dal Lemma 1.14.1 che se ν_1 e ν_2 hanno rispettivamente densità f_1 e f_2 rispetto alla stessa base μ, $\nu_1 = f_1 \cdot \mu$ e $\nu_2 = f_2 \cdot \mu$, allora ν_1 e ν_2 sono eguali se, e solo se, le loro densità sono eguali μ–q.o..

Se μ e ν sono misure sullo stesso spazio misurabile $(\Omega, \mathcal{F})$, si dice che ν è *assolutamente continua* rispetto a μ, e si scrive $\nu \ll \mu$, se si ha $\nu(A) = 0$ per ogni insieme $A \in \mathcal{F}$ tale che $\nu(A) = 0$. È evidente che la (1.14.1) definisce una misura assolutamente continua rispetto a μ; è notevole il fatto che, con questa costruzione, si ottengano *tutte* le misure assolutamente continue rispetto a μ: se ν è assolutamente continua rispetto a μ esiste una densità f tale che $\nu = f \cdot \mu$. è questo il contenuto del teorema di Radon–Nikodym del quale daremo di seguito la dimostrazione. Si tratta di un risultato fondamentale per la Probabilità e la Statistica.

Ricordiamo che si chiama *spazio di Hilbert* uno spazio vettoriale H dotato di *prodotto interno* $\langle \cdot, \cdot \rangle : H \times H \to \mathcal{C}$ e completo rispetto alla topologia della norma indotta dal prodotto interno, $\|x\| := \langle x, x \rangle^{1/2}$ $(x \in H)$. Esempî di spazî di Hilbert sono dati dagli spazî euclidei $\mathbb{R}^n$ e dallo spazio $L^2(\Omega, \mathcal{F}, \mu)$.

Dato uno spazio lineare V, si chiama *duale (topologico)* V' di V l'insieme, che è anch'esso uno spazio lineare, dei funzionali lineari e continui definiti su V, cioè

$$V' := \{F : V \to \mathbb{R} \quad \text{oppure} \quad \mathbb{C},\ F \text{ lineare e continuo}\}.$$

Il seguente, fondamentale, teorema sostanzialmente identifica il duale H' di uno spazio di Hilbert H con H stesso.

Teorema 1.14.2. (Riesz–Fréchet). *Sia H uno spazio di Hilbert; allora, ogni funzionale lineare e continuo F su H può essere rappresentato in un'unica maniera come prodotto interno con un opportuno vettore $z \in H$:*

$$\forall F \in H' \quad \exists z \in H \quad \forall x \in H \qquad F(x) = \langle x, z \rangle; \tag{1.14.3}$$

inoltre $\|F\| = \|z\|$.

Dimostrazione. Se F è identicamente nullo si ponga $z = 0$. Se invece F non è identicamente nullo, il nucleo M di F,

$$M = \ker F := \{x \in H : F(x) = 0\} = F^{-1}(\{0\}),$$

non coincide con tutto H, e, poiché F è lineare e continuo, è un sottospazio proprio chiuso di H. Pertanto il complemento ortogonale $M^\perp$ non si riduce al solo vettore nullo; qui,

$$M^\perp := \{y \in H : \forall x \in M\ \langle x, y \rangle = 0\}.$$

Sia ora $z_0 \in M^\perp$ con z_0 diverso dal vettore nullo; allora,

$$\forall x \in H \qquad F\left(x - \frac{F(x)}{F(z_0)}\, z_0\right) = F(x) - F(x) = 0,$$

cioè

$$x - \frac{F(x)}{F(z_0)}\, z_0$$

appartiene a M per ogni $x \in H$. Perciò,

$$\forall x \in H \qquad 0 = \langle x - \frac{F(x)}{F(z_0)}\, z_0, z_0 \rangle = \langle x, z_0 \rangle - \frac{F(x)}{F(z_0)}\, \langle z_0, z_0 \rangle,$$

che equivale a

$$F(x) = \frac{F(z_0)}{\|z_0\|^2} \langle x, z_0 \rangle ,$$

per ogni $x \in H$. Si ha perciò la (1.14.3) ponendo

$$z := \overline{\frac{F(z_0)}{\|z_0\|^2}}\, z_0 .$$

Si è cosí dimostrata l'esistenza di un vettore con la proprietà dell'asserto. Per provarne l'unicità, si supponga che sia, per ogni $x \in H$,

$$\langle x, z \rangle = \langle x, z' \rangle, \qquad \text{cioè} \qquad \langle x, z - z' \rangle = 0 ;$$

in particolare, per $x = z - z'$, si ha $\|z - z'\| = 0$, sicché $z = z'$.

Infine, segue dalla diseguaglianza di Schwarz che

$$|F(x)| = |\langle x, z \rangle| \leq \|x\| \, \|z\| ,$$

onde

$$\frac{|F(x)|}{\|x\|} \leq \|z\| ,$$

sicché $\|F\| := \sup\{|F(x)| : x \in H, \ \|x\| \leq 1\} \leq \|z\|$.

D'altro canto, l'ultimo asserto $\|F\| = \|z\|$ scende ora dall'eguaglianza $|F(z)| = |\langle z, z \rangle| = \|z\|^2 = \|z\| \, \|z\|$. □

Siamo ora in grado di dimostrare il teorema di Radon–Nikodym.

Teorema 1.14.3. (Radon–Nikodym). *Siano λ e μ due misure sullo spazio misurabile $(\Omega, \mathcal{F})$ e siano λ finita e μ σ–finita. Sono allora equivalenti le seguenti condizioni:*

(a) *λ è assolutamente continua rispetto a μ, $\lambda \ll \mu$;*

(b) *esiste un'unica funzione $h \in L^1(\mu)$ con $h \geq 0$ μ–q.o. tale che $\lambda = h \cdot \mu$, cioè tale che si abbia, per ogni $A \in \mathcal{F}$,*

$$\lambda(A) = \int_A h \, \mathrm{d}\mu .$$

Dimostrazione. L'implicazione (b) $\Longrightarrow$ (a) è ovvia. Basta, perciò, dimostrare l'implicazione (a) $\Longrightarrow$ (b). Si supponga dapprima che μ sia una misura finita. Ponendo $\nu := \lambda + \mu$, si definisce una nuova misura finita su $(\Omega, \mathcal{F})$. Sia ora $f \in L^2(\nu)$; la diseguaglianza di Schwarz dà

$$\begin{aligned} \left| \int f \, \mathrm{d}\lambda \right| &\leq \int |f| \, \mathrm{d}\lambda \leq \int |f| \, \mathrm{d}\nu \\ &\leq \left\{ \int |f|^2 \, \mathrm{d}\nu \right\}^{1/2} \left\{ \int d\nu \right\}^{1/2} = \|f\|_{L^2(\nu)} \sqrt{\nu(\Omega)} . \end{aligned}$$

Perciò l'applicazione $f \mapsto \int f \, d\lambda$ definisce un funzionale lineare e continuo di $L^2(\nu)$. Per il Teorema di Riesz–Fréchet, esiste allora $g \in L^2(\nu)$ tale che, per ogni f di $L^2(\nu)$, risulti

$$\int f \, d\lambda = \int f \, g \, d\nu \, . \tag{1.14.4}$$

Se $A \in \mathcal{F}$ e $\nu(A) > 0$ si ha, ricordando che $\nu = \lambda + \mu$ e che, come conseguenza, $\nu \ll \mu$,

$$0 \leq \lambda(A) = \int \mathbf{1}_A \, d\lambda = \int_A g \, d\nu \quad \text{e} \quad \lambda(A) \leq \nu(A) \, ,$$

donde, per ogni insieme $A \in \mathcal{F}$ con $\nu(A) > 0$,

$$0 \leq \frac{1}{\nu(A)} \int_A g \, d\nu \leq 1 \, . \tag{1.14.5}$$

Quest'ultima relazione implica $0 \leq g \leq 1$ ν–q.o.. Infatti, posto $A := \{g > 1\}$, si ha

$$A = \bigcup_{n \in \mathbb{N}} A_n = \bigcup_{n \in \mathbb{N}} \left\{ g > 1 + \frac{1}{n} \right\} \, .$$

Se esistesse un indice $n \in \mathbb{N}$ tale che $\nu(A_n) > 0$, la (1.14.5) implicherebbe

$$\frac{1}{\nu(A_n)} \int_{A_n} g \, d\nu \geq 1 + \frac{1}{n} \, ,$$

una contraddizione, sicché si ha $\nu(A_n) = 0$ per ogni $n \in \mathbb{N}$, e, di qui, $\nu(g > 1) = 0$. Analogamente si dimostra che $\mu(g < 0) = 0$. Si può perciò assumere, senza perdita di generalità, che sia $0 \leq g(\omega) \leq 1$, per ogni $\omega \in \Omega$. Segue ora dalla (1.14.4) che, per ogni $f \in L^2(\nu)$, si ha

$$\int f \, d\lambda = \int f \, g \, d\nu = \int f \, g \, d\lambda + \int f \, g \, d\mu \, ,$$

cioè

$$\int f(1 - g) \, d\lambda = \int f \, g \, d\mu \, . \tag{1.14.6}$$

Se $S := \{\omega \in \Omega : g(\omega) = 1\}$, la (1.14.6) dà, per $f = \mathbf{1}_S$,

$$\mu(S) = \int \mathbf{1}_S \, g \, d\mu = \int_S g \, d\mu = \int_S (1 - g) \, d\lambda = 0 \, .$$

Per l'assoluta continuità di λ rispetto a μ si ha $\lambda(S) = 0$ e $\nu(S) = 0$. Si può, quindi supporre che sia $g(\omega) < 1$ per ogni $\omega \in \Omega$ e porre, nella (1.14.6), $f = \mathbf{1}_A \sum_{k=0}^n g^k$ $(A \in \mathcal{F})$ per ottenere

$$\int_A (1 - g^{n+1}) \, d\lambda = \int_A g \left(\sum_{k=0}^n g^k \right) d\mu \, .$$

Si applichi, ad entrambi i membri, il Teorema di Beppo Levi 1.10.2,

$$\lambda(A) = \int_A h \, d\mu \, ,$$

ove

$$h := \lim_{n\to\infty} g\left(\sum_{k=0}^{n} g^k\right) = g\left(\sum_{k=0}^{\infty} g^k\right) = \frac{g}{1-g}\,.$$

Si osservi che $h \geq 0$ e che $\int |h|\,\mathrm{d}\mu = \lambda(\Omega) < +\infty$, sicché h è integrabile, $h \in L^1(\mu)$. Per dimostrare l'unicità di h, a meno di equivalenze, si supponga che esista h' in $L^1(\mu)$ tale che per ogni $A \in \mathcal{F}$ sia

$$\lambda(A) = \int_A h'\,\mathrm{d}\mu\,.$$

Ma allora

$$\forall A \in \mathcal{F} \qquad \int_A h'\,\mathrm{d}\mu = \int_A h\,\mathrm{d}\mu\,,$$

sicché $h = h'$ μ–q.o..

