2.7 Il valor medio

La nozione di *media aritmetica* di un insieme finito di numeri reali $\{x_1, x_2, \ldots, x_n\}$ è nota e molto naturale. Una delle sue possibili interpretazioni è quella che si ottiene associando ad ogni x_i un punto materiale; tali punti materiali vengono posizionati su una retta, ognuno nel punto corrispondente alla coordinata x_i. Se tali punti materiali hanno tutti la stessa massa, il punto di coordinate $\frac{1}{n}\sum_1^n x_i$ è il baricentro di tale sistema di punti materiali. Nel caso in cui i punti non abbiano tutti la stessa massa, il baricentro si ottiene attraverso una media "pesata": se m_i è la massa del punto materiale in x_i, scegliendo l'unità di misura per la massa in modo tale che $\sum_1^n m_i = 1$, il baricentro del sistema ha coordinata

$$\mu = \sum_{i=1}^{n} m_i x_i.$$

In ambito probabilistico, la nozione di **valor medio** corrisponde alla nozione di baricentro, una volta interpretate le x_i come i valori assunti da una variabile aleatoria, e le m_i come le rispettive probabilità che i corrispondenti valori vengano assunti. Questo ci conduce alla seguente definizione.

Definizione 2.7.1 Sia X una variabile aleatoria reale discreta definita su uno spazio di probabilità $(\Omega, \mathcal{A}, P)$ e a valori in $E \subset \mathbb{R}$. Si consideri la somma

$$\sum_{x \in E} |x| P\{X = x\} = \sum_{x \in E} |x| p_X(x). \tag{2.7.1}$$

Se tale somma ha un valore finito allora diremo che la variabile aleatoria X ammette **valor medio**. In tal caso, la quantità

$$E[X] := \sum_{x \in E} x P\{X = x\} = \sum_{x \in E} x p_X(x) \tag{2.7.2}$$

si dice **valore medio** della variabile aleatoria X.

Alcuni sinonimi di valore medio sono **media**, **valore atteso**, **speranza matematica** (o più semplicemente **speranza**).

Osservazione 2.7.1 Nel caso in cui E sia un insieme finito, la somma (2.7.1) ha, ovviamente, valore finito. Nel caso in cui E sia numerabile, la condizione di finitezza della serie in (2.7.1) corrisponde all'assoluta convergenza della serie in (2.7.2).

Esempio 2.7.2 (variabile aleatoria di Bernoulli) Prendiamo $X \sim Be(p)$ e calcoliamone la speranza. In base alla definizione, abbiamo:

$$E[X] = 0 \cdot P\{X = 0\} + 1 \cdot P\{X = 1\} = 0 \cdot (1 - p) + 1 \cdot p = p$$

Osservazione 2.7.3 Se X è una variabile aleatoria che assume solo valori reali positivi, la somma (2.7.2) è sempre definita, anche se può assumere il valore $+\infty$. In questo caso denoteremo con $E[X]$ il valore della somma, anche quando questo è $+\infty$. Con questa convenzione, per ogni variabile aleatoria a valori reali o complessi, la somma in (2.7.1) è $E[|X|]$. Dunque, talvolta scriveremo "$E[|X|] < +\infty$" in luogo di "la variabile X ammette valor medio". Si noti anche che, dalla definizione, X ammette valor medio se e solo se $|X|$ ammette valor medio.

Esempio 2.7.4 (variabile aleatoria di Poisson) Prendiamo $X \sim Po(\lambda)$ e calcoliamone la speranza. Per la definizione di speranza e per quanto appena visto nell'ultima nota, abbiamo

$$E[|X|] = E[X] = \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda} \lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda < +\infty$$

e quindi X ammette speranza finita e uguale a λ.

Le seguenti proprietà formali del valor medio derivano immediatamente dalla precedente definizione.

Proposizione 2.7.2 *Siano X, Y due variabili aleatorie reali discrete definite sullo spazio di probabilità $(\Omega, \mathcal{A}, P)$ e che ammettano valore medio. Allora valgono le seguenti proprietà:*

1. *(Monotonia) Se $P\{X \geq 0\} = 1$, allora $E[X] \geq 0$.*
2. *Se X e Y sono i.d., allora $E[X] = E[Y]$.*
3. $$|E[X]| \leq E[|X|].$$
4. *(Omogeneità di grado 1) Se $a \in \mathbb{R}$, allora aX ammette valor medio e $E[aX] = aE[X]$.*
5. *Se $P\{X = c\} = 1$, allora $E[X] = c$.*
6. *(Additività) La variabile aleatoria $X + Y$ ammette valor medio e*
$$E[X + Y] = E[X] + E[Y].$$

Dimostrazione.

1. Abbiamo
$$E[X] = \sum_{x \in E} |x| p_X(x)$$
Allora per ipotesi $E \subseteq \mathbb{R}^+$, e tutti gli addendi sono non negativi.
2. Se X e Y sono i.d., allora $p_X \equiv p_Y$ e abbiamo
$$E[X] = \sum_{x \in E} |x| p_X(x) = \sum_{x \in E} |x| p_Y(x) = E[Y]$$
3. Abbiamo
$$|E[X]| = \left| \sum_{x \in E} |x| p_X(x) \right| \leq \sum_{x \in E} |x| p_X(x) = E[|X|]$$
4. Per dimostrare che aX ammette valor medio, innanzitutto notiamo che aX assume valori in $aE := \{ax \mid x \in E\}$, e bisogna calcolare
$$\sum_{y \in aE} |y| P\{aX = y\} = \sum_{x \in E} |ax| P\{aX = ax\} = |a| \sum_{x \in E} |x| P\{X = x\} < +\infty$$
poichè X ha speranza finita, e quindi anche aX ammette speranza finita, e si ha
$$E[aX] = \sum_{y \in aE} y P\{aX = y\} = \sum_{x \in E} ax P\{aX = ax\} = a \sum_{x \in E} x P\{X = x\} = aE[X]$$

5. Abbiamo
$$E[X] = cP\{X = c\} = c \cdot 1 = c$$

6. Se X assume valori in E ed Y in F, con $E, F \subset \mathbb{R}$, allora $X+Y$ assumerà valori in $E+F := \{x+y \mid x \in E, y \in F\}$; innanzitutto dimostriamo che $X+Y$ ammette valor medio:
$$\begin{aligned}
\sum_{z \in E+F} |z| P\{X+Y=z\} &= \sum_{x\in E}\sum_{y \in F} |x+y| P\{X=x, Y=y\} \leq \\
&\leq \sum_{x\in E}\sum_{y \in F} |x| P\{X=x, Y=y\} + \sum_{x\in E}\sum_{y \in F} |y| P\{X=x, Y=y\} = \\
&= \sum_{x\in E} |x| \sum_{y \in F} P\{X=x, Y=y\} + \sum_{y\in F} |y| \sum_{x \in E} P\{X=x, Y=y\} = \\
&= \sum_{x\in E} |x| P\{X=x\} + \sum_{y\in F} |y| P\{Y=y\} = E[|X|] + E[|Y|] < +\infty
\end{aligned}$$
Possiamo quindi calcolare
$$\begin{aligned}
E[X+Y] &= \sum_{z\in E+F} zP\{X+Y=z\} = \sum_{x\in E}\sum_{y\in F}(x+y)P\{X=x,Y=y\} = \\
&= \sum_{x\in E} x \sum_{y\in F} P\{X=x,Y=y\} + \sum_{y\in F} y \sum_{x\in E} P\{X=x,Y=y\} = \\
&= \sum_{x\in E} xP\{X=x\} + \sum_{y\in F} yP\{Y=y\} = E[X] + E[Y]
\end{aligned}$$

∎

Grazie a queste proprietà possiamo agevolmente calcolare la speranza di una variabile aleatoria binomiale.

Esempio 2.7.5 (Variabili aleatorie binomiali) Consideriamo una variabile aleatoria $X \sim \mathcal{B}(n,p)$, e calcoliamone la speranza. Siccome X è distribuita come una variabile aleatoria della forma $S_n = \sum_{i=1}^n X_i$, con $X_i, i = 1, \ldots, n$ i.i.d. di legge $Be(p)$, allora:
$$E[X] = E[S_n] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n p = np$$
dove la prima uguaglianza segue dalla proprietà 2), la terza dalla proprietà 6) e la quarta da quanto visto nell'esempio 2.7.2.

La definizione di valor medio che abbiamo data risulta poco operativa quando dobbiamo calcolare la media di una funzione della variabile X. In particolare, se X è una variabile aleatoria discreta a valori in E (che potrebbe non essere un sottoinsieme di $\mathbb{R}$), e $f : E \to \mathbb{R}$, allora $f(X)$ è una variabile aleatoria reale discreta.

Proposizione 2.7.3 *Sia X una variabile aleatoria discreta a valori in un insieme generico E, sia p_X la sua densità e sia $f : E \to \mathbb{R}$. La variabile aleatoria $f(X)$ ammette valor medio se e solo se*
$$\sum_{x\in E} |f(x)| p_X(x) < +\infty, \tag{2.7.3}$$

In questo caso

$$E[f(X)] = \sum_{x \in E} f(x) p_X(x). \tag{2.7.4}$$

Dimostrazione. Dalla definizione di valor medio, posto $F := f(E) = \{f(x) \mid x \in E\}$, abbiamo che $f(X)$ ammette valor medio se e solo se

$$\sum_{y \in F} |y| P\{f(X) = y\} < +\infty. \tag{2.7.5}$$

Cominciamo col mostrare che la somma in (2.7.3) coincide con quella in (2.7.5). Si noti anzitutto, che la famiglia $(\{x \in E : f(x) = y\})_{y \in F}$ costituisce una partizione di E. Ma allora

$$\begin{aligned}
\sum_{x \in E} |f(x)| p_X(x) &= \sum_{y \in F} \sum_{x : f(x) = y} |f(x)| p_X(x) = \\
&= \sum_{y \in F} |y| \sum_{x : f(x) = y} P\{X = x\} = \\
&= \sum_{y \in F} |y| P\{f(X) = y\}.
\end{aligned}$$

Ciò mostra che (2.7.3) equivale al fatto che $f(X)$ ammetta valor medio. Per concludere la dimostrazione, occorre mostrare che, se $f(X)$ ammette valor medio, allora la somma in (2.7.4) coincide con

$$E[f(X)] = \sum_{x \in E} f(x) P\{X = x\}.$$

Ma per questo basta ripetere l'argomento appena usato per mostrare che la somma in (2.7.3) coincide con quella in (2.7.5). ∎

Esercizi

Esercizio 42 In una procedura di controllo di produzione, n processori prodotti da un processo industriale vengono sottoposti a controllo. Si assuma che ogni pezzo, indipendentemente dagli altri, abbia probabilità $p \in (0, 1)$ di essere difettoso. Se un processore è funzionante supera sicuramente il test di controllo, se il processore è difettoso fallisce il test con probabilità $q \in (0, 1)$, indipendentemente dagli altri. Sia X = numero di processori che hanno fallito il test. Determinare la distribuzione di X.

