Appunti del corso di Probabilità e Processi Stocastici

(in continua revisione e aggiornamento)

Michele Gianfelice
Dipartimento di Matematica
Università della Calabria
Campus di Arcavacata
Ponte Pietro Bucci - cubo 30B
87036 Arcavacata di Rende (CS)
gianfelice@mat.unical.it

a.a. 2007/2008

Indice

Parte I

Elementi di Calcolo delle Probabilità

Capitolo 1

Alcune nozioni di teoria della misura

1.1 Insiemi misurabili

Sia Ω un insieme e $\mathcal{P}(\Omega)$ l'insieme delle parti di Ω, $\forall A \subseteq \Omega$ sia $A^c := \Omega \backslash A$.

Definizione 1 *Una collezione $\mathcal{A}$ d'insiemi di Ω è detta* algebra di Boole *se:*

1. $\Omega \in \mathcal{A}$;
2. *se* $A \in \mathcal{A}$ *allora* $A^c \in \mathcal{A}$;
3. *se* $\forall n \in \mathbb{N},\ A_i \in \mathcal{A},\ i = 1, .., n,$ *allora* $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$.

Definizione 2 *Un'algebra di Boole $\mathcal{A}$ è detta* σalgebra *se, data $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ una collezione numerabile d'insiemi di Ω allora $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.*

Osservazione 3 *Osserviamo che dalla precedente definizione segue che anche l'intersezione numerabile di elementi di $\mathcal{A}$ appartiene ad $\mathcal{A}$, poiché se $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$, dove $\forall n \in \mathbb{N}$, $A_n \in \mathcal{A}$, allora per definizione di algebra di Boole $\mathcal{A} \ni \left(\bigcup_{n\in\mathbb{N}} A_n\right)^c = \bigcap_{n\in\mathbb{N}} A_n^c$, ma $A_n^c \in \mathcal{A}, \forall n \in \mathbb{N}$.*

Inoltre notiamo che $\mathcal{P}(\Omega)$ verifica tutti gli assiomi che definiscono una σalgebra, pertanto l'insieme delle σalgebre non è vuoto.

Teorema 4 *Data una collezione $\mathcal{G}$ di sottoinsiemi di Ω esiste la più piccola σalgebra $\sigma(\mathcal{G})$ contenente $\mathcal{G}$ e $\sigma(\mathcal{G})$ è detta* σalgebra generata da $\sigma(\mathcal{G})$.

Dimostrazione: Sia $\mathcal{I}(\mathcal{G})$ la collezione di tutte le σalgebre di Ω contenenti $\mathcal{G}$, allora $\sigma(\mathcal{G}) := \bigcap_{\mathcal{B}\in\mathcal{I}(\mathcal{G})} \mathcal{B}$. ■

Esempio 1 Se Ω coincide con $\mathbb{R}$ munito della topologia euclidea e $\mathcal{G}$ coincide con la collezione degli aperti di $\mathbb{R}$ rispetto a tale topologia, $\sigma(\mathcal{G})$ è detta *σalgebra dei boreliani* di $\mathbb{R}$ e si denota col simbolo $\mathcal{B}(\mathbb{R})$.

Esercizio 1 Dimostrare che se $\mathcal{G} = \{(-\infty, q]\}_{q\in\mathbb{Q}}$ allora, $\sigma(\mathcal{G}) = \mathcal{B}(\mathbb{R})$.

Definizione 5 *Siano $\mathcal{A}$ e $\mathcal{B}$ due σalgebre d'insiemi di Ω. Allora, se ogni elemento A di $\mathcal{A}$ è anche elemento di $\mathcal{B}$ diremo che $\mathcal{A}$ è contenuta in $\mathcal{B}$, ovvero $\mathcal{A} \subset \mathcal{B}$, il che permette di definire una relazione d'ordine parziale tra σalgebre di sottoinsiemi di Ω.*

Osservazione 6 *Dalla definizione precedente segue che ogni σalgebra $\mathcal{A}$ di sottoinsiemi di Ω è contenuta in $\mathcal{P}(\Omega)$ e pertanto questa è detta σalgebra* massimale.

Definizione 7 *Sia $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{P}(\Omega)$ una collezione numerabile d'insiemi di Ω. Allora, se $\forall n \in \mathbb{N}:$*

- $A_n \subseteq A_{n+1}$, $\{A_n\}_{n\in\mathbb{N}}$ *è detta* crescente, $A_\infty := \bigcup_{n\in\mathbb{N}} A_n \equiv \lim_{n\uparrow\infty} A_n$ *e la collezione* $\{A_n\}_{n\in\mathbb{N}}$ *è detta accrescere a A_∞, in simboli $A_n \uparrow A_\infty$.*
- $A_n \supseteq A_{n+1}$, $\{A_n\}_{n\in\mathbb{N}}$ *è detta* decrescente, $A_\infty := \bigcap_{n\in\mathbb{N}} A_n \equiv \lim_{n\uparrow\infty} A_n$ *e la collezione* $\{A_n\}_{n\in\mathbb{N}}$ *è detta decrescere a A_∞, in simboli $A_n \downarrow A_\infty$.*

  Inoltre, una collezione numerabile d'insiemi di Ω che sia crescente o decrescente è detta monotona.

Definizione 8 *Una collezione $\mathcal{M}$ di sottoinsiemi di Ω tale che, se $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{M}$ è una collezione monotona, allora $A_\infty \in \mathcal{M}$ è detta* classe monotona *di elementi di $\mathcal{P}(\Omega)$.*

Proposizione 9 *Una σalgebra $\mathcal{A}$ di sottoinsiemi di Ω è una classe monotona.*

Dimostrazione: Se $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$ è una successione crescente, allora $A_\infty = \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$ per definizione di σalgebra. Inoltre, considerando $\{A_n^c\}_{n\in\mathbb{N}}$, che è decrescente, si ha

$$A_\infty^c = \bigcap_{n\in\mathbb{N}} A_n^c = \left(\bigcup_{n\in\mathbb{N}} A_n\right)^c \in \mathcal{A} \,, \tag{1.1}$$

ovvero che $\mathcal{A}$ è una classe monotona. ■

Osservazione 10 *Se $\mathcal{G} \subset \mathcal{P}(\Omega)$, dalla definizione di σalgebra generata da $\mathcal{G}$, resta allora definita $\mathcal{M}(\mathcal{G})$ la classe monotona generata da $\mathcal{G}$, ovvero la più piccola classe monotona contenente gli elementi di $\mathcal{G}$.*

Teorema 11 *Se $\mathcal{B}$ è un'algebra di Boole, $\mathcal{M}(\mathcal{B})$ e $\sigma(\mathcal{B})$ coincidono.*

Dimostrazione: $\sigma(\mathcal{B})$ è una classe monotona, quindi $\sigma(\mathcal{B}) \supset \mathcal{M}(\mathcal{B})$.
D'altra parte, se dato $A \in \mathcal{P}(\Omega)$

$$\mathcal{A}(A;\mathcal{B}) := \{B \in \mathcal{P}(\Omega) : A \cup B, A \cap B, A \triangle B \in \mathcal{M}(\mathcal{B})\}\,, \tag{1.2}$$

poiché le applicazioni

$$\cup, \cap, \triangle : \mathcal{P}(\Omega) \times \mathcal{P}(\Omega) \longmapsto \mathcal{P}(\Omega) \tag{1.3}$$

sono simmetriche, $B \in \mathcal{A}(A;\mathcal{B}) \iff A \in \mathcal{A}(B;\mathcal{B})$. Inoltre, $\forall A \in \mathcal{P}(\Omega), \mathcal{A}(A;\mathcal{B})$ è una classe monotona, infatti, se $\{B_n\}_{n\in\mathbb{N}} \subset \mathcal{A}(A;\mathcal{B})$ è crescente (decrescente), come pure $\{A \cup B_n\}_{n\in\mathbb{N}}$, mentre $\{A \cap B_n\}_{n\in\mathbb{N}}, \{A \backslash B_n\}_{n\in\mathbb{N}}$ sono decrescenti (crescenti), $\{B_n \backslash A\}_{n\in\mathbb{N}}$ è crescente (decrescente) ed il loro limite è in $\mathcal{M}(\mathcal{B})$, quindi lo è anche il limite di $\{A \triangle B_n\}_{n\in\mathbb{N}}$.

Siccome $\mathcal{B} \subset \mathcal{M}(\mathcal{B})$, se $A \in \mathcal{B}$ allora, $\forall B \in \mathcal{B}$, $B \in \mathcal{A}(A;\mathcal{B})$. Perciò $\mathcal{B} \subset \mathcal{A}(A;\mathcal{B})$ e $\mathcal{A}(A;\mathcal{B}) \supset \mathcal{M}(\mathcal{B})$. Tuttavia, $\forall B \in \mathcal{M}(\mathcal{B})$, se $A \in \mathcal{B}$, $A \in \mathcal{A}(B;\mathcal{B})$, ovvero $\mathcal{B} \subset \mathcal{A}(B;\mathcal{B})$. Dunque, $\forall B, B' \in \mathcal{M}(\mathcal{B})$, $B' \cup B, B' \cap B, B' \triangle B \in \mathcal{M}(\mathcal{B})$. Quindi, posto $B' = \Omega$, $B \in \mathcal{M}(\mathcal{B})$ implica $B^c \in \mathcal{M}(\mathcal{B})$, ovvero $\mathcal{M}(\mathcal{B})$ è un'algebra di Boole. Siccome però se $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{M}(\mathcal{B})$, $\{B_n\}_{n\in\mathbb{N}}$ tale che $\forall n \in \mathbb{N}$, $B_n := \bigcup_{k=0}^{n} A_k$ è crescente $\lim_{n\uparrow\infty} A_n = \lim_{n\uparrow\infty} B_n \in \mathcal{M}(\mathcal{B})$ e dunque $\mathcal{M}(\mathcal{B})$ è una σalgebra, perciò $\mathcal{M}(\mathcal{B}) \supset \sigma(\mathcal{B})$. ■

Se $\mathcal{A}$ e $\mathcal{B}$ sono due σalgebre indichiamo con $\mathcal{A} \vee \mathcal{B}$ la σalgebra $\sigma\{\mathcal{A}, \mathcal{B}\}$, ovvero la quella generata dagli eventi di $\mathcal{A}$ e di $\mathcal{B}$. Inoltre, se $\mathcal{I}$ è un insieme di indici e $\{\mathcal{B}\}_{i\in\mathcal{I}}$ una famiglia di σalgebre indicizzata da $\mathcal{I}$ poniamo $\bigvee_{i\in\mathcal{I}} \mathcal{B}_i$ la σalgebra generata dagli eventi degli elementi di $\{\mathcal{B}\}_{i\in\mathcal{I}}$.

Definizione 12 *Dato un insieme Ω ed una σalgebra di suoi sottoinsiemi $\mathcal{A}$, la coppia $(\Omega, \mathcal{A})$ è detta* spazio misurabile.

1.2 Generalità sulle misure

1.2.1 Misure positive

Definizione 13 *Sia E un insieme ed $\mathcal{E}$ un'algebra di Boole di suoi sottoinsiemi. Una* misura positiva finitamente additiva *μ su $(E, \mathcal{E})$ è un'applicazione*

$$\mathcal{E} \ni A \longmapsto \mu(A) \in \overline{\mathbb{R}^+} \tag{1.4}$$

che per ogni $A, B \in \mathcal{E}$ disgiunti, cioè $A \cap B = \emptyset$, $\mu(A \cup B) = \mu(A) + \mu(B)$.

Definizione 14 *Sia E un insieme ed $\mathcal{E}$ una σalgebra dei suoi sottoinsiemi. Una* misura positiva *μ su $(E, \mathcal{E})$ è un'applicazione*

$$\mathcal{E} \ni A \longmapsto \mu(A) \in \overline{\mathbb{R}^+} \tag{1.5}$$

σadditiva*, ovvero tale che, per ogni collezione numerabile $\{A_n\}_{n\geq 1}$ di elementi di $\mathcal{E}$ i cui termini sono insiemi mutuamente disgiunti, cioè $A_i \cap A_j = \emptyset$ se $i \neq j$, $\mu\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n\geq 1} \mu(A_n)$.*

Data una misura positiva μ su $(E, \mathcal{E})$, la terna $(E, \mathcal{E}, \mu)$ è detta *spazio di misura.* Inoltre, sia

$$\overline{\mathcal{E}}^{\mu} := \{A \subseteq E : \exists A_-, A_+ \in \mathcal{E} \text{ tale che } A_- \subseteq A \subseteq A_+ \text{ e } \mu(A_+ \backslash A_-) = 0\}. \tag{1.6}$$

$\overline{\mathcal{E}}^{\mu} \supseteq \mathcal{E}$ ed è una σalgebra, pertanto $\overline{\mathcal{E}}^{\mu}$ è detta *completamento di $\mathcal{E}^{\mu}$ rispetto a μ.* Ne segue che μ si può estendere ad $\overline{\mathcal{E}}^{\mu}$ ponendo $\forall A \in \overline{\mathcal{E}}^{\mu}, \bar{\mu}(A) = \mu(A_-)$ e quindi considerare lo spazio di misura $\left(E, \overline{\mathcal{E}}^{\mu}, \bar{\mu}\right)$ *completamento* di $(E, \mathcal{E}, \mu)$. Dunque, se $\overline{\mathcal{E}}^{\mu} = \mathcal{E}$, $(E, \mathcal{E}, \mu)$ è detto *completo.*

Sia ora E uno spazio topologico di Haussdorff:

- *localmente compatto*, ovvero tale che ogni punto di E ammette un intorno compatto;
- *σcompatto o numerabile all'infinito*, ovvero tale che esiste una successione crescente di sottoinsiemi compatti di E, $\{K_n\}_{n\geq 1}$ tali che $\forall n \geq 1$, $K_n \subseteq K_{n+1}$, esaustiva di E, cioè $E = \bigcup_{n\geq 1} K_n$.

Data $\mathcal{E} = \mathcal{B}(E)$ la σalgebra dei boreliani di E, ovvero quella generata dagli aperti di E, una misura positiva μ su $(E, \mathcal{B}(E))$ è detta localmente finita se $\forall K \subset E$ compatto, $\mu(K) < \infty$.

Il teorema di Radon-Riesz afferma che esiste una corrispondenza biunivoca tra i funzionali lineari positivi sullo spazio lineare $C_0(E)$ delle funzioni continue a supporto compatto su E a valori reali ed un sottoinsieme $\mathcal{M}^+(E)$ delle misure localmente finite positive su $(E, \mathcal{B}(E))$, dette *misure di Radon positive*, che sono il completamento di restrizioni a $(E, \mathcal{B}(E))$ di determinate misure definite su $(E, \mathcal{E}')$, con $\mathcal{E}'$ σalgebra contenente $\mathcal{B}(E)$. In generale, se μ^ϕ è una misura su $(E, \mathcal{E}')$ che origina da un funzionale positivo ϕ su $C_0(E)$ tramite la costruzione di Radon-Riesz, il completamento $\left(E, \overline{\mathcal{B}(E)}^{\bar{\mu}^\phi}, \bar{\mu}^\phi\right)$ di $\left(E, \mathcal{B}(E), \mu_0^\phi\right)$, con μ_0^ϕ restrizione a $(E, \mathcal{B}(E))$ di μ^ϕ, non coincide con $\left(E, \mathcal{E}', \mu^\phi\right)$. Ciò è però vero nel caso in cui E sia anche metrizzabile. Infatti in tal caso data una qualsiasi misura positiva μ_0 localmente finita su $(E, \mathcal{B}(E))$ esiste una misura positiva di Radon $\bar{\mu}$ coincidente con μ_0 su $(E, \mathcal{B}(E))$. Inoltre, $\bar{\mu}$ è *regolare*, ovvero $\forall B \in \overline{\mathcal{B}(E)}^{\bar{\mu}}, \varepsilon > 0$, $\exists C, A \in \mathcal{B}(E)$ con C chiuso e A aperto, tali che $C \subset B \subset A$ e $\bar{\mu}(A \backslash C) < \varepsilon$.

1.2.2 Misure segnate

Se E è metrizzabile e compatto, è uno spazio di Hausdorff compatto, quindi localmente compatto e σcompatto. Munendo $C_0(E) \equiv C(E)$ della norma $\|\cdot\|_\infty$, mediante la decomposizione di Jordan si può costruire lo spazio lineare $\mathcal{M}(E)$ delle misure di Radon segnate che risulta essere in corrispondenza biunivoca con il duale di $C(E)$.

Sia quindi E metrizzabile, localmente compatto e σcompatto. Consideriamo gli spazi lineari di funzioni reali su E:

- $C_0(E)$, spazio delle funzioni continue a supporto compatto;
- $C_\infty(E)$, spazio delle funzioni continue *nulle all'*∞, ovvero tali che $\forall \varepsilon > 0$, $\exists K_\varepsilon \subset E$ compatto tale che $|f(x)| < \varepsilon \, \forall x \in K_\varepsilon^c$;
- $C_b(E)$, spazio delle funzioni continue limitate.

Chiaramente, $C_0(E) \subset C_\infty(E) \subset C_b(E)$, valendo l'uguaglianza se E è compatto. Munendo questi spazi della norma $\|\cdot\|_\infty$ si ottengono altrettanti spazi normati tali che:

- $C_b(E)$ è di Banach;
- $C_\infty(E)$ è un sottospazio chiuso di $C_b(E)$;
- $C_0(E)$ è denso in $C_\infty(E)$.

Tramite la compattificazione ad un punto di E, è possibile costruire un isomorfismo tra $\mathcal{M}\left(\dot{E}\right)$, con $\dot{E}$ compattificazione ad un punto di E, ed il sottoinsieme $\mathcal{M}^1(E)$ delle misure di Radon segnate finite di $\mathcal{M}(E)$, il quale risulta essere in corrispondenza biunivoca con il duale di $C_\infty(E)$.

1.3 Successioni di misure

Sia E metrizzabile, localmente compatto e σcompatto.

1.3.1 Convergenza forte

Indicando con $\mathcal{M}^{1,+}(E)$ il sottoinsieme di $\mathcal{M}^1(E)$ collezione delle misure positive finite su E e definendo su $\mathcal{M}^1(E)$ la norma $\|\mu\|_{\mathcal{M}^1} := |\mu| = \mu^+(E) + \mu^-(E)$, dove $\mu^+, \mu^- \in \mathcal{M}^{1,+}(E)$ sono gli elementi della decomposizione di Hahn-Jordan di $\mu \in \mathcal{M}^1(E)$, $\{\mu_n\}_{n\geq 1} \subset \mathcal{M}^1(E)$ converge *fortemente* a $\mu \in \mathcal{M}^1(E)$, se converge in norma $\|\cdot\|_{\mathcal{M}^1}$.

1.3.2 Convergenza debole e dualità

Per quanto appena esposto, è possibile allora definire la seguenti nozioni di convergenza debole:

1. $\{\mu_n\}_{n\geq 1} \subset \mathcal{M}(E)$ converge *vagamente* a $\mu \in \mathcal{M}(E)$, se $\forall f \in C_0(E)$, $\{\mu_n(f)\}_{n\geq 1}$ converge a $\mu(f)$, ovvero $\{\mu_n\}_{n\geq 1}$ converge nella topologia *-debole di $C_0^*(E)$.
2. $\{\mu_n\}_{n\geq 1} \subset \mathcal{M}^1(E)$ converge *debolmente* a $\mu \in \mathcal{M}^1(E)$, se $\forall f \in C_\infty(E)$, $\{\mu_n(f)\}_{n\geq 1}$ converge a $\mu(f)$, ovvero $\{\mu_n\}_{n\geq 1}$ converge nella topologia *-debole di $C_\infty^*(E)$.
3. $\{\mu_n\}_{n\geq 1} \subset \mathcal{M}^1(E)$ converge *strettamente* a $\mu \in \mathcal{M}^1(E)$, se $\forall f \in C_b(E)$, $\{\mu_n(f)\}_{n\geq 1}$ converge a $\mu(f)$.

Vale il seguente diagramma di relazioni tra le nozioni di convergenza appena definite:

convergenza in norma $\|\cdot\|_{\mathcal{M}^1}$ $\Longrightarrow$ convergenza stretta $\Longrightarrow$ convergenza debole $\Longrightarrow$ convergenza vaga.

Definizione 15 *Una misura $\mu \in \mathcal{M}^1(E)$ è detta* tesa *se $\forall \varepsilon > 0$, $\exists K_\varepsilon \subset E$ compatto tale che $|\mu|(K_\varepsilon^c) < \varepsilon$.*

Equivalentemente μ è tesa se il suo supporto è σcompatto, *ovvero può essere rappresentato come unione al più numerabile di compatti.*

Definizione 16 *Sia $\mathcal{I}$ un insieme di indici. Una famiglia di misure $\{\mu_i\}_{i\in\mathcal{I}} \subset \mathcal{M}^1(E)$ è detta* tesa *se $\forall \varepsilon > 0$, $\exists K_\varepsilon \subset E$ compatto tale che $\forall i \in \mathcal{I}$, $|\mu_i|(K_\varepsilon^c) < \varepsilon$.*

Teorema 17 *Sia $\{\mu_n\}_{n\geq 1} \subset \mathcal{M}^1(E)$. $\{\mu_n\}_{n\geq 1}$ converge debolmente se e solo se $\exists c > 0$ ed un sottoinsieme D denso in $C_\infty(E)$ tale che $\|\mu_n\|_{\mathcal{M}^1} \leq c$, $\forall n \geq 1$ e la successione $\{\mu_n(g)\}_{n\geq 1}$ converge $\forall g \in D$.*

Dimostrazione: L'implicazione diretta nel caso generale è più difficile da dimostrare di quella in cui si suppone che $\exists c > 0 : \|\mu_n\|_{\mathcal{M}^1} \leq c$, $\forall n \geq 1$. Poiché nel seguito si considereranno soltanto misure di probabilità, ci restingeremo a questo caso.

$\Rightarrow$ Ovviamente se $\{\mu_n\}_{n\geq 1}$ converge debolmente la successione $\{\mu_n(g)\}_{n\geq 1}$ converge $\forall g \in D$ con D denso in $C_\infty(E)$. Inoltre, la palla di raggio c nella topologia debole di $C^*_\infty(E)$, che è isomorfo a $\mathcal{M}^1(E)$, è compatta, quindi, se μ_∞ è il limite di $\{\mu_n\}_{n\geq 1}$ nella topologia debole, $\|\mu_\infty\|_{\mathcal{M}^1} \leq c$.

$\Leftarrow$ La successione di funzionali lineari su $C_\infty(E)$ $\{l_n\}_{n\geq 1}$ tali che $\forall n \geq 1, g \in C_\infty(E)$, $l_n(g) := \mu_n(g)$ è uniformemente limitata ed equicontinua, infatti $\forall n \geq 1$,

$$\|l_n\|_{C^*_\infty(E)} = \|\mu_n\|_{\mathcal{M}^1(E)} \leq c\,; \quad |l_n(g-f)| \leq c\,\|g-f\|_\infty\,, \tag{1.7}$$

inoltre è convergente su un sottoinsieme D denso in $C_\infty(E)$. Quindi, per il teorema di Ascoli-Arzelà, è convergente su tutto $C_\infty(E)$. Allora, l_∞ limite di $\{l_n\}_{n\geq 1}$ definisce una misura di Radon $\mu_\infty \in \mathcal{M}^1(E)$ e dunque $\{\mu_n\}_{n\geq 1}$ converge debolmente a μ_∞.

■

Teorema 18 *Sia $\{\mu_n\}_{n\geq 1} \subset \mathcal{M}^1(E)$. $\{\mu_n\}_{n\geq 1}$ converge strettamente se solo se $\{\mu_n\}_{n\geq 1}$ converge debolmente e $\forall \varepsilon > 0$, $\exists K_\varepsilon \subset E$ compatto e $N_\varepsilon > 0$ tale che, $\forall n \geq N_\varepsilon$, $|\mu_n|(K^c_\varepsilon) < \varepsilon$, ovvero la famiglia $\{\mu_n\}_{n\geq N_\varepsilon+1}$ è tesa.*

Dimostrazione:

$\Rightarrow$ $\{\mu_n\}_{n\geq 1}$ converge strettamente allora converge debolmente e, dall'implicazione diretta del teorema precedente, $\exists c > 0 : \|\mu_n\|_{\mathcal{M}^1} \leq c$, $\forall n \geq 1$. Quindi se $\{\mu_n\}_{n\geq 1}$ converge debolmente a μ_∞ allora $|\mu_\infty|(E) = \|\mu_\infty\|_{\mathcal{M}^1} \leq c$. Dunque, $\forall \varepsilon > 0$, $\exists K_\varepsilon : |\mu_\infty|(K^c_\varepsilon) < \frac{\varepsilon}{2}$. Allora, tramite $\dot{E}$, è possibile trovare $f_\varepsilon \in C_b(E) : 0 < f_\varepsilon \leq 1$ e *supp* $(f_\varepsilon) \subset K^c_\varepsilon$,

$$|\mu_\infty|(f_\varepsilon) \leq |\mu_\infty|(K^c_\varepsilon) < \frac{\varepsilon}{2}. \tag{1.8}$$

Inoltre, $\exists N_\varepsilon : \forall n > N_\varepsilon$, $||\mu_n|(f_\varepsilon) - |\mu_\infty|(f_\varepsilon)| < \frac{\varepsilon}{2}$. Pertanto,

$$|\mu_n|(f_\varepsilon) < |\mu_\infty|(f_\varepsilon) + \frac{\varepsilon}{2} < |\mu_\infty|(K^c_\varepsilon) + \frac{\varepsilon}{2} < \varepsilon\,. \tag{1.9}$$

Poiché E è localmente compatto e σcompatto, Esiste una successione $\{H_n\}_{n\geq 1}$ di sottoinsiemi compatti di E, esaustiva di E, tale che $\forall n \geq 1$, $H_n \subset \overset{\circ}{H}_{n+1}$. Allora, sia n_ε tale che $\overline{H^c_{n_\varepsilon}} \subset$ *supp* (f_ε). Si ha che $\forall n \geq 1$ $|\mu_n|\left(\overline{H^c_{n_\varepsilon}}\right) \leq |\mu_n|(f_\varepsilon)$. Quindi, $\forall n > N_\varepsilon$, $|\mu_n|\left(\overline{H^c_{n_\varepsilon}}\right) < \varepsilon$.

$\Leftarrow$ Dato $\varepsilon > 0$, sia $K_\varepsilon \subset E$ compatto e $N_\varepsilon > 0$ tale che, $\forall n \geq N_\varepsilon$, $|\mu_n|(K^c_\varepsilon) < \varepsilon$. Se $g \in C_0(E)$ tale che $g(x) = 1$, $\forall x \in K_\varepsilon$, $\forall f \in C_b(E)$, $f = gf + u$ con u tale che *supp* $(u) \subset K^c_\varepsilon$. Quindi, $\forall n \geq 1$,

$$\mu_n(f) = \mu_n(fg) + \mu_n(u)\,. \tag{1.10}$$

Poiché per ipotesi $\{\mu_n\}_{n\geq 1}$ converge debolmente, indicando con μ_∞ il suo limite, $\{\mu_n(fg)\}_{n\geq 1}$ converge a $\mu_\infty(fg)$. Inoltre, $|\mu_n(u)| \leq \|u\|_\infty |\mu_n|(K^c_\varepsilon) \leq \varepsilon\,\|f\|_\infty$.

■

Teorema 19 *Sia* $\{\mu_n\}_{n\geq 1} \subset \mathcal{M}^{1,+}(E)$ *tale che* $\{\mu_n\}_{n\geq 1}$ *converge debolmente a* μ *e* $\{\mu_n(E)\}_{n\geq 1}$ *converge a* $\mu(E) < \infty$*. Allora* $\{\mu_n\}_{n\geq 1}$ *converge strettamente.*

Dimostrazione: Basta dimostrare che $\forall \varepsilon > 0$, $\exists K_\varepsilon \subset E$ compatto e $N_\varepsilon > 0$ tale che, $\forall n \geq N_\varepsilon$, $|\mu_n|(K_\varepsilon^c) < \varepsilon$ e la tesi segue dall'implicazione inversa del teorema precedente. Dato $\varepsilon > 0$, sia $K_\varepsilon \subset E$ compatto tale che $\mu_n\left(\overline{K_\varepsilon^c}\right) < \varepsilon$ e sia $f \in C_0(E)$ tale che $supp(f) \subset K_\varepsilon$, $0 \leq f \leq 1$ e $\mu(f) > \|\mu\|_{\mathcal{M}^1} - \varepsilon$. Sia $N_\varepsilon > 0$ tale che, $\forall n \geq N_\varepsilon$,

$$|\mu_n(E) - \mu(E)| < \varepsilon, \quad |\mu_n(f) - \mu(f)| < \varepsilon\,. \tag{1.11}$$

Allora, $\mu_n(E) \geq \mu_n(f) + \mu_n\left(\overline{K_\varepsilon^c}\right)$, pertanto,

$$\mu_n\left(\overline{K_\varepsilon^c}\right) \leq \mu_n(E) - \mu_n(f) \leq \mu_n(E) + |\mu(f) - \mu_n(f)| - \mu(f) \leq |\mu_n(E) - \mu(E)| + 2\varepsilon. \tag{1.12}$$

■

Teorema 20 *Sia* E *uno spazio metrico completo e separabile, allora ogni* $\mu \in \mathcal{M}^1(E)$ *è tesa.*

Dimostrazione: È sufficiente restringersi al caso in cui $\mu \in \mathcal{M}^{1,+}(E)$ e $\mu(E) = 1$. Poiché E è separabile, esiste una successione $\{B_k^{(n)}\}_{k\geq 1}$ di palle aperte di E di raggio $\frac{1}{n}$ che ricopre E. Allora, $\forall \varepsilon > 0$ esiste $N_\varepsilon^{(n)}$ tale che $\mu\left(\bigcup_{k\leq N_\varepsilon^{(n)}} B_k^{(n)}\right) > 1 - \frac{\varepsilon}{2^n}$. Quindi, poiché E è completo, l'insieme $\bigcap_{n\geq 1}\bigcup_{k\leq N_\varepsilon^{(n)}} B_k^{(n)}$ è totalmente limitato ed ha chiusura compatta K. Pertanto, $\nu(K) > 1 - \varepsilon$. ■

Osservazione 21 *Quest'ultimo risultato è valido anche sotto le ipotesi più deboli per cui* E *sia* topologicamente completo*, ovvero* E *ammette una metrica equivalente rispetto a cui è compatto, e che* μ *sia* separabile*, ovvero abbia supporto separabile. In generale è sufficiente restringersi al caso in cui* E *sia uno spazio topologico* polacco*, ovvero omeomorfo ad uno spazio metrico completo separabile. Tuttavia in molte applicazioni si assume inoltre che* E *sia localmente compatto e* σ*compatto. La mancanza di questi due ultimi requisiti infatti non garantisce la possibilità di definire una misura finita per mezzo del teorema di Radon-Riesz, in quanto in tal caso ogni misura di finita su* $(E, \mathcal{B}(E))$ *rappresenta un funzionale lineare continuo positivo su* $C_0(E)$ *ma il viceversa non è in genarale vero. Inoltre,* $C_0(E)$ *non è in generale denso in* $C_\infty(E)$*.*

1.4 Misure di probabilità

Definizione 22 *Sia* $(\Omega, \mathcal{A})$ *uno spazio misurabile. Una misura positiva* σ*additiva* $\mathbb{P}$ *su* $(\Omega, \mathcal{A})$ *tale che* $\mathbb{P}(\Omega) = 1$ *è detta* misura di probabilità *su* $(\Omega, \mathcal{A})$*.*

Osservazione 23 *Una misura di probabilità ha le seguenti proprietà:*

1. $\mathbb{P}(\emptyset) = \mathbb{P}(\Omega^c) = 0.$

2. *Se* $A, B \in \mathcal{A}$, $A \subseteq B$,

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \backslash A) \geq \mathbb{P}(A) \ . \tag{1.13}$$

3. *Se* $A, B \in \mathcal{A}$,

$$\mathbb{P}(B \cup A) = \mathbb{P}(A \backslash B) + \mathbb{P}(B \backslash A) + \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \tag{1.14}$$

$$\leq \mathbb{P}(A) + \mathbb{P}(B) \ , \tag{1.15}$$

il che implica

$$\mathbb{P}\left(\bigcup_{i \in \mathbb{N}} A_i\right) \leq \sum_{i \in \mathbb{N}} \mathbb{P}(A_i) \ . \tag{1.16}$$

4. *Se* $\{A_n\}_{n \in \mathbb{N}}$ *è crescente, allora*

$$\mathbb{P}(A_\infty) = \mathbb{P}\left(\lim_{n \uparrow \infty} A_n\right) = \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \mathbb{P}\left(A_0 \bigvee \bigvee_{n \geq 1} (A_n \backslash A_{n-1})\right) \tag{1.17}$$

$$= \mathbb{P}(A_0) + \sum_{n \geq 1} \mathbb{P}(A_n \backslash A_{n-1}) = \mathbb{P}(A_0) + \sum_{n \geq 1} (\mathbb{P}(A_n) - (A_{n-1})) = \lim_{n \uparrow \infty} \mathbb{P}(A_n) \ .$$

Esercizio 2 Dimostrare che se $\{A_n\}_{n \in \mathbb{N}}$ è decrescente, allora

$$\mathbb{P}(A_\infty) = \mathbb{P}\left(\lim_{n \uparrow \infty} A_n\right) = \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} A_n\right) = \lim_{n \uparrow \infty} \mathbb{P}(A_n)$$

ed inoltre che, se $A_\infty = \emptyset$, allora ciò è equivalente alla proprietà 2 della definizione di misura di probabilità.

1.4.1 Spazi di Probabilità

Sia $\mathfrak{P}(E, \mathcal{E}) \subset \mathcal{M}^{+,1}(E)$ la collezione delle misure di probabilità su $(E, \mathcal{E})$. Scelta, $\mu \in \mathfrak{P}(E, \mathcal{E})$, la terna $(E, \mathcal{E}, \mu)$ è detta spazio *di probabilità (probabilizzato).* Inoltre, se $\overline{\mathcal{E}}^\mu$ è la σalgebra completamento di $\mathcal{E}$ rispetto a μ definita nella (1.6), ne segue che μ si può estendere a $\overline{\mathcal{E}}^\mu$ e quindi considerare lo spazio di probabilità $\left(E, \overline{\mathcal{E}}^\mu, \mu\right)$ completamento di $(E, \mathcal{E}, \mu)$. Se $\mathcal{E} = \overline{\mathcal{E}}^\mu$, $(E, \mathcal{E}, \mu)$ è uno spazio di probabilità completo.

1.4.2 Convergenza in distanza variazionale

Definizione 24 *Sia* $(E, \mathcal{E})$ *uno spazio misurabile, si definisce* distanza in variazione *tra due misure di probabilità* $\mu, \nu \in \mathfrak{P}(E, \mathcal{E})$ *la quantità*

$$Var(\mu - \nu) := \frac{1}{2} \sup_{A \in \mathcal{E}} |\mu(A) - \nu(A)| \, . \tag{1.18}$$

Dunque, una successione di misure di probabilità $\{\mu_n\}_{n \geq 1} \subset \mathfrak{P}(E, \mathcal{E})$ *converge a* μ in variazione *se* $\lim_{n \to \infty} Var(\mu_n - \mu) = 0$.

Osservazione 25 *Se E è localmente compatto e σcompatto ed $\mathcal{E}=\mathcal{B}(E)$, $\forall\nu_1,\nu_2\in\mathfrak{P}(E)$*

$$Var(\nu_1-\nu_2)=\frac{1}{2}\|\nu_2-\nu_2\|_{\mathcal{M}^1}. \tag{1.19}$$

Infatti, posto $\mu:=\nu_1-\nu_2\in\mathcal{M}^1(E)$, gli elementi della cui decomposizione di Hahn-Jordan denotiamo con μ^+,μ^-, poiché $\forall A\in\mathcal{B}(E)$

$$\nu_1(A)-\nu_2(A)=\nu_1(A^c)-\nu_2(A^c), \tag{1.20}$$

si ha

$$\|\mu\|_{\mathcal{M}^1}=\mu^+(E)+\mu^-(E)=(\nu_1-\nu_2)(A)+(\nu_1-\nu_2)(A^c)\leq 2\sup_{A\in\mathcal{B}(E)}|\nu_1(A)-\nu_2(A)|. \tag{1.21}$$

Inoltre, $\forall A\in\mathcal{B}(E)$

$$\begin{aligned}2|\nu_1(A)-\nu_2(A)|&=|\nu_1(A)-\nu_2(A)|+|\nu_1(A^c)-\nu_2(A^c)|\\&=\mu^+(A)+\mu^-(A)+\mu^+(A^c)+\mu^-(A^c)\leq\|\mu\|_{\mathcal{M}^1}.\end{aligned} \tag{1.22}$$

Dunque, in tal caso la nozione di convergenza in variazione equivale a quella di convergenza forte in $\mathcal{M}^1(E)$.

Osservazione 26 *Nel caso in cui $(E,\mathcal{B}(E))=(\mathbb{R},\mathcal{B}(\mathbb{R}))$ e μ,ν sono a.c. rispetto alla misura di Lebesgue, con densità rispettivamente f e g,*

$$\|\mu-\nu\|_{\mathcal{M}^1}=\int_{\mathbb{R}}dx\,|f(x)-g(x)|. \tag{1.23}$$

Analogamente se E è discreto e $\{\mu_i\}_{i\geq1}$, $\{\nu_i\}_{i\geq1}$ sono i valori assunti rispettivamente da μ e ν sugli atomi di E,

$$\|\mu-\nu\|_{\mathcal{M}^1}=\sum_{i\geq1}|\mu_i-\nu_i|. \tag{1.24}$$

1.4.3 Convergenza debole

Teorema 27 *Sia E sia uno spazio topologico polacco, localmente compatto e σcompatto. Una successione $\{\mu_n\}_{n\geq1}\subset\mathfrak{P}(E)$ converge vagamente se e solo se converge strettamente.*

Dimostrazione: Poiché $C_0(E)$ è denso in $C_\infty(E)$ e $\forall n\geq1$, $\|\mu_n\|_{\mathcal{M}^1}=1$, dall'implicazione inversa del Teorema 17 segue che $\{\mu_n\}_{n\geq1}$ converge debolmente a μ_∞ tale che

$$\mu_\infty(E)=1=\mu_n(E)\;. \tag{1.25}$$

Pertanto per il Teorema 19 si ha che $\{\mu_n\}_{n\geq1}$ converge strettamente. ■

Osservazione 28 *Se E è polacco, localmente compatto e σcompatto, per il precedente risultato la convergenza stretta di $\{\mu_n\}_{n\geq 1} \subset \mathfrak{P}(E)$ è equivalente a quella vaga e quindi a quella debole. Pertanto nel linguaggio comune del Calcolo delle Probabilità e della Statistica Matematica tali nozioni di convergenza per le misure di probabilità vanno sotto il nome di convergenza debole.*

Definizione 29 *Sia E sia uno spazio metrico. Una successione di misure di probabilità $\{\mathbb{P}_n\}_{n\geq 1}$ converge* in generale *ad una misura di probabilità $\mathbb{P}$ se, $\forall A \in \mathcal{B}(E)$ tale che $\mathbb{P}(\partial A) = 0$, la successione numerica $\{\mathbb{P}_n(A)\}_{n\geq 1}$ converge a $\mathbb{P}(A)$.*

Teorema 30 *Sia E sia uno spazio metrico. Data una successione di misure di probabilità $\{\mathbb{P}_n\}_{n\geq 1}$, le seguenti affermazioni sono equivalenti:*

1. *$\{\mathbb{P}_n\}_{n\geq 1}$ converge debolmente a $\mathbb{P}$;*
2. *$\overline{\lim}_n \mathbb{P}_n(A) \leq \mathbb{P}(A) \quad A \in \mathcal{B}(E)$ chiuso;*
3. *$\underline{\lim}_n \mathbb{P}_n(A) \geq \mathbb{P}(A) \quad A \in \mathcal{B}(E)$ aperto;*
4. *$\{\mathbb{P}_n\}_{n\geq 1}$ converge in generale a $\mathbb{P}$.*

Dimostrazione:

1. $\Rightarrow$ 2. Sia ρ una metrica equivalente, $A \in \mathcal{B}(E)$ chiuso e

$$f_A^{(\varepsilon)}(x) := \left(1 - \frac{\rho(x, A)}{\varepsilon}\right) \vee 0, \tag{1.26}$$

$$A_\varepsilon := \{x \in E : f_A^{(\varepsilon)}(x) > 0\} = \{x \in E : \rho(x, A) < \varepsilon\}, \tag{1.27}$$

con $E \ni x \longmapsto \rho(x, A) := \inf_{y\in A} \rho(x, y) \in \mathbb{R}^+$. Allora, poiché

$$\mathbb{P}_n(A) = \int_E \mathbb{P}_n(dx)\, \mathbf{1}_A(x) \leq \int_E \mathbb{P}_n(dx)\, f_A^{(\varepsilon)}(x)\ , \tag{1.28}$$

si ha

$$\overline{\lim}_n \mathbb{P}_n(A) \leq \overline{\lim}_n \int_E \mathbb{P}_n(dx)\, f_A^{(\varepsilon)}(x) = \int_E \mathbb{P}(dx)\, f_A^{(\varepsilon)}(x) \leq \mathbb{P}(A_\varepsilon) \underset{\varepsilon\downarrow 0}{\longrightarrow} \mathbb{P}(A)\,. \tag{1.29}$$

2. $\Rightarrow$ 3. Se A è aperto, A^c è chiuso, perciò

$$\underline{\lim}_n \mathbb{P}_n(A) = \underline{\lim}_n (1 - \mathbb{P}_n(A^c)) = 1 - \overline{\lim}_n \mathbb{P}_n(A^c) \geq 1 - \mathbb{P}(A^c) = \mathbb{P}(A)\,. \tag{1.30}$$

3. $\Rightarrow$ 4. Sia $\mathring{A} := A \backslash \partial A$. Dato che per ipotesi $\mathbb{P}(\partial A) = 0$ e passando al complementare 3. $\Rightarrow$ 2., si ha

$$\overline{\lim}_n \mathbb{P}_n(A) \leq \overline{\lim}_n \mathbb{P}_n\left(\overline{A}\right) \leq \mathbb{P}\left(\overline{A}\right) = \mathbb{P}(A)\ , \tag{1.31}$$

$$\underline{\lim}_n \mathbb{P}_n(A) \geq \underline{\lim}_n \mathbb{P}_n\left(\mathring{A}\right) \geq \mathbb{P}\left(\mathring{A}\right) = \mathbb{P}(A)\ . \tag{1.32}$$

Quindi, $\forall A \in \mathcal{B}(E)$ tale che $\mathbb{P}(\partial A) = 0$, $\{\mathbb{P}_n(A)\}_{n\geq 1}$ converge a $\mathbb{P}(A)$.

4. $\Rightarrow$ 1. Sia $f \in C_b(E)$. Allora $\exists M > 0 : |f| \leq M$. Posto

$$D := \{t \in \mathbb{R} : \mathbb{P}\{x \in E : f(x) = t\} \neq 0\} , \tag{1.33}$$

si consideri la partizione $T_k := \{[t_0, t_1), .., [t_{tk-2}, t_{k-1}), [t_{k-1}, t_k]\}$ di $[-M, M]$, con $k \geq 1$, tale che $-M = t_0 < .. < t_k = M$, e $\forall i = 0, .., k$, $t_i \notin D$. Sia $B_i := \{x \in E : f(x) \in [t_i, t_{i+1})\}$; dalla continuità di f segue che $f^{-1}(t_i, t_{i+1})$ è aperto e $\partial B_i \subseteq f^{-1}\{t_i\} \cup f^{-1}\{t_{i+1}\}$. Siccome $t_i \notin D$, $\mathbb{P}(\partial B_i) = 0$, perciò dalla 4. segue che

$$\lim_{n\to\infty} \sum_{i=0}^{k-1} t_i \mathbb{P}_n(B_i) = \sum_{i=0}^{k-1} t_i \mathbb{P}(B_i). \tag{1.34}$$

Allora,

$$\begin{aligned} \left| \int_E \mathbb{P}_n(dx) f(x) - \int_E \mathbb{P}(dx) f(x) \right| &\leq \left| \int_E \mathbb{P}_n(dx) f(x) - \sum_{i=1}^{k-1} t_i \mathbb{P}_n(B_i) \right| + \\ + \left| \sum_{i=1}^{k-1} t_i \mathbb{P}_n(B_i) - \sum_{i=1}^{k-1} t_i \mathbb{P}(B_i) \right| &+ \left| \sum_{i=1}^{k-1} t_i \mathbb{P}(B_i) - \int_E \mathbb{P}(dx) f(x) \right| \leq \\ &\leq 2 \max_{0 \leq i \leq k-1} |t_{i+1} - t_i| + \left| \sum_{i=1}^{k-1} t_i \mathbb{P}_n(B_i) - \sum_{i=1}^{k-1} t_i \mathbb{P}(B_i) \right| \end{aligned}$$

e la tesi segue dall'arbitrarietà di T_k.

■

Teorema 31 *Sia E sia uno spazio metrico localmente compatto e σcompatto. Data una successione di misure di probabilità $\{\mathbb{P}_n\}_{n\geq 1}$, le seguenti affermazioni sono equivalenti:*

1. *$\forall f$ uniformemente continua e limitata su E, $\left\{\int_E d\mathbb{P}_n(x) f(x)\right\}_{n\geq 1}$ converge a $\int_E d\mathbb{P}(x) f(x)$;*
2. *$\forall f$ semicontinua superiormente e limitata su E, $\overline{\lim}_n \int_E d\mathbb{P}_n f(x) \leq \int_E d\mathbb{P} f(x)$;*
3. *$\forall f$ semicontinua inferiormente e limitata su E, $\underline{\lim}_n \int_E d\mathbb{P}_n f(x) \geq \int_E d\mathbb{P} f(x)$.*

Dimostrazione: Dal teorema di Baire sull'approssimazione di funzioni semicontinue per mezzo di una successione di funzioni continue si ha che 1. implica 2. e 3.. Cambiando di segno a f si ha che 2. implica 3. e viceversa. Poiché $\forall A \in \mathcal{B}(E), \mathbf{1}_A$ è semicontinua superiormente, ripetendo l'argomento 3. $\Longrightarrow$ 4. del teorema precedente si ottiene che $\{\mathbb{P}_n\}_{n\geq 1}$ converge in generale a $\mathbb{P}$, quindi debolmente e per il Teorema 19 anche strettamente, ma poiché ogni funzione uniformemente continua e limitata su E appartiene a $C_b(E)$ si ha che 3. $\Longrightarrow$ 1.. ■

Definizione 32 *Sia $(E, \mathcal{E})$ uno spazio misurabile. Una collezione d'insiemi $\mathcal{K}_0(E) \subseteq \mathcal{E}$ è detta* classe determinante *se, date due misure di probabilità $\mathbb{P}$ e $\mathbb{Q}$ su $(E, \mathcal{E})$ che soddisfano*

$$\mathbb{P}(A) = \mathbb{Q}(A) , \quad \forall A \in \mathcal{K}_0(E) \tag{1.35}$$

ne segue che sono identiche, ovvero $\mathbb{P}(A) = \mathbb{Q}(A), \forall A \in \mathcal{E}$.

Definizione 33 *Sia E uno spazio metrico. Una collezione d'insiemi $\mathcal{K}_1(E) \subseteq \mathcal{B}(E)$ è detta* classe determinante convergenza *se, data una successione di misure di probabilità $\{\mathbb{P}_n\}_{n\geq 1}$ convergente a $\mathbb{P}$ si ha che se $\{\mathbb{P}_n(A)\}_{n\geq 1}$ converge a $\mathbb{P}(A)$ per ogni $A \in \mathcal{K}_1(E)$ tale che $\mathbb{P}(\partial A) = 0$, allora $\{\mathbb{P}_n(A)\}_{n\geq 1}$ converge a $\mathbb{P}(A), \forall A \in \mathcal{B}(E)$ tale che $\mathbb{P}(\partial A) = 0$.*

Osservazione 34 *Nel caso in cui $E = \mathbb{R}^n$, $1 \leq n < \infty$, dalla definizione di funzione di distribuzione di un vettore aleatorio data nella sezione seguente, la collezione degli insiemi*

$$\mathcal{K} := \{\times_{i=1}^{n}(-\infty, x_i] : (x_1, .., x_n) \in \mathbb{R}^n\} \tag{1.36}$$

risulta essere una classe determinante e per il Teorema 131 è anche una classe determinante convergenza.

Nel caso in cui $E = \mathbb{R}^{\mathbb{N}}$, $\mathcal{K}_0(\mathbb{R}^{\mathbb{N}})$ e $\mathcal{K}_1(\mathbb{R}^{\mathbb{N}})$ coincidono con la collezione degli insiemi cilindrici, ovvero la collezione di sottoinsiemi A di $\mathbb{R}^{\mathbb{N}}$ per cui esistono $n \geq 1$ e $B \in \mathcal{B}(\mathbb{R}^n)$ $\mathcal{K}$ tali che, posto $x = \{x_i\}_{i\in\mathbb{N}}$, A coincide con $\{x \in \mathbb{R}^{\mathbb{N}} : (x_1, .., x_n) \in B\}$.

Nel caso in cui $E = C([0, 1])$ invece si può scegliere come $\mathcal{K}_0(C([0, 1]))$ la collezione degli insiemi cilindrici, tuttavia $\mathcal{K}_1(C([0, 1])) \supset \mathcal{K}_0(C([0, 1]))$. Infatti, se $\{\mathbb{P}_n\}_{n\geq 1} \subset \mathfrak{P}(C([0, 1]))$ è la successione tale che $\forall n \geq 1$, $\mathbb{P}_n := \delta_{x_n(t)}$ con

$$x_n(t) := nt\mathbf{1}_{[0,\frac{1}{n}]}(t) - (nt - 2)\mathbf{1}_{[\frac{1}{n},\frac{2}{n}]}(t) \tag{1.37}$$

e $\mathbb{P} := \delta_0$, allora, $\forall A \in \mathcal{K}_0(C([0, 1])), n \geq 1, \mathbb{P}_n(A) = 0$ o 1. Inoltre se $\mathbb{P}(\partial A) = 0$, $\{\mathbb{P}_n(A)\}_{n\geq 1}$ converge a $\mathbb{P}(A)$, mentre se

$$B := \{x(t) \in C([0, 1]) : |x(t)| \leq \frac{1}{2}\}, \tag{1.38}$$

$B \in \mathcal{B}(C([0, 1]))$ che coincide con la σalgebra generata dagli elementi di $\mathcal{K}_0(C([0, 1]))$ e $\mathbb{P}(\partial B) = \mathbb{P}_n(B) = 0, \forall n \geq 1$, ma $\mathbb{P}(B) = 1$. Quindi, $\{\mathbb{P}_n(B)\}_{n\geq 1}$ non converge a $\mathbb{P}(B)$.

Definizione 35 *Sia E sia uno spazio topologico polacco e $\mathcal{I}$ un insieme di indici. Una famiglia di misure di probabilità $\{\mathbb{P}_i\}_{i\in\mathcal{I}} \subseteq \mathfrak{P}(E)$ è detta* relativamente compatta *se ogni successione $\{\mathbb{P}_n\}_{n\geq 1} \subseteq \{\mathbb{P}_i\}_{i\in\mathcal{I}}$ contiene una sottosuccessione debolmente convergente ad un elemento di $\mathfrak{P}(E)$.*

Teorema 36 *(di Prokhorov) Sia E sia uno spazio topologico polacco e $\mathcal{I}$ un insieme di indici. Una famiglia di misure di probabilità $\{\mathbb{P}_i\}_{i\in\mathcal{I}} \subseteq \mathfrak{P}(E)$ è relativamente compatta se e solo se è tesa.*

Dimostrazione: Ci restringiamo al caso in cui $E = \mathbb{R}$ che è anche localmente compatto e σcompatto.

⇒ Se $\{\mathbb{P}_i\}_{i\in\mathcal{I}}$ è relativamente compatta ma non tesa, $\exists\varepsilon > 0$ tale che $\forall K \subset \mathbb{R}$ compatto $\sup_{i\in\mathcal{I}} \mathbb{P}_i (K^c) > \varepsilon$. Da ciò segue che, $\forall n \geq 1$,

$$\sup_{i\in\mathcal{I}} \mathbb{P}_i ((-n,n)^c) \geq \sup_{i\in\mathcal{I}} \mathbb{P}_i ([-n,n]^c) > \varepsilon \tag{1.39}$$

e quindi che $\forall n \geq 1$, $\exists\mathbb{P}_{i_n} \in \{\mathbb{P}_i\}_{i\in\mathcal{I}}$ per cui $\mathbb{P}_{i_n} ((-n,n)^c) > \varepsilon$. Dato che $\{\mathbb{P}_i\}_{i\in\mathcal{I}}$ è relativamente compatta, da $\{\mathbb{P}_{i_n}\}_{n\geq 1}$ si può estrarre una sottosuccessione $\{\mathbb{P}_{i_{n_k}}\}_{k\geq 1}$ debolmente convergente a $\mathbb{Q} \in \mathfrak{P}(\mathbb{R})$. Allora, per il Teorema 30, $\forall n \geq 1$,

$$0 < \varepsilon < \overline{\lim}_k \mathbb{P}_{i_{n_k}} ((-n,n)^c) \leq \mathbb{Q}((-n,n)^c) \ , \tag{1.40}$$

ma poiché $\lim_{n\to\infty} \mathbb{Q}((-n,n)^c) = \mathbb{Q}(\cap_{n\geq 1} (-n,n)^c) = 0$ ciò è impossibile.

⇐ $\{\mathbb{P}_i\}_{i\in\mathcal{I}}$ è tesa e $\{\mathbb{P}_n\}_{n\geq 1} \subset \{\mathbb{P}_i\}_{i\in\mathcal{I}}$. Sia $\{F_n\}_{n\geq 1}$ la corrispondente successione di funzioni di distribuzione (la definizione di funzione di distribuzione associata ad una misura di probabilità su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ è data nella (1.42)). Per il Teorema di Helly (Teorema 42) $\exists\{F_{n_k}\}_{k\geq 1} \subseteq \{F_n\}_{n\geq 1}$ ed una funzione di distribuzione generalizzata G (cf. Definizione 40) tale che $\{F_{n_k}(x)\}_{k\geq 1}$ converge a $G(x)$, $\forall x \in \mathcal{C}_G$. Scelto $\varepsilon > 0$, sia $I_\varepsilon = (a_\varepsilon, b_\varepsilon] \subset \mathbb{R}$ tale che $\forall n \geq 1$, $\mathbb{P}_n (I_\varepsilon^c) < \varepsilon$, ovvero tale che $\mathbb{P}_n (I_\varepsilon) > 1 - \varepsilon$. Se $a, b \in \mathcal{C}_G$ tali che $a < a_\varepsilon < b_\varepsilon < b$, allora

$$1 - \varepsilon < \mathbb{P}_{n_k}(a_\varepsilon, b_\varepsilon] \leq \mathbb{P}_{n_k}(a,b] = F_{n_k}(b) - F_{n_k}(a) \underset{k\to\infty}{\longrightarrow} G(b) - G(a) \ . \tag{1.41}$$

Inoltre, $\forall k \geq 1$, $\mathbb{P}_{n_k}(-\infty, +\infty) = 1$ da cui segue, passando al limite per $k \to \infty$ che $G(+\infty) - G(-\infty) = 1$. Siccome però G è una funzione di distribuzione generalizzata, $0 \leq G(-\infty) \leq G(+\infty) \leq 1$, quindi $G(-\infty) = 0$ e $G(+\infty) = 1$, ovvero G è una funzione di distribuzione. Perciò, dalla Definizione 43 segue che $\{F_{n_k}\}_{k\geq 1}$ converge in generale a G e quindi dal Teorema 131 segue che la successione associata di misure di probabilità $\{\mathbb{P}_{n_k}\}_{k\geq 1}$ converge debolmente a $\mathbb{Q}$, distribuzione di probabilità associata a G.

■

Funzioni di distribuzione e convergenza in legge

Sia $(E, \mathcal{B}(E)) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ e $\mathbb{P} \in \mathfrak{P}(\mathbb{R})$. La funzione

$$\mathbb{R} \ni x \longmapsto F(x) := \mathbb{P}(-\infty, x] \in [0,1] \tag{1.42}$$

ha le seguenti proprietà:

1. F è non decrescente, ovvero, $\forall x, y \in \mathbb{R} : x \leq y$, $F(x) \leq F(y)$;

2. $\lim_{x\downarrow-\infty} F(x) = 0\,;\ \lim_{x\uparrow\infty} F(x) = 1;$

3. F è continua a destra, ovvero $\forall x \in \mathbb{R}$, $\lim_{y\downarrow x} F(y) = F(x)$, ed ammette limite a sinistra, ovvero $\forall x \in \mathbb{R}$, $\lim_{y\uparrow x} F(y) = F(x^-)$.

Definizione 37 *Ogni funzione $F : \mathbb{R} \longrightarrow \mathbb{R}$ che verifica le proprietà 1,2,3 precedentemente descritte è detta* funzione di distribuzione.

Teorema 38 *(di Carathéodory) Sia Ω un insieme, $\mathcal{A}$ un'algebra di Boole di suoi sottoinsiemi e sia μ_0 una misura σfinita su $(\Omega, \mathcal{A})$. Allora, esiste un'unica misura $\mu \in \mathfrak{P}(\Omega, \sigma(\mathcal{A}))$ estensione di μ_0, ovvero tale che $\forall A \in \mathcal{A}$, $\mu(A) = \mu_0(A)$.*

Teorema 39 *Sia F una funzione di distribuzione. Esiste un'unica misura di probabilità $\mathbb{P}$ su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ tale che $\forall a, b \in \mathbb{R} : a < b$, $\mathbb{P}(a, b] = F(b) - F(a)$.*

Dimostrazione: Sia $\mathcal{A}$ l'algebra dei sottoinsiemi di $\mathbb{R}$ generata da insiemi $A \subset \mathbb{R}$ che sono unioni disgiunte di intervalli del tipo $(a, b]$ con $a, b \in \mathbb{R} : a < b$, ovvero

$$A := \sum_{k=1}^{n} (a_k, b_k], \tag{1.43}$$

e sia $\mathbb{P}_0$ tale che $\forall A \in \mathcal{A}$

$$\mathbb{P}_0(A) := \sum_{k=1}^{n} F(b_k) - F(a_k)\,. \tag{1.44}$$

$\mathbb{P}_0$ così definita è unica e finitamente additiva. Resta da dimostrarne la σadditività e la tesi segue dal teorema di Carathéodory. A tal fine basta mostrare che $\mathbb{P}_0$ è continua a $\emptyset$, ovvero, che se $\{A_n\}_{n\geq1}$ è una successione decrescente a $\emptyset$ di insiemi di $\mathcal{A}$, cioè tale che $A_{n+1} \subseteq A_n$ e $A \downarrow \emptyset$, allora $\mathbb{P}_0(A_n) \downarrow \emptyset$. Supponiamo innanzi tutto che $\forall n \geq 1$, $A_n \in [-N, N]$ con $N < \infty$. Dato che $F(x)$ è continua a destra ed ogni A_n è somma finita d'intervalli aperti a sinistra e chiusi a destra, $\forall n \geq 1, \varepsilon > 0$, $\exists B_n \in \mathcal{A}$ tale che $\overline{B_n} \subseteq A_n$ e $\mathbb{P}_0(A_n) - \mathbb{P}_0(B_n) \leq \varepsilon 2^{-n}$. Siccome $[-N, N]$ è compatto e $\bigcap_{n\geq1} A_n = \bigcap_{n\geq1} \overline{B_n} = \emptyset$, $\{[-N, N] \backslash \overline{B_n}\}_{n\geq1}$ è un ricoprimento aperto di $[-N, N]$. Allora, per il teorema di Heine-Borel si può estrarre un sottoricoprimento finito. Perciò, $\exists n_0(\varepsilon) > 0$ tale che

$$[-N, N] = \bigcup_{n=1}^{n_0(\varepsilon)} \left([-N, N] \backslash \overline{B_n}\right) = [-N, N] \backslash \bigcap_{n=1}^{n_0(\varepsilon)} \overline{B_n}\,, \tag{1.45}$$

ovvero $\bigcap_{n=1}^{n_0(\varepsilon)} \overline{B_n} = \emptyset$. Quindi, siccome $A_1 \supseteq \cdots \supseteq A_{n_0(\varepsilon)}$,

$$\mathbb{P}_0\left(A_{n_0(\varepsilon)}\right) = \mathbb{P}_0\left(A_{n_0(\varepsilon)} \backslash \bigcap_{n=1}^{n_0(\varepsilon)} B_n\right) + \mathbb{P}_0\left(\bigcap_{n=1}^{n_0(\varepsilon)} B_n\right) \tag{1.46}$$

$$= \mathbb{P}_0\left(\bigcup_{n=1}^{n_0(\varepsilon)} \left(A_{n_0(\varepsilon)} \backslash B_n\right)\right) \leq \mathbb{P}_0\left(\bigcup_{n=1}^{n_0(\varepsilon)} (A_n \backslash B_n)\right) \tag{1.47}$$

$$\leq \sum_{n=1}^{n_0(\varepsilon)} \mathbb{P}_0(A_n \backslash B_n) \leq \varepsilon \sum_{n=1}^{n_0(\varepsilon)} 2^{-n} = \varepsilon\left(1 - 2^{-(n_0(\varepsilon)+1)}\right) < \varepsilon\,, \tag{1.48}$$

ovvero $\lim_{n\to\infty} \mathbb{P}_0(A_n) = 0$. In generale è sufficiente scegliere $\forall\varepsilon > 0$ un $N(\varepsilon) > 0$ tale che $\mathbb{P}_0[-N(\varepsilon), N(\varepsilon)] > 1-\varepsilon$. Allora, poiché

$$\begin{aligned} A_n &= ([-N(\varepsilon), N(\varepsilon)] \cap A_n) \cup ([-N(\varepsilon), N(\varepsilon)]^c \cap A_n), \\ \mathbb{P}_0(A_n) &\leq \mathbb{P}_0([-N(\varepsilon), N(\varepsilon)] \cap A_n) + \varepsilon. \end{aligned}$$

Perciò, ripetendo lo stesso argomento precedente esposto per A_n per $([-N(\varepsilon), N(\varepsilon)] \cap A_n)$, per n sufficientemente grande, si ottiene $\mathbb{P}_0([-N(\varepsilon), N(\varepsilon)] \cap A_n) < \varepsilon$ da cui la tesi. ■

Definizione 40 *Una funzione $G : \mathbb{R} \longrightarrow \mathbb{R}$ tale che:*

1. *G è non decrescente, ovvero, $\forall x, y \in \mathbb{R} : x \leq y$, $G(x) \leq G(y)$;*
2. $\lim_{x\downarrow-\infty} G(x) \geq 0$; $\lim_{x\uparrow\infty} G(x) \leq 1$;
3. *G è continua a destra, ovvero $\forall x \in \mathbb{R}$, $\lim_{y\downarrow x} G(y) = G(x)$.*

è detta funzione di distribuzione generalizzata*.*

Ovviamente dalla definizione di funzione di distribuzione si ha che questa classe è contenuta in quella delle funzioni di distribuzione generalizzate.

