Capitolo 4

Spazi di probabilità generali. Variabili casuali assolutamente continue

4.1 σ-algebre. Misure di Probabilità

Con gli spazi di probabilità discreti, abbiamo visto molti aspetti importanti del Calcolo delle Probabilità, senza dover ricorrere a strumenti di Analisi Matematica troppo sofisticati. La teoria fin qui sviluppata non ci permette, però, di affrontare due argomenti fondamentali, sia dal punto di vista teorico che applicativo. Anzitutto la definizione di variabili casuali l'insieme dei cui valori sia non numerabile. Molte grandezze che trattiamo quotidianamente (tempi, masse, lunghezze,...) possono assumere qualunque valore di un'intervallo di $\mathbb{R}$. È ovviamente impossibile, in uno spazio di probabilità discreto, definire una variabile casuale X per cui $X(\Omega)$ sia un intervallo di $\mathbb{R}$, dato che $X(\Omega)$ è sempre al più numerabile. L'altra questione riguarda lo studio delle *successioni* di variabili casuali. Le prime applicazioni "moderne" del calcolo delle probabilità (de Moivre, Laplace), riguardarono il calcolo "approssimato" di certe probabilità. Una formulazione rigorosa di tali approssimazioni conduce a diverse nozioni di *convergenza* di successioni di variabili casuali. Uno spazio di probabilità discreto è troppo "povero" perchè in esso si possano definire successioni interessanti di variabili casuali, ad esempio successioni di variabili casuali indipendenti con la stessa distribuzione.

Risulta quindi naturale cercare di definire spazi di probabilità in cui lo spazio campionario sia non numerabile. Per avere un'idea del tipo di problemi che si affrontano, vediamo due esempi significativi.

Esempio 4.1 Il concetto di probabilità uniforme è chiaro e naturale se lo spazio campionario è finito. È possibile estendere tale nozione ad uno spazio campionario continuo, ad esempio un intervallo limitato di $\mathbb{R}$? In altre parole, dato un intervallo I di $\mathbb{R}$, si può formalizzare l'idea di "scegliere a caso" un punto di I?

Per fissare le idee, sia $I = \Omega = [0, 1]$. Se P è la probabilità "uniforme" che stiamo cercando di definire, è naturale assumere che, se $0 \leq a \leq b \leq 1$ allora

$$P([a, b]) = b - a. \tag{4.1}$$

In tal modo, oltre a verificarsi il fatto che $P(\Omega) = 1$, si ha che la probabilità di un intervallo dipende solo dalla sua lunghezza geometrica, e non dalla sua "posizione" in $[0, 1]$. Si noti che, da (4.1), segue che $P(\{x\}) = P([x, x]) = 0$ per ogni $x \in [0, 1]$. Il problema è: è possibile estendere la P definita in (4.1) a tutti i sottoinsiemi di $[0, 1]$, in modo che l'estensione verifichi la proprietà di σ-additività? Ricordiamo che quasi tutti i risultati concernenti gli spazi di probabilità discreti sono basati sulla σ-additività.

Esempio 4.2 Nell'esempio 1.32 abbiamo costruito un modello probabilistico per N prove ripetute indipendenti per le quali $p \in (0, 1)$ è la probabilità di successo. Pensiamo ora di effettuare una successione infinita di prove ripetute, cioè $N = +\infty$. La scelta naturale per lo spazio campionario è allora

$$\Omega = \{0, 1\}^{\mathbb{N}\setminus\{0\}},$$

cioè, se $\omega \in \Omega$, allora $\omega = (\omega_1, \omega_2, \ldots)$ dove $\omega_i \in \{0, 1\}$. È ben noto che Ω *non* è numerabile. La probabilità P che vogliamo costruire sui sottoinsiemi di Ω dovrà soddisfare un requisito del tutto naturale: se consideriamo un evento che dipende solo dagli esiti delle prime N prove, con $N < +\infty$, allora la sua probabilità dovrà essere uguale a quella calcolata, nell'Esempio 1.32, con N fissato. In altre parole, se $x_1, x_2, \ldots, x_N \in \{0, 1\}$, dovrà essere

$$P(\{\omega \in \Omega : \omega_1 = x_1, \ldots, \omega_N = x_N\}) = p^{\sum_{i=1}^N x_i}(1-p)^{N-\sum_{i=1}^N x_i}. \tag{4.2}$$

Come nell'esempio precedente, il problema è di stabilire se è possibile estendere P a tutti i sottoinsiemi di Ω. Si noti che, se tale estensione (σ-additiva) esiste, allora $P(\{\eta\}) = 0$ per ogni $\eta \in \Omega$. Infatti

$$\{\eta\} = \bigcap_{N \in \mathbb{N}\setminus\{0\}} \{\omega : \omega_1 = \eta_1, \ldots, \omega_N = \eta_N\}.$$

Quest'ultima è l'intersezione di una famiglia decrescente di eventi. Poichè abbiamo assunto che P sia σ-additiva e osservando che la Proposizione 1.10 non utilizza la numerabilità dello spazio campionario, abbiamo

$$P(\{\eta\}) \;=\; \lim_{N \to +\infty} p^{\sum_{i=1}^N \eta_i}\,(1-p)^{N-\sum_{i=1}^N \eta_i}.$$

Osserviamo ora che se $(\sum_{i=1}^N \eta_i) \geq N/2$ si ha $p^{\sum_{i=1}^N \eta_i} \leq p^{N/2}$, mentre se $(\sum_{i=1}^N \eta_i) < N/2$ si ha $(1-p)^{N-\sum_{i=1}^N \eta_i} \leq (1-p)^{N/2}$; in ogni caso

$$p^{\sum_{i=1}^N \eta_i}\,(1-p)^{N-\sum_{i=1}^N \eta_i} \;\leq\; \big(\max\{p, (1-p)\}\big)^{N/2} \;\xrightarrow{N\to\infty}\; 0\,,$$

poiché $p \in (0, 1)$.

In entrambi gli esempi appena visti, la funzione $P(\cdot)$ viene definita dapprima in una famiglia di insiemi "semplici". Si può dimostrare, in entrambi i casi, che P *non* si può estendere a tutto $\mathcal{P}(\Omega)$ in modo che la P estesa risulti σ-additiva. Dobbiamo dunque ridimensionare l'obbiettivo iniziale di estendere P a tutto $\mathcal{P}(\Omega)$. Questo conduce alla seguente definizione.

Definizione 4.3 Sia Ω un insieme, e $\mathcal{A} \subseteq \mathcal{P}(\Omega)$. Diciamo che $\mathcal{A}$ è una *σ-algebra* se

(i) $\emptyset \in \mathcal{A}$.

(ii) Se $A \in \mathcal{A}$ allora $A^c \in \mathcal{A}$.