Esercizio 43 Due dadi truccati sono tali che la probabilità di ottenere un sei è il doppio della probabilità di ottenere ogni altro punteggio. Qual è la media del punteggio ottenuto lanciando i due dadi?

Esercizio 44 Da un'urna contenente r palline rosse e v palline verdi, si estraggono successivamente, senza reintroduzione, k palline, con $k \leq \min(r, v)$. Per $i = 1, 2, \ldots, k$, sia

$$X_i = \begin{cases} 1 & \text{se l'}i\text{-ma pallina estratta è rossa} \\ 0 & \text{altrimenti,} \end{cases}$$

e $X = X_1 + \cdots + X_k$.

a. Determinare la distribuzione di X.

b. Determinare le distribuzioni delle X_i.

c.* Mostrare che la densità congiunta delle X_i è data da

$$p_{X_1,\dots,X_k}(x_1,\dots,x_k) = \frac{r(r-1)\cdots(r-\sum_{i=1}^k x_i+1)v(v-1)\cdots(v-k+\sum_{i=1}^k x_i+1)}{(r+v)(r+v-1)\cdots(r+v-k+1)}$$

d. Calcolare $E[X]$.

Esercizio 45 Per $n \geq 1$, sia X_n una variabile aleatoria che assume, con la stessa probabilità, i valori $\frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1$. Se f è una funzione continua, sia

$$m_n = E[f(X_n)].$$

Mostrare che

$$\lim_{n\to+\infty} m_n = \int_0^1 f(x)dx.$$

Esercizio 46 Siano X, Y variabili aleatorie a valori in $\mathbb{N}$, definite sullo stesso spazio di probabilità. Giustificare l'identità

$$P\{X+Y=n\} = \sum_{k=0}^{+\infty} P\{X=k, Y=n-k\}.$$

Esercizio 47 * Sia X una variabile aleatoria a valori in $\mathbb{N}$. Allora

$$E[X] = \sum_{n=1}^{+\infty} P\{X \geq n\}.$$

Esercizio 48 Siano X, Y variabili aleatorie discrete, a valori in $\mathbb{N}$, con densità congiunta

$$p_{X,Y}(n,m) = c\frac{\lambda^n \mu^m \nu^{nm}}{n!m!}$$

dove $\lambda, \mu > 0$, $0 < \nu \leq 1$, e c è un'opportuna costante (quella per cui $\sum_{n,m} p_{X,Y}(n,m) = 1$).

a. Calcolare le densità marginali di X e Y.

b. Calcolare le probabilità condizionate $P\{X=n|Y=m\}$.

c.* Mostrare che gli eventi $\{X=n\}$, $\{Y=m\}$ sono indipendenti per ogni coppia $n, m \in \mathbb{N}$ se e solo se $\nu = 1$.

Esercizio 49 Si consideri la seguente classica strategia per il gioco della roulette. Gioco sempre sul rosso. Alla prima giocata punto un euro. Se perdo raddoppio la giocata, se vinco smetto. In ogni caso, dato che il mio capitale iniziale è 1023 euro, se perdo 10 volte di seguito devo smettere. Sia X la differenza tra il mio capitale alla fine del gioco e il capitale iniziale. Calcolare $E[X]$.

Esercizio 50 Un gioco a premi ha un montepremi di 512 Euro. Vengono poste ad un concorrente 10 domande. Ad ogni risposta errata il montepremi viene dimezzato. Alla prima risposta esatta il concorrente vince il montepremi rimasto. Se non si da alcuna risposta esatta non si vince nulla. Un certo concorrente risponde esattamente ad una domanda con probabilità $p \in (0,1)$, indipendentemente dalle risposte alle altre domande. Sia X la vincita di questo concorrente.

a. Determinare la densità p_X di X.

b.* Calcolare $E[X]$

Esercizio 51 In un concorso vengono assegnate le idoneità per un dato servizio. Si assuma che ogni partecipante, indipendentemente dagli altri, abbia probabilità $p = \frac{3}{4}$ di ottenere l'idoneità. Al termine del concorso, a 10 tra gli idonei viene assegnato un posto di lavoro (se gli idonei sono meno di 10 vengono assegnati tanti posti di lavoro quanti sono gli idonei). Supponiamo che al concorso partecipino 15 persone, e sia X il numero dei partecipanti che ottengono l'idoneità ma non il posto di lavoro.

a. Determinare la distribuzione di X.

b. Calcolare $E(X)$.

Esercizio 52 Si considerino 5 urne identiche, ognuna contenente una pallina rossa e quattro palline verdi. Ogni urna viene assegnata ad uno di cinque giocatori, e ogni giocatore estrae una pallina dalla propria urna. Un montepremi di 3000 Euro viene diviso tra i giocatori che estraggono la pallina rossa.

a. Sia X il numero di Euro vinti da ogni giocatore vincente ($X = 0$ se nessun giocatore estrae la pallina rossa). Determinare la densità e la media di X.

b. Si supponga di considerare uno dei cinque giocatori, chiamiamolo Tizio, e sia Y il numero di Euro vinti da Tizio. Si determinino la densità e la media di Y.

Esercizio 53 Si sceglie "a caso" un campione di 5 oggetti da un lotto di 100 di cui 10 sono difettosi per effettuare un controllo di qualità. Sia X il numero di oggetti difettosi contenuti nel campione. Determinare la densità discreta di X.

Esercizio 54 Si lancia una coppia di dadi non truccati e si sommano i risultati.

1. Si calcoli la probabilità che occorrano meno di 6 lanci per ottenere almeno un 7.

2. Si calcoli la probabilità che occorrano più di 6 lanci per ottenere almeno un 7.

2.8 Varianza e covarianza

Una quantità centrale del calcolo delle probabilità è la cosiddetta **varianza**, definita da $\mathrm{Var}[\cdot]$:

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

Se X ammette speranza finita, poichè la variabile aleatoria $(X - E[X])^2$ è non negativa la varianza è sempre ben definita, tuttavia può essere finita o infinita. In questo secondo caso indichiamo $\mathrm{Var}[X] = +\infty$.

Esempio 2.8.1 (Variabile aleatoria di Bernoulli) Prendiamo $X \sim Be(p)$ e calcoliamone la varianza. In base alla formula appena vista, abbiamo:

$$\begin{aligned}\mathrm{Var}[X] &= (0-p)^2 \cdot P\{X=0\} + (1-p)^2 \cdot P\{X=1\} = p^2(1-p) + (1-p)^2 p = \\ &= p(1-p)(p+1-p) = p(1-p)\end{aligned}$$

Si osservi che, a differenza della speranza, la varianza non è un operatore lineare. Infatti valgono le seguenti proprietà.

Proposizione 2.8.1 *Siano X, Y due variabili aleatorie reali discrete definite sullo spazio di probabilità $(\Omega, \mathcal{A}, P)$ e che ammettano speranza e varianza finite. Allora valgono le seguenti proprietà:*

1. $\mathrm{Var}[X] \geq 0$, *e* $\mathrm{Var}[X] = 0$ *se e solo se esiste* $c \in \mathbb{R}$ *tale che* $P\{X = c\} = 1$.
2. *Se X e Y sono i.d., allora* $\mathrm{Var}[X] = \mathrm{Var}[Y]$.
3. *(Omogeneità di grado 2) Se $a \in \mathbb{R}$, allora aX ammette varianza finita e* $\mathrm{Var}[aX] = a^2\mathrm{Var}[X]$.
4. *La variabile aleatoria $X + Y$ ammette varianza finita e*

$$\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\mathrm{Cov}(X, Y).$$

dove definiamo **covarianza** *tra X e Y la quantità*

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

Dimostrazione.

1. Siccome $(X - E[X])^2$ è una variabile aleatoria non negativa, abbiamo $\mathrm{Var}[X] = E[(X - E[X])^2] \geq 0$. Ora, supponiamo che $P\{X = c\} = 1$ per qualche $c \in \mathbb{R}$; allora chiaramente $E[X] = c$, e $\mathrm{Var}[X] = E[(X - E[X])^2] = (c - c)^2 \cdot 1 = 0$.

   Per ora riusciamo a dimostrare il viceversa solo se X è discreta: in questo caso infatti

$$0 = \mathrm{Var}[X] = \sum_x (x - E[X])^2 p_X(x)$$

   Siccome ogni addendo $(x - E[X])^2 p_X(x)$ è non negativo, devono essere tutti uguali a 0, cioè per ogni x si ha $x = E[X]$ oppure $p_X(x) = 0$. Il primo caso si deve verificare una ed una sola volta con probabilità 1, mentre il secondo caso si verifica tutte le altre volte.

2. Se X e Y sono i.d., abbiamo che $E[X] = E[Y]$, e anche

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

3. Per ogni $a \in \mathbb{R}$ abbiamo

$$\mathrm{Var}[aX] = E[(aX - E[aX])^2] = E[(aX - aE[X])^2] = a^2 E[(X - E[X])^2] = a^2\mathrm{Var}[X]$$

4. Il fatto che $X + Y$ ammetta varianza finita è rimandato ad un risultato successivo. Dando per vero che questo avviene, si ha

$$\begin{aligned}
\mathrm{Var}[X + Y] &= E[(X + Y - E[X + Y])^2] = E[(X - E[X] + Y - E[Y])^2] = \\
&= E[(X - E[X])^2] + E[(Y - E[Y])^2] + 2E[(X - E[X])(Y - E[Y])] = \\
&= \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\mathrm{Cov}(X, Y)
\end{aligned}$$

■

Esempio 2.8.2 (variabile aleatoria di Poisson) Prendiamo $X \sim Po(\lambda)$ e calcoliamone la varianza usando la formula $\mathrm{Var}[X] = E[X^2] - E[X]^2$. Innanzitutto calcoliamo

$$\begin{aligned} E[X^2] &= E[X(X-1)] + E[X] = \sum_{k=0}^{\infty} k(k-1)e^{-\lambda}\frac{\lambda^k}{k!} + \lambda = e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^k}{(k-2)!} + \lambda = \\ &= e^{-\lambda}\lambda^2\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!} + \lambda = e^{-\lambda}\lambda^2\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} + \lambda = \lambda^2 + \lambda \end{aligned}$$

quindi

$$\mathrm{Var}[X] = E[X^2] - E[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

2.9 Speranza e varianza di variabili aleatorie indipendenti

Una proprietà notevole della speranza si ha quando consideriamo prodotti di variabili aleatorie indipendenti.