Definizione 41 *Sia $\mathcal{I}$ un insieme di indici. Una famiglia di funzioni di distribuzione generalizzate $\{G_i\}_{i\in\mathcal{I}}$ su $\mathbb{R}$, è detta* relativamente compatta *se, $\forall\{G_n\}_{n\geq1} \subset \{G_i\}_{i\in\mathcal{I}}$, $\exists G \in \{G_i\}_{i\in\mathcal{I}}$ ed una sottosuccessione $\{n_k\}_{k\geq1} \subset \{n\}$ tale che, $\forall x \in \mathcal{C}_G$, insieme dei punti di continuità di G, $\lim_{k\to\infty} G_{n_k}(x) = G(x)$.*

Teorema 42 *(di Helly) La classe delle funzioni di distribuzione generalizzate è sequenzialmente compatta.*

Dimostrazione: Sia $\{G_n\}_{n\geq1}$ una successione di funzioni generalizzate e $T := \{x_1, x_2, ..\} \subset \mathbb{R}$ un sottoinsieme numerabile denso in $\mathbb{R}$. Dalla limitatezza della successione numerica $\{G_n(x_1)\}_{n\geq1}$ segue che esiste $\{n_k^{(1)}\}_{k\geq1} \subset \{n\}$ tale che $\{G_{n_k^{(1)}}(x_1)\}_{k\geq1}$ ha limite g_1 per $k \to \infty$. Allora, è possibile estrarre da $\{n_k^{(1)}\}_{k\geq1}$ una sottosuccessione $\{n_k^{(2)}\}_{k\geq1}$ tale che $\{G_{n_k^{(2)}}(x_2)\}_{k\geq1}$ ha limite g_2 per $k \to \infty$ e così via. In questo modo resta definita su T una funzione $G_T(x)$ tale che $G_T(x_i) = g_i$ e, considerando la successione diagonale $\{n_k^{(k)}\}_{k\geq1}$, $\forall x_i \in T$, si ha $\lim_{k\to\infty} G_{n_k^{(k)}}(x_i) = g_i$. Sia allora

$$\mathbb{R} \ni x \longmapsto G(x) := \inf\{G_T(y) : y \in T,\, y > x\}. \tag{1.49}$$

- *G è non decrescente.* Infatti, dato che gli elementi della successione $\{G_n\}_{n\geq1}$ sono funzioni non decrescenti, si ha che $G_{n_k^{(k)}}(x) \leq G_{n_k^{(k)}}(y)$, $\forall x, y \in T : x \leq y$ e quindi, passando al limite per $k \to \infty$, ciò vale anche per G_T. Dunque, dalla sua definizione si ha che G è non decrescente.

- *G è continua a destra.* Sia $\{x_k\}_{k\geq 1} \downarrow x$ e $d := \lim_{k\to\infty} G(x_k)$. Chiaramente, $G(x) \leq d$, ma se fosse $G(x) < d$, per la definizione di G esisterebbe $y \in T : x < y,\, G(y) < d$. Quindi, poiché $\{x_k\}_{k\geq 1} \downarrow x$, esisterebbe $k(y) : \forall k > k(y),\, x < x_k < y$, dunque $G(x_k) \leq G(y) < d$ che implica $\lim_{k\to\infty} G(x_k) < d$, contro l'ipotesi che tale limite è proprio uguale a d.

Pertanto G è una funzione di distribuzione generalizzata. Resta da dimostrare che, $\forall x_0 \in \mathcal{C}_G$, $\lim_{m\to\infty} G_{n_m^{(m)}}(x_0) = G(x_0)$.
Infatti, $\forall y \in T : y > x_0$,

$$\overline{\lim}_m G_{n_m^{(m)}}(x_0) \leq \overline{\lim}_m G_{n_m^{(m)}}(y) = G_T(y) \,, \tag{1.50}$$

quindi

$$\overline{\lim}_m G_{n_m^{(m)}}(x_0) \leq \inf\{G_T(y) : y \in T,\, y > x_0\} = G(x_0) \,. \tag{1.51}$$

D'altra parte sia $x' < y < x_0$ con $y \in T$. Allora,

$$G(x') \leq G(y) = G_T(y) = \lim_{m\to\infty} G_{n_m^{(m)}}(y) = \underline{\lim}_m G_{n_m^{(m)}}(y) \leq \underline{\lim}_m G_{n_m^{(m)}}(x_0) \,, \tag{1.52}$$

ma se $x' \uparrow x_0$ si ha che $G\left(x_0^-\right) \leq \underline{\lim}_m G_{n_m^{(m)}}(x_0)$. Pertanto, se $x_0 \in \mathcal{C}_G$, $G\left(x_0^-\right) = G(x_0)$, dunque

$$G(x_0) \leq \underline{\lim}_m G_{n_m^{(m)}}(x_0) \leq \lim_{m\to\infty} G_{n_m^{(m)}}(x_0) \leq \overline{\lim}_m G_{n_m^{(m)}}(x_0) = G(x_0) \,. \tag{1.53}$$

■

Definizione 43 *Una successione di funzioni di distribuzione $\{F_n\}_{n\geq 1}$ definite su $\mathbb{R}$ converge* in generale *alla funzione di distribuzione $F : \mathbb{R} \longrightarrow \mathbb{R}$ se, $\forall x \in \mathcal{C}_F \subset \mathbb{R}$, insieme dei punti di continuità di F, la successione numerica $\{F_n(x)\}_{n\geq 1}$ converge ad $F(x)$.*

In generale, sia $n \geq 1$. Consideriamo lo spazio misurabile $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, se $\mathfrak{P}(\mathbb{R}^n)$ indica la collezione delle misure di probabilità su $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, $\forall \mathbb{P}^n \in \mathfrak{P}(\mathbb{R}^n)$. La funzione

$$\mathbb{R}^n \ni x = (x_1, .., x_n) \longmapsto F(x) = F(x_1, .., x_n) := \mathbb{P}\left(\bigcap_{i=1}^n \{y \in \mathbb{R}^n : y_i \leq x_i\}\right) \in [0, 1] \tag{1.54}$$

ha le seguenti proprietà:

1. $\lim_{x_i\downarrow -\infty} F(x_1, .., x_n) = 0 \,\forall i = 1, .., n$;
2. $\lim_{x\uparrow\infty} F(x) = 1$;
3. F è continua a destra in ogni argomento, ovvero

$$\forall i = 1, .., n,\ \lim_{y_i\downarrow x_i} F(x_1, .., x_{i-1}, y, x_{i+1}, .., x_n) = F(x_1, .., x_{i-1}, x_i, x_{i+1}, .., x_n) \tag{1.55}$$

   e quindi collettivamente, cioè $\forall x \in \mathbb{R}^n$, $\lim_{y\downarrow x} F(y) = F(x)$.

Definizione 44 *Ogni funzione $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ che verifica le proprietà 1,2,3 precedentemente descritte è detta* funzione di distribuzione multivariata n-dimensionale. *Nel caso unidimensionale F è detta funzione di distribuzione* univariata.

Se F su $\mathbb{R}^n$, $n \geq 2$,è una funzione di distribuzione, scelto $i = 1, .., n$, passando al limite per $x_i \rightarrow +\infty$ nell'espressione di $F(x_1, .., x_n)$ si ottiene una funzione di distribuzione $F_{(i)} : \mathbb{R}^{n-1} \longrightarrow \mathbb{R}$, per cui, se $y \in \mathbb{R}^{n-1}$ è tale che $y_j = x_j$, con $j \in \{1, .., n\} \backslash \{i\}$,

$$F_{(i)}(y_1, .., y_{n-1}) = \lim_{x_i \rightarrow +\infty} F(x_1, .., x_n) \ . \tag{1.56}$$

Nel linguaggio del Calcolo delle Probabilità e della Statistica, le $F_{(i)}$ sono dette funzioni di distribuzione marginali $(n-1)$-dimensionali *della funzione di distribuzione F.*

Osservazione 45 *Allo stesso modo, scelto un sottoinsieme $\{i_1, .., i_k\}$ di $k \leq n-1$ elementi in $\{1, .., n\}$ la funzione $F_{(i_1, .., i_k)} : \mathbb{R}^{n-k} \longrightarrow \mathbb{R}$ per cui, se $y \in \mathbb{R}^{n-k}$ è tale che $y_j = x_j$, con $j \in \{1, .., n\} \backslash \{i_1, .., i_k\}$,*

$$F_{(i_1, .., i_k)}(y_1, .., y_{n-k}) = \lim_{\left(x_{i_1}, .., x_{i_k}\right) \rightarrow (+\infty, .., +\infty)} F(x_1, .., x_n) \ , \tag{1.57}$$

è detta funzione di distribuzione marginale $(n-1)$-dimensionale *della funzione di distribuzione F.*

Analogamente al caso delle funzioni di distribuzione univariate, il Teorema di Carathéodory implica il seguente risultato

Teorema 46 *Sia F una funzione di distribuzione multivariata n-dimensionale. Allora esiste un unica misura di probabilità $\mathbb{P}$ su $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ tale che se $\forall i = 1, .., n$*

$$\Delta_{a_i, b_i} F_n(x_1, .., x_n) := F_n(x_1, .., x_{i-1}, b_i, x_{i+1}, ..x_n) - F_n(x_1, .., x_{i-1}, a_i, x_{i+1}, ..x_n) \ , \tag{1.58}$$

$$\mathbb{P}\left(\times_{i=1}^n (a_i, b_i]\right) = \prod_{i=1}^n \Delta_{a_i, b_i} F_n(x_1, .., x_n) \ . \tag{1.59}$$

Nel seguito, considerato lo spazio di probabilità $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathbb{P})$, $n \geq 1$, per ogni funzione f a valori reali misurabile su $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ tale che

$$\left(\int_{\mathbb{R}^n} \mathbb{P}(dx)(f(x) \vee 0)\right) \wedge \left(\int_{\mathbb{R}^n} \mathbb{P}(dx)(f(x) \wedge 0)\right) < \infty \, , \tag{1.60}$$

porremo

$$\int_{\mathbb{R}^n} \mathbb{P}(dx) f(x) = \int_{\mathbb{R}^n} dF(x) f(x) \ . \tag{1.61}$$

Caratterizzazione della convergenza debole tramite la trasformata di Fourier-Stieltjes della distribuzione di probabilità

Funzione caratteristica di una distribuzione di probabilità su $\mathbb{R}^n$, $n \geq 1$

Definizione 47 *Si definisce* funzione caratteristica *di una distribuzione di probabilità* $\mathbb{P} \in \mathfrak{P}(\mathbb{R}^n)$ *la funzione*

$$\mathbb{R}^n \ni t \longmapsto \varphi(t) := \int_{\mathbb{R}^n} \mathbb{P}(dx)\, e^{i(t,x)} = \int_{\mathbb{R}^n} dF(x)\, e^{i(t,x)} \in \mathbb{C}\,. \tag{1.62}$$

Dunque φ è la trasformata di Fourier-Stieltjes della funzione di distribuzione F che coincide con la trasformata di Fourier della densità di probabilità associata, se tale distribuzione di probabilità è assolutamente continua rispetto alla misura di Lebesgue su $\mathbb{R}^n$.

Qui di seguito ci restringeremo al caso delle funzioni di distribuzione univariate, tuttavia quanto esposto resta valido anche nel caso delle funzioni di distribuzione multivariate.

Teorema 48 *(di unicità) Siano F e G due funzioni di distribuzione su $\mathbb{R}$ tali che le funzioni carattesistiche delle distribuzioni di probabilità associate coincidono $\forall t \in \mathbb{R}$. Allora anche F e G coincidono su tutto $\mathbb{R}$.*

Dimostrazione: Sia I un intervallo di $\mathbb{R}$ e $\forall \varepsilon > 0$, sia

$$f^{(\varepsilon)}(x) := \left(1 - \frac{\inf_{y\in I}|x-y|}{\varepsilon}\right) \vee 0\,. \tag{1.63}$$

$f^{(\varepsilon)}$ è continua e $\left|\left|f^{(\varepsilon)}\right|\right|_{\infty} \leq 1$. Inoltre, $\forall \varepsilon > 0$, $\exists N_{\varepsilon} > 0 : \forall N > N_{\varepsilon}$, $I \subset supp f^{(\varepsilon)} \subseteq [-N, N]$. Poiché $f^{(\varepsilon)}(-N) = f^{(\varepsilon)}(N) = 0$, per il teorema di Stone-Weierstrass, $f^{(\varepsilon)}$ può essere approssimata.uniformemente per mezzo di polinomi trigonometrici. Ovvero, esiste:

- una successione di funzioni $\{f_{N,n}^{(\varepsilon)}\}_{n\geq 1}$ della forma $f_{N,n}^{(\varepsilon)}(x) = \sum_{k=0}^{n} a_k e^{i\pi x \frac{k}{N}}$;
- un indice $n(N) \geq 1$ e una successione $\{\delta_N\}_{N\geq 1}$ tale che $\delta_N \downarrow 0$,

per cui, per ogni $\forall n \geq n(N)$

$$\sup_{x\in[-N,N]} \left|f_{N,n}^{(\varepsilon)}(x) - f^{(\varepsilon)}(x)\right| < \delta_N\,; \quad \sup_{x\in\mathbb{R}} \left|f_{N,n}^{(\varepsilon)}(x)\right| < 2\,. \tag{1.64}$$

Allora. poiché per ipotesi

$$\int_{\mathbb{R}} dF(x)\, e^{itx} = \int_{\mathbb{R}} dG(x)\, e^{itx} \tag{1.65}$$

$\forall \varepsilon > 0$,

$$\int_{\mathbb{R}} dF(x)\, f_{N,n}^{(\varepsilon)}(x) = \int_{\mathbb{R}} dG(x)\, f_{N,n}^{(\varepsilon)}(x)\ . \tag{1.66}$$

Dunque,

$$\left|\int_{\mathbb{R}} dF\left(x\right) f^{(\varepsilon)}\left(x\right) - \int_{\mathbb{R}} dG\left(x\right) f^{(\varepsilon)}\left(x\right)\right| = \left|\int_{-N}^{N} dF\left(x\right) f^{(\varepsilon)}\left(x\right) - \int_{-N}^{N} dG\left(x\right) f^{(\varepsilon)}\left(x\right)\right| \tag{1.67}$$

$$\leq \int_{-N}^{N} \left|f^{(\varepsilon)}\left(x\right) - f_{N,n}^{(\varepsilon)}\left(x\right)\right| dF\left(x\right) + \int_{-N}^{N} \left|f^{(\varepsilon)}\left(x\right) - f_{N,n}^{(\varepsilon)}\left(x\right)\right| dG\left(x\right) + \left|\int_{-N}^{N} dF\left(x\right) f_{N,n}^{(\varepsilon)}\left(x\right) - \int_{-N}^{N} dG\left(x\right) f_{N,n}^{(\varepsilon)} (\right.$$

$$\leq 2\delta_N + \left|\int_{\mathbb{R}} dF\left(x\right) f_{N,n}^{(\varepsilon)}\left(x\right)\left(1 - \mathbf{1}_{[-N,N]^c}\left(x\right)\right) - \int_{\mathbb{R}} dG\left(x\right) f_{N,n}^{(\varepsilon)}\left(x\right)\left(1 - \mathbf{1}_{[-N,N]^c}\left(x\right)\right)\right|$$

$$= 2\delta_N + \left|\int_{\mathbb{R}} dG\left(x\right) f_{N,n}^{(\varepsilon)}\left(x\right) \mathbf{1}_{[-N,N]^c}\left(x\right) - \int_{\mathbb{R}} dF\left(x\right) f_{N,n}^{(\varepsilon)}\left(x\right) \mathbf{1}_{[-N,N]^c}\left(x\right)\right|$$

$$\leq 2\delta_N + 2\left(1 - F\left(N\right) + F\left(-N\right)\right) + 2\left(1 - G\left(N\right) + G\left(-N\right)\right) .$$

Passando al limite per $N \to \infty$ si ha

$$\int_{\mathbb{R}} dF\left(x\right) f^{(\varepsilon)}\left(x\right) = \int_{\mathbb{R}} dG\left(x\right) f^{(\varepsilon)}\left(x\right) \tag{1.68}$$

che, per $\varepsilon \to 0$ implica $\int_I dF = \int_I dG$ per ogni intervallo $I \subset \mathbb{R}$ e quindi $\forall A \in \mathcal{B}\left(\mathbb{R}\right)$. ■

Teorema 49 *(formula d'inversione) Sia F una funzione di distribuzione su $\mathbb{R}$ e φ la funzione caratteristica della distribuzione di probabilità associata. Allora:*

1. *$\forall a, b \in \mathbb{R}$ tale che $a < b$ e F è continua in a e b,*

$$F\left(b\right) - F\left(a\right) = \lim_{c\to\infty} \frac{1}{2\pi}\int_{-c}^{c} dt \frac{e^{-ita} - e^{-itb}}{it}\varphi\left(t\right) . \tag{1.69}$$

2. *Se $\varphi \in L^1\left(\mathbb{R}\right)$, F ha densità*

$$f\left(x\right) = \frac{1}{2\pi}\int_{\mathbb{R}} dt e^{-itx}\varphi\left(t\right) . \tag{1.70}$$

Dimostrazione:

1. Sia

$$\frac{1}{2\pi}\int_{-c}^{c} dt \frac{e^{-ita} - e^{-itb}}{it}\varphi\left(t\right) = \frac{1}{2\pi}\int_{-c}^{c} dt \frac{e^{-ita} - e^{-itb}}{it}\int_{\mathbb{R}} dF\left(x\right) e^{itx} = \tag{1.71}$$

$$\int_{\mathbb{R}} dF\left(x\right) \frac{1}{2\pi}\int_{-c}^{c} dt \frac{e^{-ita} - e^{-itb}}{it} e^{itx} = \int_{\mathbb{R}} dF\left(x\right) \psi_c\left(x\right) .$$

Poiché

$$\left|\frac{e^{-ita} - e^{-itb}}{it} e^{itx}\right| = \left|\int_a^b dx e^{-itx}\right| \leq b - a , \tag{1.72}$$

$$\left|\psi_c\left(x\right)\right| = \left|\frac{1}{2\pi}\int_{-c}^{c} dt \frac{e^{-ita} - e^{-itb}}{it} e^{itx}\right| \leq \frac{c\left(b-a\right)}{\pi} < \infty .$$

Inoltre, $\frac{1}{t}$ è una funzione dispari, quindi

$$\psi_c(x) = \frac{1}{2\pi}\int_{-c}^{c} dt \frac{\sin t(x-a) - \sin t(x-b)}{t} = \frac{1}{2\pi}\int_{-c(x-a)}^{c(x-a)} du \frac{\sin u}{u} - \frac{1}{2\pi}\int_{-c(x-b)}^{c(x-b)} du \frac{\sin u}{u} . \tag{1.73}$$

Ma la funzione $\int_x^y du \frac{\sin u}{u}$ è continua in x e y e $\int_{\mathbb{R}} du \frac{\sin u}{u} = \pi$. Quindi, uniformemente in $x \in \mathbb{R}$

$$\lim_{c\to\infty} \psi_c = \mathbf{1}_{(a,b)} + \frac{1}{2}\left(\delta_{\{a\}} + \delta_{\{b\}}\right) . \tag{1.74}$$

Sia ora $\mathbb{P}$ la misura di probabilità su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ associata a F, ovvero tale che $\forall a, b \in \mathbb{R}$ con $a < b$, $\mathbb{P}(a,b] = F(b) - F(a)$. Per il teorema di Lebesgue della convergenza dominata,

$$\begin{aligned} \lim_{c\to\infty}\int_{\mathbb{R}} dF(x)\,\psi_c(x) &= \int_{\mathbb{R}} dF(x)\,\psi_c(x) = \mathbb{P}(a,b) + \frac{1}{2}\left(\mathbb{P}\{a\} + \mathbb{P}\{b\}\right) \\ &= F\left(b^-\right) - F(a) + \frac{1}{2}\left(F(a) - F\left(a^-\right)\right) + \frac{1}{2}\left(F(b) - F\left(b^-\right)\right) \\ &= \frac{F(b) + F(b^-)}{2} - \frac{F(a) + F(a^-)}{2} . \end{aligned} \tag{1.75}$$

Ma se a e b sono punti di continuità di F, allora

$$\lim_{c\to\infty}\int_{\mathbb{R}} dF(x)\,\psi_c(x) = F(b) - F(a) .$$

2. Se $\varphi \in L^1(\mathbb{R})$, sia $f(x)$ la sua antitrasfomata di Fourier. Allora, f è uniformemente continua su $\mathbb{R}$ e $\lim_{x\to+\infty} f(x) = \lim_{x\to-\infty} f(x) = 0$. Dunque, f è integrabile su un intervallo finito di $\mathbb{R}$. Quindi, se $-\infty < a < b < +\infty$, per il teorema di Fubini,

$$\begin{aligned} \int_a^b dx f(x) &= \int_a^b dx \frac{1}{2\pi}\int_{\mathbb{R}} dt e^{-itx}\varphi(t) = \int_{\mathbb{R}} \frac{dt}{2\pi}\varphi(t)\int_a^b dx e^{-itx} \\ &= \lim_{c\to\infty}\int_{-c}^{c} dt \frac{1}{2\pi}\frac{e^{-ita} - e^{-itb}}{it}\varphi(t) . \end{aligned} \tag{1.76}$$

Se

$$F(b) - F(a) = \lim_{c\to\infty}\frac{1}{2\pi}\int_{-c}^{c} dt \frac{e^{-ita} - e^{-itb}}{it}\varphi(t) \tag{1.77}$$

e a e b sono punti di continuità di F, allora, $\forall x \in \mathbb{R}$,

$$F(x) = \int_{-\infty}^{x} dy f(y) \tag{1.78}$$

e poiché F è non decrescente e $\lim_{x\to\infty} F(x) = 1$, f è sommabile, quindi la distribuzione di probabilità associata ad F è a.c. rispetto alla misura di Lebesgue. Dunque, la seconda affermazione del teorema segue dalla prima.

■

Osservazione 50 *La formula d'inversione provvede ad un'altra dimostrazione del teorema di unicità.*

Teorema 51 *(di Bochner-Khinchin) Sia $\varphi \in C(\mathbb{R};\mathbb{C})$ tale che $\varphi(0) = 1$. Condizione necessaria e sufficiente affinche φ sia la funzione caratteristica di una distribuzione di probabilità su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ è che la forma quadratica associata, ovvero*

$$\sum_{i,j=1}^{n} \varphi(t_i - t_j)\, \lambda_i \bar{\lambda}_j \,, \quad \forall n \geq 1, \ e\ \{t_i\}_{i=1}^{n} \subset \mathbb{R},\ \{\lambda_i\}_{i=1}^{n} \subset \mathbb{C}\,, \tag{1.79}$$

sia definita positiva.

Dimostrazione: Essendo la dimostrazione della sufficienza particolarmente complicata, dimostriamo solo l'implicazione diretta.

Se $\varphi(t) = \int_{\mathbb{R}} dF(x)\, e^{itx}$ allora, $\forall n \geq 1$, e $\{t_i\}_{i=1}^{n} \subset \mathbb{R}$, $\{\lambda_i\}_{i=1}^{n} \subset \mathbb{C}$,

$$\sum_{i,j=1}^{n} \int_{\mathbb{R}} dF(x)\, e^{ix(t_i - t_j)} \lambda_i \bar{\lambda}_j = \int_{\mathbb{R}} dF(x) \left| \sum_{i=1}^{n} e^{it_i x} \lambda_i \right|^2 \geq 0\,. \tag{1.80}$$

■

Teorema 52 *(di Pólya) Se $\varphi \in C(\mathbb{R})$ è una funzione pari tale che:*

1. $\varphi(t) \geq 0$, $\forall t \in \mathbb{R}^+$;
2. $\varphi(0) = 1$;
3. $\lim_{t\to\infty} \varphi(t) = 0$;
4. *φ è convessa su $\mathbb{R}^+$,*

allora φ è una funzione caratteristica.

Teorema 53 *(di Marcinkiewicz) Se $\mathbb{R} \ni t \longmapsto P(t) \in \mathbb{C}$ è un polinomio e $\varphi = e^P$ una funzione caratteristica, allora P ha al più grado 2.*

Convergenza debole di una successione di misure di probabilità su $\mathbb{R}$ e convergenza puntuale della successione associata di funzioni caratteristiche

Lemma 54 *Sia* $\{\mathbb{P}_n\}_{n\geq 1} \subset \mathfrak{P}(\mathbb{R})$ *una successione tesa di misure di probabilità tale che ogni sua sottosuccessione debolmente convergente ammette lo stesso limite* $\mathbb{P} \in \mathfrak{P}(\mathbb{R})$ *. Allora l'intera successione* $\{\mathbb{P}_n\}_{n\geq 1}$ *converge debolmente a* $\mathbb{P}$*.*

Dimostrazione: Se $\{\mathbb{P}_n\}_{n\geq 1}$ non converge debolmente a $\mathbb{P}$, $\exists f \in C_b(\mathbb{R})$ tale che $\{\mathbb{E}(f)_n\}_{n\geq 1}$ non converge a $\mathbb{E}(f)$. Quindi, $\exists\varepsilon > 0$ e $\{n'_\varepsilon\} \subseteq \{n\}$ tale che $\left|\mathbb{E}_{n'_\varepsilon}(f) - \mathbb{E}(f)\right| \geq \varepsilon > 0$, però essendo $\{\mathbb{P}_k\}_{k\in\{n'_\varepsilon\}}$ tesa e dunque, per il teorema di Prokhorov, relativamente compatta, $\exists\{n''_\varepsilon\} \subset \{n'_\varepsilon\}$ tale che $\{\mathbb{P}_k\}_{k\in\{n''_\varepsilon\}}$ converge debolmente a $\mathbb{Q} \in \mathfrak{P}(\mathbb{R})$. Ma, per ipotesi $\mathbb{P} = \mathbb{Q}$ quindi $\{\mathbb{E}_k(f)\}_{k\in\{n''_\varepsilon\}}$ converge a $\mathbb{E}(f)$ contro l'ipotesi $\left|\mathbb{E}_{n''_\varepsilon}(f) - \mathbb{E}(f)\right| \geq \varepsilon$. ■

Lemma 55 *Sia* $\{\mathbb{P}_n\}_{n\geq 1} \subset \mathfrak{P}(\mathbb{R})$ *una successione tesa di misure di probabilità e sia* $\{\varphi_n\}_{n\geq 1}$ *la corrispondente successione di funzioni caratteristiche. Allora,* $\{\mathbb{P}_n\}_{n\geq 1}$ *converge debolmente se e solo se* $\{\varphi_n\}_{n\geq 1}$ *converge puntualmente.*

Dimostrazione:

$\Rightarrow$ Segue dalla definizione di convergenza stretta, considerando $\forall t \in \mathbb{R}$ le funzioni $\cos tx$, $\sin tx$.

$\Leftarrow$ $\{\mathbb{P}_n\}_{n\geq 1}$ è tesa quindi, per il teorema di Prokhorov, $\exists\{n'\} \subset \{n\}$ tale che $\{\mathbb{P}_k\}_{k\in\{n'\}}$ converge debolmente a $\mathbb{P} \in \mathfrak{P}(\mathbb{R})$. Se $\{\mathbb{P}_n\}_{n\geq 1}$ non converge debolmente a $\mathbb{P}$, per il lemma precedente $\exists\{n''\} \subset \{n'\}$ tale che $\{\mathbb{P}_k\}_{k\in\{n''\}}$ converge debolmente a $\mathbb{Q} \neq \mathbb{P}$. Poiché $\forall t \in \mathbb{R}$, $\exists \lim_{n\to\infty} \varphi_n(t)$ allora,

$$\int_{\mathbb{R}} \mathbb{P}(dx)\, e^{itx} = \lim_{n'\to\infty} \varphi_{n'}(t) = \lim_{n''\to\infty} \varphi_{n''}(t) = \int_{\mathbb{R}} \mathbb{Q}(dx)\, e^{itx} . \tag{1.81}$$

Quindi, per il teorema di unicità, $\mathbb{P} = \mathbb{Q}$, contro l'ipotesi che $\mathbb{P} \neq \mathbb{Q}$.

■

Lemma 56 *Sia F una funzione di distribuzione e φ la funzione caratteristica associata. Allora* $\exists c > 0$ *tale che,* $\forall a > 0$

$$\mathbb{E}\left[\mathbf{1}_{[-\frac{1}{a},\frac{1}{a}]^c}\right] \leq \frac{c}{a}\int_0^a dt\,(1 - \operatorname{Re}\varphi(t)) . \tag{1.82}$$

Dimostrazione: Siccome

$$\operatorname{Re}\varphi(t) = \int_{\mathbb{R}} dF(x)\cos tx , \tag{1.83}$$

Per teorema di Fubini

$$\begin{aligned} \frac{1}{a}\int_0^a dt\,(1 - \operatorname{Re}\varphi(t)) &= \frac{1}{a}\int_0^a dt \int_{\mathbb{R}} dF(x)\,(1 - \cos tx) = \int_{\mathbb{R}} dF(x)\,\frac{1}{a}\int_0^a dt\,(1 - \cos tx) \\ &= \int_{\mathbb{R}} dF(x)\left(1 - \frac{\sin ax}{ax}\right) \geq \inf_{y\in\mathbb{R}\,:\,|y|>1}\left(1 - \frac{\sin y}{y}\right)\int_{\{x\in\mathbb{R}\,:\,|ax|>1\}} dF(x) \\ &\geq c^{-1}\int_{\{x\in\mathbb{R}\,:\,|x|>\frac{1}{a}\}} dF(x) . \end{aligned} \tag{1.84}$$

■

Teorema 57 *(di continuità) Sia $\{F_n\}_{n\geq 1}$ una successione di funzioni di distribuzione e sia $\{\varphi_n\}_{n\geq 1}$ la corrispondente successione di funzioni caratteristiche.*

1. *Se $\{F_n\}_{n\geq 1}$ converge in generale alla funzione di distribuzione F, allora $\{\varphi_n\}_{n\geq 1}$ converge puntualmente a φ funzione caratteristica di F.*

2. *Se $\{\varphi_n\}_{n\geq 1}$ converge puntualmente ed il limite φ è una funzione continua in $t = 0$, allora φ è la funzione caratteristica di una distribuzione di probabilità $\mathbb{P}$ e la successione $\{F_n\}_{n\geq 1}$ converge in generale a F funzione di distribuzione associata a $\mathbb{P}$.*

Dimostrazione:

1. Poiché per il Teorema 131 la convergenza in generale della successione $\{F_n\}_{n\geq 1}$ è equivalente alla convergenza stretta delle distribuzioni di probabilità associate, considerando $\forall t \in \mathbb{R}$ le funzioni $\sin tx$ e $\cos tx$ si ha la tesi.

2. Sia $\{\varphi_n\}_{n\geq 1}$ convergente puntualmente a φ continua in $t = 0$ e sia, $\forall n \geq 1$, $\mathbb{P}_n$ la distribuzione di probabilità associata a F_n. $\forall a > 0$, per il lemma precedente

$$\mathbb{P}_n\left([-\frac{1}{a}, \frac{1}{a}]^c\right) \leq \frac{c}{a}\int_0^a dt\,[1 - \operatorname{Re}\varphi_n(t)] \tag{1.85}$$

e per teorema di Lebesgue della convergenza dominata

$$\lim_{n\to\infty}\mathbb{P}_n\left([-\frac{1}{a}, \frac{1}{a}]^c\right) \leq \frac{c}{a}\lim_{n\to\infty}\int_0^a dt\,[1 - \operatorname{Re}\varphi_n(t)] = \frac{c}{a}\int_0^a dt\,[1 - \operatorname{Re}\varphi(t)]\,. \tag{1.86}$$

Ma essendo φ continua in $t = 0$, $\forall\varepsilon > 0$, $\exists a_\varepsilon > 0$ tale che $\forall n \geq 1$

$$\mathbb{P}_n\left([-\frac{1}{a_\varepsilon}, \frac{1}{a_\varepsilon}]^c\right) \leq \varepsilon\,. \tag{1.87}$$

Perciò la successione $\{\mathbb{P}_n\}_{\geq 1}$ è tesa, quindi, per il teorema di Prokhorov, relativamente compatta e, per i due precedenti lemmi, convergente debolmente a $\mathbb{P} \in \mathfrak{P}(\mathbb{R})$. Pertanto, $\forall t \in \mathbb{R}$

$$\lim_{n\to\infty}\varphi_n(t) = \lim_{n\to\infty}\int_{\mathbb{R}}\mathbb{P}_n(dx)\,e^{itx} = \int_{\mathbb{R}}\mathbb{P}(dx)\,e^{itx}\,, \tag{1.88}$$

ma per ipotesi, $\forall t \in \mathbb{R}$, $\lim_{n\to\infty}\varphi_n(t) = \varphi(t)$. Quindi, $\forall t \in \mathbb{R}$,

$$\varphi(t) = \int_{\mathbb{R}}\mathbb{P}(dx)\,e^{itx}\,. \tag{1.89}$$

■

Osservazione 58 *La topologia della convergenza debole è metrizzabile, ciò verrà discusso per il caso delle misure su $\mathbb{R}$ quando si tratterà la convergenza in distribuzione di successioni di variabili aleatorie.*

The Title The Author The Date

1.5 Applicazioni tra spazi misurabili e variabili aleatorie

Definizione 59 *Data una coppia di spazi misurabili* $(\Omega,\mathcal{F})$ *e* $(E,\mathcal{E})$, *un'applicazione* $X:\Omega\longrightarrow E$ *è detta* misurabile *se* $X^{-1}(\mathcal{E})\subseteq\mathcal{F}$, *ovvero se* $\forall B\in\mathcal{E},\ X^{-1}(B)\in\mathcal{F}$.

Definizione 60 *Sia* $(\Omega,\mathcal{F},\mathbb{P})$ *uno spazio di probabilità.*

- *Un'applicazione misurabile* ξ *da* $(\Omega,\mathcal{F})$ *a* $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ *è detta* variabile aleatoria*;*
- *Un'applicazione misurabile* ξ *da* $(\Omega,\mathcal{F})$ *a* $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$ *è detto* vettore aleatorio.

In generale, una generica applicazione misurabile tra spazi misurabile è detta elemento aleatorio.

Nel seguito supporremo che tutti gli elementi aleatori abbiano valori in uno spazio E polacco, localmente compatto e σcompatto, e dunque che

$$\mathfrak{P}(E):=\mathfrak{P}(E,\mathcal{B}(E))\subset\mathcal{M}^{1,+}(E)\ . \tag{1.90}$$

Osservazione 61 *Inoltre, useremo l'abbreviazione* v.a. *per indicare sia i vettori che le variabili aleatorie e, considerato lo spazio di probabilità* $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n),\mathbb{P})$, $n\geq 1$, *per ogni* f *misurabile su* $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$ *in sé tale che*

$$\left(\int_{\mathbb{R}^n}\mathbb{P}(dx)(f(x)\vee 0)\right)\wedge\left(\int_{\mathbb{R}^n}\mathbb{P}(dx)(f(x)\wedge 0)\right)<\infty\,, \tag{1.91}$$

porremo

$$\mathbb{E}(f(\xi)):=\int_{\mathbb{R}^n}\mathbb{P}(dx)f(x)=\int_{\mathbb{R}^n}dF(x)f(x)\ . \tag{1.92}$$

Definizione 62 *Sia* $(\Omega,\mathcal{F},\mathbb{P})$ *uno spazio di probabilità,* ξ *una v.a. e* η *un* vettore aleatorio n-dimensionale.

La misura $\mathbb{P}_\xi:=\mathbb{P}\circ\xi^{-1}\in\mathfrak{P}(\mathbb{R})$ *è detta distribuzione di probabilità o* legge *della variabile aleatoria* ξ *e la funzione*

$$\mathbb{R}\ni x\longmapsto F_\xi(x):=\mathbb{P}\{\omega\in\Omega:\xi(\omega)\leq x\}=\mathbb{P}_\xi(-\infty,x]\in[0,1] \tag{1.93}$$

è detta funzione di distribuzione o ripartizione *della variabile aleatoria* ξ. *Allo stesso modo,* $\mathbb{P}_\eta:=\mathbb{P}\circ\eta^{-1}\in\mathfrak{P}(\mathbb{R}^n)$ *è detta distribuzione di probabilità o* legge *della vettore aleatorio* η *e la funzione*

$$\begin{aligned}\mathbb{R}^n\ni(x_1,..,x_n)=x\longmapsto F_\eta(x)=F_\eta(x_1,..,x_n):=\mathbb{P}&\left(\bigcap_{i=1}^n\{\omega\in\Omega:\eta_i(\omega)\leq x_i\}\right)\\ =\mathbb{P}_\eta&\left(\bigcap_{i=1}^n\{y\in\mathbb{R}^n:y_i\leq x_i\}\right)\in[0,1]\end{aligned}$$

è detta funzione di distribuzione o ripartizione *del vettore aleatorio* η.

1.5.1 Stabilità della collezione delle v.a. rispetto alla topologia della convergenza puntuale

Definizione 63 *Una successione di v.a.* $\{\xi_n\}_{n\geq 1}$ *definite su* $(\Omega, \mathcal{F})$ *converge puntualmente se,* $\forall \omega \in \Omega$, *la successione numerica* $\{\xi_n(\omega)\}_{n\geq 1}$ *è convergente. Se inoltre esiste* $\xi : \Omega \longmapsto \mathbb{R}$ *tale che,* $\forall \omega \in \Omega$, $\lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)$, *allora* $\{\xi_n(\omega)\}_{n\geq 1}$ *è detta convergere puntualmente a* ξ.

Teorema 64 *Sia* $\{\xi_n\}_{n\geq 1}$ *una successione di v.a. definite su* $(\Omega, \mathcal{F})$ *convergente puntualmente a* ξ. *Allora,* ξ *è una v.a..*

Dimostrazione: Basta verificare che, $\forall c \in \mathbb{R}$, $\{\omega \in \Omega : \xi(\omega) < c\}$ è misurabile. Tuttavia quest'insieme coincide con l'insieme

$$M_c := \bigcup_{k\geq 1} \bigcup_{n\geq 1} \bigcap_{m\geq n} \left\{\omega \in \Omega : \xi_m(\omega) < c - \frac{1}{k}\right\} \in \mathcal{F}. \tag{1.94}$$

Infatti, dato $\omega' \in \Omega$, se $\xi(\omega') < c$, esiste $k \geq 1$ tale che $\xi(\omega') < c - \frac{1}{k}$, ma poiché $\{\xi_n(\omega')\}_{n\geq 1}$ converge a $\xi(\omega')$ esiste n sufficientemente grande tale che $\forall m \geq n$, $\xi_m(\omega') < c - \frac{1}{k}$. Quindi, $\omega' \in M_c$. Viceversa, se $\omega' \in M_c$, allora esiste $m' \geq 1$ tale che $\forall m \geq m'$, $\xi_m(\omega') < c - \frac{1}{k}$, per qualche $k \geq 1$, cioè $\xi(\omega') < c$. ■

1.6 Generalità sull'attesa di una variabile aleatoria

D'ora in poi, laddove non espressamente specificato, considereremo $(\Omega, \mathcal{F}, \mathbb{P})$ come spazio di probabilità completo, cioè tale che ogni sottoinsieme di un insieme trascurabile rispetto a $\mathbb{P}$, ovvero di probabilità nulla, risulti misurabile.

Sia ξ una v.a. non negativa, data la successione di v.a.

$$\xi_n(\omega) := \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbf{1}_{\xi^{-1}\left(\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)\right)}(\omega) + n\mathbf{1}_{\xi^{-1}([n,+\infty))}(\omega), \tag{1.95}$$

$\forall \omega \in \Omega$, $n \geq 1$, si ha $\xi_n(\omega) \leq \xi_{n+1}(\omega) \leq \xi(\omega)$ e $\{\xi_n(\omega)\}_{n\geq 1}$ converge ad $\xi(\omega)$.

Ponendo, $\forall n \geq 1$,

$$\mathbb{E}(\xi_n) := \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathbb{P}\left\{\omega \in \Omega : \xi(\omega) \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)\right\} + n\mathbb{P}\{\omega \in \Omega : \xi(\omega) \geq n\}, \tag{1.96}$$

dato che $\mathbb{E}(\xi_n) \leq \mathbb{E}(\xi_{n+1})$, $\lim_{n\to\infty} \mathbb{E}(\xi_n)$ esiste ed eventualmente uguaglia $+\infty$. Sia allora

$$\mathbb{E}(\xi) := \lim_{n\to\infty} \mathbb{E}(\xi_n). \tag{1.97}$$

Tale limite è detto *valore atteso* della variabile aleatoria ξ.

Data una qualsiasi variabile aleatoria ξ, ponendo $\xi^+ := \xi \vee 0$, $\xi^- := -(\xi \wedge 0)$, $\xi = \xi^+ - \xi^-$, il valore atteso di ξ, $\mathbb{E}(\xi) = \mathbb{E}(\xi^+) - \mathbb{E}(\xi^-)$, risulta definito se $\mathbb{E}(\xi^+) \wedge \mathbb{E}(\xi^-) < +\infty$ ed è finito se $\mathbb{E}(\xi^+), \mathbb{E}(\xi^-) < +\infty$, ovvero se $\mathbb{E}(|\xi|) = \mathbb{E}(\xi^+) + \mathbb{E}(\xi^-) < +\infty$.

Teorema 65 *(della convergenza monotona) Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità e su questo siano definite le v.a.* η *e* ξ. *Sia inoltre* $\{\xi_n\}_{n\geq 1}$ *una successione di v.a. su* $(\Omega, \mathcal{F}, \mathbb{P})$ *convergente puntualmente a* ξ.

1. *Se* $\{\xi_n\}_{n\geq 1}$ *e* η *sono tali che* $\mathbb{E}(\eta) > -\infty$ *e* $\forall n \geq 1$, *si ha* $\eta \leq \xi_n \leq \xi_{n+1} \leq \xi$, *allora la successione numerica crescente* $\{\mathbb{E}(\xi_n)\}_{n\geq 1}$ *converge a* $\mathbb{E}(\xi)$.
2. *Se* $\{\xi_n\}_{n\geq 1}$ *e* η *sono tali che* $\mathbb{E}(\eta) < \infty$ *e* $\forall n \geq 1$, *si ha* $\eta \geq \xi_n \geq \xi_{n+1} \geq \xi$, *allora la successione numerica decrescente* $\{\mathbb{E}(\xi_n)\}_{n\geq 1}$ *converge a* $\mathbb{E}(\xi)$.

Dimostrazione:

1. Supponiamo dapprima che $\eta \geq 0$. $\forall n \geq 1$, sia $\{\xi_n^{(k)}\}_{k\geq 1}$ una successione crescente di v.a. semplici convergenti puntualmente a ξ_n. Allora, posto $\zeta_1 := \xi_1^{(1)}$ e $\forall k \geq 2, \zeta_k := \max_{1\leq n\leq k} \xi_n^{(k)}$, si ha che $\{\zeta_k\}_{k\geq 1}$ è una successione crescente di funzioni semplici tali che $\forall k \geq 1$,
$$\zeta_k = \max_{1\leq n\leq k} \xi_n^{(k)} \leq \max_{1\leq n\leq k} \xi_n = \xi_k \leq \xi \ . \tag{1.98}$$
Sia $\zeta := \lim_{n\uparrow\infty} \zeta_n$. Poiché per $1 \leq n \leq k, \xi_n^{(k)} \leq \zeta_k$, passando al limite per $k \uparrow \infty$ dalla precedente disuguaglianza $\forall n \geq 1$ si ottiene
$$\xi_n \leq \zeta \leq \xi \tag{1.99}$$
ovvero $\zeta = \xi$. Inoltre, siccome $\{\zeta_k\}_{k\geq 1}$ è una successione crescente di funzioni semplici
$$\mathbb{E}(\xi) = \mathbb{E}(\zeta) = \lim_{n\uparrow\infty} \mathbb{E}(\zeta_n) \leq \lim_{n\uparrow\infty} \mathbb{E}(\xi_n) \ . \tag{1.100}$$
D'altra parte, poiché $\{\xi_k\}_{k\geq 1}$ è crescente e $\forall n \geq 1, \xi_n \leq \xi$, segue $\lim_{n\uparrow\infty} \mathbb{E}(\xi_n) \leq \mathbb{E}(\xi)$ dunque la tesi.

   Sia η è una generica v.a. tale che $\mathbb{E}(\eta) > -\infty$.

   (a) Se $\mathbb{E}(\eta) = \infty$, allora $\forall n \geq 1, \mathbb{E}(\xi_n) = \infty$ dunque di conseguenza anche $\mathbb{E}(\xi) = \infty$.

   (b) Se $\mathbb{E}(\eta) < \infty$, allora $\mathbb{E}(|\eta|) < \infty$ e la successione di v.a. $\{\xi_n - \eta\}_{n\geq 1}$ ha termini non negativi, è crescente e coverge puntualmente a $\xi - \eta$. Percui, per quanto prima dimostrato, la successione numerica $\{\mathbb{E}(\xi_n - \eta)\}_{n\geq 1}$ è a termini positivi, crescente e convergente a $\mathbb{E}(\xi) - \mathbb{E}(\eta)$. Ma poiché $\mathbb{E}(|\eta|) < \infty$ ciò implica che la successione numerica $\{\mathbb{E}(\xi_n)\}_{n\geq 1}$ è crescente e convergente a $\mathbb{E}(\xi)$.

2. Considerando $-\eta, -\xi$ e $\{-\xi_k\}_{k\geq 1}$ la tesi segue dal punto precedente.

■

1.6.1 Disuguaglianza di Markov e sue generalizzazioni

Proposizione 66 *(Disuguaglianza di Markov) Sia X una v.a. non negativa definita su* $(\Omega, \mathcal{F}, \mathbb{P})$ *.* $\forall \lambda > 0$ *vale*

$$\mathbb{P}\{\omega \in \Omega : X(\omega) > \lambda\} \leq \frac{\mathbb{E}(X)}{\lambda} \tag{1.101}$$

Dimostrazione: $\forall \lambda > 0$ si ha

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}\left(X \mathbf{1}_{[0,\lambda]}(X)\right) + \mathbb{E}\left(X \mathbf{1}_{(\lambda,+\infty)}(X)\right) \geq \mathbb{E}\left(X \mathbf{1}_{(\lambda,+\infty)}(X)\right) \\ &\geq \lambda \mathbb{E}\left(\mathbf{1}_{(\lambda,+\infty)}(X)\right) = \lambda \mathbb{P}\{\omega \in \Omega : X(\omega) > \lambda\}. \end{aligned} \tag{1.102}$$

■

Corollario 67 *Se f è una funzione non negativa, strettamente crescente e X una v.a. definita su* $(\Omega, \mathcal{F}, \mathbb{P})$ *.* $\forall \lambda > 0$ *si ha*

$$\mathbb{P}\{\omega \in \Omega : X(\omega) > \lambda\} \leq \frac{\mathbb{E}(f(X))}{f(\lambda)}. \tag{1.103}$$

Dimostrazione: Poiché f è strettamente crescente

$$\mathbb{P}\{\omega \in \Omega : X(\omega) > \lambda\} = \mathbb{P} \circ X^{-1}[\lambda, +\infty) = \mathbb{P} \circ X^{-1}\{x \in \mathbb{R} : f(x) > f(\lambda)\}, \tag{1.104}$$

e siccome f è non negativa, dalla disuguaglianza di Markov segue

$$\mathbb{P} \circ X^{-1}\{x \in \mathbb{R} : f(x) > f(\lambda)\} \leq \frac{\int \mathbb{P} \circ X^{-1}(dx) f(x)}{f(\lambda)} = \frac{\mathbb{E}(f \circ X)}{f(\lambda)}.$$

■

Osservazione 68 *(Disuguagluanza di Markov esponenziale) Il precedente risultato implica che* $\forall r, \lambda > 0$,

$$\mathbb{P}\{\omega \in \Omega : X(\omega) > \lambda\} \leq e^{-r\lambda} \mathbb{E}\left(e^{rX}\right). \tag{1.105}$$

Quindi

$$\mathbb{P}\{\omega \in \Omega : X(\omega) > \lambda\} \leq \inf_{r>0} e^{-r\lambda} \mathbb{E}\left(e^{rX}\right) = e^{-\sup_{r>0}[r\lambda - \log \mathbb{E}\left(e^{rX}\right)]}. \tag{1.106}$$

Corollario 69 *(Disuguaglianza di Chebichev) Sia X una v.a. definita su* $(\Omega, \mathcal{F}, \mathbb{P})$ *.* $\forall \varepsilon > 0$ *vale*

$$\mathbb{P}\{\omega \in \Omega : |X(\omega) - \mathbb{E}(X)| > \varepsilon\} \leq \frac{Var(X)}{\varepsilon^2}. \tag{1.107}$$

Dimostrazione: Si ponga $Y := |X - \mathbb{E}(X)|$ e si applichi il corollario precedente considerando $f = x^2$. ■

Esercizio 3 Dimostrare la disuguaglianza di Chebichev generalizzata

$$\mathbb{P}\{\omega \in \Omega : |X(\omega) - \mathbb{E}(X)| > \varepsilon\} \leq \frac{\mathbb{E}(|X - \mathbb{E}(X)|^p)}{\varepsilon^p}, \quad p > 0. \tag{1.108}$$

1.7 Funzione caratteristica di una variabile aleatoria

Definizione 70 *Si definisce* funzione caratteristica *di un v.a.* $\xi \in \mathbb{R}^n$ *definito sullo spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$,

$$\varphi_\xi(t) := \int_{\mathbb{R}^n} \mathbb{P}_\xi(dx)\, e^{i(t,x)} = \int_{\mathbb{R}^n} dF_\xi(x)\, e^{i(t,x)} = \mathbb{E}\left(e^{i(t,\xi)}\right) \tag{1.109}$$

la funzione caratteristica della distribuzione di probabilità $\mathbb{P}_\xi \in \mathfrak{P}(\mathbb{R}^n)$ *.*

1.7.1 Proprietà della funzione caratteristica di una v.a.

Consideriamo il caso in cui ξ sia una v.a. su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, allora $\varphi_\xi(t)$ ha le seguenti proprietà:

1. $\left|\varphi_\xi(t)\right| \leq \varphi_\xi(0) = 1.$

2. $\overline{\varphi_\xi(t)} = \varphi_\xi(-t)$.

3. $\varphi_\xi(t)$ è uniformemente continua in $\mathbb{R}$. Infatti, $\forall t, h \in \mathbb{R}$,

$$\left|\varphi_\xi(t+h) - \varphi_\xi(t)\right| = \left|\mathbb{E}\left[e^{i(t+h)\xi} - e^{it\xi}\right]\right| \leq \mathbb{E}\left|e^{ith} - 1\right| \leq 2\,. \tag{1.110}$$

   quindi per il teorema di Lebesgue della convergenza dominata $\lim_{h\to 0} \mathbb{E}\left|e^{ih\xi} - 1\right| = 0$.

4. Se $\forall B \in \mathcal{B}(\mathbb{R})$

$$\int_B dF_\xi = \int_B dF_{-\xi}\,, \tag{1.111}$$

   F_ξ è detta simmetrica. Allora, $\varphi_\xi(t) \in \mathbb{R} \iff F_\xi$ è simmetrica. Infatti, se F_ξ è simmetrica $\mathbb{E}[\sin tx] = 0$, perciò $\varphi_\xi(t) = \mathbb{E}[\cos tx] \in \mathbb{R}$. Viceversa, poiché

$$\varphi_{-\xi}(t) = \varphi_\xi(-t) = \overline{\varphi_\xi(t)} = \varphi_\xi(t)\,, \tag{1.112}$$

   dal teorema di unicità si ha che ξ e $-\xi$ hanno la stessa funzione di distribuzione.

5. Se $\exists n \geq 1$ tale che $\mathbb{E}(|\xi|^n) < \infty$, allora $\varphi_\xi \in C^r(\mathbb{R})$, $\forall r \leq n$ e

$$\varphi_\xi(t) = \sum_{k=0}^{n} \varphi_\xi^{(k)}(0)\, \frac{(it)^k}{k!} + o(t)\ . \tag{1.113}$$

   Infatti, per la disuguaglianza di Lyapunov, $\forall r \leq n$, $(\mathbb{E}\,|\xi|^r)^{\frac{1}{r}} \leq (\mathbb{E}\,|\xi|^n)^{\frac{1}{n}}$. Inoltre,

$$\frac{\varphi_\xi(t+h) - \varphi_\xi(t)}{h} = \mathbb{E}\left[e^{it\xi}\frac{e^{ih\xi} - 1}{h}\right] \tag{1.114}$$

e poiché

$$\left|\mathbb{E}\left[e^{it\xi}\frac{e^{ith}-1}{h}\right]\right| \leq \mathbb{E}\left|e^{it\xi}\frac{e^{ih\xi}-1}{h}\right| \leq \mathbb{E}\left|\xi\right| < \infty \tag{1.115}$$

per il teorema di Lebesgue della convergenza dominata

$$\lim_{h\to 0}\frac{\varphi_{\xi}(t+h)-\varphi_{\xi}(t)}{h} = \mathbb{E}\left[i\xi e^{it\xi}\right] = \varphi'_{\xi}(t)\ . \tag{1.116}$$

Procedendo per induzione si ha che $\varphi_{\xi}^{(r)}(t)$ esiste $\forall 1 < r \leq n,\ t \in \mathbb{R}$.

6. Se per qualche $n \geq 1$, $\exists \varphi_{\xi}^{(2n)}(0)$ allora $\mathbb{E}\left(\xi^{2n}\right) < \infty$.

7. Se $\forall n \geq 1$, $\mathbb{E}\left|\xi\right|^{n} < \infty$ e $\overline{\lim}_{n}\frac{(\mathbb{E}|\xi|^{n})^{\frac{1}{n}}}{n} = \frac{1}{eR} < \infty$ allora

$$\varphi_{\xi}(t) = \sum_{n=0}^{\infty}\varphi_{\xi}^{(n)}(0)\frac{(it)^{n}}{n!}\ ,\quad \forall t \in \mathbb{R} : |t| < R\,. \tag{1.117}$$

Infatti, se $s \in (0,R)$ allora, $\forall t \in \mathbb{R} : |t| < s$,

$$\varphi_{\xi}(t) = \sum_{n=0}^{\infty}\mathbb{E}\left(\xi^{n}\right)\frac{(it)^{n}}{n!} \tag{1.118}$$

ovvero,

$$\sum_{n=0}^{\infty}\mathbb{E}\left(|\xi|^{n}\right)\frac{(es)^{n}}{n^{n}} < \infty\,, \tag{1.119}$$

cioè $\overline{\lim}_{n}\left(\mathbb{E}\left|\xi\right|^{n}\right)^{\frac{1}{n}}\frac{es}{n} < 1$.

1.7.2 Problema dei momenti

Definizione 71 *Sia F una funzione di distribuzione di probabilità su $\mathbb{R}$. Il numero*

$$m_{n} := \int_{\mathbb{R}} dF(x)\, x^{n} \tag{1.120}$$

è detto momento n-simo o di ordine n *della distribuzione di probabilità associata ad F.*

Sia ξ una v.a. definita su uno spazio di probabilità $(\Omega,\mathcal{F},\mathbb{P})$ la cui funzione di distribuzione non sia nota, ma della quale esistono tutti i momenti. Poiché per teorema d'inversione della funzione caratteristica φ_{ξ} determina univocamente la distribuzione di probabilità della v.a. ξ e, poiché se esistono tutti i momenti di ξ, φ_{ξ} si può scivere come serie di Taylor in un opportuno intorno di $t = 0$, diamo alcune condizioni sufficienti perché i momenti di ξ ne determinino le legge.

Teorema 72 *Sia F una funzione di distribuzione di probabilità su $\mathbb{R}$ e, $\forall n \geq 1$, sia*

$$\mu_n := \int_{\mathbb{R}} dF(x) |x|^n . \tag{1.121}$$

Se $\overline{\lim}_n \frac{\mu_n^{\frac{1}{n}}}{n} < \infty$, allora i momenti della distribuzione di probabilità associata ad F la determinano univocamente.

Dimostrazione: Segue dalla proprietà 7. delle funzioni caratteristiche e dal teorema d'inversione. ■

Corollario 73 *Se una distribuzione di probabilità è concentrata su un intervallo finito di $\mathbb{R}$, i momenti la determinano univocamente.*

Dimostrazione: Segue dal teorema precedente considerando il fatto che se I è l'intervallo su cui è concentrata la distribuzione di probabilità in considerazione e F ne è la funzione di distribuzione

$$\mu_n = \int_I dF(x) |x|^n \leq \sup_{x \in I} |x|^n < \infty . \tag{1.122}$$

■

Corollario 74 *Condizione necessaria e sufficiente affinché una distribuzione di probabilità sia determinata univocamente dai suoi momenti è che $\overline{\lim}_n \frac{(m_{2n})^{\frac{1}{2n}}}{2n} < \infty$.*

Dimostrazione: Segue dal teorema precedente considerando il fatto che per la disuguaglianza di Cauchy-Schwarz i momenti di ordine dispari possono essere stimati tramite quelli di ordine pari ■

Corollario 75 *(Test di Carleman per l'unicità del problema dei momenti) Sia $\{m_n\}_{n\geq 1}$ la collezione dei momenti di una distribuzione di probabilità.*

1. *Se*

$$\sum_{n\geq 1} \frac{1}{m_{2n}^{\frac{1}{2n}}} = \infty , \tag{1.123}$$

allora i momenti determinano univocamente la distribuzione di probabilità.

2. *Se la distribuzione di probabilità è concentrata in $[0, +\infty)$ questa è determinata univocamente dai suoi momenti se*

$$\sum_{n\geq 1} \frac{1}{m_n^{\frac{1}{2n}}} = \infty . \tag{1.124}$$

È inoltre possibile, tramite la funzione caratteristica, conoscere se la v.a. ad essa associata è degenere o, se è discreta, se ha valori su un reticolo.

Teorema 76 *Sia φ_ξ la funzione caratteristica della v.a. ξ definita su $(\Omega, \mathcal{F}, \mathbb{P})$.*

1. *Se $\exists t_0 \in \mathbb{R}\backslash\{0\} : \left|\varphi_\xi(t_0)\right| = 1$, allora $\exists a \in \mathbb{R}$ tale che $\xi \in \{a + \frac{2\pi}{t_0}n\}_{n\in\mathbb{Z}}$, ovvero*
$$\sum_{n\in\mathbb{Z}} \mathbb{P}\left\{\omega \in \Omega : \xi(\omega) = a + \frac{2\pi}{t_0}n\right\} = 1\,. \tag{1.125}$$

2. *Se $\exists \alpha \in \mathbb{R}\backslash\mathbb{Q}$ e $t_0 \in \mathbb{R}\backslash\{0\}$ tale che $\left|\varphi_\xi(t_0)\right| = \left|\varphi_\xi(\alpha t_0)\right| = 1$, allora ξ è degenere e $\exists a \in \mathbb{R}$ tale che $\mathbb{P}\{\omega \in \Omega : \xi(\omega) = a\} = 1$.*

3. *Se $\forall t \in \mathbb{R}$, $\left|\varphi_\xi(t)\right| = 1$, allora ξ è degenere.*

Dimostrazione:

1. Se $\exists t_0 \in \mathbb{R}\backslash\{0\} : \left|\varphi_\xi(t_0)\right| = 1$, allora $\exists a \in \mathbb{R}$ tale che $\varphi_\xi(t_0) = e^{it_0 a}$. Ma,
$$e^{it_0 a} = \int_{\mathbb{R}} dF(x)\, e^{it_0 x}\,. \tag{1.126}$$
Perciò
$$\int_{\mathbb{R}} dF(x)\, e^{it_0(x-a)} = 1, \tag{1.127}$$
ovvero
$$\int_{\mathbb{R}} dF(x) \cos t_0 (x-a) = 1\,, \tag{1.128}$$
cioè
$$\int_{\mathbb{R}} dF(x) \left[\cos t_0(x-a) - 1\right] = 0\,. \tag{1.129}$$
Poiché $\cos t_0(x-a) - 1 \leq 0$, allora dev'essere $\cos t_0(\xi - a) = 1$ $\mathbb{P}$-q.c..