(iii) Se $(A_n)_{n\geq 0}$ è una successione di elementi di $\mathcal{A}$, allora $\bigcup_n A_n \in \mathcal{A}$.

Si noti che, per (i) e (ii), $\Omega \in \mathcal{A}$. Inoltre $\mathcal{A}$ è chiuso per unione finita (ogni famiglia finita $\{A_1, \ldots, A_n\}$ di elementi di $\mathcal{A}$ può essere completata in una successione ponendo $A_k = \emptyset$ per $k > n$, senza modificarne l'unione). Infine usando l'identità $\bigcap_n A_n = (\bigcup_n A_n^c)^c$, si vede che una σ algebra è chiusa per intersezione, sia di una famiglia finita sia di una successione.

Nel seguito, una coppia $(\Omega, \mathcal{A})$ formata da un insieme e da una σ-algebra di suoi sottoinsiemi verrà chiamata *spazio misurabile.*

Definizione 4.4 Sia $(\Omega, \mathcal{A})$ uno spazio misurabile. Una funzione

$$P : \mathcal{A} \to [0, 1]$$

si dice *probabilità* (o *misura di probabilità*) se valgono le seguenti proprietà:

(P1)
$$P(\Omega) = 1.$$

(P2) (*σ-additività*) Per ogni successione $(A_n)_{n\in\mathbb{N}}$ di elementi di $\mathcal{A}$ a due a due disgiunti, si ha
$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n).$$

In altre parole, la nozione di probabilità su un insieme generale Ω è definita dallo stesso sistema di assiomi della probabilità su spazi discreti, con la differenza che la probabilità è definita su una σ algebra di sottoinsiemi di Ω che non necessariamente coincide con $\mathcal{P}(\Omega)$. La terna $(\Omega, \mathcal{A}, P)$, dove $(\Omega, \mathcal{A})$ è uno spazio misurabile e P è una probabilità su $(\Omega, \mathcal{A})$, verrà chiamata *spazio di probabilità.* In analogia a quanto fatto nel caso discreto, Ω sarà *chiamato spazio campionario* e gli elementi di $\mathcal{A}$ saranno chiamati *eventi.* Esattamente come nel caso discreto, si mostra che in uno spazio di probabilità vale l'additività finita: per ogni $A_1, A_2, \ldots, A_N \in \mathcal{A}$ a due a due disgiunti,

$$P\left(\bigcup_{n=0}^{N} A_n\right) = \sum_{n=0}^{N} P(A_n). \tag{4.3}$$

I risultati contenuti nel paragrafo 1.2 continuano a valere nel contesto più generale or ora introdotto, a patto di restringere i risultati alla σ-algebra degli eventi. Per completezza, rienunciamo i risultati principali, le dimostrazioni dei quali sono identiche a quelle date per spazi di probabilità discreti.

Proposizione 4.5 *Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità e siano $A, B \in \mathcal{A}$. Allora valgono le seguenti proprietà:*

(i)
$$P(A^c) = 1 - P(A).$$

(ii) *Se* $A \subseteq B$ *allora*

$$P(B \setminus A) = P(B) - P(A).$$

In particolare

$$P(A) \leq P(B).$$

(iii)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

In particolare

$$P(A \cup B) \leq P(A) + P(B).$$

Proposizione 4.6 *Sia* $(\Omega, \mathcal{A})$ *uno spazio misurabile, e* $P : \mathcal{A} \to [0,1]$ *una funzione che soddisfa (P1) e l'additività in (4.3). Allora le seguenti proprietà sono equivalenti:*

(a) P *è* σ*-additiva.*

(b) *Per ogni successione crescente di eventi* $(A_n)_{n\geq 1}$ *(tale cioè che* $A_n \subseteq A_{n+1}$ *per ogni* $n \geq 1$*) si ha*

$$P\left(\bigcup_{n\geq 1} A_n\right) = \lim_{n\to+\infty} P(A_n).$$

(c) *Per ogni successione decrescente di eventi* $(A_n)_{n\geq 1}$ *(tale cioè che* $A_n \supseteq A_{n+1}$ *per ogni* $n \geq 1$*) si ha*

$$P\left(\bigcap_{n\geq 1} A_n\right) = \lim_{n\to+\infty} P(A_n).$$

Torniamo ora al problema enunciato negli esempi 4.1 e 4.2. In entrambi i casi avevamo definito una funzione $P : \mathcal{I} \to [0,1]$, dove $\mathcal{I}$ = famiglia degli intervalli in $[0,1]$ nell'Esempio 4.1, e $\mathcal{I}$ = famiglia degli insiemi del tipo $\{\omega : \omega_1 = x_1, \ldots, \omega_N = x_N\}$ nell'Esempio 4.2. È facile vedere che in entrambi i casi $\mathcal{I}$ *non* è una σ-algebra. Il problema è allora di trovare una σ-algebra $\mathcal{A}$ contenente $\mathcal{I}$, e una probabilità su $(\Omega, \mathcal{A})$ che estenda la P originaria.

La scelta della σ-algebra $\mathcal{A}$ si può fare in modo "canonico", grazie al seguente risultato.

Proposizione 4.7 *Sia* Ω *un insieme arbitrario.*

(i) *Se* $\{\mathcal{A}_\alpha : \alpha \in I\}$ *è una famiglia di* σ*-algebre di sottoinsiemi di* Ω *indicizzata da un insieme arbitrario* I*, allora* $\bigcap_{\alpha\in I} \mathcal{A}_\alpha$ *è una* σ*-algebra di sottoinsiemi di* Ω*.*

(ii) *Sia* $\mathcal{I} \subseteq \mathcal{P}(\Omega)$*. Allora esiste una* minima σ*-algebra* $\mathcal{A}$ *contenente* $\mathcal{I}$*, ossia* $\mathcal{I} \subseteq \mathcal{A}$*, e se* $\mathcal{A}'$ *è una* σ*-algebra contenente* $\mathcal{I}$ *allora* $\mathcal{A} \subseteq \mathcal{A}'$*. Tale* σ*-algebra è denotata con* $\sigma(\mathcal{I})$*, e chiamata la* σ-algebra generata da $\mathcal{I}$*.*

Dimostrazione.

(i) La dimostrazione è semplice, ed è lasciata al lettore come esercizio.

(ii) Sia

$$\Xi = \{\mathcal{A} \subseteq \mathcal{P}(\Omega) : \mathcal{I} \subseteq \mathcal{A},\ \mathcal{A} \text{ è una } \sigma\text{-algebra}\}.$$

Notare che $\Xi \neq \emptyset$, essendo $\mathcal{P}(\Omega) \in \Xi$. Per (i),

$$\mathcal{A} = \bigcap_{\mathcal{A}' \in \Xi} \mathcal{A}'$$

è una σ-algebra contenente $\mathcal{I}$, e, per definizione, $\mathcal{A} \subseteq \mathcal{A}'$ per ogni σ-algebra $\mathcal{A}'$ contenente $\mathcal{I}$.

