# Capitolo 3

# Condizionamento e martingale

## 3.1 Condizionamento

Sia $(\Omega, \mathscr{F}, \mathbb{P})$ uno spazio di probabilità. Dunque, ricordiamo,

$\Omega$ è un insieme (lo *spazio campionario*);

$\mathscr{F}$ è una $\sigma$-algebra di sottoinsiemi di $\Omega$ (gli *eventi*) ;

$\mathbb{P}$ è una probabilità su $(\Omega, \mathscr{F})$, cioè $\mathbb{P}$ è un'applicazione $\mathscr{F} \to [0,1]$ t.c.

$\mathbb{P}(\Omega) = 1$,

se $A_n \cap A_m = \emptyset$, $n \neq m$, $\mathbb{P}\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$ ($\sigma$-*additività*).

Ricordiamo che, dati due eventi $A, B \in \mathscr{F}$ con $\mathbb{P}(B) > 0$, la probabilità condizionata di $A$ dato $B$ è definita da

$$\mathbb{P}(A \,|\, B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Intuitivamente, $\mathbb{P}(A \,|\, B)$ rappresenta la probabilità del verificarsi dell'evento $A$, una volta noto il fatto che l'evento $B$ si è verificato.

Vogliamo ora definire la *media condizionale* di una v.a. $X$ data una $\sigma$-algebra $\mathscr{G} \subset \mathscr{F}$ e, in seguito, la *probabilità condizionale* di un evento qualsiasi $A \in \mathscr{F}$ data una $\sigma$-algebra $\mathscr{G}$. Poiché si tratta di oggetti alquanto delicati e spesso inizialmente di difficile comprensione, procederemo per gradi.

### 3.1.1 Media condizionale: il caso discreto

Cominciamo con un caso semplice, anche se tipico di condizionamento.

**Esempio 3.1.1.** Supponiamo che $(X, Y)$ sia una coppia di v.a. discrete, con densità (discreta) congiunta $p_{X\,Y}$ e densità marginali $p_X$ e $p_Y$. Siano $E_X$ ed $E_Y$ i possibili valori per $X$ e $Y$ rispettivamente, e supponiamo che $p_Y(y) > 0$ per ogni $y \in E_Y$. Per $X$, supponiamo che abbia speranza matematica finita.

Abbiamo visto che la densità discreta di $X$ condizionata a $\{Y = y\}$, è data da

$$p_{X\,|\,Y}(x\,|\,y) = \frac{p_{XY}(x,y)}{p_Y(y)}, \quad \text{purché } y \in E_Y.$$

La media condizionale di $X$ dato $\{Y = y\}$ è semplicemente la media di $X$ quando per $X$ si consideri la distribuzione condizionale:

$$\begin{aligned}
\mathbb{E}(X\,|\,Y = y) = & \textstyle\sum_{x\in E_X} x\, p_{X\,|\,Y}(x\,|\,y) = \displaystyle\frac{1}{p_Y(y)} \sum_{x\in E_X} x\, p_{X\,Y}(x,y) \\
= & \frac{1}{p_Y(y)} \sum_{(x,y)\in E_X\times\{y\}} x\, p_{X\,Y}(x,y) \\
= & \frac{1}{p_Y(y)}\, \mathbb{E}(X\, 1_{\{Y=y\}}).
\end{aligned}$$

Se poniamo, per $y \in E_Y$, $h_X(y) = \mathbb{E}(X\,1_{\{Y=y\}})/p_Y(y)$, la v.a. $Z = h_X(Y)$ dà la media di $X$ dato il risultato (aleatorio) $Y$. Ad essa può quindi essere dato il significato di *media condizionale di $X$ dato $Y$.*

Osserviamo che $Z = h_X(Y)$ verifica le seguenti proprietà.

1. $Z = h_X(Y)$ è $\sigma(Y)$-misurabile. Infatti è una funzione (misurabile) di $Y$.

2. Per ogni $B \subset E_Y$, si ha

   $$\mathbb{E}(h_X(Y)\,1_{\{Y\in B\}}) = \mathbb{E}(X\,1_{\{Y\in B\}})$$

   Infatti,

   $$\mathbb{E}(h_X(Y)\,1_{\{Y\in B\}}) = \sum_{y\in B} h_X(y)\,p_Y(y) = \sum_{y\in B} \mathbb{E}(X\,1_{\{Y=y\}}) = \mathbb{E}(X\,1_{\{Y\in B\}}).$$

Infine, osserviamo che la v.a. $Z = h_X(Y)$ è l'unica v.a. che soddisfa alle condizioni 1. e 2. Infatti, supponiamo che $\tilde{Z}$ sia un'altra v.a. che soddisfa a 1. e 2. Poiché da 1. $\tilde{Z}$ è $\sigma(Y)$-misurabile dev'essere (cfr. Proposizione 2.1.16) $\tilde{Z} = \tilde{h}(Y)$, per qualche funzione (misurabile) $\tilde{h}$. Ora, usando 2. con la scelta $B = \{Y = y\} = Y^{-1}(y) = \{Y = y\}(\in \sigma(Y))$, si ha

$$\mathbb{E}(X\,1_{\{Y=y\}}) = \mathbb{E}(\tilde{Z}\,1_{\{Y=y\}}) = \mathbb{E}(\tilde{h}(Y)\,1_{\{Y=y\}}) = \tilde{h}(y)\,\mathbb{E}(1_{\{Y=y\}})$$

e cioè $\tilde{h}(y) = \mathbb{E}(X\,1_{\{Y=y\}})/\mathbb{P}(Y = y) = h_X(y)$.

Analizziamo ora un caso un po' più generale.

**Esempio 3.1.2.** Sia $X$ una v.a. con speranza matematica finita e sia $\mathscr{G}$ una $\sigma$-algebra, con $\mathscr{G} \subset \mathscr{F}$. Supponiamo che $\mathscr{G}$ sia generata da una partizione al più numerabile: $\mathscr{G} = \sigma(\mathscr{C})$, con $\mathscr{C} = \{C_i\}_{i \in I}$. Supponiamo anche che $\mathbb{P}(C_i) > 0$ per ogni $i \in I$.

Osserviamo che nell'Esempio 3.1.1 abbiamo implicitamente lavorato con questo tipo di ingredienti. Infatti,

$$\sigma(Y) = \mathscr{G} = \sigma(\mathscr{C}_Y), \quad C \in \mathscr{C}_Y \text{ se e solo se } C = Y^{-1}(\{y\}), \text{ con } y \in E_Y.$$

Ricordiamo inoltre che la media condizionale di $X$ dato $Y$ è

$$Z(\omega) = h_X(Y(\omega)) = \frac{\mathbb{E}(X\,1_{\{Y=y\}})}{\mathbb{P}(Y=y)}, \quad \text{quando } \omega \in \{Y = y\}.$$

Ora, gli eventi $\{Y = y\}$, con $y \in E_Y$, sono gli atomi della partizione $\mathscr{C}_Y$ che genera $\mathscr{G} = \sigma(Y)$. Volendo quindi generalizzare ad una qualsiasi $\sigma$-algebra $\mathscr{G}$ generata da una partizione $\mathscr{C}$, viene naturale considerare la v.a.

$$Z(\omega) = \sum_{i \in I} \frac{\mathbb{E}(X\,1_{C_i})}{\mathbb{P}(C_i)}\, 1_{\{\omega \in C_i\}} = \frac{\mathbb{E}(X\,1_{C_k})}{\mathbb{P}(C_k)} \quad \text{quando } \omega \in C_k$$

Come vedremo in seguito, la v.a. $Z$ è la media condizionale di $X$ data la $\sigma$-algebra $\mathscr{G}$. Osserviamo che abbiamo fatto un paio di passi in avanti rispetto all'Esempio 3.1.1: abbiamo una $\sigma$-algebra $\mathscr{G}$ (almeno apparentemente) più generale e, soprattutto, non stiamo facendo ipotesi speciali sulla v.a. $X$: l'importante è che abbia senso la quantità $\mathbb{E}(X\,1_{C_k})$, dunque l'importante è che $X$ abbia media.

La v.a. $Z$ verifica le seguenti proprietà, analoghe a quelle già viste nell'Esempio 3.1.1.

1. $Z$ è $\mathscr{G}$-misurabile. Infatti, $1_{C_i}$ è $\mathscr{G}$-misurabile per ogni $i$, ed essendo $Z$ una combinazione lineare delle funzioni $1_{C_i}$, è $\mathscr{G}$-misurabile.

2. Per ogni $G \in \mathscr{G}$, si ha

   $$\mathbb{E}(Z\,1_G) = \mathbb{E}(X\,1_G).$$

   Infatti, preso $G \in \mathscr{G}$, sia $I_G \subset I$ tale che $G = \cup_{i \in I_G} C_i$. Allora, $Z\,1_G = \sum_{i \in I_G} \frac{\mathbb{E}(X\,1_{C_i})}{\mathbb{P}(C_i)}\, 1_{C_i}$ e quindi

   $$\begin{aligned} \mathbb{E}(Z\,1_G) &= \sum_{i \in I_G} \frac{\mathbb{E}(X\,1_{C_i})}{\mathbb{P}(C_i)}\, \mathbb{P}(C_i) = \sum_{i \in I_G} \mathbb{E}(X\,1_{C_i}) \\ &= \mathbb{E}(X\,1_{\cup_{i \in I_G} C_i}) = \mathbb{E}(X\,1_G). \end{aligned}$$

Infine, osserviamo che anche in questo caso la v.a. $Z$ è l'unica v.a. che soddisfa alle condizioni 1. e 2. Infatti, da 1. $Z$ è $\mathscr{G}$-misurabile, dunque per la Proposizione 2.1.18 è costante sugli elementi della partizione: $Z(\omega) = z_i$ per ogni $\omega \in C_i$. Allora, da 2. si ha

$$\mathbb{E}(X\,1_{C_i}) = \mathbb{E}(Z\,1_{C_i}) = z_i\,\mathbb{E}(1_{C_i}) = z_i\,\mathbb{P}(C_i),$$

cioè $z_i = \mathbb{E}(X\,1_{C_i})/\mathbb{P}(C_i)$, da cui la tesi.

