7. Cenni di integrazione astratta e integrale di Lebesgue

Nel paragrafo successivo introdurremo (senza dimostrazioni) il concetto di integrale astratto, ossia definiremo l'integrale astratto di una funzione misurabile

$$\xi : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})$$

rispetto ad una misura positiva σ-additiva μ su $\mathcal{E}$. Tale integrale verrà indicato (quando esiste) con

$$\int_E \xi(e)\mu(de).$$

Le dimostrazioni dei risultati qui riportati si possono trovare in un qualunque buon libro di teoria della misura. Si veda in alternativa [3].

7.1. **Integrale astratto rispetto ad una misura positiva.** In questa sezione consideriamo uno spazio di misure $(E, \mathcal{E}, \mu)$, con μ misura σ-finita e una funzione misurabile

$$\xi : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\bar{\mathbb{R}})).$$

Descriviamo la costruzione per uno spazio astratto $(E, \mathcal{E})$ e una generica funzione misurabile $\xi : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. In questa sezione usiamo la lettera ξ per indicare una funzione misurabile e non la lettera X, e la coppia $(E, \mathcal{E})$ e non $(\Omega, \mathcal{F})$ per indicare lo spazio di misura sottostante nella speranza che questa differente notazione ricordi allo studente che la costruzione non è limitata al caso di uno spazio di probabilità e ξ variabile aleatoria reale.

In seguito, infatti, useremo i risultati di questa sezione in tre casi distinti

- $(E, \mathcal{E}, \mu) = (\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D), \mathrm{Leb}_D)$ (Sezione 7.2);
- $(E, \mathcal{E}, \mu) = (\Omega, \mathcal{F}, \mathbb{P})$ e $\xi(e) = X(\omega)$ con $\mathbb{P}$ misura di probabilità (Capitolo 9);
- $(E, \mathcal{E}, \mu) = (\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D), P_{\mathbf{X}})$, $\xi(e) = g(x)$ dove $P_{\mathbf{X}}$ (o P_X) è la legge di un vettore aleatorio (Sezione 9.3).

Definizione 7.1. *Una funzione misurabile $S : E \to \mathbb{R}$ si dice* **semplice in forma canonica** *se*

$$S(e) = \sum_{i=1}^{M} c_i \mathbb{I}_{A_i}(e) \tag{18}$$

con $A_1, \dots, A_M$ insiemi a due a due incompatibili di $\mathcal{E}$ tali che $\cup_{i=1}^M A_i = E$, $c_i \in \mathbb{R}$ con $c_i \neq c_j$ per $i \neq j$. Una funzione S come in (18) dove $A_1, \dots, A_M$ insiemi di $\mathcal{E}$ (non necessariamente incompatibili) e $c_i \in \mathbb{R}$ non necessariamente distinti, si dice **semplice**. *Una funzione semplice si dice positiva se $S(e) \geqslant 0$ per ogni e.*

Data una funzione semplice S è sempre possibile riscriverla come una funzione semplice in forma canonica S'. Inoltre se S è una funzione semplice in forma canonica, essa è positiva se e solo se $c_i \geqslant 0$.

Definizione 7.2. *Data una misura μ su $\mathcal{E}$ e una funzione semplice positiva S, l'integrale di S rispetto μ, $\int_E S(e)\mu(de)$ in simboli, si definisce come il valore della somma a termini positivi*

$$\int_E S(e)\mu(de) := \sum_i c_i \mu(A_i).$$

Tale somma può assumere il valore $+\infty$. L'integrale $\int_E S(e)\mu(de)$ si dice finito se la somma che lo definisce è finita.

Si noti che, come conseguenza dell'additività finita della misura μ, la definizione data non dipende dalla rappresentazione di S (a patto di prendere tutti i $c_i \geqslant 0$). In particolare in Definizione 7.2 si può assumere che S sia in forma canonica.

Proposizione 7.1. *Sia $\xi : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\bar{\mathbb{R}}))$ una funzione misurabile a valori reali positivi, ossia $\xi(e) \geqslant 0$ per ogni e. Allora esiste una successione di funzioni semplici $S_n(e) = \sum_{i=1}^{M_n} c_{in} \mathbb{I}_{A_{i,n}}(e)$ tale che $\lim_n S_n = \xi$ e $S_n(e) \leqslant S_{n+1}(e)$ per ogni $n \geqslant 1$ e e in E. Inoltre se S_n e S'_n sono due successioni di funzioni misurabili semplici, positive e monotone crescenti tali che $\lim_n S_n = \lim_n S'_n$, allora $\lim_n \int S_n(e)\mu(de) = \lim_n \int S'_n(e)\mu(de)$ per qualunque misura μ.*