□

Tornando agli Esempi 4.1 e 4.2, si vuole mostrare che esiste una probabilità che estende P definita *almeno* in $\sigma(\mathcal{I})$. Tale problema di estensione è altamente non banale, e al di là degli scopi di questo corso. È possibile dimostrare che, per entrambi gli esempi, esiste effettivamente un'unica probabilità che estende P a $\sigma(\mathcal{I})$. Inoltre, è possibile estendere P ad una σ-algebra più grande di $\sigma(\mathcal{I})$, ma *non* a tutto $\mathcal{P}(\Omega)$.

Osservazione 4.8 Se (Ω, P) è uno spazio di probabilità discreto, allora si può identificare con lo spazio di probabilità $(\Omega, \mathcal{P}(\Omega), P)$.

4.2 Variabili casuali

Definizione 4.9 Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità, e sia $(E, \mathcal{E})$ uno spazio misurabile. Una funzione

$$X : \Omega \to E$$

si dice *variabile casuale* a valori in $(E, \mathcal{E})$ se per ogni $C \in \mathcal{E}$ si ha

$$X^{-1}(C) \in \mathcal{A}.$$

Dunque, se X è una variabile casuale e $C \in \mathcal{E}$, allora $P(X \in C) := P(X^{-1}(C))$ è ben definita.

Definizione 4.10 Se X è una variabile casuale a valori in $(E, \mathcal{E})$, la mappa

$$\begin{array}{rccc} \mu_X : & \mathcal{E} & \to & [0,1] \\ & C & \mapsto & P(X \in C) \end{array}$$

si dice *distribuzione* della variabile casuale X.

La dimostrazione del seguente risultato è lasciata per esercizio al lettore.

Proposizione 4.11 *La distribuzione μ_X di una variabile casuale X a valori in $(E, \mathcal{E})$ è una probabilità su $(E, \mathcal{E})$.*

La nozione di variabile casuale dipende dalla scelta della σ-algebra sull'insieme E. Nel caso in cui $E = \mathbb{R}, \mathbb{R}^d$ o $\mathbb{C}$ sceglieremo come σ-algebra la cosidetta σ-algebra di Borel $\mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}^d), \mathcal{B}(\mathbb{C})$, definita come la σ-algebra generata dai sottoinsiemi aperti. Tale scelta verrà sempre sottointesa nel seguito. Come nel caso discreto, chiameremo *scalari, vettoriali e complesse* le variabili casuali a valori, rispettivamente, in $\mathbb{R}, \mathbb{R}^d, \mathbb{C}$.

In completa analogia col caso discreto, diamo la seguente definizione.

Definizione 4.12 Siano $X_1, X_2, \ldots, X_n$ variabili casuali definite sullo stesso spazio di probabilità $(\Omega, \mathcal{A}, P)$ a valori rispettivamente negli spazi misurabili $(E_1, \mathcal{E}_1)$, $(E_2, \mathcal{E}_2)$, $\ldots$, $(E_n, \mathcal{E}_n)$. Esse si dicono indipendenti se per ogni scelta di $A_1 \in \mathcal{E}_1$, $A_2 \in \mathcal{E}_2$, $\ldots$, $A_n \in \mathcal{E}_n$ si ha

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) \;=\; \prod_{i=1}^{n} P(X_i \in A_i)\,. \tag{4.4}$$

Infine, per una variabile casuale scalare X, definiamo la funzione di ripartizione

$$F_X(x) = P(X \leq x).$$

Le proprietà della funzione di ripartizione contenute nella Proposizione 3.57 continuano a valere nel caso generale, come si vede dal fatto che la dimostrazione della Proposizione 3.57 è corretta anche in spazi di probabilità generali. Inoltre, la stessa dimostrazione usata nella Proposizione 3.58 mostra che

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

Si noti come l'identità precedente, nel caso di variabili casuali discrete, permetta di mostrare che se due variabili casuali hanno la stessa funzione di ripartizione, allora hanno la stessa distribuzione. Tale affermazione è vera anche nel caso generale, ma omettiamo la dimostrazione di questo fatto.

Proposizione 4.13 *Siano X e Y due variabili casuali a valori in $\mathbb{R}$, tali che $F_X = F_Y$. Allora $\mu_X = \mu_Y$.*

4.3 Valor medio (cenni)

La nozione generale di valor medio esula dagli scopi di questo corso. Pertanto, in questo paragrafo, verranno date le definizioni e i risultati principali, senza dettagli e dimostrazioni.

Per garantire la coerenza di quanto vedremo ora con quanto visto per spazi di probabilità discreti, introduciamo ora la nozione di variabile casuale discreta nel contesto più generale.

Definizione 4.14 Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità, $(E, \mathcal{E})$ uno spazio misurabile tale che, per ogni $x \in E$ si abbia $\{x\} \in \mathcal{E}$, e sia X una variabile casuale a valori in $(E, \mathcal{E})$. Diremo che X è una *variabile casuale discreta* se $X(\Omega)$ è finito o numerabile.

Per una variabile casuale discreta è possibile definire la densità

$$p_X(x) = P(X = x) = P(X^{-1}(\{x\})).$$

Se $E = \mathbb{R}$, non ha però alcun senso definire il valor medio di X tramite la formula $E(X) = \sum_{\omega \in \Omega} X(\omega) P(\{\omega\})$: negli esempi 4.1 e 4.2 abbiamo infatti visto che $P(\{\omega\}) = 0$ per ogni $\omega \in \Omega$, e quindi ogni valor medio sarebbe nullo. Tuttavia, avendo a disposizione la densità, è naturale definire $E(X)$ tramite la relazione

$$E(X) := \sum_{x \in X(\omega)} x\, p_X(x)\,, \tag{4.5}$$

se la serie in (4.5) converge assolutamente; in caso contrario diremo che la variabile casuale X non ammette valor medio.

Vediamo ora come si estenda la nozione di valor medio a variabili casuali generali. Nel seguito, se X è una variabile casuale scalare, definiamo $X^+ = \max(X, 0)$ e $X^- = -\min(X, 0)$. Si noti che $X^+, X^- \geq 0$ e $X = X^+ - X^-$.

Definizione 4.15 Sia X una variabile casuale scalare. Se $X \geq 0$, definiamo

$$E(X) = \sup\{E(Y) : 0 \leq Y \leq X,\ Y \text{ è una variabile casuale discreta}\} \in [0, +\infty].$$

(in particolare, $E(X)$ è dato da (4.5) se X è discreta). Per una X generale, diremo che X ammette valor medio se $E(X^+) < +\infty$ e $E(X^-) < +\infty$, e in tal caso

$$E(X) = E(X^+) - E(X^-).$$

Infine, se X è una variabile casuale complessa, posto $X = Re(X) + iIm(X)$, diciamo che X ammette valor medio se $Re(X)$ e $Im(X)$ ammettono entrambe valor medio, e in questo caso poniamo

$$E(X) = E(Re(X)) + iE(Im(X)).$$

Contrariamente al caso discreto, le dimostrazioni di alcune delle proprietà fondamentali del valor medio sono non banali, e non verranno trattate in questo corso.