Proposizione 2.9.1 *Siano X, Y variabili aleatorie reali discrete indipendenti, definite nello stesso spazio di probabilità e che ammettono valor medio. Allora XY ammette valor medio e*

$$E[XY] = E[X]E[Y].$$

Dimostrazione. Cominciamo col mostrare che XY ammette valor medio. Usando il fatto che $p_{X,Y}(x,y) = p_X(x)p_Y(y)$ si ha:

$$E[|XY|] = \sum_{x,y}|x||y|p_{X,Y}(x,y) = \sum_x |x|p_X(x)\sum_y |y|p_Y(y) = E[|X|]E[|Y|] < +\infty.$$

In modo analogo, rimuovendo i valori assoluti si dimostra che $E[XY] = E[X]E[Y]$. ∎

Corollario 2.9.2 *Siano X, Y due variabili aleatorie discrete indipendenti a valori rispettivamente in E ed F, e siano $f : E \to \mathbb{R}$ e $g : F \to \mathbb{R}$ due funzioni misurabili tali che le variabili aleatorie $f(X)$ e $g(Y)$ ammettano valor medio. Allora $E[f(X)g(Y)] = E[f(X)]E[g(Y)]$.*

Proposizione 2.9.3 *asta notare che le due variabili aleatorie $f(X)$ e $g(Y)$ sono reali, discrete e indipendenti, e applicare la Proposizione 2.9.1.*

Corollario 2.9.4 *Siano X, Y due variabili aleatorie reali discrete indipendenti, che ammettono valor medio. Allora $\mathrm{Cov}(X, Y)$ è ben definita, e vale zero.*

Il Corollario 2.9.4 segue immediatamente dalla Proposizione 2.9.1. In particolare, ne segue che due variabili aleatorie reali discrete indipendenti che ammettono media sono scorrelate. Il viceversa non è necessariamente vero, come mostra l'esempio che segue.

Esempio 2.9.1 Sia Z una variabile aleatoria tale che

$$p_Z(0) = p_Z(\pi/2) = p_Z(\pi) = \frac{1}{3}.$$

Posto $X = \sin(Z)$ e $Y = \cos(Z)$, si vede subito che $XY \equiv 0$ e $E[Y] = 0$, da cui

$$\mathrm{Cov}(X, Y) = 0$$

segue immediatamente. Ma

$$P\{X = 1, Y = 1\} = 0 \neq \frac{1}{9} = P\{X = 1\}P\{Y = 1\}.$$

Infine una semplice dimostrazione per induzione mostra quanto segue.

Corollario 2.9.5 *Siano* $X_1, X_2, \ldots, X_n$ *variabili aleatorie che ammettono varianza finita e tali che* $\mathrm{Cov}(X_i, X_j) = 0$ *se* $i \neq j$. *Allora*

$$\mathrm{Var}[X_1 + \cdots + X_n] = \sum_{i=1}^{n} \mathrm{Var}[X_i].$$

Per il Corollario 2.9.4, il risultato del Corollario 2.9.5 vale, in particolare, se le variabili $X_1, X_2, \ldots, X_n$ sono indipendenti.

Esempio 2.9.2 (variabili aleatorie binomiali) Se $X \sim B(n, p)$, allora X ha la stessa legge di $Y := Y_1 + \ldots + Y_n$, con $(Y_i)_i$ i.i.d. con legge $Be(p)$. Allora

$$\mathrm{Var}[X] = \mathrm{Var}[Y] = \mathrm{Var}\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n} \mathrm{Var}[Y_i] = \sum_{i=1}^{n} p(1-p) = np(1-p)$$

2.10 Disintegrazione

Come per gli eventi a volte può essere agevole fornire probabilità condizionate ad altri eventi, anche per le variabili aleatorie la legge può essere espressa condizionatamente ai risultati di altre variabili aleatorie. Se ad esempio sullo spazio probabilizzato $(\Omega, \mathcal{A}, P)$ abbiamo due variabili aleatorie discrete $X : \Omega \to E$ ed $Y : \Omega \to F$, per la variabile aleatoria X possiamo definire la sua densità discreta rispetto a P come di consueto

$$p_X(x) := P\{X = x\} \quad \forall x \in E$$

Possiamo inoltre definire la sua densità discreta condizionata ad Y in questo modo. Per ogni $y \in F_0$, con

$$F_0 := \{y \in F \mid P\{Y = y\} > 0\}$$

definiamo la probabilità $Q_y := P(\cdot \mid \{Y = y\})$. Allora ogni Q_y è una probabilità su $(\Omega, \mathcal{A})$, e quindi possiamo considerare la legge e la densità discreta di X rispetto a queste nuove probabilità; in particolare definiamo

$$p_{X|Y}(x|y) := Q_y\{X = x\} = P\{X = x \mid Y = y\} = \frac{p_{XY}(x, y)}{p_Y(y)} \quad \forall x \in E, y \in F_0$$

che si chiama **densità condizionata di** X **rispetto a** Y.

Come per gli eventi c'è una relazione che lega le probabilità rispetto a P alle probabilità condizionate (la formula della probabilità totale), anche per le speranze di variabili aleatorie esiste la seguente relazione analoga, chiamata **formula della disintegrazione** rispetto alla legge di Y.

Proposizione 2.10.1 *iano X ed Y variabili aleatorie discrete, e $g : E \to \mathbb{R}^+$ misurabile. Allora*

$$E[g(X)] = \sum_{y \in F_0} E_y[g(X)] p_Y(y)$$

dove, per ogni $y \in F_0$, E_y è la speranza calcolata rispetto alla probabilità Q_y.

Dimostrazione. Siccome $g(X)$ è una variabile aleatoria non negativa, la sua speranza è sempre ben definita (eventualmente uguale a $+\infty$) per qualunque misura di probabilità che noi mettiamo su $(\Omega, \mathcal{A})$. Abbiamo inoltre che, per ogni $y \in F_0$,

$$E_y[g(X)] = \sum_{x \in E} g(x) Q_y\{X = x\} = \sum_{x \in E} g(x) \frac{p_{XY}(x,y)}{p_Y(y)}$$

Quindi

$$\begin{aligned} \sum_{y \in F_0} E_y[g(X)] p_Y(y) &= \sum_{y \in F_0} \sum_{x \in E} g(x) \frac{p_{XY}(x,y)}{p_Y(y)} p_Y(y) = \sum_{x \in E} g(x) \sum_{y \in F_0} p_{XY}(x,y) = \\ &= \sum_{x \in E} g(x) p_X(x) = E[g(X)] \end{aligned}$$

∎

Esempio 2.10.1 (Media e varianza di variabili aleatorie geometriche) Consideriamo una successione $(X_n)_n$ i.i.d. con legge $Be(p)$, e indichiamo con $T := \min\{n \mid X_n = 1\}$ l'istante del primo successo. Vogliamo ora calcolare media e varianza di T, che ci forniranno quindi la media e la varianza di una variabile aleatoria geometrica. A questo scopo consideriamo la disintegrazione rispetto alla variabile aleatoria X_1, e avremo

$$E[T] = E_1[T] P\{X_1 = 1\} + E_0[T] P\{X_1 = 0\}$$

Sotto Q_1, T coincide con la costante 1, difatti abbiamo $Q_1\{T = 1\} = P\{T = 1 \mid X_1 = 1\} = 1$ poichè $\{T = 1\} = \{X_1 = 1\}$, e quindi $E_1[T] = 1$. Calcoliamo ora la legge di T sotto Q_0: per ogni $k \geq 2$ abbiamo

$$Q_0\{T > k\} = P\{T > k \mid X_1 = 0\} = P\{T > k - 1 + 1 \mid T > 1\} = P\{T > k - 1\} = (1-p)^{k-1}$$

che significa che, sotto Q_0, $T + 1 \sim Ge(p)$. Abbiamo allora

$$E[T] = 1 \cdot p + E_0[T + 1 - 1](1-p) = p + (1-p)(E[T] + 1)$$

e quindi, risolvendo l'equazione di primo grado, si ha

$$E[T] = \frac{1}{p}$$

Per calcolare la varianza di T, innanzitutto calcoliamo

$$\begin{aligned} E[T^2] &= E_1[T^2] P\{X_1 = 1\} + E_0[T^2] P\{X_1 = 0\} = p + (1-p) E_0[(T + 1 - 1)^2] = \\ &= p + (1-p)(E_0[(T+1)^2] - 2E_0[T+1] + 1) = p + (1-p)\left(E[T^2] - \frac{2}{p} + 1\right) \end{aligned}$$

Risolvendo l'equazione di primo grado si ha

$$E[T^2] = \frac{2-p}{p^2}$$

e quindi $\mathrm{Var}[T] = E[T^2] - E[T]^2 = \frac{1-p}{p^2}$.

Esercizi

Esercizio 55 Siano $X, Y \sim Be(p)$ indipendenti, con $p \in (0,1)$. Mostrare che le variabili aleatorie $X+Y$ e $X-Y$ sono scorrelate ma non indipendenti.

Esercizio 56 In un canale di trasmissione vengono trasmessi simboli binari. I disturbi sul canale fanno si che ogni simbolo trasmesso ha la probabilità del 2% di essere ricevuto errato, indipendentemente dagli altri simboli. I messaggi vengono trasmessi in "parole" composte da 50 simboli. Qual è la probabilità che una parola venga ricevuta con almeno due simboli errati? (Usare l'approssimazione di Poisson)

Esercizio 57 Sia X una variabile aleatoria a valori in $\mathbb{N}$, tale che

$$P(X = k) = \frac{e^{-2} 2^k}{2k!}(1 + \lambda k),$$

con $\lambda > 0$.

a) Determinare il valore di λ

b) Sia $Y \sim Po(2)$. Mostrare che

$$P(X = k) = \frac{1}{2}[P(Y = k) + P(Y + 1 = k)].$$

c) Usando il risultato in (b), calcolare media e varianza di X

Esercizio 58 Sia $\lambda > 0$ e $c : \mathbb{N} \to [0, +\infty)$ una funzione tale che $c(0) = 0$ e $\inf_{n>0} c(n) > \lambda$. Sia X una variabile aleatoria a valori in $\mathbb{N}$ con densità

$$p_X(n) = \begin{cases} \frac{1}{Z} & \text{se } n = 0 \\ \frac{1}{Z} \frac{\lambda^n}{c(1)c(2)\cdots c(n)} & \text{se } n \geq 1, \end{cases}$$

dove $Z = 1 + \sum_{n=1}^{+\infty} \frac{\lambda^n}{c(1)c(2)\cdots c(n)}$.

a. Mostrare che $E[c(X)] = \lambda$.

b. Mostrare che per ogni funzione $f : \mathbb{N} \to \mathbb{R}$ tale che la variabile aleatoria $f(X+1)$ ammette valor medio, si ha

$$E[c(X)f(X)] = \lambda E[f(X+1)].$$

c.* Si assuma che, per ogni $n \geq 0$, $|c(n+1) - c(n)| \leq 1$. Sia $Y \sim Po(\lambda)$. Mostrare, per induzione su k, che per ogni $k \geq 1$

$$E[c(X)^k] \leq E[Y^k].$$

(Suggerimento: usare, dopo averla verificata, l'uguaglianza $E[Y^{k+1}] = \lambda \sum_{i=0}^{k} \binom{k}{i} E[Y^i]$)

Esercizio 59 Siano $X_1, X_2 \sim Po(\lambda)$ indipendenti.

a) Fissati $n \in \mathbb{N}$ e $0 \leq k \leq n$, calcolare

$$P\{X_1 = k \mid X_1 + X_2 = n\}. \tag{2.10.0}$$

b. Supponiamo che n sia fissato, e chiamiamo $q(k)$ il valore dell'espressione in (2.10.0). Mostrare che $q(\cdot)$ è il valore della densità di una variabile aleatoria binomiale, e determinarne i parametri.

c. Siano ora $X_1, X_2, \ldots, X_m \sim Po(\lambda)$ indipendenti, con $m > 2$. Calcolare

$$q_m(k) = P\{X_1 = k \mid X_1 + X_2 + \cdots + X_m = n\},$$

e determinare i parametri della variabile aleatoria binomiale la cui densità è $q_m(\cdot)$.