Si supponga ora che μ sia σ–finita: in questo caso $\Omega = \cup_{n\in\mathbb{N}} E_n$ con $\mu(E_n) < +\infty$ per ogni $n \in \mathbb{N}$. Non è restrittivo supporre che gli insiemi E_n siano disgiunti; se cosí non è, si può sostituire la successione (E_n) con l'altra (F_n) ove $F_1 = E_1$ e, per $n \geq 2$,

$$F_n = E_n \setminus \left(\bigcup_{k=1}^{n-1} E_k\right)\,.$$

Si considerino ora gli spazî misurabili $(F_n, \mathcal{F}_n)$ ove $\mathcal{F}_n$ è la traccia di $\mathcal{F}$ su $\mathcal{F}_n$, vale a dire $\mathcal{F}_n = \mathcal{F} \cap F_n := \{B \cap F_n : B \in \mathcal{F}\}$ e siano λ_n e μ_n le restrizioni di λ e μ a $(F_n, \mathcal{F}_n)$. Poiché è, evidentemente, $\lambda_n \ll \mu_n$, a tali misure si può applicare la prima parte della dimostrazione; si ottiene cosí una funzione $\varphi_n \in L^1(\mu_n)$ tale che per ogni $E \in \mathcal{F}$ sia

$$\lambda\left(E \bigcap F_n\right) = \lambda\left(E \bigcap F_n\right) = \int_{E\cap F_n} \varphi_n\,\mathrm{d}\mu_n = \int_{E\cap F_n} \varphi_n\,\mathrm{d}\mu\,.$$

Si definisca ora $\varphi := \sum_{n\in\mathbb{N}} \varphi_n\,\mathbf{1}_{F_n}$. Poiché in ogni punto $\omega \in \Omega$ in solo termine è diverso da zero, la funzione φ è integrabile, $\varphi \in L^1(\mu)$ e, per ogni $E \in \mathcal{F}$, risulta

$$\lambda(E) = \int_E \varphi\,\mathrm{d}\mu\,,$$

che conclude la dimostrazione. □

1.15 Misura prodotto

Siano $(\Omega_1, \mathcal{F}_1, \mu_1)$ e $(\Omega_2, \mathcal{F}_2, \mu_2)$ due spazî di misura. La famiglia dei *rettangoli misurabili*

$$\mathcal{F}_1 \times \mathcal{F}_2 := \{A \times B : A \in \mathcal{F}_1, B \in \mathcal{F}_2\}$$

non è una tribú (si veda l'Esercizio 1.28). Si consideri allora la tribú $\mathcal{F}(\mathcal{F}_1 \times \mathcal{F}_2)$ generata da $\mathcal{F}_1 \times \mathcal{F}_2$; questa si dice *tribú prodotto* e la si denota con $\mathcal{F}_1 \otimes \mathcal{F}_2$. Ci si pone il problema di stabilire in quali ipotesi si possa definire in $\mathcal{F}_1 \otimes \mathcal{F}_2$ una misura μ con la proprietà

$$\forall A \in \mathcal{F}_1 \quad \forall B \in \mathcal{F}_2 \qquad \mu(A \times B) = \mu_1(A)\, \mu_2(B)\,. \tag{1.15.1}$$

Ha interesse anche stabilire quando la misura μ, se esiste, sia unica.

Definizione 1.15.1. Per ogni sottoinsieme E di $\Omega_1 \times \Omega_2$ e per ogni punto $x \in \Omega_1$, l'insieme $E_x := \{y \in \Omega_2 : (x, y) \in E\} \subseteq \Omega_2$ si dice *sezione di E in x*. In maniera analoga si definisce la sezione $E_y \subseteq \Omega_1$ di E in $y \in \Omega_2$. $\diamond$

Teorema 1.15.1. *Le sezioni di ogni insieme E della tribú prodotto $\mathcal{F}_1 \otimes \mathcal{F}_2$ sono misurabili, vale a dire, E_x appartiene a $\mathcal{F}_2$ per ogni $x \in \Omega_1$ e E_y appartiene a $\mathcal{F}_1$ per ogni $y \in \Omega_2$.*

Dimostrazione. Posto

$$\mathcal{G} := \{E \subseteq \Omega_1 \times \Omega_2 : \forall\, x \in \Omega_1\ E_x \in \mathcal{F}_2,\ \forall\, y \in \Omega_2\ E_y \in \mathcal{F}_1\}\ ,$$

risulta $\mathcal{F}_1 \times \mathcal{F}_2 \subseteq \mathcal{G}$, cioè $\mathcal{G}$ contiene i rettangoli misurabili; infatti

$$(A \times B)_x = \begin{cases} B\,, & \text{se } x \in A, \\ \emptyset\,, & \text{se } x \notin A, \end{cases}$$

sicché $(A \times B)_x \in \mathcal{F}_2$ per ogni $x \in \Omega_1$. Analogamente si procede per le sezioni rispetto ai punti $y \in \Omega_2$. Per dimostrare il teorema basterà stabilire l'inclusione $\mathcal{G} \supset \mathcal{F}_1 \otimes \mathcal{F}_2$; a tal fine si mostrerà che $\mathcal{G}$ è una tribú. Data l'evidente simmetria tra le sezioni in $x \in \Omega_1$ e quelle in $y \in \Omega_2$, tratteremo solo le prime.

Evidentemente, tanto $\emptyset$ quanto $\Omega_1 \times \Omega_2$ appartengono a $\mathcal{G}$. Se E è in $\mathcal{G}$, allora $(E^c)_x = (E_x)^c \in \mathcal{F}_2$ sicché E^c appartiene a $\mathcal{G}$ e $\mathcal{G}$ è stabile rispetto alla complementazione. Sia, infine, (E_n) una successione di insiemi di $\mathcal{G}$:

$$\left(\bigcup_{n\in\mathbb{N}} E_n\right)_x = \bigcup_{n\in\mathbb{N}} (E_n)_x \in \mathcal{F}_2\,,$$

sicché $\cup_{n\in\mathbb{N}} E_n$ appartiene a $\mathcal{G}$. $\square$

Teorema 1.15.2. *Se le misure μ_1 e μ_2 sono σ–finite, allora, per ogni insieme A di $\mathcal{F}_1 \otimes \mathcal{F}_2$, la funzione $x \mapsto \mu_2(A_x)$ è $\mathcal{F}_1$–misurabile mentre la funzione $y \mapsto \mu_2(A_y)$ è $\mathcal{F}_2$–misurabile.*

Dimostrazione. Sia $\mathcal{C}$ la famiglia di insiemi A della tribú prodotto $\mathcal{F}_1 \otimes \mathcal{F}_2$ per i quali è vero che $x \mapsto \mu_2(A_x)$ è $\mathcal{F}_1$–misurabile. $\mathcal{C}$ include la famiglia $\mathcal{F}_1 \times \mathcal{F}_2$ dei rettangoli misurabili, $\mathcal{F}_1 \times \mathcal{F}_2 \subseteq \mathcal{C}$; infatti se $A = E \times F$ con $E \in \mathcal{F}_1$ e $F \in \mathcal{F}_2$, si ha $\mu_2(A_x) = \mu_2(F)\, \mathbf{1}_E(x)$ che è, ovviamente, $\mathcal{F}_1$–misurabile. Se $A^1, A^2, \dots, A^n$ sono rettangoli misurabili disgiunti, $A^i = E^i \times F^i$ con $E^i \in \mathcal{F}_1$ e $F^i \in \mathcal{F}_2$ $\quad (i = 1, 2, \dots, n)$, allora,

$$\mu_2\left[\left(\bigcup_{i=1}^n A^i\right)_x\right] = \mu_2\left(\bigcup_{i=1}^n A^i_x\right) = \sum_{i=1}^n \mu_2(A^i_x) = \sum_{i=1}^n \mu_2(F^i)\, \mathbf{1}_E(x)\,,$$

che è misurabile rispetto a $\mathcal{F}_1$, in quanto somma di un numero finito di funzioni misurabili. Perciò $\mathcal{C}$ contiene l'anello $\mathcal{R}$ delle unioni finite e disgiunte di rettangoli misurabili. Ma $\mathcal{C}$ è una classe monotona perché se A_n è in $\mathcal{C}$ per ogni $n \in \mathbb{N}$ e se $A_n \uparrow A$, allora $(A_n)_x \uparrow A_x$ onde $\mu_2\left[(A_n)_x\right] \to \mu_2(A_x)$: perciò $\mu_2(A_x)$, come limite di funzioni misurabili, è misurabile. Si ricorre allo stesso ragionamento se (A_n) è decrescente, anziché crescente. Si può quindi concludere che $\mathcal{C} = \mathcal{F}_1 \otimes \mathcal{F}_2$. □

Teorema 1.15.3. *Se le misure μ_1 e μ_2, su $(\Omega_1, \mathcal{F}_1)$ e $(\Omega_2, \mathcal{F}_2)$ rispettivamente, sono σ–finite, esiste un'unica misura σ–finita μ definita in $\mathcal{F}_1 \otimes \mathcal{F}_2$ che verifichi la* (1.15.1)*; essa si dice* misura prodotto *e si denota con $\mu = \mu_1 \otimes \mu_2$.*

Dimostrazione. Si definisca $\mu : \mathcal{F}_1 \otimes \mathcal{F}_2 \to \overline{\mathbb{R}}_+$ mediante

$$\mu(A) := \int_{\Omega_1} \mu_2(A_x)\, \mathrm{d}\mu_1(x)\,; \tag{1.15.2}$$

μ sarà la misura cercata. La definizione (1.15.2) è lecita in virtú del teorema precedente.