Proposizione 4.16 *Siano X, Y due variabili casuali scalari o complesse, definite nello stesso spazio di probabilità $(\Omega, \mathcal{A}, P)$. Allora valgono le seguenti proprietà:*

(i) *(Monotonia) Se X, Y sono a valori reali, entrambe ammettono valor medio e $X(\omega) \leq Y(\omega)$, allora $E(X) \leq E(Y)$.*

(ii) *X ammette valor medio se e solo se $|X|$ ammette valor medio, e in tal caso*

$$|E(X)| \leq E(|X|).$$

(iii) *(Linearità) Se X e Y ammettono valor medio e $a, b \in \mathbb{C}$, allora la variabile casuale $aX + bY$ definita da*

$$(aX + bY)(\omega) = aX(\omega) + bY(\omega),$$

ammette valor medio e

$$E(aX + bY) = aE(X) + bE(Y).$$

Sia ora $g : \mathbb{R} \to \mathbb{R}$, e X una variabile casuale a valori in $\mathbb{R}$. In generale non è detto che $g(X) := g \circ X$ sia una variabile casuale. Per averne la garanzia, occorre assumere che g sia una *funzione misurabile*, cioè

- per ogni $A \in \mathcal{B}(\mathbb{R})$ si ha che $g^{-1}(A) \in \mathcal{B}(\mathbb{R})$.

La misurabilità è una proprietà piuttosto debole; ad esempio tutte le funzioni continue a tratti sono misurabili. Si rimanda ad un testo più avanzato per una discussione più approfondita sull'argomento.

Se g è misurabile, ha senso chiedersi se $g(X)$ ammette valor medio. È possibile dimostrare che tanto il fatto che $g(X)$ ammetta valor medio, quanto eventualmente il valore di tale valor medio, dipendono solo dalla distribuzione di X. In altre parole, *se X e Y hanno la stessa distribuzione e g è misurabile, allora $g(X)$ ammette valor medio se e solo se $g(Y)$ ammette valor medio, e in caso affermativo i due valori medi sono uguali.*

Poichè i polinomi sono funzioni continue, e perciò misurabili, analogamente al caso discreto, si possono definire Varianza, Covarianza e momenti e funzione generatrice dei momenti di una variabile casuale. I risultati contenuti nei paragrafi 3.5 e 3.7 sono validi in questo contesto più generale.

4.4 Variabili casuali assolutamente continue

Cominciamo con l'introdurre alcune notazioni, e alcune precisazioni tecniche. In quanto segue daremo per nota la nozione di *integrale di Riemann*

$$\int_a^b f(x)\,dx\,,$$

dove $f : [a, b] \to \mathbb{R}$. Vale tuttavia la pena di ricordare alcuni fatti fondamentali.

1. L'integrale "proprio" di Riemann viene definito per una classe di funzioni limitate, dette *Riemann-integrabili* (talvolta diremo, semplicemente, integrabili). Tutte le funzioni $f : [a, b] \to \mathbb{R}$ limitate tali che per ogni $y \in [a, b]$ i limiti destro e sinistro

$$\lim_{x\downarrow y} f(x)\,, \qquad \lim_{x\uparrow y} f(x)\,,$$

   esistono finiti (solo il primo se $y = a$ e solo il secondo se $y = b$), sono Riemann-integrabili. In particolare sono integrabili le funzioni continue e quelle monotone.

2. Sia $f : [a, b] \to \mathbb{R}$ una funzione *continua a tratti*: supponiamo cioè che esista un sottoinsieme finito $N \subseteq [a, b]$ tale che f è continua in ogni punto $x \in [a, b]\setminus N$. Se inoltre i limiti destro e sinistro di f esistono finiti nei punti di N, allora f è Rieman-integrabile e, definita la funzione integrale $F : [a, b] \to \mathbb{R}$ mediante

$$F(x) := \int_a^x f(t)dt\,,$$

   si ha che F è continua su $[a, b]$ e derivabile in ogni punto di $[a, b] \setminus N$, con

$$F'(x) \;=\; f(x)\,, \qquad \forall x \in [a, b] \setminus N\,.$$

   Viceversa, sia $F : [a, b] \to \mathbb{R}$ un funzione $\mathcal{C}^1$ *a tratti*: supponiamo cioè che F sia continua su $[a, b]$ e che esista un sottoinsieme finito $N \subseteq [a, b]$ tale che F sia derivabile in ogni punto di $[a, b] \setminus N$, e che F' sia continua in ogni punto di $[a, b] \setminus N$. Assumiamo inoltre che per ogni $y \in N$ esistano finiti i limiti

$$\lim_{x\downarrow y} F'(x)\,, \qquad \lim_{x\uparrow y} F'(x)\,.$$

Allora se definiamo

$$f(x) := \begin{cases} F'(x) & \text{se } x \in [a,b] \setminus N \\ \text{un valore arbitrario} & \text{altrimenti,} \end{cases}$$

si ha che f è Riemann-integrabile e vale inoltre la relazione

$$F(x) = \int_a^x f(t)dt\,, \qquad \forall x \in [a,b]\,.$$

I risultati appena enunciati costituiscono il *Teorema fondamentale del calcolo integrale.*

3. Si possono definire varie forme "generalizzate" di integrale di Riemann. Ad esempio se $f : (a,b] \to \mathbb{R}$ è Riemann integrabile su ogni intervallo $[a+\epsilon, b]$ con $\epsilon \in (0, b-a)$, e il limite

$$\lim_{\epsilon \downarrow 0} \int_{a+\epsilon}^b f(x)dx$$

esiste finito, diciamo che f è *integrabile in senso generalizzato* su $[a,b]$, e il suo integrale

$$\int_a^b f(x)dx$$

è dato dal limite precedente. In modo analogo definiamo l'integrale generalizzato di $f : [a,b) \to \mathbb{R}$. Se f è integrabile in senso generalizzato su $[a,c]$ e su $[c,b]$ diremo che è integrabile in senso generalizzato su $[a,b]$ e poniamo

$$\int_a^b f := \int_a^c f + \int_c^b f.$$

In altre parole, con questa definizione otteniamo, quando possibile, l'integrale di una funzione $f : [a,b] \setminus \{c\} \to \mathbb{R}$ che, ad esempio, non ammetta limite finito in c. Questo procedimento si estende in modo evidente a funzioni $f : [a,b] \setminus N \to \mathbb{R}$, dove N è un insieme finito.