### 3.1.2 Il caso generale

Negli Esempi 3.1.1 e 3.1.2, abbiamo visto che le proprietà 1. e 2. si rivelano caratteristiche della v.a. $Z$, che intuitivamente abbiamo chiamato *speranza condizionale* di $X$ data la v.a. $Y$ oppure data la $\sigma$-algebra $\mathscr{G}$. La definizione generale prende infatti in considerazione proprio queste due proprietà caratteristiche:

**Definizione 3.1.3.** Siano $(\Omega, \mathscr{F}, \mathbb{P})$ uno spazio di probabilità, $X$ una v.a. avente speranza matematica finita e $\mathscr{G}$ una sotto $\sigma$-algebra di $\mathscr{F}$. Una versione $\mathbb{E}(X\,|\,\mathscr{G})$ della *media (o speranza) condizionale di $X$ dato $\mathscr{G}$* è una v.a. tale che:

(i) $\mathbb{E}(X\,|\,\mathscr{G})$ è $\mathscr{G}$-misurabile e integrabile;

(ii) per ogni $G \in \mathscr{G}$,
$$\mathbb{E}(\mathbb{E}(X\,|\,\mathscr{G})\,1_G) = \mathbb{E}(X\,1_G)$$

Qualora $\mathscr{G} = \sigma(Y)$, con $Y$ v.a., allora $\mathbb{E}(X\,|\,\sigma(Y))$ prende il nome di speranza condizionale di $X$ dato $Y$ ed è denotata con il simbolo $\mathbb{E}(X\,|\,Y)$.

Abbiamo visto negli esempi che la speranza condizionale esiste se la $\sigma$-algebra $\mathscr{G}$ è generata da una partizione al più numerabile. Lasceremo a corsi più avanzati la dimostrazione dell'esistenza quando la $\sigma$-algebra è più generale. Del resto, in questo corso il caso di $\sigma$-algebre generate da una partizione sarà quello che interessa.

Nella Definizione 3.1.3 si parla di *versione* perché è abbastanza chiaro che non c'è unicità. Infatti, se $Z$ è una v.a. $\mathscr{G}$-misurabile e avente speranza matematica finita e tale che $\mathbb{E}(1_G Z) = \mathbb{E}(1_G X)$ per ogni $G \in \mathscr{G}$ e $Z'$ è un'altra v.a., sempre $\mathscr{G}$-misurabile, ma che differisce da $Z$ solo su un evento di probabilità 0, allora si ha anche $\mathbb{E}(1_G Z') = \mathbb{E}(1_G X)$ per ogni $G \in \mathscr{G}$.

Quando $\mathscr{G} = \sigma(Y)$, il calcolo della speranza condizionale prende un aspetto particolare. Abbiamo infatti visto (Proposizione 2.1.20) che ogni v.a. $\sigma(Y)$-misurabile è della forma $h(Y)$, dove $h$ è un'opportuna funzione misurabile. Quindi esiste una funzione misurabile $\Psi_X$ tale che

$$\mathbb{E}(X\,|\,Y) = \Psi_X(Y).$$

Il calcolo della speranza condizionale, in questo caso, si riconduce quindi a quello della funzione $\Psi_X$. Con abuso di linguaggio, denoteremo

$$\Psi_X(y) = \mathbb{E}(X \mid Y = y),$$

ma attenzione a non confondere la notazione "$\mathbb{E}(X \mid Y = y)$" con la media condizionale di $X$ dato l'evento $\{Y = y\}$ (che, osserviamo, in molti casi di interesse è un evento di probabilità nulla...). Nell'Esempio 3.1.1, abbiamo visto che

$$\Psi_X(y) = h_X(y) = \frac{\mathbb{E}(X\, 1_{\{Y=y\}})}{\mathbb{P}(Y = y)}.$$

In generale, la funzione $\Psi_X$ resta determinata dalla relazione

$$\mathbb{E}(X 1_{\{Y\in B\}}) = \mathbb{E}(\Psi_X(Y) 1_{\{Y\in B\}}), \quad \text{per ogni boreliano } B \tag{3.1}$$

(ricordiamo infatti che $\sigma(Y) = \{Y^{-1}(B)\,;\, B \in \mathscr{B}(E)\}$). Osserviamo che, se $X, Y$ sono v.a. a valori reali, allora siamo in grado di calcolare le due speranze matematiche che compaiono nella (3.1) non appena conosciamo le distribuzioni congiunte di $X$ e $Y$. Vediamo, nel prossimo esempio, come questo fatto si può sfruttare per calcolare $\mathbb{E}(X \mid Y)$.

**Esempio 3.1.4.** Siano $X$ e $Y$ due v.a. reali congiuntamente assolutamente continue, con $X$ avente speranza matematica finita. Indichiamo con $f$ la densità di probabilità congiunta e con $f_X$ e $f_Y$ le due densità marginali. Supponiamo, ma solo per semplicità, che $f_Y(y) > 0$.

La speranza matematica a destra nella (3.1) vale dunque

$$\mathbb{E}(X 1_{\{Y\in B\}}) = \int_B dy \int_{\mathbb{R}} x f(x, y)\, dx$$

mentre

$$\mathbb{E}(\Psi_X(Y) 1_{\{Y\in B\}}) = \int_B \Psi_X(y) f_Y(y)\, dy$$

Queste due quantità risultano uguali per ogni scelta del boreliano $B$ se scegliamo

$$\Psi_X(y) = \frac{1}{f_Y(y)} \int_{\mathbb{R}} x f(x, y)\, dx = \int_{\mathbb{R}} x f_{X\mid Y}(x \mid y)\, dx$$

dove

$$f_{X\mid Y}(x \mid y) = \frac{f(x, y)}{f_Y(y)}$$

è la (vecchia!) densità condizionale. Per concludere dobbiamo solo verificare che la v.a. $\Psi_X(Y)$ ha speranza matematica finita. Infatti

$$\int_{\mathbb{R}} |\Psi_X(y)|\, f_Y(y)\, dy \le \int_{\mathbb{R}} dx \int_{\mathbb{R}} |x|\, f(x, y)\, dy = \int_{\mathbb{R}} |x|\, f_X(x)\, dx < \infty$$

poiché supponiamo che $X$ abbia speranza matematica finita.

In conclusione, la media di $X$ condizionata a $Y$ è la media di $X$ fatta sotto la densità condizionale $f_{X\mid Y}(x \mid y)$, sostituendo poi ad $y$ il suo valore generico (aleatorio) $Y$: esattamente come vorrebbe l'intuizione!

### 3.1.3 Proprietà della speranza condizionale

Vediamo ora alcune proprietà delle speranze condizionali. Avvertenza: nel seguito l'abbreviazione "q.c." sta per "quasi certamente". Significa che le proprietà in questione (uguaglianze, disuguaglianze etc.) sono verificate per $\omega \in \Omega \setminus \mathcal{N}$, dove $\mathcal{N}$ denota un evento tale che $\mathbb{P}(\mathcal{N}) = 0$.

**1.** *Se* $X = c$, *allora* $\mathbb{E}(X \,|\, \mathscr{G}) = c$ *q.c.* Immediato!

**2.** $\mathbb{E}(\mathbb{E}(X \,|\, \mathscr{G})) = \mathbb{E}(X)$.

Infatti basta prendere $G = \Omega \in \mathscr{G}$ nel punto (ii) della Definizione 3.1.3.

**3.** (Linearità) *Se* $X$ *e* $Y$ *hanno media allora* $\mathbb{E}(\alpha X + \beta Y \,|\, \mathscr{G}) = \alpha\mathbb{E}(X \,|\, \mathscr{G}) + \beta\mathbb{E}(Y \,|\, \mathscr{G})$ *q.c., per ogni* $\alpha, \beta \in \mathbb{R}$.

Per ipotesi $\mathbb{E}(X \,|\, \mathscr{G})$ e $\mathbb{E}(Y \,|\, \mathscr{G})$ sono $\mathscr{G}$-misurabili ed hanno speranza matematica finita. Quindi lo stesso vale per $Z := \alpha\mathbb{E}(X \,|\, \mathscr{G}) + \beta\mathbb{E}(Y \,|\, \mathscr{G})$. Inoltre, se $G \in \mathscr{G}$,

$$\begin{aligned}\mathbb{E}(Z\, 1_G) &= \alpha\mathbb{E}(\mathbb{E}(X \,|\, \mathscr{G})\, 1_G) + \beta\mathbb{E}(\mathbb{E}(Y \,|\, \mathscr{G})\, 1_G) \\ &= \alpha\mathbb{E}(X\, 1_G) + \beta\mathbb{E}(Y\, 1_G) = \mathbb{E}((\alpha X + \beta Y)\, 1_G),\end{aligned}$$

dunque $Z := \alpha\mathbb{E}(X \,|\, \mathscr{G}) + \beta\mathbb{E}(Y \,|\, \mathscr{G})$ è una versione della speranza condizionale di $\alpha X + \beta Y$ dato $\mathscr{G}$.