Esercizio 60 Siano X, Z e W variabili aleatorie indipendenti con $X \sim Be(p)$, $Z, W \sim Po(\lambda)$. Definiamo

$$Y = XZ + W.$$

a. Determinare le densità $p_{X,Y}$ e p_Y.

b. Utilizzando la densità calcolata al punto a., determinare $E[Y]$ e $\mathrm{Var}[Y]$.

c. Calcolare $E[Y]$ e $\mathrm{Var}[Y]$ *senza* utilizzare p_Y.

Esercizio 61 Siano X e Y due variabili aleatorie a valori in $\mathbb{N}$ aventi la seguente densità congiunta:

$$P\{X = k, Y = n\} = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} e^{-\lambda} \frac{\lambda^n}{n!} & \text{se } 0 \leq k \leq n \\ 0 & \text{altrimenti,} \end{cases}$$

dove $p \in (0, 1)$ e $\lambda > 0$ sono due parametri fissati.

a. Determinare le densità marginali di X e Y. Mostrare, in particolare, che $X \sim Po(p\lambda)$ e $Y \sim Po(\lambda)$.

b. Calcolare il coefficente di correlazione $\rho_{X,Y}$.

Esercizio 62 Un'urna contiene 100 palline numerate da 0 a 99. Si estrae una pallina, e si denotano con X e Y le due cifre del numero estratto (la cifra delle decine, X, si considera uguale a zero per numeri minori di 10). Mostrare che X e Y sono indipendenti, e determinarne la distribuzione.

Esercizio 63 Dimostrare che le seguenti affermazioni sono equivalenti:

i) Gli eventi A e B sono indipendenti.

ii) Le variabili aleatorie 1_A e 1_B sono indipendenti.

iii) Le variabili aleatorie 1_A e 1_B sono scorrelate (cioè $\mathrm{Cov}(1_A, 1_B) = 0$).

Esercizio 64 Siano $X \sim Ge(p)$ e $Y \sim Ge(q)$, Determinare la densità di $\min(X, Y)$.

Capitolo 3

Spazi di probabilità generali. Variabili aleatorie assolutamente continue

3.1 Il discreto non basta

Con le variabili aleatorie discrete, 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 aleatorie 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 descrivere una tale grandezza tramite una variabile aleatoria discreta: se infatti imponiamo che $X(\Omega)$ sia un intervallo di $\mathbb{R}$, non potremo mai ottenere una tale situazione se X è discreta, dato che in questo caso $X(\Omega)$ è sempre al più numerabile. L'altra questione riguarda lo studio delle *successioni* di variabili aleatorie. 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 aleatorie. Uno spazio di probabilità discreto è troppo "povero" perchè in esso si possano definire successioni interessanti di variabili aleatorie, ad esempio successioni di variabili aleatorie indipendenti. Inoltre, anche se lo spazio probabilizzato fosse sufficientemente ricco da ospitare una successione infinita di variabili aleatorie, anche nel caso più semplice in cui queste fossero di Bernoulli si riesce a dimostrare molto semplicemente che si può costruire una variabile aleatoria che non può essere discreta (vedi l'Esempio 3.1.2 sotto).

Risulta quindi naturale lavorare con variabili aleatorie che non siano necessariamente discrete, e che ovviamente saranno definite su spazi di probabilità in cui Ω sia non numerabile. Per avere un'idea del tipo di problemi che si affrontano, vediamo due esempi significativi.

Esempio 3.1.1 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 volere che, se $0 \leq a \leq b \leq 1$ allora

$$P([a, b]) = b - a. \tag{3.1.1}$$

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 (3.1.1), segue che $P(\{x\}) = P([x,x]) = 0$ per ogni $x \in [0,1]$. Il problema è: è possibile estendere la P definita in (3.1.1) 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 3.1.2 Nell'esempio 1.5.4 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

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

cioè, se $\omega \in \Omega$, allora $\omega = (\omega_1, \omega_2, \dots)$ 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 1.5.4, con N fissato. In altre parole, se $x_1, x_2, \dots, x_N \in \{0,1\}$, dovrà essere

$$P\{\omega \in \Omega : \omega_1 = x_1, \dots, \omega_N = x_N\} = p^{\sum_{i=1}^N x_i}(1-p)^{N-\sum_{i=1}^N x_i}. \tag{3.1.2}$$

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

$$\{\eta\} = \bigcap_n \{\omega : \omega_1 = \eta_1, \dots, \omega_n = \eta_n\}.$$

Quest'ultima è l'intersezione di una famiglia decrescente di eventi. Poichè abbiamo assunto P σ-additiva e osservando che la Proposizione 1.2.2 non utilizza la numerabilità dello spazio campionario, abbiamo

$$P\{\eta\} = \lim_{n \to +\infty} p^{\sum_{i=1}^n \eta_i}(1-p)^{n-\sum_{i=1}^n \eta_i} = 0$$

se $p \in (0,1)$ (verificarlo!). Si vede quindi che su uno spazio di questo genere tutti i singoli punti hanno probabilità nulla. Nonostante questo, si possono comunque definire molto facilmente delle variabili aleatorie discrete. Se infatti definiamo $X_n : \Omega \to \{0,1\}$ come la proiezione canonica

$$X_n(\omega) := \omega_n \quad \forall n \in \mathbb{N}^*$$

allora le $(X_n)_n$ sono delle variabili aleatorie di Bernoulli indipendenti di parametro p: infatti per ogni $n \in \mathbb{N}^*$ e per ogni $x_1, \dots, x_n \in \{0,1\}$ si ha che

$$P\{X_1 = x_1, \dots, X_n = x_n\} = P\{\omega \in \Omega \mid \omega_1 = x_1, \dots, \omega_n = x_n\} = p^{\sum_{i=1}^n x_i}(1-p)^{n-\sum_{i=1}^n x_i}$$

che è il prodotto di n densità marginali di variabili aleatorie di Bernoulli. Dato che ogni sottoinsieme finito di $\mathbb{N}^*$ è contenuto in un insieme del tipo $\{1, \dots, n\}$, per la caratterizzazione dell'indipendenza tramite la densità congiunta questo basta a dimostrare che le $(X_n)_n$ sono indipendenti.

Un altro tipo di variabili aleatorie che si possono costruire su uno spazio probabilizzato di questo tipo sono quelle geometriche. Se infatti definiamo Y come l'istante del primo successo nelle prove ripetute

$$Y := \begin{cases} \min\{n \mid X_n = 1\} \\ +\infty \text{ se l'insieme sopra è vuoto} \end{cases}$$

allora si dimostra facilmente che $Y \sim Ge(p)$. Infatti:

$$P\{Y > n\} = P\{X_1 = X_2 = \ldots = X_n = 0\} = \prod_{i=1}^{n} P\{X_i = 0\} = (1-p)^n$$

che caratterizza una variabile aleatoria geometrica (come abbiamo visto nella Proposizione 2.5.10). Si ottiene quindi in particolare che

$$P\{Y = +\infty\} = 1 - P\{Y < +\infty\} = 1 - \sum_{n=1}^{\infty} p(1-p)^{n-1} = 1 - 1 = 0$$

e quindi l'insieme nella definizione di Y è quasi certamente non vuoto.

Sembrerebbe quindi che su questo spazio probabilizzato tutte le variabili aleatorie "significative" siano discrete. Non è così: infatti si riescono a costruire abbastanza facilmente variabili aleatorie che non sono discrete. Definiamo ad esempio $Z : \Omega \in [0,1)$ come

$$Z := \sum_{n=1}^{\infty} \frac{X_n}{2^n}$$

che intuitivamente corrisponde ad identificare un generico elemento di Ω con lo sviluppo in cifre binarie di un numero reale compreso tra 0 e 1. Si vede facilmente che questa variabile aleatoria non è discreta: infatti per ogni $x \in [0,1)$, si può scrivere il suo sviluppo in cifre binarie come $x = 0.x_1x_2x_3\ldots$, o più formalmente come $x = \sum_{n=1}^{\infty} x_n/2^n$; allora

$$P\{Z = x\} = P\left\{\sum_{n=1}^{\infty} \frac{X_n}{2^n} = \sum_{n=1}^{\infty} \frac{x_n}{2^n}\right\} = P\{X_n = x_n \ \forall n\} = P\left(\bigcap_n \{\omega : \omega_1 = x_1, \ldots, \omega_n = x_n\}\right) = 0$$

e quindi Z non può essere una variabile aleatoria discreta.

In entrambi gli esempi appena visti, la funzione $P(\cdot)$ viene definita dapprima su 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. In questo ci viene in aiuto la restrizione che abbiamo imposto fin dall'inizio, cioè di considerare la P definita solo su una famiglia "significativa" di eventi $\mathcal{A}$, che abbiamo imposto essere una σ-algebra. Infatti, si può dimostrare che, sebbene non si riesca ad estendere P ad ogni sottoinsieme di Ω in modo che risulti σ-additiva, questo si può fare su una opportuna σ-algebra, costruita a partire da eventi "elementari" quali quelli visti negli esempi 3.1.1 e 3.1.2 (rispettivamente, gli intervalli e gli insiemi della forma $\{x_1\} \times \ldots \times \{x_N\} \times \{0,1\}^{\mathbb{N}^* \setminus \{1,\ldots,N\}}$).