2. Dall'affermazione 1. segue che
$$\sum_{n\in\mathbb{Z}} \mathbb{P}\left\{\omega \in \Omega : \xi(\omega) = a + \frac{2\pi}{t_0}n\right\} = \sum_{n\in\mathbb{Z}} \mathbb{P}\left\{\omega \in \Omega : \xi(\omega) = b + \frac{2\pi}{\alpha t_0}n\right\}\,. \tag{1.130}$$
Quindi se ξ non è degenere devono esistere almeno due punti coincidenti negli insiemi $\{a + \frac{2\pi}{t_0}n\}_{n\in\mathbb{Z}}$ e $\{b + \frac{2\pi}{\alpha t_0}n\}_{n\in\mathbb{Z}}$, ovvero $\exists m_1, m_2, n_1, n_2 \in \mathbb{Z}$ per cui
$$\begin{aligned} a + \frac{2\pi}{t_0}n_1 &= b + \frac{2\pi}{\alpha t_0}m_1\,, \\ a + \frac{2\pi}{t_0}n_2 &= b + \frac{2\pi}{\alpha t_0}m_2\,, \end{aligned} \tag{1.131}$$
cioè
$$\frac{2\pi}{t_0}(n_1 - n_2) = \frac{2\pi}{\alpha t_0}(m_1 - m_2)\,. \tag{1.132}$$
Ma ciò implica che $\alpha = \frac{m_1 - m_2}{n_1 - n_2}$ contro l'ipotesi che α sia irrazionale.

3. Segue dall'affermazione 2..

■

1.8 Attesa condizionata rispetto ad una sigma-algebra

Se $\mathcal{B} \subseteq \mathcal{F}$ è una subσalgebra di $\mathcal{F}$, indichiamo con $L(\mathcal{B})$ la collezione delle classi di equivalenza delle funzioni $\mathcal{B}$-misurabili da Ω in $\mathbb{R}$ che coincidono $\mathbb{P}$-q.c. e con $E(\mathcal{B})$ il sottoinsieme di $L(\mathcal{B})$ costituito dalle funzioni semplici integrabili.

Sia ξ una variabile aleatoria non negativa tale che $\mathbb{E}(\xi)$ sia definita e sia

$$\mathcal{F} \ni A \longmapsto \mathbb{E}(\mathbf{1}_A \xi) = \int_A d\mathbb{P}\xi = \mathbb{Q}_\xi(A) \in \overline{\mathbb{R}^+} . \tag{1.133}$$

- $\mathbb{Q}_\xi$ è numerabilmente additiva.

  Infatti se $A = \bigvee_{n\geq 1} A_n$, $\mathbf{1}_A = \sum_{n\geq 1} \mathbf{1}_{A_n}$ e

$$\mathbb{Q}_\xi(A) = \mathbb{E}(\mathbf{1}_A \xi) = \mathbb{E}\left(\sum_{n\geq 1} \mathbf{1}_{A_n}\xi\right) = \sum_{n\geq 1} \mathbb{E}(\mathbf{1}_{A_n}\xi) = \sum_{n\geq 1} \mathbb{Q}_\xi(A_n) . \tag{1.134}$$

- $\mathbb{Q}_\xi$ è assolutamente continua rispetto a $\mathbb{P}$, ovvero $\forall A \in \mathcal{F}$ tale che $\mathbb{P}(A) = 0$, $\mathbb{Q}_\xi(A) = 0$.

  Infatti ciò è vero per una variabile aleatoria semplice. Quindi se $\{\xi_n\}_{n\geq 1}$ è una successione monotona crescente di variabili aleatorie semplici convergente a ξ, per il teorema della convergenza monotona

$$\mathbb{Q}_\xi(A) = \mathbb{E}(\xi \mathbf{1}_A) = \lim_{n\to\infty} \mathbb{E}(\xi_n \mathbf{1}_A) = 0 , \tag{1.135}$$

  poiché $\forall n \geq 1$, $\mathbb{E}(\xi_n \mathbf{1}_A) = 0$ se $\mathbb{P}(A) = 0$.

Se ξ una variabile aleatoria tale che $\mathbb{E}(\xi)$ sia definita, allora, $\mathbb{Q}_\xi(A) = \mathbb{Q}_\xi^+(A) - \mathbb{Q}_\xi^-(A)$ con $\mathbb{Q}_\xi^\pm(A) = \int_A d\mathbb{P}\xi^\pm$.

Teorema 77 *(di Radon-Nikodým) Se μ una misura numerabilmente additiva e finita su uno spazio misurabile $(\Omega, \mathcal{F})$ e λ una misura assolutamente continua rispetto a μ. Allora esiste una funzione $f \in L(\mathcal{F})$ integrabile rispetto a μ, tale che $\forall A \in \mathcal{F}$, $\lambda(A) = \int_A \mu(d\omega) f(\omega)$.*

Perciò, considerando al posto di $(\Omega, \mathcal{F})$, $(\Omega, \mathcal{B})$, con $\mathcal{B}$ subσalgebra di $\mathcal{F}$, si ha che $\mathbb{Q}_\xi$ è assolutamente continua rispetto a $\mathbb{P}$. Pertanto esiste $f_\xi \in L(\mathcal{B})$ tale che, $\forall B \in \mathcal{B}$, $\mathbb{Q}_\xi(B) = \int_B d\mathbb{P} f_\xi$. f_ξ è detta *attesa condizionata* di ξ rispetto alla σalgebra $\mathcal{B}$ ed in generale si indica $\mathbb{E}(\xi|\mathcal{B})$. Una qualsiasi funzione $\mathcal{B}$-mis. h tale che $h = \mathbb{E}(\xi|\mathcal{B})$ $\mathbb{P}$-q.c. è detta *versione* della probabilità condizionata di ξ rispetto a $\mathcal{B}$.

Se $\xi = \mathbf{1}_A$, $A \in \mathcal{F}$, allora $\mathbb{E}(\mathbf{1}_A|\mathcal{B})$ è detta *probabilità condizionata* di A rispetto alla σalgebra $\mathcal{B}$ e si indica $\mathbb{P}(A|\mathcal{B})$. Inoltre, per definizione di attesa condizionata, si ha che, $\forall B \in \mathcal{B}$,

$$\mathbb{P}(A \cap B) = \int_{A\cap B} d\mathbb{P} = \int_B d\mathbb{P}\mathbf{1}_A \int_B d\mathbb{P}\mathbb{P}(A|\mathcal{B}) . \tag{1.136}$$

1.8.1 Attesa condizionata rispetto ad una v.a.

Data una v.a. ξ, sia $\mathcal{F}_\xi$ la subσalgebra di $\mathcal{F}$ *generata da* ξ, ovvero quella generata dalla collezione di eventi $\{\omega \in \Omega : \xi(\omega) \in B\}$, $B \in \mathcal{B}(\mathbb{R})$.

Teorema 78 *η è una v.a. $\mathcal{F}_\xi$-mis. $\Longleftrightarrow$ $\exists$ una funzione $\mathcal{B}(\mathbb{R})$-mis. ϕ, tale che $\eta = \phi \circ \xi$.*

Dimostrazione:

$\Longleftarrow$ Basta osservare che $\forall B \in \mathcal{B}(\mathbb{R})$,

$$\{\omega \in \Omega : \eta(\omega) \in B\} = \{\omega \in \Omega : (\phi \circ \xi)(\omega) \in B\} = \{\omega \in \Omega : \xi(\omega) \in \phi^{-1}(B)\} \in \mathcal{F}_\xi\,, \tag{1.137}$$

poiché $\phi^{-1}(B) \in \mathcal{B}(\mathbb{R})$.

$\Longrightarrow$ Sia Φ_ξ la classe delle v.a. $\mathcal{F}_\xi$-mis. e $\bar{\Phi}_\xi$ la classe delle v.a. $\mathcal{F}_\xi$-mis. per cui $\forall \eta \in \bar{\Phi}_\xi$, $\exists \phi$ $\mathcal{B}(\mathbb{R})$-mis. tale che $\eta = \phi \circ \xi$. Dunque, $\bar{\Phi}_\xi \subseteq \Phi_\xi$. Poiché $\forall A \in \mathcal{F}_\xi$, $\xi(A) \in \mathcal{B}(\mathbb{R})$, la v.a. $\mathbf{1}_A = \mathbf{1}_{\xi(A)} \circ \xi \in \bar{\Phi}_\xi$ che equivale ad affermare che tutte le v.a. semplici appartengono a $\bar{\Phi}_\xi$. Poiché ogni v.a. $\mathcal{F}_\xi$-mis. si può rappresentare come limite puntuale di v.a. semplici, data una generica v.a. $\mathcal{F}_\xi$-mis. η, esiste una successione $\{\eta_n\}_{n\geq 1}$ di v.a. semplici convergente puntualmente a η ed inoltre, $\forall n \geq 1$, esiste una funzione $\mathcal{B}(\mathbb{R})$-mis. ϕ_n tale che $\eta_n = \phi_n \circ \xi$. Quindi, poiché $B := \{x \in \mathbb{R} : \exists \lim_{n\to\infty} \phi_n(x)\} \in \mathcal{B}(\mathbb{R})$, ponendo $\phi(x) = \mathbf{1}_B(x) \lim_{n\to\infty} \phi_n(x)$, che è $\mathcal{B}(\mathbb{R})$-mis., si ha che $\forall \omega \in \Omega$, $\lim_{n\to\infty} (\phi_n \circ \xi)(\omega) = (\phi \circ \xi)(\omega)$, ovvero $\bar{\Phi}_\xi = \Phi_\xi$.

■

Dal risultato precedente segue che esiste una funzione $\mathcal{B}(\mathbb{R})$-mis. ϕ_η tale che l'attesa condizionata di una v.a. η rispetto alla σalgebra generata da un'altra v.a. ξ, $\mathbb{E}(\eta|\mathcal{F}_\xi)$ che di solito si indica $\mathbb{E}(\eta|\xi)$ è uguale a $\phi_\eta \circ \xi$. Allora, $\forall B \in \mathcal{B}(\mathbb{R})$, $\{\omega \in \Omega : \xi(\omega) \in B\} \in \mathcal{F}_\xi$ ed indicando con $\mathbb{P}_\xi$ la misura di probabilità indotta da ξ su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ segue

$$\int_{\{\omega\in\Omega\,:\,\xi(\omega)\in B\}} d\mathbb{P}\eta = \int_{\{\omega\in\Omega\,:\,\xi(\omega)\in B\}} d\mathbb{P}\mathbb{E}(\eta|\xi) = \int_{\{\omega\in\Omega\,:\,\xi(\omega)\in B\}} d\mathbb{P}\phi_\eta \circ \xi = \int_B d\mathbb{P}_\xi \phi_\eta\,. \tag{1.138}$$

Perciò ϕ_η è la derivata di Radon-Nikodým della misura segnata $\mathbb{Q}_\eta(B) = \int_B d\mathbb{P}_\xi \phi_\eta$ rispetto a $\mathbb{P}_\xi$. Allo stesso modo, nel caso in cui $\eta = \mathbf{1}_A$, $A \in \mathcal{F}$, si ha che esiste una funzione $\mathcal{B}(\mathbb{R})$-mis. ϕ_A tale che la probabilità condizionata $\mathbb{P}(A|\mathcal{F}_\xi) = \mathbb{P}(A|\xi)$ è uguale a $\phi_A \circ \xi$. Perciò, $\forall B \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}\left(A \cap \xi^{-1}(B)\right) = \int_{\{\omega\in\Omega\,:\,\xi(\omega)\in B\}} d\mathbb{P}\mathbf{1}_A = \int_{\{\omega\in\Omega\,:\,\xi(\omega)\in B\}} d\mathbb{P}\mathbb{P}(A|\xi) = \int_B d\mathbb{P}_\xi \phi_A\,. \tag{1.139}$$

Solitamente per indicare le funzioni $\mathbb{R} \ni x \longmapsto \phi_\eta(x), \phi_A(x) \in \mathbb{R}$ si usa rispettivamente la notazione $\mathbb{E}(\eta|\xi = x), \mathbb{P}(A|\xi = x)$.

Esempio 2 Se ξ è una v.a. discreta, sia $n_\xi \in \mathbb{N}$, tale che $\Omega \ni \omega \longmapsto \xi(\omega) \in \{x_i\}_{i=1}^{n_\xi}$. Allora, ponendo $\mathbb{P}\{\omega \in \Omega : \xi(\omega) = x_i\} = p_i$, si ha

$$\int_{\{\omega\in\Omega\,:\,\xi(\omega)\in B\}} d\mathbb{P}\eta = \int_{\{\omega\in\Omega\,:\,\xi(\omega)\in B\}} d\mathbb{P}\mathbb{E}(\eta|\xi) = \sum_{x_k\in B} p_k \mathbb{E}(\eta|\xi = x_k) \quad B \in \mathcal{B}(\mathbb{R})\ , \tag{1.140}$$

in particolare, se $B=\{x_i\}$,

$$\mathbb{E}(\eta|\xi=x_k)=\frac{1}{p_i}\int_{\{\omega\in\Omega\,:\,\xi(\omega)=x_i\}}d\mathbb{P}\eta\,, \tag{1.141}$$

e se $\eta=\mathbf{1}_A$, $A\in\mathcal{F}$,

$$\mathbb{P}(A|\xi=x_i)=\frac{\mathbb{P}(A\cap\{\omega\in\Omega\,:\,\xi(\omega)=x_i\})}{p_i}\,. \tag{1.142}$$

Esempio 3 Se $\Omega\ni\omega\longmapsto X(\omega)=(\xi(\omega),\eta(\omega))\in\mathbb{R}^2$ è un vettore aleatorio la distribuzione di probabilità delle cui componeneti $\mathbb{P}\{\omega\in\Omega\,:\,\xi(\omega)\leq x,\,\eta(\omega)\leq y\}$ è a.c. rispetto alla misura di Lebesgue su $(\mathbb{R}^2,\mathcal{B}(\mathbb{R}^2))$ ed ha densità $f_X(x,y)$, siano $f_\xi(x)$ e $f_\eta(y)$ le densità marginali di ξ e η. Allora, ponendo $f_{\xi|\eta}(x|y):=\frac{f_X(x,y)}{f_\eta(y)}\mathbf{1}_{\{y\in\mathbb{R}\,:\,f_\eta(y)>0\}}$, $\forall A,B\in\mathcal{B}(\mathbb{R})$, si ha

$$\mathbb{P}\{\omega\in\Omega\,:\,\xi(\omega)\in A,\,\eta(\omega)\in B\}=\int_{\{\omega\in\Omega\,:\,\eta(\omega)\in B\}}d\mathbb{P}\mathbf{1}_{\xi^{-1}(A)}=\int_{\{\omega\in\Omega\,:\,\eta(\omega)\in B\}}d\mathbb{P}\mathbb{P}\left(\xi^{-1}(A)\,|\eta\right) \tag{1.143}$$
$$=\int_B dyf_\eta(y)\,\mathbb{P}\left(\xi^{-1}(A)\,|\eta=y\right)=\int_B dyf_\eta(y)\,\mathbb{P}\left(\{\omega\in\Omega:\xi(\omega)\in A\}|\eta=y\right)\,,$$

ma

$$\int_B dyf_\eta(y)\int_A dxf_{\xi|\eta}(x|y)=\int_{A\times B}dxdyf_X(x,y)=\mathbb{P}\{\omega\in\Omega\,:\,X(\omega)\in A\times B\} \tag{1.144}$$
$$=\mathbb{P}\{\omega\in\Omega\,:\,\xi(\omega)\in A,\,\eta(\omega)\in B\}.$$

Perciò,

$$\mathbb{P}\left(\xi^{-1}(A)\,|\eta=y\right)=\int_A dxf_{\xi|\eta}(x|y)=\int_{\mathbb{R}}dx\mathbf{1}_A(x)\,f_{\xi|\eta}(x|y)\;\;\mathbb{P}\text{-q.c. }. \tag{1.145}$$

Procedendo allo stesso modo, $\forall g\in L^1(\mathbb{R},\mathcal{B}(\mathbb{R}),\mathbb{P}_\xi)$, si ottiene

$$\mathbb{E}(g(\xi)\,|\eta=y)=\int_{\mathbb{R}}dxg(x)\,f_{\xi|\eta}(x|y)\;\;\mathbb{P}\text{-q.c. }. \tag{1.146}$$

1.8.2 Versione regolare di una probabilità condizionata

Definizione 79 *Sia* $(\Omega,\mathcal{F},\mathbb{P})$ *uno spazio di probabilità e* $\mathcal{B}$ *una sub*σ*algebra di* $\mathcal{F}$*. La funzione* $\mathcal{F}\times\Omega\ni(A,\omega)\longmapsto P(A|\omega)\in[0,1]$ *è detta* versione regolare *della probabilità condizionata* $\mathbb{P}(\cdot|\mathcal{B})$ *se*

- $\forall\omega\in\Omega$, $\mathcal{F}\ni A\longmapsto P(A|\omega)\in[0,1]$ *è una misura di probabilità su* $(\Omega,\mathcal{F})$*;*
- $\forall A\in\mathcal{F}$, $\Omega\ni\omega\longmapsto P(A|\omega)\in[0,1]$ *è una v.a.* $\mathcal{B}$*-mis. e* $P(A|\cdot)=\mathbb{P}(A|\mathcal{B})$ $\mathbb{P}$*-q.c..*

Proposizione 80 *Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità e* $\mathcal{B}$ *una sub*σ*algebra di* $\mathcal{F}$. *Data una v.a.* ξ *per cui* $\mathbb{E}(\xi)$ *esiste e* P *una versione regolare della probabilità condizionata rispetto a* $\mathcal{B}$, *si ha*

$$\mathbb{E}(\xi|\mathcal{B})(\omega) = \int P(d\omega'|\omega)\,\xi(\omega') \ \ \mathbb{P}\text{-}q.c.. \tag{1.147}$$

Dimostrazione: Se, dato $A \in \mathcal{F}$, $\xi = \mathbf{1}_A$, la definizione precedente implica

$$\mathbb{E}(\mathbf{1}_A|\mathcal{B})(\omega) = \mathbb{P}(A|\mathcal{B})(\omega) = P(A|\omega) = \int_A P(d\omega'|\omega) \ \ \mathbb{P}\text{-q.c.}. \tag{1.148}$$

Conseguentemente, l'affermazione contenuta nella tesi è valida per le v.a. semplici. Poiché $\xi = \xi^+ - \xi^-$ è sufficiente considerare $\xi \geq 0$. In tal caso si può costruire una successione $\{\xi\}_{n\geq 1}$ di v.a. semplici monotona crescente convergente puntualmente a ξ. Quindi, per teorema sulla convergenza monotona

$$\lim_{n\to\infty} \mathbb{E}(\xi_n|\mathcal{B})(\omega) = \lim_{n\to\infty} \int P(d\omega'|\omega)\,\xi_n(\omega') = \int P(d\omega'|\omega)\,\xi(\omega') = \mathbb{E}(\xi|\mathcal{B}) \ \ \mathbb{P}\text{-q.c.}. \tag{1.149}$$

■

Esempio 4 Se il vettore aleatorio $X = (\xi, \eta)$ ha distribuzione a.c. rispetto alla misura di Lebesgue, considerando lo spazio di probabilità $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_\xi)$, con $\frac{\mathbb{P}_\xi(dx)}{dx} = f_\xi(x)$ distribuzione di probabilità marginale della componente ξ di X e $\mathcal{B}_\eta$ la subσalgebra di $\mathcal{F}$ generata dalla componente η di X,

$$\mathcal{B}(\mathbb{R}) \times \mathbb{R} \ni (A, y) \longmapsto P(A|y) = \int_A dx f_{\xi|\eta}(x|y) \in [0,1] \tag{1.150}$$

è una versione regolare della probabilità condizionata $\mathbb{P}\left(\xi^{-1}(A)|\mathcal{B}_\eta\right)$ che è anche a.c. rispetto alla misura di Lebesgue su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Pertanto, $\forall g \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_\xi)$,

$$\mathbb{E}(g(\xi)|\eta = y) = \int_{\mathbb{R}} dx g(x) f_{\xi|\eta}(x|y) \ \ \mathbb{P}_\eta - q.c., \tag{1.151}$$

dove $\mathbb{P}_\eta$ è la distribuzione di probabilità marginale della componente η di X.

Una versione regolare della probabilità condizionata rispetto ad una σalgebra non è altro che un nucleo di probabilità.

Definizione 81 *Data una coppia di spazi misurabili* $(\Omega, \mathcal{F})$ *e* $(E, \mathcal{E})$, *la funzione* $\mathcal{E} \times \Omega \ni (A, \omega) \longmapsto q(A|\omega) \in [0,1]$ *è detta* nucleo di probabilità da $(\Omega, \mathcal{F})$ a $(E, \mathcal{E})$ *se*

- $\forall \omega \in \Omega$, $\mathcal{E} \ni A \longmapsto q(A|\omega) \in [0,1]$ *è una misura di probabilità su* $(E, \mathcal{E})$;
- $\forall A \in \mathcal{E}$, $\Omega \ni \omega \longmapsto q(A|\omega) \in [0,1]$ *è una v.a.* $\mathcal{F}$*-mis. e* $q(E|\cdot) = 1$.

Un nucleo di probabilità da $(\Omega, \mathcal{F})$ a $(E, \mathcal{E})$ mappa la misura di probabilità $\mathbb{P}$ su $(\Omega, \mathcal{F})$ nella misura di probabilità $\mathbb{Q}$ su $(E, \mathcal{E})$ tale che

$$\mathbb{Q}(A) = \int d\mathbb{P}(\omega)\,q(A|\omega), \quad A \in \mathcal{E}. \tag{1.152}$$

Esempio 5 Se $(\Omega, \mathcal{F}, \mathbb{P})$ è una spazio di probabilità e $\Omega \ni \omega \longmapsto \xi(\omega) \in E$ è un applicazione misurabile, allora

$$\mathcal{E} \times \Omega \ni (A, \omega) \longmapsto q_\xi (A|\omega) = (\mathbf{1}_A \circ \xi)(\omega) \in [0,1] \tag{1.153}$$

è il nucleo di probabilità che mappa la misura di probabilità $\mathbb{P}$ su $(\Omega, \mathcal{F})$ nella misura di probabilità $\mathbb{P}_\xi$ su $(E, \mathcal{E})$.

Esempio 6 Se il vettore aleatorio $X = (\xi, \eta)$ ha distribuzione a.c. rispetto alla misura di Lebesgue di densità $f_X(x,y)$, considerando lo spazio di probabilità $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_\eta)$, con $\frac{\mathbb{P}_\eta(dy)}{dy} = f_\eta(y)$ distribuzione di probabilità marginale della componente η di X, l'applicazione

$$\mathcal{B}(\mathbb{R}) \times \mathbb{R} \ni (A, y) \longmapsto q_\xi (A|y) = \int_A dx f_{\xi|\eta}(x|y) \in [0,1] \tag{1.154}$$

è il nucleo di probabilità che mappa la misura di probabilità $\mathbb{P}_\eta$ nella misura di probabilità $\mathbb{P}_\xi$ su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ a.c. rispetto alla misura di Lebesgue di densità $f_\xi(x)$. Infatti,

$$\int_{\mathbb{R}} d\mathbb{P}_\eta q_\xi (A|\cdot) = \int_{\mathbb{R}} dy f_\eta(y) \int_A dx f_{\xi|\eta}(x|y) = \int_{A\times\mathbb{R}} dxdy f_X(x,y) = \int_A dx f_\xi(x) = \mathbb{P}_\xi(A) . \tag{1.155}$$

Definizione 82 *Data $\mathcal{B}$ una subσalgebra di $\mathcal{F}$, un nucleo di probabilità q da $(\Omega, \mathcal{B})$ a $(\Omega, \mathcal{F})$ è detto* proprio *se $\forall B \in \mathcal{B}, \omega \in \Omega$, $q(B|\omega) = \mathbf{1}_B(\omega)$.*

Proposizione 83 *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità e $\mathcal{B}$ una subσalgebra di $\mathcal{F}$. Un nucleo di probabilità q proprio da $(\Omega, \mathcal{B})$ a $(\Omega, \mathcal{F})$ è una versione regolare della probabilità condizionata $\mathbb{P}(\cdot|\mathcal{B})$ se e solo se $\forall A \in \mathcal{F}$, $\mathbb{E}(q(A|\cdot)) = \mathbb{P}(A)$.*

Dimostrazione:

$\Longrightarrow$ Poiché q è proprio ed è per definizione una misura di probabilità su $(\Omega, \mathcal{F})$, dati $A \in \mathcal{F}$, $B \in \mathcal{B}$, si ha

$$\begin{aligned} q((A \cap B)|\omega) &\leq q(A|\omega) \wedge q(B|\omega) = q(A|\omega)\mathbf{1}_B(\omega) , \\ q((A \cap B^c)|\omega) &\leq q(A|\omega) \wedge q(B^c|\omega) = q(A|\omega)\mathbf{1}_{B^c}(\omega) , \end{aligned} \tag{1.156}$$

ma,

$$\begin{aligned} q(A|\omega) &= q((A \cap B) \cup (A \cap B^c)|\omega) = q((A \cap B)|\omega) + q((A \cap B^c)|\omega) \\ &= q(A|\omega)\mathbf{1}_B(\omega) + q(A|\omega)\mathbf{1}_{B^c}(\omega) . \end{aligned} \tag{1.157}$$

Quindi

$$\mathbb{P}(A \cap B) = \mathbb{E}(\mathbf{1}_B q(A|\cdot)) = \mathbb{E}(q(A \cap B|\cdot)) . \tag{1.158}$$

$\Longleftarrow$ Se $\forall A \in \mathcal{F}$, $\mathbb{E}(q(A|\cdot)) = \mathbb{P}(A)$, in particolare vale che $\forall B \in \mathcal{B}$, $\mathbb{P}(A \cap B) = \mathbb{E}(q(A \cap B|\cdot))$, ovvero $q(A|\cdot)$ è una versione di $\mathbb{P}(A|\mathcal{B})$ che è regolare perché è per definizione una misura di probabilità su $(\Omega, \mathcal{F})$.

■

Osservazione 84 *Dato uno spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$ *ed una* σ*algebra* $\mathcal{B} \subset \mathcal{F}$, *se si compone il nucleo di probabilità proprio che rappresenta una versione regolare della probabilità condizionata* $\mathbb{P}(\cdot|\mathcal{B})$ *con il nucleo di probabilità* q_ξ *di cui all'Esempio 4 si ottiene una versione regolare della distribuzione di probabilità condizionata dell'elemento aleatorio* ξ *a valori in* $\mathcal{E}$ *definito su* $(\Omega, \mathcal{F}, \mathbb{P})$ *rispetto alla* σ*algebra* $\mathcal{B}$. *Ovvero,*

$$\mathcal{E} \times \Omega \ni (A, \omega) \longmapsto p_\xi (A|\omega) = \int q(d\omega'|\omega) (\mathbf{1}_A \circ \xi)(\omega') = \int q(d\omega'|\omega) \mathbf{1}_{\xi^{-1}(A)}(\omega') = q\left(\xi^{-1}(A)|\omega\right) , \tag{1.159}$$

$$\int_B d\mathbb{P} p_\xi (A|\cdot) = \int d\mathbb{P} \mathbf{1}_B q\left(\xi^{-1}(A)|\cdot\right) = \int d\mathbb{P} q\left(\xi^{-1}(A) \cap B|\cdot\right) = \mathbb{P}\left(\xi^{-1}(A) \cap B\right) \quad B \in \mathcal{B} , \tag{1.160}$$

cioè $p_\xi (A|\cdot) = \mathbb{P}(\xi \in A|\mathcal{B})$ $\mathbb{P}$*-q.c..*

Definizione 85 *Sia* ξ *una v.a. definita su uno spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$ *e* $\mathcal{B}$ *una sub*σ*algebra di* $\mathcal{F}$. *La funzione* $\mathbb{R} \times \Omega \ni (x, \omega) \longmapsto F(x|\omega) \in [0, 1]$ *è detta* funzione di distribuzione regolare *di* ξ *rispetto a* $\mathcal{B}$ *se:*

- $\forall \omega \in \Omega$, $\mathbb{R} \ni x \longmapsto F(x|\omega) \in [0, 1]$ *è una funzione di distribuzione, ovvero:*
  - $F(\cdot|\omega)$ *è non decrescente;*
  - $\lim_{x \downarrow -\infty} F(x|\omega) = 0$, $\lim_{x \uparrow +\infty} F(x|\omega) = 1$;
  - $F(\cdot|\omega)$ *è continua a destra e* $\forall x \in \mathbb{R}\, \exists \lim_{y \uparrow x} F(y|\omega)$.
- $\forall x \in \mathbb{R}$, $F(x|\cdot) = p_\xi ((-\infty, x]|\cdot)$ $\mathbb{P}$*-q.c..*

Teorema 86 *Sia* ξ *una v.a. definita su uno spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$ *e* $\mathcal{B}$ *una sub*σ*algebra di* $\mathcal{F}$. *Una funzione di distribuzione regolare di* ξ *rispetto a* $\mathcal{B}$ *esiste sempre.*

Dimostrazione: $\forall x \in \mathbb{R}$, sia $F_x = \mathbb{P}(\xi \leq x|\mathcal{B})$ una qualsiasi versione della probabilità condizionata dell'evento $\xi^{-1}((-\infty, x])$, rispetto a $\mathcal{B}$.

- Posto $\mathbb{Q} = \{r_i\}_{i \in \mathbb{Z}}$, siccome $\forall i \in \mathbb{Z}, \{\xi \leq r_i\} \subset \{\xi \leq r_{i+1}\}$, $\mathbf{1}_{\{\xi \leq r_i\}} \leq \mathbf{1}_{\{\xi \leq r_{i+1}\}}$ perciò $F_{r_i} \leq F_{r_{i+1}}$ $\mathbb{P}$-q.c.. Allora, se
$$A_i := \{\omega \in \Omega : F_{r_i} > F_{r_{i+1}}\} , \tag{1.161}$$
$\mathbb{P}(A_i) = 0$ così come $\mathbb{P}(A)$ se $A := \bigcup_{i \in \mathbb{Z}} A_i$ è l'insieme degli $\omega \in \Omega$ per cui $F_r(\omega)$ non è monotona in r.

- Allo stesso modo, sia $\forall i \in \mathbb{Z}$,

$$B_i := \left\{\omega \in \Omega : \lim_{n\to\infty} F_{r_i+\frac{1}{n}}(\omega) \neq F_{r_i}(\omega)\right\}. \tag{1.162}$$

  Poiché $\forall i \in \mathbb{Z}$, la successione $\{\mathbf{1}_{\{\xi\leq r_i+\frac{1}{n}\}}\}_{n\geq 1}$ è monotona decrescente e convergente a $\mathbf{1}_{\{\xi\leq r_i\}}$, per il teorema di convergenza monotona $\lim_{n\to\infty} F_{r_i+\frac{1}{n}}(\omega) = F_{r_i}(\omega)$ $\mathbb{P}$-q.c., perciò $\mathbb{P}(B_i) = 0$ così come $\mathbb{P}(B)$ se $B := \bigcup_{i\in\mathbb{Z}} B_i$ è l'insieme degli $\omega \in \Omega$ per cui $F_r(\omega)$ non è continua a destra in $r \in \mathbb{Q}$.

- Ponendo inoltre

$$C := \left\{\omega \in \Omega : \lim_{r\to+\infty} F_r(\omega) \neq 1\right\} \cup \left\{\omega \in \Omega : \lim_{r\to-\infty} F_r(\omega) > 0\right\}, \tag{1.163}$$

  si ha $\mathbb{P}(C) = 0$, poiché $\{\omega \in \Omega : \xi \leq r\} \uparrow \Omega$ per $r \to +\infty$ e $\{\omega \in \Omega : \xi \leq r\} \downarrow \emptyset$ per $r \to -\infty$.

Definiamo allora $F(x|\omega) = \mathbf{1}_{A^c\cap B^c\cap C^c}(\omega)\lim_{r\downarrow x} F_r(\omega) + \mathbf{1}_{A\cup B\cup C}(\omega) F(x)$ dove F è la funzione di distribuzione di ξ. Per costruzione $F(\cdot|\omega)$ risulta una funzione di distribuzione, inoltre $\forall r \in \mathbb{Q}$, $F(r|\cdot) = \mathbb{P}(\xi \leq r|\mathcal{B})$, ma per la continuità a destra della F nel primo argomento, $\forall\omega \in \Omega$, $\lim_{r\downarrow x} F(r|\omega) = F(x|\omega)$ d'altra parte, per il teorema di convergenza monotona

$$\lim_{r\downarrow x} \mathbb{P}(\xi \leq r|\mathcal{B}) = \mathbb{P}(\xi \leq x|\mathcal{B}) \ \mathbb{P}\text{-q.c.}\ . \tag{1.164}$$

Dunque $F(x|\cdot) = \mathbb{P}(\xi \leq x|\mathcal{B})$ $\mathbb{P}$-q.c.. $\blacksquare$

Formula di Bayes generalizzata

Se $X = (\xi, \eta)$ è un vettore aleatorio che ha distribuzione a.c. rispetto alla misura di Lebesgue di densità $f_X(x, y)$ e $g \in L^1(\Omega, \mathcal{B}(\mathbb{R}), \mathbb{P}_\xi)$ poiché $\forall B \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}_\eta(B) = \int_{\mathbb{R}} d\mathbb{P}_\xi \mathbb{P}(B|\xi) = \int_{\mathbb{R}} dx f_\xi(x) \int_B dy f_{\eta|\xi}(y|x) \tag{1.165}$$

nonché,

$$\begin{aligned}\mathbb{Q}_g(B) &= \int_B d\mathbb{P}_\eta \mathbb{E}(g|\eta) = \int_B dy f_\eta(y) \int_{\mathbb{R}} dx g(x) f_{\xi|\eta}(x|y) = \int_{\mathbb{R}} dx f_\xi(x) g(x) \int_B dy f_{\eta|\xi}(y|x) \\ &= \int_{\mathbb{R}} dx f_\xi(x) g(x) \mathbb{P}(B|\xi = x) = \int_B dy \int_{\mathbb{R}} dx f_\xi(x) f_{\eta|\xi}(y|x) g(x)\ ,\end{aligned} \tag{1.166}$$

allora

$$\mathbb{E}(g|\eta = y) = \frac{d\mathbb{Q}_g}{d\mathbb{P}_\eta}(y) = \frac{\frac{d\mathbb{Q}_g}{dy}}{\frac{d\mathbb{P}_\eta}{dy}}(y) = \frac{\int_{\mathbb{R}} dx f_\xi(x) f_{\eta|\xi}(y|x) g(x)}{\int_{\mathbb{R}} dx f_\xi(x) f_{\eta|\xi}(y|x)} \quad dy\text{-q.c.}\,. \tag{1.167}$$

1.9 Attesa condizionata come operatore di proiezione su un sottospazio chiuso di $L^2(\Omega, \mathcal{F}, \mathbb{P})$

Se $\mathcal{B} \subseteq \mathcal{F}$ è una subσalgebra di $\mathcal{F}$, poniamo $\forall p \in [1, +\infty]$, $L^p(\Omega, \mathcal{B}, \mathbb{P}) := L^p(\mathcal{B})$.

Lemma 87 *Se $\mathcal{B} \subset \mathcal{F}$, $\forall p \in [1, +\infty]$, $L^p(\mathcal{B})$ può essere identificato come sottospazio lineare chiuso di $L^p(\mathcal{F})$.*

Dimostrazione: Poiché $\mathcal{B} \subset \mathcal{F}$, ogni funzione $\mathcal{B}$-mis. è anche $\mathcal{F}$-mis. inoltre, poiché la misura di probabilità su $\mathcal{B}$ è una restrizione di quella data su $\mathcal{F}$, l'integrale sulle funzioni semplici integrabili $E(\mathcal{B})$ è dato dalla restrizione dell'integrale definito su $E(\mathcal{F})$. Munendo $E(\mathcal{B})$ della norma $\|\cdot\|_{L^p}$, si ottiene un'applicazione isometrica da $E(\mathcal{B})$ in $E(\mathcal{F})$. Poiché $E(\mathcal{B})$ è denso in $L^p(\mathcal{B})$ e $L^p(\mathcal{F})$ è completo, questa isometria si estende ad un isometria da $L^p(\mathcal{B})$ in $L^p(\mathcal{F})$. Siccome l'immagine tramite un isometria di uno spazio completo è completa, l'immagine di $L^p(\mathcal{B})$ in $L^p(\mathcal{F})$ è completa ed in particolare è chiusa. ■

Definizione 88 *Sia $\mathbb{E}^{\mathcal{B}}$ la proiezione ortogonale di $L^2(\mathcal{F})$ in $L^2(\mathcal{B})$. Dato $f \in L^2(\mathcal{F})$, $\mathbb{E}^{\mathcal{B}}(f)$ è detta* attesa condizionata *di f rispetto a $\mathcal{B}$.*

1.9.1 Proprietà dell'attesa condizionata

Poiché per definizione $\mathbb{E}^{\mathcal{B}}$ è un proiettore ortogonale, ne possiede tutte le proprietà, in particolare $\forall f \in L^2(\mathcal{F})$:

1. $\mathbb{E}^{\mathcal{B}}(f) \in L^2(\mathcal{B})$;
2. $\left\|\mathbb{E}^{\mathcal{B}}(f)\right\|_{L^2} \leq \|f\|_{L^2}$.

Inoltre se $\mathcal{C}, \mathcal{B}$ sono subσalgebre di $\mathcal{F}$ tali che $\mathcal{C} \subset \mathcal{B}$ allora,

3. $\mathbb{E}^{\mathcal{C}}\mathbb{E}^{\mathcal{B}} = \mathbb{E}^{\mathcal{C}}$ e $\mathbb{E}\mathbb{E}^{\mathcal{B}} = \mathbb{E}$.
   Infatti, poiché $\mathcal{B} \supset \mathcal{C}$ allora $L^2(\mathcal{B}) \supset L^2(\mathcal{C})$ quindi $\forall f \in L^2(\mathcal{F})$
   $$f = u + v, \quad u = \mathbb{E}^{\mathcal{B}}(f), v \in \left(L^2(\mathcal{B})\right)^{\perp}, \tag{1.168}$$
   ma
   $$u = w + h, \quad w = \mathbb{E}^{\mathcal{C}}(f), h \in \left(L^2(\mathcal{C})\right)^{\perp}. \tag{1.169}$$
   Dunque $f = w + h + v$, ma $(L^2(\mathcal{C}))^{\perp} \supset (L^2(\mathcal{B}))^{\perp}$ quindi $v + h \in (L^2(\mathcal{C}))^{\perp}$, perciò
   $$\mathbb{E}^{\mathcal{C}}(f) = w = \mathbb{E}^{\mathcal{C}}(u) = \mathbb{E}^{\mathcal{C}}\left(\mathbb{E}^{\mathcal{B}}(f)\right). \tag{1.170}$$
   Se $\mathcal{A}$ è la σalgebra generata da $\{\Omega, \emptyset\}$, sostituendo $\mathcal{A}$ a $\mathcal{C}$ si ottiene $\mathbb{E}^{\mathcal{A}}\mathbb{E}^{\mathcal{B}} = \mathbb{E}^{\mathcal{A}}$. Notiamo inoltre che f è $\mathcal{A}$-mis. se e solo se f è costante, da cui segue che $(L^2(\mathcal{A}))^{\perp}$ è il sottospazio di $L^2(\mathcal{F})$ corrispondente alle funzioni tali che $\mathbb{E}(f) = 0$. Poiché $\forall f \in L^2(\mathcal{F})$, $f = \mathbb{E}(f)\mathbf{1}_{\Omega} + h$ con h tale che $\mathbb{E}(h) = 0$, $\mathbb{E}^{\mathcal{A}}(f) = \mathbb{E}(f)\mathbf{1}_{\Omega}$ e $\mathbb{E}\mathbb{E}^{\mathcal{A}}(f) = \mathbb{E}(f)$. Quindi, siccome
   $\mathbb{E}^{\mathcal{B}}(f) \in L^2(\mathcal{F})$, $\mathbb{E}\mathbb{E}^{\mathcal{B}}(f) = \mathbb{E}\mathbb{E}^{\mathcal{A}}\mathbb{E}^{\mathcal{B}}(f) = \mathbb{E}(f)$.

Se $\phi \in L^\infty(\mathcal{B})$, allora

4. $\forall f \in L^2(\mathcal{F}),\ \mathbb{E}^{\mathcal{B}}(\phi f) = \phi \mathbb{E}^{\mathcal{B}}(f)$.

   Infatti, sia
   $$L^2(\mathcal{F}) \ni f \longmapsto M_\phi f = \phi f \in L^2(\mathcal{F}) \tag{1.171}$$
   l'operatore lineare limitato corrispondente alla moltiplicazione per ϕ. Poiché ϕ è misurabile e $L(\mathcal{F})$ è un algebra, $M_\phi(L^2(\mathcal{B})) \subset L^2(\mathcal{B})$. Inoltre, M_ϕ è hermitiano, ovvero $\forall f, g \in L^2(\mathcal{F}),\ (M_\phi f, g)_{L^2} = (f, M_\phi g)_{L^2}$ in altri termini $\mathbb{E}(f(\phi g)) = \mathbb{E}((\phi f), g)$. Poiché, $L^2(\mathcal{B})$ è invariante per M_ϕ lo stesso vale per $(L^2(\mathcal{B}))^\perp$. Quindi, data $f \in L^2(\mathcal{F}),\ f = u + v$ con $u = \mathbb{E}^{\mathcal{B}}(f) \in L^2(\mathcal{B}), v \in (L^2(\mathcal{B}))^\perp$ e $M_\phi f = M_\phi u + M_\phi v$, dove $M_\phi u \in L^2(\mathcal{B})$, $M_\phi v \in (L^2(\mathcal{B}))^\perp$. Allora $\mathbb{E}^{\mathcal{B}}(M_\phi f) = M_\phi u = M_\phi \mathbb{E}^{\mathcal{B}}(f)$.

5. Se $f \geq 0$ allora $\mathbb{E}^{\mathcal{B}}(f) \geq 0$.

   Infatti, $\forall B \in \mathcal{B}$, $\mathbf{1}_B \in L^\infty(\mathcal{B})$ quindi, per la proprietà precedente,
   $$\mathbb{E}\left(\mathbb{E}^{\mathcal{B}}(f)\mathbf{1}_B\right) = \mathbb{E}\left(\mathbb{E}^{\mathcal{B}}(f\mathbf{1}_B)\right) = \mathbb{E}(f\mathbf{1}_B) \geq 0\ . \tag{1.172}$$
   Posto $u = \mathbb{E}^{\mathcal{B}}(f)$ e $\forall n \geq 1, B_n := \{\omega \in \Omega : u(\omega) < -\frac{1}{n}\} \in \mathcal{B}$, $0 \leq \mathbb{E}[\mathbb{E}^{\mathcal{B}}(f\mathbf{1}_{B_n})] \leq -\frac{1}{n}\mathbb{P}(B_n)$ il che implica $\mathbb{P}(B_n) = 0, \forall n \geq 1$. Dunque, $\mathbb{P}\left(\bigcup_{n\geq 1} B_n\right) = \lim_{n\to\infty} \mathbb{P}(B_n) = 0$.

   Da ciò segue:

   5.1 $\forall f, g \in L^2(\mathcal{F})$ tali che $f > g$, $\mathbb{E}^{\mathcal{B}}(f) > \mathbb{E}^{\mathcal{B}}(g)$;

   5.2 poiché $-|f| \leq f \leq |f|$, $\left|\mathbb{E}^{\mathcal{B}}(f)\right| \leq \mathbb{E}^{\mathcal{B}}(|f|)$.

1.9.2 Estensione dell'attesa condizionata a $L^1(\Omega, \mathcal{F}, \mathbb{P})$

Teorema 89 *L'operatore lineare limitato $\mathbb{E}^{\mathcal{B}} : L^2(\mathcal{F}) \longrightarrow L^2(\mathcal{B}) \subset L^2(\mathcal{F})$ ammette un'estensione continua $\mathbb{E}_{\mathcal{B}}$ definita su $L^1(\mathcal{F})$ ed a valori in $L^1(\mathcal{B})$ che è una versione di $\mathbb{E}(\cdot|\mathcal{B})$. Tale estensione ha le seguenti proprietà:*

1. *$\forall f \in L^1(\mathcal{B}),\ \mathbb{E}_{\mathcal{B}}(f) = f$;*

2. *$\|\mathbb{E}_{\mathcal{B}}(f)\|_{L^1} \leq \|f\|_{L^1}$;*

3. *se $\mathcal{C}, \mathcal{B}$ sono subσalgebre di $\mathcal{F}$ tali che $\mathcal{C} \subset \mathcal{B}$, allora $\mathbb{E}_{\mathcal{C}}\mathbb{E}_{\mathcal{B}} = \mathbb{E}_{\mathcal{C}}$ ed in particolare $\mathbb{E}\mathbb{E}_{\mathcal{B}}(f) = \mathbb{E}(f)$;*

4. *se $\phi \in L^\infty(\mathcal{B})$, allora $\forall f \in L^1(\mathcal{F}),\ \mathbb{E}_{\mathcal{B}}(\phi f) = \phi \mathbb{E}_{\mathcal{B}}(f)$.*

Dimostrazione: $\forall f\in L^{2}(\mathcal{F})$,

$$\mathbb{E}\left|\mathbb{E}^{\mathcal{B}}(f)\right|\leq\mathbb{E}\left(\mathbb{E}^{\mathcal{B}}(|f|)\right)=\mathbb{E}\mathbb{E}^{\mathcal{A}}\left(\mathbb{E}^{\mathcal{B}}(|f|)\right)=\mathbb{E}(|f|)=\|f\|_{L^{1}}\ , \tag{1.173}$$

ovvero $\left\|\mathbb{E}^{\mathcal{B}}(f)\right\|_{L^{1}}\leq\|f\|_{L^{1}}$. Quindi $\mathbb{E}^{\mathcal{B}}$ è un operatore lineare limitato su $L^{2}(\mathcal{F})$ munito della norma L^{1}. Poiché esiste un insieme denso in $L^{2}(\mathcal{F})$ che è denso anche in $L^{1}(\mathcal{F})$ e $L^{1}(\mathcal{B})$ è completo, si può considerare l'estensione di $\mathbb{E}^{\mathcal{B}}$ come operatore da $L^{1}(\mathcal{F})$ in $L^{1}(\mathcal{B})$ che denotiamo con $\mathbb{E}_{\mathcal{B}}(\cdot)$. Perciò, $\mathbb{E}_{\mathcal{B}}(\cdot)$, in qualità di estensione per continuità di $\mathbb{E}^{\mathcal{B}}$, ne eredita le proprietà. ■

Inoltre, dalle ultime due proprietà $\forall f\in L^{1}(\mathcal{F})$, si ha

$$\mathbb{E}\left(\mathbf{1}_{B}\mathbb{E}_{\mathcal{B}}(f)\right)=\mathbb{E}\left(\mathbb{E}_{\mathcal{B}}(\mathbf{1}_{B}f)\right)=\mathbb{E}\left(\mathbf{1}_{B}f\right)=\mathbb{E}\left(\mathbf{1}_{B}\mathbb{E}(f|\mathcal{B})\right)\quad B\in\mathcal{B}\ , \tag{1.174}$$

cioè $\mathbb{E}_{\mathcal{B}}=\mathbb{E}(\cdot|\mathcal{B})$ $\mathbb{P}$-q.c..

Osservazione 90 *Poiché $\mathbb{E}[\eta|\xi]$ coincide in $L^{2}(\mathcal{F})$ con $\mathbb{E}^{\mathcal{F}^{\xi}}[\eta]$ minimizza la distanza di η da $L^{2}\left(\mathcal{F}^{\xi}\right)$ definita dalla norma indotta dal prodotto scalare in $L^{2}(\mathcal{F})$, cioè $\inf_{\zeta\in L^{2}\left(\mathcal{F}^{\xi}\right)}\|\eta-\zeta\|_{L^{2}}^{2}$. Infatti, siccome*

$$\mathbb{E}\left[(\eta-\mathbb{E}[\eta|\xi])\,|\mathcal{F}_{\xi}\right]=\mathbb{E}[\eta|\xi]-\mathbb{E}[\eta|\xi]=0\ , \tag{1.175}$$

si ha

$$\begin{aligned}\|\eta-\zeta\|_{L^{2}}^{2} &= \|(\eta-\mathbb{E}[\eta|\xi])-(\zeta-\mathbb{E}[\eta|\xi])\|_{L^{2}}^{2}\\
&= \|\eta-\mathbb{E}[\eta|\xi]\|_{L^{2}}^{2}+\|\zeta-\mathbb{E}[\eta|\xi]\|_{L^{2}}^{2}\\
&\quad -2\mathbb{E}\left[(\eta-\mathbb{E}[\eta|\xi])(\zeta-\mathbb{E}[\eta|\xi])\right]\\
&= \|\eta-\mathbb{E}[\eta|\xi]\|_{L^{2}}^{2}+\|\zeta-\mathbb{E}[\eta|\xi]\|_{L^{2}}^{2}\\
&\quad -2\mathbb{E}\left[\mathbb{E}\left[(\eta-\mathbb{E}[\eta|\xi])(\zeta-\mathbb{E}[\eta|\xi])\,|\mathcal{F}_{\xi}\right]\right]\\
&= \|\eta-\mathbb{E}[\eta|\xi]\|_{L^{2}}^{2}+\|\zeta-\mathbb{E}[\eta|\xi]\|_{L^{2}}^{2}\\
&\quad -2\mathbb{E}\left[(\zeta-\mathbb{E}[\eta|\xi])\,\mathbb{E}\left[(\eta-\mathbb{E}[\eta|\xi])\,|\mathcal{F}_{\xi}\right]\right]\\
&= \|\eta-\mathbb{E}[\eta|\xi]\|_{L^{2}}^{2}+\|\zeta-\mathbb{E}[\eta|\xi]\|_{L^{2}}^{2}\\
&\geq \|\eta-\mathbb{E}[\eta|\xi]\|_{L^{2}}^{2}\ ,\end{aligned} \tag{1.176}$$

dove l'uguaglianza è raggiunta soltanto se $\zeta=\mathbb{E}[\eta|\xi]$.

Quanto appena esposto, si può riformulare nel linguaggio della statistica affermando che il miglior stimatore in media quadratica della v.a. η, potendo osservare le occorrenze della v.a. ξ, è la v.a. $\varphi^{*}(\xi)$ con $\varphi^{*}:\mathbb{R}\longrightarrow\mathbb{R},\mathcal{B}(\mathbb{R})$-misurabile, che è una versione di $\mathbb{E}[\eta|\xi]$. Infatti, l'estremo inferiore su tutte le funzioni $\mathcal{B}(\mathbb{R})$-misurabili $\varphi:\mathbb{R}\longrightarrow\mathbb{R}$ di $\|\eta-\varphi(\xi)\|_{L^{2}}^{2}$ è uguale a $\inf_{\zeta\in L^{2}\left(\mathcal{F}_{\xi}\right)}\|\eta-\zeta\|_{L^{2}}^{2}$.

Il grafico della funzione $\mathbb{R}\ni x\longmapsto\varphi^{*}(x)\in\mathbb{R}$ è detta curva di regressione di η rispetto a x.

Proposizione 91 *$\forall f\in L^{1}(\mathcal{F})$, esiste una successione crescente di subσalgebre finite $\mathcal{B}_{1}\subset\cdots\subset\mathcal{B}_{k}\subset\cdots\subset\mathcal{B}\subset\mathcal{F}$ tale che*

$$\|\mathbb{E}(f|\mathcal{B}_{k})-\mathbb{E}(f|\mathcal{B})\|_{L^{1}}\underset{k\rightarrow\infty}{\longrightarrow}0\,. \tag{1.177}$$

Dimostrazione: Sia $u = \mathbb{E}(f|\mathcal{B})$. Poiché $u \in L^1(\mathcal{B})$ esiste una successione $\{u_k\}_{k\geq 1}$ di funzioni semplici in $L^1(\mathcal{B})$ tale che u ne è il limite. Sia $\mathcal{B}_k$ la σalgebra generata dalle u_l, $l \leq k$. Allora $\mathcal{B}_k \subset \mathcal{B}$, u_k è $\mathcal{B}_k$-mis., $\mathbb{E}(u_k|\mathcal{B}_k) = u_k$ e $\|\mathbb{E}(u - u_k|\mathcal{B}_k)\|_{L^1} \leq \|u - u_k\|_{L^1}$. Quindi

$$\begin{aligned} \|\mathbb{E}(u|\mathcal{B}_k) - u\|_{L^1} &\leq \|\mathbb{E}(u|\mathcal{B}_k) - \mathbb{E}(u_k|\mathcal{B}_k)\|_{L^1} + \|\mathbb{E}(u_k|\mathcal{B}_k) - u\|_{L^1} \\ &= \|\mathbb{E}(u - u_k|\mathcal{B}_k)\|_{L^1} + \|u_k - u\|_{L^1} \leq 2\|u_k - u\|_{L^1} \ . \end{aligned} \tag{1.178}$$

Ma $\mathbb{E}(u|\mathcal{B}_k) = \mathbb{E}(\mathbb{E}(f|\mathcal{B})|\mathcal{B}_k) = \mathbb{E}(f|\mathcal{B}_k)$, perciò

$$\|\mathbb{E}(f|\mathcal{B}_k) - \mathbb{E}(f|\mathcal{B})\|_{L^1} \leq 2\|u_k - u\|_{L^1} \underset{k\to\infty}{\longrightarrow} 0\,. \tag{1.179}$$

■

Proposizione 92 *Se $\mathcal{B}$ è una subσalgebra finita di $\mathcal{F}$ i cui atomi sono $e_1, .., e_n$, allora $\forall f \in L^1(\mathcal{F})$*

$$\mathbb{E}(f|\mathcal{B}) = \sum_{k=1}^{n} \frac{\mathbb{E}(f\mathbf{1}_{e_k})}{\mathbb{P}(e_k)}\mathbf{1}_{e_k} \ . \tag{1.180}$$

Dimostrazione: Se $f \in L^2(\mathcal{F})$, $\mathbb{E}^{\mathcal{B}}(f) \in L^2(\mathcal{B}) = span\{v_k,\, k = 1, .., n\}$, $v_k := \frac{\mathbf{1}_{e_k}}{\sqrt{\mathbb{E}(\mathbf{1}_{e_k})}}$. Perciò $\mathbb{E}^{\mathcal{B}}(f) = \sum_{k=1}^{n} \mathbb{E}(fv_k)\,v_k$. La tesi segue dal fatto che $L^2(\mathcal{B})$ è denso in $L^1(\mathcal{B})$. ■

Definizione 93 *Dato $B \in \mathcal{F}$, la misura di probabilità su $(\Omega, \mathcal{F})$*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{E}(\mathbf{1}_A\mathbf{1}_B)}{\mathbb{E}(\mathbf{1}_B)} \quad A \in \mathcal{F} \tag{1.181}$$

è detta probabilità condizionata rispetto all'evento B.

Quindi, in accordo con quanto affermato nell'Esempio 1

$$\mathbb{E}(f|B) := \frac{\mathbb{E}(f\mathbf{1}_B)}{\mathbb{P}(B)} = \frac{\int_{\{\omega\in\Omega\,:\,\mathbf{1}_B(\omega)=1\}} d\mathbb{P}f}{\mathbb{P}(B)} = \int_\Omega \mathbb{P}(d\omega|\mathbf{1}_B = 1)\,f(\omega) \ . \tag{1.182}$$

Definizione 94 *Sia $\{A_k\}_{k\geq 1}$ una successione di eventi esaustiva di Ω, ovvero tale che $\forall k \geq 1$, $A_k \subseteq A_{k+1}$, $\mathbb{P}(A_k) < \infty$, $\bigcup_{k\geq 1} A_k = \Omega$ e, fissato $n \geq 0$, sia*

$$\mathbb{R} \ni x \longmapsto \phi_n = \begin{cases} n & x > n \\ x & -n \leq x \leq n \\ -n & x < -n \end{cases} \in \mathbb{R}\,. \tag{1.183}$$

Allora, $\forall n \geq 0$ e f misurabile, l'applicazione $T_n f := \phi_n(f)\,\mathbf{1}_{A_n} \in L^\infty(\mathcal{F})$ è detta operatore di troncamento.

Proposizione 95 *Sia $p \in [1, +\infty]$ e $\mathcal{B}$ subσalgebra di $\mathcal{F}$. Se $f \in L^p(\mathcal{F})$, $\mathbb{E}(f|\mathcal{B}) \in L^p(\mathcal{B})$ e*

1. $\|\mathbb{E}(f|\mathcal{B})\|_{L^p} \leq \|f\|_{L^p}$;

2. $\forall f \in L^p(\mathcal{F}), \, g \in L^q(\mathcal{B}) : \frac{1}{p} + \frac{1}{q} = 1, \, \mathbb{E}(fg|\mathcal{B}) = g\mathbb{E}(f|\mathcal{B})$;

3. $\forall f \in L^p(\mathcal{F}), \, g \in L^q(\mathcal{F}) : \frac{1}{p} + \frac{1}{q} = 1, \, \mathbb{E}(g\mathbb{E}(f|\mathcal{B})) = \mathbb{E}(f\mathbb{E}(g|\mathcal{B})) = \mathbb{E}(\mathbb{E}(f|\mathcal{B})\,\mathbb{E}(g|\mathcal{B}))$.

Dimostrazione:

1. Poiché la funzione $\mathbb{R}^+ \ni x \longmapsto x^p \in \mathbb{R}^+$ per $p \geq 1$ è convessa, per la disuguaglianza di Jensen si ha che $\|\mathbb{E}(f|\mathcal{B})\|_{L^p} \leq \|f\|_{L^p}$. Inoltre, se $f \in L^\infty(\mathcal{F})$, esiste una successione crescente di subσalgebre finite $\{\mathcal{B}_k\}_{k\geq 1}$ che accresce a $\mathcal{B}$ tale che la successione $\mathbb{E}(f|\mathcal{B}_k)$ converge a $\mathbb{E}(f|\mathcal{B})$ in L^1. Da $\{\mathbb{E}(f|\mathcal{B}_k)\}_{k\geq 1}$ si può estrarre una sottosuccessione $\{\mathbb{E}(f|\mathcal{B}_{k_l})\}_{k_l\geq 1}$ convergente a $\mathbb{E}(f|\mathcal{B})$ $\mathbb{P}$-q.c.. Poiché, $\|\mathbb{E}(f|\mathcal{B}_k)\|_{L^\infty} \leq \|f\|_{L^\infty}$ e $|\mathbb{E}(f|\mathcal{B}_{k_l})| \leq \|f\|_{L^\infty}$ la 1. vale anche per $f \in L^\infty(\mathcal{F})$.

2. Sia $\{A_k\}_{k\geq 1}$ una successione di eventi esaustiva di Ω. Allora, posto $\forall n \geq 0$,
$$f_n := T_n(f) \in L^\infty(\mathcal{F}) \tag{1.184}$$
e la successione $\{f_n\}_{n\geq 1}$ converge a f in L^p. Da ciò segue che anche $\{\mathbb{E}(f_n|\mathcal{B})\}_{n\geq 1}$ converge a $\mathbb{E}(f|\mathcal{B})$ in L^p. Allo stesso modo si ha che la successione $\{g_n\}_{n\geq 1}$ costruita come $\{f_n\}_{n\geq 1}$ converge a g in L^q. Inoltre, siccome $g_n \in L^\infty(\mathcal{B})$, $\mathbb{E}(g_n f_n|\mathcal{B}) = g_n\mathbb{E}(f_n|\mathcal{B})$ e $\{g_n\mathbb{E}(f_n|\mathcal{B})\}_{n\geq 1}$ converge a $g\mathbb{E}(f|\mathcal{B})$ in L^1. Ma poiché $\{g_n f_n\}_{n\geq 1}$ converge a fg in L^1 anche $\{\mathbb{E}(g_n f_n|\mathcal{B})\}_{n\geq 1}$ converge a $\mathbb{E}(gf|\mathcal{B})$ in L^1.

3. Se $p = q = \frac{1}{2}$, l'affermazione è evidente come pure se $p = \infty, q = 1$ o viceversa. Infatti, se $f \in L^\infty(\mathcal{F})$ e $g \in L^1(\mathcal{F})$ per la 1, $\mathbb{E}(f|\mathcal{B}) \in L^\infty(\mathcal{F})$ quindi
$$\mathbb{E}(g\mathbb{E}(f|\mathcal{B})) = \mathbb{E}(\mathbb{E}((g\mathbb{E}(f|\mathcal{B}))|\mathcal{B})) = \mathbb{E}(\mathbb{E}(f|\mathcal{B})\,\mathbb{E}(g|\mathcal{B})) \ . \tag{1.185}$$
Inoltre, $L^\infty(\mathcal{F}) \subset L^2(\mathcal{F})$ e $L^2(\mathcal{F})$ è denso in $L^1(\mathcal{F})$ ed il caso generale si dimostra usando le successioni di funzioni $\{f_n\}_{n\geq 1}, \{g_n\}_{n\geq 1}$ come in precedenza.

■

1.10 Correlazione tra coppie d'eventi

Definizione 96 *Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità. Dato* $B \in \mathcal{F}$*, tale che* $\mathbb{P}(B) > 0$*, la misura di probabilità su* $(\Omega, \mathcal{F})$
$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{E}(\mathbf{1}_A \mathbf{1}_B)}{\mathbb{E}(\mathbf{1}_B)} \quad A \in \mathcal{F} \tag{1.186}$$
è detta probabilità condizionata *rispetto all'evento* B.