**4.** (Monotonia) *Se* $\mathbb{P}(X \geq Y) = 1$ *allora* $\mathbb{E}(X \,|\, \mathscr{G}) \geq \mathbb{E}(Y \,|\, \mathscr{G})$ *q.c.*

Poniamo $G = \{\omega\,;\, \mathbb{E}(X \,|\, \mathscr{G}) < \mathbb{E}(Y \,|\, \mathscr{G})\}$. Evidentemente $G \in \mathscr{G}$. Se fosse $\mathbb{P}(G) > 0$ allora si avrebbe $0 > \mathbb{E}((\mathbb{E}(X \,|\, \mathscr{G}) - \mathbb{E}(Y \,|\, \mathscr{G}))\, 1_G)$, mentre invece si ha

$$\mathbb{E}((\mathbb{E}(X \,|\, \mathscr{G}) - \mathbb{E}(Y \,|\, \mathscr{G}))\, 1_G) = \mathbb{E}(\mathbb{E}(X - Y \,|\, \mathscr{G})\, 1_G) = \mathbb{E}((X - Y)\, 1_G) \geq 0,$$

il che è assurdo.

**5.** (Disuguaglianza di Jensen) *Se* $\Phi$ *è una funzione continua convessa e tale che* $\Phi(X)$ *abbia speranza matematica finita, allora*

$$\mathbb{E}(\Phi(X) \,|\, \mathscr{G}) \geq \Phi(\mathbb{E}(X \,|\, \mathscr{G})) \quad q.c.$$

*In particolare, se* $p \geq 1$ *e* $\mathbb{E}(|X|^p) < +\infty$, *allora possiamo scegliere* $\Phi(x) = |x|^p$ *e dunque* $\mathbb{E}(|X|^p \,|\, \mathscr{G}) \geq |\mathbb{E}(X \,|\, \mathscr{G})|^p$, *q.c.*

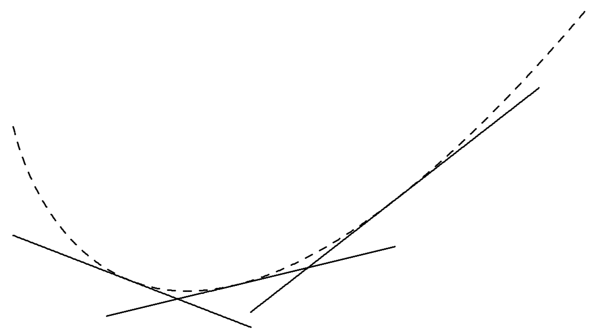

**Figura 3.1** Una funzione convessa e continua è sempre inviluppo superiore delle funzioni lineari affini che la minorano.

Se $\Phi$ è convessa allora si può dimostrare che $\Phi(x) = \sup_{f \in A_\Phi} f(x)$, dove $A_\Phi$ è l'insieme delle funzioni lineari affini $f$ che minorano $\Phi$: $A_\Phi = \{f\,;\, f(x) = ax + b \text{ e } f(x) \le \Phi(x)\}$ (vedi figura). Allora,

$$\mathbb{E}(\Phi(X)\,|\,\mathscr{G}) \ge \mathbb{E}(f(X)\,|\,\mathscr{G}) = f(\mathbb{E}(X\,|\,\mathscr{G})),$$

da cui segue che

$$\mathbb{E}(\Phi(X)\,|\,\mathscr{G}) \ge \sup_{f \in A_\Phi} f(\mathbb{E}(X\,|\,\mathscr{G})) = \Phi(\mathbb{E}(X\,|\,\mathscr{G})).$$

**6.** *Se $Y$ è $\mathscr{G}$-misurabile allora $\mathbb{E}(XY\,|\,\mathscr{G}) = Y\,\mathbb{E}(X\,|\,\mathscr{G})$, q.c. In particolare, $\mathbb{E}(Y\,|\,\mathscr{G}) = Y$, q.c.*

Faremo la dimostrazione solo nel caso in cui $Y$ è prende un numero finito di valori, $y_1, \dots, y_n$: $Y = \sum_{i \le n} y_i\, 1_{\{Y = y_i\}}$. Se poniamo $G_i = \{Y = y_i\}$, allora $G_i \in \mathscr{G}$ perché $Y$ è $\mathscr{G}$-misurabile. Ora, preso $G \in \mathscr{G}$, si ha:

$$\begin{aligned}\mathbb{E}(XY1_G) &= \sum_{i \le n} y_i \mathbb{E}(X\, 1_{G_i \cap G}) = \sum_{i \le n} y_i \mathbb{E}(\mathbb{E}(X\,|\,\mathscr{G})\, 1_{G_i \cap G}) \\ &= \mathbb{E}\Big(\sum_{i \le n} y_i\, 1_{G_i}\, \mathbb{E}(X\,|\,\mathscr{G})\, 1_G\Big) = \mathbb{E}(Y\,\mathbb{E}(X\,|\,\mathscr{G})\, 1_G).\end{aligned}$$

**7.** *Siano $\mathscr{A}, \mathscr{G} \subset \mathscr{F}$ due $\sigma$-algebre tali che $\mathscr{A} \subset \mathscr{G}$. Allora: $\mathbb{E}(X\,|\,\mathscr{A}) = \mathbb{E}(\mathbb{E}(X\,|\,\mathscr{A})\,|\,\mathscr{G}) = \mathbb{E}(\mathbb{E}(X\,|\,\mathscr{G})\,|\,\mathscr{A})$, q.c.*

$\mathbb{E}(X\,|\,\mathscr{A})$ è $\mathscr{A}$-misurabile e quindi è $\mathscr{G}$-misurabile perché $\mathscr{A} \subset \mathscr{G}$. Dunque, per il punto precedente, $\mathbb{E}(X\,|\,\mathscr{A}) = \mathbb{E}(\mathbb{E}(X\,|\,\mathscr{A})\,|\,\mathscr{G})$. Dimostriamo ora che $\mathbb{E}(X\,|\,\mathscr{A}) = \mathbb{E}(\mathbb{E}(X\,|\,\mathscr{G})\,|\,\mathscr{A})$: preso $A \in \mathscr{A}$, allora $A$ appartiene anche a $\mathscr{G}$ e

$$\mathbb{E}(\mathbb{E}(X\,|\,\mathscr{G})\, 1_A) = \mathbb{E}(X\, 1_A).$$

**8.** *Sia $X$ indipendente da $G$. Allora,* $\mathbb{E}(X\,|\,\mathscr{G}) = \mathbb{E}(X)$*, q.c.*

Se $G \in \mathscr{G}$ allora $X$ e $Y = 1_G$ sono indipendenti, quindi

$$\mathbb{E}(X\,1_G) = \mathbb{E}(X)\mathbb{E}(1_G) = \mathbb{E}(\mathbb{E}(X)\,1_G).$$

Ora sappiamo che la speranza condizionale $\mathbb{E}(X\,|\,\mathscr{G})$ si calcola facilmente se $X$ è indipendente da $\mathscr{G}$ (vedi 8.) oppure se $X$ è già $\mathscr{G}$-misurabile. C'è un'altra situazione, molto frequente e che combina queste due situazioni, in cui il calcolo è abbastanza facile.

**9.** *Siano $X$ una v.a. $\mathscr{G}$-misurabile e $Y$ una v.a. indipendente da $\mathscr{G}$. Per ogni $f$ misurabile tale che $f(X,Y)$ ha speranza finita si ha* $\mathbb{E}(f(X,Y)\,|\,\mathscr{G}) = \bar{f}(X)$ *q.c., dove* $\bar{f}(x) = \mathbb{E}(f(x,Y))$.

Lo dimostreremo solo nel caso in cui la funzione $f$ è della forma $f(x,y) = f_1(x)f_2(y)$. Il caso generale si dimostra usando il fatto che ogni funzione $f(x,y)$ si può approssimare con delle combinazioni lineari di funzioni del tipo $f_1(x)f_2(y)$ e ne lasceremo la dimostrazione ad un corso più avanzato.

Se $f$ è della forma indicata, allora $\bar{f}(x) = f_1(x)\mathbb{E}(f_2(Y))$. Usando 6. e 8.,

$$\begin{aligned}\mathbb{E}(f(X,Y)\,|\,\mathscr{G}) = \mathbb{E}(f_1(X)\,f_2(Y)\,|\,\mathscr{G}) = f_1(X)\,\mathbb{E}(f_2(Y)\,|\,\mathscr{G}) =\\ = f_1(X)\,\mathbb{E}(f_2(Y)) = \bar{f}(X).\end{aligned}$$

**10.** *Si ha, per ogni v.a. $Z$ che sia $\mathscr{G}$-misurabile e di quadrato integrabile,*

$$\mathbb{E}[(X - \mathbb{E}(X\,|\,\mathscr{G}))^2] \le \mathbb{E}[(X - Z)^2] \tag{3.2}$$

*e, per di più la disuguaglianza precedente è stretta a meno che non sia $Z = \mathbb{E}(X\,|\,\mathscr{G})$ q.c. In altre parole $\mathbb{E}(X\,|\,\mathscr{G})$ è la v.a. $\mathscr{G}$-misurabile che meglio approssima $X$ nel senso dei minimi quadrati.*

È questa una delle proprietà più importanti della speranza condizionale. Si noti l'analogia con $\mathbb{E}(X)$: sia $\mathbb{E}(X)$ che $\mathbb{E}(X\,|\,\mathscr{G})$ sono le migliori stime di $X$ nel senso dei minimi quadrati, nel primo caso nella classe delle costanti mentre nel secondo caso nello spazio delle v.a. con momento secondo finito e $\mathscr{G}$-misurabili.