Sia (A^n) una successione di insiemi disgiunti di $\mathcal{F}_1 \otimes \mathcal{F}_2$. Poiché

$$\left[\bigcup_{n\in\mathbb{N}} A^n\right]_x = \bigcup_{n\in\mathbb{N}} A^n_x\,,$$

e poiché anche le sezioni sono disgiunte, si ha

$$\begin{aligned}\mu\left(\bigcup_{n\in\mathbb{N}} A^n\right) &= \int_{\Omega_1} \mu_2\left[\left(\bigcup_{n\in\mathbb{N}} A^n\right)_x\right] \mathrm{d}\mu_1(x) = \int_{\Omega_1} \sum_{n\in\mathbb{N}} \mu_2(A^n_x)\, \mathrm{d}\mu_1(x) \\ &= \sum_{n\in\mathbb{N}} \int_{\Omega_1} \mu_2(A^n_x)\, \mathrm{d}\mu_1(x) = \sum_{n\in\mathbb{N}} \mu(A^n)\,,\end{aligned}$$

ciò che prova che μ è una misura. Per verificare che μ soddisfaccia alla (1.15.1) si ricordi che $\mu_2\left[(A \times B)_x\right] = \mu_2(B)\, \mathbf{1}_A(x)$ se A appartiene a $\mathcal{F}_1$ e B a $\mathcal{F}_2$; perciò

$$\mu(A \times B) = \int \mu_2(B)\, \mathbf{1}_A(x)\, \mathrm{d}\mu_1(x) = \mu_2(B) \int \mathbf{1}_A\, \mathrm{d}\mu = \mu_1(A)\, \mu_2(B)\,.$$

Rimane dunque da controllare che μ sia σ–finita. Siccome tanto μ_1 quanto μ_2 sono σ–finite, esistono successioni (E_n) di insiemi di $\mathcal{F}_1$ e (F_r) di insiemi di $\mathcal{F}_2$ tali che

$$\bigcup_{n\in\mathbb{N}} E_n = \Omega_1\,, \quad \bigcup_{r\in\mathbb{N}} F_r = \Omega_2\,, \quad \text{e} \quad \mu_1(E_n) < +\infty\,,\ \mu_2(F_r) < +\infty$$

per ogni n e per ogni r di $\mathbb{N}$. Perciò

$$\Omega_1 \times \Omega_2 = \bigcup_{n,r\in\mathbb{N}} E_n \times F_r$$

e

$$\forall\, n, r \in \mathbb{N} \qquad \mu(E_n \times F_r) = \mu_1(E_n)\, \mu_2(F_r) < +\infty\,.$$

Si noti che, se $\mathcal{R}$ è l'anello delle unioni finite e disgiunte di rettangoli misurabili, la restrizione ν di μ a $\mathcal{R}$ è una misura tale che $\nu(E \times F) = \mu_1(E)\, \mu_2(F)$ per ogni $E \in \mathcal{F}_1$ e per ogni $F \in \mathcal{F}_2$; per il teorema di Carathéodory è unica l'estensione di ν a $\mathcal{F}_1 \otimes \mathcal{F}_2$ che è la tribú generata da $\mathcal{R}$. Ciò stabilisce l'unicità di μ. □

Teorema 1.15.4. (Fubini–Tonelli). *Siano dati due spazî mensurali σ–finiti*

$$(\Omega_1, \mathcal{F}_1, \mu_1) \qquad e \qquad (\Omega_2, \mathcal{F}_2, \mu_2).$$

(a) *Sia $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}_+$ una funzione positiva misurabile rispetto alla tribú prodotto $\mathcal{F} := \mathcal{F}_1 \otimes \mathcal{F}_2$; allora, per ogni ogni $x \in \Omega_1$, la funzione $y \mapsto f(x,y)$ è $\mathcal{F}_2$–misurabile, la funzione $x \mapsto \int f(x,y)\,\mathrm{d}\mu_2(y)$ è $\mathcal{F}_1$–misurabile, ed inoltre si ha, se μ è la misura prodotto $\mu = \mu_1 \otimes \mu_2$,*

$$\int_{\Omega_1\times\Omega_2} f\,\mathrm{d}\mu = \int_{\Omega_1} \mathrm{d}\mu_1(x) \int_{\Omega_2} f(x,y)\,\mathrm{d}\mu_2(y). \tag{1.15.3}$$

(b) *Se $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ è $\mathcal{F}$–misurabile ed integrabile rispetto alla misura prodotto μ, allora la funzione $y \mapsto f(x,y)$ è integrabile rispetto a μ_2 per quasi ogni (rispetto alla misura μ_1) $x \in \Omega_1$, la funzione definita μ_1–quasi ovunque da $x \mapsto \int f(x,y)\,\mathrm{d}\mu_2(y)$ è integrabile rispetto a μ_1 e vale ancora la (1.15.3).*

Dimostrazione. Sia $f = \mathbf{1}_A$ con $A \in \mathcal{F} = \mathcal{F} \otimes \mathcal{F}$, allora ogni sezione A_x è in $\mathcal{F}_2$ e la funzione

$$y \mapsto \mathbf{1}_A(x,y) = \mathbf{1}_{A_x}(y)$$

è $\mathcal{F}_2$–misurabile per ogni $x \in \Omega_1$. Inoltre, ricordando la (1.15.2), si ha

$$\int_{\Omega_1} \mathrm{d}\mu_1(x) \int_{\Omega_2} \mathbf{1}_A(x,y)\,\mathrm{d}\mu_2(y) = \int_{\Omega_1} \mu_2(A_x)\,\mathrm{d}\mu_1(x) = \mu(A) = \int_{\Omega_1\times\Omega_2} \mathbf{1}_A\,\mathrm{d}\mu\,;$$

la (1.15.3) vale perciò per le funzioni indicatrici di insiemi di $\mathcal{F}_1 \otimes \mathcal{F}_2$. Per la linearità degli integrali, l'asserto vale anche per le funzioni $\mathcal{F}$–semplici. Sia ora (s_n) una successione di funzioni positive $\mathcal{F}$–semplici che tende crescendo alla funzione misurabile positiva f. L'applicazione $y \mapsto f(x,y)$ è $\mathcal{F}_2$–misurabile perché $f(x,y) = \sup_{n\in\mathbb{N}} s_n(x,y)$. Analogamente la funzione

$$x \mapsto \int_{\Omega_2} f(x,y)\,\mathrm{d}\mu_2(y) = \sup_{n\in\mathbb{N}} \int_{\Omega_2} s_n(x,y)\,\mathrm{d}\mu_2(y)$$

risulta $\mathcal{F}_1$–misurabile; il teorema di Beppo Levi dà ora

$$\begin{aligned}
\int_{\Omega_1\times\Omega_2} f\,\mathrm{d}\mu &= \sup_{n\in\mathbb{N}} \int_{\Omega_1\times\Omega_2} s_n\,\mathrm{d}\mu = \sup_{n\in\mathbb{N}} \int_{\Omega_1} \mathrm{d}\mu_1(x) \int_{\Omega_2} s_n(x,y)\,\mathrm{d}\mu_2(y) \\
&= \int_{\Omega_1} \mathrm{d}\mu_1(x) \left\{ \sup_{n\in\mathbb{N}} \int_{\Omega_2} s_n(x,y)\,\mathrm{d}\mu_2(y) \right\} \\
&= \int_{\Omega_1} \mathrm{d}\mu_1(x) \int_{\Omega_2} \sup_{n\in\mathbb{N}} s_n(x,y)\,\mathrm{d}\mu_2(y) = \int_{\Omega_1} \mathrm{d}\mu_1(x) \int_{\Omega_2} f(x,y)\,\mathrm{d}\mu_2(y).
\end{aligned}$$

(b) Poiché f è integrabile, tali sono anche f^+ e f^-. Perciò la funzione

$$x \mapsto \int f^+(x,y)\,\mathrm{d}\mu_2(y)$$

è finita per quasi ogni $x \in \Omega_1$ (rispetto a μ_1) per il Teorema 1.10.1(c); similmente si argomenta per f^-. Perciò la differenza

$$\int\limits_{\Omega_2} f(x,y)\,\mathrm{d}\mu_2(y) = \int\limits_{\Omega_2} f^+(x,y)\,\mathrm{d}\mu_2(y) - \int\limits_{\Omega_2} f^-(x,y)\,\mathrm{d}\mu_2(y)$$

è definita per quasi ogni $x \in \Omega_1$. Basta ora applicare la parte (a) separatamente alle funzioni positive f^+ e f^-. □

Naturalmente, nelle stesse ipotesi, vale anche la relazione simmetrica

$$\int\limits_{\Omega_1\times\Omega_2} f\,\mathrm{d}\mu = \int\limits_{\Omega_2} d\mu_2(y) \int\limits_{\Omega_1} f(x,y)\,\mathrm{d}\mu_1(x)\,.$$

Si osservi che, per poter scrivere la (1.15.3), occorre verificare che la funzione f sia integrabile rispetto alla misura prodotto. Si vedano, a tal fine, gli esercizî.

Il procedimento sin qui esposto si adatta, senza particolari difficoltà, al prodotto di un numero finito di misure, cioè alla misura prodotto sullo spazio misurabile

$$(\Omega_1 \times \Omega_2 \times \cdots \times \Omega_n, \mathcal{F}_1 \otimes \mathcal{F}_2 \otimes \cdots \otimes \mathcal{F}_n)\,.$$

Si presenta, invece, qualche difficoltà nel definire la misura prodotto nel caso di un prodotto numerabile (o di cardinalità maggiore) di spazî mensurali a meno che le misure non siano tutte di probabilità, caso che tratteremo di seguito.

Definizione 1.15.2. Sia $((\Omega_n, \mathcal{F}_n))_{n\in\mathbb{N}}$ una successione di spazî misurabili e sia $\Omega = \prod_{n\in\mathbb{N}} \Omega_n$ l'insieme di tutte le successioni

$$(\omega_1, \omega_2, \dots, \omega_n, \dots)$$

tali che ω_n appartenga a Ω_n per ogni $n \in \mathbb{N}$. Se E_j è in $\mathcal{F}_j$ per $j = 1, 2, \dots, n$, l'insieme $E \subseteq \Omega$ definito da

$$E = E_1 \times E_2 \times \cdots \times E_n \times \prod_{j=n+1}^{\infty} \Omega_j \tag{1.15.4}$$

si dice *cilindro di base n–dimensionale* $E_1 \times E_2 \times \cdots \times E_n$,

Un cilindro con base n–dimensionale come quello della (1.15.4) può essere pensato come avente base di dimensione maggiore; per esempio, il cilindro E definito dalla (1.15.4) può essere pensato come un cilindro di base $(n+1)$–dī-mensionale $E_1 \times E_2 \times \cdots \times E_n \times \Omega_{n+1}$.

Lemma 1.15.1. *La famiglia $\mathcal{C}$ dei cilindri è un semi–anello di sottoinsiemi del prodotto cartesiano $\Omega = \prod_{n\in\mathbb{N}} \Omega_n$.*

Dimostrazione. Si considerino i due cilindri

$$E = E_1 \times E_2 \times \cdots \times E_n \times \prod_{j=n+1}^{\infty} \Omega_j\,,$$

$$F = F_1 \times F_2 \times \cdots \times F_k \times \prod_{j=k+1}^{\infty} \Omega_j\,,$$

ove $E_j \in \mathcal{F}_j$ per $j = 1, 2, \dots, n$ e $F_i \in \mathcal{F}_i$ per $i = 1, 2, \dots, k$. Non è restrittivo supporre, per fissare le idee, che sia $k < n$; si può quindi supporre che E e F abbiano entrambi la base di dimensione n e che sia $F_j = \Omega_j$ se $j = k+1, \dots, n$.