Sia $(\Omega, \mathcal{F})$ uno spazio misurabile e $\mathbb{P}$ una misura di probabilità su $(\Omega, \mathcal{F})$. Dati $A, B \in \mathcal{F}$, l'evento A si dice essere *positivamente correlato* a B rispetto a $\mathbb{P}$, se

$$\mathbb{P}(A|B) > \mathbb{P}(A) \ . \tag{1.187}$$

Analogamente, A si dice essere *negativamente correlato* a B rispetto a $\mathbb{P}$, se

$$\mathbb{P}(A|B) < \mathbb{P}(A) \ . \tag{1.188}$$

Nel caso si abbia

$$\mathbb{P}(A|B) = \mathbb{P}(A) \ , \tag{1.189}$$

si dice che A e B sono *stocasticamente indipendenti* o *non correlati* rispetto a $\mathbb{P}$. Pertanto, se $\mathbb{P}(A)$ e $\mathbb{P}(B)$ non sono nulle, quanto appena esposto può riassumersi nella seguente

Definizione 97 *Dato* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità ed* A, B *due eventi tali che* $\mathbb{P}(A)$, $\mathbb{P}(B) > 0$, *allora* A *e* B *si dicono*

- *positivamente correlati se* $\mathbb{P}(A \cap B) > \mathbb{P}(A)\mathbb{P}(B)$;
- *negativamente correlati se* $\mathbb{P}(A \cap B) < \mathbb{P}(A)\mathbb{P}(B)$;
- *non correlati o indipendenti se* $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Osserviamo che se A è positivamente correlato a B, $A^c = \Omega \backslash A$ è negativamente correlato a B e viceversa, in quanto

$$\mathbb{P}(A|B) = 1 - \mathbb{P}(A^c|B) > \mathbb{P}(A) = 1 - \mathbb{P}(A^c) \ . \tag{1.190}$$

Quindi, $\mathbb{P}(A^c|B) < \mathbb{P}(A^c)$. Chiaramente, se A è indipendente da B lo è anche A^c.

Esempio 7 Considerando un urna contenente N palline di cui H bianche, si effettuano due estrazioni. Indicando con E_1 ed E_2 gli eventi in cui sia stata estratta una pallina bianca rispettivamente alla prima ed alla seconda estrazione, nel caso di estrazioni

con reimbussolamento: la composizione dell'urna alla seconda estrazione è identica a quella che si aveva in occasione della prima estrazione. Pertanto,

$$\mathbb{P}(E_1) = \mathbb{P}(E_2) = \frac{H}{N} \ . \tag{1.191}$$

Inoltre, $\mathbb{P}(E_1 \cap E_2) = \left(\frac{H}{N}\right)^2 = \mathbb{P}(E_1)\mathbb{P}(E_2)$, cioè le estrazioni risultano indipendenti;

senza reimbussolamento: la composizione dell'urna alla seconda estrazione risulta variata rispetto all'estrazione precedente. Pertanto,

$$\mathbb{P}(E_1) = \frac{H}{N} \tag{1.192}$$

$$\begin{aligned} \mathbb{P}(E_2) &= \mathbb{P}(E_2|E_1)\mathbb{P}(E_1) + \mathbb{P}(E_2|E_1^c)\mathbb{P}(E_1^c) \\ &= \frac{H-1}{N-1}\frac{H}{N} + \frac{H}{N-1}\frac{N-H}{N} = \frac{H}{N} \ , \end{aligned} \tag{1.193}$$

ma, essendo $H < N, \mathbb{P}(E_2|E_1) = \frac{H-1}{N-1} < \frac{H}{N} = \mathbb{P}(E_2)$. Inoltre, siccome $\mathbb{P}(E_1|E_2) = \mathbb{P}(E_2|E_1)\frac{\mathbb{P}(E_1)}{\mathbb{P}(E_2)} = \mathbb{P}(E_2|E_1)$, si ha $\mathbb{P}(E_1|E_2) < \mathbb{P}(E_1)$.

Le nozioni d'indipendenza e correlazione tra eventi precedentemente introdotte possono essere generalizzate al caso in cui si consideri al posto di $\mathbb{P}$ la probabilità condizionata rispetto ad una σalgebra $\mathcal{B}$, $\mathbb{P}(\cdot|\mathcal{B})$.

Definizione 98 *Dato* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità* $\mathcal{B}$ *una subσalgebra di* $\mathcal{F}$ *ed* A, B *due eventi tali che* $\mathbb{P}(A|\mathcal{B}), \mathbb{P}(B|\mathcal{B}) > 0$, *allora* A *e* B *si dicono*

- *positivamente correlati condizionatamente a* $\mathcal{B}$ *se* $\mathbb{P}(A \cap B|\mathcal{B}) > \mathbb{P}(A|\mathcal{B})\mathbb{P}(B|\mathcal{B})$;
- *negativamente correlati condizionatamente a* $\mathcal{B}$ *se* $\mathbb{P}(A \cap B|\mathcal{B}) < \mathbb{P}(A|\mathcal{B})\mathbb{P}(B|\mathcal{B})$;
- *non correlati o indipendenti condizionatamente a* $\mathcal{B}$ *se* $\mathbb{P}(A \cap B|\mathcal{B}) = \mathbb{P}(A|\mathcal{B})\mathbb{P}(B|\mathcal{B})$.

Esempio 8 Considerando un urna contenente N palline tra bianche e nere, la cui composizione sia incognita, sia H la v.a. che conta il numero di palline bianche nell'urna e $\mathcal{B}$ la σalgebra generata dalla partizione di Ω associata agli eventi A_i, $i = 0, .., N$, corrispondenti ai differenti valori assunti da H. Indicando con E_1 ed E_2 gli eventi in cui sia stata estratta una pallina bianca rispettivamente alla prima ed alla seconda estrazione, dall'esempio precedente segue che, nel caso di estrazione con reinbussolamento, $\forall i = 0, .., N$, subordinatamente all'evento A_i, E_1ed E_2 sono indipendenti ovvero, $\mathbb{P}(E_1 \cap E_2|A_i) = \mathbb{P}(E_1|A_i)\mathbb{P}(E_2|A_i)$. Tuttavia,

$$\begin{aligned}
\mathbb{P}(E_j) &= \sum_{i=0}^{N} \mathbb{P}(E_j|A_i)\mathbb{P}(A_i) = \sum_{i=0}^{N} \frac{i}{N}\mathbb{P}(A_i), \quad j = 1, 2\ , && (1.194) \\
\mathbb{P}(E_1 \cap E_2) &= \sum_{i=0}^{N} \mathbb{P}(E_1 \cap E_2|A_i)\mathbb{P}(A_i) = \sum_{i=0}^{N} \mathbb{P}(E_1|A_i)\mathbb{P}(E_2|A_i)\mathbb{P}(A_i) && (1.195) \\
&= \sum_{i=0}^{N} \left(\frac{i}{N}\right)^2 \mathbb{P}(A_i)\ .
\end{aligned}$$

Nel caso in cui H sia uniformemente distribuita, $\mathbb{P}(H = i) = \frac{1}{N+1}$, si ha

$$\begin{aligned}
\mathbb{P}(E_1) &= \mathbb{P}(E_2) = \sum_{i=0}^{N} \frac{i}{N}\frac{1}{N+1} = \frac{1}{2}\ , && (1.196) \\
\mathbb{P}(E_1 \cap E_2) &= \sum_{i=0}^{N} \left(\frac{i}{N}\right)^2 \frac{1}{N+1} = \frac{1}{3} + \frac{1}{6N} > \frac{1}{4} = \mathbb{P}(E_1)\mathbb{P}(E_2)\ . && (1.197)
\end{aligned}$$

Quindi, l'indipendenza stocastica rispetto ad una σalgebra non implica quella stocastica rispetto a $\mathbb{P}$.

1.11 Indipendenza stocastica ed ortogonalità tra sottospazi di $L^2(\Omega,\mathcal{F},\mathbb{P})$

Definizione 99 *Sia $(\Omega,\mathcal{F},\mathbb{P})$ uno spazio di probabilità. Due subσalgebre di $\mathcal{F}$, $\mathcal{B}$ e $\mathcal{C}$, si dicono* indipendenti *relativamente a $\mathbb{P}$ se $\forall f\in L^2(\mathcal{B})\,,\ g\in L^2(\mathcal{C})$ tali che $\mathbb{E}(f)=\mathbb{E}(g)=0$, $\mathbb{E}(fg)=0$.*

Una volta fissata la misura di probabilità $\mathbb{P}$ su $(\Omega,\mathcal{F})$ quindi, ci siferirà a tale proprietà semplicemente come proprietà d'indipendenza. Nel seguito, laddove non differentemente specificato, considereremo fissato lo spazio di probabilità $(\Omega,\mathcal{F},\mathbb{P})$.

Osservazione 100 *Quanto affermato nella precedente definizione si può tradurre nel modo seguente. Sia $\mathcal{H}:=\{f\in L^2(\mathcal{F}):\mathbb{E}(f)=0\}$ ovvero il sottospazio di $L^2(\mathcal{F})$ ortogonale a $\mathbf{1}_\Omega$. Poiché per la proprietà 3 dell'attesa condizionata ($\forall\mathcal{B}\subset\mathcal{F}$, $\mathbb{E}\mathbb{E}^{\mathcal{B}}=\mathbb{E}$), si ha $\mathbb{E}^{\mathcal{B}}(\mathcal{H})\subset\mathcal{H}$, affermare che $\mathcal{B}$ e $\mathcal{C}$ sono indipendenti rispetto a $\mathbb{P}$ risulta equivalente ad affermare che $\mathcal{H}\cap L^2(\mathcal{B})$ e $\mathcal{H}\cap L^2(\mathcal{C})$ sono tra di loro ortogonali, ovvero che*

$$L^2(\mathcal{B})\cap L^2(\mathcal{C})=\mathcal{H}^\perp:=\{f\in L^2(\mathcal{F}):f=\mathbb{E}(f)\mathbf{1}_\Omega\}\ , \tag{1.198}$$

ovvero, poiché $L^2(\mathcal{B})\cap L^2(\mathcal{C})=L^2(\mathcal{B}\cap\mathcal{C})$, che $\mathcal{B}\cap\mathcal{C}=\mathcal{A}$, la σalgebra generata da $\{\Omega,\emptyset\}$.

Notiamo inoltre che se $\mathcal{B},\mathcal{C}$ sono due subσalgebre di $\mathcal{F}$ indipendenti rispetto a $\mathbb{P}$, $\mathcal{B}'\subset\mathcal{B}$ e $\mathcal{C}'\subset\mathcal{C}$ sono tra di loro indipendenti rispetto a $\mathbb{P}$. Infatti, $\mathcal{H}\cap L^2(\mathcal{B}')\subset\mathcal{H}\cap L^2(\mathcal{B})$ e $\mathcal{H}\cap L^2(\mathcal{C}')\subset\mathcal{H}\cap L^2(\mathcal{C})$ quindi l'ortogonalità tra $\mathcal{H}\cap L^2(\mathcal{B})$ e $\mathcal{H}\cap L^2(\mathcal{C})$ implica quella tra $\mathcal{H}\cap L^2(\mathcal{B}')$ e $\mathcal{H}\cap L^2(\mathcal{C}')$.

Proposizione 101 *Siano $\mathcal{B},\mathcal{C}$ subσalgebre di $\mathcal{F}$, le seguenti affermazioni sono equivalenti:*

1. *$\mathcal{B}$ e $\mathcal{C}$ sono indipendenti;*
2. *$\mathbb{E}(\xi\eta)=\mathbb{E}(\xi)\,\mathbb{E}(\eta)\quad\forall\xi\in L^2(\mathcal{B})\,,\eta\in L^2(\mathcal{C})$.*

Dimostrazione: Poiché ξ e η possono essere decomposte nel modo seguente $\xi=u+\mathbb{E}(\xi)\mathbf{1}_\Omega,\eta=v+\mathbb{E}(\eta)\mathbf{1}_\Omega$ si ha che $u\in\mathcal{H}\cap L^2(\mathcal{B})\,,v\in\mathcal{H}\cap L^2(\mathcal{C})$ perciò l'affermazione 2. è equivalente all'affermazione

$$\mathbb{E}(uv)=0\quad\forall u\in\mathcal{H}\cap L^2(\mathcal{B})\,,v\in\mathcal{H}\cap L^2(\mathcal{C}) \tag{1.199}$$

che implica, per definizione, l'indipendenza di $\mathcal{B},\mathcal{C}$. ■

Teorema 102 *Sia $(\Omega,\mathcal{F},\mathbb{P})$ uno spazio di probabilità e siano $\mathcal{B},\mathcal{C}\subset\mathcal{F}$. Le seguenti affermazioni sono equivalenti:*

1. *$\mathcal{B}$ e $\mathcal{C}$ sono indipendenti;*
2. *$\forall f\in L^1(\mathcal{C})\,,\ \mathbb{E}(f|\mathcal{B})=\mathbb{E}(f)$.*

Dimostrazione:

1. $\Longrightarrow$ 2. Sia $f \in L^2(\mathcal{C})$ e sia $u = f - \mathbb{E}(f)\mathbf{1}_\Omega$. Allora, $u \in \mathcal{H} \cap L^2(\mathcal{C})$ ed è per ipotesi ortogonale ad $L^2(\mathcal{B}) \cap \mathcal{H}$. Dunque, poiché $\mathbb{E}(\cdot|\mathcal{B})$ coincide $\mathbb{P} - q.c.$ con l'estensione in $L^1(\mathcal{B})$ di $\mathbb{E}^{\mathcal{B}}$, proiettore ortogonale su $L^2(\mathcal{B})$, e $L^2(\mathcal{B}) \subset L^1(\mathcal{B})$,
$$0 = \mathbb{E}(u|\mathcal{B}) = \mathbb{E}(f - \mathbb{E}(f)\mathbf{1}_\Omega|\mathcal{B}) = \mathbb{E}(f|\mathcal{B}) - \mathbb{E}(f)\mathbf{1}_\Omega \ , \tag{1.200}$$
cioè $\mathbb{E}(f|\mathcal{B}) = \mathbb{E}(f)\mathbf{1}_\Omega$ ovvero vale la tesi. Nel caso in cui $f \in L^1(\mathcal{C})$, posto $\forall n \geq 1$, $f_n = T_n(f) \in L^\infty(\mathcal{C}) \subset L^2(\mathcal{C})$, $L^2(\mathcal{C})$ è denso in $L^1(\mathcal{C})$ e la successione $\{f_n\}_{n\geq 1}$ converge a f in $L^1(\mathcal{C})$. Quindi, applicando quanto detto in precedenza a f_n e passando al limite per $n \uparrow \infty$, si ha la tesi.

2. $\Longrightarrow$ 1. Sia $f \in L^2(\mathcal{C}) \cap \mathcal{H}$, allora per ipotesi $\mathbb{E}(f|\mathcal{B}) = \mathbb{E}(f) = 0$, da cui segue che $L^2(\mathcal{C}) \cap \mathcal{H}$ è ortogonale a $L^2(\mathcal{B}) \cap \mathcal{H}$ e cioé che $\mathcal{B}$ e $\mathcal{C}$ sono indipendenti.

$\blacksquare$

Corollario 103 *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità e siano $\mathcal{B}$ e $\mathcal{C}$ due subσalgebre di $\mathcal{F}$ tra di loro indipendenti. Allora, $\forall f \in L^1(\mathcal{F})$, $\mathbb{E}(\mathbb{E}(f|\mathcal{C})|\mathcal{B}) = \mathbb{E}(f)$.*

Dimostrazione: Sia $f \in L^1(\mathcal{F})$, allora $\mathbb{E}(f|\mathcal{C}) \in L^1(\mathcal{C})$. Posto $u = \mathbb{E}(f|\mathcal{C})$, $\mathbb{E}(u) = \mathbb{E}(\mathbb{E}(f|\mathcal{C})) = \mathbb{E}(f)$. Ma poiché $u \in L^1(\mathcal{C})$ per il teorema precedente $\mathbb{E}(u|\mathcal{B}) = \mathbb{E}(u) = \mathbb{E}(f)$. $\blacksquare$

Definizione 104 *Due v.a. ξ_1, ξ_2 definite su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ si dicono* mutuamente indipendenti *rispetto a $\mathbb{P}$ se le σalgebre $\mathcal{B}_i := \xi_i^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathcal{F}$, $i = 1, 2$, sono indipendenti rispetto a $\mathbb{P}$.*

Se $\mathcal{B}_1, \mathcal{B}_2 \subset \mathcal{F}$ sono indipendenti, le v.a. X_1, X_2 rispettivamente $\mathcal{B}_1$-mis. e $\mathcal{B}_2$-mis. sono mutuamente indipendenti, in quanto $X_i^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathcal{B}_i, i = 1, 2$. Inoltre, se f_1, f_2 sono $\mathcal{B}(\mathbb{R})$-mis. allora le v.a. $Y_i = f_i \circ X_i, i = 1, 2$ sono anch'esse mutuamente indipendenti, poiché $\forall i = 1, 2$, $Y_i^{-1}(\mathcal{B}(\mathbb{R})) \subset X_i^{-1}\left(f_i^{-1}(\mathcal{B}(\mathbb{R}))\right) \subset X_i^{-1}(\mathcal{B}(\mathbb{R}))$.

Teorema 105 *Siano ξ_1, ξ_2 due v.a. definite su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ ed X il vettore aleatorio di componenti (ξ_1, ξ_2). Le seguenti affermazioni sono equivalenti:*

1. *ξ_1 e ξ_2 sono v.a. indipendenti;*

2. *$\forall \phi_1, \phi_2$ $\mathcal{B}(\mathbb{R})$-mis. e limitate,*
$$\mathbb{E}(\phi_1(\xi_1)\phi_2(\xi_2)) = \mathbb{E}(\phi_1(\xi_1))\mathbb{E}(\phi_2(\xi_2)) \ ; \tag{1.201}$$

3. *$\mathbb{P}_X = \mathbb{P}_{\xi_1} \otimes \mathbb{P}_{\xi_2}$.*

Dimostrazione:

1. $\Longleftrightarrow$ 2. Sia $\mathcal{B}_i = \xi_i^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathcal{F}$, $i = 1, 2$. L'indipendenza di ξ_1, ξ_2 è equivalente a quella di $\mathcal{B}_1, \mathcal{B}_2$. Quindi, se $f_i \in L^2(\mathcal{B}_i)$, $i = 1, 2$, per il teorema della dipendenza funzionale, $\exists \psi_1, \psi_2$ $\mathcal{B}(\mathbb{R})$-mis. tali che $f_i = \psi_i(\xi_i)$, $i = 1, 2$ e l'affermazione 2 è equivalente a

$$\mathbb{E}(\psi_1(\xi_1))\mathbb{E}(\psi_2(\xi_2)) = \mathbb{E}(\psi_1(\xi_1)\psi_2(\xi_2)) \tag{1.202}$$

per ψ_1, ψ_2 tali che $\psi_i \in L^2(\mathbb{R}), i = 1, 2$. Ma, poiché

$$\mathbb{E}(f_i) = \lim_{n\to\infty} \mathbb{E}\left(f_i^{(n)}\right) = \lim_{n\to\infty} \mathbb{E}\left(\psi_i^{(n)}(\xi_i)\right) \quad i = 1, 2 \;, \tag{1.203}$$

$$\mathbb{E}(f_1 f_2) = \lim_{n\to\infty} \mathbb{E}\left(f_1^{(n)} f_2^{(n)}\right) = \lim_{n\to\infty} \mathbb{E}\left(\psi_1^{(n)}(\xi_1)\psi_2^{(n)}(\xi_2)\right) \;, \tag{1.204}$$

con $f_i^{(n)} := T_n(f_i) = \psi_i^{(n)}(f_i)$, dove dalla (1.183) $\psi_i^{(n)} := \phi_n \circ \psi_i$, siccome $f_i^{(n)} \in L^\infty(\mathcal{B}_i) \subset L^2(\mathcal{B}_i), i = 1, 2$, la condizione precedente è equivalente a quella più restrittiva in cui ψ_1 e ψ_2 sono scelte $\mathcal{B}(\mathbb{R})$-mis. e limitate.

2. $\Longrightarrow$ 3. Ponendo, $\forall A_1, A_2 \in \mathcal{B}(\mathbb{R})$, $\phi_1 = \mathbf{1}_{A_1}, \phi_2 = \mathbf{1}_{A_2}$ si ha

$$\mathbb{P}_X(A_1 \times A_2) = \mathbb{E}(\mathbf{1}_{A_1}(\xi_1)\mathbf{1}_{A_2}(\xi_2)) = \mathbb{E}(\mathbf{1}_{A_1}(\xi_1))\mathbb{E}(\mathbf{1}_{A_2}(\xi_2)) = \mathbb{P}_{\xi_1}(A_1)\mathbb{P}_{\xi_2}(A_2) \;. \tag{1.205}$$

Ma poiché $\mathbb{P}_X$ è una misura su $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$, si ha la tesi.

3. $\Longrightarrow$ 2. $\forall \phi_1, \phi_2$ $\mathcal{B}(\mathbb{R})$-mis. e limitate, per il teorema di Fubini, si ha

$$\mathbb{E}(\phi_1(\xi_1)\phi_2(\xi_2)) = \int_{\mathbb{R}^2} d\mathbb{P}_{\xi_1} \otimes d\mathbb{P}_{\xi_2} \phi_1(\xi_1)\phi_2(\xi_2) \tag{1.206}$$

$$= \int_{\mathbb{R}} d\mathbb{P}_{\xi_1}\phi_1(\xi_1) \int_{\mathbb{R}} d\mathbb{P}_{\xi_2}\phi_2(\xi_2) = \mathbb{E}(\phi_1(\xi_1))\mathbb{E}(\phi_2(\xi_2)) \;,$$

cioè la 2..

■

Definizione 106 *Sia $\{\mathcal{B}_i\}_{i=1}^n$ una collezione di subσalgebre di $\mathcal{F}$ e $\forall H \in \mathcal{P}\{1, .., n\}$ sia $\mathcal{B}_H$ la σalgebra generata da $\{\mathcal{B}_i\}_{i\in H}$. Allora gli elementi della collezione $\{\mathcal{B}_i\}_{i=1}^n$ sono detti essere mutuamente indipendenti (rispetto a $\mathbb{P}$) se $\forall H \in \mathcal{P}\{1, .., n\}$, $\mathcal{B}_H$ e $\mathcal{B}_{H^c}$ sono indipendenti.*

Definizione 107 *Sia $\{X_i\}_{i=1}^n$ una collezione di v.a. definite su $(\Omega, \mathcal{F}, \mathbb{P})$, queste si diranno indipendenti se le σalgebre $\mathcal{B}_i := X_i^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathcal{F}$, $i = 1, .., n$, sono mutuamente indipendenti.*

Allora, se $\{\mathcal{B}_i\}_{i=1}^n$ è una famiglia di subσalgebre di $\mathcal{F}$ mutuamente indipendenti, gli elementi di una collezione di v.a. $\{X_i\}_{i=1}^n$ tali che $\forall i = 1, .., n$, X_i sia $\mathcal{B}_i$-mis. sono mutuamente indipendenti, in quanto $\forall i = 1, .., n$, $X_i^{-1}(\mathcal{B}(\mathbb{R})) \subset \mathcal{B}_i$. Inoltre, se $\{X_i\}_{i=1}^n$ è una collezione di v.a. indipendenti e $\{f_i\}_{i=1}^n$ una collezione di funzioni $\mathcal{B}(\mathbb{R})$-mis. allora le v.a. appartenenti alla famiglia $\{Y_i\}_{i=1}^n$ tali che $\forall i = 1, .., n$, $Y_i = f_i \circ X_i$ sono anch'esse mutuamente indipendenti, poiché $\forall i = 1, .., n$, $Y_i^{-1}(\mathcal{B}(\mathbb{R})) \subset X_i^{-1}\left(f_i^{-1}(\mathcal{B}(\mathbb{R}))\right) \subset X_i^{-1}(\mathcal{B}(\mathbb{R}))$.

Teorema 108 *Sia* $\{\mathcal{B}_i\}_{i=1}^n$ *una collezione di subσalgebre di* $\mathcal{F}$*, le seguenti affermazioni sono equivalenti:*

1. *gli elementi di* $\{\mathcal{B}_i\}_{i=1}^n$ *sono mutuamente indipendenti;*
2. $\mathbb{E}\left(\prod_{i=1}^n f_i\right) = \prod_{i=1}^n \mathbb{E}\left(f_i\right) \quad \forall f_i \in L^\infty\left(\mathcal{B}_i\right)\ i = 1, .., n.$

Dimostrazione:

1. $\Longrightarrow$ 2. Procedendo per induzione, supponiamo che la tesi valga per $m < n$ e sia $h = \prod_{i=2}^n f_i$. Sia $\mathcal{B}_H$ la subσalgebra di $\mathcal{F}$ generata dalla collezione di σalgebre $\{f_i^{-1}(\mathcal{B}(\mathbb{R}))\}_{i=2}^n$. Allora, $h \in L^\infty(\mathcal{B}_H)$ e poiché per ipotesi $f_i^{-1}(\mathcal{B}(\mathbb{R}))$ e $\mathcal{B}_H$ sono indipendenti, per il teorema precedente, $\mathbb{E}(hf_1) = \mathbb{E}(h)\,\mathbb{E}(f_1)$ e la tesi segue usando l'induzione su h.

2. $\Longrightarrow$ 1. Sia $H \subset \{1, .., n\}$ e $\mathcal{C}, \mathcal{C}'$ le subσalgebre generate rispettivamente da $\{\mathcal{B}_i\}_{i\in H}$ e $\{\mathcal{B}_i\}_{i\in H^c}$. Bisogna dimostrare che $\mathcal{C}, \mathcal{C}'$ sono indipendenti. A tal fine è sufficiente dimostrare che

$$\mathbb{E}(gg') = \mathbb{E}(g)\,\mathbb{E}(g') \quad \forall g \in L^2(\mathcal{C}),\ g' \in L^2(\mathcal{C}')\ . \tag{1.207}$$

Ma ciò segue dal fatto che, siccome, $\forall K \subset \{1, .., n\}$, se $\mathcal{A}$ è la σalgebra generata da $\{\mathcal{B}_i\}_{i\in K}$,

$$V_{\mathcal{A}} := span\left\{ f \in L^\infty(\mathcal{A}) : \exists m \in \mathbb{N} : f = \sum_{i=1}^m \prod_{j\in K} f_i^{(j)},\ f_i^{(j)} \in L^\infty(\mathcal{B}_j) \right\} \tag{1.208}$$

è denso in $L^2(\mathcal{A})$ e che per ipotesi $\forall f_i \in L^\infty(\mathcal{B}_i)\ i = 1, .., n$,

$$\mathbb{E}\left(\prod_{i\in H} f_i \prod_{i\in H'} f_i\right) = \mathbb{E}\left(\prod_{i=1}^n f_i\right) = \prod_{i=1}^n \mathbb{E}(f_i) = \prod_{i\in H} \mathbb{E}(f_i) \prod_{i\in H'} \mathbb{E}(f_i)\ . \tag{1.209}$$

■

Allo stesso modo, procedendo come nel caso di un vettore aleatorio di due componenti, si ha:

Teorema 109 *Sia* $\{\xi_i\}_{i=1}^n$ *una collezione di v.a. definite su uno spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$ *ed* X *il vettore aleatorio di componenti* $(\xi_1, .., \xi_n)$*. Le seguenti affermazioni sono equivalenti:*

1. *Le v.a.* ξ_i*,* $i = 1, .., n$*, sono tra loro indipendenti;*
2. *per ogni collezione di* $\{\phi_i\}_{i=1}^n$ *di funzioni* $\mathcal{B}(\mathbb{R})$*-mis. e limitate,*

$$\mathbb{E}\left(\prod_{i=1}^n \phi_i(\xi_i)\right) = \prod_{i=1}^n \mathbb{E}(\phi_i(\xi_i))\ ; \tag{1.210}$$

3. $\mathbb{P}_X = \bigotimes_{i=1}^n \mathbb{P}_{\xi_i}.$

Capitolo 2

Convergenza di successioni di variabili aleatorie

Quanto qui di seguito esposto resta valido, laddove non specificato differentemente, anche nel caso in cui si considerino elementi aleatori.

Abbiamo già notato che il limite puntuale di una successione di v.a. sia una v.a. (cfr. Teorema 64).

In questo capitolo illustreremo le altre nozioni di convergenza per successioni di v.a. che sono utili nel Calcolo delle Probabilità e nelle sue applicazioni.

2.1 Convergenza quasi certa

Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità completo.

Definizione 110 *Una successione di v.a.* $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ *converge a* $\xi \in L(\mathcal{F})$ *quasi certamente rispetto a* $\mathbb{P}$ *(*$\mathbb{P}$*-q.c.) se* $\mathbb{P}\{\omega \in \Omega : \lim_{n\to\infty} \xi_n(\omega) \neq \xi(\omega)\} = 0$.

Teorema 111 *La successione di v.a.* $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ *converge a* $\xi \in L(\mathcal{F})$ $\mathbb{P}$*-q.c. se e solo se* $\forall \varepsilon > 0$, $\lim_{n\to\infty} \mathbb{P}\{\omega \in \Omega : \sup_{k\geq n} |\xi_k(\omega) - \xi(\omega)| > \varepsilon\} = 0$.

Dimostrazione: $\forall \varepsilon > 0$ sia $A_k^\varepsilon := \{\omega \in \Omega : |\xi_k(\omega) - \xi(\omega)| > \varepsilon\}$ e $A^\varepsilon := \bigcap_{n\geq 1} \bigcup_{k\geq n} A_k^\varepsilon$. Quindi, $\{\omega \in \Omega : \lim_{n\to\infty} \xi_n(\omega) \neq \xi(\omega)\} = \bigcup_{\varepsilon>0} A^\varepsilon = \bigcup_{m\geq 1} A^{\frac{1}{m}}$. Allora,

$$0 = \mathbb{P}\{\omega \in \Omega : \lim_{n\to\infty} \xi_n(\omega) \neq \xi(\omega)\} = \mathbb{P}\left(\bigcup_{\varepsilon>0} A^\varepsilon\right) = \mathbb{P}\left(\bigcup_{m\geq 1} A^{\frac{1}{m}}\right) \Longleftrightarrow \forall m \geq 1,\, \mathbb{P}\left(A^{\frac{1}{m}}\right) = 0$$

$$\Longleftrightarrow \forall \varepsilon > 0,\, \mathbb{P}(A^\varepsilon) = 0 \Longleftrightarrow \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k\geq n} A_k^\varepsilon\right) = 0 \Longleftrightarrow \lim_{n\to\infty} \mathbb{P}\{\omega \in \Omega : \sup_{k\geq n} |\xi_k(\omega) - \xi(\omega)| > \varepsilon\} = 0.$$

■

Teorema 112 *(di Egorov) La successione di v.a.* $\{\xi_n\}_{n\geq1} \subset L(\mathcal{F})$ *converge a* $\xi \in L(\mathcal{F})$ $\mathbb{P}$*-q.c. se e solo se* $\forall \varepsilon > 0$, $\exists K_\varepsilon \in \mathcal{F}$ *compatto tale che* $\mathbb{P}(K_\varepsilon^c) < \varepsilon$ *e su cui* $\{\xi_n\}_{n\geq1}$ *converge uniformemente a* ξ*.*

Dimostrazione:

$\Leftarrow$ $\{\xi_n\}_{n\geq1}$ converge puntualmente su $G = \bigcup_{m\geq1} K_{\frac{1}{m}}$ quindi, $\forall m \geq 1$, $\mathbb{P}(G^c) \leq \mathbb{P}\left(K^c_{\frac{1}{m}}\right) < \frac{1}{m}$, ovvero $\mathbb{P}(G^c) = 0$.

$\Rightarrow$ $\forall q \geq 1$, sia $A_n^q := \{\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| > \frac{1}{q}\}$, e $B_m^q := \bigcup_{n\geq m} A_n^q$. Poiché $B_m^q \supseteq B_{m+1}^q$ e poiché per ipotesi $\{\xi_i\}_{i\geq1}$ converge a ξ $\mathbb{P}$-q.c., $\forall q \geq 1$,

$$\lim_{m\to\infty} \mathbb{P}(B_m^q) = \mathbb{P}\left(\bigcap_{m\geq1} B_m^q\right) = 0. \tag{2.1}$$

Quindi, $\forall q \geq 1$ e $\varepsilon > 0$ esiste m_ε^q tale che $\mathbb{P}\left(B^q_{m^q_\varepsilon}\right) < \frac{\varepsilon}{2^q}$. Posto $K_\varepsilon := \bigcap_{q\geq1}\left(B^q_{m^q_\varepsilon}\right)^c$, si ha

$$\mathbb{P}(K_\varepsilon^c) = \mathbb{P}\left(\bigcup_{q\geq1} B^q_{m^q_\varepsilon}\right) \leq \sum_{q\geq1} \mathbb{P}\left(B^q_{m^q_\varepsilon}\right) < \sum_{q\geq1} \frac{\varepsilon}{2^q} < \varepsilon, \tag{2.2}$$

inoltre, $\forall \omega \in K_\varepsilon$, $q \geq 1$ e $k \geq m_\varepsilon^q$, $|\xi_k(\omega) - \xi(\omega)| \leq \frac{1}{q}$.

■

Teorema 113 *La successione di v.a.* $\{\xi_n\}_{n\geq1} \subset L(\mathcal{F})$ *è fondamentale* $\mathbb{P}$*-q.c. se e solo se* $\forall \varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left\{\omega \in \Omega : \sup_{\substack{k\geq n\\ l\geq n}} |\xi_k(\omega) - \xi_l(\omega)| > \varepsilon\right\} = 0 \tag{2.3}$$

oppure, equivalentemente, se

$$\lim_{n\to\infty} \mathbb{P}\{\omega \in \Omega : \sup_{k\geq0} \left|\xi_{n+k}(\omega) - \xi_n(\omega)\right| > \varepsilon\} = 0 \ . \tag{2.4}$$

Dimostrazione: Innanzitutto, l'equivalenza delle due affermazioni segue dalla disuguaglianza

$$\begin{aligned}\sup_{k\geq0}\left|\xi_{n+k}(\omega) - \xi_n(\omega)\right| &= \sup_{k\geq0}\left|\xi_{m+l+k}(\omega) - \xi_{m+l}(\omega)\right| \leq \sup_{\substack{k\geq0\\ l\geq0}}\left|\xi_{m+l+k}(\omega) - \xi_{m+l}(\omega)\right| \\ &= \sup_{\substack{k\geq0\\ l\geq0}}\left|\xi_{m+k}(\omega) - \xi_{m+l}(\omega)\right| \leq 2\sup_{k\geq0}\left|\xi_{n+k}(\omega) - \xi_n(\omega)\right|\end{aligned} \tag{2.5}$$

e dal fatto che

$$\sup_{\substack{k\geq0\\ l\geq0}}\left|\xi_{m+k}(\omega) - \xi_{m+l}(\omega)\right| = \sup_{\substack{k\geq m\\ l\geq m}}|\xi_k(\omega) - \xi_l(\omega)| . \tag{2.6}$$

Inoltre, ponendo $A^\varepsilon_{k,l} := \{\omega \in \Omega : |\xi_k(\omega) - \xi_l(\omega)| > \varepsilon\}$, $A^\varepsilon := \bigcap_{n\geq1}\bigcup_{\substack{k\geq n\\ l\geq n}} A^\varepsilon_{k,l}$, si ha che $\{\omega \in \Omega : \{\xi_i(\omega)\}_{i\geq1}$ non è fondamentale$\} = \bigcup_{\varepsilon>0} A^\varepsilon$. Procedendo come nel teorema precedente si ha la tesi. ■

Teorema 114 *La successione di v.a.* $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ *converge ad una v.a.* ξ $\mathbb{P}$*-q.c. se e solo se è fondamentale* $\mathbb{P}$*-q.c.*

Dimostrazione:

$\Rightarrow$ Poiché $\forall \omega \in \Omega$,

$$\sup_{\substack{k\geq n \\ l\geq n}} |\xi_k(\omega) - \xi_l(\omega)| \leq \sup_{k\geq n} |\xi_k(\omega) - \xi(\omega)| + \sup_{l\geq n} |\xi_l(\omega) - \xi(\omega)| \ , \tag{2.7}$$

$\forall \varepsilon > 0$,

$$\left\{ \omega \in \Omega : \sup_{\substack{k\geq n \\ l\geq n}} |\xi_k(\omega) - \xi_l(\omega)| \leq \varepsilon \right\}^c \subset \left\{ \omega \in \Omega : \sup_{k\geq n} |\xi_k(\omega) - \xi_l(\omega)| \leq \frac{\varepsilon}{2} \right\}^c \ . \tag{2.8}$$

Dunque se $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ converge a $\xi \in L(\mathcal{F})$ $\mathbb{P}$-q.c. la tesi segue dal teorema precedente.

$\Leftarrow$ Sia $\{\xi_n\}_{n\geq 1}$ fondamentale $\mathbb{P}$-q.c. e sia $N := \{\omega \in \Omega : \{\xi_n(\omega)\}_{n\geq 1}\text{non è fondamentale}\}$. Poiché una successione numerica è convergente se e solo se è fondamentale, sia

$$\xi(\omega) := \begin{cases} \lim_{n\to\infty} \xi_n(\omega) & \omega \in \Omega\backslash N \\ 0 & \omega \in N \end{cases} \ . \tag{2.9}$$

ξ è una v.a. e $\lim_{n\to\infty} \xi_n = \xi$ $\mathbb{P}$-q.c..

■

Teorema 115 *Condizione sufficiente affinché la successione di v.a.* $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ *converga a* $\xi \in L(\mathcal{F})$ $\mathbb{P}$*-q.c. è che* $\forall \varepsilon > 0$, $\sum_{k\geq 1} \mathbb{P}\{\omega \in \Omega : |\xi_k(\omega) - \xi(\omega)| > \varepsilon\} < \infty$.

Dimostrazione:

$$\begin{aligned} &\lim_{n\to\infty} \mathbb{P}\{\omega \in \Omega : \sup_{k\geq n} |\xi_k(\omega) - \xi(\omega)| > \varepsilon\} = \\ &\lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k\geq n} \{\omega \in \Omega : |\xi_k(\omega) - \xi(\omega)| > \varepsilon\} \right) \leq \\ &\lim_{n\to\infty} \sum_{k\geq n} \mathbb{P}\{\omega \in \Omega : |\xi_k(\omega) - \xi(\omega)| > \varepsilon\} = 0, \end{aligned} \tag{2.10}$$

poiché la serie $\sum_{k\geq 1} \mathbb{P}\{\omega \in \Omega : |\xi_k(\omega) - \xi(\omega)| > \varepsilon\}$ è convergente per ipotesi. ■

In realtà, posto $A_k^\varepsilon := \{\omega \in \Omega : |\xi_k(\omega) - \xi(\omega)| > \varepsilon\}$, la convergenza della serie $\sum_{k\geq 1} \mathbb{P}(A_k^\varepsilon)$ implica che $\mathbb{P}\left(\bigcap_{n\geq 1} \bigcup_{k\geq n} A_k^\varepsilon\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k\geq n} A_k^\varepsilon\right) = 0$. Questo risultato è infatti una conseguenza del risultato più generale che segue:

Lemma 116 *(di Borel-Cantelli) Sia* $\{A_n\}_{n\geq 1} \subset \mathcal{F}$ *e sia* $\{A_n \, i.s.\} := \overline{\lim}_n A_n = \bigcap_{n\geq 1} \bigcup_{k\geq n} A_k$. *Allora,*

1. *se* $\sum_{n\geq 1} \mathbb{P}(A_n) < \infty$, *ne segue che* $\mathbb{P}\{A_n \, i.s.\} = 0$;
2. *se* $\sum_{n\geq 1} \mathbb{P}(A_n) = \infty$ *e inoltre gli eventi* A_n *sono mutuamente indipendenti, ne segue che* $\mathbb{P}\{A_n \, i.s.\} = 1$.

Dimostrazione:

1. Se $\sum_{n\geq 1} \mathbb{P}(A_n) < \infty$, allora
$$\mathbb{P}\left(\bigcap_{n\geq 1} \bigcup_{k\geq n} A_k\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k\geq n} A_k\right) \leq \lim_{n\to\infty} \sum_{k\geq n} \mathbb{P}(A_k) = 0 \; . \tag{2.11}$$

2. $\{A_n \, i.s.\}^c = \bigcup_{n\geq 1} \bigcap_{k\geq n} A_k^c$. Sia $N > n$. Poiché gli eventi A_n^c sono indipendenti, $\mathbb{P}\left(\bigcap_{k=n}^{N} A_k^c\right) = \prod_{k=n}^{N} \mathbb{P}(A_k^c)$. Ma
$$\prod_{k=n}^{N} \mathbb{P}(A_k^c) = e^{\sum_{k=n}^{N} \log \mathbb{P}\left(A_k^c\right)} = e^{\sum_{k=n}^{N} \log(1-\mathbb{P}(A_k))} \leq e^{-\sum_{k=n}^{N} \mathbb{P}(A_k)} \; . \tag{2.12}$$
Quindi, se $\sum_{n\geq 1} \mathbb{P}(A_n) = \infty$, ne consegue che $\mathbb{P}\left(\bigcap_{k\geq n} A_k^c\right) = 0$, perciò $\mathbb{P}\{A_n \, i.s.\}^c = 0$.

■

Corollario 117 *Sia* $\{\varepsilon_n\}_{n\geq 1}$ *una successione di numeri positivi tale che* $\lim_{n\to\infty} \varepsilon_n = 0$ *e* $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$. *Se* $\exists \xi \in L(\mathcal{F})$ *tale che* $\sum_{n\geq 1} \mathbb{P}\{\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon_n\} < \infty$, *allora la successione di v.a.* $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ *converge a* $\xi \in L(\mathcal{F})$ $\mathbb{P}$*-q.c..*

Dimostrazione: Sia $A_n := \{\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon_n\}$. Allora, per il Lemma di Borel-Cantelli, $\mathbb{P}\{A_n \, i.s.\} = 0$. Quindi, per $\mathbb{P}$ quasi ogni $\omega \in \Omega$, $\exists N(\omega)$ tale che $\forall n > N(\omega)$, $|\xi_n(\omega) - \xi(\omega)| \leq \varepsilon_n$. Ma poiché $\varepsilon_n \downarrow 0$, $\xi_n(\omega)$ converge a $\xi(\omega)$ per $\mathbb{P}$ quasi ogni $\omega \in \Omega$. ■

2.2 Convergenza in probabilità

Definizione 118 *La successione di v.a.* $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ *converge in probabilità a* $\xi \in L(\mathcal{F})$ *se,* $\forall \varepsilon > 0$, $\lim_{n\to\infty} \mathbb{P}\{\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon\} = 0$.

Lemma 119 *L'applicazione*
$$L(\mathcal{F}) \times L(\mathcal{F}) \ni (\xi, \eta) \longmapsto \rho(\xi, \eta) := \inf\{\varepsilon > 0 : \mathbb{P}\{\omega \in \Omega : |\xi(\omega) - \eta(\omega)| > \varepsilon\} < \varepsilon\} \in \mathbb{R}^+, \tag{2.13}$$
definisce una distanza su $L(\mathcal{F})$.

Dimostrazione: ρ è simmetrica. Inoltre, se $\xi = \eta\,\mathbb{P}$-q.c., allora, $\forall\varepsilon > 0$,

$$\mathbb{P}\{\omega \in \Omega : |\xi(\omega) - \eta(\omega)| > \varepsilon\} = 0 \tag{2.14}$$

e dunque $\rho(\xi, \eta) = 0$. Sia

$$E(\xi, \eta) := \{\varepsilon > 0 : \mathbb{P}\{\omega \in \Omega : |\xi(\omega) - \eta(\omega)| > \varepsilon\} < \varepsilon\}. \tag{2.15}$$

Se $\varepsilon_1 \in E(\xi, \eta)$, allora anche qualsiasi $\varepsilon_2 > \varepsilon_1$ appartiene a $E(\xi, \eta)$. Infatti,

$$\{\omega \in \Omega : |\xi(\omega) - \eta(\omega)| > \varepsilon_2\} \subset \{\omega \in \Omega : |\xi(\omega) - \eta(\omega)| > \varepsilon_1\} \tag{2.16}$$

perciò,

$$\mathbb{P}\{\omega \in \Omega : |\xi(\omega) - \eta(\omega)| > \varepsilon_2\} \leq \mathbb{P}\{\omega \in \Omega : |\xi(\omega) - \eta(\omega)| > \varepsilon_1\} < \varepsilon_1 < \varepsilon_2 \,. \tag{2.17}$$

Allora, per definizione di ρ, se $\rho(\xi, \eta) = 0$ esiste $\{\varepsilon_n\}_{n\geq 1} \subset E$, decrescente tale che $\varepsilon_n \downarrow 0$ percui $\rho(\xi, \eta) = \lim_{n\to\infty} \varepsilon_n$. Poiché la successione di eventi $\{\Omega_{\varepsilon_n}\}_{n\geq 1} \subset \mathcal{F}$, con

$$\Omega_{\varepsilon_n} := \{\omega \in \Omega : |\xi(\omega) - \eta(\omega)| > \varepsilon_n\} \tag{2.18}$$

è crescente,

$$\mathbb{P}\left(\lim_{n\to\infty} \Omega_{\varepsilon_n}\right) = \mathbb{P}\left(\bigcup_{n\geq 1} \Omega_{\varepsilon_n}\right) = \mathbb{P}\{\omega \in \Omega : |\xi(\omega) - \eta(\omega)| > 0\} = \lim_{n\to\infty} \mathbb{P}(\Omega_{\varepsilon_n}) \leq \lim_{n\to\infty} \varepsilon_n = 0 \;, \tag{2.19}$$

ovvero $\xi = \eta\,\mathbb{P}$-q.c.. Resta da dimostrare la validità della disuguaglianza triangolare. Poiché $|\xi(\omega) - \eta(\omega)| \leq |\xi(\omega) - \zeta(\omega)| + |\zeta(\omega) - \eta(\omega)|$, $\forall\varepsilon_1, \varepsilon_2 > 0$,

$$\begin{gathered}\{\omega \in \Omega : |\xi(\omega) - \zeta(\omega)| \leq \varepsilon_1\} \cap \{\omega \in \Omega : |\eta(\omega) - \zeta(\omega)| \leq \varepsilon_2\} \subset \\ \{\omega \in \Omega : |\xi(\omega) - \eta(\omega)| \leq \varepsilon_1 + \varepsilon_2\} \;.\end{gathered} \tag{2.20}$$

Quindi, se $\varepsilon_1 \in E(\xi, \zeta)$, $\varepsilon_2 \in E(\eta, \zeta)$

$$\begin{gathered}\{\omega \in \Omega : |\xi(\omega) - \eta(\omega)| > \varepsilon_1 + \varepsilon_2\} \subset \\ \{\omega \in \Omega : |\xi(\omega) - \zeta(\omega)| > \varepsilon_1\} \cup \{\omega \in \Omega : |\eta(\omega) - \zeta(\omega)| > \varepsilon_2\} \,,\end{gathered} \tag{2.21}$$

$$\begin{gathered}\mathbb{P}\{\omega \in \Omega : |\xi(\omega) - \eta(\omega)| > \varepsilon_1 + \varepsilon_2\} \leq \mathbb{P}\{\omega \in \Omega : |\xi(\omega) - \zeta(\omega)| > \varepsilon_1\} + \\ + \mathbb{P}\{\omega \in \Omega : |\zeta(\omega) - \eta(\omega)| > \varepsilon_2\} < \varepsilon_1 + \varepsilon_2 \,,\end{gathered} \tag{2.22}$$

il che implica $\varepsilon_1 + \varepsilon_2 \in E(\xi, \eta)$. Posto

$$G(\xi, \eta; \zeta) := \{\varepsilon > 0 : \varepsilon = \varepsilon_1 + \varepsilon_2 \text{ tali che } \varepsilon_1 \in E(\xi, \zeta)\,, \varepsilon_2 \in E(\eta, \zeta)\} \subset E(\xi, \eta) \;, \tag{2.23}$$

si ha

$$\inf E(\xi, \eta) \leq \inf G(\xi, \eta; \zeta) = \inf E(\xi, \zeta) + \inf E(\zeta, \eta) \;. \tag{2.24}$$

■

Dotando $L(\mathcal{F})$ della distanza ρ si ha il seguente risultato:

Proposizione 120 *La successione di v.a.* $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ *converge in probabilità a* $\xi \in L(\mathcal{F})$ *se e solo se* $\lim_{n\to\infty} \rho(\xi_n, \xi) = 0$.

Dimostrazione:

$\Rightarrow$ Se, $\forall \varepsilon > 0$, $\lim_{n\to\infty} \mathbb{P}\{\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon\} = 0$, allora $\exists n_\varepsilon \geq 1$ tale che $\forall n \geq n_\varepsilon$, $\mathbb{P}\{\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon\} < \varepsilon$. Quindi, $\rho(\xi_n, \xi) \leq \varepsilon$.

$\Leftarrow$ Se $\lim_{n\to\infty} \rho(\xi_n, \xi) = 0$, $\forall \delta > 0\, \exists n_\delta \geq 1 : \forall n \geq n_\delta$, $\rho(\xi_n, \xi) < \delta$. Quindi, poiché se $\varepsilon_1 < \varepsilon_2$

$$\{\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon_1\} \supset \{\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon_2\} , \tag{2.25}$$

allora $\exists \varepsilon_\delta : \forall \varepsilon > \varepsilon_\delta$, $\mathbb{P}\{\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon\} < \delta$.

■

Esercizio 1 Dimostrare che la distanza testè introdotta non è associata ad alcuna norma.

Definizione 121 *Una successione di v.a.* $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ *è fondamentale in probabilità se* $\forall \varepsilon > 0$,

$$\lim_{\substack{n\to\infty \\ m\to\infty}} \mathbb{P}\{\omega \in \Omega : |\xi_n(\omega) - \xi_m(\omega)| > \varepsilon\} = 0 \tag{2.26}$$

ovvero se $\lim_{\substack{n\to\infty \\ m\to\infty}} \rho(\xi_n, \xi_m) = 0$.

Teorema 122 *Una successione di v.a.* $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ *fondamentale in probabilità contiene una sottosuccessione fondamentale* $\mathbb{P}$*-q.c..*

Dimostrazione: Poiché una successione di v.a. è fondamentale $\mathbb{P}$-q.c. se e solo se è convergente $\mathbb{P}$-q.c., è sufficiente dimostrare che $\{\xi_i\}_{i\geq 1}$ contiene una sottosuccessione convergente $\mathbb{P}$-q.c.. Sia $n_1 = 1$ e sia $\forall k \geq 2$,

$$n_k := \min\left\{n > n_{k-1} : \frac{1}{2^k} \in E(\xi_l, \xi_m), \forall l, m \geq n\right\} \tag{2.27}$$

con $E(\xi_l, \xi_m)$ definito nella (2.15). Allora,

$$\sum_{k\geq 1} \mathbb{P}\left\{\omega \in \Omega : \left|\xi_{n_{k+1}}(\omega) - \xi_{n_k}(\omega)\right| > \frac{1}{2^k}\right\} \leq \sum_{k\geq 1} \frac{1}{2^k} < \infty . \tag{2.28}$$

Quindi, per il Lemma di Borel-Cantelli,

$$\mathbb{P}\left\{\omega \in \Omega : \left|\xi_{n_{k+1}}(\omega) - \xi_{n_k}(\omega)\right| > \frac{1}{2^k} \quad i.s.\right\} = 0 , \tag{2.29}$$

perciò $\sum_{k\geq 1} \left|\xi_{n_{k+1}} - \xi_{n_k}\right| < \infty\, \mathbb{P}$-q.c.. Sia $N := \left\{\omega \in \Omega : \sum_{k\geq 1} \left|\xi_{n_{k+1}}(\omega) - \xi_{n_k}(\omega)\right| = \infty\right\}$, allora, posto

$$\xi(\omega) := \begin{cases} \xi_{n_1}(\omega) + \sum_{k\geq 1} \left(\xi_{n_{k+1}}(\omega) - \xi_{n_k}(\omega)\right) & \omega \in \Omega \backslash N \\ 0 & \omega \in N \end{cases} \tag{2.30}$$

si ha che $\left\{\xi_{n_k}\right\}_{k\geq 1}$ converge a ξ $\mathbb{P}$-q.c.. Infatti, per la (2.22), poiché $\left|\xi_{n_k}-\xi\right|\leq\sum_{l\geq k+1}\left|\xi_{n_{l+1}}-\xi_{n_l}\right|$ e $\sum_{l\geq k+1}\frac{1}{2^l}=\frac{1}{2^k}$,

$$\begin{aligned}\mathbb{P}\left\{\omega\in\Omega:\left|\xi_{n_k}(\omega)-\xi(\omega)\right|>\frac{1}{2^k}\right\}&\leq\sum_{l\geq k+1}\mathbb{P}\left\{\omega\in\Omega:\left|\xi_{n_{l+1}}(\omega)-\xi_{n_l}(\omega)\right|>\frac{1}{2^l}\right\}\\&<\sum_{l\geq k+1}\frac{1}{2^l}=\frac{1}{2^k}\end{aligned}\tag{2.31}$$

e la tesi segue dal Corollario 117. ■

Teorema 123 *La successione di v.a.* $\{\xi_n\}_{n\geq 1}\subset L(\mathcal{F})$ *converge in probabilità se e solo se è fondamentale in probabilità.*

Dimostrazione:

⇒ Sia $\xi\in L(\mathcal{F})$ il limite di $\{\xi_n\}_{n\geq 1}$. Allora, per la disuguaglianza triangolare si ha $\rho(\xi_n,\xi_m)\leq\rho(\xi_n,\xi)+\rho(\xi,\xi_m)$. Quindi $\{\xi_n\}_{n\geq 1}$ è fondamentale in probabilità.

⇐ Per il teorema precedente esiste una sottosuccessione $\left\{\xi_{n_k}\right\}_{k\geq 1}$ di $\{\xi_n\}_{n\geq 1}$ convergente $\mathbb{P}$-q.c.. ad una v.a. $\xi\in L(\mathcal{F})$. Allora, $\forall\varepsilon>0$,

$$\begin{aligned}\mathbb{P}\{\omega\in\Omega:|\xi_n(\omega)-\xi(\omega)|>\varepsilon\}&\leq\mathbb{P}\left\{\omega\in\Omega:\left|\xi_n(\omega)-\xi_{n_k}(\omega)\right|>\frac{\varepsilon}{2}\right\}+\\+\mathbb{P}\left\{\omega\in\Omega:\left|\xi_{n_k}(\omega)-\xi(\omega)\right|>\frac{\varepsilon}{2}\right\}&\leq\mathbb{P}\left\{\omega\in\Omega:\left|\xi_n(\omega)-\xi_{n_k}(\omega)\right|>\frac{\varepsilon}{2}\right\}+\\&+\mathbb{P}\left(\bigcup_{k\,:\,n_k\geq n}\left\{\omega\in\Omega:\left|\xi_{n_k}(\omega)-\xi(\omega)\right|>\frac{\varepsilon}{2}\right\}\right).\end{aligned}\tag{2.32}$$

Nel limite di $n\to\infty$, il primo termine dell'ultima disuguaglianza tende a zero perché per ipotesi $\{\xi_n\}_{n\geq 1}$ è fondamentale, il secondo perché $\left\{\xi_{n_k}\right\}_{k\geq 1}$ converge a ξ $\mathbb{P}$-q.c.. Pertanto $\{\xi_n\}_{n\geq 1}$ converge in probabilità.

■

Osservazione 124 *Dal precedente risultato segue che* $L(\mathcal{F})$ *munito della metrica* ρ *è uno spazio metrico completo.*

2.3 Convergenza in media di ordine p

Definizione 125 *La successione di v.a.* $\{\xi_n\}_{n\geq 1}\subset L(\mathcal{F})$ *converge a* $\xi\in L(\mathcal{F})$ *in media di ordine p, per* $p\in(0,\infty)$, *se* $\lim_{n\to\infty}\mathbb{E}(|\xi_n-\xi|^p)=0$.

Osservazione 126 *Nel caso in cui $p \geq 1$, l'applicazione*

$$L(\mathcal{F}) \ni \xi \longmapsto \|\xi\|_p = \left(\mathbb{E}\left(|\xi|^p\right)\right)^{\frac{1}{p}} \in \mathbb{R} \ , \tag{2.33}$$

definisce una norma su $L(\mathcal{F})$. Sia $L^p(\mathcal{F}) := L^p(\Omega, \mathcal{F}, \mathbb{P})$ la chiusura di $L(\mathcal{F})$ rispetto alla norma $\|\cdot\|_p$. Allora, la nozione di convergenza in media di ordine p, per $p \geq 1$, coincide con quella usuale di convergenza in L^p. Pertanto, si può considerare anche la nozione di convergenza in $L^\infty(\mathcal{F}) := L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ chiusura $L(\mathcal{F})$ nella norma

$$\|\xi\|_\infty = ess \sup |\xi| := \inf\{c > 0 : \mathbb{P}\{\omega \in \Omega : |\xi(\omega)| > c\} = 0\} \ . \tag{2.34}$$

Lemma 127 *(di Fatou) Sia $\eta \in L^1(\mathcal{F})$ e $\{\xi_n\}_{n\geq 1}$ una successione di v.a. definite su $(\Omega, \mathcal{F}, \mathbb{P})$ convergente puntualmente e tale che, $\forall n \geq 1$, $|\xi_n| \leq \eta$. Allora,*

$$-\infty < \mathbb{E}\left(\underline{\lim}_n \xi_n\right) \leq \underline{\lim}_n \mathbb{E}(\xi_n) \leq \overline{\lim}_n \mathbb{E}(\xi_n) \leq \mathbb{E}\left(\overline{\lim}_n \xi_n\right) < +\infty \ . \tag{2.35}$$

Dimostrazione: Poiché $-\eta \leq \xi_n \leq \eta$, $\xi_n + \eta > 0$. Sia allora

$$\zeta_n := \inf_{k\geq n}(\xi_k + \eta) \leq 2\eta \ . \tag{2.36}$$

$\{\zeta_n\}_{n\geq 1}$ è una successione crescente di v.a. il cui limite puntuale è la v.a.

$$\zeta = \underline{\lim}_n \zeta_n := \sup_{n\geq 1} \xi_n + \eta \ . \tag{2.37}$$

Inoltre, $\forall n \geq 1, \mathbb{E}(\zeta_n) \leq \mathbb{E}\left(\zeta_{n+1}\right) \leq 2\mathbb{E}(\eta)$, pertanto la successione numerica $\{\mathbb{E}(\zeta_n)\}_{n\geq 1}$ è crescente e

$$\lim_{n\to\infty} \mathbb{E}(\zeta_n) = \underline{\lim}_n \mathbb{E}(\zeta_n) \leq \underline{\lim}_n \mathbb{E}(\xi_n + \eta) \leq 2\mathbb{E}(\eta) \ . \tag{2.38}$$

Resta dunque da dimostrare che $\lim_{n\to\infty} \mathbb{E}(\zeta_n) = \mathbb{E}(\lim_{n\to\infty} \zeta_n)$ per cui si avrebbe che

$$\mathbb{E}\left(\underline{\lim}_n (\xi_n + \eta)\right) = \mathbb{E}\left(\lim_{n\to\infty} \zeta_n\right) = \lim_{n\to\infty} \mathbb{E}(\zeta_n) = \underline{\lim}_n \mathbb{E}(\zeta_n) \leq \underline{\lim}_n \mathbb{E}(\xi_n + \eta) \tag{2.39}$$

ovvero,

$$\mathbb{E}\left(\underline{\lim}_n \xi_n\right) + \mathbb{E}(\eta) \leq \underline{\lim}_n \mathbb{E}(\xi_n) + \mathbb{E}(\eta) \tag{2.40}$$

e, poiché $\eta \in L^1(\mathcal{F})$, $\mathbb{E}\left(\underline{\lim}_n \xi_n\right) \leq \underline{\lim}_n \mathbb{E}(\xi_n)$.

$\forall k \geq 1$, sia $\{\zeta_k^{(n)}\}_{n\geq 1}$ una successione crescente di funzioni semplici convergente puntualmente a ζ_k. Posto $\theta_n := \max_{1\leq k\leq n} \zeta_k^{(n)}$, si ha che

$$\theta_n \leq \theta_{n+1} = \max_{1\leq k\leq n+1} \zeta_k^{(n+1)} \leq \max_{1\leq k\leq n+1} \zeta_k = \zeta_{n+1} \tag{2.41}$$

ovvero, $\{\theta_n\}_{n\geq 1}$ è una successione crescente di funzioni semplici convergente puntualmente ad una v.a. θ. Ma, poiché $\forall k, n \geq 1$, $\zeta_k^{(n)} \leq \theta_n \leq \zeta_n$, passando al limite per $n \to \infty$, si ha, $\forall k \geq 1$, $\zeta_k \leq \theta \leq \zeta$, cioè $\theta = \zeta$ e dunque

$$\mathbb{E}(\zeta) = \mathbb{E}(\theta) = \mathbb{E}\left(\lim_{n\to\infty} \theta_n\right) = \lim_{n\to\infty} \mathbb{E}(\theta_n) \leq \lim_{n\to\infty} \mathbb{E}(\zeta_n) \leq \mathbb{E}(\zeta) \tag{2.42}$$

ovvero, $\lim_{n\to\infty} \mathbb{E}(\zeta_n) = \mathbb{E}(\lim_{n\to\infty} \zeta_n)$.