Esaminiamo in dettaglio questo problema, suddividendolo in due sottoproblemi:

1. trovare una opportuna σ-algebra $\mathcal{A}$ che contenga tutti gli eventi che riteniamo "significativi";
2. costruire una misura di probabilità P su $\mathcal{A}$ che assuma dei valori fissati sugli eventi "significativi".

Per fissare le idee, ci aiutamo con gli stessi Esempi 3.1.1 e 3.1.2. 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 3.1.1, e $\mathcal{I}$ = famiglia degli insiemi del tipo $\{\omega : \omega_1 = x_1, \ldots, \omega_N = x_N\}$ nell'Esempio 3.1.2.

È facile vedere che $\mathcal{I}$ *non* è una σ-algebra. Come abbiamo detto, il primo problema è allora di trovare una σ-algebra $\mathcal{A}$ contenente $\mathcal{I}$. La scelta della σ-algebra $\mathcal{A}$ si può fare in modo "canonico", grazie al seguente risultato.

Proposizione 3.1.1 *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 $\mathcal{I}$.

∎

Ora che abbiamo risolto il primo problema, nel seguito una coppia $(\Omega, \mathcal{A})$ formata da un insieme e da una σ-algebra di suoi sottoinsiemi verrà chiamata **spazio misurabile** in analogia con quanto visto nel Capitolo 2.

Riguardo al secondo problema, 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. Basterà qui dire che per tutti gli scopi di questo libro (e in particolare per gli Esempi 3.1.1 e 3.1.2) è possibile dimostrare che esiste un'unica probabilità che estende P a $\sigma(\mathcal{I})$.

Osservazione 3.1.3 Se $(\Omega, \mathcal{A}, P)$ è uno spazio di probabilità discreto, allora si può identificare con lo spazio di probabilità $(\Omega, \mathcal{P}(\Omega), P)$.

3.2 Variabili aleatorie reali

Nel contesto più generale in cui ci stiamo mettendo, in cui uno spazio probabilizzato e le variabili aleatorie definite su di esso possono non essere discreti, è più che mai importante ricordare la definizione di variabile aleatoria e della sua legge.

Ricordiamo quindi che, dato uno spazio di probabilità $(\Omega, \mathcal{A}, P)$ e uno spazio misurabile (come ora possiamo permetterci di chiamarlo!) $(E, \mathcal{E})$, diciamo che l'applicazione $X : \Omega \to E$ è una **variabile aleatoria** a valori in $(E, \mathcal{E})$ se per ogni $C \in \mathcal{E}$ si ha $X^{-1}(C) = \{X \in C\} \in \mathcal{A}$. Ricordiamo poi che, se X è una variabile aleatoria e $C \in \mathcal{E}$, allora $P\{X \in C\}$ è ben definita: per questo possiamo definire **distribuzione**, o **legge**, della variabile aleatoria X la misura di probabilità definita da
$$\mu_X(C)\ (= P_X(C)\ ) := P\{X \in C\} \quad \forall C \in \mathcal{E}$$

La nozione di variabile aleatoria dipende dalla scelta della σ-algebra sull'insieme E. Nel caso in cui $E = \mathbb{R}$, vorremmo che sia ben definita la probabilità che queste variabili aleatorie assumano valori in un intervallo, ad esempio affinchè la funzione di ripartizione

$$F_X(t) = P\{X \leq t\} \quad \forall t \in \mathbb{R}.$$

sia ben definita. A questo scopo è necessario dotare $\mathbb{R}$ di una σ-algebra che contenga le semirette $(-\infty, t]$, $t \in \mathbb{R}$, o equivalentemente gli insiemi (a, b) (che, per quanto visto quando abbiamo trattato la funzione di ripartizione, si possono ottenere con un numero finito di operazioni ammissibili per una σ-algebra quali unione numerabile e complemento). A questo scopo, usando la Proposizione 3.1.1, definiamo

$$\mathcal{B}(\mathbb{R}) := \sigma(\tau)$$

dove τ è la topologia euclidea su $\mathbb{R}$: questa σ-algebra si indica col nome di σ-algebra **boreliana** (o **dei boreliani**) di $\mathbb{R}$.

Nel seguito chiameremo X variabile aleatoria reale solo se $\mathbb{R}$ sarà dotato della σ-algebra dei boreliani, e quindi se e solo se

$$X^{-1}(B) = \{X \in B\} \in \mathcal{A} \quad \forall B \in \mathcal{B}(\mathbb{R})$$

Notiamo che a $\mathcal{B}(\mathbb{R})$ appartengono tutti gli aperti, i chiusi e anche insiemi più generali come qualunque insieme B misurabile secondo Riemann, tale cioè che l'integrale di Riemann

$$\int_{\mathbb{R}} \mathbf{1}_B(t)\ dt$$

(eventualmente improprio se B non è limitato) sia ben definito. In realtà $\mathcal{B}(\mathbb{R})$ contiene anche insiemi che non sono integrabili secondo Riemann, quali ad esempio $B := \cup_{x \in Q}\{x\}$, che è ottenibile come unione numerabile di punti (che sono insiemi chiusi nella topologia euclidea) ma non è misurabile secondo Riemann.

Molte delle definizioni e delle proprietà relative a variabili aleatorie date nel Capitolo 2 continuano a valere nella loro piena generalità. Ad esempio, le proprietà della funzione di ripartizione contenute nella Proposizione 2.6.2 continuano a valere nel caso generale, come si vede dal fatto che la dimostrazione della Proposizione 2.6.2 è corretta anche in spazi di probabilità generali. Inoltre, la stessa dimostrazione usata nella Proposizione 2.6.3 mostra che

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

Ovviamente questo risultato sarà uguale a 0 su tutti gli $x \in \mathbb{R}$ tranne che su al più un'infinità numerabile di punti, e anzi si avrà anche che $\sum_{x \in \mathbb{R}}(F_X(x) - F_X(x^-)) \leq 1$ (perchè?). Si può dimostrare che X è una variabile aleatoria discreta se e solo se F_X è costante in ogni punto che non appartenga a questo insieme.

Un altro esempio è dato dalla nozione di variabili aleatorie indipendenti: le variabili aleatorie reali $\{X_i : i \in I\}$ si dicono **indipendenti** se, per ogni $J \subseteq I$ finito e per ogni scelta di $A_j \in \mathcal{B}(\mathbb{R})$, $j \in J$, si ha

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

Ovviamente in generale non varrà la caratterizzazione di indipendenza che abbiamo dato con le densità discrete nella Proposizione 2.5.3, dato che in generale non si potrà parlare di densità discreta. Nel caso di variabili aleatorie reali non discrete, quindi, per il resto di questo corso ci si limiterà ad usare l'indipendenza come regola di calcolo.

3.3 Variabili aleatorie assolutamente continue

Una famiglia molto importante di variabili aleatorie non discrete è quella delle variabili aleatorie assolutamente continue. Va precisato che una formulazione rigorosa della teoria del Calcolo delle Probabilità dovrà far uso del cosiddetto *integrale di Lebesgue*, che è una generalizzazione dell'integrale di Riemann. Poichè però questo va oltre lo scopo di questo corso, da qui in poi ci limiteremo a formulare enunciati usando l'integrale di Riemann, rimandando la generalizzazione a corsi successivi.

Definizione 3.3.1 Sia X una variabile aleatoria reale. Diremo che X è **assolutamente continua** se esiste una funzione $f_X : \mathbb{R} \to [0, +\infty)$ integrabile secondo Riemann e tale che per ogni I intervallo (non necessariamente limitato) di $\mathbb{R}$ si abbia

$$P\{X \in I\} = \int_I f_X(x)\, dx. \tag{3.3.1}$$

La funzione f_X si dice **densità** della variabile aleatoria X. La definizione è in realtà equivalente a dire che la relazione (3.3.1) vale per ogni insieme I misurabile secondo Riemann, tale cioè che la funzione $\mathbf{1}_C$ sia integrabile secondo Riemann (eventualmente in senso improprio).

Osservazione 3.3.1 Si noti che, scegliendo $I = \mathbb{R}$ in (3.3.1), si deve avere che

$$\int_{\mathbb{R}} f_X(x)\, dx = 1.$$

È possibile dimostrare, con qualche strumento di analisi matematica che qui non trattiamo, che, data una funzione $f : \mathbb{R}^d \to [0, +\infty)$ integrabile su $\mathbb{R}^d$ e tale che

$$\int_{\mathbb{R}^d} f(x)\, dx = 1,$$

allora esiste una variabile aleatoria assolutamente continua X, definita in un opportuno spazio di probabilità, la cui densità è f.

Osservazione 3.3.2 È naturale chiedersi se la Definizione 3.3.1 identifichi univocamente la densità f_X, ossia se esista un'unica funzione f_X che verifica (3.3.1). La risposta è negativa: per esempio se si modifica il valore di f_X in un numero finito di punti, o anche più in generale in un insieme di misura di Riemann nulla, gli integrali

$$\int_I f_X(x)\, dx$$

non vengono modificati. Conviene allora definire la seguente relazione di equivalenza nell'insieme delle funzioni integrabili su $\mathbb{R}^d$: $f \sim g$ se

$$\int_I f(x)\, dx = \int_I g(x)\, dx$$

per ogni scelta di I intervallo (o, più in generale, misurabile secondo Riemann). Dunque, per una variabile aleatoria assolutamente continua X, tutte le funzioni di una determinata classe di equivalenza per la relazione precedente sono densità di X. Una variabile aleatoria assolutamente continua identifica univocamente tale classe di equivalenza, ma tutti gli elementi di tale classe sono densità per X. Quindi, ogni qual volta useremo espressioni del tipo "f è la densità di X", intenderemo “f è un rappresentante della classe di equivalenza delle densità di X” (si usa allora dire che “f è una **versione** della densità di X). Analogamente, se X è una variabile aleatoria assolutamente continua di densità f_X, scrivendo “$f_X = g$” intenderemo “$f_X \sim g$”. Non vi saranno problemi di ambiguità nell'uso di queste convenzioni.

Osservazione 3.3.3 Si può dimostrare (ma non lo faremo nè lo useremo mai nel corso) che da (3.3.1) segue che

$$P\{X \in C\} = \int_C f_X(x)\ dx$$

per ogni $C \in \mathcal{B}(\mathbb{R}^d)$.