Dimostriamo la (3.2). Sia $X$ una v.a. con momento secondo finito. Intanto, osserviamo che grazie alla disuguaglianza di Jensen, sappiamo che anche $\mathbb{E}(X\,|\,\mathscr{G})$ ha momento secondo finito (si scelga $p = 2$ nel punto 5.). Ora, presa una qualunque v.a. $Z$ che sia $\mathscr{G}$-misurabile e di quadrato integrabile, si ha

$$\begin{aligned}\mathbb{E}[(X - Z)^2] = \ & \mathbb{E}[(X - \mathbb{E}(X\,|\,\mathscr{G}))^2] + \mathbb{E}[(Z - \mathbb{E}(X\,|\,\mathscr{G}))^2]\\ & -2\mathbb{E}[(X - \mathbb{E}(X\,|\,\mathscr{G}))(Z - \mathbb{E}(X\,|\,\mathscr{G}))].\end{aligned}$$

Ma $Z - \mathbb{E}(X \mid \mathscr{G})$ è una v.a. $\mathscr{G}$-misurabile e dunque

$$\begin{aligned}
&\mathbb{E}[(X - \mathbb{E}(X \mid \mathscr{G}))(Z - \mathbb{E}(X \mid \mathscr{G}))] \\
&= \mathbb{E}\big[\mathbb{E}[(X - \mathbb{E}(X \mid \mathscr{G}))(Z - \mathbb{E}(X \mid \mathscr{G}))] \,\big|\, \mathscr{G}\big] \\
&= (Z - \mathbb{E}(X \mid \mathscr{G}))\mathbb{E}\big[\mathbb{E}[X - \mathbb{E}(X \mid \mathscr{G}) \,\big|\, \mathscr{G}]\big] = 0.
\end{aligned}$$

Dunque,

$$\mathbb{E}[(X - Z)^2] \geq \mathbb{E}[(X - \mathbb{E}(X \mid \mathscr{G}))^2]$$

e si ha uguaglianza se e solo se $Z = \mathbb{E}(X \mid \mathscr{G})$, da cui la tesi.

### 3.1.4 Probabilità condizionale

**Definizione 3.1.5.** Sia $A \in \mathscr{F}$ e sia $\mathscr{G}$ una una sotto-$\sigma$-algebra di $\mathscr{F}$. Una versione della *probabilità condizionale* di $A$ dato $\mathscr{G}$ è data da

$$\mathbb{P}(A \mid \mathscr{G}) = \mathbb{E}(1_A \mid \mathscr{G}).$$

Così come nel caso della media condizionale, quando $\mathscr{G} = \sigma(Y)$ useremo la notazione $\mathbb{P}(A \mid Y)$.

Ricordiamo che, anche in questo caso, occorre parlare di "versione": analogamente alla media condizionale, la probabilità condizionale è definita a meno di insiemi di probabilità nulla, ovvero è definita q.c.

**Esempio 3.1.6.** Riprendendo gli esempi 3.1.1, 3.1.2 e 3.1.4, possiamo scrivere:

- se $(X, Y)$ è una coppia di v.a. discrete, di densità discreta congiunta $p_{X\,Y}$ e densità (discrete) marginali $p_X$ e $p_Y$ allora q.c.

  $$\mathbb{P}(X \in B \mid Y) = h_B(Y), \quad \text{dove } h_B(y) = \mathbb{P}(X \in B \mid Y = y) = \sum_{x \in B} \frac{p_{X\,Y}(x, y)}{p_Y(y)}$$

  per ogni $B \subset E_X$ (abbiamo scelto in particolare $A = \{X \in B\}$);

- se $\mathscr{G} = \sigma(\mathscr{C})$, con $\mathscr{C} = \{C_i\}_{i \in I}$ partizione al più numerabile di $\Omega$, e se $A \in \mathscr{F}$ allora q.c.

  $$\mathbb{P}(A \mid \mathscr{G}) = \sum_i \frac{\mathbb{P}(A \cap C_i)}{\mathbb{P}(C_i)} \, 1_{C_i};$$

- se $(X, Y)$ è una coppia di v.a. assolutamente continue, di densità continua congiunta $p_{X\,Y}$ e densità (continue) marginali $p_X$ e $p_Y$ allora q.c.

  $$\mathbb{P}(X \in B \mid Y) = \Psi_B(Y), \quad \text{dove } \Psi_B(y) = \int_B \frac{p_{X\,Y}(x, y)}{p_Y(y)} \, dy,$$

  per ogni $B \in \mathscr{E}$ (abbiamo scelto in particolare $A = \{X \in B\}$).

La probabilità condizionale verifica le due proprietà basilari della probabilità:

**(i)** per ogni $A \in \mathscr{F}$, $\mathbb{P}(A^c \,|\, \mathscr{G}) = 1 - \mathbb{P}(A \,|\, \mathscr{G})$, q.c.: immediata conseguenza del fatto che $1 = 1_A + 1_{A^c}$.

**(ii)** per ogni $A_1, A_2, \ldots \in \mathscr{F}$ tali che $A_n \cap A_m = \emptyset$ per $n \neq m$ allora $\mathbb{P}(\cup_n A_n \,|\, \mathscr{G}) = \sum_n \mathbb{P}(A_n \,|\, G)$, q.c.

Lo dimostriamo solo nel caso in cui $\mathscr{G}$ è generata da una partizione $\mathscr{C} = \{C_i\}_{i \in I}$ al più numerabile. In tal caso, ponendo $I_{\mathcal{N}} = \{i \in I : \mathbb{P}(C_i = 0)\}$, allora ovviamente $\mathbb{P}(\cup_{i \in I_{\mathcal{N}}} C_i) = 0$ ed è quindi immediato vedere che

$$\mathbb{E}(X \,|\, \mathscr{G}) = \sum_{i \notin I_{\mathcal{N}}} \frac{\mathbb{E}(X \, 1_{C_i})}{\mathbb{P}(C_i)} \, 1_{C_i}, \quad \text{q.c.}$$

Ma allora, q.c. si hanno le seguenti uguaglianze:

$$\begin{aligned} \mathbb{P}(\cup_n A_n \,|\, \mathscr{G})(\omega) &= \sum_{i \notin I_{\mathcal{N}}} \frac{\mathbb{P}((\cup_n A_n) \cap C_i)}{\mathbb{P}(C_i)} \, 1_{C_i} = \sum_{i \notin I_{\mathcal{N}}} \sum_n \frac{\mathbb{P}(A_n \cap C_i)}{\mathbb{P}(C_i)} \, 1_{C_i} \\ &= \sum_n \sum_{i \notin I_{\mathcal{N}}} \frac{\mathbb{P}(A_n \cap C_i)}{\mathbb{P}(C_i)} \, 1_{C_i} = \sum_n \mathbb{P}(A_n \,|\, \mathscr{G}), \end{aligned}$$

da cui la tesi.

Si potrebbe allora pensare di poter costruire una v.a.

$$\mathbb{P}(\cdot, \cdot) \, : \, \mathscr{F} \times \Omega \longrightarrow \mathbb{R}$$

tale che

1. fissato $A \in \mathscr{F}$, $\omega \mapsto \mathbb{P}(A, \omega)$ è una versione di $\mathbb{P}(A \,|\, \mathscr{G})(\omega)$;
2. fissato $\omega \in \Omega$, $A \mapsto \mathbb{P}(A, \omega)$ è una probabilità su $(\Omega, \mathscr{F})$.

Qualora sia possibile, $\mathbb{P}(A, \omega)$ prende il nome di *versione regolare della probabilità condizionata*. In generale però, tale funzione $\mathbb{P}(A, \omega)$ non si può costruire. Infatti, abbiamo visto che, fissato $A \in \mathscr{F}$, $\mathbb{P}(A \,|\, \mathscr{G})$ è definita q.c.: esiste $\mathcal{N}_A \in \mathscr{F}$ tale che $\mathbb{P}(\mathcal{N}_A) = 0$ e $\mathbb{P}(A \,|\, \mathscr{F})(\omega)$ è ben definita per $\omega \notin \mathcal{N}_A$. Ne segue che, in generale,

$$\omega \mapsto \mathbb{P}(A \,|\, \mathscr{F})(\omega) \text{ è ben definita per } \omega \notin \mathcal{N} = \cup_{A \in \mathscr{F}} \mathcal{N}_A.$$

Ora, perché il tutto abbia senso occorre che $\mathcal{N}$ sia "trascurabile", cioè $\mathbb{P}(\mathcal{N}) = 0$. Ma, poiché $\mathscr{F}$ non è in generale numerabile e dunque $\mathcal{N} = \cup_{A \in \mathscr{F}} \mathcal{N}_A$ non è in generale un'unione numerabile di eventi,

- non è detto che $\mathcal{N}$ sia un evento, ovvero $\mathcal{N} \in \mathscr{F}$;

- se anche $\mathcal{N} \in \mathscr{F}$, non è detto che $\mathbb{P}(\mathcal{N}) = 0$.