Considerando poi la successione di v.a. $\{-\xi_n\}_{n\geq 1}$ si ha la tesi. ■

Il seguente risultato afferma che $\forall p \geq 1$, $L^p(\mathcal{F})$ è uno spazio di Banach, ovvero è uno spazio lineare normato completo nella metrica indotta dalla norma.

Teorema 128 *Sia $p \geq 1$. La successione di v.a. $\{\xi_n\}_{n\geq 1} \subset L^p(\mathcal{F})$ converge in media di ordine p ad una v.a. $\xi \in L^p(\mathcal{F})$ se e solo se è fondamentale in $L^p(\mathcal{F})$, ovvero se* $\lim_{n,m\to\infty} \|\xi_n - \xi_m\|_p = 0$.

Dimostrazione:

$\Rightarrow$ Sia $\xi \in L^p(\mathcal{F})$ il limite di $\{\xi_n\}_{n\geq 1}$. Allora, per la disuguaglianza triangolare si ha $\|\xi_n - \xi_m\|_p \leq \|\xi_n - \xi\|_p + \|\xi - \xi_m\|_p$. Quindi $\{\xi_n\}_{n\geq 1}$ è fondamentale in $L^p(\mathcal{F})$.

$\Leftarrow$ Sia $n_1 = 1$ e sia $\forall k \geq 2$

$$n_k := \inf\left\{n > n_{k-1} : ||\xi_m - \xi_l||_p \leq \frac{1}{2^{2k}}, \forall m, l \geq n\right\}. \tag{2.43}$$

Posto, $\forall k \geq 1$, $A_k := \left\{\omega \in \Omega : \left|\xi_{n_{k+1}}(\omega) - \xi_{n_k}(\omega)\right| > \frac{1}{2^k}\right\}$, per la disuguaglianza di Chebichev

$$\mathbb{P}(A_k) \leq \frac{\mathbb{E}\left(\left|\xi_{n_{k+1}} - \xi_{n_k}\right|^p\right)}{\frac{1}{2^{pk}}} = \frac{2^{-2pk}}{2^{-pk}} = 2^{-pk}. \tag{2.44}$$

Quindi, $\sum_{k\geq 1} \mathbb{P}(A_k) \leq \sum_{k\geq 1} \frac{1}{2^{pk}} < \infty$, perciò, per il Lemma di Borel-Cantelli, $\mathbb{P}\{A_k\, i.s.\} = 0$ ovvero, $\sum_{k\geq 1} \left|\xi_{n_{k+1}} - \xi_{n_k}\right| < \infty$ e $\left\{\xi_{n_k}\right\}_{k\geq 1} \subset \{\xi_n\}_{n\geq 1}$ converge $\mathbb{P}$-q.c. a

$$\xi(\omega) := \begin{cases} \xi_{n_1}(\omega) + \sum_{k\geq 1}\left(\xi_{n_{k+1}}(\omega) - \xi_{n_k}(\omega)\right) & \omega \in \Omega \backslash N \\ 0 & \omega \in N \end{cases}, \tag{2.45}$$

con $N := \left\{\omega \in \Omega : \sum_{k\geq 1}\left|\xi_{n_{k+1}}(\omega) - \xi_{n_k}(\omega)\right| = \infty\right\} \subseteq \{A_k\, i.s.\}$. Poiché per ipotesi $\{\xi_n\}_{n\geq 1}$ è fondamentale in $L^p(\mathcal{F})$, $\forall \varepsilon > 0$, $\exists N_\varepsilon > 0 : \forall n, m > N_\varepsilon$, $\|\xi_n - \xi_m\|_p^p < \varepsilon$. Allora, per il Lemma di Fatou,

$$\begin{aligned} \mathbb{E}(|\xi_n - \xi|^p) &= \mathbb{E}\left(\lim_{k\to\infty}\left|\xi_n - \xi_{n_k}\right|^p\right) = \mathbb{E}\left(\underline{\lim}_k \left|\xi_n - \xi_{n_k}\right|^p\right) \\ &\leq \underline{\lim}_k \mathbb{E}\left(\left|\xi_n - \xi_{n_k}\right|^p\right) = \underline{\lim}_k \left\|\xi_n - \xi_{n_k}\right\|_p^p < \varepsilon. \end{aligned} \tag{2.46}$$

Dunque, $\lim_{n\to\infty} \|\xi_n - \xi\|_p^p = 0$. Inoltre, $\|\xi\|_p \leq \|\xi_n - \xi\|_p + \|\xi_n\|_p < \infty$, ovvero $\xi \in L^p(\mathcal{F})$.

■

2.4 Convergenza in distribuzione (legge)

Definizione 129 *Sia $\{\xi_n\}_{n\geq 1}$ una successione di v.a. ciascuna definita su uno spazio di probabilità $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ e sia, $\forall n \geq 1$,*

$$\mathbb{R} \ni x \longmapsto F_n(x) := \mathbb{P}_n \{\omega \in \Omega_n : \xi_n(\omega) \leq x\} \in [0,1] \tag{2.47}$$

la funzione di distribuzione (ripartizione) di ξ_n.

La successione $\{\xi_n\}_{n\geq 1}$ converge in distribuzione (legge) *ad una v.a. ξ definita su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ se, indicando con $F(x)$ la funzione di distribuzione di ξ, la successione $\{F_n(x)\}_{n\geq 1}$ converge ad $F(x)$ $\forall x \in \mathcal{C}_F \subset \mathbb{R}$, insieme dei punti di continuità di F.*

Equivalentemente,

la successione di v.a. $\{\xi_n\}_{n\geq 1}$, converge in distribuzione (legge) *ad una v.a. ξ definita su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ se, $\forall f \in C_b(\mathbb{R})$, la successione numerica $\{\mathbb{E}(f(\xi_n))\}_{n\geq 1}$ converge a $\mathbb{E}(f(\xi))$, ovvero se $\{\mathbb{P}_{\xi_n}\}_{n\geq 1}$, successione associata delle distribuzioni di probabilità degli elementi di $\{\xi_n\}_{n\geq 1}$, converge debolmente a $\mathbb{P}_\xi$ distribuzione di probabilità della v.a. ξ.*

È facile convincersi, prendendo ad esempio il teorema di de Moivre-Laplace () e la legge dei grandi numeri per gli schemi di Bernoulli (), che le nozioni di convergenza debole delle distribuzioni di probabilità e convergenza in generale delle funzioni di distribuzione coincidono.

Definizione 130 *Una successione di elementi aleatori $\{\xi_n\}_{n\geq 1}$, ciascuno definito su uno spazio di probabilità $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ ed a valori su uno spazio topologico polacco E, è detta convergere in distribuzione, o debolmente, all'elemento aleatorio ξ definito sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, se la successione di misure di probabilità $\{\mathbb{P}_{\xi_n}\}_{n\geq 1} \subset \mathfrak{P}(E)$ associata a $\{\xi_n\}_{n\geq 1}$ converge debolmente a $\mathbb{P}_\xi \in \mathfrak{P}(E)$.*

Teorema 131 *Sia $\{\xi_n\}_{n\geq 1}$ una successione di v.a. ciascuna definita su uno spazio di probabilità $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ e siano $\{\mathbb{P}_{\xi_n}\}_{n\geq 1}, \{F_{\xi_n}\}_{n\geq 1}$ rispettivamente la successione di distribuzioni probabilità e la successione associata di funzioni di distribuzione. Le seguenti affermazioni sono equivalenti:*

1. *$\{\xi_n\}_{n\geq 1}$ convergente in distribuzione alla v.a. ξ definita su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$;*

2. *$\{\mathbb{P}_{\xi_n}\}_{n\geq 1}$ converge debolmente a $\mathbb{P}_\xi$;*

3. *$\{\mathbb{P}_{\xi_n}\}_{n\geq 1}$ converge in generale a $\mathbb{P}_\xi$;*

4. *$\{F_{\xi_n}\}_{n\geq 1}$ converge in generale a F_ξ.*

Dimostrazione: 1. $\Leftrightarrow$ 2. segue dalla definizione di convergenza in distribuzione. 2. $\Leftrightarrow$ 3. segue dal Teorema 30. Resta da dimostrare 3. $\Leftrightarrow$ 4.

3. $\Rightarrow$ 4. Se $\{\mathbb{P}_{\xi_n}\}_{n\geq 1}$ converge in generale a $\mathbb{P}_\xi$, in particolare $\forall x : \mathbb{P}_\xi\{x\} = 0$,

$$\lim_{n\to\infty} \mathbb{P}_{\xi_n}(-\infty, x] = \mathbb{P}_\xi(-\infty, x], \tag{2.48}$$

ovvero $\{F_{\xi_n}\}_{n\geq 1}$ converge in generale a F_ξ.

4. $\Rightarrow$ 3. Per l'affermazione 3. del Teorema 30 è sufficiente dimostrare che $\forall A \in \mathcal{B}(\mathbb{R})$ aperto $\underline{\lim}_n \mathbb{P}_{\xi_n}(A) \geq \mathbb{P}_\xi(A)$.

Poiché A è aperto, esiste una collezione numerabile d'intervalli aperti disgiunti $\{I_k\}_{k\geq 1}$, con $I_k = (a_k, b_k)$, $k \geq 1$, tale che $A = \bigvee_{k\geq 1} I_k$. Inoltre, poiché i punti di discontinuità di F_ξ formano un insieme al più numerabile, $\forall \varepsilon > 0, k \geq 1$, $\exists a'_k, b'_k \in \mathcal{C}_{F_\xi}$ tali che se $I'_k = (a'_k, b'_k]$, $I'_k \subset I_k$ e $\mathbb{P}_\xi(I'_k) \geq \mathbb{P}_\xi(I_k) - \varepsilon 2^{-k}$. Allora, per il Lemma di Fatou,

$$\begin{aligned}
\underline{\lim}_n \mathbb{P}_{\xi_n}(A) &= \underline{\lim}_n \sum_{k\geq 1} \mathbb{P}_{\xi_n}(I_k) \geq \sum_{k\geq 1} \underline{\lim}_n \mathbb{P}_{\xi_n}(I_k) \geq \sum_{k\geq 1} \underline{\lim}_n \mathbb{P}_{\xi_n}(I'_k) \\
&= \sum_{k\geq 1} \underline{\lim}_n \left(F_{\xi_n}(b'_k) - F_{\xi_n}(a'_k)\right) = \sum_{k\geq 1} \left(F_\xi(b'_k) - F_\xi(a'_k)\right) \\
&= \sum_{k\geq 1} \mathbb{P}_\xi(I'_k) \geq \sum_{k\geq 1} \left(\mathbb{P}_\xi(I_k) - \varepsilon 2^{-k}\right) = \mathbb{P}_\xi(A) - \varepsilon .
\end{aligned} \tag{2.49}$$

■

Osservazione 132 *A differenza delle altre nozioni di convergenza, quella di convergenza in legge non necessita della condizione che i termini della successione di v.a. considerata siano definiti sullo stesso spazio di probabilità, in quanto questa concerne in definitiva la convergenza della successione delle funzioni di distribuzione associata alla sequenza di v.a.. Tuttavia è sempre possibile ridurre anche la nozione di convergenza in distribuzione al caso in cui gli elementi della successione di v.a. considerata siano tutti definiti sullo stesso spazio probabilità. Infatti, se F è la funzione di distribuzione associata alla v.a. ξ, sia*

$$[0,1] \ni u \longmapsto Q(u) := \inf\{x \in \mathbb{R} : F(x) \geq u\} \in \mathbb{R}, \tag{2.50}$$

la funzione quantile *di F. Considerando Q come v.a. su $([0,1], \mathcal{B}([0,1]), du)$ si ha che*

$$\int_{\{u\in[0,1]:Q(u)\leq x\}} du = \int_{\{u\in[0,1]:u\leq F(x)\}} du = F(x), \tag{2.51}$$

ovvero Q è equivalente in distribuzione a ξ. A questo punto, se la successione di v.a. $\{\xi_n\}_{n\geq 1}$ converge a ξ in distribuzione e, $\forall n \geq 1$, F_n, F indicano rispettivamente le funzioni di distribuzione di ξ_n e ξ, i cui quantili sono rispettivamente Q_n e Q, ne segue che la successione di v.a. $\{Q_n\}_{n\geq 1}$ definite su $([0,1], \mathcal{B}([0,1]), du)$ converge in distribuzione alla v.a. Q definita sullo stesso spazio di probabilità ed equivalente in distribuzione a ξ.

In generale se una successione $\{\xi_n\}_{n\geq 1}$ di elementi aleatori a valori su uno spazio topologico polacco E converge in distribuzione ad un elemento aleatorio ξ è sempre possibile trovare uno spazio di probabilità $(X, \mathcal{X}, \mathbb{Q})$ ed una successione $\{X_n\}_{n\geq 1}$ di elementi aleatori su di esso definiti e a valori in E, ciascuno equivalente in distribuzione ad un elemento della successione $\{\xi_n\}_{n\geq 1}$, tale che $\{X_n\}_{n\geq 1}$ converge $\mathbb{Q}$-q.c. all'elemento aleatorio X equivalente in distribuzione a ξ.

2.4.1 Metrizzabilità della convergenza debole

Sia E sia uno spazio metrico.

$\forall \mathbb{P}, \mathbb{Q} \in \mathfrak{P}(E)$ sia

$$\sigma(\mathbb{P}, \mathbb{Q}) := \inf \left\{ \varepsilon > 0 : \mathbb{P}\left(\overline{A}\right) \leq \mathbb{Q}\left(\overline{A}_\varepsilon\right) + \varepsilon, \forall A \in \mathcal{B}(E) \right\} , \tag{2.52}$$

con A_ε definito nella (1.27). L'applicazione:

$$\mathfrak{P}(E) \times \mathfrak{P}(E) \ni (\mathbb{P}, \mathbb{Q}) \longmapsto L(\mathbb{P}, \mathbb{Q}) := \sigma(\mathbb{P}, \mathbb{Q}) \vee \sigma(\mathbb{Q}, \mathbb{P}) \in \mathbb{R}^+ \tag{2.53}$$

è detta *distanza di Lévy-Prokhorov.*

Infatti:

- L è simmetrica.
- $\forall \mathbb{P}, \mathbb{Q} \in \mathfrak{P}(E)$, $\sigma(\mathbb{P}, \mathbb{Q}), \sigma(\mathbb{Q}, \mathbb{P}) \geq 0$ e se $\mathbb{P} = \mathbb{Q}$, $\sigma(\mathbb{P}, \mathbb{Q}) = 0$. Quindi, se $L(\mathbb{P}, \mathbb{Q}) = 0$, allora, $\forall \varepsilon > 0$, $\overline{A} \in \mathcal{B}(E)$, $\mathbb{P}\left(\overline{A}\right) \leq \mathbb{Q}\left(\overline{A}_\varepsilon\right) + \varepsilon$. Passando al limite per $\varepsilon \downarrow 0$, poiché $\overline{A}_\varepsilon \downarrow \overline{A}$ si ha che $\mathbb{P}\left(\overline{A}\right) = \mathbb{Q}\left(\overline{A}\right)$, ovvero $\mathbb{P}\left(\overline{A}^c\right) = \mathbb{Q}\left(\overline{A}^c\right)$ e dunque essendo $\overline{A}$ chiuso e $\overline{A}^c$ aperto, $\mathbb{P}(A) = \mathbb{Q}(A)$, $\forall A \in \mathcal{B}(E)$, cioè $\mathbb{P} = \mathbb{Q}$.
- Se $\mathbb{P}, \mathbb{P}', \mathbb{Q} \in \mathfrak{P}(E)$ sono tali che $L(\mathbb{P}, \mathbb{Q}) = \varepsilon$, $L(\mathbb{P}', \mathbb{Q}) = \eta$, allora

$$\mathbb{P}\left(\overline{A}\right) \leq \mathbb{Q}\left(\overline{A}_\varepsilon\right) + \varepsilon \leq \mathbb{P}'\left(\left(\overline{A}_\varepsilon\right)_\eta\right) + \eta + \varepsilon \leq \mathbb{P}'\left(\overline{A}_{\varepsilon+\eta}\right) + \eta + \varepsilon . \tag{2.54}$$

  Perciò $\sigma(\mathbb{P}, \mathbb{P}') \leq \varepsilon + \eta$. Procedendo allo stesso modo si ha $\sigma(\mathbb{P}', \mathbb{P}) \leq \varepsilon + \eta$, ovvero

$$L(\mathbb{P}, \mathbb{P}') \leq \varepsilon + \eta = L(\mathbb{P}, \mathbb{Q}) + L(\mathbb{P}', \mathbb{Q}) . \tag{2.55}$$

Lemma 133 *Sia $\mathcal{G} \subset C_b(E)$ una collezione di funzioni uniformemente continua ed equilimitata. Se $\{\mathbb{P}_n\}_{n\geq1} \subset \mathfrak{P}(E)$ è una successione di misure di probabilità convergente debolmente a $\mathbb{P} \in \mathfrak{P}(E)$, allora* $\lim_{n\to\infty} \sup_{g\in\mathcal{G}} |\mathbb{E}_n(g) - \mathbb{E}(g)| = 0$.

Dimostrazione: Se $\{\mathbb{P}_n\}_{n\geq1}$ converge debolmente a $\mathbb{P}$ ciò è equivalente a dire che la successione di elementi aleatori $\{X_n\}_{n\geq1}$ associata a $\{\mathbb{P}_n\}_{n\geq1}$ converge in distribuzione all'elemento aleatorio X la cui distribuzione di probabilità è $\mathbb{P}$. Inoltre, per il Teorema 27 $\{\mathbb{P}_n\}_{n\geq1}$ converge a $\mathbb{P}$ anche strettamente. Allora, poiché E è omeoforfo ad uno spazio metrico separabile, esiste uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{Q})$ ed una successione di elementi aleatori $\{\xi_n\}_{n\geq1}$ ivi definiti convergente $\mathbb{Q}$-q.c. a ξ e tali che, $\forall n \geq 1$, ξ_n è equivalente in distribuzione a X_n come ξ lo è a X. Allora, se ϕ è l'applicazione che mappa $\{X_n\}_{n\geq1}, X$ in $\{\xi_n\}_{n\geq1}, \xi$, posto $\forall g \in \mathcal{G}$,

$$\Gamma := \phi_* g = g \circ \phi^{-1} ; \quad \mathbb{M}(\Gamma(\xi)) := \int_\Omega d\mathbb{Q}(\omega) (\Gamma \circ \xi)(\omega) . \tag{2.56}$$

Denotando con ρ una metrica di equivalente su E, poiché $\exists c > 0 : \forall g \in \mathcal{G}$, $|g| < c$ e $\forall \varepsilon > 0$, $\exists \delta_\varepsilon > 0 : \forall x, y \in E$ tale che $\rho(x,y) < \delta_\varepsilon$, $|g(x) - g(y)| < \varepsilon$, si ha

$$\begin{aligned}|\mathbb{M}(\Gamma(\xi_n)) - \mathbb{M}(\Gamma(\xi))| &\leq \mathbb{M}(|\Gamma(\xi_n) - \Gamma(\xi)|) \leq \mathbb{M}\left(|\Gamma(\xi_n) - \Gamma(\xi)| \mathbf{1}_{\{\omega\in\Omega : \rho(\xi_n(\omega),\xi(\omega))\geq\delta_\varepsilon\}}\right) \\ &+ \mathbb{M}\left(|\Gamma(\xi_n) - \Gamma(\xi)| \mathbf{1}_{\{\omega\in\Omega : \rho(\xi_n(\omega),\xi(\omega))<\delta_\varepsilon\}}\right) \leq 2c\mathbb{Q}\{\omega\in\Omega \,:\, \rho(\xi_n(\omega),\xi(\omega)) \geq \delta_\varepsilon\} + \varepsilon\,.\end{aligned}$$

Ma, siccome $\{\xi_n\}_{n\geq1}$ converge a ξ $\mathbb{Q}$-q.c., allora

$$\mathbb{Q}\{\omega\in\Omega \,:\, \rho(\xi_n(\omega),\xi(\omega)) \geq \delta_\varepsilon\} \underset{n\to\infty}{\longrightarrow} 0\,. \tag{2.57}$$

Quindi, $\forall \varepsilon > 0$, $\exists N_\varepsilon > 0 : \forall n > N_\varepsilon$ e $\forall g \in \mathcal{G}$

$$|\mathbb{E}_n(g) - \mathbb{E}(g)| = |\mathbb{M}(\Gamma(\xi_n)) - \mathbb{M}(\Gamma(\xi))| \leq 2\varepsilon\,. \tag{2.58}$$

■

Teorema 134 *Una successione di misure di probabilità* $\{\mathbb{P}_n\}_{n\geq1} \subset \mathfrak{P}(E)$ *converge a* $\mathbb{P} \in \mathfrak{P}(E)$ *nella metrica di Lévy-Prokhorov se e solo se* $\{\mathbb{P}_n\}_{n\geq1}$ *converge debolmente a* $\mathbb{P}$.

Dimostrazione:

- $\Rightarrow$ Se $\lim_{n\to\infty} L(\mathbb{P}_n, \mathbb{P}) = 0$, allora, $\forall A \in \mathcal{B}(E)$ e $\forall \varepsilon > 0$, $\overline{\lim}_n \mathbb{P}_n\left(\overline{A}\right) \leq \mathbb{P}\left(\overline{A}_\varepsilon\right) + \varepsilon$. Passando al limite per $\varepsilon \downarrow 0$ si ha che $\overline{\lim}_n \mathbb{P}_n\left(\overline{A}\right) \leq \mathbb{P}\left(\overline{A}\right)$, ovvero che $\{\mathbb{P}_n\}_{n\geq1}$ converge in generale a $\mathbb{P}$ che, per il Teorema 30, è equivalente a dire che $\{\mathbb{P}_n\}_{n\geq1}$ converge debolmente a $\mathbb{P}$.

- $\Leftarrow$ $\forall \varepsilon > 0$, sia $\mathcal{G}_\varepsilon := \left\{f_A^{(\varepsilon)}\right\}_{A\in\mathcal{B}(E)}$ la collezione delle funzioni definite nella (1.26). Poiché $\forall A \in \mathcal{B}(E), \varepsilon > 0$, $\mathbf{1}_A \leq f_A^{(\varepsilon)} \leq \mathbf{1}_{A_\varepsilon}$, con A_ε definito nella (1.27), e, denotando con ρ una metrica equivalente,

$$\left|f_A^{(\varepsilon)}(x) - f_A^{(\varepsilon)}(y)\right| \leq \frac{1}{\varepsilon}|\rho(x,A) - \rho(y,A)| \leq \frac{1}{\varepsilon}\rho(x,y)\,. \tag{2.59}$$

Pertanto $\forall \varepsilon > 0$, $\mathcal{G}_\varepsilon$ è una famiglia uniformemente continua ed equicontinua quindi, per il lemma precedente

$$\lim_{n\to\infty} \sup_{A\in\mathcal{B}(E)} \left|\mathbb{E}_n\left(f_A^{(\varepsilon)}\right) - \mathbb{E}\left(f_A^{(\varepsilon)}\right)\right| = 0. \tag{2.60}$$

Posto $\forall \varepsilon > 0$,

$$\Delta_n^\varepsilon := \sup_{A\in\mathcal{B}(E)} \left|\mathbb{E}_n\left(f_A^{(\varepsilon)}\right) - \mathbb{E}\left(f_A^{(\varepsilon)}\right)\right|\,, \tag{2.61}$$

$\exists N_\varepsilon > 0 : \forall n > N_\varepsilon$, $\Delta_n^\varepsilon < \varepsilon$. Allora, $\forall n > N_\varepsilon$,

$$\mathbb{P}(A_\varepsilon) = \mathbb{E}(\mathbf{1}_{A_\varepsilon}) \geq \mathbb{E}\left(f_A^{(\varepsilon)}\right) \geq \mathbb{E}_n\left(f_A^{(\varepsilon)}\right) - \Delta_n^\varepsilon \geq \mathbb{E}_n(\mathbf{1}_A) - \varepsilon = \mathbb{P}_n(A) - \varepsilon\,, \tag{2.62}$$

ovvero, per definizione di distanza di Lévy-Prokhorov, ciò è equivalente a dire che $\forall n > N_\varepsilon$, $L(\mathbb{P}_n, \mathbb{P}) \leq \varepsilon$.

■

Consideriamo lo spazio delle funzioni limitate e lipshitziane su E, $BL(E) \subset C_b(E)$, munito della norma

$$\|f\|_{BL} := \sup_{x\in E} |f(x)| + \sup_{x,y\in E\,:\,x\neq y} \frac{|f(x) - f(y)|}{\rho(x,y)} \tag{2.63}$$

e consideriamo su $\mathfrak{P}(E)$ la metrica

$$\|\mathbb{P} - \mathbb{Q}\|^*_{BL} := \sup_{f\in BL(E)\,:\,\|f\|_{BL}\leq 1} \left| \int d\mathbb{P}f - \int d\mathbb{Q}f \right|, \tag{2.64}$$

cioè quella del duale di $BL(E)$.

Teorema 135 *Una successione di misure di probabilità* $\{\mathbb{P}_n\}_{n\geq 1} \subset \mathfrak{P}(E)$ *converge debolmente a* $\mathbb{P} \in \mathfrak{P}(E)$ *se e solo se* $\lim_{n\to\infty} \|\mathbb{P}_n - \mathbb{P}\|^*_{BL} = 0$.

Dimostrazione:

$\Rightarrow$ Segue dal fatto che $C_b(E) \supset BL(E)$.

$\Leftarrow$ Segue dal fatto che, poiché per ipotesi $\lim_{n\to\infty} \|\mathbb{P}_n - \mathbb{P}\|^*_{BL} = 0$, per n sufficientemente grande vale la disuguaglianza $\|\mathbb{P}_n - \mathbb{P}\|^*_{BL} \geq \frac{2}{3}L^2(\mathbb{P}_n, \mathbb{P})$.

■

Osservazione 136 *Quindi per il teorema precedente se al posto di* $C_b(\mathbb{R})$ *si considera come spazio di funzioni test* $BL(\mathbb{R})$, *la successione di v.a.* $\{\xi_n\}_{n\geq 1}$ *converge in distribuzione alla v.a.* ξ *se e solo se la successione* $\{F_n\}_{n\geq 1}$ *converge nella metrica*

$$\|F' - F\|^*_{BL} := \sup_{f\in BL(\mathbb{R})\,:\,\|f\|_{BL}\leq 1} \left| \int_{\mathbb{R}} dF'(x) f(x) - \int_{\mathbb{R}} dF(x) f(x) \right|. \tag{2.65}$$

Alternativamente, si può affermare che la successione di v.a. $\{\xi_n\}_{n\geq 1}$ *converge in distribuzione alla v.a.* ξ *se e solo se la successione* $\{F_n\}_{n\geq 1}$ *converge nella metrica di Lévy-Prokhorov*

$$\rho^{LP}(F', F) := LP(F; F') \vee LP(F'; F), \tag{2.66}$$

dove

$$LP(F; F') := \inf\{\varepsilon > 0 : F'(x-\varepsilon) - \varepsilon \leq F(x) \leq F'(x+\varepsilon) + \varepsilon, \forall x \in \mathbb{R}\}. \tag{2.67}$$

Inoltre, dal Teorema di continuità, segue che la convergenza in legge di una successione di v.a. è equivalente alla convergenza puntuale delle funzioni caratteristiche associate.

2.5 Connessioni tra le varie nozioni di convergenza di successioni di v.a.

2.5.1 Relazione tra convergenza in distribuzione e le altre nozioni di convergenza

Teorema 137 *Sia $\{\xi_n\}_{n\geq 1}$ una successione di v.a. convergente in distribuzione alla v.a. ξ. Allora la successione dei quantili associati $\{Q_n\}_{n\geq 1}$ converge q.c. rispetto alla misura di Lebesgue al quantile di ξ, Q.*

Dimostrazione: Per l'Osservazione 132, la successione $\{Q_n\}_{n\geq 1}$ converge in distribuzione a Q, inoltre, per ipotesi, $\{F_n(x)\}_{n\geq 1}$ converge ad $F(x)$ $\forall x \in \mathcal{C}_F \subset \mathbb{R}$, quindi, per definizione di quantile, $\{Q_n\}_{n\geq 1}$ converge puntualmente a Q in tutti i punti di continuità di Q. Poiché i punti di discontinuità di $Q : [0,1] \longmapsto \mathbb{R}$ sono al più numerabili il loro insieme è trascurabile rispetto alla misura di Lebesgue. Quindi, $\{Q_n\}_{n\geq 1}$ converge Lebsgue-q.c. a Q. ■

Osservazione 138 *Se la successione di v.a. $\{\xi_n\}_{n\geq 1}$ converge in distribuzione ad una v.a. degenere, allora converge anche in probabilità. Infatti, se ξ è degenere $\exists c \in \mathbb{R}$ tale che $\forall f \in C_b(\mathbb{R})$, $\lim_{n\to\infty} \mathbb{E}(f(\xi_n)) = f(c)$. In particolare, posto $\forall \varepsilon > 0$,*

$$f_\varepsilon(x) := \mathbf{1}_{[0,\varepsilon]}(|x-c|) + \left(2 - \frac{|x-c|}{\varepsilon}\right) \mathbf{1}_{(\varepsilon,2\varepsilon]}(|x-c|), \tag{2.68}$$

si ha che $f_\varepsilon \in C_b(\mathbb{R})$ e $f_{\frac{\varepsilon}{2}} < \mathbf{1}_{[c-\varepsilon,c+\varepsilon]}$. Perciò, $\forall \varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\{\omega \in \Omega : |\xi_n(\omega) - c| \leq \varepsilon\} = \lim_{n\to\infty} \mathbb{E}\left(\mathbf{1}_{[c-\varepsilon,c+\varepsilon]}(\xi_n)\right) \geq \lim_{n\to\infty} \mathbb{E}\left(f_{\frac{\varepsilon}{2}}(\xi_n)\right) = 1, \tag{2.69}$$

ovvero $\lim_{n\to\infty} \mathbb{P}\{\omega \in \Omega : |\xi_n(\omega) - c| > \varepsilon\} = 0$.

2.5.2 Relazione tra convergenza in probabilità e le altre nozioni di convergenza

Proposizione 139 *Se la successione di v.a. $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ converge a $\xi \in L(\mathcal{F})$ in probabilità, allora $\{\xi_n\}_{n\geq 1}$ converge a ξ in distribuzione.*

Dimostrazione: $\forall \varepsilon > 0$, sia $N_\varepsilon > 0$ tale che $\mathbb{P}\{\omega \in \Omega : |\xi(\omega)| > N_\varepsilon\} < \varepsilon$. Poiché $[-N_\varepsilon, N_\varepsilon]$ è compatto, ogni funzione continua su $[-N_\varepsilon, N_\varepsilon]$ è limitata e uniformemente continua, quindi $\forall \varepsilon' > 0$, $\exists \delta_{\varepsilon'} > 0$ tale che $\forall x, y \in [-N_\varepsilon, N_\varepsilon] : |x-y| < \delta_{\varepsilon'}$, $|f(x) - f(y)| < \varepsilon'$. Allora, $\forall f \in C_b(\mathbb{R})$ e $\delta > 0$

$$\begin{aligned}|\mathbb{E}[f(\xi_n)] - \mathbb{E}[f(\xi)]| \leq \mathbb{E}[|f(\xi_n) - f(\xi)|] = \\ \mathbb{E}\left[|f(\xi_n) - f(\xi)| \mathbf{1}_{\{\omega\in\Omega : |\xi_n(\omega)-\xi(\omega)|\leq\delta\}}\right] + \mathbb{E}\left[|f(\xi_n) - f(\xi)| \mathbf{1}_{\{\omega\in\Omega : |\xi_n(\omega)-\xi(\omega)|>\delta\}}\right].\end{aligned}$$

Dunque $\forall\varepsilon > 0$,

$$\begin{aligned}\mathbb{E}\left[\left|f\left(\xi_n\right)-f\left(\xi\right)\right|\right] &= \mathbb{E}\left[\left|f\left(\xi_n\right)-f\left(\xi\right)\right|\mathbf{1}_{\{\omega\in\Omega\,:\,|\xi_n(\omega)-\xi(\omega)|\leq\delta_\varepsilon\}}\mathbf{1}_{\xi^{-1}([-N_\varepsilon,N_\varepsilon]^c)}\right]+ \\ &+\mathbb{E}\left[\left|f\left(\xi_n\right)-f\left(\xi\right)\right|\mathbf{1}_{\{\omega\in\Omega\,:\,|\xi_n(\omega)-\xi(\omega)|\leq\delta_\varepsilon\}}\mathbf{1}_{\xi^{-1}([-N_\varepsilon,N_\varepsilon])}\right]+ \\ &+\mathbb{E}\left[\left|f\left(\xi_n\right)-f\left(\xi\right)\right|\mathbf{1}_{\{\omega\in\Omega\,:\,|\xi_n(\omega)-\xi(\omega)|>\delta_\varepsilon\}}\right] \\ &\leq 2\sup_{x\in\mathbb{R}}\left|f\left(x\right)\right|\varepsilon+\varepsilon+2\sup_{x\in\mathbb{R}}\left|f\left(x\right)\right|\mathbb{P}\{\omega\in\Omega\,:\,\left|\xi_n\left(\omega\right)-\xi\left(\omega\right)\right|>\delta_\varepsilon\},\end{aligned} \tag{2.70}$$

ma poiché per ipotesi $\lim_{n\to\infty}\mathbb{P}\{\omega\in\Omega\,:\,\left|\xi_n\left(\omega\right)-\xi\left(\omega\right)\right|>\delta_\varepsilon\}=0$, esiste $c>0$ tale che $\mathbb{E}\left[\left|f\left(\xi_n\right)-f\left(\xi\right)\right|\right]\leq c\varepsilon$. ■

Osservazione 140 (La convergenza in probabilità non implica quella quasi certa) *Si consideri sullo spazio di Probabilità* $\left([0,1],\mathcal{B}\left([0,1]\right),dx\right)$ *la successione di v.a.* $\{\xi_n\}_{n\geq1}=\{\{\eta_k^i\}_{i=1}^k\}_{k\geq1}$ *tali che* $\forall k\geq1, 1\leq i\leq k$, $\eta_k^i:=\mathbf{1}_{[\frac{i-1}{k},\frac{i}{k}]}$. *Allora,* $\{\xi_n\}_{n\geq1}$ *converge alla v.a. degenere identicamente uguale a* 0 *in probabilità ed in media di ordine* p, $\forall p>0$, *ma non converge puntualmente in nessun punto di* $[0,1]$.

Osservazione 141 (La convergenza in probabilità non implica quella in media di ordine p) *Si consideri sullo spazio di Probabilità* $\left([0,1],\mathcal{B}\left([0,1]\right),dx\right)$ *la successione di v.a.* $\{\xi_n\}_{n\geq1}$ *tale che* $\forall n\geq1$, $\xi_n:=e^n\mathbf{1}_{[0,\frac{1}{n}]}$. *Allora* $\{\xi_n\}_{n\geq1}$ *converge Lebesgue-q.c., quindi in probabilità, ma non in media di ordine p poiché* $\lim_{n\to\infty}\mathbb{E}\left(|\xi_n|^p\right)=\lim_{n\to\infty}\frac{e^{np}}{n}=\infty$, $\forall p>0$.

2.5.3 Relazione tra convergenza in media di ordine p e le altre nozioni di convergenza

Proposizione 142 *Se la successione di v.a.* $\{\xi_n\}_{n\geq1}\subset L\left(\mathcal{F}\right)$ *converge a* $\xi\in L\left(\mathcal{F}\right)$ *in media di ordine p,* $p>0$, *allora* $\{\xi_n\}_{n\geq1}$ *converge a* ξ *in probabilità.*

Dimostrazione: Per la disuguaglianza di Chebichev generalizzata, $\forall\varepsilon>0$, si ha

$$\mathbb{P}\left\{\omega\in\Omega:\left|\xi_n\left(\omega\right)-\xi\left(\omega\right)\right|>\varepsilon\right\}\leq\frac{\mathbb{E}\left[|\xi_n-\xi|^p\right]}{\varepsilon^p}. \tag{2.71}$$

■

Osservazione 143 *Se la successione di v.a.* $\{\xi_n\}_{n\geq1}\subset L\left(\mathcal{F}\right)$ *converge a* $\xi\in L\left(\mathcal{F}\right)$ *in media di ordine* $p>0$, *poiché* $\{\xi_n\}_{n\geq1}$ *converge a* ξ *anche in probabilità, allora:*

1. *per il Teorema 122,* $\{\xi_n\}_{n\geq1}$ *ammette una sottosuccessione* $\{\xi_{n_k}\}_{k\geq1}$ *convergente a* ξ $\mathbb{P}$*-q.c.;*
2. *per la Proposizione 145,* ξ *converge anche in distribuzione.*

Osservazione 144 (La convergenza in media di ordine p non implica quella quasi certa) *Dall'Osservazione 140, s'evince che la convergenza in media di ordine p non implica quella quasi certa. Un'ulteriore esempio di ciò è la successione* $\{\xi_n\}_{\geq 0}$ *di v.a. di Bernoulli, definite sullo spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$ *tali che* $\mathbb{P}\{\omega \in \Omega : \xi_n(\omega) = 1\} = p_n$. *Allora,* $\{\xi_n\}_{\geq 0}$ *converge alla v.a. degenere identicamente uguale a* 0 *in media di ordine* p, $\forall p > 0$, *e quindi in probabilità, se* $\lim_{n\to\infty} p_n = 0$, *mentre convegerebbe allo stesso limite* $\mathbb{P}$*-q.c. se* $\sum_{n\geq 1} p_n < \infty$, *in quanto*

$$\lim_{n\to\infty} \mathbb{P}\{\omega \in \Omega : \sup_{k\geq n} \xi_n(\omega) = 1\} = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k\geq n} \{\omega \in \Omega : \xi_n(\omega) = 1\}\right) \leq \sum_{k\geq n} p_n \ . \tag{2.72}$$

In particolare se $p_n = \frac{1}{n}$, $\{\xi_n\}_{\geq 0}$ *non convergerebbe* $\mathbb{P}$*-q.c..*

2.5.4 Relazione tra convergenza $\mathbb{P}$-q.c. e le altre nozioni di convergenza

Proposizione 145 *Se la successione di v.a.* $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ *converge a* $\xi \in L(\mathcal{F})$ $\mathbb{P}$*-q.c., allora* $\{\xi_n\}_{n\geq 1}$ *converge a* ξ *in probabilità.*

Dimostrazione: $\{\xi_n\}_{n\geq 1}$ converge a ξ $\mathbb{P}$-q.c. se e solo se $\forall \varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left\{\omega \in \Omega : \sup_{k\geq n} |\xi_k(\omega) - \xi(\omega)| > \varepsilon\right\} = 0\,, \tag{2.73}$$

il che implica che $\forall \varepsilon > 0$, $\lim_{n\to\infty} \mathbb{P}\{\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| > \varepsilon\} = 0$. ■

Osservazione 146 *Dall'Osservazione 141 s'evince che la convergenza quasi certa non implica quella in media di ordine p.*

Relazione tra convergenza $\mathbb{P}$-q.c. e convergenza L^1

Definizione 147 *Una famiglia di v.a.* $\{\xi_i\}_{i\in\mathcal{I}}$ *definite su* $(\Omega, \mathcal{F}, \mathbb{P})$ *è detta* uniformemente integrabile *se* $\forall \varepsilon > 0, \exists \delta_\varepsilon > 0$ *tale che* $\forall A \in \mathcal{F}$ *per cui* $\mathbb{P}(A^c) < \delta_\varepsilon, \forall i \in \mathcal{I}$, $\mathbb{E}(|\xi_i| \mathbf{1}_{A^c}) < \varepsilon$.

Proposizione 148 *Una famiglia di v.a.* $\{\xi_i\}_{i\in\mathcal{I}}$ *definite su* $(\Omega, \mathcal{F}, \mathbb{P})$ *è uniformemente integrabile se e solo se* $\lim_{N\to\infty} \sup_{i\in\mathcal{I}} \mathbb{E}\left(|\xi_i| \mathbf{1}_{\xi_i^{-1}([-N,N]^c)}\right) = 0$.

Dimostrazione:

$\Rightarrow$ Poiché $\{\xi_i\}_{i\in\mathcal{I}}$ è uniformemente integrabile $\forall \varepsilon > 0$, $\exists \delta_\varepsilon, N_\varepsilon > 0$ tali che $\forall i \in \mathcal{I}$, $N > N_\varepsilon$, $\mathbb{P}\left(\xi_i^{-1}([-N,N]^c)\right) < \delta_\varepsilon$ e $\mathbb{E}\left(|\xi_i| \mathbf{1}_{\xi_i^{-1}([-N,N]^c)}\right) < \varepsilon$. Allora, $\forall i \in \mathcal{I}$,

$$\mathbb{E}(|\xi_i|) = \mathbb{E}\left(|\xi_i| \left(\mathbf{1}_{\xi_i^{-1}([-N,N]^c)} + \mathbf{1}_{\xi_i^{-1}([-N,N])}\right)\right) \leq \varepsilon + N\ , \tag{2.74}$$

ovvero $\exists c > 0$ tale che $\sup_{i\in\mathcal{I}} \mathbb{E}(|\xi_i|) < c$. Per la disuguaglianza di Markov, $\forall i \in \mathcal{I}$, si ha

$$\mathbb{P}\left(\xi_i^{-1}[-M,M]^c\right) \leq \frac{\mathbb{E}(|\xi_i|)}{M} \leq \frac{c}{M} \underset{M\to\infty}{\longrightarrow} 0\,. \tag{2.75}$$

Quindi, $\forall \varepsilon > 0$, $\exists \delta_\varepsilon > 0 : \forall M > \frac{c}{\delta_\varepsilon}$, $\mathbb{P}\left(\xi_i^{-1}([-M,M]^c)\right) < \delta_\varepsilon$ e $\mathbb{E}\left(|\xi_i| \mathbf{1}_{\xi_i^{-1}([-M,M]^c)}\right) < \varepsilon$.

$\Leftarrow$ Poiché $\forall A \in \mathcal{F}$ e $M > 0$

$$\begin{aligned} \mathbb{E}(|\xi_i| \mathbf{1}_{A^c}) &= \mathbb{E}\left(|\xi_i| \mathbf{1}_{A^c}\left(\mathbf{1}_{\xi_i^{-1}([-M,M]^c)} + \mathbf{1}_{\xi_i^{-1}([-M,M])}\right)\right) \\ &\leq \mathbb{E}\left(|\xi_i| \mathbf{1}_{\xi_i^{-1}([-M,M]^c)}\right) + M\mathbb{P}(A^c)\,, \end{aligned} \tag{2.76}$$

$\forall \varepsilon > 0$, $\exists M_\varepsilon > 0$ tale che $\forall M > M_\varepsilon$, $\sup_{i\in\mathcal{I}} \mathbb{E}\left(|\xi_i| \mathbf{1}_{\xi_i^{-1}([-M,M]^c)}\right) < \varepsilon$. Quindi, $\forall A \in \mathcal{F}$ tale che $\mathbb{P}(A^c) < \frac{\varepsilon}{M}$, $\sup_{i\in\mathcal{I}} \mathbb{E}(|\xi_i| \mathbf{1}_{A^c}) < 2\varepsilon$.

■

Proposizione 149 *La chiusura nella topologia della convergenza $\mathbb{P}$-q.c. della famiglia di v.a. uniformemente integrabile $\{\xi_i\}_{i\in\mathcal{I}}$ definite su $(\Omega, \mathcal{F}, \mathbb{P})$ è anch'essa uniformemente integrabile.*

Dimostrazione: Sia $\{\xi_n\}_{n\geq 1} \subset \{\xi_i\}_{i\in\mathcal{I}}$ una successione di v.a. convergente $\mathbb{P}$-q.c. a ξ. Per Lemma di Fatou, $\forall A \in \mathcal{F}$,

$$\mathbb{E}(|\xi| \mathbf{1}_{A^c}) = \mathbb{E}\left(\lim_{n\to\infty} |\xi_n| \mathbf{1}_{A^c}\right) = \mathbb{E}\left(\underline{\lim}_n |\xi_n| \mathbf{1}_{A^c}\right) \leq \underline{\lim}_n \mathbb{E}(|\xi_n| \mathbf{1}_{A^c}) \leq \sup_{i\in\mathcal{I}} \mathbb{E}(|\xi_i| \mathbf{1}_{A^c})\,. \tag{2.77}$$

■

Teorema 150 *(della convergenza dominata - generalizzato) Sia $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ una successione di v.a. uniformemente integrabile e convergente $\mathbb{P}$-q.c. a $\xi \in L(\mathcal{F})$. Allora, $\{\xi_n\}_{n\geq 1}$ converge a ξ in L^1.*

Dimostrazione: Per il Teorema di Egorov, $\forall \varepsilon > 0$, $\exists K_\varepsilon \in \mathcal{F}$ compatto tale che $\mathbb{P}(K_\varepsilon^c) < \varepsilon$ e $\{\xi_n\}_{n\geq 1}$ converge uniformemente a ξ su K_ε. Allora,

$$\begin{aligned} \mathbb{E}(|\xi_n - \xi|) &\leq \mathbb{E}\left(|\xi_n - \xi| \left(\mathbf{1}_{K_\varepsilon} + \mathbf{1}_{K_\varepsilon^c}\right)\right) \leq \mathbb{E}(|\xi_n - \xi| \mathbf{1}_{K_\varepsilon}) + \mathbb{E}\left(|\xi_n| \mathbf{1}_{K_\varepsilon^c}\right) + \mathbb{E}\left(|\xi| \mathbf{1}_{K_\varepsilon^c}\right) \\ &\leq \sup_{\omega\in K_\varepsilon} |\xi_n(\omega) - \xi(\omega)| + \mathbb{E}\left(|\xi_n| \mathbf{1}_{K_\varepsilon^c}\right) + \mathbb{E}\left(|\xi| \mathbf{1}_{K_\varepsilon^c}\right)\,. \end{aligned} \tag{2.78}$$

Poiché ξ_n converge uniformemente a ξ e la chiusura di $\{\xi_n\}_{n\geq 1}$ nella topologia della convergenza $\mathbb{P}$-q.c. è uniformemente integrabile, il membro destro della precedente disuguaglianza tende a zero. Inoltre, essendo $\{\xi_n\}_{n\geq 1}$ uniformemente integrabile $\forall \varepsilon > 0$, $\exists \delta_\varepsilon, N_\varepsilon > 0$ tali che $\forall n \geq 1$, $N > N_\varepsilon, \mathbb{P}\left(\xi_n^{-1}([-N,N]^c)\right) < \delta_\varepsilon$ e $\mathbb{E}\left(|\xi_n| \mathbf{1}_{\xi_n^{-1}([-N,N]^c)}\right) < \varepsilon$. Allora, $\forall n \geq 1$,

$$\mathbb{E}(|\xi_n|) = \mathbb{E}\left(|\xi_n| \left(\mathbf{1}_{\xi_n^{-1}([-N,N]^c)} + \mathbf{1}_{\xi_n^{-1}([-N,N])}\right)\right) \leq \varepsilon + N \tag{2.79}$$

ovvero, $\xi_n \in L^1(\mathcal{F})$ e quindi lo è anche ξ in quanto $\mathbb{E}(|\xi|) \leq \mathbb{E}(|\xi_n - \xi|) + \mathbb{E}(|\xi_n|) < \infty$. ■

Corollario 151 *(Teorema di Lebesgue della convergenza dominata) Sia* $\{\xi_n\}_{n\geq 1} \subset L(\mathcal{F})$ *una successione di v.a. convergente* $\mathbb{P}$*-q.c. a* $\xi \in L(\mathcal{F})$ *e tale che* $\exists \varphi \in L^1(\mathcal{F}), \varphi > 0$ *per cui* $\forall n \geq 1, |\xi_n| \leq \varphi$*. Allora,* $\xi \in L^1(\mathcal{F})$ *e* $\{\xi_n\}_{n\geq 1}$ *converge a* ξ *in* L^1*.*

Dimostrazione: Poiché $\forall n \geq 1, |\xi_n| \leq \varphi, \forall N > 0$

$$\{\omega \in \Omega : |\xi_n(\omega)| > N\} \subset \{\omega \in \Omega : \varphi(\omega) > N\} . \tag{2.80}$$

Dunque, essendo $\varphi \in L^1(\mathcal{F}), \forall \varepsilon > 0, \exists N_\varepsilon > 0$ tale che $\forall N > N_\varepsilon$, $\mathbb{E}\left(|\varphi| \mathbf{1}_{\varphi^{-1}([-N,N]^c)}\right) < \varepsilon$. Allora,

$$\sup_{n\geq 1} \mathbb{E}\left(|\xi_n| \mathbf{1}_{\xi_n^{-1}([-N,N]^c)}\right) \leq \mathbb{E}\left(|\varphi| \mathbf{1}_{\varphi^{-1}([-N,N]^c)}\right) < \varepsilon. \tag{2.81}$$

Quindi, $\{\xi_n\}_{n\geq 1}$ è uniformemente integrabile e così la sua chiusura rispetto alla topologia della convergenza $\mathbb{P}$-q.c.. Pertanto la tesi segue dal teorema precedente. ■

Teorema 152 *(di Beppo Levi) Sia* $\{\xi_n\}_{n\geq 0} \subset L(\mathcal{F})$ *una successione non decrescente di v.a. tali che* $\exists K > 0$ *per cui* $\forall n \geq 0, |\mathbb{E}(\xi_n)| < K$*. Allora,*

1. *esiste una v.a.* $\xi \in L(\mathcal{F})$ *tale che* $\{\xi_n\}_{n\geq 0}$ *converge a* ξ $\mathbb{P}$*-q.c.;*
2. $\xi \in L^1(\mathcal{F})$ *;*
3. $\{\xi_n\}_{n\geq 0}$ *converge a* ξ *in* L^1*.*

Dimostrazione: Consideriamo la successione di v.a. $\{\eta_n\}_{n\geq 1}$ tale che $\eta_n := \xi_n - \xi_0 \geq 0$, allora $\{\eta_n\}_{n\geq 1}$ è crescente e $\forall n \geq 1, \eta_n \in L^1(\mathcal{F})$, poiché $\mathbb{E}(\eta_n) < 2K$. Pertanto è sufficiente dimostrare la validità della tesi per $\{\eta_n\}_{n\geq 1}$.

1. Sia $N := \{\omega \in \Omega : \eta_n(\omega) \to \infty\}$ e $N_n^c := \{\omega \in \Omega : \eta_n(\omega) > c\}$. Poiché $\{\eta_n\}_{n\geq 1}$ è crescente anche la successione di eventi $\{N_n^c\}_{n\geq 1}$ è crescente, inoltre $N = \bigcap_{c\geq 0} \bigcup_{n\geq 1} N_n^c$. Per la disuguaglianza di Markov, $\forall n \geq 1$,

$$\mathbb{P}(N_n^c) \leq \frac{1}{c}\mathbb{E}(\eta_n) \leq \frac{2K}{c} \underset{c\to\infty}{\longrightarrow} 0. \tag{2.82}$$

   Allora, $\mathbb{P}\left(\bigcup_{n\geq 1} N_n^c\right) \leq \frac{2K}{c} \underset{c\to\infty}{\longrightarrow} 0$ e poiché $N \subset \bigcup_{n\geq 1} N_n^c$, $\mathbb{P}(N) = 0$.

2. $\forall l \geq 1$, sia $A_l := \{\omega \in \Omega : l - 1 \leq \eta(\omega) < l\}$, e sia

$$\mathbb{R}^+ \ni x \longmapsto \varphi_l(x) := x\mathbf{1}_{[0,l]}(x) + l\mathbf{1}_{[l,+\infty)}(x) . \tag{2.83}$$

   Posto, $\forall l \geq 1$, $T_l(\eta_n) := (\varphi_l \circ \eta_n) \mathbf{1}_{A_l}$, per la continuità delle φ_l si ha che

$$\lim_{n\to\infty} T_l(\eta_n) = T_l\left(\lim_{n\to\infty} \eta_n\right) = T_l(\eta) . \tag{2.84}$$

   Inoltre, $\forall l, n \geq 1$, $\mathbb{E}(T_l(\eta_n)) \leq \mathbb{E}(\eta_n) \leq 2K$. Quindi $\eta \in L^1$.

3. Poiché $\{\eta_n\}_{n\geq 1}$ è crescente

$$\eta_n(\omega) \leq \eta(\omega) = \sum_{l\geq 1} T_l(\eta)(\omega) \leq \sum_{l\geq 1} l\mathbf{1}_{A_l}(\omega) \leq \eta(\omega) + 1. \tag{2.85}$$

Dunque, siccome $\eta \in L^1$, $\{\eta_n\}_{n\geq 1}$ soddisfa la ipotesi del teorema di Lebesgue della convergenza dominata, perciò converge in L^1.

■

Parte II

Introduzione alla teoria dei processi stocastici

Capitolo 3

Martingale

3.1 Martingale a tempo discreto

D'ora in poi, laddove non espressamente specificato, $\forall p \geq 1$, una successione di v.a. su $(\Omega, \mathcal{F}, \mathbb{P})$ convergente in media di ordine p si dirà *convergente in L^p.*

Definizione 153 *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità. Una successione $\{\mathcal{F}_n\}_{n\geq 0}$ non decrescente di subσalgebre di $\mathcal{F}$, ovvero tali che $\mathcal{F}_0 \subseteq .. \subseteq \mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq .. \subseteq \mathcal{F}$, è detta* filtrazione *di $(\Omega, \mathcal{F}, \mathbb{P})$.*

La σ algebra $\mathcal{F}_\infty$ generata da tutti i termini della successione $\{\mathcal{F}_n\}_{n\geq 0}$ è detta σalgebra limite *della filtrazione e, se $\mathcal{F}_\infty = \mathcal{F}$, la fitrazione è detta convergere a $\mathcal{F}$.*

Definizione 154 *Uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ munito di una filtrazione $\{\mathcal{F}_n\}_{n\geq 0}$ è detto* filtrato *ed è così indicato: $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$.*

Definizione 155 *Una successione di v.a. $\{\xi_n\}_{n\geq 0}$ definita sullo spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ è detta* adattata *alla filtrazione $\{\mathcal{F}_n\}_{n\geq 0}$ se $\forall n \geq 0$, $\xi_n^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{F}_n$.*

Definizione 156 *Una successione di v.a. $\{\xi_n\}_{n\geq 0}$ definita sullo spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ e adattata alla filtrazione $\{\mathcal{F}_n\}_{n\geq 0}$ è detta* prevedibile *se $\forall n \geq 0$, $\xi_n^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{F}_{n-1}$.*

Osservazione 157 *Data una successione di v.a. $\{\xi_n\}_{n\geq 0}$ definita sullo spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ se $\forall n \geq 0$, $\mathcal{F}_n^\xi$ è la σalgebra generata dalla collezione di σalgebre $\{\xi_k^{-1}(\mathcal{B}(\mathbb{R}))\}_{k=0}^n$, allora $\{\xi_n\}_{n\geq 0}$ è adattata a $\{\mathcal{F}_n\}_{n\geq 0}$ se e solo se, $\forall n \geq 0$, $\mathcal{F}_n^\xi \subseteq \mathcal{F}_n$.*

Definizione 158 *Una successione di v.a. $\{\xi_n\}_{n\geq 0}$ definita sullo spazio filtrato $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ e adattata alla filtrazione $\{\mathcal{F}_n\}_{n\geq 0}$ è detta* martingala, submartingala, supermartingala *se $\forall n \geq 0$, $\xi_n \in L^1(\mathcal{F})$ e rispettivamente si ha:*

$$\begin{array}{l} \mathbb{E}\left(\xi_{n+1}|\mathcal{F}_n\right) = \xi_n \\ \mathbb{E}\left(\xi_{n+1}|\mathcal{F}_n\right) \geq \xi_n \\ \mathbb{E}\left(\xi_{n+1}|\mathcal{F}_n\right) \leq \xi_n \end{array} \quad \mathbb{P}\text{-}q.c. \; . \tag{3.1}$$

Osservazione 159 *Se la successione di v.a.* $\{\xi_n\}_{n\geq 0}$ *definita sullo spazio filtrato* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *è una submartingala, la successione di v.a.* $\{\eta_n\}_{n\geq 0}$ *tale che* $\forall n \geq 0$, $\eta_n := -\xi_n$ *è una supermartingala. Pertanto submartingale e supermartingale condividono le stesse proprietà.*

Osservazione 160 *Dalla Definizione 158 si ha che, se la successione di v.a.* $\{\xi_n\}_{n\geq 0}$ *definita su* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *è una martingala,* $\forall n \geq 0, A \in \mathcal{F}_n$,

$$\mathbb{E}\left(\mathbf{1}_A \xi_{n+1}\right) = \mathbb{E}\left(\mathbb{E}\left(\mathbf{1}_A \xi_{n+1} | \mathcal{F}_n\right)\right) = \mathbb{E}\left(\mathbf{1}_A \mathbb{E}\left(\xi_{n+1} | \mathcal{F}_n\right)\right) = \mathbb{E}\left(\mathbf{1}_A \xi_n\right) . \tag{3.2}$$

Quindi, poiché $\{\mathcal{F}_n\}_{n\geq 0}$ *è crescente,* $\forall n \geq 0$,

$$\mathbb{E}\left(\xi_{n+1}\right) = \mathbb{E}\left(\mathbb{E}\left(\xi_{n+1} | \mathcal{F}_n\right)\right) = \mathbb{E}\left(\xi_n\right) \tag{3.3}$$

ovvero, $\mathbb{E}(\xi_n) = \mathbb{E}(\xi_0)$. *Inoltre,* $\forall k \geq 1, n \geq 0$,

$$\mathbb{E}\left(\xi_{n+k} | \mathcal{F}_n\right) = \mathbb{E}\left(\mathbb{E}\left(\xi_{n+k} | \mathcal{F}_{n+k-1}\right) | \mathcal{F}_n\right) = \mathbb{E}\left(\xi_{n+k-1} | \mathcal{F}_n\right) \tag{3.4}$$

che iterata k volte implica $\mathbb{E}\left(\xi_{n+k} | \mathcal{F}_n\right) = \xi_n$ $\mathbb{P}$*-q.c..*

Analogamente, se la successione di v.a. $\{\xi_n\}_{n\geq 0}$ *definita su* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *è una submartingala,* $\forall n \geq 0, A \in \mathcal{F}_n$,

$$\mathbb{E}\left(\mathbf{1}_A \xi_{n+1}\right) \geq \mathbb{E}\left(\mathbf{1}_A \xi_n\right) , \mathbb{E}\left(\xi_n\right) \geq \mathbb{E}\left(\xi_0\right) \tag{3.5}$$

e $\forall k \geq 1$, $\mathbb{E}\left(\xi_{n+k} | \mathcal{F}_n\right) \geq \xi_n$ $\mathbb{P}$*-q.c..*

Osservazione 161 *Notiamo che, se la successione di v.a.* $\{\xi_n\}_{n\geq 0}$ *definita su* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *è una submartingala e* $f : \mathbb{R} \longrightarrow \mathbb{R}$ *è una funzione convessa* $\mathcal{B}(\mathbb{R})$ *-mis., allora, per la disuguaglianza di Jensen, si ha che la successione di v.a.* $\{\eta_n\}_{n\geq 0}$ *tale che* $\forall n \geq 0$, $\eta_n := f(\xi_n)$, *è ancora una submartingala. In particolare ciò resta vero se* $\{\xi_n\}_{n\geq 0}$ *è una martingala.*

Esempio 1 Sia $\{\xi_n\}_{n\geq 1}$ una successione di v.a. indipendenti definite su $(\Omega, \mathcal{F}, \mathbb{P})$, tali che, $\forall n \geq 1$, $\xi_n \in L^1(\mathcal{F})$, $\mathbb{E}(\xi_n) = 0$ e sia $\{\mathcal{F}_n\}_{n\geq 0}$ la filtrazione tale che, $\mathcal{F}_0 := \{\emptyset, \Omega\}$ e $\forall n \geq 1$, $\mathcal{F}_n := \mathcal{F}_n^\xi$. La successione di v.a. $\{S_n\}_{n\geq 0}$ tale che, $S_0 := 0$ e $\forall n \geq 1$, $S_n := \sum_{k=1}^n \xi_k$ è una martingala su $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$. Infatti, $\{S_n\}_{n\geq 0}$ è adattata a $\{\mathcal{F}_n\}_{n\geq 0}$, $\forall n \geq 0$, $\mathbb{E}(|S_n|) \leq \sum_{k=0}^n \mathbb{E}(|\xi_n|) < \infty$ e

$$\mathbb{E}\left(S_{n+1} | \mathcal{F}_n\right) = \mathbb{E}\left(S_n + \xi_{n+1} | \mathcal{F}_n\right) = S_n + \mathbb{E}\left(\xi_{n+1} | \mathcal{F}_n^\xi\right) = S_n + \mathbb{E}\left(\xi_{n+1}\right) = S_n \quad \mathbb{P}\text{-q.c.} . \tag{3.6}$$

Esempio 2 Sia $\{\xi_n\}_{n\geq 1}$ una successione di v.a.i.i.d definite su $(\Omega, \mathcal{F}, \mathbb{P})$, tali che $\mathbb{E}(\xi_1) = \mu$, $\mathbb{E}\left((\xi_1 - \mu)^2\right) = \sigma^2 < \infty$. Allora, considerando la filtrazione $\{\mathcal{F}_n\}_{n\geq 0}$ definita nell'Esempio 1 e la successione di v.a. $\{S_n\}_{n\geq 0}$ tale che, $S_0 := 0$ e $\forall n \geq 1$, $S_n := \sum_{k=1}^n \xi_k$, si ha che le seguenti succesioni di v.a.:

1. $\{X_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $X_n := S_n - n\mu$;
2. $\{Y_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $Y_n := X_n^2 - n\sigma^2$;

3. *(martingala di Wald)* $\{Z_n(t)\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $Z_n := \varphi^{-n}(t)\, e^{tS_n}$, se $\exists t \neq 0$ per cui $\varphi(t) := \mathbb{E}\left(e^{t\xi_1}\right) \geq 1$;

sono martingale su $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$. Infatti:

1. segue dall'esempio precedente, considerando al posto di $\{\xi_n\}_{n\geq 0}$, la successione di v.a. $\{\eta_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $\eta_n := \xi_n - \mu$, i cui elementi appartengono ad $L^2(\mathcal{F}) \subset L^1(\mathcal{F})$, poiché $\forall n \geq 0$, $\mathbb{E}(\eta_n^2) = \sigma^2 < \infty$.