Ora
$$E \cap F = \prod_{j=1}^{n} (E_j \cap F_j) \times \prod_{j=n+1}^{\infty} \Omega_j \,,$$
che è ancora un cilindro. D'altro canto,
$$E \setminus F = E \cap F^c = \prod_{j=1}^{n} (E_j \cap F_j^c) \times \prod_{j=n+1}^{\infty} \Omega_j \,,$$
che è anch'esso un cilindro. □

Teorema 1.15.5. *Se $((\Omega_n, \mathcal{F}_n, \mathbb{P}_n))_{n\in\mathbb{N}}$ è una successione di spazî di probabilità, esiste un'unica misura di probabilità sullo spazio misurabile $(\Omega, \mathcal{F})$, ove Ω è il prodotto cartesiano $\prod_{n\in\mathbb{N}} \Omega_n$ e $\mathcal{F}$ è la tribú, $\mathcal{F} = \mathcal{F}(\mathcal{C})$, generata dai cilindri di $\prod_{n\in\mathbb{N}} \Omega_n$, tale che sul cilindro*
$$E = E_1 \times E_2 \times \dots \times E_n \times \prod_{j=n+1}^{\infty} \Omega_j$$
assuma il valore
$$\mathbb{P}\left(\prod_{j=1}^{n} E_j \times \prod_{j=n+1}^{\infty} \Omega_j \right) = \mathbb{P}_1(E_1)\, \mathbb{P}_2(E_2) \dots \mathbb{P}_n(E_n) \,.$$

Dimostrazione. Si definisca $\mathbb{P}$ sul semi–anello $\mathcal{C}$ mediante
$$\mathbb{P}(E) := \mathbb{P}_1(E_1)\, \mathbb{P}_2(E_2) \dots \mathbb{P}_n(E_n) \,,$$
se E è il cilindro (1.15.4).

Per verificare che la funzione d'insieme P cosí definita su $\mathcal{C}$ è finitamente additiva, si supponga che F sia un cilindro che si può scrivere come unione disgiunta di altri due cilindri, $F = F_1 \cup F_2$. Esiste allora un numero naturale N tale che F, F_1 e F_2 possano essere scritti nella forma
$$E_1 \times E_2 \times \dots \times E_N \times \prod_{j=N+1}^{\infty} \Omega_j \,.$$

Si può ora applicare il Teorema 1.15.3 per ottenere $\mathbb{P}(F) = \mathbb{P}(F_1) + \mathbb{P}(F_2)$ e quindi estendere $\mathbb{P}$ all'anello $\mathcal{R}$ generato da $\mathcal{C}$. Per poter applicare il teorema di estensione di Carathéodory occorre mostrare che $\mathbb{P}$ è una misura su $\mathcal{R}$. A tal fine, basta provare che, per ogni successione decrescente (A_n) di insiemi di $\mathcal{R}$ tale che $\mathbb{P}(A_n) \to \varepsilon > 0$, si ha $\cap_n A_n \neq \emptyset$. Sia (A_n) una successione siffatta. Si ponga $H_n := \prod_{j=n+1}^{\infty} \Omega_j$. Operando come abbiamo appena fatto, si può definire una funzione d'insieme prodotto $\nu^{(n)}$ definita sulla classe $\mathcal{R}^{(n)}$ delle unioni finite di cilindri di $\mathcal{C}(\mathcal{F}_{n+1}, \mathcal{F}_{n+2}, \dots)$. Cosí, per ogni $n \in \mathbb{N}$ si può costruire $\mathbb{P}$ sul semi–anello $\mathcal{C}$ dei cilindri come prodotto delle $n+1$ misure

$$\mathbb{P}_1, \mathbb{P}_2, \dots, \mathbb{P}_n \text{ e } \nu^{(n)} .$$

Sia $A_n(\omega_1)$ la sezione di A_n in $\omega_1 \in \Omega_1$,

$$A_n(\omega_1) := \{\omega \in H_1 : (\omega_1, \omega) \in A_n\} ;$$

per ogni $\omega_1 \in \Omega_1$ si ha $A_n(\omega_1) \in \mathcal{R}^{(1)}$. Posto

$$B_n := \left\{\omega_1 : \nu^{(1)}\left(A_n(\omega_1)\right) > \frac{1}{2}\,\varepsilon\right\} ,$$

B_n è un sottoinsieme di Ω_1 e, di piú, è l'unione finita di insiemi di $\mathcal{F}_1$, sicché è esso stesso in $\mathcal{F}_1$. Dalla decomposizione

$$A_n = (A_n \cap (B_n \times H_1)) \bigcup (A_n \cap (B_n^c \times H_1))$$

segue

$$\varepsilon \leq \mathbb{P}(A_n) \leq \mathbb{P}_1(B_n) + \frac{1}{2}\,\varepsilon\, [1 - \mathbb{P}_1(B_n)]$$

e di qui

$$\mathbb{P}_1(B_n) > \frac{1}{2}\,\varepsilon$$

per ogni $n \in \mathbb{N}$. Poiché (A_n) è decrescente, tale è anche la successione $\{B_n\}$. Perciò

$$\mathbb{P}_1\left(\cap_{n\in\mathbb{N}} B_n\right) \geq \frac{1}{2}\,\varepsilon .$$

Ora $\mathbb{P}_1$ è una misura; di conseguenza, esiste $\omega_1 \in \omega_1$ tale che, per ogni $n \in \mathbb{N}$,

$$\nu^{(1)}\left(A_n(\omega_1)\right) > \frac{1}{2} .$$

Si fissi ω_1 e si ripeta il ragionamento per la successione $(A_n(\omega_1))$ di sottoinsiemi di H_1. Si ha cosí un punto $\omega_2 \in \Omega_2$ tale che, per ogni $n \in \mathbb{N}$,

$$\nu^{(2)}\left(A_n(\omega_1, \omega_2)\right) > \frac{1}{4}\,\varepsilon .$$

Per induzione si ottiene un punto $(\omega_1, \omega_2, \dots)$ di $\prod_{n\in\mathbb{N}} \Omega_n$, tale che, per ogni coppia di numeri naturali k e n, sia

$$A_n(\omega_1, \omega_2, \dots, \omega_k) \neq \emptyset .$$

Ricordando però la definizione di A_n, che è in $\mathcal{R}$, si vede che $(\omega_1, \omega_2, \dots)$ appartiene ad A_n per ogni n, sicché

$$\cap_{n\in\mathbb{N}} A_n \neq \emptyset ,$$

e, dunque, $\mathbb{P}$ è una misura. □

1.16 Completamento di misure

La necessità di tenere conto dei sottoinsiemi di insiemi di misura nulla induce a considerare il cosiddetto completamento di una misura.

Teorema 1.16.1. *Sia μ una misura su $(\Omega, \mathcal{F})$ e si definisca con $\mathcal{E}(\mu)$ la classe dei sottoinsiemi degli insiemi di μ–misura nulla*

$$\mathcal{E}(\mu) := \{E \subseteq \Omega : \exists N \in \mathcal{F},\ E \subseteq N,\ \mu(N) = 0\} ,$$

e si ponga

$$\mathcal{F}(\mu) := \{A \cup E : A \in \mathcal{F},\ E \in \mathcal{E}(\mu)\} .$$

Allora:

(a) *$\mathcal{F}(\mu)$ è una tribú;*

(b) *la funzione $\overline{\mu} : \mathcal{F}(\mu) \to \overline{\mathbb{R}}_+$ definita da $\overline{\mu}(A \cup E) := \mu(A)$ è una misura su $(\Omega, \mathcal{F}(\mu))$;*

(c) *$\overline{\mu}$ è* completa*, nel senso che, se $F \in \mathcal{F}(\mu)$ e $\overline{\mu}(F) = 0$, allora appartiene a $\mathcal{F}(\mu)$ ogni insieme $B \subseteq F$ (e necessariamente è $\mu(B) = 0$).*

Dimostrazione. (a) $\mathcal{F}(\mu)$ è stabile rispetto alla complementazione; infatti, sia $F \in \mathcal{F}(\mu)$, e $F = A \cup E$, con $A \in \mathcal{F}$, $E \in \mathcal{E}(\mu)$, $E \subseteq N$, $N \in \mathcal{F}$, $\mu(N) = 0$; allora:

$$\begin{aligned} F^c = A^c \cap E^c &= [(A^c \cap N) \cup (A^c \cap N^c)] \cap E^c \\ &= (A^c \cap N \cap E^c) \cup (A^c \cap N^c \cap E^c) \\ &= (A^c \cap N \cap E^c) \cup (A^c \cap N^c) . \end{aligned}$$

Ora $A^c \cap N^c \in \mathcal{F}$ e $A^c \cap N \cap E^c \subseteq N$, sicché $F^c \in \mathcal{F}(\mu)$. Se, poi, per ogni $n \in \mathbb{N}$, è $F_n \in \mathcal{F}(\mu)$, con $F_n = A_n \cup E_n$, $A_n \in \mathcal{F}$, $E_n \in \mathcal{E}(\mu)$, $E \subseteq N_n$, $N_n \in \mathcal{F}$, $\mu(N_n) = 0$, si ha

$$\cup_{n\in\mathbb{N}} F_n = (\cup_{n\in\mathbb{N}} A_n) \bigcup (\cup_{n\in\mathbb{N}} E_n) ,$$

ove

$$\cup_{n\in\mathbb{N}} A_n \in \mathcal{F} \quad \text{e} \quad \cup_{n\in\mathbb{N}} E_n \subseteq \cup N_n \in \mathcal{F}$$

e

$$\mu(\cup_{n\in\mathbb{N}} N_n) \leq \sum_{n\in\mathbb{N}} \mu(N_n) = 0 ;$$

dunque $\cup_{n\in\mathbb{N}} F_n$ è in $\mathcal{F}$ e $\mathcal{F}(\mu)$ è stabile rispetto all'unione numerabile.

(b) Basta controllare che quella data sia una buona definizione. Siano, allora, $A_1 \cup E_1 = A_2 \cup E_2$ due diverse rappresentazioni dell'insieme $F \in \mathcal{F}(\mu)$:

$$F = A_1 \cup E_1 = A_2 \cup E_2 ,$$

con $A_i \in \mathcal{F}$, $E_i \in \mathcal{E}(\mu)$, $E_i \subseteq N_i \in \mathcal{F}$, $\mu(N_i) = 0 \quad (i = 1, 2)$. Si ha $A_1 \setminus A_2 \subseteq E_2$ e quindi

$$\mu(A_1) = \mu(A_1 \cap A_2) + \mu(A_1 \setminus A_2) = \mu(A_1 \cap A_2) \leq \mu(A_2) ;$$

simmetricamente si ha $\mu(A_2) \leq \mu(A_1)$ e dunque $\mu(A_1) = \mu(A_2)$.