Sia ora $f : [a, +\infty) \to \mathbb{R}$ una funzione che assumiamo Riemann-integrabile (anche nel senso generalizzato appena visto) su ogni intervallo $[a,c]$, e per cui il limite

$$\lim_{c \to +\infty} \int_a^c f(x)dx$$

esiste finito. Allora diciamo che f è integrabile in senso generalizzato su $[a, +\infty)$ e poniamo

$$\int_a^{+\infty} f(x)dx := \lim_{c \to +\infty} \int_a^c f(x)dx.$$

In modo analogo definiamo l'integrabilità su semirette del tipo $(-\infty, b]$. Infine, se $f : \mathbb{R} \to \mathbb{R}$ è integrabile su $(-\infty, a]$ e su $[a, +\infty)$ allora diciamo che f è integrabile in senso generalizzato su $\mathbb{R}$ e poniamo

$$\int_{-\infty}^{+\infty} f(x)dx := \int_{-\infty}^a f(x)dx + \int_a^{+\infty} f(x)dx.$$

4. Non è difficile fornire generalizzazioni del Teorema fondamentale del calcolo integrale a funzioni integrabili nel senso generalizzato appena descritto. Ad esempio, se $f : \mathbb{R} \to \mathbb{R}$ è *continua a tratti* ed è integrabile in senso generalizzato su $\mathbb{R}$, allora, posto

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

si ha che $F'(x) = f(x)$ per ogni x in cui f è continua. Viceversa, se $F : \mathbb{R} \to \mathbb{R}$ è una funzione $\mathcal{C}^1$ *a tratti*, allora la funzione F' (definita in modo arbitrario nei punti in cui F non è derivabile) è integrabile su $\mathbb{R}$ in senso generalizzato e, per ogni $x \in \mathbb{R}$, si ha

$$F(x) = \int_{-\infty}^{x} F'(t)dt.$$

Definizione 4.17 Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità, e $X : \Omega \to \mathbb{R}$ una variabile casuale. Diciamo che X è assolutamente continua se esiste $f_X : \mathbb{R} \in [0, +\infty)$ integrabile su $\mathbb{R}$ in senso generalizzato tale che, se $F_X(x) := P(X \leq x)$ è la funzione di ripartizione di X, si ha

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt. \tag{4.6}$$

Una tale funzione f_X viene detta *densità* di X.

Osservazione 4.18 Prendendo il limite $x \to \infty$ nella relazione (4.6) e usando la continuità dal basso della probabilità, si ottiene che la densità f_X soddisfa la seguente relazione:

$$\int_{-\infty}^{+\infty} f_X(x)\, dx \;=\; 1\,. \tag{4.7}$$

Osservazione 4.19 Se X è una variabile casuale assolutamente continua, la definizione 4.17 non identifica la sua densità f_X in modo unico. Infatti, se f_X è *una densità* di X, ogni funzione g per cui

$$\int_{-\infty}^{x} g(t)dt = \int_{-\infty}^{x} f_X(t)dt \tag{4.8}$$

per ogni $x \in \mathbb{R}$, è una densità di X. Ad esempio, se g è ottenuta da f_X modificandone il valore in un numero finito di punti, allora (4.8) vale. Questa "ambiguità" nella nozione di densità di una variabile casuale assolutamente continua non porterà tuttavia alcun problema. Spesso diremo, impropriamente, che una certa funzione f è *la densità* di X.

Dalla definizione 4.17 seguono facilmente alcune proprietà delle variabili casuali assolutamente continue, che raccogliamo nelle seguenti osservazioni.

Osservazioni 4.20 1. Come abbiamo visto, la relazione

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

vale per tutte le variabili casuali a valori in $\mathbb{R}$. Se X è assolutamente continua, la sua funzione di ripartizione

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

è, per il Teorema fondamentale del calcolo integrale, una funzione continua, e pertanto $F_X(x) = F_X(x^-)$ per ogni $x \in \mathbb{R}$. Quindi, se X è assolutamente continua,

$$P(X = x) = 0$$

per ogni $x \in \mathbb{R}$.

2. Per ogni $a, b \in [-\infty, +\infty]$, $a \leq b$, si ha

$$P(X \in (a, b]) = P(X \leq b) - P(X \leq a) = \int_a^b f_X(t)dt.$$

Si noti che

$$P(X \in [a, b]) = P(X \in (a, b]) + P(X = a) = P(X \in (a, b]),$$

per quanto abbiamo appena visto. In modo analogo, si mostra che

$$P(X \in [a, b]) = P(X \in (a, b]) = P(X \in [a, b)) = P(X \in (a, b)) = \int_a^b f_X(t)dt.$$

3. Sia X una variabile casuale a valori in $\mathbb{R}$ tale che la funzione di ripartizione F_X sia $\mathcal{C}^1$ a tratti. Dal Teorema fondamentale del calcolo integrale, richiamato in precedenza, segue che X è una variabile casuale assolutamente continua, e ogni funzione f tale che $f(x) = F'_X(x)$ per ogni x in cui F'_X è continua è una densità di X.

Enunciamo ora il seguente risultato fondamentale, omettendo la dimostrazione, che richiede tecniche più avanzate.

Teorema 4.21 *Sia X una variabile casuale assolutamente continua con densità f_X. Sia inoltre $g : \mathbb{R} \to \mathbb{R}$ una funzione misurabile tale che $g(x)f_X(x)$ e $|g(x)|f_X(x)$ siano integrabili in senso generalizzato su $\mathbb{R}$. Allora $g(X)$ ammette valor medio e*

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x)f_X(x)dx.$$

Enunciamo ora senza dimostrazione un altro risultato, che rappresenta l'analogo della formula (3.44), dimostrata per le variabili casuali discrete.

Teorema 4.22 *In uno stesso spazio di probabilità $(\Omega, \mathcal{A}, P)$, siano definite X e Y due variabili casuali assolutamente continue, indipendenti, con densità rispettivamente f_X e f_Y. Allora, per ogni $z \in \mathbb{R}$, le funzioni $x \mapsto f_X(x)f_Y(z-x)$ e $x \mapsto f_X(z-x)f_Y(x)$ sono integrabili in senso generalizzato su $\mathbb{R}$. Inoltre la variabile casuale $X + Y$ è assolutamente continua, e una sua densità è data da*

$$f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx = \int_{-\infty}^{+\infty} f_X(z-x)f_Y(x)dx.$$

4.5 Esempi di variabili casuali assolutamente continue

4.5.1 Variabili casuali uniformi

Siano $a, b \in \mathbb{R}$ con $a < b$. Una variabile casuale assolutamente continua X a valori in $\mathbb{R}$ si dice *uniforme* su $[a, b]$, e scriviamo

$$X \sim U(a, b),$$

se

$$f_X(x) = \frac{1}{b-a} 1_{[a,b]}(x).$$

È naturale interpretare una variabile casuale $X \sim U(a, b)$ come "un punto scelto a caso su $[a, b]$". Le variabili uniformi possono essere "simulate" al calcolatore, da speciali programmi chiamati *generatori di numeri casuali*. Essi sono alla base di numerosi algoritmi, detti *algoritmi stocastici*.