2. Considerando la successione di v.a.i.i.d. $\{\eta_n\}_{n\geq 0}$ definita al punto precedente, si ha che $\mathbb{P}$-q.c.,

$$\mathbb{E}(|Y_n|) \leq \mathbb{E}\left(X_n^2\right) + n\sigma^2 = \mathbb{E}\left((S_n - n\mu)^2\right) + n\sigma^2 = \mathbb{E}\left(\left(\sum_{k=1}^{n} \eta_k\right)^2\right) + n\sigma^2 = 2n\sigma^2 . \quad (3.7)$$

Inoltre,

$$\begin{aligned} \mathbb{E}(Y_{n+1}|\mathcal{F}_n) &= \mathbb{E}\left(\left(X_n + \left(\xi_{n+1} - \mu\right)\right)^2 - (n+1)\,\sigma^2|\mathcal{F}_n\right) \quad (3.8)\\ &= \mathbb{E}\left(\left(X_n + \eta_{n+1}\right)^2 - (n+1)\,\sigma^2|\mathcal{F}_n\right) \\ &= \mathbb{E}\left(X_n^2 + 2\eta_{n+1}X_n + \eta_{n+1}^2 - (n+1)\,\sigma^2|\mathcal{F}_n\right) \\ &= X_n^2 + 2X_n\mathbb{E}\left(\eta_{n+1}|\mathcal{F}_n\right) - n\sigma^2 + \mathbb{E}\left(\eta_{n+1}^2 - \sigma^2|\mathcal{F}_n\right) \\ &= X_n^2 + 2X_n\mathbb{E}\left(\eta_{n+1}\right) - n\sigma^2 + \mathbb{E}\left(\eta_{n+1}^2 - \sigma^2\right) = X_n^2 - n\sigma^2 = Y_n \quad \mathbb{P}\text{-q.c. .} \end{aligned}$$

3. $\forall n \geq 0$, $\mathbb{E}(|Z_n|) < \infty$. Inoltre,

$$\begin{aligned} \mathbb{E}(Z_{n+1}|\mathcal{F}_n) &= \varphi^{-(n+1)}\mathbb{E}\left(e^{t\left(S_n + \xi_{n+1}\right)}|\mathcal{F}_n\right) = \varphi^{-n}(t)\, e^{tS_n}\mathbb{E}\left(\varphi^{-1}(t)\, e^{t\xi_{n+1}}|\mathcal{F}_n\right) \quad (3.9)\\ &= Z_n\mathbb{E}\left(\varphi^{-1}(t)\, e^{t\xi_{n+1}}\right) = Z_n \quad \mathbb{P}\text{-q.c. .} \end{aligned}$$

Esempio 3 Sia $\{\xi_n\}_{n\geq 0}$ una successione di v.a. definite su $(\Omega, \mathcal{F}, \mathbb{P})$ formante una catena di Markov omogenea a stati finiti o numerabili. Indicando con $\mathcal{S}$ lo spazio degli stati della catena, sia $(P_{ij})_{i,j\in\mathcal{S}}$ la matrice di tansizione di probabilità, ovvero, $\forall i, j \in \mathcal{S}, n \geq 0$,

$$P_{i,j} = \mathbb{P}\left(\{\omega \in \Omega : \xi_{n+1}(\omega) = j\}|\{\omega \in \Omega : \xi_n(\omega) = i\}\right). \quad (3.10)$$

Se $\phi : \mathcal{S} \longmapsto \mathbb{R}$ è un autovettore destro di $(P)_{i,j}$, associato all'autovalore $\lambda > 0$, la successione di v.a. $\{\eta_n\}_{n\geq 0}$ definita sullo spazio filtrato $\left(\Omega, \mathcal{F}, \{\mathcal{F}_n^\xi\}_{n\geq 0}, \mathbb{P}\right)$ tale che $\forall n \geq 0$, $\eta_n := \lambda^{-n}\phi(\xi_n)$ e $\eta_n \in L^1(\mathcal{F})$ è una martingala. Infatti, $\{\eta_n\}_{n\geq 0}$ è adattata e, per la proprietà di Markov,

$$\begin{aligned} \mathbb{E}\left(\eta_{n+1}|\mathcal{F}_n^\xi\right) &= \mathbb{E}\left(\eta_{n+1}|\xi_n\right) = \lambda^{-n-1}\mathbb{E}\left(\phi\left(\xi_{n+1}\right)|\xi_n\right) = \lambda^{-n-1}\sum_{i\in\mathcal{S}} P_{\xi_n,i}\phi(i) \quad (3.11)\\ &= \lambda^{-n}\phi(\xi_n) = \eta_n \quad \mathbb{P}\text{-q.c..} \end{aligned}$$

Osservazione 162 *Più in generale, se la successione di v.a.* $\{\xi_n\}_{n\geq 0}$ *definita su* $(\Omega, \mathcal{F}, \mathbb{P})$ *è un processo di Markov omogeneo a tempo discreto a valori reali e* $\mathbb{R}^2 \ni (x, y) \longmapsto F(x|y) \in [0, 1]$ *una funzione di distribuzione regolare di* ξ_1 *rispetto a* $\mathcal{F}_{\xi_0}$*, allora questa ne è la funzione di distribuzione di transizione, in quanto*

$$F(y|\cdot) = \mathbb{P}\left(\{\omega \in \Omega : \xi_{n+1}(\omega) \leq y\}|\mathcal{F}_{\xi_n}\right) = \mathbb{P}\left(\{\omega \in \Omega : \xi_{n+1}(\omega) \leq y\}|\xi_n\right) \ . \tag{3.12}$$

La successione di v.a. $\{\eta_n\}_{n\geq 0}$ *tale che* $\forall n \geq 0$, $\eta_n \in L^1(\mathcal{F})$ *e* $\eta_n := \lambda^{-n}\phi(\xi_n)$ *dove,* $\forall x \in \mathbb{R}, \lambda\phi(x) = \int_{\mathbb{R}} dF(y|x)\phi(y)$ *è una martingala su* $\left(\Omega, \mathcal{F}, \{\mathcal{F}_n^{\xi}\}_{n\geq 0}, \mathbb{P}\right)$*. La martingala di Wald di cui al punto 3 dell'Esempio 2, infatti, può essere letta in questo modo se si considera il processo di Markov a tempo discreto* $\{S_n\}_{n\geq 0}$ *il cui termine n-simo è la somma parziale degli elementi della successione di v.a.i.i.d.* $\{\xi_n\}_{n\geq 0}$*. Allora, se G è la funzione di distribuzione di* ξ_1*,*

$$F(y|x) := \mathbb{P}(\{\omega \in \Omega : S_{n+1}(\omega) \leq y\}|S_n)(x) = G(y - x) \tag{3.13}$$

e, poiché per definizione $\varphi(t) := \mathbb{E}\left(e^{t\xi_1}\right) = \int_{\mathbb{R}} dG(x)e^{tx}$*, si ha*

$$\int_{\mathbb{R}} dF(y|x)e^{ty} = \int_{\mathbb{R}} dG(y-x)e^{ty} = \int_{\mathbb{R}} dG(z)e^{t(z+x)} = e^{tx}\varphi(t) \ . \tag{3.14}$$

3.1.1 Caratterizzazione e proprietà

Definizione 163 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato. Una v.a.* $\Omega \ni \omega \longmapsto \tau(\omega) \in \mathbb{Z}$ *è detta* tempo di Markov *rispetto a* $\{\mathcal{F}_n\}_{n\geq 0}$ *se* $\forall n \geq 0$, $\{\omega \in \Omega : \tau(\omega) = n\} \in \mathcal{F}_n$*. Se inoltre* $\mathbb{P}\{\omega \in \Omega : \tau(\omega) < \infty\} = 1$*, allora* τ *è detto* tempo d'arresto.

Proposizione 164 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato e* τ *un tempo di Markov rispetto a* $\{\mathcal{F}_n\}_{n\geq 0}$*.Allora,*

$$\mathcal{F}_\tau := \{A \in \mathcal{F} : A \cap \{\omega \in \Omega : \tau(\omega) = n\} \in \mathcal{F}_n, \forall n \geq 0\} \tag{3.15}$$

è una σ*algebra.*

Dimostrazione: $\Omega, \emptyset \in \mathcal{F}_\tau$ e $\mathcal{F}_\tau$ è chiusa rispetto ad unioni numerabili d'insiemi. Inoltre, se $A \in \mathcal{F}_\tau$, $\forall n \geq 0$,

$$\begin{aligned} A^c \cap \{\omega \in \Omega : \tau(\omega) = n\} &= \{\omega \in \Omega : \tau(\omega) = n\} \setminus A \\ &= \{\omega \in \Omega : \tau(\omega) = n\} \setminus (A \cap \{\omega \in \Omega : \tau(\omega) = n\}) \in \mathcal{F}_n \ , \end{aligned} \tag{3.16}$$

quindi $A^c \in \mathcal{F}_n$ e dunque $\mathcal{F}_\tau$ è una σalgebra. ■

Proposizione 165 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato,* $\{\xi_n\}_{n\geq 0}$ *una successione di v.a. adattata a* $\{\mathcal{F}_n\}_{n\geq 0}$ *e* τ *un tempo di Markov. Allora,* $\xi_\tau := \sum_{n\geq 0}\xi_n\mathbf{1}_{\{\omega\in\Omega : \tau(\omega)=n\}}$ *è una v.a..*

Dimostrazione: $\forall A \in \mathcal{B}(\mathbb{R})$, si ha

$$\{\omega \in \Omega : \xi_\tau \in A\} = \bigvee_{n\geq 0} \{\omega \in \Omega : \xi_n(\omega) \in A\} \cap \{\omega \in \Omega : \tau(\omega) = n\} \in \mathcal{F} \ . \tag{3.17}$$

■

Osservazione 166 *Poiché*

$$\begin{aligned}\{\omega\in\Omega:\tau(\omega)=n\} &= \{\omega\in\Omega:n-1<\tau(\omega)\leq n\} \\ &= \{\omega\in\Omega:\tau(\omega)>n-1\}\setminus\{\omega\in\Omega:\tau(\omega)>n\}\in\mathcal{F}_n\,, \end{aligned} \tag{3.18}$$

$$\{\omega\in\Omega:\tau(\omega)\leq n\} = \bigvee_{k=1}^{n}\{\omega\in\Omega:k-1<\tau(\omega)\leq k\}\ , \tag{3.19}$$

ma, $\forall 1\leq k\leq n$, $\{\omega\in\Omega:k-1<\tau(\omega)\leq k\}\in\mathcal{F}_k\subseteq\mathcal{F}_n$. *Quindi,* $\{\omega\in\Omega:\tau(\omega)\leq n\}\in\mathcal{F}_n$ *così come* $\{\omega\in\Omega:\tau(\omega)>n\}$.

Osservazione 167 *Sia* $(\Omega,\mathcal{F},\{\mathcal{F}_n\}_{n\geq0},\mathbb{P})$ *uno spazio filtrato e* τ,σ *due tempi di Markov rispetto a* $\{\mathcal{F}_n\}_{n\geq0}$. *Allora* $\sigma\vee\tau$, $\sigma\wedge\tau$ *e* $\sigma+\tau$ *sono tempi di Markov. Infatti,* $\forall n>0$,

$$\{\omega\in\Omega:\tau(\omega)\vee\sigma(\omega)\leq n\} = \{\omega\in\Omega:\tau(\omega)\leq n\}\cap\{\omega\in\Omega:\sigma(\omega)\leq n\}\in\mathcal{F}_n\,; \tag{3.20}$$

$$\{\omega\in\Omega:\tau(\omega)\wedge\sigma(\omega)>n\} = \{\omega\in\Omega:\tau(\omega)>n\}\cap\{\omega\in\Omega:\sigma(\omega)>n\}\in\mathcal{F}_n\,; \tag{3.21}$$

$$\{\omega\in\Omega:\tau(\omega)+\sigma(\omega)=n\} = \bigvee_{k=0}^{n}\{\omega\in\Omega:\tau(\omega)=n-k\}\cap\{\omega\in\Omega:\sigma(\omega)=k\}\in\mathcal{F}_n\,. \tag{3.22}$$

Quindi se $\{\tau_n\}_{\geq0}$ *è una successione di tempi di Markov, anche* $\inf_{n\geq0}\tau_n$, $\sup_{n\geq0}\tau_n$, $\overline{\lim}_n\tau_n$, $\underline{\lim}_n\tau_n$, *sono tempi di Markov. Inoltre, dato un qualsiasi* $N\in\mathbb{N}$, *anche* $\tau_N=N\delta_{\tau,N}$ *è un tempo di Markov.*

Esempio 4 Sia $(\Omega,\mathcal{F},\{\mathcal{F}_n\}_{n\geq0},\mathbb{P})$ uno spazio filtrato, $\{\xi_n\}_{n\geq0}$ una successione di v.a. adattata a $\{\mathcal{F}_n\}_{n\geq0}$. $\forall\ A\in\mathcal{B}(\mathbb{R})$, $\tau_A:=\inf\{n\geq0:\xi_n\in A\}$ è un tempo di Markov rispetto a $\{\mathcal{F}_n\}_{n\geq0}$ in quanto,

$$\{\omega\in\Omega:\tau_A(\omega)=n\} = \{\omega\in\Omega:\xi_n(\omega)\in A\}\cap\left(\bigcap_{k=0}^{n-1}\{\omega\in\Omega:\xi_k(\omega)\notin A\}\right)\in\mathcal{F}_n\ . \tag{3.23}$$

Definizione 168 *Sia* $(\Omega,\mathcal{F},\{\mathcal{F}_n\}_{n\geq0},\mathbb{P})$ *uno spazio filtrato,* $\{\xi_n\}_{n\geq0}$ *una martingala e* τ *un tempo di Markov rispetto a* $\{\mathcal{F}_n\}_{n\geq0}$. *La successione di v.a.* $\{\xi_{n\wedge\tau}\}_{n\geq0}$ *tale che,* $\forall n\geq0$,

$$\xi_{n\wedge\tau} := \sum_{k=0}^{n-1}\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)=k\}}\xi_k+\xi_n\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)\geq n\}} \tag{3.24}$$

è detta martingala troncata *o* arrestata.

Infatti, poiché dalla (3.18), $\forall n \geq 1$, $\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)=n\}} = \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>n-1\}} - \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>n\}}$ e

$$
\begin{aligned}
\xi_{n\wedge\tau} &= \sum_{k=0}^{n} \mathbf{1}_{\{\omega\in\Omega:\tau(\omega)=k\}}\xi_k + \xi_n \mathbf{1}_{\{\omega\in\Omega:\tau(\omega)>n\}} = \qquad (3.25)\\
&= \xi_0 \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)=0\}} + \sum_{k=1}^{n} \xi_k \left(\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>k-1\}} - \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>k\}} \right) + \xi_n \mathbf{1}_{\{\omega\in\Omega:\tau(\omega)>n\}} = \\
&= \xi_0 \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)=0\}} + \xi_1 \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>0\}} + \sum_{k=1}^{n-1} \left(\xi_{k+1} - \xi_k \right) \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>k\}} = \\
&= \xi_0 \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)\geq 0\}} + \sum_{k=1}^{n-1} \left(\xi_{k+1} - \xi_k \right) \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>k\}} = \\
&= \xi_0 + \sum_{k=0}^{n-1} \left(\xi_{k+1} - \xi_k \right) \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>k\}}\,,
\end{aligned}
$$

dove l'ultimo passaggio segue dal fatto che $\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)\geq 0\}} = \mathbf{1}_{\Omega}$. Allora, $\forall n \geq 0$,

$$
\begin{aligned}
\mathbb{E}\left(\xi_{(n+1)\wedge\tau} - \xi_{n\wedge\tau} | \mathcal{F}_n \right) &= \mathbb{E}\left(\left(\xi_{n+1} - \xi_n \right) \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>n\}} | \mathcal{F}_n \right) \qquad (3.26)\\
&= \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>n\}} \mathbb{E}\left(\xi_{n+1} - \xi_n | \mathcal{F}_n \right) = 0\,,
\end{aligned}
$$

ovvero $\mathbb{E}\left(\xi_{(n+1)\wedge\tau} | \mathcal{F}_n \right) = \mathbb{E}\left(\xi_{n\wedge\tau} | \mathcal{F}_n \right)$. Ma per definizione $\xi_{n\wedge\tau}$ è $\mathcal{F}_n$-mis., quindi $\mathbb{E}\left(\xi_{n\wedge\tau} | \mathcal{F}_n \right) = \xi_{n\wedge\tau}$ cioè $\{\xi_{n\wedge\tau}\}_{n\geq 0}$ è una martingala.

Definizione 169 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato e* $\{\tau_k\}_{k\geq 0}$, *tale che* $\forall k \geq 0$, $\tau_k \leq \tau_{k+1}$, *una successione non decrescente di tempi di Markov per cui* $\lim_{k\to\infty} \tau_k = +\infty$ $\mathbb{P}$*-q.c. La martingala* $\{\xi_n\}_{n\geq 0}$ *è detta* martingala locale *se* $\forall k \geq 0$, $\{\xi_{n\wedge\tau_k} \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>0\}}\}_{n\geq 0}$ *è una martingala.*

Definizione 170 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato,* $\{\xi_n\}_{n\geq 0}$ *una martingala e* $\{\eta_n\}_{n\geq 0}$ *una successione di v.a. prevedibile. La successione di v.a.* $\{\theta_n\}_{n\geq 0}$ *tale che,* $\forall n \geq 0$,

$$
\theta_n = \xi_n \cdot \eta_n := \xi_0 \eta_0 + \sum_{m=1}^{n} \eta_m \left(\xi_m - \xi_{m-1} \right)\,, \qquad (3.27)
$$

è detta trasformata martingala *di* $\{\xi_n\}_{n\geq 0}$ *tramite* $\{\eta_n\}_{n\geq 0}$.

Teorema 171 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato. Una successione di v.a.* $\{\theta_n\}_{n\geq 0}$ *adattata a* $\{\mathcal{F}_n\}_{n\geq 0}$ *è una martingala locale se e solo se esiste una successione di v.a. prevedibile* $\{\eta_n\}_{n\geq 0}$ *ed una martingala* $\{\xi_n\}_{n\geq 0}$ *tale che* $\{\theta_n\}_{n\geq 0}$ *è la trasformata martingala di* $\{\xi_n\}_{n\geq 0}$ *tramite* $\{\eta_n\}_{n\geq 0}$.

Dimostrazione:

$\Rightarrow$ Sia $\{\theta_n\}_{n\geq 0}$ una martingala locale, $\forall n \geq 1$, si ponga:

$$\eta_n := \mathbb{E}\left(|\theta_n - \theta_{n-1}|\,|\mathcal{F}_{n-1}\right)\ ; \tag{3.28}$$

$$\zeta_n := \frac{1}{\eta_n}\mathbf{1}_{\{\omega\in\Omega\,:\,\eta_n\neq 0\}}\ . \tag{3.29}$$

Allora, posto $\xi_0 := \theta_0$ e, $\forall n \geq 1$, $\xi_n := \sum_{k=1}^{n}\zeta_k\left(\theta_k - \theta_{k-1}\right)$, poiché per ipotesi η_n, e dunque anche ζ_n, è $\mathcal{F}_{n-1}$-mis., $\forall n \geq 1$, si ha che $\mathbb{P}$-q.c.,

$$\mathbb{E}\left(\xi_n - \xi_{n-1}|\mathcal{F}_{n-1}\right) = \mathbb{E}\left(\zeta_n\left(\theta_n - \theta_{n-1}\right)|\mathcal{F}_{n-1}\right) = \zeta_n\mathbb{E}\left(\theta_n - \theta_{n-1}|\mathcal{F}_{n-1}\right) = 0\,, \tag{3.30}$$

$$\begin{aligned}\mathbb{E}\left(\left|\xi_n - \xi_{n-1}\right|\,|\mathcal{F}_{n-1}\right) &= \mathbb{E}\left(\frac{|\theta_n - \theta_{n-1}|}{\mathbb{E}\left(|\theta_n - \theta_{n-1}|\,|\mathcal{F}_{n-1}\right)}\mathbf{1}_{\{\omega\in\Omega\,:\,\mathbb{E}(|\theta_n-\theta_{n-1}||\mathcal{F}_{n-1})\neq 0\}}|\mathcal{F}_{n-1}\right) \\ &\leq 1\ .\end{aligned} \tag{3.31}$$

Perciò, $\forall n \geq 1$,

$$\begin{aligned}\mathbb{E}\left(|\xi_n|\right) &\leq \mathbb{E}\left(|\theta_0|\right) + \sum_{k=1}^{n}\mathbb{E}\left(\left|\xi_k - \xi_{k-1}\right|\right) = \mathbb{E}\left(|\theta_0|\right) + \sum_{k=1}^{n}\mathbb{E}\left(\mathbb{E}\left(\left|\xi_k - \xi_{k-1}\right|\,|\mathcal{F}_{k-1}\right)\right) \\ &\leq \mathbb{E}\left(|\theta_0|\right) + n < \infty\,,\end{aligned} \tag{3.32}$$

ovvero, $\{\xi_n\}_{n\geq 0}$ è una martingala rispetto a $\{\mathcal{F}_n\}_{n\geq 0}$. Inoltre, posto $\eta_0 := 1$,

$$\begin{aligned}\xi_n\cdot\eta_n &:= \xi_0\eta_0 + \sum_{m=1}^{n}\eta_m\left(\xi_m - \xi_{m-1}\right) = \theta_0 + \sum_{m=1}^{n}\eta_m\zeta_m\left(\theta_m - \theta_{m-1}\right) \\ &= \theta_0 + \sum_{m=1}^{n}\left(\theta_m - \theta_{m-1}\right)\ .\end{aligned} \tag{3.33}$$

$\Leftarrow$ Siano $\{\eta_n\}_{n\geq 0}$ una successione di v.a. prevedibile e $\{\xi_n\}_{n\geq 0}$ una martingala. Posto $\{\theta_n\}_{n\geq 0}$ la trasformata martingala di $\{\xi_n\}_{n\geq 0}$ tramite $\{\eta_n\}_{n\geq 0}$, $\forall k \geq 1$, sia

$$\tau_k := \begin{cases}\inf\{n\geq 0 : |\eta_n| > k\} & se \quad \{\omega\in\Omega : |\eta_n(\omega)| > k\} \neq \emptyset \\ \infty & se \quad \{\omega\in\Omega : |\eta_n(\omega)| > k\} = \emptyset\end{cases}\ . \tag{3.34}$$

Poiché $\forall n \geq 0$, η_n è $\mathcal{F}_{n-1}$-mis., le v.a. τ_k sono anch'esse $\mathcal{F}_{n-1}$-mis., quindi $\mathcal{F}_n$-mis., e perciò sono tempi di Markov. Inoltre, dalla definizione di τ_k si ha che $\tau_{k+1} \geq \tau_k$, ovvero $\{\tau_k\}_{k\geq 0}$ è una successione crescente di tempi di Markov per cui $\lim_{k\to\infty}\tau_k = +\infty\ \mathbb{P}$-q.c. e, dalla (3.25), si ha che

$$\begin{aligned}\xi_{n\wedge\tau_k}\cdot\eta_{n\wedge\tau_k} &= \left(\xi_0 + \sum_{m=0}^{n-1}\left(\xi_{m+1} - \xi_m\right)\mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>m\}}\right)\cdot\left(\eta_0 + \sum_{m=0}^{n-1}\left(\eta_{m+1} - \eta_m\right)\mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>m\}}\right) \\ &= \xi_0\eta_0 + \sum_{m=1}^{n}\eta_{m\wedge\tau_k}\left(\xi_{m\wedge\tau_k} - \xi_{(m-1)\wedge\tau_k}\right) \\ &= \xi_0\eta_0 + \sum_{m=0}^{n-1}\eta_{m+1}\left(\xi_{m+1} - \xi_m\right)\mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>m\}} = \theta_{n\wedge\tau_k}\,,\end{aligned} \tag{3.35}$$

Quindi, $\forall n, k \geq 0$,

$$\mathbb{E}\left(\left(\theta_{(n+1)\wedge\tau_k} - \theta_{n\wedge\tau_k}\right) \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>0\}} | \mathcal{F}_n\right) = \tag{3.36}$$
$$\mathbb{E}\left(\eta_{n+1}\left(\xi_{n+1} - \xi_n\right) \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>n\}} \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>0\}} | \mathcal{F}_n\right) =$$
$$\mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>0\}} \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>n\}} \mathbb{E}\left(\eta_{n+1}\left(\xi_{n+1} - \xi_n\right) | \mathcal{F}_n\right) =$$
$$\mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>0\}} \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>n\}} \eta_{n+1} \mathbb{E}\left(\left(\xi_{n+1} - \xi_n\right) | \mathcal{F}_n\right) = 0 \quad \mathbb{P}\text{-q.c. .}$$

Poiché dalla definizione di τ_k segue che $\left|\eta_{n\wedge\tau_k}\right| \leq k$, allora,

$$\mathbb{E}\left(\left|\theta_{n\wedge\tau_k}\right| \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>0\}}\right) = \mathbb{E}\left(\left|\xi_{n\wedge\tau_k} \cdot \eta_{n\wedge\tau_k}\right| \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>0\}}\right) \leq$$
$$\mathbb{E}\left(\left|\xi_0\eta_0 + \sum_{m=1}^{n} \eta_{m\wedge\tau_k}\left(\xi_{m\wedge\tau_k} - \xi_{(m-1)\wedge\tau_k}\right)\right|\right) \leq$$
$$\mathbb{E}\left(\left|\xi_0\eta_0\right|\right) + \mathbb{E}\left(\sum_{m=1}^{n} \left|\eta_{m\wedge\tau_k}\right| \left|\xi_{m\wedge\tau_k} - \xi_{(m-1)\wedge\tau_k}\right|\right) =$$
$$\mathbb{E}\left(\left|\xi_0\eta_0\right|\right) + \mathbb{E}\left(\sum_{m=1}^{n} \left|\eta_{m\wedge\tau_k}\right| \left|\xi_m - \xi_{m-1}\right| \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>m\}}\right) \leq$$
$$\mathbb{E}\left(\left|\xi_0\eta_0\right|\right) + k\sum_{m=1}^{n} \mathbb{E}\left(\left|\xi_m - \xi_{m-1}\right| \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>m\}}\right) < \infty\,.$$

Dunque, $\left\{\theta_{n\wedge\tau_k}\mathbf{1}_{\{\omega\in\Omega\,:\,\tau_k(\omega)>0\}}\right\}_{n\geq 0}$ è una martingala e quindi $\{\theta_n\}_{n\geq 0}$ è una martingala locale.

■

Esempio 5 Sia $\{\xi_n\}_{n\geq 1}$ una successione di v.a.i.i.d. su $(\Omega, \mathcal{F}, \mathbb{P})$ distribuite secondo la legge di Bernoulli di parametro $p \in (0,1)$ tali che $\forall n \geq 1$, $\xi_n \in \{-1, 1\}$. Se l'evento $\{\omega \in \Omega : \xi_n(\omega) = 1\}$ rappresenta il successo di un giocatore al'n-simo turno di gioco, supponendo che η_n rappresenti l'ammontare della scommessa del giocatore al turno n-simo, il guadagno totale di un giocatore all'n-simo turno di gioco risulta

$$\theta_n = \sum_{i=1}^{n} \eta_i\xi_i = \theta_{n-1} + \eta_n\xi_n\,. \tag{3.37}$$

Se $\forall n \geq 1$, η_n è una funzione di $\theta_1, .., \theta_{n-1}$, ovvero la puntata all'$n$-simo turno di gioco dipende dai guadagni ottenuti ai turni precedenti, ponendo $\mathcal{F}_0 = \{\emptyset, \Omega\}$ e $\forall n \geq 1$, $\mathcal{F}_n = \mathcal{F}_n^\theta = \mathcal{F}_n^\xi$, allora η_n è una v.a. $\mathcal{F}_{n-1}$-mis., cioé la successione di v.a. $\{\eta_n\}_{n\geq 0}$, con $\eta_0 = 0$, è prevedibile rispetto alla filtrazione $\{\mathcal{F}_n\}_{n\geq 0}$. Ponendo $\zeta_n := \sum_{i=1}^n \xi_i$ e $\zeta_0 = 0$, si ha $\theta_n = \sum_{i=1}^n \eta_i\left(\zeta_i - \zeta_{i-1}\right)$.

Notiamo che dal punto di vista del giocatore, posto $\mathbb{E}\left(\theta_{n+1} - \theta_n | \mathcal{F}_n\right) = x$, il gioco risulta favorevole se $x > 0$, sfavorevole se $x < 0$, equo se $x = 0$. Ma poiché le ξ_i sono v.a. indipendenti,

$$x = \mathbb{E}\left(\theta_{n+1} - \theta_n | \mathcal{F}_n\right) = \mathbb{E}\left(\eta_{n+1}\xi_{n+1} | \mathcal{F}_n\right) = \eta_{n+1}\mathbb{E}\left(\xi_{n+1} | \mathcal{F}_n\right) = \eta_{n+1}\mathbb{E}\left(\zeta_{n+1} - \zeta_n | \mathcal{F}_n\right) \tag{3.38}$$
$$= \eta_{n+1}\mathbb{E}\left(\xi_1\right) = \eta_{n+1}\left(p - q\right) = \eta_{n+1}\left(2p - 1\right)\,.$$

Quindi, poiché $\forall n \geq 0$, $\eta_{n+1} > 0$, il gioco risulta:

favorevole, se e solo se $p > \frac{1}{2}$ ovvero, se la successione di v.a. $\{\theta_n\}_{n\geq 0}$, e dunque $\{\zeta_n\}_{n\geq 0}$, è una submartingala;

sfavorevole, se e solo se $p < \frac{1}{2}$ ovvero, se la successione di v.a. $\{\theta_n\}_{n\geq 0}$, e dunque $\{\zeta_n\}_{n\geq 0}$, è una supermartingala;

equo, se e solo se $p = \frac{1}{2}$ ovvero, se la successione di v.a. $\{\theta_n\}_{n\geq 0}$, e dunque $\{\zeta_n\}_{n\geq 0}$, è una martingala.

In quest'ultimo caso, $\{\theta_n\}_{n\geq 0}$ è la trasformata martingala di $\{\zeta_n\}_{n\geq 0}$ tramite $\{\eta_n\}_{n\geq 0}$.

Nel caso in cui $\{\eta_n\}_{n\geq 0}$ sia tale che $\eta_1 = 1$ e, $\forall n \geq 2$, $\eta_n := 2^{n-1} \prod_{k=1}^{n-1} \delta_{\xi_k,-1}$, ovvero la strategia di gioco consista nel raddoppiare la posta ad ogni turno di gioco e lasciare il gioco alla prima vittoria, allora, se il turno vincente è l'$(n+1)$-simo, si ha che il guadagno totale risulta

$$\theta_{n+1} = \theta_n + \eta_{n+1} = \sum_{i=1}^{n} \eta_i \xi_i + \eta_{n+1} = \sum_{i=1}^{n} 2^{i-1} (-1) + 2^n = -(2^n - 1) + 2^n = 1 \,. \tag{3.39}$$

Sia $\tau := \inf\{n \geq 1 : \theta_n = 1\}$. Nel caso di gioco equo si ha

$$\mathbb{P}\{\omega \in \Omega : \tau(\omega) = n\} = \mathbb{P}\{\omega \in \Omega : \xi_n(\omega) = 1, \xi_k(\omega) = -1 \quad \forall 1 \leq k \leq n-1\} = \left(\frac{1}{2}\right)^n . \tag{3.40}$$

Perciò,

$$\mathbb{P}\{\omega \in \Omega : \tau(\omega) < \infty\} = \sum_{n\geq 1} \mathbb{P}\{\omega \in \Omega : \tau(\omega) = n\} = \sum_{n\geq 1} \left(\frac{1}{2}\right)^n = 1 \,. \tag{3.41}$$

Inoltre, considerando θ_τ, per definizione si ha

$$\theta_\tau = \sum_{n\geq 1} \theta_n \mathbf{1}_{\{\omega\in\Omega : \tau(\omega)=n\}} = \sum_{n\geq 1} \mathbf{1}_{\{\omega\in\Omega : \tau(\omega)=n\}} = \mathbf{1}_{\{\omega\in\Omega : \tau(\omega)<\infty\}} \,, \tag{3.42}$$

per cui $\mathbb{E}(\theta_\tau) = \mathbb{P}\{\omega \in \Omega : \tau(\omega) < \infty\} = 1$. Pertanto, mediante tale strategia di gioco, se questo è equo, è possibile incrementare il proprio capitale di un unità in tempo finito quasi certamente. Nell'ambiente delle scommesse, la strategia appena descritta prende il nome di martingala. Osserviamo inoltre, che, nel caso di gioco equo, poiché $\{\theta_n\}_{n\geq 0}$ è una martingala, $\forall n \geq 1$, $\mathbb{E}(\theta_n) = \mathbb{E}(\theta_0) = 0$, mentre $\mathbb{E}(\theta_\tau) = 1 \neq \mathbb{E}(\theta_n)$.

Definizione 172 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato. Una successione di v.a.* $\{\xi_n\}_{n\geq 0}$ *adattata a* $\{\mathcal{F}_n\}_{n\geq 0}$ *tale che,* $\forall n \geq 0$, $\xi_n \in L^1(\mathcal{F})$ *e* $\mathbb{E}(\xi_{n+1}|\mathcal{F}_n) = 0\,\mathbb{P}$*-q.c. è detta* differenza martingala.

Notiamo che se la successione di v.a. $\{\eta_n\}_{n\geq 0}$ è una martingala su $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$, allora la successione di v.a. $\{\xi_n\}_{n\geq 0}$ tale che, $\xi_0 := \eta_0$ e $\forall n \geq 1$, $\xi_n := \eta_n - \eta_{n-1}$, è una differenza martingala. D'altra parte, se la successione di v.a. $\{\xi_n\}_{n\geq 0}$ è una differenza martingala su $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$, la successione di v.a. $\{\eta_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $\eta_n := \sum_{k=0}^{n} \xi_k$ è una martingala.

Teorema 173 (Decomposizione di Doob-Meyer) *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato e* $\{\xi_n\}_{n\geq 0}$ *una submartingala. Allora esiste una martingala* $\{m_n\}_{n\geq 0}$ *ed una successione prevedibile non decrescente* $\{A_n\}_{n\geq 0}$*, detta* compensatore *di* $\{\xi_n\}_{n\geq 0}$*, tale che* $\forall n \geq 0$*, vale la decomposizione*

$$\xi_n = m_n + A_n \quad \mathbb{P}\text{-}q.c. \ , \tag{3.43}$$

che è unica, e quindi, $\mathbb{E}(\xi_n) = \mathbb{E}(m_0) + \mathbb{E}(A_n)$*.*

Dimostrazione: $\forall n \geq 1$,

$$\begin{aligned} \xi_n &= \xi_0 + \sum_{k=1}^{n} \left(\xi_k - \xi_{k-1}\right) = \xi_0 + \sum_{k=1}^{n} \left(\xi_k - \mathbb{E}(\xi_k|\mathcal{F}_{k-1})\right) + \sum_{k=1}^{n} \left(\mathbb{E}(\xi_k|\mathcal{F}_{k-1}) - \xi_{k-1}\right) \\ &= \xi_0 + \sum_{k=1}^{n} \left(\xi_k - \mathbb{E}(\xi_k|\mathcal{F}_{k-1})\right) + \sum_{k=1}^{n} \left(\mathbb{E}(\xi_k|\mathcal{F}_{k-1}) - \mathbb{E}\left(\xi_{k-1}|\mathcal{F}_{k-1}\right)\right) \\ &= \xi_0 + \sum_{k=1}^{n} \left(\xi_k - \mathbb{E}(\xi_k|\mathcal{F}_{k-1})\right) + \sum_{k=1}^{n} \mathbb{E}\left(\xi_k - \xi_{k-1}|\mathcal{F}_{k-1}\right) \ . \end{aligned} \tag{3.44}$$

Posto

$$A_n := \sum_{k=1}^{n} \mathbb{E}\left(\xi_k - \xi_{k-1}|\mathcal{F}_{k-1}\right), \quad A_0 := 0 \ , \tag{3.45}$$

si ha che A_n è $\mathcal{F}_{n-1}$-mis. ovvero, la successione di v.a. $\{A_n\}_{n\geq 0}$ è prevedibile. Inoltre, poiché $\{\xi_n\}_{n\geq 0}$ è una submartingala, $\forall n \geq 0$,

$$A_{n+1} - A_n = \mathbb{E}\left(\xi_{n+1} - \xi_n|\mathcal{F}_n\right) \geq 0 \, , \tag{3.46}$$

cioè $\{A_n\}_{n\geq 0}$ è non decrescente. D'altra parte, posto

$$m_n := \sum_{k=1}^{n} \left(\xi_k - \mathbb{E}(\xi_k|\mathcal{F}_{k-1})\right), \quad m_0 := \xi_0 \ , \tag{3.47}$$

la successione di v.a. $\{m_n\}_{n\geq 0}$ è una martingala, in quanto la successione di v.a. $\{\zeta_n\}_{n\geq 0}$ tale che $\zeta_0 := m_0$ e $\forall n \geq 1$, $\zeta_n := \xi_n - \mathbb{E}(\xi_n|\mathcal{F}_{n-1})$, è una differenza martingala.

L'unicità della decomposizione segue dal fatto che, se esistono: una martingala $\{m'_n\}_{n\geq 0}$ ed una successione di v.a. prevedibile $\{A'_n\}_{n\geq 0}$ tali che $m'_0 := \xi_0$, $A'_0 := 0$ e $\forall n \geq 1$, $\xi_n = m'_n + A'_n$, allora,

$$\xi_{n+1} - \xi_n = (m_{n+1} - m_n) + (A_{n+1} - A_n) = \left(m'_{n+1} - m'_n\right) + \left(A'_{n+1} - A'_n\right) \ , \tag{3.48}$$

perciò, poiché $\mathbb{E}(m_{n+1} - m_n|\mathcal{F}_n) = \mathbb{E}\left(m'_{n+1} - m'_n|\mathcal{F}_n\right) = 0 \, \mathbb{P}$-q.c.,

$$\begin{aligned} \mathbb{E}(m_{n+1} - m_n|\mathcal{F}_n) + \mathbb{E}(A_{n+1} - A_n|\mathcal{F}_n) &= \mathbb{E}\left(m'_{n+1} - m'_n|\mathcal{F}_n\right) + \mathbb{E}\left(A'_{n+1} - A'_n|\mathcal{F}_n\right) \\ \mathbb{E}(A_{n+1} - A_n|\mathcal{F}_n) &= \mathbb{E}\left(A'_{n+1} - A'_n|\mathcal{F}_n\right) \\ A_{n+1} - A_n &= A'_{n+1} - A'_n \quad \mathbb{P}\text{-q.c.}. \end{aligned} \tag{3.49}$$

Ma siccome $A_0 = A'_0 = 0$, $\forall n \geq 0$ si ha che $\mathbb{P}$-q.c. $A_n = A'_n$ e quindi $m_n = m'_n$. ■

Esempio 6 Sia $\{\xi_n\}_{n\geq 1}$ una successione di v.a.i.i.d definite su $(\Omega, \mathcal{F}, \mathbb{P})$, tali che $\mathbb{E}(\xi_1) = \mu$, $\mathbb{E}\left((\xi_1 - \mu)^2\right) = \sigma^2 < \infty$. Allora, considerando la filtrazione $\{\mathcal{F}_n\}_{n\geq 0}$ definita nell'Esempio 1 e la successione di v.a. $\{S_n\}_{n\geq 0}$ tale che, $S_0 := 0$ e $\forall n \geq 1$, $S_n := \sum_{k=1}^n \xi_k$, dall'Esempio 2 si ha che la successione di v.a. $\{X_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $X_n := S_n - n\mu$, è una martingala. Poiché la funzione $\mathbb{R} \ni x \longmapsto f(x) := x^2 \in \mathbb{R}$ è convessa, la successione di v.a. $\{X_n^2\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $X_n^2 := (S_n - n\mu)^2$, è una submartingala e siccome

$$\begin{aligned}
\mathbb{E}\left(X_n^2 - X_{n-1}^2 | \mathcal{F}_{n-1}\right) &= \mathbb{E}\left((X_{n-1} + \xi_n - \mu)^2 - (X_{n-1})^2 | \mathcal{F}_{n-1}\right) \\
&= \mathbb{E}\left((X_{n-1})^2 + 2X_{n-1}(\xi_n - \mu) + (\xi_n - \mu)^2 - (X_{n-1})^2 | \mathcal{F}_{n-1}\right) \\
&= 2X_{n-1}\mathbb{E}\left((\xi_n - \mu) | \mathcal{F}_{n-1}\right) + \mathbb{E}\left((\xi_n - \mu)^2 | \mathcal{F}_{n-1}\right) \\
&= \mathbb{E}\left((\xi_n - \mu)^2\right) = \sigma^2 ,
\end{aligned} \tag{3.50}$$

dalla (3.45) il compensatore di $\{X_n^2\}_{n\geq 0}$ è la successione di v.a. degeneri $\{A_n\}_{n\geq 0}$ tale che $A_0 = 0$ e, $\forall n \geq 1$, $A_n := n\sigma^2$. Per cui, $\{Y_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $Y_n := X_n^2 - n\sigma^2$ è la martingala associata alla decomposizione di Doob-Meyer di $\{X_n^2\}_{n\geq 0}$.

Conservazione della proprietà di martingala rispetto a tempi d'arresto

Teorema 174 *Sia $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ uno spazio filtrato, $\{\xi_n\}_{n\geq 0}$ una martingala e τ_1, τ_2 due tempi d'arresto tali che, $\forall i = 1, 2$*

1. $\mathbb{E}\left(\left|\xi_{\tau_i}\right|\right) < \infty$;
2. $\underline{\lim}_n \mathbb{E}\left(|\xi_n| \mathbf{1}_{\{\omega\in\Omega : \tau_i(\omega) > n\}}\right) = 0$.

Allora, $\mathbb{P}$-q.c.,

$$\mathbb{E}\left(\xi_{\tau_2}\mathbf{1}_{\{\omega\in\Omega : \tau_2(\omega)\geq\tau_1(\omega)\}} | \mathcal{F}_{\tau_1}\right) = \xi_{\tau_1}\mathbf{1}_{\{\omega\in\Omega : \tau_2(\omega)\geq\tau_1(\omega)\}} . \tag{3.51}$$

Dimostrazione: Per la definizione di probabilità condizionata, basta mostrare che, $\forall A \in \mathcal{F}_{\tau_1}$,

$$\mathbb{E}\left(\xi_{\tau_2}\mathbf{1}_{\{\omega\in\Omega : \tau_2(\omega)\geq\tau_1(\omega)\}}\mathbf{1}_A\right) = \mathbb{E}\left(\xi_{\tau_1}\mathbf{1}_{\{\omega\in\Omega : \tau_2(\omega)\geq\tau_1(\omega)\}}\mathbf{1}_A\right) , \tag{3.52}$$

ovvero, poiché

$$\begin{aligned}
\mathbb{E}\left(\xi_{\tau_2}\mathbf{1}_{\{\omega\in\Omega : \tau_2(\omega)\geq\tau_1(\omega)\}}\mathbf{1}_A\right) &= \sum_{n\geq 0} \mathbb{E}\left(\xi_{\tau_2}\mathbf{1}_{\{\omega\in\Omega : \tau_2(\omega)\geq\tau_1(\omega)\}}\mathbf{1}_{\{\omega\in\Omega : \tau_1(\omega)=n\}}\mathbf{1}_A\right) \\
\mathbb{E}\left(\xi_{\tau_1}\mathbf{1}_{\{\omega\in\Omega : \tau_2(\omega)\geq\tau_1(\omega)\}}\mathbf{1}_A\right) &= \sum_{n\geq 0} \mathbb{E}\left(\xi_{\tau_1}\mathbf{1}_{\{\omega\in\Omega : \tau_2(\omega)\geq\tau_1(\omega)\}}\mathbf{1}_{\{\omega\in\Omega : \tau_1(\omega)=n\}}\mathbf{1}_A\right) \\
&= \sum_{n\geq 0} \mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega : \tau_2(\omega)\geq n\}}\mathbf{1}_{\{\omega\in\Omega : \tau_1(\omega)=n\}}\mathbf{1}_A\right) ,
\end{aligned} \tag{3.53}$$

che

$$\mathbb{E}\left(\xi_{\tau_2}\mathbf{1}_{\{\omega\in\Omega : \tau_2(\omega)\geq n\}}\mathbf{1}_{\{\omega\in\Omega : \tau_1(\omega)=n\}}\mathbf{1}_A\right) = \mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega : \tau_2(\omega)\geq n\}}\mathbf{1}_{\{\omega\in\Omega : \tau(\omega)=n\}}\mathbf{1}_A\right) . \tag{3.54}$$

Quindi, posto $B := A \cap \{\omega \in \Omega : \tau_1(\omega) = n\} \in \mathcal{F}_n$, deve valere

$$\mathbb{E}\left(\xi_{\tau_2}\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)\geq n\}}\mathbf{1}_B\right) = \mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)\geq n\}}\mathbf{1}_B\right) . \tag{3.55}$$

Ma,

$$\begin{aligned}\mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)\geq n\}}\mathbf{1}_B\right) &= \mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)=n\}}\mathbf{1}_B\right) + \mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)>n\}}\mathbf{1}_B\right) \\ &= \mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)=n\}}\mathbf{1}_B\right) + \mathbb{E}\left(\mathbb{E}\left(\xi_{n+1}|\mathcal{F}_n\right)\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)>n\}}\mathbf{1}_B\right) \\ &= \mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)=n\}}\mathbf{1}_B\right) + \mathbb{E}\left(\mathbb{E}\left(\xi_{n+1}\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)>n\}}\mathbf{1}_B|\mathcal{F}_n\right)\right) \\ &= \mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)=n\}}\mathbf{1}_B\right) + \mathbb{E}\left(\xi_{n+1}\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)\geq n+1\}}\mathbf{1}_B\right) .\end{aligned} \tag{3.56}$$

Iterando m volte, si ha

$$\begin{aligned}\mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)\geq n\}}\mathbf{1}_B\right) &= \sum_{k=n}^{n+m}\mathbb{E}\left(\xi_k\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)=k\}}\mathbf{1}_B\right) + \mathbb{E}\left(\xi_{n+m}\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)>n+m\}}\mathbf{1}_B\right) \\ &= \mathbb{E}\left(\xi_{\tau_2}\mathbf{1}_{\{\omega\in\Omega:\,n\leq\tau_2(\omega)\leq n+m\}}\mathbf{1}_B\right) + \mathbb{E}\left(\xi_{n+m}\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)>n+m\}}\mathbf{1}_B\right) .\end{aligned} \tag{3.57}$$

Perciò, siccome $\forall m \geq 0$,

$$\{\omega\in\Omega\,:\,n\leq\tau_2(\omega)\leq n+m\} \subset \{\omega\in\Omega\,:\,n\leq\tau_2(\omega)\leq n+m+1\} , \tag{3.58}$$

$$\begin{aligned}\{\omega\in\Omega\,:\,\tau_2(\omega)\geq n\} &= \bigcup_{m\geq 0}\{\omega\in\Omega\,:n\leq\tau_2(\omega)\leq n+m\} \\ &= \lim_{m\uparrow\infty}\{\omega\in\Omega\,:\,n\leq\tau_2(\omega)\leq n+m\} ,\end{aligned} \tag{3.59}$$

otteniamo

$$\mathbb{E}\left(\xi_{\tau_2}\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)\geq n\}}\mathbf{1}_B\right) = \lim_{m\to\infty}\mathbb{E}\left(\xi_{\tau_2}\mathbf{1}_{\{\omega\in\Omega:\,n\leq\tau_2(\omega)\leq n+m\}}\mathbf{1}_B\right) ,$$

ovvero,

$$\begin{aligned}\mathbb{E}\left(\xi_{\tau_2}\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)\geq n\}}\mathbf{1}_B\right) &= \lim_{m\to\infty}\left[\mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)\geq n\}}\mathbf{1}_B\right) - \mathbb{E}\left(\xi_{n+m}\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)>n+m\}}\mathbf{1}_B\right)\right] \\ &= \mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)\geq n\}}\mathbf{1}_B\right) - \underline{\lim}_m\mathbb{E}\left(\xi_m\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)>m\}}\mathbf{1}_B\right) \\ &= \mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)\geq n\}}\mathbf{1}_B\right) ,\end{aligned} \tag{3.60}$$

pioché, per ipotesi,

$$\begin{aligned}\underline{\lim}_m\mathbb{E}\left(\xi_m\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)>m\}}\mathbf{1}_B\right) &\leq \underline{\lim}_m\left|\mathbb{E}\left(\xi_m\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)>m\}}\mathbf{1}_B\right)\right| \\ &\leq \underline{\lim}_m\mathbb{E}\left(|\xi_m|\,\mathbf{1}_{\{\omega\in\Omega:\,\tau_2(\omega)>m\}}\mathbf{1}_B\right) = 0 .\end{aligned} \tag{3.61}$$

■

Teorema 175 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato,* $\{\xi_n\}_{n\geq 0}$ *una martingala e* τ *un tempo d'arresto tale che* $\mathbb{E}(\tau) < \infty$*. Allora, se esiste* $c > 0$ *tali che,* $\forall n > 0$,

$$\mathbb{E}\left(\left|\xi_{n+1}-\xi_n\right|\mathbf{1}_{\{\omega\in\Omega:\,\tau(\omega)\geq n\}}|\mathcal{F}_n\right) \leq c\mathbf{1}_{\{\omega\in\Omega:\,\tau(\omega)\geq n\}} \quad \mathbb{P}\text{-}q.c. , \tag{3.62}$$

1. $\mathbb{E}\left(|\xi_\tau|\right) < \infty$;

2. $\mathbb{E}\left(\xi_\tau\right) = \mathbb{E}\left(\xi_0\right)$.

Dimostrazione: Sia $\{\eta_n\}_{n\geq 0}$, la differenza martingala associata a $\{\xi_n\}_{n\geq 0}$. Allora,

$$\begin{aligned}\mathbb{E}\left(|\xi_\tau|\right) &= \mathbb{E}\left(\left|\sum_{n=0}^{\tau}\eta_n\right|\right) \leq \mathbb{E}\left(\sum_{n=0}^{\tau}|\eta_n|\right) = \mathbb{E}\left(\sum_{n\geq 0}\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)=n\}}\sum_{k=0}^{n}|\eta_k|\right) \\ &= \mathbb{E}\left(\sum_{k\geq 0}\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)\geq k\}}|\eta_k|\right) = \mathbb{E}\left(|\eta_0|\right) + \mathbb{E}\left(\mathbb{E}\left(\sum_{k\geq 1}\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)\geq k\}}|\eta_k|\,\middle|\,\mathcal{F}_k\right)\right) \\ &= \mathbb{E}\left(|\xi_0|\right) + \mathbb{E}\left(\sum_{k\geq 1}\mathbb{E}\left(|\eta_k|\,\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)\geq k\}}|\mathcal{F}_k\right)\right) \leq \mathbb{E}\left(|\xi_0|\right) + c\sum_{k\geq 1}\mathbb{P}\left\{\omega\in\Omega : \tau(\omega)\geq k\right\} \\ &= \mathbb{E}\left(|\xi_0|\right) + c\mathbb{E}\left(\tau\right) < \infty\,. \end{aligned} \tag{3.63}$$

Inoltre,

$$\begin{aligned}\mathbb{E}\left(|\xi_n|\,\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>n\}}\right) &= \mathbb{E}\left(\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>n\}}\left|\sum_{k=0}^{n}\eta_k\right|\right) \leq \mathbb{E}\left(\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>n\}}\sum_{k=0}^{n}|\eta_k|\right) \\ &\leq \mathbb{E}\left(\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>n\}}\sum_{k=0}^{\tau}|\eta_k|\right)\,. \end{aligned} \tag{3.64}$$

Ma poiché $\mathbb{E}\left(\sum_{k=0}^{\tau}|\eta_k|\right) \leq \mathbb{E}\left(|\xi_0|\right) + c\mathbb{E}\left(\tau\right) < \infty$ e la successione di eventi $\{\{\omega\in\Omega : \tau(\omega) > n\}\}_{n\geq 0}$ è decrescente, quindi convergente a $\emptyset$, per il teorema di Lebesgue della convergenza dominata, si ha che

$$\underline{\lim}_n\mathbb{E}\left(|\xi_n|\,\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>n\}}\right) \leq \underline{\lim}_n\mathbb{E}\left(\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)>n\}}\sum_{k=0}^{\tau}|\eta_k|\right) = 0\,. \tag{3.65}$$

Dunque le ipotesi del teorema precedente sono soddisfatte e la tesi segue per $\tau_2 = \tau$ e $\tau_1 = 0$. ∎

Corollario 176 (Identità di Wald) *Sia $\{\xi_n\}_{n\geq 1}$ una successione di v.a.i.i.d. definita su $(\Omega, \mathcal{F}, \mathbb{P})$, tale che $\mathbb{E}\left(|\xi_1|\right) < \infty$, e τ un tempo d'arresto rispetto alla filtrazione $\{\mathcal{F}_n\}_{n\geq 0}$ tale che, $\mathcal{F}_0 := \{\emptyset, \Omega\}$ e $\forall n \geq 1$, $\mathcal{F}_n := \mathcal{F}_n^{\xi}$, per cui $\mathbb{E}\left(\tau\right) < \infty$. Allora,*

$$\mathbb{E}\left(\sum_{i=1}^{\tau}\xi_i\right) = \mathbb{E}\left(\tau\right)\mathbb{E}\left(\xi_1\right)\,. \tag{3.66}$$

Se inoltre $\mathbb{E}\left(\xi_1^2\right) < \infty$, allora

$$\mathbb{E}\left(\left(\sum_{i=1}^{\tau}\xi_i - \tau\mathbb{E}\left(\xi_1\right)\right)^2\right) = \mathbb{E}\left(\tau\right)\mathbb{E}\left(\left(\xi_1 - \mathbb{E}\left(\xi_1\right)\right)^2\right)\,. \tag{3.67}$$

Dimostrazione: La successione di v.a. $\{\eta_n\}_{n\geq 0}$ tale che, $\eta_0 = 0$ e $\forall n \geq 1$, $\eta_n := \sum_{i=1}^n \xi_i - n\mathbb{E}(\xi_1)$, è una martingala rispetto a $\{\mathcal{F}_n\}_{n\geq 0}$ (cfr Esempio 2). Allora, poiché

$$\mathbb{E}\left(\left|\eta_{n+1} - \eta_n\right| |\mathcal{F}_n\right) = \mathbb{E}\left(\left|\xi_{n+1} - \mathbb{E}(\xi_1)\right| |\mathcal{F}_n\right) \leq 2\mathbb{E}(|\xi_1|) < \infty\,, \tag{3.68}$$

le ipotesi del teorema precedente sono soddisfatte e

$$\mathbb{E}(\eta_\tau) = \mathbb{E}\left(\sum_{i=1}^{\tau} \xi_i - \tau\mathbb{E}(\xi_1)\right) = \mathbb{E}(\eta_0) = 0\,. \tag{3.69}$$

Le stesse considerazioni valgono per la martingala $\{\zeta_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $\zeta_n := \eta_n - n\mathbb{E}\left((\xi_1 - \mathbb{E}(\xi_1))^2\right)$. ■

Esempio 7 *(Tempi di arresto per passeggiate aleatorie)* Sia $\{\xi_n\}_{n\geq 1}$ una successione di v.a.i.i.d. su $(\Omega, \mathcal{F}, \mathbb{P})$ distribuite secondo la legge di Bernoulli di parametro $p \in (0,1)$ tali che $\forall n \geq 1$, $\xi_n \in \{-1, 1\}$ e $\{\mathcal{F}_n\}_{n\geq 0}$ la filtrazione tale che, $\mathcal{F}_0 := \{\emptyset, \Omega\}$ e $\forall n \geq 1$, $\mathcal{F}_n := \mathcal{F}_n^\xi$. Se $a, b > 0$, $\tau := \min\{n \geq 1 : S_n = -a, b\}$ è un tempo di Markov. Supponendo che sia un tempo d'arresto e che $\mathbb{E}(\tau) < \infty$, posto $S_0 = 0$ e, $\forall n \geq 1$, $S_n := \sum_{i=1}^n \xi_i$, le successioni di v.a.

1. $\{X_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $X_n := S_n - n(p-q)$;
2. $\{Y_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $Y_n := X_n^2 - 4npq$;
3. $\{Z_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $Z_n := \left(\frac{q}{p}\right)^{S_n}$;

sono martingale rispetto a $\{\mathcal{F}_n\}_{n\geq 0}$. Infatti $\{X_n\}_{n\geq 0}$ e $\{Y_n\}_{n\geq 0}$ sono state già introdotte nell'Esempio 2, mentre per quanto riguarda $\{Z_n\}_{n\geq 0}$, osservando che, $\forall n \geq 0$, $Z_{n+1} = Z_n \left(\frac{q}{p}\right)^{\xi_{n+1}}$, vale

$$\begin{aligned}\mathbb{E}(Z_{n+1}|\mathcal{F}_n) &= Z_n\mathbb{E}\left(\left(\frac{q}{p}\right)^{\xi_{n+1}} |\mathcal{F}_n\right) = Z_n\mathbb{E}\left(\left(\frac{q}{p}\right)^{\xi_1}\right) = Z_n\mathbb{E}\left(\left(\frac{q}{p}\right)^{-1} q + \left(\frac{q}{p}\right) p\right) \\ &= Z_n \quad \mathbb{P}\text{-q.c.}\end{aligned} \tag{3.70}$$

e $\mathbb{E}(|Z_n|) \leq \left(\frac{q}{p} \vee \frac{p}{q}\right)^n < \infty$.

Allora, se $\alpha := \mathbb{P}\{\omega \in \Omega : S_\tau = -a\}$, $\beta := \mathbb{P}\{\omega \in \Omega : S_\tau = b\}$, poiché $\alpha + \beta = 1$, le identità di Wald ed il precedente teorema implicano:

$$\mathbb{E}(S_\tau) = \mathbb{E}(\tau)(p-q) \Longrightarrow \mathbb{E}(\tau) = \frac{\mathbb{E}(S_\tau)}{p-q} = \frac{\beta b - \alpha a}{p-q}\,, \tag{3.71}$$

$$\begin{cases} 1 = \mathbb{E}(Z_0) = \mathbb{E}(Z_\tau) = \mathbb{E}\left(\left(\frac{q}{p}\right)^{S_\tau}\right) = \alpha\left(\frac{q}{p}\right)^{-a} + \beta\left(\frac{q}{p}\right)^b \\ 1 = \alpha + \beta \end{cases} \Longrightarrow \begin{cases} \alpha = \frac{\left(\frac{q}{p}\right)^b - 1}{\left(\frac{q}{p}\right)^b - \left(\frac{p}{q}\right)^a} \\ \beta = \frac{1 - \left(\frac{p}{q}\right)^a}{\left(\frac{q}{p}\right)^b - \left(\frac{p}{q}\right)^a} \end{cases}. \tag{3.72}$$

Nel caso in cui $p = q$,

$$\begin{cases} 0 = \mathbb{E}(S_\tau) = -\alpha a + \beta b \\ 1 = \alpha + \beta \end{cases} \Longrightarrow \begin{cases} \alpha = \frac{b}{a+b} \\ \beta = \frac{a}{a+b} \end{cases} . \tag{3.73}$$

$$0 = \mathbb{E}(Y_0) = \mathbb{E}(Y_\tau) = \mathbb{E}\left(S_\tau^2\right) - \mathbb{E}(\tau) \Longrightarrow \begin{cases} \mathbb{E}(\tau) = \alpha a^2 + \beta b^2 \\ \alpha + \beta = 1 \end{cases} \Longrightarrow \mathbb{E}(\tau) = ab\,, \tag{3.74}$$

Poiché $\{S_n\}_{n\geq 0}$ è una catena di Markov a valori in $\mathbb{Z}$, nel caso $p = q$ rappresenta la passeggiata aleatoria simmetrica sugli interi che è ricorrente. Per cui, se $\sigma := \min\{n \geq 1 : S_n = 1\}$, $\mathbb{P}\{\omega \in \Omega : \sigma(\omega) < \infty\} = 1$ e

$$S_\sigma = \sum_{n\geq 1} S_n \mathbf{1}_{\{\omega\in\Omega\,:\,\sigma(\omega)=n\}} = \sum_{n\geq 1} \mathbf{1}_{\{\omega\in\Omega\,:\,\sigma(\omega)=n\}} = 1\ \mathbb{P} - q.c.\ . \tag{3.75}$$

Perciò, sebbene $\forall n \geq 0$,

$$\mathbb{E}\left(|S_{n+1} - S_n|\,|\mathcal{F}_n\right) = \mathbb{E}\left(\left|\xi_{n+1}\right|\,|\mathcal{F}_n\right) = \mathbb{E}\left(|\xi_1|\right) = 1 < \infty\,, \tag{3.76}$$

$\mathbb{E}(S_\sigma) \neq \mathbb{E}(S_0) = 0$. Quindi, $\mathbb{E}(\sigma) = \infty$.