(c) Sia $B \subseteq F \in \mathcal{F}(\mu)$ e $\overline{\mu}(F) = 0$. Scritto F nella forma $F = A \cup E$ con $A \in \mathcal{F}$, $E \in \mathcal{E}(\mu)$, $E \subseteq N \in \mathcal{F}$ e $\mu(N) = 0$, si ha $\mu(A) = 0$, sicché $B \subseteq A \cup N \in \mathcal{F}$ con $\mu(A \cup N) = 0$ onde $B \in \mathcal{E}(\mu)$ e dunque anche B appartiene a $\mathcal{F}(\mu)$. □

Si dice che lo spazio di misura $(\Omega, \mathcal{F}(\mu), \overline{\mu})$ è il *completamento* di $(\Omega, \mathcal{F}, \mu)$ e che la tribú $\mathcal{F}(\mu)$ è il *completamento della tribú $\mathcal{F}$ rispetto alla misura μ.*

Teorema 1.16.2. *Se $(\Omega, \mathcal{F}, \mu)$ è completo e $f = g$ q.o, allora f è misurabile se, e solo se, g è misurabile.*

Dimostrazione. Per ogni $c \in \mathbb{R}$ si ha $\{f < c\} \Delta \{g < c\} \subseteq \{f \neq g\}$; giacché, per ipotesi, $\mu(f \neq g) = 0$, gli insiemi $\{f < c\}$ e $\{g < c\}$ differiscono per un sottoinsieme di misura nulla che è in $\mathcal{F}$, perché $\mathcal{F}$ è completa. Perciò

$$\{f < c\} \in \mathcal{F} \quad \Longrightarrow \quad \{g < c\} \in \mathcal{F}\,,$$

che dà l'asserto. □

Il completamento della tribú di Borel è la tribú $\mathcal{L}(\mathbb{R})$ degli insiemi *misurabili secondo Lebesgue*, mentre il completamento della misura di Borel su $\mathcal{B}(\mathbb{R})$ si dice *misura di Lebesgue*. Le misure di Borel e di Lebesgue dunque coincidono su tutti gli insiemi boreliani, in particolare, sugli intervalli.

1.17 Appendice

In questa appendice rivedremo brevemente le proprietà delle funzioni convesse che servono per lo studio della probabilità.

Definizione 1.17.1. Sia $I =]a, b[$ con $-\infty \leq a < b \leq +\infty$ un intervallo aperto di $\mathbb{R}$. Una funzione $\varphi : I \to \mathbb{R}$ che, per ogni coppia di punti x e y di I, e per ogni $\alpha \in [0, 1]$ soddisfaccia alla diseguaglianza

$$\varphi(\alpha\, x + (1-\alpha)\, y) \leq \alpha\, \varphi(x) + (1-\alpha)\, \varphi(y)\,. \tag{1.17.1}$$

si dice *convessa*. Si dice, invece, che φ è *concava* se è convessa la funzione $-\varphi$. ◇

Si controlla facilmente per induzione che vale la seguente generalizzazione della (1.17.1); per ogni scelta di n punti x_j $(j = 1, 2, \dots, n)$ in I e per ogni scelta di n numeri positivi α_j $(j = 1, 2, \dots, n)$, $\alpha_j \geq 0$, con $\sum_{j=1}^n \alpha_j = 1$, si ha

$$\varphi\left(\sum_{j=1}^n \alpha_j\, x_j\right) \leq \sum_{j=1}^n \alpha_j\, \varphi(x_j)\,.$$

Si considerino ora $x \in I$ e h_1 e h_2 sufficientemente piccoli affinché $x + h_2$ e $x - h_2$ appartengano a I e tali che $0 < h_1 < h_2$; scelto $\alpha = (h_2 - h_1)/h_2$ nella (1.17.1), si ha

$$\frac{h_2 - h_1}{h_2}\, \varphi(x) + \frac{h_1}{h_2}\, \varphi(x \pm h_2) \geq \varphi\left(\frac{h_2 - h_1}{h_2}\, x + \frac{h_1}{h_2}\,(x \pm h_2)\right) = \varphi(x \pm h_1)\,.$$

Di qui si ricavano facilmente le diseguaglianze

$$\begin{aligned} \frac{\varphi(x+h_1) - \varphi(x)}{h_1} &\leq \frac{\varphi(x+h_2) - \varphi(x)}{h_2}\,, \\ \frac{\varphi(x-h_2) - \varphi(x)}{-h_2} &\leq \frac{\varphi(x-h_1) - \varphi(x)}{-h_1}\,. \end{aligned}$$

Queste ultime due relazioni mostrano che il rapporto incrementale destro e quello sinistro di φ sono rispettivamente crescente e decrescente, sicché esistono le derivate sinistra $D^-\varphi(x)$ e destra $D^+\varphi(x)$ di φ in $x \in I$. Inoltre per $h > 0$ e $h' > 0$ scende dalla (1.17.1)

$$\frac{\varphi(x-h) - \varphi(x)}{-h} \leq \frac{\varphi(x+h') - \varphi(x)}{h'} ,$$

che mostra che le derivate $D^-\varphi(x)$ e $D^+\varphi(x)$ sono finite. Di conseguenza ogni funzione convessa è continua in I. Segue da quanto detto sopra che si può scrivere

$$D^+\varphi(x) = \inf_{y>x} \frac{\varphi(y) - \varphi(x)}{y - x} , \tag{1.17.2}$$

$$D^-\varphi(x) = \sup_{y<x} \frac{\varphi(x) - \varphi(y)}{x - y} . \tag{1.17.3}$$

Perciò, se $x_1 < x_2$ è

$$\begin{aligned} D^-\varphi(x_1) \leq D^+\varphi(x_1) &\leq \frac{\varphi(x_2) - \varphi(x_1)}{x_2 - x_1} \\ &= \frac{\varphi(x_1) - \varphi(x_2)}{x_1 - x_1} \leq D^-\varphi(x_2) \leq D^+\varphi(x_2) ; \end{aligned}$$

ne consegue che le funzioni

$$t \mapsto D^+\varphi(t) \quad \text{e} \quad t \mapsto D^-\varphi(t)$$

sono entrambe isotone, e, pertanto, ciascuna di esse ammette al piú un'infinità numerabile di punti di discontinuità.

Scende dalla (1.17.2) e dalla (1.17.3) che valgono le diseguaglianze

$$\begin{aligned} \varphi(y) &\geq \varphi(x) + D^+\varphi(x)\,(y - x), \qquad \text{se } y > x, \\ \varphi(y) &\geq \varphi(x) + D^-\varphi(x)\,(y - x), \qquad \text{se } y < x, \end{aligned}$$

onde, per ogni $y \in I$ e per ogni $\alpha_x \in [D^-\varphi(x), D^+\varphi(x)]$,

$$\varphi(y) \geq \varphi(x) + \alpha_x\,(y - x) . \tag{1.17.4}$$

In un punto x nel quale la funzione φ sia derivabile, $D^+\varphi(x) = D^-\varphi(x) = D\varphi(x)$, si ottiene la relazione elementare ben nota

$$\varphi(y) \geq \varphi(x) + D\varphi(x)\,(y - x) ,$$

che esprime il fatto geometrico che il grafico di φ giace sopra la retta tangente al grafico in x.

Posto, per x e s in I, $g_s(x) := \varphi(s) + \alpha_s\,(x - s)$, si ha, evidentemente, $\varphi(x) \geq g_s(x)$ e

$$\varphi(x) = \sup\{g_s(x) : s \in I\} .$$

Si noti che, dal punto di vista geometrico, g_s è una retta che si dice *linea di supporto di* φ. In probabilità è utile il seguente teorema, detto *della linea di supporto*.

Teorema 1.17.1. *Data una funzione convessa φ nell'intervallo aperto I esistono due successioni (a_n) e (b_n) di numeri reali tali che, per ogni $x \in I$, valga la rappresentazione*

$$\varphi(x) = \sup_{n\in\mathbb{N}}\{a_n\, x + b_n\}\,,$$

sicché φ è l'estremo superiore di una famiglia numerabile di linee di supporto.

Dimostrazione. Si consideri l'insieme $\mathbb{Q} \cap I$ dei numeri razionali di I; per ogni numero $q \in \mathbb{Q} \cap I$ è

$$\varphi(x) \geq g_q(x) = \varphi(q) + \alpha_q\,(x-q)\,,$$

ove

$$D^-\varphi(q) \leq \alpha_q \leq D^+\varphi(q)\,.$$

Poiché $D^+\varphi$ e $D^-\varphi$ sono limitate in ogni intervallo chiuso contenuto in I, la successione (α_q) è limitata. Pertanto esiste una successione (α_{q_n}) estratta da (α_q) che converge a $\alpha \in [D^-\varphi, D^+\varphi]$ e

$$\lim_{n\to\mathbb{N}} g_{q_n}(x) = \varphi(x)\,.$$

Poiché $\varphi(x) \geq g_{q_n}(x)$ per ogni $n \in \mathbb{N}$ si ha

$$\varphi(x) = \sup_{n\in\mathbb{N}}\{\varphi(q_n) + \alpha_{q_n}\,(x-q_n)\}\,,$$

che prova l'asserto quando si sia posto

$$a_n := \alpha_{q_n} \qquad \text{e} \qquad b_n := \varphi(q_n) - \alpha_{q_n} q_n$$

per ogni $n \in \mathbb{N}$. □

1.18 Note al Capitolo 1

La teoria della misura come è qui succintamente presentata è scritta nello spirito del libro (Halmos, 1950). L'approccio di questo libro non è l'unico possibile; un altro, a prima vista completamente differente, si trova in (Bourbaki, 1965) e nei libri dei seguaci di questo "autore". Ritengo che il primo sia preferibile per gli usi del calcolo delle probabilità in questa scelta confortato dal notare che è anche quella di un maestro delle probabilità come Doob; si veda (Doob, 1994).

Quello che presento è il bagaglio minimo che deve essere a disposizione di uno studioso di probabilità; il solo argomento essenziale che ometto qui è l'equiintegrabilità, la cui introduzione rimando al Capitolo sulle martingale. Naturalmente esulano dallo scopo di queste lezioni lo studio delle misure reali, vale a dire delle funzioni $\mu : \mathcal{F} \to \mathbb{R}$ che siano σ–additive, e quello degli spazî di misura infinita, $\mu(\Omega) = +\infty$.

Nello scrivere queste lezioni mi sono basato soprattutto sui libri (Kingman & Taylor, 1966), (Ash, 1972) e (Rao, 1987).

Per la storia della teoria della misura di Lebesgue si veda (Hawkins, 1975). In generale, per la storia della misura e dell'integrazione, si vedano (Pier, 1994.b) e (Dieudonné, 1978.b). Un riferimento piú recente, ed eccellente, tanto per la storia quanto per tutti gli aspetti qui trattati, sono i due volumi (Pap, 2002). Un vero trattato, che dedica grande attenzione agli aspetti della teoria che interessano la probabilità sono i due volumi (Bogachev, 2007).