Osservazione 3.3.4 Sia X una variabile aleatoria reale assolutamente continua. Per la definizione 3.3.1 si ha

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt. \tag{3.3.2}$$

che in un certo senso è l'analogo continuo dell'Equazione (2.6.1). Dal Teorema Fondamentale del Calcolo Integrale segue che la funzione F_X è derivabile tranne al più in un insieme N di misura di Riemann zero, e per ogni $x \in \mathbb{R} \setminus N$

$$F'_X(x) = f_X(x). \tag{3.3.3}$$

In particolare, se f_X è continua a tratti, ossia f_X è continua su $\mathbb{R} \setminus N$ dove N è finito, allora (3.3.3) vale per $x \notin N$. Viceversa, usando di nuovo il Teorema Fondamentale del Calcolo Integrale, se una variabile aleatoria X è tale che F_X è $\mathcal{C}^1$ a tratti, cioè è continua ovunque e derivabile con continuità su $\mathbb{R} \setminus N$ dove N è finito, allora X è assolutamente continua e la densità f_X è data da

$$f_X(x) = \begin{cases} F'(x) & \text{se } x \notin N \\ \text{arbitrario} & \text{se } x \in N. \end{cases} \tag{3.3.4}$$

Infine, per ogni intervallo $I \subset \mathbb{R}$ di estremi a e b si ha

$$P\{X \in I\} = F_X(b) - F_X(a)$$

indipendentemente da se l'intervallo sia aperto o chiuso a destra o a sinistra.

Esempio 3.3.5 (variabili aleatorie uniformi) Sia C intervallo di $\mathbb{R}$ di estremi a e b: una variabile aleatoria X si dice **uniforme** su C (e scriviamo $X \sim U(a, b)$), se

$$f_X(x) = \frac{1}{m(C)} \mathbf{1}_C(x).$$

dove $m(C) := b - a$ è la misura di Riemann di C. La variabile aleatoria X si interpreta di solito come “un punto scelto a caso su C”. Notare che, se A è un intervallo di $\mathbb{R}$, allora

$$P\{X \in A\} = \frac{1}{m(C)} \int_A \mathbf{1}_C(x)\ dx = \frac{m(A \cap C)}{m(C)}.$$

Talvolta C verrà chiamato **supporto** della distribuzione di X. Va osservato che variabili aleatorie uniformi su $[a, b]$, (a, b), $[a, b)$, $(a, b]$ hanno la stessa distribuzione, in quanto le relative densità differiscono in al più due punti. Non c'è alcun bisogno, dunque, di distinguere tra tali alternative.

Esempio 3.3.6 (variabili aleatorie esponenziali) Diciamo che X è una variabile aleatoria **esponenziale** di parametro λ (e scriveremo $X \sim \text{Exp}(\lambda)$) se

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

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

Proposizione 3.3.2 *Sia $X \sim \text{Exp}(\lambda)$. Allora per ogni $s, t > 0$*

$$P\{X > s + t \mid X > s\} = P\{X > t\}.$$

Dimostrazione. Osservando che

$$P\{X > s + t \mid X > s\} = \frac{P\{X > s + t\}}{P\{X \geq s\}},$$

e che

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

il risultato desiderato segue immediatamente. ∎

Come risultato ausiliario, la precedente dimostrazione ci fornisce il valore della funzione di ripartizione: per ogni $t \geq 0$,

$$F_X(t) = P\{X \leq t\} = 1 - P\{X > t\} = 1 - e^{-\lambda t}$$

Come per le variabili aleatorie geometriche, si può dimostrare che le variabili aleatorie esponenziali sono le uniche variabili aleatorie assolutamente continue a valori reali positivi per cui vale la precedente proprietà di assenza di memoria. La dimostrazione di tale fatto è omessa. Le variabili aleatorie 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, ...

3.4 Massimi, minimi e somme di variabili aleatorie assolutamente continue indipendenti

Siano X e Y variabili aleatorie reali indipendenti. Vogliamo ora mostrare che per il massimo, il minimo e la somma di X e Y valgono risultati analoghi a quelli per le variabili aleatorie discrete.

Definiamo innanzitutto $Z := \max(X, Y)$ e $W := \min(X, Y)$. Vogliamo determinare la distribuzione di Z e W. Quanto visto nel caso discreto nella Proposizione 2.6.4 riguardo alla funzione di ripartizione del massimo e del minimo di variabili aleatorie indipendenti continua a valere per variabili aleatorie generiche, in quanto non veniva usata la discretezza delle variabili aleatorie ma solo le proprietà della funzione di ripartizione, che abbiamo visto valere in generale. Perciò

$$\begin{aligned} F_Z(x) &= F_X(x) F_Y(x), \\ F_W(x) &= 1 - (1 - F_X(x))(1 - F_Y(x)) \end{aligned}$$

Esempio 3.4.1 Consideriamo $X \sim \text{Exp}(\lambda)$ e $Y \sim \text{Exp}(\lambda')$ indipendenti; allora $W := \min(X, Y) \sim \text{Exp}(\lambda + \lambda')$: infatti per ogni $x \geq 0$

$$F_W(x) = 1 - (1 - F_X(x))(1 - F_Y(x)) = 1 - e^{-\lambda x} e^{-\lambda' x} = 1 - e^{-(\lambda+\lambda')x}$$

che è per l'appunto la funzione di ripartizione di una variabile aleatoria $\text{Exp}(\lambda + \lambda')$.

Per la somma di variabili aleatorie assolutamente continue, presentiamo ora una generalizzazione dell'Equazione (2.5.3), vista per variabili aleatorie discrete, che non verrà dimostrata poichè la dimostrazione fa uso di strumenti tecnici oltre la portata di questo corso. Se definiamo $Z := X+Y$, allora Z è una variabile aleatoria continua di densità

$$f_Z(z) = \int_{\mathbb{R}} f_X(z-y) f_Y(y) dy = \int_{\mathbb{R}} f_X(y) f_Y(z-y) dy \tag{3.4.0}$$

Spesso quest'ultima formula si scrive

$$f_{X+Y} = f_X * f_Y,$$

dove "$*$" denota il cosiddetto **prodotto di convoluzione**, definito appunto da

$$g * h(z) := \int_{\mathbb{R}} g(z-y) h(y) dy.$$

Come si verifica facilmente, il prodotto di convoluzione è commutativo, cioè

$$g * h(z) = h * g(z) = \int_{\mathbb{R}} h(z-y) g(y) dy.$$

Esempio 3.4.2 (variabili aleatorie Gamma) Una variabile aleatoria reale assolutamente continua X è detta **variabile aleatoria Gamma** di parametri $a, \lambda > 0$ (e scriveremo $X \sim \Gamma(\alpha, \lambda)$), se

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

dove, per ogni $a > 0$, la quantità $\Gamma(a)$ è la cosiddetta **funzione Gamma di Eulero**, definita da

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

Il valore di $\Gamma(\alpha)$ è noto esplicitamente per vari valori di a. Infatti, osservando che

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

e che, integrando per parti

$$\Gamma(a+1) = \int_0^{+\infty} x^a e^{-x} \, dx = \left[-x^a e^{-x}\right]_0^{+\infty} + a \int_0^{+\infty} x^{a-1} e^{-x} \, dx = a\Gamma(a),$$

si ha

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

per $n \in \mathbb{N} \setminus \{0\}$.

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

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

ossia

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

e questo assicura che $\Gamma(a)$ è la quantità giusta affinchè f_X sia una densità. In particolare, se $a = 1$, si ottiene che $X \sim \mathrm{Exp}(\lambda)$.

Mostriamo ora che le variabili aleatorie Gamma sono stabili per somma indipendente.

Proposizione 3.4.1 *Siano $X \sim \Gamma(a, \lambda)$ e $Y \sim \Gamma(b, \lambda)$ indipendenti. Allora $X + Y \sim \Gamma(a + b, \lambda)$.*

Dimostrazione. Basta applicare l'Equazione (3.4.0): per ogni $z \geq 0$,

$$\begin{aligned} f_{X+Y}(z) &= \int_{\mathbb{R}} f_X(z-y) f_Y(y) dy = \\ &= \int_{\mathbb{R}} \frac{1}{\Gamma(a)} \lambda^a (z-y)^{a-1} e^{-\lambda(z-y)} \mathbf{1}_{[0,+\infty)}(z-y) \frac{1}{\Gamma(b)} \lambda^b y^{b-1} e^{-\lambda y} \mathbf{1}_{[0,+\infty)}(y)\, dy = \\ &= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} e^{-\lambda z} \int_0^z (z-y)^{a-1} y^{b-1}\, dy \end{aligned}$$

Facendo il cambio di variabile $x := y/z$, si ottiene

$$f_{X+Y}(z) = \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} e^{-\lambda z} \int_0^1 z^{a-1}(1-x)^{a-1} z^{b-1} x^{b-1} z\, dx = \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} z^{a+b-1} e^{-\lambda z} B(a,b)$$

dove

$$B(a,b) := \int_0^1 (1-x)^{a-1} x^{b-1}\, dx$$

è una quantità che non dipende più da z, chiamata **funzione Beta**. Si ha quindi che, a meno di una costante di normalizzazione, f_{X+Y} è la densità di una variabile aleatoria $\Gamma(a + b)$, e segue la tesi. ∎

Osservazione 3.4.3 Come sottoprodotto della dimostrazione, siccome questa costante di normalizzazione dovrà essere per forza $\frac{\lambda^{a+b}}{\Gamma(a+b)}$, si ottiene l'identità

$$\frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)} B(a,b) = \frac{\lambda^{a+b}}{\Gamma(a+b)}$$

che si può riscrivere come

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

Esempio 3.4.4 (variabili aleatorie normali o gaussiane) Una variabile aleatoria reale assolutamente continua X si dice **normale** o **gaussiana**, e si scrive $X \sim N(\mu, \sigma^2)$, se

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

Innanzitutto notiamo che la funzione $e^{-1/2x^2}$ (come le sue trasformazioni affini) non ammette primitiva, e quindi il fatto che la costante di normalizzazione giusta sia proprio $\frac{1}{\sqrt{2\pi\sigma^2}}$ non è facilmente verificabile con l'integrazione in una variabile (cosa che comunque si dimostra agevolmente con l'integrazione in più variabili). Una conseguenza del fatto che la densità non ammette primitiva è che la funzione di ripartizione è una funzione integrale non esplicitabile in termini di funzioni elementari.

Le variabili aleatorie gaussiane hanno molte proprietà notevoli. La prima che vediamo è la sua invarianza per traslazioni affini. Difatti, se $Y \sim N(0,1)$, $\mu \in \mathbb{R}$ e $\sigma > 0$, e definiamo

$$X := \sigma Y + \mu,$$

allora la funzione di ripartizione di X è

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

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

$$f_X(x) = F'_X(x) = \frac{1}{\sigma}F'_Y\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma}f_Y\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3.4.1}$$

e quindi $X \sim N(\mu, \sigma^2)$. Allo stesso modo 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

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

Nonostante il fatto che la funzione di ripartizione non sia calcolabile esplicitamente, le variabili aleatorie 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. Per il calcolo di probabilità relative a gaussiane, abbiamo visto che si può ricondurre il calcolo alla funzione di ripartizione di una $N(0,1)$: per questo motivo, in molti libri sono presenti tavole numeriche per il calcolo della funzione di ripartizione, così come in quasi tutti i linguaggi di programmazione è implementata una funzione opportuna per eseguire lo stesso calcolo.