Calcoliamo ora media e varianza di $X \sim U(a, b)$.

$$E(X) = \frac{1}{b-a} \int_a^b x dx = \frac{a+b}{2}.$$

$$E(X^2) = \frac{1}{b-a} \int_a^b x^2 dx = \frac{a^2 + ab + b^2}{3},$$

da cui

$$Var(X) = E(X^2) - E^2(X) = \frac{(b-a)^2}{12}.$$

La funzione generatrice dei momenti vale

$$\gamma_X(t) \;=\; E(e^{tX}) \;=\; \frac{1}{b-a} \int_a^b e^{tx}\, dx \;=\; \frac{e^{tb} - e^{ta}}{t(b-a)}\,.$$

Le variabili casuali uniformi possono essere usate come ingredienti per "costruire" variabili casuali a valori in $\mathbb{R}$ con distribuzione arbitraria. Infatti, come mostra il seguente risultato, data una qualunque variabile casuale X a valori in $\mathbb{R}$, esiste una funzione misurabile $f : [0, 1] \to \mathbb{R}$ tale che, se $Y \sim U(0, 1)$, X e $f(Y)$ hanno la stessa distribuzione.

Proposizione 4.23 *Sia X una variabile casuale a valori in $\mathbb{R}$, e sia F_X la sua funzione di ripartizione. Consideriamo la funzione $f : \mathbb{R} \to \mathbb{R}$ così definita*

$$f(x) := \begin{cases} 0 & \text{se } x \le 0 \text{ oppure } x \ge 1 \\ \inf\{z : F_X(z) \ge x\} & \text{se } x \in (0,1) \end{cases}.$$

Allora f è una funzione misurabile e, se $Y \sim U(0, 1)$, le variabili casuali X e $f(Y)$ hanno la stessa distribuzione.

Dimostrazione. Omettiamo la dimostrazione della misurabilità di f. Sia $Z := f(Y)$. Pe mostrare che X e Z hanno la stessa distribuzione, è sufficiente mostrare che $F_X = F_Z$. Sia allora $x \in \mathbb{R}$ arbitrario, e mostriamo che

$$F_X(x) = F_Z(x), \tag{4.9}$$

cioè

$$P(X \leq x) = P(Z \leq x).$$

Per dimostrarlo, consideriamo $y \in (0,1)$. Notare che

$$f(y) \leq x \iff \inf\{z : F_X(z) \geq y\} \leq x.$$

Poiché F_X è crescente

$$\inf\{z : F_X(z) \geq y\} \leq x \iff F_X(x) \geq y.$$

Poiché $P(Y \in (0,1)) = 1$, abbiamo allora che

$$\begin{aligned} F_Z(x) = P(f(Y) \leq x) = P(\{f(Y) \leq x\} \cap \{Y \in (0,1)\}) &= P(\{F_X(x) \geq Y\} \cap \{Y \in (0,1)\}) \\ &= P(F_X(x) \geq Y) = F_X(x), \end{aligned}$$

dove abbiamo usato il semplice fatto che, se $Y \sim U(0,1)$ e $c \in [0,1]$, allora $P(Y \leq c) = c$. □

4.5.2 Variabili casuali Gamma. Variabili casuali esponenziali

Cominciamo col ricordare la definizione della funzione Gamma di Eulero:

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx,$$

ben definita per $\alpha > 0$ (perché?). Il valore di $\Gamma(\alpha)$ è noto esplicitamente per vari valori di α. Infatti, osservando che

$$\Gamma(1) = \int_0^{+\infty} e^{-x} dx = 1$$

e che, integrando per parti

$$\Gamma(\alpha+1) = \int_0^{+\infty} x^{\alpha} e^{-x} dx = -x^{\alpha} e^{-x} \Big|_0^{+\infty} + \alpha \int_0^{+\infty} x^{\alpha-1} e^{-x} dx = \alpha\Gamma(\alpha),$$

si ha

$$\Gamma(n) = (n-1)!$$

per $n \in \mathbb{N} \setminus \{0\}$. Inoltre, se $\alpha = \frac{1}{2}$, usando il cambiamento di variabile $x = y^2$,

$$\Gamma(1/2) = \int_0^{+\infty} x^{-1/2} e^{-x} dx = 2 \int_0^{+\infty} e^{-y^2} dy = \int_{-\infty}^{+\infty} e^{-y^2} dy = \sqrt{\pi},$$

ove si è usato il noto valore dell'integrale di Gauss

$$\int_{-\infty}^{+\infty} e^{-y^2} dy = \sqrt{\pi}.$$

Dunque, usando ancora la relazione $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, si trova, per $n \in \mathbb{N} \setminus \{0\}$

$$\Gamma\left(n + \frac{1}{2}\right) = \sqrt{\pi} \prod_{k=0}^{n-1} \left(k + \frac{1}{2}\right).$$

Notiamo anche che se nella definizione di Γ operiamo il cambio di variabili $x = \lambda y$, con $\lambda > 0$, otteniamo

$$\Gamma(\alpha) = \int_0^{+\infty} \lambda^\alpha y^{\alpha-1} e^{-\lambda y} dy,$$

ossia

$$\int_0^{+\infty} \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x} dx = 1.$$

Ciò garantisce la bontà della seguente definizione. Una variabile casuale scalare assolutamente continua X è detta variabile casuale Gamma di parametri $\alpha, \lambda > 0$, se

$$f_X(x) = \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x} 1_{[0,+\infty)}(x).$$

In tal caso scriveremo $X \sim \Gamma(\alpha, \lambda)$. Si ha, se $X \sim \Gamma(\alpha, \lambda)$,

$$E(X) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} x^\alpha e^{-\lambda x} dx = \frac{1}{\lambda \Gamma(\alpha)} \int_0^{+\infty} \lambda^{\alpha+1} x^\alpha e^{-\lambda x} dx = \frac{1}{\lambda \Gamma(\alpha)} \Gamma(\alpha+1) = \frac{\alpha}{\lambda},$$

ricordando che $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. Analogamente

$$E(X^2) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} x^{\alpha+1} e^{-\lambda x} dx = \frac{1}{\lambda^2 \Gamma(\alpha)} \int_0^{+\infty} \lambda^{\alpha+2} x^{\alpha+1} e^{-\lambda x} dx = \frac{1}{\lambda^2 \Gamma(\alpha)} \Gamma(\alpha+2) = \frac{\alpha(\alpha+1)}{\lambda^2},$$

da cui, facilmente,

$$Var(X) = \frac{\alpha}{\lambda^2}.$$

Tutti questi calcoli si sarebbero potuti fare facilmente calcolando la funzione generatrice dei momenti. Infatti

$$\gamma_X(t) = \int_{-\infty}^{+\infty} e^{tx} f_X(x) dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} x^{\alpha-1} e^{-(\lambda-t)x} dx$$

Quest'ultimo integrale è finito solo per $t < \lambda$, e in tal caso

$$\int_0^{+\infty} x^{\alpha-1} e^{-(\lambda-t)x} dx = \frac{\Gamma(\alpha)}{(\lambda-t)^\alpha},$$

e quindi

$$\gamma_X(t) = \begin{cases} \left(\frac{\lambda}{\lambda-t}\right)^\alpha & \text{per } t < \lambda \\ +\infty & \text{altrimenti.} \end{cases}$$

Il seguente risultato illustra una proprietà molto importante delle variabili casuali Gamma.