Le idee fondamentali della teoria della misura sono dovute, in gran parte, a Lebesgue che, nella sua tesi di dottorato (Lebesgue, 1902), le introdusse in $\mathbb{R}^n$. Il concetto di misura svincolato dal sostegno costituito da $\mathbb{R}^n$, e chiamato talvolta "misura astratta", fu introdotto e sviluppato da Fréchet (1915). Questi notò che non è essenziale che gli insiemi E per i quali è definita la misura $\mu(E)$ siano misurabili nel senso di Lebesgue: basta che essi costituiscano una tribú $\mathcal{F}$ e che μ sia numerabilmente additiva su $\mathcal{F}$. Questo svincolava la costruzione della misura dalla topologia dello spazio $\mathbb{R}^n$ e consentiva cosí di definire la misura su insiemi "astratti" qualsiasi Ω; è il punto di vista moderno.

La bibliografia sulle probabilità vere e proprie è molto vasta; di seguito dò una selezione dei libri sulla Probabilità che ho tenuto presenti nello scrivere queste lezioni. Si tratta di un elenco parziale limitato alla letteratura in italiano, inglese e francese; su molti punti ho fatto riferimento ad altre fonti, che citerò nelle note ai varî capitoli. Qui ricordo:

(Ash, 1972), (Bauer, 1981 e 1986), (Breiman, 1968), (Chow & Teicher, 1978), (Chung, 1968), (Cramér, 1946), (Dudley, 1989), (Durrett, 1991), (Feller, 1971), (Fristedt & Gray, 1997), (Jacod & Protter, 2000), (Klenke, 2008), (Kolmogorov, 1933), (Koralov & Sinai, 2007), (Laha & Rohatgi, 1979), (Loève, 1963), (Métivier, 1968), (Neveu, 1964), (Rao, 1984), (Rényi, 1966), (Stromberg, 1994), (Tucker, 1967), (Williams, 1991).

Non si può tacere che, tra i libri citati sopra, tre hanno avuto una grande importanza dal punto di vista storico: (Kolmogorov, 1933), monografia densissima nella quale la probabilità ha trovato il suo assetto moderno e che riporta i risultati fondamentali, (Cramér, 1946) e (Feller, 1950) che sono stati, per lungo tempo, i soli testi di riferimento per gli studiosi.

Un approccio differente da quello tradizionale, basato sull'assiomatizzazione delle speranze anziché delle probabilità si puó trovare in (Whittle, 1992).

Sarà bene tenere presenti le voci dell'enciclopedia (Kotz & Johnson, 1982). Si consultino anche i libri di esempî e controesempî, che sono sempre utili per mettere a cimento le proprie conoscenze, (Romano & Siegel, 1986), (Stoyanov, 1987) e (Székely, 1986).

Sezione 1.3 Per il Teorema 1.3.1 si vedano (Tarski, 1938) e (Horn & Tarski, 1948). Per tutta la problematica riguardante le misure finitamente additive si consulti (Bhaskara Rao & Bhaskara Rao, 1983). Di seguito dò la dimostrazione del Teorema di Tarski basata sul Teorema di Hahn–Banach.

Dimostrazione. Sia B lo spazio di Banach di tutte le funzioni a valori reali limitate definite in Ω, munito della norma della convergenza uniforme

$$\|f\| := \sup_{\omega\in\Omega} |f(\omega)| .$$

Sia S il sottospazio di B costituito dalle funzioni semplici e si definisca il funzionale $I : S \to \mathbb{R}$ mediante

$$I(f) = I\left(\sum_{i=1}^{n} c_i\, \mathbf{1}_{A_i}\right) := \sum_{i=1}^{n} c_i\, \mu\left(A_i\right) .$$

Tale funzionale è omogeneo: se α è un numero reale, si ha $I(\alpha\, f) = \alpha\, I(f)$. Se f e g sono due funzioni semplici

$$f = \sum_{i=1}^{n} c_i\, \mathbf{1}_{A_i} \qquad \text{e} \qquad g = \sum_{j=1}^{k} d_j\, \mathbf{1}_{B_j}\,,$$

ove supporremo che sia

$$A_i \cap A_j = B_i \cap B_j = \emptyset \quad (i \neq k) \quad \text{e} \quad \bigcup_{i=1}^{n} A_i = \bigcup_{j=1}^{k} B_j = \Omega\,,$$

si può scrivere

$$f + g = \sum_{i=1}^{n} \sum_{j=1}^{k} (c_i + d_j)\ \mathbf{1}_{A_i \cap B_j}\,,$$

onde

$$\begin{aligned} I(f+g) &= \sum_{i=1}^{n} \sum_{j=1}^{k} (c_i + d_j)\ \mu\,(A_i \cap B_j) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{k} c_i\, \mu\,(A_i \cap B_j) + \sum_{i=1}^{n} \sum_{j=1}^{k} d_j\, \mu\,(A_i \cap B_j) \\ &= \sum_{i=1}^{n} c_j\, \mu \left[A_j \cap \left(\cup_{j=1}^{k} B_j\right)\right] + \sum_{j=1}^{k} d_j\, \mu\,[(\cup_{i=1}^{n} A_i) \cap B_j] \\ &= \sum_{i=1}^{n} c_j\, \mu\,(A_i) + \sum_{j=1}^{k} d_j\, \mu\,(B_j) = I(f) + I(g)\,. \end{aligned}$$

Il funzionale I è dunque lineare; di piú, esso è limitato, perché

$$|I(f)| \leq \left(\max_{i=1,2,\dots,n} c_i\right) \sum_{i=1}^{n} \mu\,(A_i) = \|f\|\, \mu(\Omega)\,;$$

di qui

$$\|I\| \leq \mu(\Omega)\,. \tag{1.18.1}$$

Per il teorema di Hahn–Banach, si può prolungare il funzionale I come funzionale L a tutto B in modo da conservarne la norma, $\|L\| = \|I\|$. Per ogni sottoinsieme A di Ω si ponga

$$\nu(A) := L\,(\mathbf{1}_A)\,.$$

La linearità di L dà l'additività semplice di ν; inoltre, se A appartiene all'algebra $\mathcal{A}$ si ha

$$\nu(A) = L\,(\mathbf{1}_A) = I\,(\mathbf{1}_A) = \mu(A)\,,$$

sicché ν estende μ. Infine dimostriamo che ν è positiva. A tal fine, si supponga, se possibile, che esista un sottoinsieme B di Ω tale che $\nu(B) < 0$. In tal caso si avrebbe per il complementare B^c di B

$$\|L\| = \|L\|\, \|\mathbf{1}_{B^c}\| \geq |L\,(\mathbf{1}_{B^c})| \geq \nu\,(B^c) = \nu(\Omega) - \nu(B) > \mu(\Omega)\,,$$

ciò che implica

$$\|I\| = \|L\| > \mu(\Omega)\,,$$

che contraddice alla (1.18.1). $\square$

Sezione 1.5 Per il Teorema 1.5.2 si vedano (Banach & Kuratowski, 1929) e (Ulam, 1930). Avverto però che le varianti di questo teorema sono numerose.

Alle funzioni d'insieme finitamente additive, note anche come *cariche* è dedicato il libro (Bhaskara Rao & Bhaskara Rao, 1983).

Uno dei problemi che mal si prestano ad un modello numerabilmente additivo è quello della prima cifra decimale; per questo, si veda (Scozzafava, 1981).

Sezione 1.7 L'integrale di Stieltjes fu introdotto in (Stieltjes, 1894); l'integrale introdotto da Stieltjes si potrebbe chiamare di Riemann–Stieltjes. Questo autore utilizzava la nozione, comune in fisica, di distribuzione di massa: $\varphi(x)$ rappresenta la massa concentrata nell'intervallo $]0,x]$ e i salti di φ rappresentano la massa concentrata nei punti di discontinuità. Si deve a Radon (1913), l'aver riconosciuto che si potevano adattare i metodi di Borel e di Lebesgue, sostituendo alla lunghezza $b-a$ dell'intervallo $]a,b]$ la quantità $\varphi(b)-\varphi(a)$ per ottenere la misura di un aperto di $\mathbb{R}$.

Sezione 1.10 I risultati di questa sezione furono tutti introdotti dagli autori dei quali portano il nome, per l'integrale di Lebesgue, e dunque in un contesto differente da quello nel quale li presentiamo. Si vedano (Levi, 1906), (Fatou, 1906), (Lebesgue, 1902).

Sezione 1.14 Gli spazî di Hilbert furono introdotti e studiati dal punto di vista geometrico da (Schmidt, 1908). Gli spazî L^p furono introdotti e studiati da Riesz (1910). Fu introdotta da Jensen (1906) la diseguaglianza che porta il suo nome. Hölder (1889) dimostrò la diseguaglianza (1.13.7) nel caso delle somme finite; egli indicò chiaramente il suo debito verso Rogers (1888), tanto che qualche autore chiama la (1.13.7) diseguaglianza di Rogers–Hölder. Anche la diseguaglianza (1.13.8) fu dimostrata nel caso delle somme finite da Minkowski (1896). Entrambe le diseguaglianze furono estese agli integrali da Riesz (1910).

In generale nel rintracciare gli autori delle diseguaglianze occorre tenere presente il monito contenuto nell'introduzione di (Hardy et al., 1934): *... we speak of the inequalities of Schwarz, Hölder and Jensen, though all these inequalitites can be traced further back; and we do not enumerate explicitly all the minor additions which are necessary for absolute completeness.*

Di fatto ci siamo occupati, sia pur brevemente, dei soli spazî L^p con $p \geq 1$, e solo di questi darò le applicazioni; ciò accade perché in uno spazio L^p con $0 < p < 1$ viene meno la diseguaglianza triangolare e il solo funzionale lineare è quello nullo. Si veda, a questo proposito, (Day, 1940).

Mitrinović & Lacković (1985) attribuiscono l'origine del termine *convesso* a Hermite che lo introdusse nel 1881.

La dimostrazione del Teorema di Fischer–Riesz che gli spazî L^p, con $p \in [1,+\infty[$, sono completi può essere consultata, per esempio, in (Dunford & Schwartz, 1958) (III.6.6).

Il teorema di rappresentazione di Riesz–Fréchet è dovuto indipendentemente a Riesz (1907) e a Fréchet (1907); esso compare in due note pubblicate nello stesso fascicolo dei *Comptes Rendus.*

Le dimostrazioni del teorema di Radon–Nikodym si possono ricondurre a tre: quella originale degli autori dei quali porta il nome, (Radon, 1913) e (Nikodym, 1930), quella data in questo capitolo e dovuta a von Neumann (1940) ed infine quella basata sul teorema di convergenza per le martingale. Tutte e tre le dimostrazioni si possono trovare in (Rao, 1987).

Il teorema di Radon–Nikodym è essenziale per molte applicazioni anche alla Statistica matematica; a questo proposito, si veda l'articolo (Halmos & Savage, 1949).