Vediamo ora che anche le variabili aleatorie gaussiane sono stabili rispetto alla somma indipendente.

Proposizione 3.4.2 *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 $\mu_1 = \mu_2 = 0$. Per quanto visto nell'Esempio **??**, si ha

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

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

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

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 $\mu_1 = \mu_2 = 0$.

Se in generale $\mu_1 \neq 0$, $\mu_2 \neq 0$, abbiamo che $(X-\mu_1)+(Y-\mu_2) \sim N(0, \sigma_1^2+\sigma_2^2)$, e per quanto appena visto abbiamo che $X+Y = (X-\mu_1)+(Y-\mu_2)+\mu_1+\mu_2 \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$. ∎

3.5 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.

Abbiamo visto che per una variabile aleatoria discreta è possibile definire la densità discreta

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

e la speranza matematica tramite

$$E[X] = \sum_{x \in E_0} x P\{X = x\}$$

dove

$$E_0 := \{x \in \mathbb{R} \mid p_X(x) \neq 0\}$$

Se però X non è una variabile aleatoria discreta, non ha alcun senso definite il valor medio di X tramite $E[X] = \sum_{x \in E} x P\{X = x\}$: abbiamo infatti visto che esistono variabili aleatorie per cui $P\{X = x\} = 0$ per ogni $x \in \mathbb{R}$ (ad esempio se X è assolutamente continua), e quindi ogni valor medio sarebbe nullo.

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

Definizione 3.5.1 Sia X una variabile aleatoria reale. Se $X \geq 0$, definiamo

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

(in particolare, $E[X]$ è dato da (2.7.2) se X è discreta). Per una X in generale non positiva o nulla, diremo che X ammette valor medio se $E[X^+] < +\infty$ e $E[X^-] < +\infty$, e in tal caso definiamo

$$E[X] = E[X^+] - E[X^-].$$

Infine, se X è una variabile aleatoria complessa, posto $X = \mathrm{Re}(X) + i\mathrm{Im}(X)$, diciamo che X ammette valor medio se $\mathrm{Re}(X)$ e $\mathrm{Im}(X)$ ammettono entrambe valor medio, e in questo caso poniamo

$$E[X] = E[\mathrm{Re}(X)] + iE[\mathrm{Im}(X)].$$

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

1. *(Monotonia) Se X, Y sono a valori reali, entrambe ammettono valor medio e $X(\omega) \leq Y(\omega)$, allora $E[X] \leq E[Y]$.*
2. *Se X e Y sono i.d., allora $E[X] = E[Y]$.*
3. *X ammette valor medio se e solo se $|X|$ ammette valor medio, e in tal caso*
$$|E[X]| \leq E[|X|].$$
4. *(Linearità) Se X e Y ammettono valor medio e $a, b \in \mathbb{C}$, allora la variabile aleatoria $aX + bY$ ammette valor medio e*
$$E[aX + bY] = aE[X] + bE[Y].$$

Analogamente al caso discreto, a partire dalla definizione di valore medio si definiscono varianza, covarianza e momenti di una variabile aleatoria. I risultati corrispondenti al caso discreto sono validi in questo contesto più generale. Infine, è possibile dimostrare che vale la seguente proprietà moltiplicativa per i valori medi di variabili aleatorie indipendenti, che generalizza la Proposizione 2.9.1.

Proposizione 3.5.3 *Siano X, Y variabili aleatorie reali indipendenti, definite nello stesso spazio di probabilità e che ammettono valor medio. Allora XY ammette valor medio e*
$$E[XY] = E[X]E[Y].$$

Un caso molto importante è quello delle variabili aleatorie continue, per cui sussistono delle formule analoghe a quelle del caso discreto. Innanzitutto è possibile dimostrare il seguente risultato (ma con strumenti oltre gli scopi di questo corso).

Proposizione 3.5.4 *Sia X una variabile aleatoria assolutamente continua a valori in $\mathbb{R}$, con densità f_X. Se $g : \mathbb{R} \to \mathbb{R}$ è misurabile, allora $g(X)$ ammette valor medio se e solo se la funzione $|g(x)|f_X(x)$ è integrabile su $\mathbb{R}$, e in questo caso*
$$E[g(X)] = \int_{\mathbb{R}} g(x) f_X(x)\, dx.$$
In particolare, se X è una variabile aleatoria reale con densità f, allora X ammette valor medio se e solo se $|x|f(x)$ è integrabile su $\mathbb{R}$ e in questo caso
$$E[X] = \int_{\mathbb{R}} x f(x)\, dx.$$

Da questa espressione generale per la media di una funzione di una variabile aleatoria è possibile ricavare le espressioni per tutte le quantità viste nel Capitolo 2.

Vediamo ora alcuni esempi.

Esempio 3.5.1 (variabili aleatorie uniformi) Calcoliamo ora media e varianza di $X \sim U(a,b)$. Innanzitutto, siccome la densità ha supporto limitato, gli integrali esistono e sono sicuramente finiti. Abbiamo quindi

$$\begin{aligned}
E[X] &= \frac{1}{b-a}\int_a^b x\,dx = \frac{a+b}{2}, \\
\mathrm{Var}[X] &= E[(X-E[X])^2] = \frac{1}{b-a}\int_a^b \left(x-\frac{a+b}{2}\right)^2 dx = \frac{1}{b-a}\left[\frac{1}{3}\left(x-\frac{a+b}{2}\right)^3\right]_a^b = \frac{(b-a)^3}{12}
\end{aligned}$$

Esempio 3.5.2 (variabili aleatorie Gamma) Se $X \sim \Gamma(a,\lambda)$, calcoliamo $E[X^n]$ per $n \geq 1$. Siccome il supporto della densità è uguale a $\mathbb{R}^+$, questa speranza è sicuramente ben definita, eventualmente uguale a $+\infty$. Abbiamo però

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

e quindi in particolare $E[X^n] < +\infty$ per ogni $n \geq 1$. Da questo facilmente segue

$$\begin{aligned}
E[X] &= \frac{\Gamma(a+1)}{\lambda\Gamma(a)} = \frac{a}{\lambda}, \\
\mathrm{Var}[X] &= E[X^2]-E[X]^2 = \frac{\Gamma(a+2)}{\lambda^2\Gamma(a)} - \left(\frac{a}{\lambda}\right)^2 = \frac{a(a+1)}{\lambda^2} - \frac{a^2}{\lambda^2} = \frac{a}{\lambda^2}.
\end{aligned}$$

In particolare, se $X \sim \mathrm{Exp}(\lambda)$, si ha

$$E[X] = \frac{1}{\lambda}, \qquad \mathrm{Var}[X] = \frac{1}{\lambda^2}$$

Esempio 3.5.3 (variabili aleatorie gaussiane) È facile vedere che, se $X \sim N(0,1)$, allora

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} |x|e^{-\frac{x^2}{2}}\,dx < +\infty$$

e quindi

$$E[X] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} xe^{-\frac{x^2}{2}}\,dx = 0$$

dato che l'integrando è una funzione dispari. Inoltre, integrando per parti,

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

Abbiamo visto che, se $X \sim N(\mu,\sigma^2)$, $\mu \in \mathbb{R}$ e $\sigma > 0$, allora $Y := \frac{X-\mu}{\sigma} \sim N(0,1)$. Questo significa che

$$\begin{aligned}
E[X] &= E[\sigma Y + \mu] = \sigma E[Y] + \mu = \mu, \\
\mathrm{Var}[X] &= \mathrm{Var}[\sigma Y + \mu] = \mathrm{Var}[\sigma Y] = \sigma^2\mathrm{Var}[Y] = \sigma^2
\end{aligned}$$

e quindi i ruoli dei parametri nella definizione di legge gaussiana sono rispettivamente di media per μ e varianza per σ^2.

3.6 Disuguaglianze

Come in molti altri settori della matematica, in probabilità le disuguaglianze giocano un ruolo fondamentale.

Proposizione 3.6.1 (Disuguaglianza di Markov) *Sia $X \in L^1(\Omega, \mathcal{A}, P)$ positiva, che ammette valor medio. Allora, per ogni $a > 0$,*

$$P\{X \geq a\} \leq \frac{E[X]}{a}.$$

Dimostrazione. Se consideriamo la variabile aleatoria $a\mathbf{1}_{\{}X \geq a\}$, allora abbiamo $X \geq a\mathbf{1}_{\{}X \geq a\}$, e quindi per la monotonia della speranza si ha

$$E[X] \geq E[a\mathbf{1}_{\{}X \geq a\}] = aP\{X \geq a\}$$

Invertendo la relazione, segue la tesi. ∎

Proposizione 3.6.2 (Disuguaglianza di Chebichev) *Sia $X \in L^2(\Omega, \mathcal{A}, P)$. Allora, per ogni $a > 0$,*

$$P\{|X - E[X]| > a\} \leq \frac{\mathrm{Var}[X]}{a^2}.$$

Dimostrazione. Poichè

$$P\{|X - E[X]| > a\} = P\{(X - E[X])^2 > a^2\},$$

è sufficiente applicare la disuguaglianza di Markov a $(X - E[X])^2$ e ad a^2. ∎

Proposizione 3.6.3 (Disuguaglianze di Chebichev unilatere) *Sia $X \in L^2(\Omega, \mathcal{A}, P)$ con media m e varianza σ^2. Allora, per ogni $a > 0$,*

$$P\{X > m + a\} \leq \frac{\sigma^2}{\sigma^2 + a^2}, \qquad P\{X < m - a\} \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

Dimostrazione. Supponiamo all'inizio che $m = 0$. Allora per ogni $b > 0$ si ha

$$P\{X > a\} = P\{X + b > a + b\} \leq P\{(X + b)^2 > (a + b)^2\} \leq \frac{E[(X + b)^2]}{(a + b)^2} = \frac{E[X^2] + b^2}{(a + b)^2} = \frac{\sigma^2 + b^2}{(a + b)^2}$$

dove la prima disuguaglianza segue dal fatto che $a + b > 0$, e la seconda dalla disuguaglianza di Markov. Siccome questa disuguaglianza vale per ogni $b > 0$, deve valere anche per il minimo di $f(b) := \frac{\sigma^2+b^2}{(a+b)^2}$. Derivando questa funzione si ottiene

$$f'(b) = \frac{2(ab - \sigma^2)}{(a + b)^3} \geq 0 \quad \text{se e solo se } b \geq \frac{\sigma^2}{a}$$

quindi

$$P\{X > a\} \leq f\left(\frac{\sigma^2}{a}\right) = \frac{\sigma^2}{\sigma^2 + a^2}$$

Se m è un generico numero reale, in generale le variabili aleatorie $X - m$ e $m - X$ hanno media nulla, e quindi applicando quanto appena dimostrato si ha

$$P\{X > m + a\} = P\{X - m > a\} \leq \frac{\sigma^2}{\sigma^2 + a^2}, \qquad P\{X < m - a\} = P\{m - X > a\} \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

∎

Esercizi

Esercizio 65 Sia X una variabile aleatoria uniformenente distribuita nell'intervallo $(-\frac{\pi}{2}, \frac{\pi}{2})$, cioè

$$f_X(x) = \frac{1}{\pi} 1_{(-\frac{\pi}{2},\frac{\pi}{2})}(x).$$

Posto $Y = \cos(X)$, trovare la distribuzione di Y.