Corollario 177 (Identità di Wald fondamentale) *Sia $\{\xi_n\}_{n\geq 1}$ una successione di v.a.i.i.d. definita su $(\Omega, \mathcal{F}, \mathbb{P})$. Posto $S_0 = 0$ e $S_n := \sum_{i=1}^n \xi_i$, se $\exists t_0 \neq 0$ tale che $\varphi(t_0) := \mathbb{E}\left(e^{t_0\xi_1}\right) \geq 1$, dato un tempo d'arresto τ rispetto alla filtrazione $\{\mathcal{F}_n\}_{n\geq 0}$, tale che, $\mathcal{F}_0 := \{\emptyset, \Omega\}$ e $\forall n \geq 1$, $\mathcal{F}_n := \mathcal{F}_n^\xi$, per cui $\mathbb{E}(\tau) < \infty$, se esiste $c > 0$ tale che, $\forall n \geq 0$,*

$$|S_n|\,\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)\geq n\}} \leq c\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)\geq n\}} \quad \mathbb{P}\text{-}q.c.\ , \tag{3.77}$$

allora,

$$\mathbb{E}\left(\frac{e^{t_0 S_\tau}}{\varphi^\tau(t_0)}\right) = 1\,. \tag{3.78}$$

Dimostrazione: La successione di v.a. $\{\eta_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $\eta_n := \frac{e^{t_0 S_n}}{\varphi^n(t_0)}$ è una martingala rispetto a $\{\mathcal{F}_n\}_{n\geq 0}$ (cfr Esempio 2). Dunque,

$$\begin{aligned} \mathbb{E}\left(\left|\eta_{n+1} - \eta_n\right|\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)\geq n\}}|\mathcal{F}_n\right) &= \eta_n \mathbb{E}\left(\left|\frac{e^{t_0\xi_{n+1}}}{\varphi(t_0)} - 1\right|\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)\geq n\}}|\mathcal{F}_n\right) \\ &\leq \eta_n \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)\geq n\}}\mathbb{E}\left(\left|\frac{e^{t_0\xi_{n+1}}}{\varphi(t_0)} - 1\right|\right) < \infty\,. \end{aligned} \tag{3.79}$$

Dunque, vale il teorema precedente e $\mathbb{E}(\eta_\tau) = \mathbb{E}(\eta_0) = 1$. ■

3.1.2 Convergenza

Martingale $\mathbf{L}^2$

Definizione 178 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato e* $\{\xi_n\}_{n\geq 0}$ *una martingala. Se* $\sup_{n\geq 0} \mathbb{E}\left(\xi_n^2\right) < \infty$, $\{\xi_n\}_{n\geq 0}$ *è detta* L^2-martingala.

Teorema 179 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato e* $\{\xi_n\}_{n\geq 0}$ *una* L^2*-martingala. Allora, esiste* $\xi_\infty \in L^2(\mathcal{F}_\infty)$, *detto* valore finale *della martingala* $\{\xi_n\}_{n\geq 0}$, *tale che* $\{\xi_n\}_{n\geq 0}$ *converge a* ξ_∞ *in* L^2 *e* $\forall n \geq 0$, $\xi_n = \mathbb{E}(\xi_\infty|\mathcal{F}_n)$. *Inoltre, se* $\{\mathcal{F}_n\}_{n\geq 0}$ *è una filtrazione di* $(\Omega, \mathcal{F}, \mathbb{P})$ *e* $\eta \in L^2(\mathcal{F}_\infty)$, *la successione di v.a.* $\{\xi_n\}_{n\geq 0}$, *tale che* $\forall n \geq 0$, $\xi_n = \mathbb{E}(\eta|\mathcal{F}_n)$ *è una* L^2*-martingala convergente in* L^2 *a* η.

Dimostrazione: Notiamo che $\forall k, n \geq 0$,

$$\mathbb{E}\left(\left(\xi_{n+k} - \xi_n\right)^2\right) = \mathbb{E}\left(\xi_{n+k}^2\right) + \mathbb{E}\left(\xi_n^2\right) - 2\mathbb{E}\left(\xi_{n+k}\xi_n\right), \tag{3.80}$$

ma

$$\mathbb{E}\left(\xi_{n+k}\xi_n\right) = \mathbb{E}\left(\mathbb{E}\left(\xi_{n+k}\xi_n|\mathcal{F}_n\right)\right) = \mathbb{E}\left(\xi_n\mathbb{E}\left(\xi_{n+k}|\mathcal{F}_n\right)\right) = \mathbb{E}\left(\xi_n^2\right) \ . \tag{3.81}$$

Pertanto,

$$\mathbb{E}\left(\xi_{n+k}^2\right) - \mathbb{E}\left(\xi_n^2\right) = \mathbb{E}\left(\left(\xi_{n+k} - \xi_n\right)^2\right) > 0 \ . \tag{3.82}$$

Ponendo $k = 1$ si ha che la successione numerica $\{a_n\}_{n\geq 0}$ tale che $\forall n \geq 0$, $a_n = \mathbb{E}\left(\xi_n^2\right)$ è crescente e, poiché per ipotesi $\sup_{n\geq 0} \mathbb{E}\left(\xi_n^2\right) < \infty$, $\{a_n\}_{n\geq 0}$ è convergente e quindi di Cauchy. Dalla (3.82) segue che anche $\{\xi_n\}_{n\geq 0}$ è di Cauchy in L^2 che è convergente in quanto $L^2(\mathcal{F})$ è completo. Inoltre, $\forall k \geq 0$, $\xi_n = \mathbb{E}\left(\xi_{n+k}|\mathcal{F}_n\right)$. Dunque, se ξ_∞ è il limite di $\{\xi_n\}_{n\geq 0}$ in L^2, poiché $\forall n \geq 0$, $L^2(\mathcal{F}_n) \subseteq L^2(\mathcal{F}_\infty)$ e $L^2(\mathcal{F}_\infty)$ è chiuso come sottospazio di $L^2(\mathcal{F})$, $\xi_\infty \in L^2(\mathcal{F}_\infty)$. Inoltre, $\forall n \geq 0$, $\lim_{k\to\infty} \left\|\xi_{n+k} - \xi_\infty\right\|_{L^2} = 0$ e, dalla limitatezza di $\mathbb{E}(\cdot|\mathcal{F}_n)$ nella topologia di L^2, in quanto estensione continua ad $L^1(\mathcal{F})$ del proiettore su $L^2(\mathcal{F}_n)$ sottospazio di $L^2(\mathcal{F})$, segue

$$\lim_{k\to\infty} \mathbb{E}\left(\xi_{n+k}|\mathcal{F}_n\right) = \mathbb{E}\left(\lim_{k\to\infty} \xi_{n+k}|\mathcal{F}_n\right) = \mathbb{E}\left(\xi_\infty|\mathcal{F}_n\right) \ . \tag{3.83}$$

D'altra parte, se $\{\mathcal{F}_n\}_{n\geq 0}$ è una filtrazione di $(\Omega, \mathcal{F}, \mathbb{P})$ e $\eta \in L^2(\mathcal{F}_\infty)$, la successione di v.a. $\{\xi_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $\xi_n = \mathbb{E}(\eta|\mathcal{F}_n)$ è una L^2-martingala. Infatti, $\forall n, k \geq 0$, poiché $\mathcal{F}_n \subseteq \mathcal{F}_{n+k}$,

$$\mathbb{E}\left(\xi_{n+k}|\mathcal{F}_n\right) = \mathbb{E}\left(\mathbb{E}\left(\eta|\mathcal{F}_{n+k}\right)|\mathcal{F}_n\right) = \mathbb{E}\left(\eta|\mathcal{F}_n\right) = \xi_n \quad \mathbb{P}\text{-q.c.}. \tag{3.84}$$

Inoltre, poiché $\mathbb{E}(\cdot|\mathcal{F}_n)$ coincide $\mathbb{P}$-q.c. con l'estensione continua ad $L^1(\mathcal{F})$ del proiettore su $L^2(\mathcal{F}_n)$, $\forall n \geq 0$, $\|\xi_n\|_{L^2} = \|\mathbb{E}(\eta|\mathcal{F}_n)\|_{L^2} \leq \|\eta\|_{L^2} < \infty$. Quindi per la prima parte del teorema, la successione $\{\xi_n\}_{n\geq 0}$ è convergente a $\xi_\infty \in L^2(\mathcal{F}_\infty)$. Resta da dimostrare che $\xi_\infty = \eta$, ma ciò deriva dal risultato seguente. ■

Lemma 180 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato. Allora,* $\forall \eta \in L^2(\mathcal{F})$ *la successione* $\{\mathbb{E}(\eta|\mathcal{F}_n)\}_{n\geq 0}$ *converge a* $\mathbb{E}(\eta|\mathcal{F}_\infty)$ *in* L^2.

Dimostrazione: Poiché, $\forall n \geq 0$, $L^2(\mathcal{F}_n) \subseteq L^2(\mathcal{F}_\infty)$, allora, $\mathcal{H} \subset L^2(\mathcal{F}_\infty)$, dove $\mathcal{H} := \overline{\bigcup_{n\geq 0} L^2(\mathcal{F}_n)}^{L^2}$ è il più piccolo sottospazio di $L^2(\mathcal{F})$ contenente ciascuno dei sottospazi $L^2(\mathcal{F}_n)$. D'altra parte, $\mathcal{F}_\infty$ è la classe monotona generata da $\bigcup_{n\geq 0}\mathcal{F}_n$ e $\forall B \in \bigcup_{n\geq 0}\mathcal{F}_n$, $\mathbf{1}_B \in \mathcal{H}$. Sia ora $\mathcal{M} := \{A \in \mathcal{P}(\Omega) : \mathbf{1}_A \in \mathcal{H}\}$. Poiché se $\{A_n\}_{n\geq 0} \subset \mathcal{M}$ è una successione crescente il cui limite è A_∞, la successione $\{\mathbf{1}_{A_n}\}_{n\geq 0}$ converge a $\mathbf{1}_{A_\infty}$ $\mathbb{P}$-q.c. e per il teorema della convergenza dominata generalizzato coneverge anche il L^2, cioè $\mathbf{1}_{A_\infty} \in \mathcal{H}$. Inoltre, siccome $\mathbb{P}(\Omega) = 1$, considerando la successione degli insiemi complementari degli A_n, che è decrescente, anche $\mathbf{1}_{A^c_\infty} \in \mathcal{H}$. Quindi anche $\mathcal{M}$ è una classe monotona, perciò $\mathcal{M} = \mathcal{F}_\infty$ e $\mathcal{H} = L^2(\mathcal{F}_\infty)$. Ma $\{\mathbb{E}(\eta|\mathcal{F}_n)\}_{n\geq 0}$ converge in L^2 a $\Pi_{\mathcal{H}}(\eta)$, dove $\Pi_{\mathcal{H}}$ è il proiettore su $\mathcal{H}$, da cui segue la tesi. ■

Proprietà delle L^2-martingale Sia $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ uno spazio filtrato e $\{\xi_n\}_{n\geq 0}$ una L^2-martingala. Allora, poiché $\mathbb{R} \ni x \longmapsto f(x) = x^2 \in \mathbb{R}^+$ è una funzione convessa, per la disuguaglianza di Jensen, la successione di v.a. $\{\eta_n\}_{n\geq 0}$ tale che $\forall n \geq 0$, $\eta_n := \xi_n^2$ è una submartingala e inoltre $\sup_{n\geq 0}\mathbb{E}(|\eta_n|) = \sup_{n\geq 0}\mathbb{E}(\xi_n^2) < \infty$.

Definizione 181 *Sia $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ uno spazio filtrato e $\{\xi_n\}_{n\geq 0}$ una L^2-martingala. La successione prevedibile di v.a. $\{\langle\xi^2\rangle_n\}_{n\geq 0}$, compensatore nella decomposizione di Doob-Meyer della submartingala $\{\xi_n^2\}_{n\geq 0}$, è detto* caratteristica quadratica *di $\{\xi_n\}_{n\geq 0}$.*

Data una L^2-martingale $\{\xi_n\}_{n\geq 0}$ su $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$, indichiamo con $\{m_n^\xi\}_{n\geq 0}$, tale che, $\forall n \geq 0$, $m_n^\xi := \xi_n^2 - \langle\xi^2\rangle_n$, la martingala associata alla submartingala $\{\xi_n^2\}_{n\geq 0}$ nella decomposizione di Doob-Meyer.

Osservazione 182 *Dalla (3.45) segue che, $\forall n \geq 0$,*

$$\langle\xi^2\rangle_n = \sum_{i=0}^{n-1}\mathbb{E}\left(\xi_{i+1}^2 - \xi_i^2|\mathcal{F}_i\right) = \sum_{i=0}^{n-1}\mathbb{E}\left(\left(\xi_{i+1} - \xi_i\right)^2|\mathcal{F}_i\right) = \sum_{i=1}^{n}\mathbb{E}\left(\left(\xi_i - \xi_{i-1}\right)^2|\mathcal{F}_{i-1}\right), \tag{3.85}$$

in quanto, essendo $\{\xi_n\}_{n\geq 0}$ una martingala, $\forall n \geq 0$,

$$\mathbb{E}\left(\xi_n\xi_{n-1}|\mathcal{F}_{n-1}\right) = \xi_{n-1}\mathbb{E}\left(\xi_n|\mathcal{F}_{n-1}\right) = \xi_{n-1}^2 \quad \mathbb{P}\text{-}q.c.. \tag{3.86}$$

Inoltre, $\forall k \leq n$,

$$\begin{aligned}\mathbb{E}\left((\xi_n - \xi_k)^2|\mathcal{F}_k\right) &= \mathbb{E}\left(\xi_n^2 - \xi_k^2|\mathcal{F}_k\right) = \mathbb{E}\left(m_n^\xi + \langle\xi^2\rangle_n - m_k^\xi + \langle\xi^2\rangle_k|\mathcal{F}_k\right) \\ &= \mathbb{E}\left(m_n^\xi - m_k^\xi|\mathcal{F}_k\right) + \mathbb{E}\left(\langle\xi^2\rangle_n - \langle\xi^2\rangle_k|\mathcal{F}_k\right) \\ &= \mathbb{E}\left(\langle\xi^2\rangle_n - \langle\xi^2\rangle_k|\mathcal{F}_k\right) \quad \mathbb{P}\text{-}q.c..\end{aligned} \tag{3.87}$$

Date $\{\xi_n\}_{n\geq 0}, \{\eta_n\}_{n\geq 0}$ due L^2-martingale su $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$, poiché anche $\{X_n\}_{n\geq 0}, \{Y_n\}_{n\geq 0}$ tali che, $\forall n \geq 0$, $X_n = \xi_n + \eta_n$ e $Y_n = \xi_n - \eta_n$ sono L^2-martingale, $\{m_n^X\}_{n\geq 0}, \{m_n^Y\}_{n\geq 0}$ sono anche martingale e quindi lo è anche $\{Z_n\}_{n\geq 0}$ tale che, $\forall n \geq 0$,

$$\begin{aligned}Z_n = m_n^X - m_n^Y &= (\xi_n + \eta_n)^2 - \langle(\xi + \eta)^2\rangle_n - (\xi_n - \eta_n)^2 + \langle(\xi - \eta)^2\rangle_n \\ &= 4\xi_n\eta_n + \langle(\xi - \eta)^2\rangle_n - \langle(\xi + \eta)^2\rangle_n .\end{aligned} \tag{3.88}$$

Definizione 183 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato e* $\{\xi_n\}_{n\geq 0}, \{\eta_n\}_{n\geq 0}$ *due* L^2*-martingale. La successione di v.a.* $\{\langle \xi, \eta\rangle_n\}_{n\geq 0}$ *su* $(\Omega, \mathcal{F}, \mathbb{P})$ *tale che,* $\forall n \geq 0$,

$$\langle \xi, \eta\rangle_n := \frac{1}{4}\left[\left\langle (\xi+\eta)^2\right\rangle_n - \left\langle (\xi-\eta)^2\right\rangle_n\right] \tag{3.89}$$

è detta caratteristica mutua *di* $\{\xi_n\}_{n\geq 0}$ *e* $\{\eta_n\}_{n\geq 0}$.

Osservazione 184 *Notiamo che, se* $\{\langle \xi, \eta\rangle_n\}_{n\geq 0}$ *è la caratteristica mutua di due* L^2*-martingale* $\{\xi_n\}_{n\geq 0}$ *e* $\{\eta_n\}_{n\geq 0}$ *su* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$, *allora, per definizione* $\{\langle \xi, \eta\rangle_n\}_{n\geq 0}$ *è prevedibile rispetto alla filtrazione* $\{\mathcal{F}_n\}_{n\geq 0}$. *Inoltre, dalla (3.88) si ottiene che la successione di v.a.* $\{\zeta_n\}_{n\geq 0}$ *tale che,* $\forall n \geq 0$,

$$\zeta_n := \frac{1}{4} Z_n = \xi_n\eta_n - \langle \xi, \eta\rangle_n \tag{3.90}$$

è una martingala. D'altra parte, poiché, $\forall k \leq n$,

$$\begin{aligned}
\mathbb{E}\left((\xi_n-\xi_k)(\eta_n-\eta_k)|\mathcal{F}_k\right) &= \mathbb{E}\left(\xi_n\eta_n - \xi_k\eta_n - \xi_n\eta_k + \xi_k\eta_k|\mathcal{F}_k\right) = \\
&= \mathbb{E}\left(\xi_n\eta_n|\mathcal{F}_k\right) - \xi_k\mathbb{E}\left(\eta_n|\mathcal{F}_k\right) - \eta_k\mathbb{E}\left(\xi_n|\mathcal{F}_k\right) + \xi_k\eta_k = \\
&= \mathbb{E}\left(\xi_n\eta_n|\mathcal{F}_k\right) - \xi_k\eta_k = \mathbb{E}\left(\xi_n\eta_n - \xi_k\eta_k|\mathcal{F}_k\right) = \\
&= \mathbb{E}\left(\zeta_n + \langle \xi, \eta\rangle_n - \zeta_k - \langle \xi, \eta\rangle_k|\mathcal{F}_k\right) = \mathbb{E}\left(\zeta_n - \zeta_k|\mathcal{F}_k\right) + \mathbb{E}\left(\langle \xi, \eta\rangle_n - \langle \xi, \eta\rangle_k|\mathcal{F}_k\right) \\
&= \mathbb{E}\left(\langle \xi, \eta\rangle_n - \langle \xi, \eta\rangle_k|\mathcal{F}_k\right) \quad \mathbb{P}\text{-}q.c. \ .
\end{aligned} \tag{3.91}$$

si ha che $\forall n \geq 0$,

$$\langle \xi, \eta\rangle_n = \sum_{i=0}^{n-1}\mathbb{E}\left(\left(\xi_{i+1}-\xi_i\right)\left(\eta_{i+1}-\eta_i\right)|\mathcal{F}_i\right) = \sum_{i=1}^{n}\mathbb{E}\left(\left(\xi_i-\xi_{i-1}\right)\left(\eta_i-\eta_{i-1}\right)|\mathcal{F}_{i-1}\right) \tag{3.92}$$

e quindi, $\langle \xi, \xi\rangle_n = \left\langle \xi^2\right\rangle_n$. *Inoltre, quanto sopra implica che gli incrementi di* $\{\xi_n\}_{n\geq 0}$ *e* $\{\eta_n\}_{n\geq 0}$, *per* $k \leq n$, *risultano scorrelati condizionatamente ad* $\mathcal{F}_k$ *se e solo se,* $\forall n \geq 0$, $\langle \xi, \eta\rangle_n = 0$ *ovvero, se e solo se la successione di v.a.* $\{W_n\}_{n\geq 0}$ *tale che,* $\forall n \geq 0$,

$$W_n := \xi_n\eta_n = \zeta_n + \langle \xi, \eta\rangle_n \tag{3.93}$$

è una martingala.

Sia $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ uno spazio filtrato e $\{\xi_n\}_{n\geq 0}, \{\eta_n\}_{n\geq 0}$ due successioni di v.a. su $(\Omega, \mathcal{F}, \mathbb{P})$, adattate alla filtrazione $\{\mathcal{F}_n\}_{n\geq 0}$. Poiché $\forall \alpha, \beta \in \mathbb{R}$ la successione di v.a. $\{\zeta_n\}_{n\geq 0}$ su $(\Omega, \mathcal{F}, \mathbb{P})$, tale che, $\forall n \geq 0, \zeta_n := \alpha\xi_n + \beta\eta_n$ è adattata a $\{\mathcal{F}_n\}_{n\geq 0}$, sia $\mathfrak{L}^2$ la chiusura nella norma $\mathfrak{L} \ni \{\xi_n\}_{n\geq 0} \longmapsto \sup_{n\geq 0}\|\xi_n\|_{L^2} \in \mathbb{R}^+$ dello spazio lineare $\mathfrak{L}$ delle successioni di v.a. $\{\xi_n\}_{n\geq 0}$ su $(\Omega, \mathcal{F}, \mathbb{P})$, adattate alla filtrazione $\{\mathcal{F}_n\}_{n\geq 0}$. Allora, $\mathfrak{M}^2$, collezione delle L^2-martingale su $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$, è un sottospazio chiuso di $\mathfrak{L}^2$ che è anche uno spazio di Banach (cfr Osservazione 199). Se $\mathfrak{M}_0^2$ indica il sottospazio $\mathfrak{M}^2$ i cui elementi sono tali che il primo termine della successione è uguale a zero, $\langle\cdot,\cdot\rangle$ risulta un'applicazione bilineare da $\mathfrak{M}_0^2 \times \mathfrak{M}_0^2$ a valori in $\mathfrak{P}_0^2$, sottospazio lineare di $\mathfrak{L}^2$ generato dalle successioni di v.a. prevedibili non decrescenti su $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$. Ovvero, indicando con $\boldsymbol{\xi}$ il generico elemento $\{\xi_n\}_{n\geq 0} \in \mathfrak{M}_0^2$ e con $\langle \boldsymbol{\xi}, \boldsymbol{\eta}\rangle$ la caratteristica mutua di $\boldsymbol{\xi}$ e $\boldsymbol{\eta}$, allora, $\langle \boldsymbol{\xi}, \boldsymbol{\xi}\rangle = \{\langle \xi^2\rangle_n\}_{n\geq 0} = \langle \boldsymbol{\xi}\rangle$, e:

1. $\forall \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathfrak{M}_0^2, \quad \langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle = \langle \boldsymbol{\eta}, \boldsymbol{\xi} \rangle$;

2. $\forall \alpha, \beta \in \mathbb{R}$ e $\boldsymbol{\xi}, \boldsymbol{\eta}, \boldsymbol{\zeta} \in \mathfrak{M}_0^2, \quad \langle \alpha \boldsymbol{\xi} + \beta \boldsymbol{\eta}, \boldsymbol{\zeta} \rangle = \alpha \langle \boldsymbol{\xi}, \boldsymbol{\zeta} \rangle + \beta \langle \boldsymbol{\eta}, \boldsymbol{\zeta} \rangle$.

Inoltre, fissato $\boldsymbol{\xi} \in \mathfrak{M}_0^2$, dall'osservazione precedente, si ha che l'operatore lineare da $\langle \boldsymbol{\xi} \cdot \rangle : \mathfrak{M}_0^2 \longmapsto \mathfrak{P}_0^2$ ha come nucleo il sottospazio di $\mathfrak{M}_0^2$ generato dalle martingale i cui incrementi, $\forall n, k \geq 0$ tali che $k \leq n$, risultano scorrelati condizionatamente ad $\mathcal{F}_k$. D'altra parte, dalla definizione di caratteristica mutua, segue che $\forall \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathfrak{M}_0^2$,

$$|\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle|^2 \leq \langle \boldsymbol{\xi}, \boldsymbol{\xi} \rangle \langle \boldsymbol{\eta}, \boldsymbol{\eta} \rangle = \langle \boldsymbol{\xi} \rangle \langle \boldsymbol{\eta} \rangle \ . \tag{3.94}$$

Definizione 185 *Siano $\{\xi_n\}_{n\geq 0}, \{\eta_n\}_{n\geq 0}$ due successioni di v.a. su $(\Omega, \mathcal{F}, \mathbb{P})$. La successione di v.a. $\{[\xi, \eta]_n\}_{n\geq 0}$ su $(\Omega, \mathcal{F}, \mathbb{P})$ tale che, $\forall n \geq 0$,*

$$[\xi, \eta]_n := \sum_{i=0}^{n-1} (\xi_{i+1} - \xi_i) (\eta_{i+1} - \eta_i) = \sum_{i=1}^{n} (\xi_i - \xi_{i-1}) (\eta_i - \eta_{i-1}) \tag{3.95}$$

è detta covariazione quadratica *di $\{\xi_n\}_{n\geq 0}$ e $\{\eta_n\}_{n\geq 0}$, mentre la successione di v.a. $\{[\xi^2]_n\}_{n\geq 0}$ su $(\Omega, \mathcal{F}, \mathbb{P})$ tale che, $\forall n \geq 0$,*

$$[\xi]_n := [\xi, \xi]_n = \sum_{i=0}^{n-1} (\xi_{i+1} - \xi_i)^2 = \sum_{i=1}^{n} (\xi_i - \xi_{i-1})^2 \tag{3.96}$$

è detta variazione quadratica *della successione di v.a. $\{\xi_n\}_{n\geq 0}$.*

Osservazione 186 *Dalla precedente definizione segue che, se $\{\xi_n\}_{n\geq 0}, \{\eta_n\}_{n\geq 0}$ sono due successioni di v.a. su $(\Omega, \mathcal{F}, \mathbb{P})$, allora, $\forall n \geq 0$,*

$$[\xi, \eta]_n = \frac{1}{4} \left[\left[(\xi + \eta)^2 \right]_n - \left[(\xi - \eta)^2 \right]_n \right] \ . \tag{3.97}$$

Osservazione 187 *Dato $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ uno spazio filtrato e $\boldsymbol{\xi}, \boldsymbol{\eta} \in \mathfrak{M}_0^2$, dalla (3.92) segue che, ponendo $\forall n \geq 0$, $[\boldsymbol{\xi}, \boldsymbol{\eta}]_n = [\xi, \eta]_n$, si ha*

$$\begin{aligned} \langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle_n = \langle \xi, \eta \rangle_n &= \sum_{i=1}^{n} \mathbb{E} \left((\xi_i - \xi_{i-1}) (\eta_i - \eta_{i-1}) | \mathcal{F}_{i-1} \right) \\ &= \sum_{i=1}^{n} \mathbb{E} \left(\xi_i \eta_i - \xi_{i-1} \eta_{i-1} | \mathcal{F}_{i-1} \right) = \sum_{i=1}^{n} \mathbb{E} \left([\xi, \eta]_i - [\xi, \eta]_{i-1} | \mathcal{F}_{i-1} \right) \\ &= \sum_{i=1}^{n} \mathbb{E} \left([\boldsymbol{\xi}, \boldsymbol{\eta}]_i - [\boldsymbol{\xi}, \boldsymbol{\eta}]_{i-1} | \mathcal{F}_{i-1} \right) \ . \end{aligned} \tag{3.98}$$

Inoltre, poiché, $\forall n \geq 0, \boldsymbol{\xi} \in \mathfrak{M}_0^2$,

$$\mathbb{E}\left([\boldsymbol{\xi}]_n \left|\mathcal{F}_{n-1}\right.\right) = \mathbb{E}\left(\sum_{i=1}^{n}\left(\xi_i - \xi_{i-1}\right)^2 \middle| \mathcal{F}_{n-1}\right) = \sum_{i=1}^{n-1}\left(\xi_i - \xi_{i-1}\right)^2 + \tag{3.99}$$
$$+\mathbb{E}\left(\left(\xi_n - \xi_{n-1}\right)^2 \middle| \mathcal{F}_{n-1}\right) = [\boldsymbol{\xi}]_{n-1} + \langle\boldsymbol{\xi}\rangle_n - \langle\boldsymbol{\xi}\rangle_{n-1} .$$

Siccome la successione $\{\langle\boldsymbol{\xi}\rangle_n\}_{n\geq 0}$ è non decrescente, $\langle\boldsymbol{\xi}\rangle_n - \langle\boldsymbol{\xi}\rangle_{n-1} \geq 0$ e

$$\mathbb{E}\left([\boldsymbol{\xi}]_n \left|\mathcal{F}_{n-1}\right.\right) = [\boldsymbol{\xi}]_{n-1} + \langle\boldsymbol{\xi}\rangle_n - \langle\boldsymbol{\xi}\rangle_{n-1} \geq [\boldsymbol{\xi}]_{n-1} , \tag{3.100}$$

ovvero, essendo

$$\mathbb{E}\left(\left|[\boldsymbol{\xi}]_n\right|\right) \leq 2n \sup_{n\geq 0} \mathbb{E}\left(\xi_n^2\right) < \infty , \tag{3.101}$$

la successione di v.a. $\{[\boldsymbol{\xi}]_n\}_{n\geq 0}$ è una submartingala su $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$. Applicando la decomposizione di Doob-Meyer a $[\boldsymbol{\xi}]$ si ha che esistono $M^{\xi} \in \mathfrak{M}_0^2$ e $A^{\xi} \in \mathfrak{P}_0^2$ tali che, $\forall n \geq 0$, $[\boldsymbol{\xi}]_n = M_n^{\xi} + A_n^{\xi}$, ma dalla (3.99) segue che

$$A_n^{\xi} = \sum_{i=1}^{n} \mathbb{E}\left([\boldsymbol{\xi}]_i - [\boldsymbol{\xi}]_{i-1} \left|\mathcal{F}_{i-1}\right.\right) = \sum_{i=1}^{n}\left(\mathbb{E}\left([\boldsymbol{\xi}]_i \left|\mathcal{F}_{i-1}\right.\right) - [\boldsymbol{\xi}]_{i-1}\right) = \tag{3.102}$$
$$= \sum_{i=1}^{n}\left(\langle\boldsymbol{\xi}\rangle_i - \langle\boldsymbol{\xi}\rangle_{i-1}\right) = \langle\boldsymbol{\xi}\rangle_n = \left\langle\xi^2\right\rangle_n .$$

Allora, $\forall n \geq 0$,

$$\xi_n^2 = m_n^{\xi} + \left\langle\xi^2\right\rangle_n = \xi_n^2 - [\boldsymbol{\xi}]_n + [\boldsymbol{\xi}]_n = \xi_n^2 - [\boldsymbol{\xi}]_n + M_n^{\xi} + \left\langle\xi^2\right\rangle_n \tag{3.103}$$

ovvero, $m_n^{\xi} = \xi_n^2 - [\boldsymbol{\xi}]_n + M_n^{\xi}$ cioè, $\xi_n^2 - [\boldsymbol{\xi}]_n = m_n^{\xi} - M_n^{\xi}$. Dunque, anche la successione di v.a. $\left\{\mu_n^{\xi}\right\}_{n\geq 0}$ tale che, $\forall n \geq 0$, $\mu_n^{\xi} := \xi_n^2 - [\boldsymbol{\xi}]_n$ è una L^2-martingala.

Martingale regolari

Definizione 188 *Sia $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ uno spazio filtrato e $\{\xi_n\}_{n\geq 0}$ una martingala. Se esiste $\eta \in L^1(\mathcal{F})$, tale che $\forall n \geq 0$, $\xi_n = \mathbb{E}(\eta|\mathcal{F}_n)$, $\{\xi_n\}_{n\geq 0}$ è detta martingala* regolare.

Osservazione 189 *Poiché $L^2(\mathcal{F}_\infty) \subseteq L^2(\mathcal{F}) \subset L^1(\mathcal{F})$ ogni L^2-martingala è regolare.*

Teorema 190 *Sia $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ uno spazio filtrato e $\{\xi_n\}_{n\geq 0}$ una martingala regolare, tale che, $\forall n \geq 0$, $\xi_n = \mathbb{E}(\eta|\mathcal{F}_n)$, con $\eta \in L^1(\mathcal{F})$. Posto $\xi_\infty := \mathbb{E}(\eta|\mathcal{F}_\infty)$, detto* valore finale *della martingala $\{\xi_n\}_{n\geq 0}$, allora, $\forall n \geq 0$, $\xi_n = \mathbb{E}(\xi_\infty|\mathcal{F}_n)$ e $\{\xi_n\}_{n\geq 0}$ converge a ξ_∞ in $L^1(\mathcal{F})$.*

Dimostrazione: *Sia, $\forall N > 0$,*

$$\mathbb{R} \ni t \longmapsto \varphi(t) := -N\mathbf{1}_{(-\infty,-N]}(t) + t\mathbf{1}_{[-N,N]}(t) + N\mathbf{1}_{[N.\infty)}(t) \in \mathbb{R} . \tag{3.104}$$

Allora, $\forall N \in \mathbb{N}^+$, posto $\eta_N := \varphi(\eta)$, la successione $\{\eta_N\}_{N\geq 1}$ converge a $\eta \in L^1$. Inoltre, posto, $\forall n \geq 0$, $\eta_{N,n} := \mathbb{E}(\eta_N|\mathcal{F}_n)$ si ha che

$$\begin{aligned}\mathbb{E}\left(\left|\eta_{N,n} - \xi_n\right|\right) &= \mathbb{E}\left(\left|\mathbb{E}(\eta_N|\mathcal{F}_n) - \xi_n\right|\right) = \mathbb{E}\left(\left|\mathbb{E}(\eta_N|\mathcal{F}_n) - \mathbb{E}(\eta|\mathcal{F}_n)\right|\right) \\ &\leq \mathbb{E}\left(\mathbb{E}\left(\left|\eta_N - \eta\right| |\mathcal{F}_n\right)\right) \leq \mathbb{E}\left(\left|\eta_N - \eta\right|\right)\ ,\end{aligned} \tag{3.105}$$

ovvero $\{\eta_{N,n}\}_{N\geq 1}$ converge a ξ_n in L^1. Quindi, poiché $\forall N \geq 1$ $\left\|\eta_{N,n}\right\|_{L^\infty} \leq N$, $\{\eta_{N,n}\}_{n\geq 0}$ è una L^2-martingala, perciò esiste $\eta_{\infty,N} \in L^2(\mathcal{F}_\infty)$ tale che $\eta_{N,n} = \mathbb{E}\left(\eta_{\infty,N}|\mathcal{F}_n\right)$. Allora, per il Lemma 180 $\eta_{\infty,N} = \mathbb{E}(\eta_N|\mathcal{F}_\infty)$ e, poiché $L^2(\mathcal{F}_\infty) \subset L^1(\mathcal{F}_\infty)$, $\{\eta_{N,n}\}_{n\geq 0}$ converge a $\eta_{\infty,N}$ anche in L^1. Inoltre, siccome $\{\eta_N\}_{N\geq 0}$ converge a $\eta \in L^1$,

$$\mathbb{E}\left(\left|\mathbb{E}(\eta_N|\mathcal{F}_\infty) - \mathbb{E}(\eta|\mathcal{F}_\infty)\right|\right) \leq \mathbb{E}\left(\mathbb{E}\left(\left|\eta_N - \eta\right| |\mathcal{F}_\infty\right)\right) = \mathbb{E}\left(\left|\eta_N - \eta\right|\right)\ , \tag{3.106}$$

cioè $\{\mathbb{E}(\eta_N|\mathcal{F}_\infty)\}_{N\geq 1}$ converge a $\mathbb{E}(\eta|\mathcal{F}_\infty)$ in L^1. ■

Proposizione 191 (Disuguaglianza massimale di Doob) *Sia $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ uno spazio filtrato e $\{\xi_n\}_{n\geq 0}$ una martingala. Allora, $\forall r > 0$ e $\forall n \geq 0$,*

$$\mathbb{P}\left\{\omega \in \Omega : \sup_{0\leq k<n} |\xi_k| > r\right\} \leq \frac{2}{r}\left(\mathbb{E}(|\xi_n|) + \mathbb{E}(|\xi_0|)\right) \tag{3.107}$$

e

$$\mathbb{P}\left\{\omega \in \Omega : \lim_{n\to\infty} \sup_{0\leq k<n} |\xi_k| > r\right\} \leq \frac{4}{r}\sup_{n\geq 0}\mathbb{E}(|\xi_n|)\ . \tag{3.108}$$

Dimostrazione: Sia $A_n^r := \left\{\omega \in \Omega : \sup_{0\leq k<n} \xi_k > r\right\}$ e sia

$$\tau := \inf\{0 \leq k < n : \xi_k > r\}\mathbf{1}_{A_n^r} + n\mathbf{1}_{(A_n^r)^c}\,. \tag{3.109}$$

Allora, $\{\omega \in \Omega : \tau(\omega) > n\} = \emptyset$ e

$$\{\omega \in \Omega : \tau(\omega) > m\} = \bigcap_{0\leq k\leq m}\{\omega \in \Omega : \xi_k(\omega) \leq r\} \in \mathcal{F}_m \subset \mathcal{F}_n\,. \tag{3.110}$$

Quindi τ è un tempo d'arresto. Considerando la martingala arrestata $\{\xi_{n\wedge\tau}\}_{n\geq 0}$,

$$\begin{aligned}\mathbb{E}(\xi_{n\wedge\tau}) &= \mathbb{E}\left(\xi_{n\wedge\tau}\mathbf{1}_{\{\omega\in\Omega:\tau(\omega)<n\}}\right) + \mathbb{E}\left(\xi_{n\wedge\tau}\mathbf{1}_{\{\omega\in\Omega:\tau(\omega)=n\}}\right) = \\ &= \mathbb{E}\left(\xi_{n\wedge\tau}\mathbf{1}_{\{\omega\in\Omega:\tau(\omega)<n\}}\right) + \mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\tau(\omega)=n\}}\right)\ .\end{aligned} \tag{3.111}$$

Dunque, poiché $\forall\omega \in \{\omega \in \Omega : \tau(\omega) < n\}$, $\xi_{n\wedge\tau} > r$,

$$\mathbb{E}(\xi_{n\wedge\tau}) > r\mathbb{P}\{\omega \in \Omega : \tau(\omega) < n\} + \mathbb{E}\left(\xi_n\mathbf{1}_{\{\omega\in\Omega:\tau(\omega)=n\}}\right)\ , \tag{3.112}$$

ma, dalla (3.110) si ha che

$$\{\omega \in \Omega : \tau(\omega) < n\} = \bigcup_{k=0}^{n-1}\{\omega \in \Omega : \xi_k(\omega) > r\} = A_n^r\,, \tag{3.113}$$

inoltre, poiché $\{\xi_{n\wedge\tau}\}_{n\geq 0}$ è una una martingala, $\mathbb{E}(\xi_{n\wedge\tau}) = \mathbb{E}(\xi_{0\wedge\tau}) = \mathbb{E}(\xi_0)$. Allora,

$$r\mathbb{P}(A_n^r) < \mathbb{E}(\xi_0) - \mathbb{E}\left(\xi_n \mathbf{1}_{\{\omega\in\Omega:\tau(\omega)=n\}}\right) \leq \mathbb{E}(\xi_0) + \mathbb{E}(|\xi_n|) \ . \tag{3.114}$$

Poiché

$$\left\{\omega \in \Omega : \sup_{0\leq k<n} |\xi_k| > r\right\} = \left\{\omega \in \Omega : \sup_{0\leq k<n} \xi_k > r\right\} \cup \left\{\omega \in \Omega : \sup_{0\leq k<n} (-\xi_k) > r\right\} , \tag{3.115}$$

considerando la successione $\{-\xi_n\}_{n\geq 0}$ si ha la (3.107).

D'altra parte, la successione di v.a. $\{\eta_n\}_{n\geq 1}$ tale che, $\forall n \geq 1$, $\eta_n := \sup_{0\leq k<n} |\xi_k|$ è crescente. Dunque, se $\lim_{n\to\infty} \eta_n > r$ allora, $\forall \varepsilon > 0$, $\exists N_\varepsilon$ tale che $\forall n > N_\varepsilon$, $\eta_n > r - \varepsilon$. Quindi,

$$\mathbb{P}\left\{\omega \in \Omega : \lim_{n\to\infty} \sup_{0\leq k<n} |\xi_k| > r\right\} = \mathbb{P}\left\{\bigcap_{\varepsilon>0} \bigcup_{N\geq 1} \bigcap_{n>N} \left\{\omega \in \Omega : \sup_{0\leq k<n} |\xi_k| > r - \varepsilon\right\}\right\} \tag{3.116}$$

$$\leq \mathbb{P}\left\{\bigcap_{\varepsilon>0} \bigcup_{n\geq 1} \left\{\omega \in \Omega : \sup_{0\leq k<n} |\xi_k| > r - \varepsilon\right\}\right\} = \mathbb{P}\left\{\bigcap_{\varepsilon>0} \left\{\omega \in \Omega : \sup_{n\geq 1} \sup_{0\leq k<n} |\xi_k| > r - \varepsilon\right\}\right\}$$

$$\leq \frac{4}{r-\varepsilon} \sup_{n\geq 0} \mathbb{E}(|\xi_n|) \ .$$

■

Proposizione 192 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato e* $\{\xi_n\}_{n\geq 0}$ *una martingala regolare di valore finale* ξ_∞*. Allora,* $\{\xi_n\}_{n\geq 0}$ *converge a* ξ_∞ $\mathbb{P}$*-q.c.*

Dimostrazione: Dato $m \geq 0$, $\forall n \geq m$, sia $\eta_n := \xi_n - \xi_m$. Allora la successione di v.a. $\{\eta_n\}_{n\geq m}$ è una martingala regolare convergente in L^1 ad una v.a. η_∞. Inoltre,

$$\sup_{n\geq m} \mathbb{E}(|\eta_n|) \leq \sup_{n\geq m} \mathbb{E}(|\xi_n - \xi_\infty| + |\xi_\infty - \xi_m|) = \sup_{n\geq m} \mathbb{E}(|\xi_n - \xi_\infty|) + \mathbb{E}(|\xi_\infty - \xi_m|) \ . \tag{3.117}$$

Pertanto, poiché $\{\xi_n\}_{n\geq 0}$ converge a ξ_∞ in L^1, $\forall \varepsilon > 0$, $\exists m_\varepsilon$ tale che $\forall n > m_\varepsilon$, $\mathbb{E}(|\xi_\infty - \xi_n|) < \varepsilon$. Allora, $\sup_{n\geq m} \mathbb{E}(|\eta_n|) < 2\varepsilon$ e, per la disuguaglianza massimale di Doob (3.108), $\forall m \geq m_\varepsilon, r > 0$,

$$\mathbb{P}\left\{\omega \in \Omega : \sup_{n\geq m} |\xi_n - \xi_m| > r\right\} \leq \frac{2\varepsilon}{r} \ . \tag{3.118}$$

Poiché la successione $\{\zeta_m\}_{m\geq 0}$ tale che $\zeta_m := \sup_{n\geq m} |\xi_n - \xi_m|$ è decrescente, $\forall r > 0$, si ha $\mathbb{P}\{\omega \in \Omega : \lim_{m\to\infty} \zeta_m > r\} = 0$ ovvero, $\lim_{m\to\infty} \zeta_m = 0$ $\mathbb{P}$-q.c., cioè che $\{\xi_n\}_{n\geq 0}$ è fondamentale $\mathbb{P}$-q.c. e quindi convergente $\mathbb{P}$-q.c. ad una v.a. ξ. Ma, poiché $\{\xi_n\}_{n\geq 0}$ converge a ξ_∞ in L^1, esiste una sottosuccessione $\{\xi_{n_k}\}_{k\geq 0} \subset \{\xi_n\}_{n\geq 0}$ convergente $\mathbb{P}$-q.c. a ξ_∞. Quindi, $\xi_\infty = \xi$. ■

Martingale $\mathbf{L}^1$

Definizione 193 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato e* $\{\xi_n\}_{n\geq 0}$ *una martingala. Se* $\sup_{n\geq 0} \mathbb{E}(|\xi_n|) < \infty$, $\{\xi_n\}_{n\geq 0}$ *è detta* L^1-martingala

Lemma 194 (Doob's upcrossing Lemma) *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato e* $\{\xi_n\}_{n\geq 0}$ *una submartingala. Dati* $a, b \in \mathbb{R}$ *tale che* $a < b$, *sia* $\{\tau_n\}_{n\geq 0}$ *una successione crescente di tempi d'arresto tali che:*

$$\tau_0 := 0,\ \tau_n := \sum_{m\geq 1} \left(\delta_{n,2m-1} \min\{n > \tau_{2m-2} : \xi_n \leq a\} + \delta_{n,2m} \min\{n > \tau_{2m-1} : \xi_n \geq b\}\right). \quad (3.119)$$

Allora, posto, $\forall n \geq 1$,

$$U_n(a,b) := \mathbf{1}_{[0,n]}(\tau_2) \max\{m \geq 1 : \tau_{2m} \leq n\}, \quad (3.120)$$

$$\mathbb{E}(U_n(a,b)) \leq \frac{\mathbb{E}((\xi_n - a) \vee 0 - (\xi_0 - a) \vee 0)}{b-a}. \quad (3.121)$$

Dimostrazione: Poiché $\{\xi_n\}_{n\geq 0}$ è una submartingala e $\mathbb{R} \ni x \longmapsto f_a(x) := (x-a) \vee 0$ è una funzione convessa, dall'Osservazione 161, $\{\eta_n\}_{n\geq 0}$ tale che, $\forall n \geq 1$, $\eta_n := f_a(\xi_n)$ è una submartingala non negativa. Dunque, è sufficiente restringersi a cansiderare il caso in cui $\{\xi_n\}_{n\geq 0}$ sia una submartingala non negativa e dimostrare

$$\mathbb{E}(U_n(0,b)) \leq \frac{\mathbb{E}(\xi_n - \xi_0)}{b}. \quad (3.122)$$

Allora,

$$bU_n(0,b) \leq \sum_{k=1}^{n} \left(\xi_k - \xi_{k-1}\right) \sum_{m\geq 1} \mathbf{1}_{(\tau_{2m-1},\tau_{2m}]}(k)\ . \quad (3.123)$$

Poiché se τ è un tempo d'arresto, $\mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)<k\}} = \sum_{l=0}^{k-1} \mathbf{1}_{\{\omega\in\Omega\,:\,\tau(\omega)=l\}}$ è $\mathcal{F}_{k-1}$-mis., siccome

$$\mathbf{1}_{(\tau_{2m-1},\tau_{2m}]}(k) = \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_{2m-1}(\omega)<k\}} \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_{2m}(\omega)\geq k\}} = \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_{2m-1}(\omega)<k\}} \left(1 - \mathbf{1}_{\{\omega\in\Omega\,:\,\tau_{2m}(\omega)<k\}}\right), \quad (3.124)$$

$\sum_{m\geq 1} \mathbf{1}_{(\tau_{2m-1},\tau_{2m}]}(k)$ è $\mathcal{F}_{k-1}$-mis.. Quindi,

$$
\begin{aligned}
b\mathbb{E}\left(U_n\left(0,b\right)\right) &\leq \sum_{k=1}^{n} \mathbb{E}\left(\left(\xi_k - \xi_{k-1}\right)\sum_{m\geq 1} \mathbf{1}_{(\tau_{2m-1},\tau_{2m}]}(k)\right) = \\
&= \sum_{k=1}^{n} \mathbb{E}\left(\mathbb{E}\left(\left(\xi_k - \xi_{k-1}\right)\sum_{m\geq 1} \mathbf{1}_{(\tau_{2m-1},\tau_{2m}]}(k)\,|\mathcal{F}_{k-1}\right)\right) = \\
&= \sum_{k=1}^{n} \mathbb{E}\left(\sum_{m\geq 1} \mathbf{1}_{(\tau_{2m-1},\tau_{2m}]}(k)\,\mathbb{E}\left(\left(\xi_k - \xi_{k-1}\right)|\mathcal{F}_{k-1}\right)\right) = \\
&= \sum_{k=1}^{n} \mathbb{E}\left(\sum_{m\geq 1} \mathbf{1}_{(\tau_{2m-1},\tau_{2m}]}(k)\left(\mathbb{E}\left(\xi_k|\mathcal{F}_{k-1}\right) - \xi_{k-1}\right)\right) \leq \\
&\leq \sum_{k=1}^{n} \mathbb{E}\left(\left(\mathbb{E}\left(\xi_k|\mathcal{F}_{k-1}\right) - \xi_{k-1}\right)\right) = \sum_{k=1}^{n}\left(\mathbb{E}\left(\xi_k\right) - \mathbb{E}\left(\xi_{k-1}\right)\right) \\
&= \mathbb{E}\left(\xi_n\right) - \mathbb{E}\left(\xi_0\right) \ .
\end{aligned}
\tag{3.125}
$$

■

Teorema 195 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato e* $\{\xi_n\}_{n\geq 0}$ *una submartingala tale che* $\sup_{n\geq 0}\mathbb{E}\left(|\xi_n|\right) < \infty$. *Allora,* $\{\xi_n\}_{n\geq 0}$ *converge* $\mathbb{P}$*-q.c. ad una v.a.* $\xi \in L^1\left(\mathcal{F}\right)$.

Dimostrazione: Supponendo che $\mathbb{P}\left\{\overline{\lim}_n \xi_n > \underline{\lim}_n \xi_n\right\} > 0$, allora, poiché

$$
\left\{\overline{\lim}_n \xi_n > \underline{\lim}_n \xi_n\right\} = \bigcup_{a,b\in\mathbb{Q}\,:\,b>a} \left\{\overline{\lim}_n \xi_n > b > a > \underline{\lim}_n \xi_n\right\}, \tag{3.126}
$$

esiste $(a,b) \in \{(a',b') \in \mathbb{Q}\times\mathbb{Q} : b' > a'\}$ tale che $\mathbb{P}\left\{\overline{\lim}_n \xi_n > b > a > \underline{\lim}_n \xi_n\right\} > 0$. Sia $U_n\left(a,b\right)$ la quantità definita nella (3.120) ovvero, il numero di volte che $\{\xi_i\}_{i=1}^n$ scavalca l'intervallo (a,b). Per il lemma precedente,

$$
\begin{aligned}
\mathbb{E}\left(U_n\left(a,b\right)\right) &\leq \frac{\mathbb{E}\left(\left(\xi_n - a\right)\vee 0 - \left(\xi_0 - a\right)\vee 0\right)}{b-a} \leq \frac{\mathbb{E}\left(|\xi_n| + |\xi_0|\right) + 2\,|a|}{b-a} \\
&\leq 2\frac{\sup_{n\geq 0}\mathbb{E}\left(|\xi_n|\right) + 2\,|a|}{b-a} \ .
\end{aligned}
\tag{3.127}
$$

Quindi, $\{U_n\left(a,b\right)\}_{n\geq 0}$ è una successione crescente di v.a. i cui termini sono uniformemente limitati da una costante, pertanto, per il Teorema di Beppo Levi, $\{U_n\left(a,b\right)\}_{n\geq 0}$ converge in L^1 contro l'ipotesi. Allora, $\{\xi_n\}_{n\geq 0}$ converge $\mathbb{P}$-q.c. ad una v.a. ξ e, per il Lemma di Fatou,

$$
\mathbb{E}\left(|\xi|\right) = \mathbb{E}\left(\left|\overline{\lim}_n \xi_n\right|\right) = \mathbb{E}\left(\overline{\lim}_n |\xi_n|\right) \leq \overline{\lim}_n \mathbb{E}\left(|\xi_n|\right) \leq \sup_{n\geq 0}\mathbb{E}\left(|\xi_n|\right) < \infty \, . \tag{3.128}
$$

■

Corollario 196 *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato e* $\{\xi_n\}_{n\geq 0}$ *una maringala non negativa. Allora,* $\{\xi_n\}_{n\geq 0}$ *converge* $\mathbb{P}$*-q.c..*

Dimostrazione: Poiché,

$$\sup_{n\geq 0} \mathbb{E}\left(|\xi_n|\right) = \sup_{n\geq 0} \mathbb{E}\left(\xi_n\right) = \mathbb{E}\left(\xi_0\right) < \infty\,, \tag{3.129}$$

la tesi segue dal teorema precedente. ■

Osservazione 197 *Sia* $\{\xi_n\}_{n\geq 1}$ *una successione di v.a.i.i.d. di Bernoulli definite su* $(\Omega, \mathcal{F}, \mathbb{P})$ *a valori in* $\{0, 2\}$ *ed equidistribuite. Allora, la successione di v.a.* $\{\eta_n\}_{n\geq 0}$ *tale che,* $\eta_0 = 1$ *e* $\forall n \geq 1$, $\eta_n := \prod_{i=1}^{n} \xi_i$, *è una martingala rispetto alla filtrazione* $\{\mathcal{F}_n\}_{n\geq 0}$ *tale che,* $\mathcal{F}_0 := \{\emptyset, \Omega\}$ *e* $\forall n \geq 1$, $\mathcal{F}_n := \mathcal{F}_n^{\xi}$. $\{\eta_n\}_{n\geq 1}$ *è non negativa e* $\mathbb{E}(\eta_1) = \mathbb{E}(\xi_1) = 1$. *Quindi, per il corollario precedente,* $\{\eta_n\}_{n\geq 1}$ *converge* $\mathbb{P}$*-q.c. a* 0. *Pertanto, per le* L^1*-martingale è in generale impossibile ricostruirne il valore a partire dal limite, come è invece possibile nel caso delle martingale regolari.*

Teorema 198 (Criterio di regolarità per L^1-martingale) *Sia* $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 0}, \mathbb{P})$ *uno spazio filtrato e* $\{\xi_n\}_{n\geq 0}$ *una* L^1*-martingala. Allora,* $\{\xi_n\}_{n\geq 0}$ *è regolare se e solo se è uniformemente integrabile.*

Dimostrazione:

$\Rightarrow$ Poiché $\{\xi_n\}_{n\geq 0}$ è regolare, sia ξ il suo valore limite. Allora, dato $N > 0$, sia $\forall n \geq 0$, $A_n^N := \{\omega \in \Omega : |\xi_n(\omega)| > N\} \in \mathcal{F}_n$. Quindi,

$$\begin{aligned}\mathbb{E}\left(|\xi_n| \mathbf{1}_{A_n^N}\right) &= \mathbb{E}\left(|\mathbb{E}(\xi|\mathcal{F}_n)| \mathbf{1}_{A_n^N}\right) \leq \mathbb{E}\left(\mathbb{E}(|\xi| |\mathcal{F}_n) \mathbf{1}_{A_n^N}\right) = \mathbb{E}\left(\mathbb{E}\left(|\xi| \mathbf{1}_{A_n^N}|\mathcal{F}_n\right)\right) \\ &\leq \mathbb{E}\left(|\xi| \mathbf{1}_{A_n^N}\right) = \mathbb{E}\left(|\xi| \mathbf{1}_{A_n^N}\mathbf{1}_{(\sqrt{N},+\infty)}(|\xi|)\right) + \mathbb{E}\left(|\xi| \mathbf{1}_{A_n^N}\mathbf{1}_{[0,\sqrt{N}]}(|\xi|)\right) \\ &\leq \mathbb{E}\left(|\xi| \mathbf{1}_{(\sqrt{N},+\infty)}(|\xi|)\right) + \sqrt{N}\mathbb{P}\left\{\omega \in \Omega : |\xi_n(\omega)| > N\right\}.\end{aligned} \tag{3.130}$$

Per la disuguaglianza di Chebichev,

$$\begin{aligned}\mathbb{E}\left(|\xi_n| \mathbf{1}_{A_n^N}\right) &\leq \mathbb{E}\left(|\xi| \mathbf{1}_{(\sqrt{N},+\infty)}(|\xi|)\right) + \sqrt{N}\frac{\mathbb{E}(|\xi_n|)}{N} \\ &= \mathbb{E}\left(|\xi| \mathbf{1}_{(\sqrt{N},+\infty)}(|\xi|)\right) + \frac{\mathbb{E}(|\mathbb{E}(\xi|\mathcal{F}_n)|)}{\sqrt{N}} \\ &\leq \mathbb{E}\left(|\xi| \mathbf{1}_{(\sqrt{N},+\infty)}(|\xi|)\right) + \frac{\mathbb{E}(|\xi|)}{\sqrt{N}}\end{aligned} \tag{3.131}$$

Poiché $\xi \in L^1(\mathcal{F})$, passando al limite per $N \to \infty$ si ha che $\forall n \geq 0$, $\lim_{N\uparrow\infty} \mathbb{E}\left(|\xi_n| \mathbf{1}_{A_n^N}\right) = 0$, da cui segue che la successione di v.a. $\{\xi_n\}_{n\geq 0}$ è uniformemente integrabile.

$\Leftarrow$ Poiché $\{\xi_n\}_{n\geq 0}$ è una successione di v.a. uniformemente integrabile, allora $\sup_{n\geq 0} \mathbb{E}(|\xi_n|) < \infty$. Quindi per il Teorema (195) $\{\xi_n\}_{n\geq 0}$ converge $\mathbb{P}$-q.c. ad una v.a. $\xi \in L^1(\mathcal{F})$. Inoltre, per il teorema della convergenza dominata generalizzato $\{\xi_n\}_{n\geq 0}$ converge a ξ in L^1. Quindi, poiché $\forall m \geq n$, $\mathbb{E}(\xi_m | \mathcal{F}_n) = \xi_n$, e poiché $\forall n \geq 0$, $\mathbb{E}(\cdot | \mathcal{F}_n)$ è un operatore lineare limitato, dunque continuo, si ha

$$\lim_{m\to\infty} \mathbb{E}(\xi_m | \mathcal{F}_n) = \mathbb{E}(\xi | \mathcal{F}_n) = \xi_n \quad \mathbb{P}\text{-q.c.}, \tag{3.132}$$

ovvero che $\{\xi_n\}_{n\geq 0}$ è regolare.

■

Esempio 8 *(Derivata di Radon-Nikodým)* Sia $f : [0,1) \longrightarrow \mathbb{R}$ una funzione lipschitziana, ovvero tale che $\exists L > 0$ per cui, $\forall x, y \in [0,1]$,

$$|f(x) - f(y)| \leq L |x - y| \ . \tag{3.133}$$

Allora f è assolutamente continua su $[0,1)$ e pertanto esiste una funzione $g : [0,1) \longrightarrow \mathbb{R}$ integrabile secondo Lebesgue tale che, $\forall x \in [0,1)$,

$$f(x) - f(0) = \int_0^x dy g(y) \ . \tag{3.134}$$

Infatti, considerando lo spazio di probabilità $([0,1), \mathcal{B}([0,1)), dx)$, sia $\{\xi_n\}_{n\geq 0}$ la successione di v.a. tale che $\forall n \geq 0$,

$$\xi_n(x) := \sum_{k=1}^{2^n} \frac{k-1}{2^n} \mathbf{1}_{[\frac{k-1}{2^n}, \frac{k}{2^n})}(x) \tag{3.135}$$

e $\{\mathcal{F}_n^\xi\}_{n\geq 0}$ la filtrazione generata dalla famiglia di v.a. $\{\xi_n\}_{n\geq 0}$, per cui, dalla definizione di ξ_n, risulta $\mathcal{F}_n^\xi = \mathcal{F}^{\xi_n}$. Sia allora $\{\eta_n\}_{n\geq 0}$ la successione di v.a. tali che $\forall n \geq 0$,

$$\eta_n := \frac{f(\xi_n + 2^{-n}) - f(\xi_n)}{2^{-n}} \ . \tag{3.136}$$

$\{\eta_n\}_{n\geq 0}$ è una martingala su $\left([0,1), \mathcal{B}([0,1)), \{\mathcal{F}_n^\xi\}_{n\geq 0}, dx\right)$. Infatti, $\forall n \geq 0, x \in [0,1)$, poiché $\xi_{n+1}(x) \in \{\xi_n(x), \xi_n(x) + 2^{-(n+1)}\}$ e

$$\begin{aligned} &\mathbb{P}\{x \in [0,1) : \xi_{n+1}(x) = \xi_n(x)\} = \qquad (3.137)\\ &\mathbb{P}\{x \in [0,1) : \xi_{n+1}(x) = \xi_n(x) + 2^{-(n+1)}\} = \frac{1}{2}, \end{aligned}$$

$$\begin{aligned} \mathbb{E}\left(\eta_{n+1} | \mathcal{F}_n^\xi\right) &= \mathbb{E}\left(\frac{f\left(\xi_{n+1} + 2^{-(n+1)}\right) - f\left(\xi_{n+1}\right)}{2^{-(n+1)}} | \mathcal{F}_n^\xi\right) \qquad (3.138)\\ &= \frac{1}{2}\frac{f\left(\xi_n + 2^{-(n+1)}\right) - f(\xi_n)}{2^{-(n+1)}} + \frac{1}{2}\frac{f\left(\xi_n + 2^{-(n+1)} + 2^{-(n+1)}\right) - f\left(\xi_n + 2^{-(n+1)}\right)}{2^{-(n+1)}} \\ &= \frac{f(\xi_n + 2^{-n}) - f(\xi_n)}{2^{-n}} = \eta_n \ . \end{aligned}$$

Inoltre, poiché f è lipschitziana, $\{\eta_n\}_{n\geq 0}$ è uniformemente integrabile, in quanto

$$\lim_{N\rightarrow\infty}\sup_{n\geq 0}\mathbb{E}\left(|\eta_n|\,\mathbf{1}_{\{\omega\in\Omega\,:\,|\eta_n(\omega)|>N\}}\right)=\lim_{N\rightarrow\infty}\sup_{n\geq 0}\mathbb{E}\left(\left|\frac{f\left(\xi_n+2^{-n}\right)-f\left(\xi_n\right)}{2^{-n}}\right|\mathbf{1}_{\{\omega\in\Omega\,:\,|\eta_n(\omega)|>N\}}\right) \tag{3.139}$$

$$\leq L\lim_{N\rightarrow\infty}\sup_{n\geq 0}\mathbb{P}\{\omega\in\Omega:|\eta_n(\omega)|>N\}\leq L\lim_{N\rightarrow\infty}\sup_{n\geq 0}\frac{\mathbb{E}\left(|\eta_n(\omega)|\right)}{N}$$

$$=L\lim_{N\rightarrow\infty}\sup_{n\geq 0}\frac{\mathbb{E}\left(\left|\frac{f\left(\xi_n+2^{-n}\right)-f(\xi_n)}{2^{-n}}\right|\right)}{N}\leq L^2\lim_{N\rightarrow\infty}\frac{1}{N}=0$$

e quindi è regolare. Allora, esiste una v.a. $g\in L^1\left(\mathcal{B}\left([0,1)\right)\right)$ tale che, $\forall n\geq 0,\eta_n=\mathbb{E}\left(g|\mathcal{F}_n^{\xi}\right)$. Inoltre, poiché $\mathcal{F}_{\infty}^{\xi}=\bigcup_{n\geq 0}\mathcal{F}_n^{\xi}=\mathcal{B}\left([0,1)\right)$, il valore finale di $\{\eta_n\}_{n\geq 0}$ è $\mathbb{E}\left(g|\mathcal{F}_{\infty}^{\xi}\right)=g$. Quindi, dato che, $\forall k,n\geq 0$, se $B=[0,\frac{k}{2^n}]\in\mathcal{F}_n^{\xi}$,

$$\begin{aligned}\int_0^{\frac{k}{2^n}}dxg(x)&=\mathbb{E}\left(g\mathbf{1}_B\right)=\mathbb{E}\left(\eta_n\mathbf{1}_B\right)=\int_0^{\frac{k}{2^n}}dx\eta_n(x)\\&=\int_0^{\frac{k}{2^n}}dx\frac{f\left(\xi_n(x)+2^{-n}\right)-f\left(\xi_n(x)\right)}{2^{-n}}\\&=\int_0^{\frac{k}{2^n}}dx\sum_{l=1}^{2^n}\frac{f\left(\frac{l-1}{2^n}+2^{-n}\right)-f\left(\frac{l-1}{2^n}\right)}{2^{-n}}\mathbf{1}_{[\frac{l-1}{2^n},\frac{l}{2^n})}(x)\\&=\sum_{l=1}^{k}\frac{f\left(\frac{l-1}{2^n}+2^{-n}\right)-f\left(\frac{l-1}{2^n}\right)}{2^{-n}}2^{-n}=f\left(\frac{k}{2^n}\right)-f(0)\;.\end{aligned} \tag{3.140}$$

si ha la (3.134).