Osserviamo infine che, poiché la probabilità condizionale di $A$ è data dalla media condizionale della v.a. $X = 1_A$, molte delle proprietà viste per le medie condizionali valgono anche in questo caso. Ad esempio, se $A \in \mathscr{G}$ allora ovviamente $\mathbb{P}(A \mid \mathscr{G}) = 1_A$ q.c., così come se $A$ è indipendente da $\mathscr{G}$ allora $\mathbb{P}(A \mid \mathscr{G}) = \mathbb{P}(A)$, q.c.

## 3.2 Martingale

### 3.2.1 Definizioni e generalità

Sia $(\Omega, \mathscr{F}, \mathbb{P})$ uno spazio di probabilità. Una *filtrazione* è una successione $(\mathscr{F}_n)_n$ *crescente* di sotto-$\sigma$-algebre di $\mathscr{F}$, cioè tale che $\mathscr{F} \supset \mathscr{F}_{n+1} \supset \mathscr{F}_n$. Presa una successione $(X_n)_n$, diremo che è *adattata* alla filtrazione $(\mathscr{F}_n)_n$, o anche che è $\mathscr{F}_n$-adattata, se $X_n$ è $\mathscr{F}_n$-misurabile, per ogni $n$. In questo paragrafo ci collocheremo sempre in questo contesto.

**Definizione 3.2.1.** Una successione $(X_n)_n$ di v.a. è una martingala (risp. una supermartingala, una submartingala ) rispetto alla filtrazione $(\mathscr{F}_n)_n$ se $(X_n)_n$ è $\mathscr{F}_n$-adattata, la v.a. $X_n$ ha speranza matematica finita per ogni $n$ e se, per ogni $n$,

$$\mathbb{E}(X_{n+1} \mid \mathscr{F}_n) = X_n \quad \text{q.c. (risp. } \leq, \geq).$$

**Esercizio 3.1.** Verificare che $(X_n)_n$ è una martingala (risp. una supermartingala, una submartingala) se e solo se, per ogni $A \in \mathscr{F}_n$,

$$\mathbb{E}(1_A X_{n+1}) = \mathbb{E}(1_A X_n) \quad \text{(risp. } \leq, \geq). \tag{3.3}$$

Si noti che $(X_n)_n$ è una supermartingala se e solo se $(-X_n)_n$ è una submartingala e $(X_n)_n$ è una martingala se e solo se $(X_n)_n$ è simultaneamente una supermartingala e una submartingala.

**Esempio 3.2.2.** Sia $(Y_n)_n$ una successione di v.a. indipendenti e sia $\mathscr{F}_n = \sigma(Y_1, \dots, Y_n)$. Supponiamo inoltre che le $Y_n$ abbiano tutte speranza matematica finita ed $= 0$ (risp. $\leq 0$, $\geq 0$). Allora: $X_n = Y_0 + Y_1 + \cdots + Y_n$ è una $\mathscr{F}_n$-martingala (risp. una supermartingala, una submartingala).

Infatti

$$\begin{aligned}\mathbb{E}(X_{n+1} \mid \mathscr{F}_n) = {} & \mathbb{E}(X_n + Y_n \mid \mathscr{F}_n) = \mathbb{E}(X_n \mid \mathscr{F}_n) + \mathbb{E}(Y_n \mid \mathscr{F}_n) \\ = {} & X_n + \mathbb{E}(Y_n) = (\text{risp. } \leq 0, \geq 0)\, X_n\end{aligned}$$

poiché $X_n$ è $\mathscr{F}_n$-misurabile, mentre $Y_n$ è indipendente da $\mathscr{F}_n$.

**Esercizio 3.2.** Sia $X$ una v.a. avente speranza matematica finita e $\{\mathscr{F}_n\}_n$ una filtrazione. Mostrare che $X_n = \mathbb{E}(X \mid \mathscr{F}_n)$ è una martingala.

**Esercizio 3.3.** Siano $Y_0, Y_1, \dots$ v.a. i.i.d., tali che $\mathbb{P}(Y_n = 1) = p = 1 - q = \mathbb{P}(Y_n = -1)$. Siano $\mathscr{F}_n = \sigma(Y_0, \dots, Y_n)$ e $X_n = Y_0 + \cdots + Y_n$. Poniamo

$$Z_n = \left(\frac{q}{p}\right)^{X_n} \quad \text{e} \quad W_n = \left(\frac{p}{q}\right)^{X_n}.$$

Dire se $(Z_n)_n$ e/o $(W_n)_n$ sono delle $\mathscr{F}_n$ martingale o submartingale o supermartingale.

**Esercizio 3.4.** Siano $Y_0, Y_2, \dots$ v.a. i.i.d., con media $\mu$. Poniamo $\mathscr{F}_n = \sigma(Y_0, \dots, Y_n)$ e $X_n = Y_0 \cdots Y_n$. Dimostrare che $(X_n)_n$ è una $\mathscr{F}_n$-martingala se e solo se $\mu = 1$. Supponiamo quindi $\mu = 1$ e che le $(Y_n)_n$ abbiano varianza $\sigma^2$: verificare che $Z_n = X_n^2 - \sigma^2 \sum_{k=0}^{n-1} X_k^2$ è una $\mathscr{F}_n$-martingala.

### 3.2.2 Prime proprietà

- *Se $(X_n)_n$ è una martingala (risp. una supermartingala, una submartingala), la successione $(\mathbb{E}(X_n))_n$ è costante (risp. decrescente, crescente).*

Basta infatti applicare la (3.3) con $A = \Omega$.

- *Se $(X_n)_n$ è una martingala (risp. una supermartingala, una submartingala), per $m < n$ si ha $\mathbb{E}(X_n \,|\, \mathscr{F}_m) = X_m$ q.c. (risp. $\leq, \geq$).*

Infatti, per ricorrenza,

$$\begin{aligned} \mathbb{E}[X_n \,|\, \mathscr{F}_m] = \;& \mathbb{E}[\mathbb{E}(X_n \,|\, \mathscr{F}_{n-1}) \,|\, \mathscr{F}_m] \\ = \;& \mathbb{E}[X_{n-1} \,|\, \mathscr{F}_m] = \cdots = \mathbb{E}[X_{m+1} \,|\, \mathscr{F}_m] = X_m. \end{aligned}$$

- *Se $(X_n)_n$ e $(Y_n)_n$ sono supermartingale, lo stesso vale per $X_n + Y_n$. Anche $X_n \wedge Y_n$ è una supermartingala.*

Infatti

$$\begin{aligned} \mathbb{E}(X_{n+1} \wedge Y_{n+1} \,|\, \mathscr{F}_n) \leq \mathbb{E}(X_{n+1} \,|\, \mathscr{F}_n) \leq X_n \\ \mathbb{E}(X_{n+1} \wedge Y_{n+1} \,|\, \mathscr{F}_n) \leq \mathbb{E}(Y_{n+1} \,|\, \mathscr{F}_n) \leq Y_n \end{aligned}$$

e dunque $\mathbb{E}(X_{n+1} \wedge Y_{n+1} \,|\, \mathscr{F}_n) \leq X_n \wedge Y_n$.

- *Se $(X_n)_n$ è una martingala (risp. una submartingala) e $f : \mathbb{R} \to \mathbb{R}$ è una funzione convessa (risp. convessa e crescente) tale che $f(X_n)$ è integrabile per ogni $n \geq 0$, allora $Y_n = f(X_n)$ è una submartingala.*

La dimostrazione segue direttamente applicando la disuguaglianza di Jensen (per la speranza condizionale).

### 3.2.3 La decomposizione di Doob

Si dice che $(A_n)_n$ è un *processo prevedibile crescente* se $A_0 = 0$, $A_n \le A_{n+1}$ per ogni $n \ge 0$ e $A_{n+1}$ è $\mathscr{F}_n$-misurabile.

Sia $(X_n)_n$ una $\mathscr{F}_n$-submartingala. Definiamo

$$A_0 = 0, \quad A_{n+1} = A_n + \mathbb{E}(X_{n+1} - X_n \,|\, \mathscr{F}_n) = \sum_{k=0}^{n} \mathbb{E}(X_{k+1} - X_k \,|\, \mathscr{F}_k).$$

Per costruzione, $(A_n)_n$ è un processo prevedibile crescente e $M_n = X_n - A_n$ verifica $\mathbb{E}(M_{n+1} - M_n \,|\, \mathscr{F}_n) = 0$ (perché $A_{n+1}$ è $\mathscr{F}_n$-misurabile!). Quindi, $(M_n)_n$ è una $\mathscr{F}_n$-martingala ed inoltre

$$X_n = M_n - A_n.$$

Questa decomposizione è, per di più, unica. Infatti, se $X_n = M'_n + A'_n$ è un'altra decomposizione di questo tipo, allora $A'_0 = 0$ e

$$A'_{n+1} - A'_n = X_{n+1} - X_n - (M'_{n+1} - M'_n)$$

da cui, condizionando rispetto a $\mathscr{F}_n$, $A'_{n+1} - A'_n = \mathbb{E}(X_{n+1} - X_n \,|\, \mathscr{F}_n)$. Dunque, $A'_n = A_n$ e $M'_n = M_n$. Abbiamo quindi dimostrato che

**Teorema 3.2.3.** *(Decomposizione di Doob) Ogni submartingala $(X_n)_n$ si può scrivere, in maniera unica, nella forma $X_n = M_n + A_n$, dove $(M_n)_n$ è una martingala e $(A_n)_n$ è un processo prevedibile crescente integrabile.*

Il processo prevedibile crescente $(A_n)_n$ del Teorema (3.2.3) si chiama il *compensatore* della submartingala $(X_n)_n$.