Dimostrazione. [Prop. 1.6 e Lemma 1.16 in [1]] traccia: è sufficiente scegliere $S_n(e) = k/2^n$ se $\xi(e) < n$ e $k/2^n \leqslant \xi(e) < (k+1)/2^n$ $k = 0, \dots, n2^n - 1$. La seconda parte è lasciata come esercizio, si deve usare la monotonia. ✠

Definizione 7.3. *Data una misura μ su $\mathcal{E}$ e $\xi : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\bar{\mathbb{R}}))$ funzione misurabile positiva (ossia $\xi(e) > 0$ per ogni $e \in E$) allora*

$$\int_E \xi(e)\mu(de) := \lim_n \int_E S_n(e)\mu(de)$$

dove S_n è data nella proposizione precedente.

Una definizione alternativa ma del tutto equivalente è

$$\int \xi(e)\mu(de) := \sup_{S \text{ semplice}: 0 \leqslant S \leqslant \xi} \int S(e)\mu(de).$$

Non è difficile dimostrare che queste due definizioni coincidono.

Veniamo ora alla definizione di integrale per una funzione misurabile di segno qualunque. Data una generica funzione misurabile $e \mapsto \xi(e)$, si scriva

$$\xi(e) = \xi^+(e) - \xi^-(e).$$

Si ha che $0 \leqslant |\xi(e)| = \xi^+(e) + \xi^-(e)$, inoltre $|\xi^\pm(e)| \leqslant |\xi(e)|$. Quindi possiamo usare la definizione (7.3) per definire $\int_E |\xi(e)|\mu(de)$, $\int_E \xi^+(e)\mu(de)$ e $\int_E \xi^-(e)\mu(de)$, inoltre si vede facilmente usando la definizione che

$$0 \leqslant \int_E \xi^\pm(e)\mu(de) \leqslant \int_E |\xi(e)|\mu(de).$$

Possiamo quindi definire l'integrale astratto per una generica ξ misurabile come segue.

Definizione 7.4. *Data una misura μ su $\mathcal{E}$ e $\xi : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\bar{\mathbb{R}}))$ funzione misurabile, si dice che ξ ammette integrale rispetto a μ finito (è μ-integrabile) se $\int_E |\xi(e)|\mu(de) < +\infty$. In questo caso si definisce*

$$\int_E \xi(e)\mu(de) = \int_E \xi(e)^+\mu(de) - \int_E \xi(e)^-\mu(de).$$

Se $\int_E \xi^+(e)\mu(de) = +\infty$ e $\int_E \xi^-(e)\mu(de) < +\infty$ allora ξ si dice avere integrale infinito (ma ben definito), e $\int_E \xi(e)\mu(de) := +\infty$. Analogamente, se $\int_E \xi^+(e)\mu(de) < +\infty$ e $\int_E \xi^-(e)\mu(de) = +\infty$ allora ξ si dice avere integrale infinito e $\int_E \xi(e)\mu(de) := -\infty$. Infine se $\int_E \xi^+(e)\mu(de) = +\infty$ e $\int_E \xi^-(e)\mu(de) = +\infty$, allora ξ si dice non ammettere integrale rispetto a μ.

Si noti che se $A \in \mathcal{E}$ e ξ è misurabile allora è misurabile anche $\mathbb{I}_A(e)\xi(e)$. Quest'osservazione consente di dar senso alla definizione seguente.

Definizione 7.5. *Data una misura μ su $\mathcal{F}$, $\xi : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\bar{\mathbb{R}}))$ funzione misurabile e $A \in \mathcal{E}$, l'integrale di ξ su A rispetto a μ, indicato con $\int_A \xi(e)\mu(de)$ è definito da*

$$\int_A \xi(e)\mu(de) := \int_E \mathbb{I}_A(e)\xi(e)\mu(de).$$

Raccogliamo nella successiva proposizione (senza dimostrazione) alcune proprietà fondamentali dell'integrale astratto rispetto ad una misura positiva.