Esercizio 66 Si scelga a caso un punto X dell'intervallo $[0, 2]$, con distribuzione uniforme di densità

$$f_X(x) = \frac{1}{2} 1_{[0,2]}(x)$$

(in altre parole, X è una variabile aleatoria con densità f_X). Qual è la probabilità che il triangolo equilatero il cui lato ha lunghezza X abbia area maggiore di 1?

Esercizio 67 * Si scelga a caso un angolo $\Theta \in (0, \frac{\pi}{2})$ con distribuzione uniforme di densità

$$f_\Theta(\theta) = \frac{2}{\pi} 1_{(0,\frac{\pi}{2})}(\theta),$$

e si consideri il punto del piano di coordinate $(\cos(\Theta), \sin(\Theta))$. Per tale punto si tracci la tangente alla circonferenza di centro $(0, 0)$ e raggio 1, e sia L la lunghezza del segmento i cui estremi sono i punti d'intersezione di tale tangente con gli assi cartesiani. Determinare la distribuzione di L.

Esercizio 68 Sia $X \sim U(-1, 1)$, e $Y = \frac{1}{1-X^2}$. Determinare la distribuzione di Y.

Esercizio 69 * Sia X una variabile aleatoria reale assolutamente continua, e tale che $P\{X > 0\} = 1$. Tramite un'opportuna integrazione per parti mostrare che

$$E[X] = \int_0^{+\infty} P\{X > t\}\, dt.$$

Esercizio 70 Siano $X_1, X_2, \ldots, X_n \sim \text{Exp}(\lambda)$ indipendenti, e

$$Y = \max(X_1, \ldots, X_n), \quad Z = \min(X_1, \ldots, X_n).$$

Determinare la distribuzione di Y e Z.

Esercizio 71 Siano $X, Y \sim N(0, 1)$ indipendenti. Determinare la densità congiunta di $(X+Y, X-Y)$, e mostrare che esse sono indipendenti. Inoltre calcolare le densità marginali di $X + Y$ e $X - Y$.

Esercizio 72 Siano X e Y due variabili aleatorie reali, assolutamente continue e indipendenti, con densità f_X e f_Y rispettivamente.

a. Mostrare che, posto $Z = X - Y$,

$$f_Z(z) = \int f_X(x+z) f_Y(x)\, dx.$$

b. Usando la formula nel punto a., determinare f_Z nel caso in cui $X \sim \text{Exp}(1)$ e $Y \sim \text{Exp}(1)$.

Esercizio 73 Siano $X \sim \text{Exp}(1)$ e $Y \sim \text{Exp}(2)$ indipendenti.

a. Scrivere la densità congiunta del vettore (X, Y).

b. Sia T_t il triangolo di vertici $(0,0)$, $(0,t)$ e (t,t), dove $t > 0$. Calcolare $P\{(X,Y) \in T_t\}$.

c. Calcolare $P\{X \leq Y\}$.

Esercizio 74 Sia T il triangolo di vertici $(0,0)$, $(0,1)$ e $(1,1)$. Sia (X,Y) un vettore aleatorio assolutamente continuo con densità congiunta

$$f_{X,Y}(x,y) = -4\log(y) 1_T(x,y).$$

a. Determinare le densità marginali di X e Y.

b. Calcolare, se esistono, media e varianza di Y.

Esercizio 75 * Sia X una variabile aleatoria reale assolutamente continua con densità f_X, e sia Y una variabile aleatoria reale discreta di densità discreta p_Y. Supponiamo che Y assuma solo un numero finito di valori, cioè l'insieme $\{x : p_Y(x) > 0\} = \{x_1, \ldots, x_n\}$ è finito, e che X e Y siano indipendenti.

a. Esprimere la funzione di ripartizione di $X + Y$ in termini di f_X e p_Y.

b. Mostrare che $X + Y$ è una variabile aleatoria assolutamente continua, e determinarne la densità in funzione di f_X e p_Y.

Esercizio 76 Sia (X,Y) un vettore aleatorio di dimensione 2 con densità uniforme

$$f_{X,Y}(x,y) = \frac{2}{\pi} 1_\Gamma(x,y),$$

dove $\Gamma = \{(x,y) : x \geq 0, x^2 + y^2 \leq 1\}$.

a. Determinare le densità marginali di X e Y.

b. Calcolare $\text{Cov}(X,Y)$

Esercizio 77 * Siano $X \sim N(0,1)$ e $T \sim \text{Exp}(1/2)$ indipendenti. Si definiscano $Y = X/\sqrt{T}$ e $S = X^2 + T$.

a. Determinare la densità congiunta di (Y, S)

b. Y e S sono variabili aleatorie indipendenti? (Motivare la risposta!)

Esercizio 78 Sia X una variabile aleatoria la cui funzione di ripartizione è

$$F_X(t) = \begin{cases} 0 & \text{per } x < 0 \\ x/2 & \text{per } 0 \leq x < 1 \\ 1 & \text{per } x \geq 1 \end{cases}$$

Sia inoltre $Y \sim U(0,1)$ indipendente da X, e si definisca

$$Z(\omega) = \begin{cases} X(\omega) & \text{se } X(\omega) < 1 \\ Y(\omega) & \text{se } X(\omega) = 1 \end{cases}$$

Determinare la distribuzione di Z (sugg: determinare la funzione di ripartizione di Z)

Esercizio 79 Un congegno elettronico è costituito da n componenti collegate in serie: esso smette di funzionare non appena una qualsiasi delle sue componenti smette di funzionare. Siano $T_1, T_2, \ldots, T_n$ i tempi di vita delle n componenti, che si assumono indipendenti e identicamente distribuiti con distribuzione esponenziale di parametro 1. Sia X_n il tempo di vita dell'intero dispositivo.

a. Determinare la distribuzione di X_n.

b. Mostrare che, per ogni $\varepsilon > 0$

$$\lim_{n \to +\infty} P\{X_n > \varepsilon\} = 0$$

Esercizio 80 * Siano $Y \sim U(-\pi/2, \pi/2)$ e X una variabile aleatoria assolutamente continua con densità

$$f_X(x) = xe^{-\frac{x^2}{2}} 1_{[0,+\infty)}(x).$$

Si assuma che X e Y siano indipendenti. Definite

$$\begin{aligned} Z &= X \cos Y \\ W &= X \sin Y \end{aligned}$$

determinare la distribuzione congiunta e le distribuzioni marginali di (Z, W).(Sugg: osservare che $X = \sqrt{Z^2 + W^2}$ e $Y = \arctan(W/Z)$).

Esercizio 81 Sia X una variabile aleatoria reale assolutamente continua con densità

$$f_X(x) = -\log(x) 1_{(0,1)}(x).$$

a. Determinare la funzione di ripartizione di X.

b. Sia $Y = -\log X$. Determinare la distribuzione di Y. In particolare, mostrare che Y è una variabile Gamma, e determinarne i parametri.

Esercizio 82 L'ufficio informazioni di Trenitalia ha due numeri verdi. Il tempo di attesa T_i, $i = 1, 2$ per parlare con l'operatore è, per entrambi i numeri, una variabile aleatoria esponenziale di media $\mu = 15$ minuti. Inoltre T_1 e T_2 si possono considerare indipendenti. Avendo a disposizione due telefoni, decido di chiamare contemporaneamente i due numeri, in modo da parlare con l'operatore che per primo risponderà.

a. Quanto tempo, in media, dovrò aspettare per parlare con un operatore?

b. Qual è la probabilità di attendere meno di 5 minuti?

Esercizio 83 * Sia X una variabile aleatoria reale, e sia F_X la sua funzione di ripartizione.

a. Supponendo che F_X sia invertibile (e perciò anche continua), determinare la distribuzione della variabile aleatoria $Y = F_X(X)$.

b*. Mostrare che il risultato al punto a. vale anche assumendo solo la continuità di F_X.

c. Mostrare, con un controesempio, che invece il risultato al punto a. può essere falso so F_X non è continua.

d. Mostrare che il risultato al punto a. è *sicuramente* falso se F_X è discontinua (pensare al supporto della distribuzione di Y).

Esercizio 84 Sia (X, Y) una vettore aleatorio assolutamente continuo con densità congiunta

$$f_{X,Y}(x, y) = 3x 1_{(0,x)}(y) 1_{(0,1)}(x).$$

a. Trovare le densità marginali f_X e f_Y.

b. Usando un opportuno cambiamento di variabili, determinare la densità della variabile aleatoria $Z = X - Y$.

Esercizio 85 Siano $X \sim \mathrm{Exp}(1)$ e $Y \sim \mathrm{Exp}(1/2)$ variabili aleatorie indipendenti. Determinare la densità di $X + Y$.

Esercizio 86 Sia (X, Y) un vettore aleatorio con densità congiunta

$$f_{X,Y}(x, y) = 2e^{-(x+y)} 1_A(x, y),$$

dove $A = \{(x, y) : 0 \leq y \leq x\}$.

a. X e Y sono indipendenti? (Giustificare la risposta!)

b. Determinare le densità marginali f_X e f_Y.

c. Calcolare $E\left[(X - Y)^2\right]$

Esercizio 87 Sia Y un numero casuale generato da un calcolatore con distribuzione uniforme in $[0, 1]$. Dopo aver generato Y, supponiamo di voler generare un nuovo numero casuale W scelto con distribuzione uniforme tra Y e 1. Per fare ciò, generiamo un numero casuale $X \sim U(0, 1)$ indipendente da Y, e poniamo

$$W = XY + 1 - X.$$

a. Determinare la densità congiunta di (Y, W).

b. Determinare la densità di W.

c*. Calcolare il coefficente di correlazione $\rho_{Y,W}$.

Esercizio 88 Siano $X, Y \sim U(0, 1)$ indipendenti.

a. Determinare la densità congiunta di $\left(\frac{X}{Y}, Y\right)$.

b. Determinare la densità di $\frac{X}{Y}$.