Presa una supermartingala (risp. martingala) $(X_n)_n$, allora $(-X_n)_n$ è una submartingala (risp. martingala). Applicando a $(-X_n)_n$ il Teorema 3.2.3, otteniamo la seguente decomposizione di Doob per supermartingale:

**Teorema 3.2.4.** *Ogni supermartingala $(X_n)_n$ si può scrivere, in maniera unica, nella forma $X_n = M_n - A_n$, dove $(M_n)_n$ è una martingala e $(A_n)_n$ è un processo prevedibile crescente integrabile.*

**Esercizio 3.5.** Sia $(X_n)_n$ la martingala dell'Esempio 3.2.2: $X_n = Y_0 + \cdots Y_n$, con $Y_1, \ldots$ i.i.d., di media 0, e $\mathscr{F}_n = \sigma(Y_0, \ldots, Y_n)$. Supponiamo che le $Y_i$ abbiano varianza $\sigma^2$. Verificare che $X_n^2$ è una submartingala e determinarne il compensatore.

### 3.2.4 Martingale trasformate

**Proposizione 3.2.5.** *Sia $(M_n)_n$ una $\mathscr{F}_n$-martingala e $(H_n)_n$ un processo prevedibile, cioè tale che $H_{n+1}$ sia $\mathscr{F}_n$ misurabile. Sia*

$$X_0 = 0 \quad \textit{e per } n > 0, \quad X_n = \sum_{j=1}^{n} H_j \,(M_j - M_{j-1})$$

*Allora $(X_n)_n$ è ancora una martingala rispetto alla stessa filtrazione $\mathscr{F}_n$.*

La martingala $(X_n)_n$ definita nella Proposizione 3.2.5 è detta la *martingala trasformata di* $(M_n)_n$ *tramite* $(H_n)_n$.

**Dimostrazione della Proposizione 3.2.5.** $X_n$ è chiaramente $\mathscr{F}_n$-misurabile e poiché $H_{n+1}$ è $\mathscr{F}_n$-misurabile, si ha

$$\mathbb{E}(H_{n+1}\,(M_{n+1} - M_n) \,|\, \mathscr{F}_n) = H_{n+1}\,\mathbb{E}((M_{n+1} - M_n) \,|\, \mathscr{F}_n) = 0.$$

Essendo $X_{n+1} = H_{n+1}\,(M_{n+1} - M_n) + X_n$, si ha

$$\begin{aligned}\mathbb{E}(X_{n+1} \,|\, \mathscr{F}_n) &= \mathbb{E}(H_{n+1}\,(M_{n+1} - M_n) \,|\, \mathscr{F}_n) + \mathbb{E}(X_n \,|\, \mathscr{F}_n) \\ &= H_{n+1}\,\mathbb{E}(M_{n+1} - M_n \,|\, \mathscr{F}_n) + X_n = X_n\end{aligned}$$

□

Presentiamo infine un'utile caratterizzazione delle martingale, che fa uso delle martingale trasformate.

**Proposizione 3.2.6.** *Sia* $(M_n)_n$ *un processo* $\mathscr{F}_n$*-adattato.* $(M_n)_{0\leq n\leq N}$ *è una martingala se e solo se per ogni processo prevedibile* $(H_n)_{0\leq n\leq N}$ *si ha*

$$\mathbb{E}\Big(\sum_{n=1}^{N} H_n(M_n - M_{n-1})\Big) = 0.$$

**Dimostrazione.** Se $(M_n)_{0\leq n\leq N}$ è una martingala, dalla Proposizione 3.2.5 anche $(X_n)_{0\leq n\leq N}$ è una martingala, con $X_n = \sum_{j=1}^{n} H_j\,(M_j - M_{j-1})$ per $n \geq 1$ e $X_0 = 0$. Quindi,

$$\mathbb{E}\Big(\sum_{n=1}^{N} H_n(M_n - M_{n-1})\Big) = \mathbb{E}(X_N) = \mathbb{E}(X_0) = 0.$$

Viceversa, per dimostrare che $(M_n)_{0\leq n\leq N}$ è una martingala basta far vedere che $\mathbb{E}(M_{n+1}\,1_A) = \mathbb{E}(M_n\,1_A)$, per ogni fissato $n = 0, 1, \ldots, N-1$ e $A \in \mathscr{F}_n$.

Fissiamo quindi $n = 0, 1, \ldots, N-1$ e $A \in \mathscr{F}_n$. Sia

$$H_j = \begin{cases} 1_A & \text{se } j = n+1 \\ 0 & \text{altrimenti.} \end{cases}$$

$H$ è ovviamente prevedibile e

$$0 = \mathbb{E}\Big(\sum_{j=1}^{N} H_j\,(M_j - M_{j-1})\Big) = \mathbb{E}\Big(1_A\,(M_{n+1} - M_n)\Big)$$

cioè $\mathbb{E}(M_{n+1}\,1_A) = \mathbb{E}(M_n\,1_A)$.

□

## 3.3 Tempi d'arresto

**Definizione 3.3.1.** **(i)** Un *tempo d'arresto rispetto alla filtrazione* $(\mathscr{F}_n)_n$, anche detto un $\mathscr{F}_n$-tempo d'arresto, è una v.a. $\tau : \Omega \to \bar{\mathbb{N}}$ (quindi può anche prendere il valore $+\infty$) tale che, per ogni $n \geq 0$,

$$\{\tau \leq n\} \in \mathscr{F}_n.$$

**(ii)** Se $\tau$ è un $\mathscr{F}_n$-tempo d'arresto, si chiama *$\sigma$-algebra degli eventi antecedenti al tempo* $\tau$, in simboli $\mathscr{F}_\tau$, la $\sigma$-algebra

$$\mathscr{F}_\tau = \{A \in \mathscr{F}_\infty;\ \text{per ogni } n \geq 0,\ A \cap \{\tau \leq n\} \in \mathscr{F}_n\},$$

dove $\mathscr{F}_\infty$ denota la $\sigma$-algebra generata da $\bigcup_n \mathscr{F}_n$.

**Esercizio 3.6.** Verificare che $\mathscr{F}_\tau$ è effettivamente una $\sigma$-algebra.

È utile osservare che nelle (i) e (ii) della Definizione 3.3.1, si possono sostituire le condizioni $\{\tau \leq n\} \in \mathscr{F}_n$ e $A \cap \{\tau \leq n\} \in \mathscr{F}_n$ con $\{\tau = n\} \in \mathscr{F}_n$ e $A \cap \{\tau = n\} \in \mathscr{F}_n$, rispettivamente, perché

$$\{\tau \leq n\} = \bigcup_{k=0}^{n} \{\tau = k\}, \quad \text{quindi } \{\tau = n\} = \{\tau \leq n\} \cap \{\tau \leq n-1\}^c.$$

**Esempio 3.3.2.** (*Esempio fondamentale*) Sia $(X_n)_n$ un processo a valori in $(E, \mathscr{E})$ e adattato alla filtrazione $(\mathscr{F}_n)_n$. Poniamo, per $A \in \mathscr{E}$,

$$\tau_A(\omega) = \inf\{n \geq 0; X_n(\omega) \in A\}, \tag{3.4}$$

con l'intesa che $\inf \emptyset = +\infty$. Allora $\tau_A$ è un tempo d'arresto, poichè

$$\{\tau_A = n\} = \{X_0 \notin A, X_1 \notin A, \ldots, X_{n-1} \notin A, X_n \in A\} \in \mathscr{F}_n;$$

$\tau_A$ si chiama *tempo di passaggio* o *d'ingresso* in $A$.

Siano $\tau_1, \tau_2$ due tempi d'arresto della filtrazione $(\mathscr{F}_n)_n$. Vediamo ora alcune proprietà interessanti, di cui ci serviremo nel seguito. La loro verifica può apparire complicata a prima vista (soprattutto le $\sigma$-algebre $\mathscr{F}_\tau$ fanno un po' di paura) ma si rivela poi abbastanza semplice e scolastica. Lasciamo per esercizio la dimostrazione.

**Esercizio 3.7.** Siano $\tau_1$ e $\tau_2$ due $\mathscr{F}_n$-tempi d'arresto. Dimostrare che $\tau_1+\tau_2$, $\tau_1 \vee \tau_2$, $\tau_1 \wedge \tau_2$ sono tempi d'arresto, rispetto alla stessa filtrazione.

**Esercizio 3.8.** Siano $\tau_1$ e $\tau_2$ due $\mathscr{F}_n$-tempi d'arresto. Dimostrare che se $\tau_1 \leq \tau_2$, allora $\mathscr{F}_{\tau_1} \subset \mathscr{F}_{\tau_2}$.

**Esercizio 3.9.** Siano $\tau_1$ e $\tau_2$ due $\mathscr{F}_n$-tempi d'arresto. Mostrare che $\mathscr{F}_{\tau_1 \wedge \tau_2} = \mathscr{F}_{\tau_1} \cap \mathscr{F}_{\tau_2}$.