Proposizione 4.24 *Siano $X \sim \Gamma(\alpha, \lambda)$ e $Y \sim \Gamma(\beta, \lambda)$ indipendenti. Allora $X + Y \sim \Gamma(\alpha + \beta, \lambda)$.*

Dimostrazione. Usando il Teorema (4.22), per $z > 0$ abbiamo:

$$\begin{aligned} f_{X+Y}(z) &= \int f_X(z-x) f_Y(x) dx \\ &= \frac{\lambda^\alpha \lambda^\beta}{\Gamma(\alpha)\Gamma(\beta)} \int 1_{(0,+\infty)}(z-x)(z-x)^{\alpha-1} e^{-\lambda(z-x)} 1_{(0,+\infty)}(x) x^{\beta-1} e^{-\lambda x} dx \\ &= \frac{\lambda^\alpha \lambda^\beta}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda z} \int_0^z (z-x)^{\alpha-1} x^{\beta-1} dx = \frac{\lambda^\alpha \lambda^\beta}{\Gamma(\alpha)\Gamma(\beta)} e^{-\lambda z} z^{\alpha+\beta-1} \int_0^1 (1-t)^{\alpha-1} t^{\beta-1} dt, \end{aligned}$$

dove nell'ultimo passaggio abbiamo usato il cambio di variabile $x = zt$. Ne segue che $f_{X+Y}(z)$ è proporzionale a $z^{\alpha+\beta-1} e^{-\lambda z}$, e quindi è necessariamente la densità di una $\Gamma(\alpha+\beta, \lambda)$. □

Osservazione 4.25 Si noti che dalla dimostrazione precedente segue la non ovvia identità

$$\Gamma(\alpha+\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\int_0^1 (1-t)^{\alpha-1} t^{\beta-1} dt}.$$

Di grande interesse applicativo sono alcuni casi particolari di variabili casuali Gamma. Se $X \sim \Gamma(1, \lambda)$, diciamo che X è una variabile casuale esponenziale di parametro λ, e scriveremo

$$X \sim Exp(\lambda).$$

Si noti che in questo caso

$$f_X(x) = \lambda e^{-\lambda x} 1_{[0,+\infty)}(x).$$

Tali variabili casuali si possono interpretare come l'analogo continuo delle variabili casuali Geometriche. Esse, infatti, soddisfano ad una proprietà di perdita di memoria del tutto analoga a quella delle variabili casuali Geometriche.

Proposizione 4.26 *Sia $X \sim Exp(\lambda)$. Allora per ogni $s, t > 0$*

$$P(X \geq s+t | X > s) = P(X > t).$$

Dimostrazione. Osservando che

$$P(X \geq s+t | X > s) = \frac{P(X \geq s+t)}{P(X \geq s)},$$

e che

$$P(X \geq s) = \int_s^{+\infty} e^{-\lambda x} dx = e^{-\lambda s},$$

il risultato desiderato segue immediatamente. □

Come per le variabili casuali Geometriche, si può dimostrare che le variabili casuali esponenziali sono le uniche variabili casuali assolutamente continue a valori reali positivi per cui vale la precedente proprietà di perdita di memoria. La dimostrazione di tale fatto è omessa. Le variabili casuali esponenziali sono normalmente usate come modelli per "tempi di attesa": tempo di decadimento di atomi radioattivi, tempo intercorrente tra due terremoti successivi, tempo intercorrente tra l'arrivo di due clienti ad uno sportello, ...

Un'altra classe di variabili Gamma che hanno grande rilevanza soprattutto in statistica sono le variabili χ^2. Se $n \geq 1$ è un numero naturale, e $X \sim \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$, diciamo che X è una variabile casuale χ^2 con n *gradi di libertà*, e scriviamo $X \sim \chi^2(n)$. Vedremo più avanti un esempio in cui appaiono variabili casuali di questo tipo.

4.5.3 Variabili casuali Normali o Gaussiane

Una variabile casuale scalare assolutamente continua X si dice Normale o Gaussiana *standard*, e si scrive $X \sim N(0,1)$, se

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Notare che si tratta di una buona definizione dato che

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}.$$

È facile vedere che, se $X \sim N(0,1)$

$$E(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{x^2}{2}} dx = 0$$

dato che l'integrando è una funzione dispari (e Riemann-integrabile in senso generalizzato). Inoltre, integrando per parti,

$$E(X^2) = Var(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \left[-x e^{-\frac{x^2}{2}} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \right] = 1.$$

Sia ora $Y \sim N(0,1)$, $\mu \in \mathbb{R}$ e $\sigma > 0$, e definiamo

$$X = \sigma Y + \mu.$$

Dalle proprietà elementari di media e varianza, si vede che

$$E(X) = \mu, \ Var(X) = \sigma^2.$$

Per determinare la densità di X procediamo come segue.

$$F_X(x) = P(X \leq x) = P(\sigma Y + \mu \geq x) = P\left(Y \leq \frac{x-\mu}{\sigma}\right) = F_Y\left(\frac{x-\mu}{\sigma}\right).$$

Essendo Y una variabile casuale assolutamente continua con densità continua, F_Y, e dunque F_X, è di classe $\mathcal{C}^1$. Perciò

$$f_X(x) = F_X'(x) = \frac{1}{\sigma} F_Y'\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma} f_Y\left(\frac{x-\mu}{\sigma}\right)$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{4.10}$$

Una variabile casuale la cui densità è data da (4.10) si dice Normale o Gaussiana di media μ e varianza σ^2, e si scrive

$$X \sim N(\mu, \sigma^2).$$

Passando per la funzione di ripartizione come nell'argomento precedente, si dimostra (esercizio!) che se $X \sim N(\mu, \sigma^2)$ e $a, b \in \mathbb{R}$, $a \neq 0$, allora

$$aX + b \sim N(a\mu + b, a^2\sigma^2).$$

In particolare

$$\frac{X-\mu}{\sigma} \sim N(0,1).$$

Le variabili casuali Normali sono senz'altro le più usate nelle applicazioni. Detto in modo grossolano, questo è perchè ogni qual volta una quantità aleatoria è la somma di molte quantità aleatorie indipendenti, allora la sua distribuzione è approssimativamente Gaussiana. Una versione parziale, ma rigorosa, di tale affermazione, verrà data nel prossimo capitolo, con il Teorema Limite Centrale.