Sezione 1.15 Lebesgue (1904) dimostrò che per una funzione limitata e misurabile f definita in un rettangolo $(a, b) \times (a, b)$ i due integrali iterati assumono lo stesso valore

$$\int_a^b \mathrm{d}x \int_a^b f(x, y)\, \mathrm{d}y = \int_a^b dy \int_a^b f(x, y)\, \mathrm{d}x\,.$$

Lebesgue trovò quindi un caso particolare del teorema di Fubini; inoltre, prendendo come f una funzione indicatrice, si prova l'esistenza della misura prodotto, almeno sui rettangoli del tipo $(a, b) \times (a, b)$.

Per il teorema di Fubini si vedano (Fubini, 1907), (Tonelli, 1909) e l'ultimo lavoro di Fubini, pubblicato postumo, (Fubini, 1949), nel quale si spiega la genesi di tale teorema. Per le cautele necessarie nell'utilizzo del teorema di Fubini si veda l'articolo (Chatterji, 1986).

1.19 Esercizî sul Capitolo 1

1.1. (a) Sia $(\mathcal{A}_n)$ una successione crescente di algebre di sottoinsiemi di Ω. Anche $\mathcal{A} := \cup_{n\in\mathbb{N}}\mathcal{A}_n$ è un'algebra.

(b) Se $(\mathcal{F}_n)$ una successione crescente di tribú di sottoinsiemi di Ω, l'unione $\cup_{n\in\mathbb{N}}\mathcal{F}_n$ non è necessariamente una tribú.

1.2. (a) La famiglia dei sottoinsiemi di $\mathbb{N}$ che o sono finiti o hanno il complementare finito, cioè

$$\mathcal{A} := \{A \subseteq \mathbb{N} : \mathrm{card}(A) < \aleph_0 \text{ o } \mathrm{card}(A^c) < \aleph_0\}$$

è un'algebra ma non una tribú.

(b) Se Ω non è finito o numerabile ($\mathrm{card}(\Omega) > \aleph_0$), la famiglia dei sottoinsiemi di Ω che o sono finiti o numerabili o che hanno il complementare finito o numerabile, cioè

$$\mathcal{F} := \{A \subseteq \Omega : \mathrm{card}(A) \leq \aleph_0 \text{ o } \mathrm{card}(A^c) \leq \aleph_0\}$$

è una tribú.

1.3. Sia $\mathcal{A}$ una famiglia di sottoinsiemi di Ω che goda delle seguenti proprietà:

(a) $\Omega \in \mathcal{A}$;

(b) $A \in \mathcal{A} \Longrightarrow A^c \in \mathcal{A}$;

(c) $\mathcal{A}$ è stabile rispetto all'unione finita ($A_1, A_2 \in \mathcal{A} \Longrightarrow A_1 \cup A_2 \in \mathcal{A}$);

(d) se $A_n \in \mathcal{A}$ per ogni $n \in \mathbb{N}$ e se gli insiemi della successione (A_n) sono a due a due disgiunti, allora $\cup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Allora $\mathcal{A}$ è una tribú.

1.4. Siano $\mathcal{F}$ una tribú e A e B insiemi di $\mathcal{F}$ tali che $A \subseteq B$ e $A \neq B$. Esiste allora una probabilità $\mathbb{P}$ su $\mathcal{F}$ tale che $\mathbb{P}(A) < \mathbb{P}(B)$.

1.5. Se per ogni n di $\mathbb{N}$ si ha $A_n \in \mathcal{F}$ e $\mathbb{P}(A_n) = 1$, allora

$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = 1 .$$

1.6. Per un'arbitraria successione $(A_n) \subseteq \mathcal{F}$ riesce

$$\mathbb{P}\left(\liminf_{n\to+\infty} A_n\right) \leq \liminf_{n\to+\infty} \mathbb{P}(A_n) \leq \limsup_{n\to+\infty} \mathbb{P}(A_n) \leq \mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) .$$

sicché $A_n \to A$ implica $\mathbb{P}(A_n) \to \mathbb{P}(A)$ anche se la successione non è monotona, in altre parole, le probabilità passano al limite lungo tutte le successioni convergenti di eventi. Si dia l'esempio di uno spazio di probabilità e di una successione di eventi per i quali le diseguaglianze precedenti siano tutte strette.

1.7. $\mathbb{P}(A_n \Delta A) \to 0$ implica sia $\mathbb{P}(A_n) \to \mathbb{P}(A)$ sia $\mathbb{P}(A_n \cap A) \to \mathbb{P}(A)$.

1.8. Sia $\mathcal{A}$ l'algebra dell'esercizio (1.2)(a). Se $\mu : \mathcal{A} \to \{0, 1\}$ è definita mediante

$$\mu(A) = \begin{cases} 0, & \text{se } \operatorname{card}(A) < \aleph_0 , \\ 1, & \text{se } \operatorname{card}(A^c) < \aleph_0 , \end{cases}$$

allora μ è finitamente additiva, ma non σ–additiva.

1.9. Nell'insieme $\mathbb{N}$ dei numeri naturali si consideri la famiglia $\mathcal{D}$ dei sottoinsiemi A di $\mathbb{N}$ per i quali esiste il limite

$$\mu(A) := \lim_{n\to+\infty} \frac{\operatorname{card}(A \bigcap \{1, \dots, n\})}{n} .$$

Tale limite, quando esiste, si dice *densità asintotica* dell'insieme A. Si mostri che

(a) μ è finitamente, ma non σ–additiva su $\mathcal{D}$;

(b) $\mathcal{D}$ contiene $\emptyset$ e $\mathbb{N}$ ed è stabile rispetto alla complementazione e alle unioni finite e disgiunte, mentre non è stabile rispetto alle unioni numerabili o alle unioni finite non disgiunte.

(c) Si considerino infine gli insiemi $P := \{2n : n \in \mathbb{N}\}$ di tutti i numeri pari e $P^c := \{2n-1 : n \in \mathbb{N}\}$ di tutti i numeri dispari. Se

$$F_n := \left(P^c \bigcap \left[2^{2n}, 2^{2n+1}\right]\right) \bigcup \left(P \bigcap \left[2^{2n+1}, 2^{2n+2}\right]\right)$$

e

$$F := \bigcup_{n \in \mathbb{Z}_+} F_n \,,$$

allora si ha tanto $P^c \in \mathcal{D}$ quanto $F \in \mathcal{D}$, ma $P^c \cap F \notin \mathcal{D}$, sicché $\mathcal{D}$ non è un'algebra.

1.10. Nell'esercizio precedente si mostri che, per ogni $\alpha \in [0,1]$, esiste un insieme $E \in \mathcal{D}$ tale che $\mu(E) = \alpha$.

1.11. Si dia l'esempio di una misura (necessariamente infinita) e di una successione (A_n) di insiemi misurabili per i quali $A_n \downarrow \emptyset$ ma tale che $(\mu(A_n))$ non converga a zero.

1.12. In uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, si scriva

$$A \sim B \quad \text{se} \quad P(A \Delta B) = 0$$

con A e B in $\mathcal{F}$; "$\sim$" è una relazione d'equivalenza su $\mathcal{F}$ compatibile con le operazioni di unione, intersezione e complementazione. Si scriva $[A]$ per indicare la classe di equivalenza di A. Inoltre, nell'insieme quoziente $\mathcal{F}^{\sim} := \mathcal{F}|_{\sim}$, la funzione $d([A],[B]) := P(A \Delta B)$ ove $A \in [A]$ e $B \in [B]$, è una metrica.

1.13. Si mostri che una misura μ su $(\Omega, \mathcal{F})$ è σ–subadditiva: se (A_n) è una successione di insiemi misurabili (non necessariamente disgiunti a due a due), allora

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mu(A_n) \,.$$

1.14. Per ogni anello $\mathcal{R}$ si ha:

(a) $\emptyset \in \mathcal{R}$;

(b) $A, B \in \mathcal{R} \Rightarrow A \cup B \in \mathcal{R} \quad \text{e} \quad A \setminus B \in \mathcal{R}$.

1.15. Un'algebra è un anello che contiene Ω. Un anello è un'algebra se, e solo se, è stabile rispetto all'operazione di complementazione.

1.16. Ogni misura localmente finita su $(\mathbb{R}, \mathcal{B})$ è σ–finita.

1.17. In uno spazio topologico Ω si indica con $\mathcal{B}(\Omega)$ la tribú di Borel, cioè quella generata dagli insiemi aperti. Se Ω e Ω' sono spazî topologici e $f : \Omega \to \Omega'$ è una funzione continua, essa è misurabile rispetto a $\mathcal{B}(\Omega)$ e a $\mathcal{B}(\Omega')$.

1.18. Sia μ una misura infinita, ma σ–finita, definita sullo spazio misurabile $(\Omega, \mathcal{F})$. Allora per ogni $k > 0$ esiste un insieme A di $\mathcal{F}$ tale che $\mu(A) \in \,]k, \infty[$.

1.19. Si dimostri che, data una funzione semplice

$$s = \sum_{j=1}^{k} \alpha_j \, \mathbf{1}_{A_j} ,$$

ove gli insiemi A_j sono disgiunti, misurabili e formano una partizione di Ω, la funzione semplice s può essere rappresentata nella forma

$$s = \sum_{j=1}^{k} \beta_j \, \mathbf{1}_{C_j} ,$$

ove gli insiemi $C_1, C_2, \dots, C_k$ sono misurabili e soddisfanno alle relazioni

$$C_1 \supset C_2 \supset \cdots \supset C_k .$$

1.20. Siano $\mathcal{C}_1$ e $\mathcal{C}_2$ semi–anelli di sottoinsiemi degli insiemi Ω_1 e Ω_2 rispettivamente. Allora $\mathcal{C}_1 \times \mathcal{C}_2$ è un semi–anello di sottoinsiemi di $\Omega_1 \times \Omega_2$.

1.21. Se, per ogni $n \in \mathbb{N}$, le funzioni f_n sono integrabili, e $f_n \downarrow 0$, allora $\int f_n \, \mathrm{d}\mu \downarrow 0$.

1.22. Si facciano discendere le diseguaglianze dell'esercizio 3) dal lemma di Fatou.

1.23. Se, per ogni $n \in \mathbb{N}$, le funzioni $f_n : \Omega \to \overline{\mathbb{R}}_+$ sono misurabili (e positive), si ha

$$\int \left(\sum_n f_n \right) \mathrm{d}\mu = \sum_n \int f_n \, \mathrm{d}\mu .$$

1.24. Se $f, g : \Omega \to \overline{\mathbb{R}}$ sono integrabili si stabilisca la diseguaglianza

$$\int (f \wedge g) \, \mathrm{d}\mu \leq \int f \, \mathrm{d}\mu \wedge \int g \, \mathrm{d}\mu .$$

Si studii il caso dell'eguaglianza.