**Esercizio 3.10.** Siano $\tau_1$ e $\tau_2$ due $\mathscr{F}_n$-tempi d'arresto. Dimostrare che $\{\tau_1 < \tau_2\}$ e $\{\tau_1 = \tau_2\}$ appartengono entrambi a $\mathscr{F}_{\tau_1} \cap \mathscr{F}_{\tau_2}$

Sia $(X_n)_n$ un processo adattato; spesso avremo bisogno di considerare la posizione del processo al tempo $\tau$, cioè $X_\tau(\omega) = X_{\tau(\omega)}(\omega)$. Possiamo quindi dire che

$$X_\tau = X_n \text{ sull'evento } \{\tau = n\}, \quad n \in \bar{\mathbb{N}}.$$

Osserviamo che la v.a. $X_\tau$ è $\mathscr{F}_\tau$-misurabile perché

$$\{X_\tau \in A\} \cap \{\tau = n\} = \{X_n \in A\} \cap \{\tau = n\} \in \mathscr{F}_n.$$

Il risultato che segue è molto utile per calcolare la speranza condizionale rispetto alla $\sigma$-algebra $\mathscr{F}_\tau$.

**Proposizione 3.3.3.** *Sia $X$ una v.a. integrabile e sia $\tau$ un tempo d'arresto. Si ponga, per ogni $n \in \bar{\mathbb{N}}$, $X_n = \mathbb{E}(X \mid \mathscr{F}_n)$. Allora sull'evento $\{\tau = n\}$, si ha $\mathbb{E}(X \mid \mathscr{F}_\tau) = \mathbb{E}(X \mid \mathscr{F}_n)$, cioè*

$$\mathbb{E}(X \mid \mathscr{F}_\tau) = X_\tau \tag{3.5}$$

**Dimostrazione.** Se $A \in \mathscr{F}_\tau$, allora $1_A\, 1_{\{\tau=n\}} = 1_{A\cap\{\tau=n\}}$ è $\mathscr{F}_n$-misurabile. Quindi,

$$\mathbb{E}(1_A\, X_\tau) = \sum_{n\in\bar{\mathbb{N}}} \mathbb{E}[1_A\, X_n\, 1_{\{\tau=n\}}] = \sum_{n\in\bar{\mathbb{N}}} \mathbb{E}[1_A\, X\, 1_{\{\tau=n\}}] = \mathbb{E}(1_A\, X),$$

e dunque $\mathbb{E}(X \mid \mathscr{F}_\tau) = X_\tau$.

□

Concludiamo con il prossimo risultato, molto semplice da dimostrare ma, come vedremo in seguito, particolarmente utile.

**Proposizione 3.3.4.** *Se $(X_n)_n$ è una martingala (risp. una supermartingala, una submartingala), lo stesso è vero anche per il processo arrestato*

$$X_n^\tau = X_{n\wedge\tau},$$

*dove $\tau$ è una tempo d'arresto della filtrazione $(\mathscr{F}_n)_n$.*

**Dimostrazione.** Per la definizione di tempo d'arresto, $\{\tau \geq n+1\} = \{\tau \leq n\}^c \in \mathscr{F}_n$; dunque

$$\begin{aligned} \mathbb{E}(X_{n+1}^\tau - X_n^\tau \mid \mathscr{F}_n) = &\ \mathbb{E}((X_{n+1} - X_n)1_{\{\tau\geq n+1\}} \mid \mathscr{F}_n) \\ = &\ 1_{\{\tau\geq n+1\}}\mathbb{E}(X_{n+1} - X_n \mid \mathscr{F}_n) = (\text{risp. } \leq, \geq)\, 0. \end{aligned} \tag{3.6}$$

□

### 3.3.1 Il teorema d'arresto

Sia $(X_n)_n$ una supermartingala. Quindi, per $m < n$, si ha $\mathbb{E}(X_n \,|\, \mathscr{F}_m) \leq X_m$ q.c. Questa proprietà resta vera se si sostituiscono $m$ e $n$ con dei tempi d'arresto? Cioè: se $\tau_1$ e $\tau_2$ sono due tempi d'arresto con $\tau_1 \leq \tau_2$, è vero che

$$\mathbb{E}(X_{\tau_2} \,|\, \mathscr{F}_{\tau_1}) \leq X_{\tau_1}? \tag{3.7}$$

Supponiamo che $\tau_1$ e $\tau_2$ siano limitati: esiste $k \in \mathbb{N}$ tale che $\tau_1 \leq \tau_2 \leq k$. Poiché $A \cap \{\tau_1 = j\} \in \mathscr{F}_j$ e $(X_n^{\tau_2})_n = (X_{\tau_2 \wedge n})_n$ è una supermartingala,

$$\mathbb{E}(X_{\tau_2} 1_A) = \sum_{j=0}^{k} \mathbb{E}(X_{\tau_2 \wedge k} 1_{A \cap \{\tau_1 = j\}}) \leq \sum_{j=0}^{k} \mathbb{E}(X_{\tau_2 \wedge j} 1_{A \cap \{\tau_1 = j\}}).$$

Ma $X_{\tau_2 \wedge j} = j$ su $\{\tau_1 = j\}$, dunque

$$\mathbb{E}(X_{\tau_2} 1_A) \leq \sum_{j=0}^{k} \mathbb{E}(X_j 1_{A \cap \{\tau_1 = j\}}) = \mathbb{E}(X_{\tau_1} 1_A).$$

Quindi abbiamo dimostrato che

**Teorema 3.3.5.** *(Teorema d'arresto) Siano $(X_n)_n$ una martingala (risp. supermartingala, submartingala) e $\tau_1, \tau_2$ due tempi d'arresto limitati tali che $\tau_1 \leq \tau_2$. Allora $\mathbb{E}(X_{\tau_2} \,|\, \mathscr{F}_{\tau_1}) = X_{\tau_1}$ (risp. $\leq, \geq$).*

Prendendo le speranze matematiche,

**Corollario 3.3.6.** *Siano $(X_n)_n$ una martingala (risp. supermartingala, submartingala) e $\tau$ un tempo d'arresto limitato. Allora $\mathbb{E}(X_\tau) = \mathbb{E}(X_0)$ (risp. $\leq, \geq$).*

## 3.4 Soluzioni

**Soluzione dell'esercizio 3.1.** La dimostrazione è immediata conseguenza della definizione di media condizionata: posto $Y_n = \mathbb{E}(X_{n+1} \,|\, \mathscr{F}_n)$ allora dev'essere

$$\mathbb{E}(1_A Y_n) = \mathbb{E}(1_A X_{n+1}) \quad \text{per ogni } A \in \mathscr{F}_n.$$

Dunque, $Y_n = X_n$ (risp. $\leq, \geq$) se e solo se per ogni $A \in \mathscr{F}_n$ si ha

$$\mathbb{E}(1_A X_{n+1}) = \mathbb{E}(1_A X_n) \quad (\text{risp. } \leq, \geq).$$

**Soluzione dell'esercizio 3.2.** Usando il fatto che $\mathscr{F}_n \subset \mathscr{F}_{n+1}$ e la proprietà 7 del Paragrafo 3.1.3, si ha

$$\mathbb{E}(X_{n+1} \,|\, \mathscr{F}_n) = \mathbb{E}\Big(\mathbb{E}(X \,|\, \mathscr{F}_{n+1}) \,\Big|\, \mathscr{F}_n\Big) = \mathbb{E}(X \,|\, \mathscr{F}_n) = X_n,$$

dunque $(X_n)_n$ è una martingala.

**Soluzione dell'esercizio 3.3.** Intanto, è immediato verificare che sia $Z_n$ che $W_n$ sono integrabili. Studiamo dapprima $(Z_n)_n$:

$$\mathbb{E}(Z_{n+1} \,|\, \mathscr{F}_n) = \mathbb{E}\Big(\Big(\frac{q}{p}\Big)^{X_n + Y_{n+1}} \,|\, \mathscr{F}_n\Big) = \Big(\frac{q}{p}\Big)^{X_n} \mathbb{E}\Big(\Big(\frac{q}{p}\Big)^{Y_{n+1}} \,|\, \mathscr{F}_n\Big)$$
$$= Z_n\Big(\frac{q}{p} \cdot \mathbb{P}(Y_{n+1} = 1) + \frac{p}{q} \cdot \mathbb{P}(Y_{n+1} = -1)\Big) = Z_n\Big(\frac{q}{p} \cdot p + \frac{p}{q} \cdot q\Big) = Z_n,$$

dunque $(Z_n)_n$ è una martingala. Ora, essendo $W_n = 1/Z_n$, con $f(z) = 1/z$ convessa, $(W_n)_n$ è una submartingala. Verifichiamolo anche direttamente:

$$\mathbb{E}(W_{n+1} \,|\, \mathscr{F}_n) = \mathbb{E}\Big(\Big(\frac{p}{q}\Big)^{X_n + Y_{n+1}} \,|\, \mathscr{F}_n\Big) = \Big(\frac{p}{q}\Big)^{X_n} \mathbb{E}\Big(\Big(\frac{p}{q}\Big)^{Y_{n+1}} \,|\, \mathscr{F}_n\Big)$$
$$= W_n\Big(\frac{p}{q} \cdot \mathbb{P}(Y_{n+1} = 1) + \frac{q}{p} \cdot \mathbb{P}(Y_{n+1} = -1)\Big) = W_n\Big(\frac{p}{q} \cdot p + \frac{q}{p} \cdot q\Big) = c\, W_n,$$

dove $c = (p^3 + q^3)/(pq) = (p^2 - pq + q^2)/(pq)$. Osserviamo che $c \geq 1$ per ogni $p, q$, con $c = 1$ se e solo se $p = q = 1/2$. Dunque, $(W_n)_n$ è una submartingala (e in particolare una-banale-martingala se $p = q$).