Anche per le variabili casuali Gaussiane è facile calcolare la funzione generatrice dei momenti. Sia $X \sim N(0,1)$.

$$\gamma_X(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{tx-\frac{1}{2}x^2} dx.$$

L'integrale precedente viene calcolato usando il metodo del *completamento dei quadrati*, molto utile per integrali gaussiani. Si tratta di osservare che

$$tx - \frac{1}{2}x^2 = -\frac{1}{2}(x-t)^2 + \frac{t^2}{2}.$$

Pertanto

$$\gamma_X(t) = e^{\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x-t)^2} dx.$$

Si osservi che $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2}$ è la densità di una variabile casuale Gaussiana di media t e varianza 1, e quindi il suo integrale è 1. Ne segue che

$$\gamma_X(t) = e^{\frac{t^2}{2}}.$$

Per calcolare la funzione generatrice di una generica Gaussiana si procede come segue. Sia $Y \sim N(\mu, \sigma^2)$. Da quanto sopra osservato, Y si può scrivere nella forma $Y = \sigma Z + \mu$, con $Z := (Y-\mu)/\sigma \sim N(0,1)$. Ma allora

$$\gamma_Y(t) = E\left(e^{t(\sigma Z+\mu)}\right) = e^{t\mu} E\left(e^{t\sigma Z}\right) = e^{t\mu}\gamma_Z(t\sigma) = e^{t\mu} e^{\frac{t^2\sigma^2}{2}}.$$

La proprietà enunciata nella seguente proposizione è di estrema importanza.

Proposizione 4.27 *Siano $X \sim N(\mu_1, \sigma_1^2)$ e $Y \sim N(\mu_2, \sigma_2^2)$ indipendenti. Allora*

$$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

Dimostrazione. Iniziamo col supporre $\mu_1 = \mu_2 = 0$. Dal Teorema (4.22), si ha

$$\begin{aligned}
f_{X+Y}(z) &= \int f_X(z-x) f_Y(x) dx \\
&= \frac{1}{2\pi\sigma_1\sigma_2} \int \exp\left[-\frac{1}{2\sigma_1^2}(z-x)^2\right] \exp\left[-\frac{1}{2\sigma_2^2}x^2\right] dx \\
&= \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[\frac{1}{2(\sigma_1^2+\sigma_2^2)}z^2\right] \int \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)(x-\xi)^2\right] dx,
\end{aligned}$$

dove $\xi = \frac{\sigma_2}{\sigma_1\sqrt{\sigma_1^2+\sigma_2^2}} z$. Utilizzando il fatto che

$$\int \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)(x-\xi)^2\right] dx = \sqrt{2\pi\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)^{-1}},$$

con un po' di calcoli si trova

$$f_{X+Y}(z) = \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} \exp\left[-\frac{1}{2(\sigma_1^2+\sigma_2^2)} z^2\right],$$

cioè $X+Y \sim N(0, \sigma_1^2+\sigma_2^2)$. La proposizione è dunque dimostrata nel caso $\mu_1 = \mu_2 = 0$.

In generale, per quanto appena visto, abbiamo che $(X-\mu_1)+(Y-\mu_2) \sim N(0, \sigma_1^2+\sigma_2^2)$, da cui segue che $X+Y = (X-\mu_1)+(Y-\mu_2)+\mu_1+\mu_2 \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

□

4.6 Calcoli con densità: alcuni esempi

Esempio 4.28 Sia $X \sim N(0,1)$ e $Y = X^2$. Allora

$$F_Y(x) = 0 \quad \text{se } x < 0.$$

Se invece $x \geq 0$

$$F_Y(x) = P(X^2 \leq x) = P(-\sqrt{x} \leq X \leq \sqrt{x}) = P(X \leq \sqrt{x}) - P(X < -\sqrt{x})$$
$$= F_X(\sqrt{x}) - F_X(-\sqrt{x}),$$

dove si è usato il fatto che

$$P(X < -\sqrt{x}) = P(X \leq -\sqrt{x}) = F_X(-\sqrt{x})$$

(perchè?). Dunque F_Y è $\mathcal{C}^1$ su $\mathbb{R} \setminus \{0\}$, continua su $\mathbb{R}$, e dunque

$$f_Y(x) = F_Y'(x) = \frac{1}{2\sqrt{x}}\left[F_X'(\sqrt{x}) + F_X'(-\sqrt{x})\right]$$

$$= \frac{1}{2\sqrt{x}}\left[f_X(\sqrt{x}) + f_X(-\sqrt{x})\right] 1_{[0,+\infty)}(x) = \frac{1}{\sqrt{2x\pi}}\, e^{-x/2} 1_{[0,+\infty)}(x),$$

ossia $Y \sim \chi^2(1) = \Gamma(\frac{1}{2}, \frac{1}{2})$.

Esempio 4.29 Siano $X_1, X_2, \ldots, X_n \sim N(0,1)$ indipendenti. Dall'Esempio 4.28 e dalla Proposizione 4.24, segue che

$$X_1^2 + X_2^2 + \cdots + X_n^2 \sim \Gamma\left(\frac{n}{2}, \frac{1}{2}\right) = \chi^2(n).$$

Esempio 4.30 Siano $X, Y \sim Exp(\lambda)$ indipendenti. Definiamo $Z = \max(X, Y)$ e $W = \min(X, Y)$. Vogliamo determinare la distribuzione di Z e W. Quanto visto nel caso discreto nella Proposizione 3.60 riguardo alla funzione di ripartizione del massimo e del minimo di variabili casuali indipendenti continua a valere per variabili casuali generiche, in quanto non veniva usata la discretezza dello spazio di probabilità. Perciò

$$F_Z(x) = F_X(x)F_Y(x) = F_X^2(x), \quad F_W(x) = 1 - (1 - F_X(x))(1 - F_Y(x)) = 1 - (1 - F_X(x))^2.$$

Essendo

$$F_X(x) = F_Y(x) = [1 - e^{-\lambda x}]1_{[0,+\infty)}(x),$$

abbiamo che F_Z e F_W sono di classe $\mathcal{C}^1$ a tratti, e dunque

$$f_Z(x) = 2F_X(x)f_X(x) = 2\lambda(1 - e^{-\lambda x})e^{-\lambda x}1_{[0,+\infty)}(x),$$

e

$$f_W(x) = 2(1 - F_X(x))f_X(x) = 2\lambda e^{-2\lambda x}1_{[0,+\infty)}(x).$$

In particolare, $W \sim Exp(2\lambda)$.