Proposizione 7.2. *Sia μ una misura su $\mathcal{E}$.*

(1) **Monotonia.** *Se ξ è una funzione misurabile a valori reali e $\xi \geqslant 0$, allora*

$$\int_E \xi(e)\mu(de) \geqslant 0.$$

(2) **Linearità.** *Dati a e b numeri reali, ξ_1 e ξ_2 funzioni reali misurabili μ-integrabili, allora $a\xi_1 + b\xi_2$ è μ-integrabile e*

$$\int_E (a\xi_1(e) + b\xi_2(e))\mu(de) = a\int_E \xi_1(e)\mu(de) + b\int_E \xi_2(e)\mu(de).$$

(3) **Insensitività.** *Se $A \in \mathcal{E}$ con $\mu(A) = 0$ e ξ è funzione reale misurabile μ-integrabile allora $\int_A \xi(e)\mu(de) = 0$. Se ξ_1 e ξ_2 sono due funzioni reali misurabili e integrabili tali che $\mu\{e : \xi_1(e) \neq \xi_2(e)\} = 0$, allora $\int_E \xi_1(e)\mu(de) = \int_E \xi_2(e)\mu(de)$.*

(4) **Convergenza monotona.** *Se $(\xi_n)_n$ è una successione di funzioni misurabili tali che per ogni n e per ogni $e \in E$ valga $0 \leqslant \xi_n(e) \leqslant \xi_{n+1}(e)$, allora*

$$\lim_{n\to+\infty} \int_E \xi_n(e)\mu(de) = \int_E \lim_{n\to+\infty} \xi_n(e)\mu(de).$$

(5) **Convergenza di serie.** *Sia $(\xi_n)_n$ una successione di funzioni reali misurabili, positive allora*

$$\sum_n \int_E \xi_n(e)\mu(de) = \int_E \Big(\sum_n \xi_n(e)\Big)\mu(de).$$

Se uno dei due membri dell'uguaglianza è infinito lo è necessariamente anche l'altro.

(6) **Finitezza.** *Se ξ è misurabile e μ-integrabile allora $\mu\{e : |\xi(e)| = +\infty\} = 0$.*

[Sec. 1.4, Thm. 1.17, Cor. 1.22 [1]]

A proposito dei punti (4) e (5) della Proposizione precedente, si ricordi che limiti di successioni misurabili sono misurabili (cf. Proposizione 5.6).

Esempio 7.1 (Integrale rispetto ad una misura discreta)**.** *Riprendiamo l'Esempio 1.3 con notazioni leggermente modificate. Su $(E, \mathcal{E})$ consideriamo*

$$\mu(A) = \sum_{n\geqslant 1} c_n \delta_{e_n}(A)$$

dove $(e_n)_{n\geqslant 1}$ è una successione (al più numerabile) di punti in E e $(c_n)_{n\geqslant 1}$ una successione di numeri $c_n > 0$. Se ξ è misurabile e positiva, usando la definizione si verifica che

$$\int_E \xi(e)\mu(de) = \sum_n \xi(e)c_n.$$

Se ξ ha segno qualunque ma $\sum_n |\xi(e_n)|c_n < +\infty$ (ossia se la serie è assolutamente convergente) allora ξ è μ-integrabile e $\int_E \xi(e)\mu(de) = \sum_n \xi(e_n)c_n$.

Esempio 7.2 (Integrale rispetto ad una misura di probabilità discreta)**.** *Sia $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ una v.a. discreta. Si verifica usando la definizione (con $E = \Omega$, $\mathcal{E} = \mathcal{F}$, $\mu = \mathbb{P}$ e $\xi = X$) che*

$$\int_\Omega X(\omega)\mathbb{P}(d\omega) = \sum_{x\in C} x\mathbb{P}\{X = x\} = \mathbb{E}[X]$$

dove C è il supporto di X e abbiamo assunto $\sum_{x\in C} |x|\mathbb{P}\{X = x\} < +\infty$. In questo caso se consideriamo $(E, \mathcal{E}, \mu) = (C, \mathcal{P}(C), P_X)$, dove P_X la misura immagine di X, vale

$$P_X(A) = \sum_{x\in C} p_X(x)\delta_x(A)$$

e quindi usando l'esempio precedente

$$\int_C xP_X(dx) = \sum_{x\in C} xP_X(x) = \sum_{x\in C} x\mathbb{P}\{X = x\} = \mathbb{E}[X].$$

In altri termini, in queste circostanze, indicato con $\mathbb{E}[X]$ il valore atteso introdotto in Sezione 6.3, valgono le seguenti identià:

$$\mathbb{E}[X] = \int_\Omega X(\omega)\mathbb{P}(d\omega) = \int_C xP_X(dx) = \sum_{x\in C} x\mathbb{P}\{X = x\}$$

assumendo $\sum_{x\in C} |x|\mathbb{P}\{X = x\} < +\infty$.

Un esempio molto importante per questo corso è il caso in cui $\mu = \mathrm{Leb}_D$, che trattiamo nel successivo paragrafo.