**Soluzione dell'esercizio 3.4.** Poiché $X_{n+1} = X_n \cdot Y_{n+1}$, si ha:

$$\mathbb{E}(X_{n+1} \,|\, \mathscr{F}_n) = \mathbb{E}(X_n\, Y_{n+1} \,|\, \mathscr{F}_n) = X_n\, \mathbb{E}(Y_{n+1} \,|\, \mathscr{F}_n) = X_n\, \mathbb{E}(Y_{n+1}) = \mu\, X_n,$$

dunque $(X_n)_n$ è una martingala se e solo se $\mu = 1$. Supponiamo ora che $\mathbb{E}(Y_n) = \mu$ e $\mathrm{Var}(Y_n) = \sigma^2$, per ogni $n$. Allora, usando che $X_{n+1}^2 = X_n^2 \cdot Y_{n+1}^2$:

$$\mathbb{E}(Z_{n+1} \,|\, \mathscr{F}_n) = \mathbb{E}\Big(X_{n+1}^2 - \sigma^2 \sum_{k=0}^{n} X_k^2 \,\Big|\, \mathscr{F}_n\Big)$$
$$= X_n^2\, \mathbb{E}(Y_{n+1}^2 \,|\, \mathscr{F}_n) - \sigma^2 \sum_{k=0}^{n} X_k^2 = X_n^2\, \mathbb{E}(Y_{n+1}^2) - \sigma^2 \sum_{k=0}^{n} X_k^2$$
$$= X_n^2(\sigma^2 + 1) - \sigma^2 \sum_{k=0}^{n} X_k^2 = X_n^2 - \sigma^2 \sum_{k=0}^{n-1} X_k^2 = Z_n$$

**Soluzione dell'esercizio 3.5.** $X_n^2$ è una submartingala perché $X_n^2 = f(X_n)$, con $f(x) = x^2$ convessa e $f(X_n)$ è integrabile. Il compensatore è $A_n = \sum_{k=0}^{n-1} \mathbb{E}(X_{k+1}^2 - X_k^2 \mid \mathscr{F}_k)$. Ora,

$$\begin{aligned} &\mathbb{E}(X_{k+1}^2 - X_k^2 \mid \mathscr{F}_k) = \mathbb{E}((X_k + Y_{k+1})^2 - X_k^2 \mid \mathscr{F}_k) \\ &= \mathbb{E}(Y_{k+1}^2 + 2X_k Y_{k+1} \mid \mathscr{F}_k) = \mathbb{E}(Y_{k+1}^2) + 2X_k \mathbb{E}(Y_{k+1}) = \sigma^2, \end{aligned}$$

dunque $A_n = n\,\sigma^2$.

**Soluzione dell'esercizio 3.6.** 1. $\Omega \in \mathscr{F}_\tau$: infatti, $\Omega \in \mathscr{F}_\infty$ ($\mathscr{F}_\infty$ è una $\sigma$-algebra) e, per ogni $n$, $\Omega \cap \{\tau \le n\} = \{\tau \le n\} \in \mathscr{F}_n$ perché $\tau$ è $\mathscr{F}_n$-misurabile.

2. Se $A \in \mathscr{F}_\tau$ allora $A^c \in \mathscr{F}_\tau$: se $A \in \mathscr{F}_\tau$ allora $A^c \in \mathscr{F}_\infty$ e $A^c \cap \{\tau \le n\} = \{\tau \le n\} \cap (A \cap \{\tau \le n\})^c \in \mathscr{F}_n$ per ogni $n$ (perchè sia $\{\tau \le n\}$ che $A \cap \{\tau \le n\}$ stanno in $\mathscr{F}_n$).

3. Se $\{A_k\} \subset \mathscr{F}_\tau$ allora $\cup_k A_k \in \mathscr{F}_\tau$: in tal caso, ovviamente $\cup_k A_k \in \mathscr{F}_\infty$ ed inoltre $(\cup_k A_k) \cap \{\tau \le n\} = \cup_k (A_k \cap \{\tau \le n\}) \in \mathscr{F}_n$

**Soluzione dell'esercizio 3.7.** Per la prima, abbiamo visto che basta verificare che, per ogni $n$, $\{\tau_1 + \tau_2 = n\} \in \mathscr{F}_n$. Ma si può scrivere

$$\{\tau_1 + \tau_2 = n\} = \bigcup_{k=0}^{n} \{\tau_1 = k, \tau_2 \le n - k\}.$$

Ma, se $0 \le k \le n$, gli eventi $\{\tau_1 = k\}$ e $\{\tau_2 \le n - k\}$ appartengono a $\mathscr{F}_n$ (ricordiamo che le $\sigma$-algebre $\mathscr{F}_n$ sono crescenti). Dunque, nell'unione a destra tutti gli eventi appartengono a $\mathscr{F}_n$.

Per la seconda, osserviamo che

$$\{\tau_1 \vee \tau_2 \le n\} = \underbrace{\{\tau_1 \le n\}}_{\in \mathscr{F}_n} \cap \underbrace{\{\tau_2 \le n\}}_{\in \mathscr{F}_n} \in \mathscr{F}_n.$$

Infine,

$$\{\tau_1 \wedge \tau_2 > n\} = \underbrace{\{\tau_1 > n\}}_{\in \mathscr{F}_n} \cap \underbrace{\{\tau_2 > n\}}_{\in \mathscr{F}_n} \in \mathscr{F}_n,$$

dunque $\{\tau_1 \wedge \tau_2 \le n\} = \{\tau_1 \wedge \tau_2 > n\}^c \in \mathscr{F}_n$.

**Soluzione dell'esercizio 3.8.** Intuitivamente questa relazione è evidente: se il tempo $\tau_1$ precede $\tau_2$, allora gli eventi antecedenti a $\tau_1$ sono anche antecedenti $\tau_2$. Verifichiamo formalmente la proprietà. Sia $A \in \mathscr{F}_{\tau_1}$. Vogliamo verificare che $A \in \mathscr{F}_{\tau_2}$, quindi che

$$A \cap \{\tau_2 \le n\} \in \mathscr{F}_n, \qquad \text{per ogni } n.$$

Ma, poiché $\tau_1 \le \tau_2$, $\{\tau_2 \le n\} \subset \{\tau_1 \le n\}$. Dunque possiamo scrivere

$$A \cap \{\tau_2 \le n\} = \underbrace{A \cap \{\tau_1 \le n\}}_{\in \mathscr{F}_n} \cap \underbrace{\{\tau_2 \le n\}}_{\in \mathscr{F}_n} \in \mathscr{F}_n.$$

**Soluzione dell'esercizio 3.9.** Infatti, $\tau_1 \wedge \tau_2$ è un tempo d'arresto tale che $\tau_1 \wedge \tau_2 \le \tau_i$, $i = 1, 2$, dunque $\mathscr{F}_{\tau_1 \wedge \tau_2} \subset \mathscr{F}_{\tau_i}$, $i = 1, 2$, e quindi $\mathscr{F}_{\tau_1 \wedge \tau_2} \subset \mathscr{F}_{\tau_1} \cap \mathscr{F}_{\tau_2}$. Viceversa, dimostriamo che $\mathscr{F}_{\tau_1 \wedge \tau_2} \supset \mathscr{F}_{\tau_1} \cap \mathscr{F}_{\tau_2}$. Preso $A \in \mathscr{F}_{\tau_1} \cap \mathscr{F}_{\tau_2}$, allora $A \cap \{\tau_i \le n\} \in \mathscr{F}_n$, per ogni $n$ e $i = 1, 2$. Quindi,

$$
\begin{aligned}
A \cap \{\tau_1 \wedge \tau_2 \le n\} &= A \cap \Big(\{\tau_1 \le n\} \cup \{\tau_2 \le n\}\Big) \\
&= \Big(A \cap \{\tau_1 \le n\}\Big) \cup \Big(A \cap \{\tau_2 \le n\}\Big) \in \mathscr{F}_n,
\end{aligned}
$$

cioè $A \in \mathscr{F}_{\tau_1 \wedge \tau_2}$.

**Soluzione dell'esercizio 3.10.** Osserviamo dapprima che

$$
\begin{aligned}
\{\tau_1 < \tau_2\} &= \cup_k \underbrace{\{\tau_1 < k\} \cap \{\tau_2 = k\}}_{\in \mathscr{F}_k \subset \mathscr{F}_\infty} \in \mathscr{F}_\infty \quad \text{e} \\
\{\tau_1 = \tau_2\} &= \cup_k \underbrace{\{\tau_1 = k\} \cap \{\tau_2 = k\}}_{\in \mathscr{F}_k \subset \mathscr{F}_\infty} \in \mathscr{F}_\infty.
\end{aligned}
$$

Inoltre, per ogni $n$,

$$
\begin{aligned}
\{\tau_1 < \tau_2\} \cap \{\tau_1 = n\} &= \{\tau_1 = n\} \cap \{\tau_2 \le n\}^c \in \mathscr{F}_n \quad \text{e} \\
\{\tau_1 < \tau_2\} \cap \{\tau_2 = n\} &= \{\tau_1 < n\} \cap \{\tau_2 = n\}^c \in \mathscr{F}_n.
\end{aligned}
$$

Dunque, $\{\tau_1 < \tau_2\} \in \mathscr{F}_{\tau_1}$ e $\{\tau_1 < \tau_2\} \in \mathscr{F}_{\tau_2}$, cioè $\{\tau_1 < \tau_2\} \in \mathscr{F}_{\tau_1} \cap \mathscr{F}_{\tau_2}$.