1.25. Sia $E = \cup_{n\in\mathbb{N}} E_n$ l'unione numerabile di insiemi disgiunti di $\mathcal{F}$. Se $f : \Omega \to \overline{\mathbb{R}}$ è integrabile, si mostri che

$$\int_E f \, \mathrm{d}\mu = \sum_{n\in\mathbb{N}} \int_{E_n} f \, \mathrm{d}\mu .$$

1.26. Sia λ la misura di Lebesgue su $(\mathbb{R}, \mathcal{B})$ e, per ogni $n \in \mathbb{N}$, si definisca $f_n : \mathbb{R} \to \mathbb{R}$ mediante $f_n := -n^2 \, \mathbf{1}_{]0,1/n[}$. Si mostri che per tale successione non vale il lemma di Fatou.

1.27. Se $\nu = \mu \, f^{-1}$, la σ–finitezza di μ è condizione necessaria, ma non sufficiente per la σ–finitezza di ν.

1.28. Si dia l'esempio di due spazî misurabili $(\Omega_1, \mathcal{F}_1)$ e $(\Omega_2, \mathcal{F}_2)$ tali che la famiglia $\mathcal{F}_1 \times \mathcal{F}_2$ dei rettangoli misurabili,

$$\mathcal{F}_1 \times \mathcal{F}_2 := \{A \times B : A \in \mathcal{F}_1, B \in \mathcal{F}_2\}$$

non sia una tribú.

1.29. Si considerino gli spazî mensurali $(\Omega_i, \mathcal{F}_i, \mu_i)$ con $i = 1, 2$, ove $\Omega_i = \mathbb{N}$, $\mathcal{F}_i = \mathcal{P}(\mathbb{N})$ e μ_i è la misura del contare (cioè $\mu_i(\{n\}) = 1$ per ogni $n \in \mathbb{N}$, $i = 1, 2$). Per la funzione $f : \mathbb{N} \times \mathbb{N} \to \mathbb{Z}$ definita da $f(n, n) = n$, $f(n, n+1) = -n$, $f(n, k) = 0$ se $k \neq n, n+1$, si mostri che

$$\int_{\mathbb{N}} \mathrm{d}\mu_1 \int_{\mathbb{N}} f \, \mathrm{d}\mu_2 \neq \int_{\mathbb{N}} \mathrm{d}\mu_2 \int_{\mathbb{N}} f \, \mathrm{d}\mu_1 .$$

1.30. Si applichi il teorema di Fubini alla funzione

$$(x, y) \mapsto \varphi(x, y) := y \exp\{-(1 + x^2)\, y^2\} \quad (x, y) \in \mathbb{R}_+^2 ,$$

per trovare che

$$\int_{\mathbb{R}_+} \exp(-t^2) \, \mathrm{d}t = \frac{\sqrt{\pi}}{2} .$$

1.31. Gli integrali iterati nel teorema di Fubini possono essere entrambi finiti ma differenti. Sia $\Omega_1 = \Omega_2 = I := [0, 1]$ e $\mu_1 = \mu_2 = \lambda$, la (restrizione della) misura di Lebesgue ai boreliani di $[0, 1]$. Sia (δ_n) una successione strettamente crescente con $\delta_1 = 0$ e $\lim_n \delta_n = 1$ e sia $g_n : I \to \mathbb{R}$ una funzione continua con supporto contenuto in $[\delta_n, \delta_{n+1}]$ e tale che $\int g_n \, \mathrm{d}\lambda = 1$ per ogni $n \in \mathbb{N}$. Si definisca $f : I \times I \to \mathbb{R}$ mediante

$$f(x, y) := \sum_{n=1}^{\infty} \{g_n(x) - g_{n+1}(x)\} g_n(y) .$$

Allora

$$\int_I \mathrm{d}x \int_I f(x, y) \, \mathrm{d}y \neq \int_I \mathrm{d}y \int_I f(x, y) \, \mathrm{d}x .$$

1.32. Siano $\Omega_1 = \Omega_2 = I := [0, 1]$, μ_1 la (restrizione della) misura di Lebesgue a(i boreliani di) I e μ_2 la misura del contare su $\mathcal{B} = \mathcal{B}(I)$, cioè $\mu_2(A) = \operatorname{card}(A)$ se l'insieme A è finito, mentre $\mu_2(A) = +\infty$ se A è infinito. Sia, infine, D la diagonale del quadrato $I \times I$, $D := \{(x, x) : x \in I\}$. Si mostri che D è in $\mathcal{B} \otimes \mathcal{B}$ e che se $f = \mathbf{1}_D$ risulta

$$\int_I d\mu_1 \int_I f \, \mathrm{d}\mu_2 \neq \int_I d\mu_2 \int_I f \, \mathrm{d}\mu_1 .$$

1.33. Siano $(\Omega, \mathcal{F})$ e $(\Omega_1, \mathcal{F}_1)$ due spazî misurabili e ν una misura su $\mathcal{F}_1$. Per quasi ogni $y \in \Omega_1$ rispetto a ν, sia μ_y una misura su $\mathcal{F}$ tale che, per ogni $E \in \mathcal{F}$, la funzione $y \mapsto \mu_y(E)$ sia $\mathcal{F}_1$–misurabile. Se $\mu : \mathcal{F} \to \mathbb{R}$ è definita da

$$\mu(A) := \int_{\Omega_1} \mu_y(A) \, \mathrm{d}\nu(y) ,$$

allora:

(a) μ è una misura su $\mathcal{F}$; ed è una probabilità, $\mu(\Omega) = 1$, se $\nu(\Omega_1) = \mu_y(\Omega) = 1$ per quasi ogni $y \in \Omega_1$;

(b) se $f : \Omega \to \bar{\mathbb{R}}_+$ è $\mathcal{F}$–misurabile, è $\mathcal{F}_1$–misurabile la funzione $g : \Omega_1 \to \mathbb{R}$ definita da

$$g(y) := \int_\Omega f \, \mathrm{d}\mu_y \,;$$

inoltre

$$\int_\Omega f \, \mathrm{d}\mu = \int_{\Omega_1} g \, \mathrm{d}\nu \,.$$

1.34. Si mostri che, per due misure μ e ν su $(\Omega, \mathcal{F})$, con ν finita, sono equivalenti le affermazioni:

(a) $\mu \ll \nu$ (μ è assolutamente continua rispetto a ν);

(b) ad ogni $\varepsilon > 0$ corrisponde $\delta = \delta(\varepsilon) > 0$ tale che $\mu(E) < \varepsilon$ per ogni insieme $E \in \mathcal{F}$ con $\nu(E) < \delta$.

È ineliminabile la richiesta che ν sia finita. Per esempio, si considerino, sullo spazio misurabile $(\Omega, \mathcal{F})$ con $\Omega = \mathbb{N}$ e $\mathcal{F} = \mathcal{P}(\Omega)$, le misure definite da

$$\mu(A) := \sum_{n \in \mathbb{N}} 2^{-n} \qquad \text{e} \qquad \nu(A) := \sum_{n \in \mathbb{N}} 2^n \,.$$

1.35. Si mostri, mediante un esempio, che la densità di una misura di probabilità P rispetto ad un'altra misura di probabilità μ, con $P \ll \mu$, non è necessariamente limitata μ–q.c..

1.36. In uno spazio mensurale finito $(\Omega, \mathcal{F}, \mu)$, si supponga che una funzione φ di $\mathcal{L}^1$ sia tale che per ogni insieme misurabile $A \in \mathcal{F}$ con $\mu(A) > 0$ risulti

$$0 \le \frac{1}{\mu(A)} \int_A \varphi \, \mathrm{d}\mu \le 1 \,;$$

allora φ assume valori in $[0, 1]$, μ–q.o.

1.37. Nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ si considerino gli insiemi misurabili A_1, A_2,..., A_n. Vale allora la *diseguaglianza di Bonferroni*

$$\mathbb{P}\left(\bigcup_{j=1}^n A_j\right) \ge \sum_{j=1}^n \mathbb{P}(A_j) - \sum_{i \ne j} \mathbb{P}\left(A_i \bigcap A_j\right) \,.$$

1.38. Quando vale l'eguaglianza nella diseguaglianza di Hölder? e nella diseguaglianza di Schwarz?

1.39. Se nello spazio $(\Omega, \mathcal{F}, \mu)$ la misura μ non è finita può non valere piú l'inclusione $L^q \subseteq L^p$ se $q > p \ge 1$.

(a) Se $\Omega = \mathbb{N}$, $\mathcal{F} = \mathcal{P}(\mathbb{N})$ e μ è la misura del contare, si studii la relazione tra gli spazî $l^p := L^p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ e $l^q := L^q(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ con $q > p \ge 1$;

(b) se $\Omega = \mathbb{R}_+$, $\mathcal{F} = \mathcal{B}(\mathbb{R}_+)$, e $\mu = \lambda$ è la misura di Lebesgue, si mostri che né $L^1 \subseteq L^2$ né $L^2 \subseteq L^1$.

1.40. Siano μ_j e ν_j due misure sullo spazio misurabile $(\Omega_j, \mathcal{F}_j)$ con $j = 1, 2$ e siano $\mu = \mu_1 \otimes \mu_2$ e $\nu = \nu_1 \otimes \nu_2$ le rispettive misure prodotto su $(\Omega, \mathcal{F})$ con $\Omega = \Omega_1 \otimes \Omega_2$ e $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$. Se μ_j è assolutamente continua rispetto a ν_j, $\mu_j \ll \nu_j$ $(j = 1, 2)$, con densità f_j $(\mu_j = f_j \cdot \nu_j)$, si mostri che μ è assolutamente continua rispetto a ν e si calcoli la densità f di μ rispetto a ν.

1.41. Si dia l'esempio di uno spazio misurabile $(\Omega, \mathcal{F})$ e di due misure definite su $(\Omega, \mathcal{F})$ che coincidono sugli insiemi di una famiglia $\mathcal{C}$ che genera $\mathcal{F}$ $(\mathcal{F}(\mathcal{C}) = \mathcal{F})$.

1.42. Si chiama *distanza in variazione* di due misure di probabilità μ e ν su $(\mathbb{R}, \mathcal{B})$ la quantità

$$d_V(\mu, \nu) := \sup_{B \in \mathcal{B}} |\mu(B) - \nu(B)| \ .$$

In questo modo si è definita una distanza sullo spazio $\mathcal{M}_1(\mathbb{R}, \mathcal{B})$ delle misure di perobabilità su $(\mathbb{R}, \mathcal{B})$; inoltre si ha

$$\forall \mu, \nu \in \mathcal{M}_1(\mathbb{R}, \mathcal{B}) \qquad 0 \le d_V(\mu, \nu) \le 1 \,.$$

Se le probabilità μ e ν sono assolutamente continue, con densità f e g rispettivamente, $\mu = f \cdot \lambda$ e $\nu = g \cdot \lambda$ (λ è la misura di Lebesgue su $(\mathbb{R}, \mathcal{B})$), allora è

$$d_V(\mu, \nu) \le \int |f - g| \, \mathrm{d}\lambda \le 2 \, \mathrm{d}_V(\mu, \nu) \,.$$