7.2. **Integrale rispetto alla misura di Lebesgue su $\mathbb{R}^D$.** Disponendo della definizione di integrale astratto possiamo definire l'integrale di Lebesgue per funzioni definite su $\mathbb{R}^D$. Basta considerare la costruzione astratta nel caso in cui $(E, \mathcal{E}) = (\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$ e $\mu = \mathrm{Leb}_D$. Data una funzione misurabile $h : (\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ il suo integrale astratto rispetto a Leb_D è detto integrale di Lebesgue e si indica con

$$\int_{\mathbb{R}^D} h(x)\mathrm{Leb}_D(dx)$$

o più semplicemente con

$$\int_{\mathbb{R}^D} h(x)dx \quad \text{oppure} \quad \int_{\mathbb{R}^D} h(x_1, \ldots, x_D)dx_1 \ldots dx_D.$$

Coerentemente con le definizioni viste si ha per ogni boreliano A

$$\int_A h(x)dx = \int_{\mathbb{R}^D} \mathbb{I}_A(x)h(x)dx.$$

Si avrà dunque che una funzione misurabile (boreliana) h è Lebesgue integrabile se

$$\int_{\mathbb{R}^D} |h(x)|dx < +\infty.$$

Una funzione boreliana di segno qualunque h è integrabile alla Lebesgue se sono integrabili alla Lebesgue h^+ e h^- e in tal caso

$$\int h(x)dx = \int h^+(x)dx - \int h^-(x)dx.$$

Nel caso $D = 1$, giova confrontare velocemente l'integrale di Lebesgue con l'integrale (improrio) di Riemann.

Si dice assolutamente integrabile per Riemann una funzione integrabile per Riemann in senso improprio tale per cui sia integrabile per Riemann in senso improprio anche il suo valore assoluto. Esistono funzioni che sono integrabili in senso improprio alle Riemann ma che non sono assolutamente integrabili alla Riemann, ad esempio $\sin(x)/x$ su $(1, +\infty)$ è integrabile in senso improprio alle Riemann ma non è assolutamente integrabile. Si noti che per l'integrale di Lebesgue questo non può accadere, visto che abbiamo definito integrabile una funzione f se e solo se $\int |f(x)|dx < +\infty$. Come conseguenza, una funzione di segno qualunque può essere integrabile in senso improprio per Riemann pur non essendo integrabile per Lebesgue, ad esempio $\sin(x)/x$ su $(1, +\infty)$ non è integrabile alla Lebesgue. Tuttavia *ogni funzione assolutamente integrabile per Riemann è integrabile per Lebesgue.* Al contrario una funzione integrabile per Lebesque può non essere integrabile per Riemann. Ad esempio la funzione $h(x) = \mathbb{I}_{\mathbb{Q}\cap[0,1]}(x)$ non è integrabile per Riemann ma lo è le Lebsque e $\int h(x)dx = 0$ [dimostrarlo].

IMPORTANTE: in tutti i casi pratici gli integrali con i quali avremo a che fare saranno integrali che si possono calcolare ricorrendo all'integrale alla Riemann (eventualmente in senso improprio). Tipicamente dovremo integrare funzioni continue a tratti.

8. Variabili aleatorie assolutamente continue

8.1. **Definizioni e proprietà.** Usando Proposizione 7.2 si dimostra il seguente

Lemma 8.1. *Se $f : \mathbb{R}^D \to [0, +\infty)$ è una funzione boreliana Lebesgue integrabile tale che $\int_{\mathbb{R}^D} f(x_1, \ldots, x_D)dx_1 \cdots dx_D = 1$ allora*

$$A \mapsto P(A) := \int_A f(x_1, \ldots, x_D)dx_1 \cdots dx_D$$

è una misura di probabilità (σ-additiva) su $\mathcal{B}(\mathbb{R}^D)$.

Dimostrazione. Il fatto che $P(A) \geqslant 0$ segue dalla positività dell'integrale, si veda (1) di Proposizione 7.2 considerando $\xi(\omega) = \mathbb{I}_A(\omega)$. L'additività completa da (5) di Proposizione 7.2 considerando $\xi_n(\omega) = \mathbb{I}_{A_n}(\omega)$. ✠

Definizione 8.1. *Un vettore aleatorio* $\mathbf{X} = (X_1, \dots, X_D)$ *ha legge assolutamente continua con densità* $f_{\mathbf{X}}$ *se esiste una funzione positiva integrabile (alla Lebesgue) su* $\mathbb{R}^D$ *tale che*

$$P\{X_1 \leqslant x_1, \dots, X_D \leqslant x_D\} = \int_{\times_{i=1}^D (-\infty, x_i]} f_{\mathbf{X}}(u_1, \dots, u_D) du_1 \cdots du_D$$

(nel senso di Lebesgue) per ogni $(x_1, \dots, x_D)$ *in* $\mathbb{R}^D$.

La funzione $f_{\mathbf{X}}$ è detta *densità* di $\mathbf{X}$. Si noti che non è unica, ad esempio se $f_{\mathbf{X}}$ viene cambiata su un numero finito di punti continua ad essere una densità per $\mathbf{X}$. In effetti $f_{\mathbf{X}}$ è unica a meno di insiemi di misura di Lebesque nulla.

Nel caso speciale $D = 1$, un numero aleatorio è detto con legge assolutamente continua se per ogni $x \in \mathbb{R}$ si ha

$$F_X(x) = \mathbb{P}\{X \leqslant x\} = \int_{-\infty}^{x} f_X(t) dt.$$

Proposizione 8.2. *Se* $\mathbf{X}$ *è un vettore aleatorio con legge assolutamente continua, allora per ogni* A *boreliano si ha*

$$\mathbb{P}\{\mathbf{X} \in A\} = \int_A f_{\mathbf{X}}(x_1, \dots, x_D) dx_1 \cdots dx_D = \int_{\mathbb{R}^D} \mathbb{I}_A(x) f_{\mathbf{X}}(x_1, \dots, x_D) dx_1 \cdots dx_D.$$

Dimostrazione. Si ponga $\tilde{P}(A) = \int_A f_{\mathbf{X}}(x_1, \dots, x_D) dx_1 \cdots dx_D$. Per il Lemma 8.1 $\tilde{P}$ è una misura di probabilità. Da definizione 8.1 si ha che $\tilde{P}(-\infty, x[= P_{\mathbf{X}}(-\infty, x]$ per ogni $x \in \mathbb{R}^D$, quindi per il Teorema 2.2 si ottiene che $\tilde{P}$ e $P_{\mathbf{X}}$ coincidono su tutti i boreliani di $\mathbb{R}^D$, ossia la tesi. ✠

Una conseguenza di Porposizione 7.2 punto 3 e del fatto che $\mathrm{Leb}_D(H) = 0$ se H è un iperpiano di dimensione $k < D$, è la seguente

Proposizione 8.3. *Sia* $\mathbf{X}$ *un vettore aleatorio assolutamente continuo, allora per ogni* $x \in \mathbb{R}^D$ *si ha* $\mathbb{P}\{\mathbf{X} = x\} = 0$. *Più in generale se* H *è un iperpiano di dimensione* $k < D$ *allora* $\mathbb{P}\{\mathbf{X} \in H\} = 0$.

Dalla proposizione precedente segue che se $\mathbf{X} = (X_1, \dots, X_D)$ è un vettore assolutamente continuo, allora per $1 \leqslant i < k \leqslant D$ si ha

$$\mathbb{P}\{X_i = X_k\} = 0.$$

Osservazione. *In dimensione* $D = 1$ *si ha che se* F_X *è continua allora*

$$\mathbb{P}\{X = x\} = F_X(x) - F_X(x^-) = 0.$$

Se X *ha legge assolutamente continua,* $F_X(x) = \int_{-\infty}^{x} f_X(t) dt$ *è una funzione continua (un'altra proprietà dell'integrale di Lebesgue) e ritroviamo quanto affermato in Proposizione 8.3. Precisiamo però che, sebbene sia vero che la funzione di ripartizione di una v.a. assolutamente continua sia continua, non vale il viceversa, ossia esistono funzioni di ripartizione continue che non si possono scrivere come integrali di densità. Le variabili aleatorie per cui* $\mathbb{P}\{X = x\} = 0$ *per ogni* $x \in \mathbb{R}^D$ *sono dette con legge diffusa (o senza parte atomica). Ogni v.a. con legge assolutamente continua è ha legge diffusa, ma come detto esistono v.a. con legge diffusa che non sono assolutamente continue. L'esempio più classico è la cosiddetta legge di Cantor.*

Da momento che per una variabile assolutamente continua X con densità f_X si ha $P\{X = x\} = 0$ per ogni x, non si deve fare l'errore di pensare che $f(x)$ sia uguale $\mathbb{P}\{X = x\}$, ossia $f_X(x) \neq \mathbb{P}\{X = x\}$. Tuttavia, per h piccolo ($h \to 0$) per il teorema della media

$$\mathbb{P}\{x - h \leqslant X \leqslant x + h\} = \int_{x-h}^{x+h} f_X(t) dt \sim 2h f_X(x)$$

ossia per $h \to 0$

$$f_X(x) \sim \frac{1}{2h} \mathbb{P}\{x - h \leqslant X \leqslant x + h\}.$$

8.2. **Esempi.**