Osservazione 199 *Sia $\mathfrak{M}$ la collezione delle martingale sullo spazio filtrato $(\Omega,\mathcal{F},\{\mathcal{F}_n\}_{n\geq 0},\mathbb{P})$. $\mathfrak{M}$ è uno spazio lineare in quanto, se $\boldsymbol{\xi},\boldsymbol{\eta}\in\mathfrak{M}$, allora, $\forall\alpha,\beta\in\mathbb{R}$, anche $\zeta:=\alpha\boldsymbol{\xi}+\beta\boldsymbol{\eta}$ è una martingala. Sia allora, $\forall p\geq 1$, le norma su $\mathfrak{M}$*

$$\mathfrak{M}\ni\boldsymbol{\xi}\longmapsto|||\boldsymbol{\xi}|||_p:=\sup_{n\geq 0}\|\xi_n\|_{L^p}\in\mathbb{R}^+. \tag{3.141}$$

Se $\mathfrak{M}^p$ è la chiusure nella norma $|||\cdot|||_p$ di $\mathfrak{M}$ allora, poiché se $p'>p,|||\cdot|||_p<|||\cdot|||_{p'},\mathfrak{M}^{p'}\subset\mathfrak{M}^p$.

Notiamo inoltre che, se $p\geq 1,\mathfrak{M}^p$ è spazio di Banach, così come tutti i suoi sottospazi chiusi. Infatti, se $\left\{\boldsymbol{\xi}^{(N)}\right\}_{N\geq 1}$ è una successione di Cauchy in $\mathfrak{M}^p$, allora, $\forall\varepsilon>0,\exists N_{\varepsilon}>0$ tale che, uniformemente in $n\geq 0,\forall N,M>N_{\varepsilon}$, $\left\|\xi_n^{(N)}-\xi_n^{(M)}\right\|_{L^p}<\varepsilon$, ovvero, $\forall n\geq 0,\left\{\xi_n^{(N)}\right\}_{N\geq 1}$ è di Cauchy in $L^p(\mathcal{F}_n)$, che è completo, e pertanto convergente ad un elemento $\xi_n\in L^p(\mathcal{F}_n)$. Inoltre, poiché $L^p(\mathcal{F}_n)\subseteq L^1(\mathcal{F}_n),\left\{\xi_{n+1}^{(N)}\right\}_{N\geq 1}$ converge in $L^1(\mathcal{F}_{n+1})$, quindi, siccome $\mathcal{F}_n\subset\mathcal{F}_{n+1}$, per la continuità dell'operatore $\mathbb{E}(\cdot|\mathcal{F}_n):L^1(\mathcal{F}_{n+1})\circlearrowleft$,

$$\xi_n=\lim_{N\rightarrow\infty}\xi_n^{(N)}=\lim_{N\rightarrow\infty}\mathbb{E}\left(\xi_{n+1}^{(N)}\middle|\mathcal{F}_n\right)=\mathbb{E}\left(\lim_{N\rightarrow\infty}\xi_{n+1}^{(N)}\middle|\mathcal{F}_n\right)=\mathbb{E}\left(\xi_{n+1}|\mathcal{F}_n\right)\;, \tag{3.142}$$

ovvero, $\boldsymbol{\xi} := \{\xi_n\}_{n\geq 0}$ *è una martingala rispetto alla filtrazione* $\{\mathcal{F}_n\}_{n\geq 0}$ *e* $\left\{\boldsymbol{\xi}^{(N)}\right\}_{N\geq 1}$ *converge a* $\boldsymbol{\xi}$.

Se $\mathfrak{M}^{reg}$ *è la chiusura in* $\mathfrak{M}^1$ *dello spazio lineare delle martingale uniformemente integrabili, quanto finora esposto può riassumersi così:*

$$\mathfrak{M}^2 \subset \mathfrak{M}^{reg} \subset \mathfrak{M}^1 \ . \tag{3.143}$$

3.2 Martingale indicizzate da un generico insieme totalmente ordinato

Il concetti ed i risultati espressi nel caso delle martingale a tempo discreto possono essere generalizzati al caso in cui si considerino famiglie di subσalgebre di uno spazio di probabilità e/o famiglie di v.a su di esso definite indicizzate da un generico insieme di indici totalmente ordinato $\mathcal{I}$. Tale è il caso se, per esempio, $\mathcal{I} = \mathbb{Z}$, $-\mathbb{N}$, $\mathbb{R}$, $\mathbb{R}^+$ o, dato un insieme Λ, $\mathcal{I} = \{A \in \mathcal{P}_\Lambda : \text{se } B, C \in \mathcal{I} \text{ allora o } C \subseteq B \text{ o } B \subseteq C\}$.

Definizione 200 *Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità e* $\mathcal{I}$ *un insieme totalmente ordinato. Una successione* $\{\mathcal{F}_i\}_{i\in\mathcal{I}}$ *di subσalgebre di* $\mathcal{F}$ *tali che, se* $i, j \in \mathcal{I}$ *e* $i < j$, $\mathcal{F}_i \subseteq \mathcal{F}_j$, *è detta filtrazione di* $(\Omega, \mathcal{F}, \mathbb{P})$.

La σ *algebra* $\mathcal{F}_\infty$ *generata da tutti i termini della successione* $\{\mathcal{F}_i\}_{i\in\mathcal{I}}$ *è detta* σ*algebra limite della filtrazione e, se* $\mathcal{F}_\infty = \mathcal{F}$, *la fitrazione è detta convergere a* $\mathcal{F}$.

Uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ *munito di una filtrazione* $\{\mathcal{F}_i\}_{i\in\mathcal{I}}$ *è detto filtrato ed è così indicato:* $(\Omega, \mathcal{F}, \{\mathcal{F}_i\}_{i\in\mathcal{I}}, \mathbb{P})$.

Definizione 201 *Una successione generalizzata di v.a.* $\{\xi_i\}_{i\in\mathcal{I}}$ *definita sullo spazio filtrato* $(\Omega, \mathcal{F}, \{\mathcal{F}_i\}_{i\in\mathcal{I}}, \mathbb{P})$ *è detta adattata alla filtrazione* $\{\mathcal{F}_i\}_{i\in\mathcal{I}}$ *se* $\forall i \in \mathcal{I}$, $\xi_i^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{F}_i$.

Se inoltre $\forall i \in \mathcal{I}$, $\exists j < i$ *tale che* $\xi_i^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{F}_j$, $\{\xi_i\}_{i\in\mathcal{I}}$ *è detta prevedibile.*

Osservazione 202 *Data una successione generalizzata di v.a.* $\{\xi_i\}_{i\in\mathcal{I}}$ *definita sullo spazio filtrato* $(\Omega, \mathcal{F}, \{\mathcal{F}_i\}_{i\in\mathcal{I}}, \mathbb{P})$ *se* $\forall i \in \mathcal{I}$, $\mathcal{F}_i^\xi$ *è la* σ*algebra generata dalla collezione di* σ*algebre* $\{\xi_j^{-1}(\mathcal{B}(\mathbb{R}))\}_{j\in\mathcal{I}:j<i}$, *allora* $\{\xi_i\}_{i\in\mathcal{I}}$ *è adattata a* $\{\mathcal{F}_i\}_{i\in\mathcal{I}}$ *se e solo se,* $\forall i \in \mathcal{I}$, $\mathcal{F}_i^\xi \subseteq \mathcal{F}_i$.

Definizione 203 *Una successione generalizzata di v.a.* $\{\xi_i\}_{i\in\mathcal{I}}$ *definita sullo spazio filtrato* $(\Omega, \mathcal{F}, \{\mathcal{F}_i\}_{i\in\mathcal{I}}, \mathbb{P})$ *e adattata alla filtrazione* $\{\mathcal{F}_i\}_{i\in\mathcal{I}}$ *è detta martingala, submartingala, supermartingala se* $\forall i \in \mathcal{I}$, $\xi_i \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ *e* $\forall j \in \mathcal{I}$ *tale che* $j > i$, *rispettivamente si ha:*

$$\begin{array}{l} \mathbb{E}\left(\xi_j|\mathcal{F}_n\right) = \xi_i \\ \mathbb{E}\left(\xi_j|\mathcal{F}_n\right) \geq \xi_i \\ \mathbb{E}\left(\xi_j|\mathcal{F}_n\right) \leq \xi_i \end{array} \quad \mathbb{P}\text{-}q.c.. \tag{3.144}$$

3.2.1 Martingale inverse

Definizione 204 *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità. Una successione $\{\mathcal{G}_n\}_{n\geq 0}$ non crescente di subσalgebre di $\mathcal{F}$, ovvero tali che $\mathcal{F} \supseteq .. \supseteq \mathcal{G}_n \supseteq \mathcal{G}_{n+1} \supseteq ...$ Una successione di v.a. $\{\eta_n\}_{n\geq 0}$ definita sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ è detta* martingala inversa, *se $\forall n \geq 0$, si ha:*

$$\begin{array}{l} \eta_n^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{G}_n \\ \eta_n \in L^1(\mathcal{F}) \\ \mathbb{E}(\eta_n | \mathcal{G}_{n+1}) = \eta_{n+1} \quad \mathbb{P}\text{-}q.c.. \end{array} \tag{3.145}$$

Osservazione 205 *Se la successione di v.a. $\{\eta_n\}_{n\geq 0}$ definita su $(\Omega, \mathcal{F}, \mathbb{P})$ è una martingala inversa rispetto alla successione $\{\mathcal{G}_n\}_{n\geq 0}$ non crescente di subσalgebre di $\mathcal{F}$, ponendo $\forall n \geq 0$, $\xi_n = \eta_{-n}$ e $\mathcal{F}_n = \mathcal{G}_{-n}$, $\{\xi_n\}_{n\in\mathbb{Z}_-}$, è una martingala su $\left(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\in\mathbb{Z}_-}, \mathbb{P}\right)$.*

Teorema 206 *Se la successione di v.a. $\{\eta_n\}_{n\geq 0}$ definita su $(\Omega, \mathcal{F}, \mathbb{P})$ è una martingala inversa rispetto alla successione $\{\mathcal{G}_n\}_{n\geq 0}$ non crescente di subσalgebre di $\mathcal{F}$, allora $\{\eta_n\}_{n\geq 0}$ converge $\mathbb{P}$-q.c. ed in $L^1(\mathcal{F})$ ad una v.a. η tale che, $\forall n \geq 0$, $\mathbb{E}(\eta_n) = \mathbb{E}(\eta)$.*

Dimostrazione: Notiamo che per la martingala $\{\xi_n\}_{n\in\mathbb{Z}_-}$ su $\left(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\in\mathbb{Z}_-}, \mathbb{P}\right)$ la disuguaglianza massimale di Doob si traduce

$$\mathbb{P}\{\omega \in \Omega : \max_{n\leq k\leq 0} |\xi_k| > c\} \leq \frac{4}{c}\mathbb{E}(|\xi_0|) \quad c > 0 \tag{3.146}$$

Inoltre, $\forall a, b \in \mathbb{R}$ tali che $a < b$, dato $n \in \mathbb{Z}_-$, sia $U_n(a, b)$ il numero di coppie di indici $k, l \in \mathbb{Z}_-$ tali che $n \leq k < l \leq 0,$per cui $\xi_l \leq a$; $\xi_m \in (a, b)$, $\forall l < m < k$; $\xi_k \geq b$. Allora, il Doob's upcrossing Lemma si traduce

$$\mathbb{E}(U_n(a, b)) \leq \frac{\mathbb{E}((\xi_0 - a) \vee 0) - \mathbb{E}((\xi_n - a) \vee 0)}{b - a} \leq \frac{\mathbb{E}((\xi_0 - a) \vee 0)}{b - a}. \tag{3.147}$$

Queste disuguaglianze sono indipendenti dall'indice n, perciò, $\forall c > 0$,

$$\mathbb{P}\{\omega \in \Omega : \sup_{k\leq 0} |\xi_k| > c\} \leq \frac{4}{c}\mathbb{E}(|\xi_0|) \tag{3.148}$$

e, poiché $U_n(a, b)$ è crescente, per il Teorema di Beppo Levi,

$$\lim_{n\to\infty} \mathbb{E}(U_n(a, b)) = \mathbb{E}\left(\lim_{n\to\infty} U_n(a, b)\right) \leq \frac{\mathbb{E}((\xi_0 - a) \vee 0)}{b - a}. \tag{3.149}$$

Quindi per il Teorema 195 $\{\xi_n\}_{n\in-\mathbb{N}}$, e dunque $\{\eta_n\}_{n\geq 0}$, converge $\mathbb{P}$-q.c. ad una v.a. $\eta \in L^1(\mathcal{F})$.

Inoltre, $\{\xi_n\}_{n\in\mathbb{Z}_-}$ è uniformemente integrabile, quindi regolare e dunque convergente in $L^1(\mathcal{F})$ a η. Infatti, $\{|\xi_n|\}_{n\in\mathbb{Z}_-}$ è una submartingala, quindi $\forall m, n \in \mathbb{Z}_-$ tali che $m \geq n$, $\mathbb{E}(|\xi_n| | \mathcal{F}_m) \geq \xi_m$ $\mathbb{P}$-q.c. ovvero, $\forall c > 0$,

$$\begin{aligned} \sup_{n\leq 0} \mathbb{E}\left(|\xi_n| \mathbf{1}_{\{|\xi_n|>c\}}\right) &\leq \sup_{n\leq 0} \mathbb{E}\left(\mathbb{E}\left(\left|\xi_{n-1}\right| |\mathcal{F}_n\right) \mathbf{1}_{\{|\xi_n|>c\}}\right) \\ &= \sup_{n\leq 0} \mathbb{E}\left(\mathbb{E}\left(\left|\xi_{n-1}\right| \mathbf{1}_{\{|\xi_n|>c\}} |\mathcal{F}_n\right)\right) = \sup_{n\leq 0} \mathbb{E}\left(\left|\xi_{n-1}\right| \mathbf{1}_{\{|\xi_n|>c\}}\right) \end{aligned} \tag{3.150}$$

Iterando n volte, poiché $\forall n \in \mathbb{Z}_{-}$, $\mathcal{F}_n \subseteq \mathcal{F}_{n-1}$

$$\sup_{n\leq 0} \mathbb{E}\left(|\xi_n| \mathbf{1}_{\{|\xi_n|>c\}}\right) \leq \sup_{n\leq 0} \mathbb{E}\left(|\xi_0| \mathbf{1}_{\{|\xi_n|>c\}}\right) \leq \mathbb{E}\left(|\xi_0| \mathbf{1}_{\{\sup_{n\leq 0}|\xi_n|>c\}}\right). \tag{3.151}$$

Ma, $\xi_0 \in L^1(\mathcal{F})$, quindi, per la (3.148), $\lim_{c\to\infty} \sup_{n\leq 0} \mathbb{E}\left(|\xi_n| \mathbf{1}_{\{|\xi_n|>c\}}\right) = 0$. Allora, si ha che $\{\eta_n\}_{n\geq 0}$ converge in $L^1(\mathcal{F})$ a η tale che, $\forall n \geq 0, \mathbb{E}(\eta) = \mathbb{E}(\eta_n)$, poiché $\forall n \geq 0$, $\mathbb{E}(\eta_n|\mathcal{G}_{n+1}) = \eta_{n+1}$ $\mathbb{P}$-q.c.. ■

Esempio 9 Sia definita su $(\Omega, \mathcal{F}, \mathbb{P})$ una successione $\{\xi_n\}_{n\geq 1}$ di v.a. $L^1(\mathcal{F})$ scambiabili, ovvero tali che, $\forall n \geq 1$, se σ è una permutazione dell'insieme di indici $\{1, .., n\}$, $(\xi_i)_{i=1}^n \stackrel{d}{=} \left(\xi_{\sigma(i)}\right)_{i=1}^n$. Sia allora $\{S_n\}_{n\geq 0}$, la successione di v.a. tali che $S_0 = 0$ e, $\forall n \geq 1, S_n := \sum_{k=1}^n \xi_k$. Data $\{\mathcal{G}_n\}_{n\geq 0}$ la successione di subσalgebre di $\mathcal{F}$ tale che $\forall n \geq 0$, $\mathcal{G}_n$ è la σalgebra generata dalla successione di v.a. $\{S_k\}_{k\geq n}$, la successione di v.a. $\{\eta_n\}_{n\geq 0}$ tali che $\eta_0 = 0$ e $\forall n \geq 1$, $\eta_n := \frac{S_n}{n}$, è una martingala inversa rispetto a $\{\mathcal{G}_n\}_{n\geq 0}$. Infatti, $\{\mathcal{G}_n\}_{n\geq 0}$ è non crescente e $\forall n \geq 0$, $\eta_n^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{G}_n$. Inoltre, $\mathbb{P}$-q.c.,

$$\mathbb{E}\left(\sum_{k=1}^{n+1} \xi_k \middle| \mathcal{G}_{n+1}\right) = \sum_{k=1}^{n+1} \mathbb{E}(\xi_1|\mathcal{G}_{n+1}) = \sum_{k=1}^{n+1} \mathbb{E}\left(\frac{k}{k}\xi_1 \middle| \mathcal{G}_{n+1}\right) = \sum_{k=1}^{n+1} \frac{1}{k}\mathbb{E}\left(\sum_{i=1}^{k} \xi_i \middle| \mathcal{G}_{n+1}\right). \tag{3.152}$$

Perciò,

$$\eta_{n+1} = \mathbb{E}\left(\eta_{n+1}|\mathcal{G}_{n+1}\right) = \frac{\mathbb{E}(S_{n+1}|\mathcal{G}_{n+1})}{n+1} = \frac{\sum_{k=1}^{n+1} \mathbb{E}(\eta_k|\mathcal{G}_{n+1})}{n+1} = \frac{(n+1)\mathbb{E}(\eta_k|\mathcal{G}_{n+1})}{n+1} \tag{3.153}$$
$$= \mathbb{E}(\eta_k|\mathcal{G}_{n+1}) \quad \forall 1 \leq k \leq n+1$$

ovvero, $\eta_{n+1} = \mathbb{E}(\eta_n|\mathcal{G}_{n+1})$ $\mathbb{P}$-q.c.. Allora, per il teorema precedente, esiste una v.a. η tale che $\{\eta_n\}_{n\geq 0}$ converge a η sia $\mathbb{P}$-q.c. che in $L^1(\mathcal{F})$, per cui, $\forall n \geq 0$, $\mathbb{E}(\eta_n) = \mathbb{E}(\eta)$.

Legge forte dei grandi numeri per v.a.i.i.d.

Lemma 207 *Sia ξ una v.a. non negativa definita su $(\Omega, \mathcal{F}, \mathbb{P})$. Allora*

$$\sum_{n\geq 1} \mathbb{P}\{\omega \in \Omega : |\xi(\omega)| \geq n\} \leq \mathbb{E}(\xi) \leq 1 + \sum_{n\geq 1} \mathbb{P}\{\omega \in \Omega : |\xi(\omega)| \geq n\} . \tag{3.154}$$

Dimostrazione: Poiché

$$\mathbb{E}(\xi) = \mathbb{E}\left[\sum_{k\geq 0} \xi \mathbf{1}_{[k,k+1)}(\xi)\right] = \sum_{k\geq 0} \mathbb{E}\left[\xi \mathbf{1}_{[k,k+1)}(\xi)\right] , \tag{3.155}$$

si ha

$$\mathbb{E}(\xi) \geq \sum_{k\geq 0} k\mathbb{P}\{\omega \in \Omega : \xi(\omega) \in [k, k+1)\} \tag{3.156}$$
$$= \sum_{n\geq 1} \sum_{k\geq n} \mathbb{P}\{\omega \in \Omega : \xi(\omega) \in [k, k+1)\}$$
$$= \sum_{n\geq 1} \mathbb{P}\{\omega \in \Omega : |\xi(\omega)| \geq n\} .$$

D'altronde

$$\begin{aligned}\mathbb{E}(\xi) &\leq \sum_{k\geq 0}(k+1)\,\mathbb{P}\{\omega\in\Omega:\xi(\omega)\in[k,k+1)\} \\ &= 1+\sum_{k\geq 0}k\mathbb{P}\{\omega\in\Omega:\xi(\omega)\in[k,k+1)\} \\ &= 1+\sum_{n\geq 1}\mathbb{P}\{\omega\in\Omega:|\xi(\omega)|\geq n\}\ .\end{aligned} \tag{3.157}$$

■

Teorema 208 *Sia definita su* $(\Omega,\mathcal{F},\mathbb{P})$ *una successione* $\{\xi_n\}_{n\geq 1}$ *di v.a.i.i.d.. Se* $\xi_1\in L^1(\mathcal{F})$ *allora la successione di v.a.* $\{\eta_n\}_{n\geq 0}$ *tali che* $\eta_0=0$ *e* $\forall n\geq 1$, $\eta_n:=\frac{S_n}{n}$, *dove* $S_n:=\sum_{k=1}^n\xi_k$, *converge a* $\mathbb{E}(\xi_1)$ $\mathbb{P}$*-q.c.. Viceversa, se* $\{\eta_n\}_{n\geq 0}$ *converge* $\mathbb{P}$*-q.c. ad una v.a. degenere* $\mu<\infty$, *allora* $\mathbb{E}(\xi_1)=\mu<\infty$.

Dimostrazione: Poiché gli elementi di $\{\xi_n\}_{n\geq 1}$ sono v.a. identicamente distribuite sono scambiabili. Inoltre se $\xi_1\in L^1(\mathcal{F})$ dall'esempio precedente si ha che $\forall m\geq 1$,

$$\eta=\lim_{n\to\infty}\eta_n=\lim_{n\to\infty}\frac{\sum_{k=m}^{n+m}\xi_k}{n}\quad \mathbb{P}\text{-q.c. .} \tag{3.158}$$

Perciò se gli elementi di $\{\xi_n\}_{n\geq 1}$ sono v.a. indipendenti, $\forall m\geq 1$, η e η_m sono indipendenti e $\forall c\in\mathbb{R}$,

$$\mathbb{P}\{\omega\in\Omega:\eta(\omega)>c,\,\eta_m(\omega)>c\}=\mathbb{P}\{\omega\in\Omega:\eta(\omega)>c\}\,\mathbb{P}\{\omega\in\Omega:\eta_m(\omega)>c\}\ . \tag{3.159}$$

Per cui,

$$\mathbb{P}\bigcup_{k=n}^{m}\{\omega\in\Omega:\eta(\omega)>c,\,\eta_k(\omega)>c\}=\mathbb{P}\{\omega\in\Omega:\eta(\omega)>c\}\,\mathbb{P}\bigcup_{k=n}^{m}\{\omega\in\Omega:\eta_k(\omega)>c\}\ , \tag{3.160}$$

ovvero

$$\begin{aligned}\mathbb{P}\left\{\omega\in\Omega:\eta(\omega)>c,\ \max_{n\leq k\leq m}\eta_k(\omega)>c\right\} &= \mathbb{P}\{\omega\in\Omega:\eta(\omega)>c\}\times \\ &\times\mathbb{P}\left\{\omega\in\Omega:\max_{n\leq k\leq m}\eta_k(\omega)>c\right\}\ ,\end{aligned} \tag{3.161}$$

da cui segue che, passando al limite per $m\to\infty$,

$$\begin{aligned}\mathbb{P}\left\{\omega\in\Omega:\eta(\omega)>c,\,\sup_{k\geq n}\eta_k(\omega)>c\right\} &= \mathbb{P}\{\omega\in\Omega:\eta(\omega)>c\}\times \\ &\times\mathbb{P}\left\{\omega\in\Omega:\sup_{k\geq n}\eta_k(\omega)>c\right\}\ ,\end{aligned} \tag{3.162}$$

cioè

$$\mathbb{P}\left\{\omega\in\Omega:\eta(\omega)>c,\ \overline{\lim}_n\eta_n(\omega)>c\right\}=\mathbb{P}\{\omega\in\Omega:\eta(\omega)>c\}\times \times\,\mathbb{P}\left\{\omega\in\Omega:\overline{\lim}_n\eta_n(\omega)>c\right\}\ . \tag{3.163}$$

Ma, poichè $\{\eta_n\}_{n\geq 0}$ converge a η, $\overline{\lim}_n\eta_n=\eta$, dunque

$$\mathbb{P}\{\omega\in\Omega:\eta(\omega)>c\}=\left(\mathbb{P}\{\omega\in\Omega:\eta(\omega)>c\}\right)^2\ , \tag{3.164}$$

ovvero, $\forall c\in\mathbb{R}$, $\mathbb{P}\{\omega\in\Omega:\eta(\omega)>c\}$ è uguale a zero o uno, cioè η è una v.a. degenere il cui valore è μ dato che, $\forall n\geq 0$, $\mathbb{E}(\eta)=\mathbb{E}(\eta_n)=\mathbb{E}(\xi_1)=\mu$.

D'altra parte, siccome

$$\eta_{n+1}-\eta_n=\frac{S_{n+1}}{n+1}-\frac{S_n}{n}=\frac{\xi_{n+1}}{n+1}-\frac{S_n}{n}\left(1-\frac{n}{n+1}\right)=\frac{\xi_{n+1}}{n+1}-\eta_n\left(1-\frac{n}{n+1}\right)\ , \tag{3.165}$$

se $\{\eta_n\}_{n\geq 0}$ converge $\mathbb{P}$-q.c. ad una v.a. $\eta\in L^\infty$,

$$\lim_{n\to\infty}\frac{\xi_n}{n}=\lim_{n\to\infty}\eta_n-\eta_{n-1}\left(\frac{n-1}{n}\right)=0\ \mathbb{P}\text{-q.c. .} \tag{3.166}$$

cioè, $\forall\varepsilon>0$,

$$\lim_{n\to\infty}\mathbb{P}\left\{\omega\in\Omega:\sup_{k\geq n}\left|\frac{\xi_k(\omega)}{k}\right|>\varepsilon\right\}=\lim_{n\to\infty}\mathbb{P}\bigcup_{k\geq n}\left\{\omega\in\Omega:\left|\frac{\xi_k(\omega)}{k}\right|>\varepsilon\right\} =\mathbb{P}\{\omega\in\Omega:|\xi_n(\omega)|>n\varepsilon\ i.s.\}=0\ . \tag{3.167}$$

In particolare, per $\varepsilon=1,\mathbb{P}\{\omega\in\Omega:|\xi_n(\omega)|>n\ i.s.\}=0$ quindi, poiché gli elementi di $\{\xi_n\}_{n\geq 1}$ sono v.a. indipendenti, gli eventi della collezione $\{\{\omega\in\Omega:|\xi_n(\omega)|>n\}\}_{n\geq 1}$ sono tra loro mutuamente indipendenti, per il Lemma di Borel-Cantelli, $\sum_{n\geq 1}\mathbb{P}\{\omega\in\Omega:|\xi_n(\omega)|>n\}<\infty$ e per il Lemma precedente $\xi_1\in L^1(\mathcal{F})$. Inoltre, se η è degenere, $\exists\mu\in\mathbb{R}$ tale che $\eta=\mu$ $\mathbb{P}$-q.c., ma per la prima parte del teorema $\lim_{n\to\infty}\eta_n=\mathbb{E}(\xi_1)$ $\mathbb{P}$-q.c., quindi $\mathbb{E}(\xi_1)=\mu$. ■

3.2.2 Martingale a tempo continuo

Sia $\mathcal{I}=\mathbb{R}^+$ e $(\Omega,\mathcal{F},\mathbb{P})$ uno spazio di probabilità. La successione $\{\mathcal{F}_t\}_{t\in\mathbb{R}^+}$ di subσalgebre di $\mathcal{F}$ tali che $\mathcal{F}_s\subseteq\mathcal{F}_t$ se $s\leq t$ è una filtrazione di $(\Omega,\mathcal{F},\mathbb{P})$, pertanto $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in\mathbb{R}^+},\mathbb{P})$ è uno spazio filtrato.

Definizione 209 *Sia* $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in\mathbb{R}^+},\mathbb{P})$ *uno spazio filtrato. Una v.a.* $\Omega\ni\omega\longmapsto\tau(\omega)\in[0,\infty]$ *è detta* tempo di Markov *rispetto alla filtrazione* $\{\mathcal{F}_t\}_{t\in\mathbb{R}^+}$, *se* $\forall t\geq 0$, $\{\omega\in\Omega:\tau(\omega)>t\}\in\mathcal{F}_t$. *Se inoltre,* $\mathbb{P}\{\omega\in\Omega:\tau(\omega)<\infty\}=1$, *allora* τ *è detto* tempo d'arresto.

Osservazione 210 *Analogamente al caso in cui si consideri una filtrazione indicizzata da* $\mathbb{N}$, *risultano essere tempi di Markov le v.a.* $\tau=t\mathbf{1}_{\{\omega\in\Omega:\tau(\omega)=t\}}$ *e* $\tau=\sigma\wedge t$ *dove* $t>0$ *e* σ *tempo di Markov;* $\tau=\tau_1\vee\tau_2,\tau=\tau_1\wedge\tau_2$ *e* $\tau=\tau_1+\tau_2$ *con* τ_1,τ_2 *tempi di Markov.*

Esempio 10 Se $\{\xi_t\}_{t\in\mathbb{R}^+}$ è un processo stocastico a traiettorie continue, ovvero tale che $\Omega \ni \omega \longmapsto \xi_t(\omega) \in C(\mathbb{R}^+)$, e $\{\mathcal{F}_t^\xi\}_{t\in\mathbb{R}^+}$ è la filtrazione generata dal processo $\{\xi_t\}_{t\in\mathbb{R}^+}$, cioè $\forall t \geq 0$, $\mathcal{F}_t^\xi$ è la σalgebra generata da $\{\xi_s\}_{s\in[0,t]}$, $\forall c \in \mathbb{R}$, $\tau_c := \inf\{t \geq 0 : \xi_t \geq c\}$ è un tempo di Markov rispetto a $\{\mathcal{F}_t^\xi\}_{t\in\mathbb{R}^+}$. Infatti, essendo ξ_t continua

$$\begin{aligned}\{\omega \in \Omega : \tau_c(\omega) > t\} &= \bigcup_{k\geq 1} \left\{\omega \in \Omega : \min_{s\in[0,t]} \xi_s(\omega) \leq c - \frac{1}{k}\right\} \\ &= \bigcup_{k\geq 1} \bigcap_{s\in[0,t]\cap\mathbb{Q}} \left\{\omega \in \Omega : \xi_s(\omega) \leq c - \frac{1}{k}\right\} .\end{aligned} \tag{3.168}$$

Ma, se $s \leq t$,

$$\left\{\omega \in \Omega : \xi_s(\omega) \leq c - \frac{1}{k}\right\} = \left\{\omega \in \Omega : \xi_s(\omega) > c - \frac{1}{k}\right\}^c \in \mathcal{F}_t \tag{3.169}$$

quindi dato che $\mathcal{F}_t^\xi$ è una σalgebra è chiusa per unioni ed intersezioni numerabili di eventi.

Osservazione 211 *Se $\{\xi_t\}_{t\in\mathbb{R}^+}$ è un processo stocastico a traiettorie continue e $\{\mathcal{F}_t^\xi\}_{t\in\mathbb{R}^+}$ è la filtrazione da esso generata, il ragionamento fatto nell'esempio precedente vale anche nel caso in cui si consideri la v.a. $\tau_A := \inf\{t \geq 0 : \xi_t \in A\}$, dove è un sottoinsieme chiuso di $\mathbb{R}$. Nel caso in cui le traiettorie di $\{\xi_t\}_{t\in\mathbb{R}^+}$ non sono continue, oppure A non è chiuso τ_A non è in generale un tempo di Markov rispetto a $\{\mathcal{F}_t^\xi\}_{t\in\mathbb{R}^+}$. Tuttavia, se le traiettorie di $\{\xi_t\}_{t\in\mathbb{R}^+}$ sono continue a destra ed hanno limite a sinistra è possibile considerare $\forall t \geq 0$ una σalgebra $\mathcal{F}_t \supseteq \mathcal{F}_t^\xi$ e dunque una filtrazione $\{\mathcal{F}_t\}_{t\in\mathbb{R}^+}$ per cui τ_A risulti un tempo di Markov $\forall A \in \mathcal{B}(\mathbb{R})$. Infatti, sia $\forall t \geq 0$, $\mathcal{F}_{t^+}^\xi := \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}^\xi$. Allora, $\mathcal{F}_{t^+}^\xi \supseteq \mathcal{F}_t^\xi$ e $\xi_t^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{F}_{t^+}^\xi$, inoltre, $\forall s < t$, $\mathcal{F}_{s^+}^\xi \subseteq \mathcal{F}_{t^+}^\xi$. Quindi, conisderando $\forall t \geq 0$, la σalgebra $\mathcal{F}_t$ completamento di $\mathcal{F}_{t^+}^\xi$ rispetto a $\mathbb{P}$, ovvero*

$$\mathcal{F}_t := \overline{\mathcal{F}^\xi_{t^+}}^{\mathbb{P}} = \{A \subseteq \Omega : \exists A_-, A_+ \in \mathcal{F}_{t^+}^\xi \text{ tale che } A_- \subseteq A \subseteq A_+ \text{ e } \mathbb{P}(A_+ \backslash A_-) = 0\} , \tag{3.170}$$

$\forall A \in \mathcal{B}(\mathbb{R})$,

$$\begin{cases} \tau_A := \inf\{t \geq 0 : \xi_t \in A\} & \text{se } \exists t \geq 0 : \xi_t \in A \\ \infty & \text{se } \forall t \geq 0,\ \xi_t \notin A \end{cases} \tag{3.171}$$

è un tempo di Markov rispetto a $\{\mathcal{F}_t\}_{t\in\mathbb{R}^+}$.

Esempio 11 Se $\{\xi_t\}_{t\in\mathbb{R}^+}$ è un moto browniano tale che $\mathbb{E}(\xi_1) = 0$, $\mathbb{E}(\xi_1^2) = \sigma^2$, allora, generalizzando quanto esposto nell'Esempio 2, sono martingale rispetto alla filtrazione generata dal moto browniano:

1. $\{\xi_t\}_{t\in\mathbb{R}^+}$;
2. $\{\eta_t\}_{t\in\mathbb{R}^+}$, tale che $\forall t \geq 0$, $\eta_t = \xi_t^2 - \sigma^2 t$;

3. $\{\zeta_t(\lambda)\}_{t\in\mathbb{R}^+}$, tale che $\forall t \geq 0, \lambda \in \mathbb{R}, \zeta_t(\lambda) = e^{\lambda\xi_t - \frac{\lambda^2}{2}\sigma^2 t}$.

   Più in generale, se $f \in C^2(\mathbb{R})$ strettamente crescente allora, $\{W_t(f,\lambda)\}_{t\in\mathbb{R}^+}$, tale che $\forall t \geq 0, \lambda \in \mathbb{R}$,

$$W_t(f,\lambda) := \exp\left\{\lambda f(\xi_t) - \frac{\sigma^2}{2}\int_0^t ds\left[\lambda^2 (f'(\xi_t))^2 + \lambda f''(\xi_t)\right]\right\} \tag{3.172}$$

   è una martingala rispetto alla filtrazione generata dal moto browniano, se $\forall t \geq 0$, $\lambda \in \mathbb{R}$, $\mathbb{E}(W_t(f,\lambda)) < \infty$.

Esempio 12 Se $\{\xi_t\}_{t\in\mathbb{R}^+}$ è un processo di Poisson di parametro λ,allora, generalizzando quanto esposto nell'Esempio 2, sono martingale rispetto alla filtrazione generata dal processo di Poisson:

1. $\{\eta_t\}_{t\in\mathbb{R}^+}$, tale che $\forall t \geq 0$, $\eta_t = \xi_t - \lambda t$; ;
2. $\{\zeta_t\}_{t\in\mathbb{R}^+}$, tale che $\forall t \geq 0$, $\zeta_t = \eta_t^2 - \lambda t$;
3. $\{v_t(\theta)\}_{t\in\mathbb{R}^+}$, tale che $\forall t \geq 0, \theta \in \mathbb{R}, v_t(\theta) = e^{\theta\xi_t - \lambda t\left(1-e^{\theta}\right)}$.

Capitolo 4

Processi stazionari e teoria ergodica

4.1 Misura sullo spazio delle successioni a valori reali

Sia $\mathbb{R}^{\mathbb{N}}$ l'insieme delle successioni a valori reali $\mathbf{x} := \{x_i\}_{i\in\mathbb{N}}$. $\forall n \in \mathbb{N}$, dato un boreliano $B \in \mathcal{B}(\mathbb{R}^n)$, sia

$$C(B) = \{\mathbf{x} \in \mathbb{R}^{\mathbb{N}} : (x_1, .., x_n) \in B\} \tag{4.1}$$

l'insieme cilindrico (*cilindro*) di base $B \in \mathcal{B}(\mathbb{R}^n)$ e $\mathcal{C}$ la σalgebra generata dagli insiemi cilindrici.

Dato uno spazio misurabile $(\Omega, \mathcal{F})$, indichiamo con $\mathfrak{P}(\Omega, \mathcal{F})$ la collezione delle misure di probabilità su $(\Omega, \mathcal{F})$.

Teorema 212 *(di Kolmogorov sull'estensione di misure su $(\mathbb{R}^{\mathbb{N}}, \mathcal{C})$)*
Sia $\{\mathbb{P}_n\}_{n\geq 1}, \mathbb{P}_n \in \mathfrak{P}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ una collezione di misure di probabilità tali che $\forall n \geq 1,\ B \in \mathcal{B}(\mathbb{R}^n)$,

$$\mathbb{P}_{n+1}(B \times \mathbb{R}) = \mathbb{P}_n(B). \tag{4.2}$$

Allora esiste un unica $\mathbb{P} \in \mathfrak{P}(\mathbb{R}^{\mathbb{N}}, \mathcal{C})$ tale che se $\forall n \geq 1,\ B \in \mathcal{B}(\mathbb{R}^n)$,

$$\mathbb{P}(C(B)) = \mathbb{P}_n(B). \tag{4.3}$$

Dimostrazione: $\forall n \geq 1,\ B \in \mathcal{B}(\mathbb{R}^n)$, assegnamo a $\mathbb{P}(C(B))$ il valore $\mathbb{P}_n(B)$. Allora poiché, $\forall k \geq 1$,

$$\begin{aligned} C(B) &= \{\omega \in \mathbb{R}^{\mathbb{N}} : (\omega_1, .., \omega_n) \in B,\ \omega_{n+1}, .., \omega_{n+k} \in \mathbb{R}\} \qquad (4.4) \\ &= C\left(B \times \underbrace{\mathbb{R} \times ... \times \mathbb{R}}_{k\,volte}\right) \end{aligned}$$

per la (4.2), la definizione di $\mathbb{P}(C(B))$ è indipendente dalla rappresentazione scelta di $C(B)$, infatti

$$\mathbb{P}_n(B) = \mathbb{P}_{n+1}(B \times \mathbb{R}) = ... = \mathbb{P}_{n+k}(B \times \mathbb{R} \times ... \times \mathbb{R}),\ k \geq 1. \tag{4.5}$$

Sia $\mathcal{C}\left(\mathbb{R}^{\mathbb{N}}\right)$ la collezione di tutti i cilindri $C(B)$ al variare di $n \geq 1$ e $B \in \mathcal{B}(\mathbb{R}^n)$. È facile vedere che $\mathcal{C}\left(\mathbb{R}^{\mathbb{N}}\right)$ è un algebra. Dato $k \geq 2$ e $A_1, .., A_k \in \mathcal{C}\left(\mathbb{R}^{\mathbb{N}}\right)$ disgiunti, $\exists n \geq 1$ e $B_1, .., B_k \in \mathbb{R}^n$ boreliani disgiunti tale che $A_i = C(B_i)$, $i = 1, .., k$. Allora,

$$\begin{aligned} \mathbb{P}\left(\bigcup_{i=1}^{k} A_i\right) &= \mathbb{P}\left(\bigcup_{i=1}^{k} C(B_i)\right) = \mathbb{P}\left(C\left(\bigcup_{i=1}^{k} B_i\right)\right) = \sum_{i=1}^{k} \mathbb{P}_n(B_i) \\ &= \sum_{i=1}^{k} \mathbb{P}(C(B_i)) = \sum_{i=1}^{k} \mathbb{P}(A_i), \end{aligned} \tag{4.6}$$

ovvero $\mathbb{P}$ è finitamente additiva sull'algebra $\mathcal{C}\left(\mathbb{R}^{\mathbb{N}}\right)$. Se $\mathbb{P}$ è anche σadditiva su $\mathcal{C}\left(\mathbb{R}^{\mathbb{N}}\right)$ allora la tesi segue dal teorema precedente. A tal fine è sufficiente dimostrare che, per una generica successione $\{C(B_n)\}_{n\geq 1}$ di elementi di $\mathcal{C}\left(\mathbb{R}^{\mathbb{N}}\right)$ tale che $C(B_n) \downarrow \varnothing, \mathbb{P}(C(B_n)) \underset{n\to\infty}{\longrightarrow} 0$. Supponiamo il contrario, ovvero $\lim_{n\to\infty} \mathbb{P}(C(B_n)) = \delta > 0$. Poiché $\forall n \geq 1$, $B_n \in \mathcal{B}(\mathbb{R}^n)$, dato $\delta > 0$ esiste $K_n \subseteq B_n$ compatto tale che $\mathbb{P}_n(B_n \backslash K_n) \leq \frac{\delta}{2^{n+1}}$, allora,

$$\mathbb{P}(C(B_n) \backslash C(K_n)) = \mathbb{P}_n(B_n \backslash K_n) \leq \frac{\delta}{2^{n+1}}. \tag{4.7}$$

Sia $\bar{C}_n := \bigcap_{k=1}^{n} C(K_k)$ e sia D_n tale che $\bar{C}_n = C(D_n)$. Poiché $\{C(B_n)\}_{n\geq 1}$ è decrescente, si ha

$$\begin{aligned} \mathbb{P}(C(B_n) \backslash C(D_n)) &\leq \sum_{k=1}^{n} \mathbb{P}(C(B_n) \backslash C(K_k)) = \sum_{k=1}^{n} \mathbb{P}(C(B_n \backslash K_k)) \\ &\leq \sum_{k=1}^{n} \frac{\delta}{2^{k+1}} = \frac{\delta}{2}, \end{aligned} \tag{4.8}$$

ma per ipotesi $\lim_{n\to\infty} \mathbb{P}(C(B_n)) = \delta > 0$ e quindi $\lim_{n\to\infty} \mathbb{P}(C(D_n)) \geq \frac{\delta}{2} > 0$. Ma ciò contraddice l'ipotesi che $C(D_n) \downarrow \varnothing$. Infatti, $\forall n \geq 1$ sia $x^{(n)} \in C(D_n)$, allora $\left(x_1^{(n)}, .., x_n^{(n)}\right) \in D_n$. Inoltre, poiché D_1 è compatto, esiste una sottosuccessione $\{n_1\}$ di $\{n\}$ tale che $x_1^{(n_1)} \to x_1^0 \in D_1$. Allo stesso modo è possibile scegliere $\{n_2\} \subseteq \{n_1\}$ tale che $\left(x_1^{(n_2)}, x_2^{(n_2)}\right) \to (x_1^0, x_2^0) \in D_2$. Iterando questa costruzione si ha

$$\forall k \geq 1, \left(x_1^{(n_k)}, .., x_k^{(n_k)}\right) \to \left(x_1^0, .., x_k^0\right) \in D_k. \tag{4.9}$$

Considerando la sottosuccessione diagonale $\{m_k\}$ segue che $\forall i \geq 1, x_i^{(m_k)} \to x_i^0$ e $(x_1^0, ..) \in C(D_n)$ $\forall n \geq 1$ contrariamente all'ipotesi che $C(D_n) \downarrow \varnothing$. ■

4.2 Leggi 0-1 per successioni di v.a. indipendenti

Sia $\{\xi_i\}_{i\geq 1}$ una successione di v.a. definita sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e, $\forall n \geq 1$, siano

- $\mathcal{F}_n$ la σalgebra generata dalla collezione di v.a. $\{\xi_i\}_{i=1}^n$, ovvero quella generata dalla collezione di eventi di $\mathcal{F}$:

$$\{\omega \in \Omega : (\xi_1(\omega), .., \xi_n(\omega)) \in B\} \quad B \in \mathcal{B}(\mathbb{R}^n) \; ; \tag{4.10}$$

- $\mathcal{F}^n$ la σalgebra generata dalla collezione di v.a. $\{\xi_i\}_{i\geq 1}$, ovvero quella generata dalla collezione di eventi di $\mathcal{F}$:

$$\{\omega \in \Omega : \big(\xi_n(\omega), \xi_{n+1}(\omega), ..\big) \in B\} \quad B \in \mathcal{C} \, . \tag{4.11}$$

La σalgebra

$$\mathcal{T} := \bigcap_{n\geq 1} \mathcal{F}^n \tag{4.12}$$

è detta *σalgebra di coda.* Ne segue che $\forall n \geq 1$, ogni evento $A \in \mathcal{T}$ è indipendente da $\{\xi_i\}_{i=1}^n$.

Teorema 213 *(Legge 0-1 di Kolmogorov) Sia $\{\xi_i\}_{i\geq 1}$ una successione di v.a. indipendenti definita sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$. $\forall A \in \mathcal{T}$, $\mathbb{P}(A)$ può assumere soltanto i valori zero e uno.*

Dimostrazione: Sia $A \in \mathcal{T}$. Allora, $A \in \mathcal{F}^1$ e pertanto $\exists A_n \in \mathcal{F}_n$ tale che

$$\lim_{n\to\infty} \mathbb{P}(A \triangle A_n) = 0, \tag{4.13}$$

ovvero

$$\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(A_n) \; , \tag{4.14}$$

ma allora

$$\lim_{n\to\infty} \mathbb{P}(A \cap A_n) = \mathbb{P}(A) \; .$$

Poiché per ipotesi $A \in \mathcal{T}$, A e A_n sono indipendenti, $\forall n \geq 1$, $\mathbb{P}(A \cap A_n) = \mathbb{P}(A)\,\mathbb{P}(A_n)$ che tende a $\mathbb{P}^2(A)$ per $n \to \infty$. Dunque, vale $\mathbb{P}(A) = \mathbb{P}^2(A)$. ■

Corollario 214 *Nelle ipotesi del teorema precedente, una v.a. η definita su $(\Omega, \mathcal{F}, \mathbb{P})$ e misurabile rispetto a $\mathcal{T}$ è degenere, ovvero $\exists c \in \mathbb{R} : \eta = c\,\mathbb{P}$-q.c..*

Dimostrazione: Poiché η è $\mathcal{T}$-misurabile, $\forall x \in \mathbb{R}$, $\{\omega \in \Omega : \eta(\omega) \leq x\} \in \mathcal{T}$. Quindi,

$$\mathbb{P}\{\omega \in \Omega : \eta(\omega) \leq x\} = F_\eta(x) \tag{4.15}$$

può assumere soltanto i valori 0 e 1. Sia

$$c := \inf\{x \in \mathbb{R} : F_\eta(x) = 1\} \, . \tag{4.16}$$

Allora, $\forall \varepsilon > 0$,

$$\mathbb{P}\{\omega \in \Omega : c - \varepsilon < \eta(\omega) \leq c + \varepsilon\} = F_\eta(c+\varepsilon) - F_\eta(c-\varepsilon) = 1 \, , \tag{4.17}$$

cioè la tesi. ■

Definizione 215 *Un'applicazione* $\mathbb{N} \ni n \longmapsto \sigma_n \in \mathbb{N}$ *è detta* permutazione finita *se* $\sigma_n = n$ *salvo un numero finito di elementi di* $\mathbb{N}$. *Sia quindi, per ogni permutazione finita* σ,

$$\mathbb{R}^{\mathbb{N}} \ni x \longmapsto \mathbf{T}_\sigma x = \{x_{\sigma_n}\}_{n\geq 1} \in \mathbb{R}^{\mathbb{N}} . \tag{4.18}$$

e $\forall B \in \mathcal{C}$

$$\mathbf{T}_\sigma^{-1}(B) := \left\{x \in \mathbb{R}^{\mathbb{N}} : \mathbf{T}_\sigma x \in B\right\} . \tag{4.19}$$

Data $\xi := \{\xi_i\}_{i\geq 1}$ una successione di v.a. definite su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, $\forall B \in \mathcal{C}$, sia

$$A_B := \{\omega \in \Omega : \xi(\omega) \in B\}, \tag{4.20}$$

e, posto $\mathbf{T}_\sigma(\xi) := \{\xi_{\sigma_i}\}_{i\geq 1}$, sia

$$A_{\mathbf{T}_\sigma^{-1}(B)} := \{\omega \in \Omega : \mathbf{T}_\sigma(\xi)(\omega) \in B\}. \tag{4.21}$$

Definizione 216 *Ogni evento* $A \in \mathcal{F}$ *è detto* simmetrico *se* $\exists B \in \mathcal{C}$ *tale che* $A = A_B$ *e per ogni permutazione finita* $\sigma, A_B = A_{\mathbf{T}_\sigma^{-1}(B)}$.

Teorema 217 *(Legge 0-1 di Hewitt-Savage) Sia* $\xi := \{\xi_i\}_{i\geq 1}$ *una successione di v.a.i.i.d.. Ogni evento simmetrico ha probabilità zero o uno.*

Dimostrazione: Sia A un evento simmetrico e sia quindi $B \in \mathcal{C}$ tale che $A = A_B$. Sia $B_n \in \mathcal{B}(\mathbb{R}^n)$ tale che, se

$$A_n := A_{C(B_n)} = \{\omega \in \Omega : (\xi_1(\omega), .., \xi_n(\omega)) \in B_n\}, \tag{4.22}$$

$\lim_{n\to\infty} \mathbb{P}(A \triangle A_n) = 0$. Poiché le ξ_i sono v.a.i.i.d., per ogni permutazione finita σ,

$$\mathbb{P}(A_n) = \mathbb{P}_\xi(C(B_n)) = \mathbb{P}_{\mathbf{T}_\sigma(\xi)}(C(B_n)) . \tag{4.23}$$

Dunque,

$$\begin{aligned} \mathbb{P}(A \triangle A_n) &= \mathbb{P}(A_B \triangle A_n) = \mathbb{P}_\xi(B \triangle C(B_n)) = \mathbb{P}_{\mathbf{T}_\sigma(\xi)}(B \triangle C(B_n)) \\ &= \mathbb{P}(\{\omega \in \Omega : \mathbf{T}_\sigma(\xi)(\omega) \in B\} \triangle \{\omega \in \Omega : \mathbf{T}_\sigma(\xi)(\omega) \in C(B_n)\}) \\ &= \mathbb{P}(\{\omega \in \Omega : \xi(\omega) \in B\} \triangle \{\omega \in \Omega : \mathbf{T}_\sigma(\xi)(\omega) \in C(B_n)\}) \\ &= \mathbb{P}\left(A \triangle A_{\mathbf{T}_\sigma^{-1}(C(B_n))}\right) . \end{aligned} \tag{4.24}$$

Allora, $\lim_{n\to\infty} \mathbb{P}\left(A \triangle A_{\mathbf{T}_\sigma^{-1}(C(B_n))}\right) = 0$, ovvero $\lim_{n\to\infty} \mathbb{P}\left(A_{\mathbf{T}_\sigma^{-1}(C(B_n))}\right) = \mathbb{P}(A)$, il che implica

$$\lim_{n\to\infty} \mathbb{P}\left(A_n \cap A_{\mathbf{T}_\sigma^{-1}(C(B_n))}\right) = \mathbb{P}(A). \tag{4.25}$$

Scelta la permutazione finita σ tale che, $\sigma_i = 2n - i + 1$ per $i = 1, .., n$ e $\sigma_i = i \,\forall i \geq n+1$,

$$\begin{aligned} &\mathbb{P}\left(A_n \cap A_{\mathbf{T}_\sigma^{-1}(C(B_n))}\right) = \\ &\qquad \mathbb{P}\left(\{\omega \in \Omega : (\xi_1(\omega), .., \xi_n(\omega)) \in B_n\} \cap \{\omega \in \Omega : \left(\xi_{2n}(\omega), ..\xi_{n+1}(\omega)\right) \in B_n\}\right) , \end{aligned}$$

ma poiché le ξ_i sono v.a.i.i.d.,

$$\mathbb{P}\left(A_n \cap A_{\mathbf{T}_\sigma^{-1}(C(B_n))}\right) = \mathbb{P}(A_n)\,\mathbb{P}\left(A_{\mathbf{T}_\sigma^{-1}(C(B_n))}\right) = \mathbb{P}^2(A_n) \underset{n\to\infty}{\longrightarrow} \mathbb{P}^2(A). \tag{4.26}$$

Perciò $\forall A \in \mathcal{F}$ simmetrico $\mathbb{P}(A) = \mathbb{P}^2(A)$. $\blacksquare$

Osservazione 218 *Se $\{\xi_i\}_{i\geq 1}$ una successione di v.a. indipendenti definita sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, sia $\{S_n\}_{n\geq 1}$ la successione di v.a. tale che $\forall n \geq 1$, $S_n := \sum_{k=1}^{n} \xi_n$. Allora, se $\mathcal{G}_n$ è la σalgebra generata dalle v.a. $\{S_m\}_{m\geq n}$, ogni evento $A \in \mathcal{T}(S) := \bigcap_{n\geq 1} \mathcal{G}_n$ è invariante per permutazioni finite, quindi, se le ξ_i sono v.a.i.i.d., $\mathbb{P}(A)$ assume soltanto i valori 0 e 1. È altrettanto vero che ogni evento $A \in \mathcal{T}$ essendo indipendente dai valori assunti da $\xi_1, .., \xi_n$, $\forall n \geq 1$, è invariante per permutazioni finite. Quindi la Legge 0-1 di Kolmogorov nel caso di una successione di v.a.i.i.d., risulta essere un corollario di quella di Hewitt-Savage.*

4.3 Successioni di v.a. stazionarie (in senso stretto)

Sia $\mathbb{R}^{\mathbb{N}} \ni \mathbf{x} \longmapsto \mathbf{S}\mathbf{x} \in \mathbb{R}^{\mathbb{N}}$ l'applicazione tale che $\forall i \geq 1$, $(\mathbf{S}\mathbf{x})_i = x_{i+1}$ e $\forall B \in \mathcal{C}$ sia

$$\mathbf{S}^{-1}B := \{\mathbf{x} \in \mathbb{R}^{\mathbb{N}} : \mathbf{S}x \in B\}. \tag{4.27}$$

Dunque, $\forall n \geq 2, \mathbf{S}^{-n}B = \mathbf{S}^{-1}\left(\mathbf{S}^{-(n-1)}B\right)$. Inoltre, se $B \in \mathcal{C}$, da (4.20) segue che, $\forall n \geq 1$,

$$A_{\mathbf{S}^{-n}B} = \{\omega \in \Omega : \mathbf{S}^n\xi(\omega) \in B\}\ . \tag{4.28}$$

Definizione 219 *Sia $\xi = \{\xi_i\}_{i\geq 1}$ una successione di v.a. definita sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$. Un evento $A \in \mathcal{F}$ è detto* invariante *relativamente alla successione ξ se $\exists B \in \mathcal{C}$ tale che $\forall n \geq 1, A = A_{\mathbf{S}^{-n}B}$. Sia inoltre $\mathcal{I}$ la σalgebra generata dalla collezione di tali insiemi.*

Definizione 220 *Una v.a. η definita sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ è detta* invariante *relativamente alla successione ξ se è $\mathcal{I}$ misurabile, ovvero se esiste una v.a. φ definita su $\left(\mathbb{R}^{\mathbb{N}}, \mathcal{C}\right)$ tale che $\eta = \varphi \circ \xi = \varphi \circ \mathbf{S}\xi$.*

Definizione 221 *Una successione di v.a. $\xi = \{\xi_i\}_{i\geq 1}$ definita su $(\Omega, \mathcal{F}, \mathbb{P})$ è detta* stazionaria (in senso stretto, rispetto a $\mathbb{P}$) *se $\forall B \in \mathcal{C}$, $n \geq 1\, \mathbb{P}(A_B) = \mathbb{P}(A_{\mathbf{S}^{-n}B})$.*

Definizione 222 *Una successione di v.a. $\xi = \{\xi_i\}_{i\geq 1}$ definita su $(\Omega, \mathcal{F}, \mathbb{P})$, stazionaria, è detta* ergodica (rispetto a $\mathbb{P}$) *se, $\forall A \in \mathcal{F}$ invariante rispetto a ξ, $\mathbb{P}(A)$ può assumere solo i valori zero e uno.*

Proposizione 223 *Ogni evento invariante per la successione di v.a. $\xi = \{\xi_i\}_{i\geq 1}$ definita su $(\Omega, \mathcal{F}, \mathbb{P})$ appartiene alla σalgebra di coda.*

Dimostrazione: $\mathcal{T} = \bigcap_{n\geq 1} \mathcal{F}^n$, ma $\mathcal{F}^n$ è la σalgebra generata dalla successione $\mathbf{S}^n\xi$ quindi $\mathcal{T} = \bigcap_{n\geq 1} \mathbf{S}^{-n}\mathcal{F}_\xi$ con $\mathcal{F}_\xi \subseteq \mathcal{F}$ la σalgebra generata dalla collezione d'insiemi $\{A_B\}_{B\in\mathcal{C}}$. Poiché se A è invariante, per definizione $\exists B \in \mathcal{C}$ tale che $\forall n \geq 1$,

$$A = \{\omega \in \Omega : \mathbf{S}^n\xi \in B\}, \tag{4.29}$$

$A \in \mathbf{S}^{-n}\mathcal{F}_\xi = \mathcal{F}^n$. ■

Osservazione 224 *Se $\xi = \{\xi_i\}_{i\geq 1}$ è una successione di v.a.i.i.d., ξ è stazionaria in senso stretto. Quindi, poiché la σalgebra degli insiemi invarianti è contenuta nella σalgebra di coda, l'ergodicità della successione ξ segue dalla Legge 0-1 di Kolmogorov.*

Teorema 225 *(Ergodico massimale) Sia $\xi = \{\xi_i\}_{i\geq 1}$ una successione di v.a. definite sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ stazionaria e tale che $\xi_1 \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Sia $\forall n \geq 1$, $S_n := \sum_{k=1}^n \xi_k$ e $M_n := \max\{0, S_1, .., S_n\}$. Allora,*

$$\mathbb{E}\left[\xi_1 \mathbf{1}_{\{\omega\in\Omega: M_n(\omega)>0\}}\right] \geq 0 \ . \tag{4.30}$$

Dimostrazione: Poiché $\forall n \geq 1$ sia M_n che S_n dipendono soltanto dal vettore aleatorio $(\xi_1, .., \xi_n)$, siano $\bar{M}_n, \bar{S}_n$ variabili aleatorie su $\left(\mathbb{R}^{\mathbb{N}}, \mathcal{C}\right)$ tali che $M_n = \bar{M}_n \circ \xi, S_n = \bar{S}_n \circ \xi$. Allora, $\forall 1 \leq k \leq n$,

$$\bar{M}_n \circ \mathbf{S}\xi \geq \bar{S}_k \circ \mathbf{S}\xi \ . \tag{4.31}$$

Perciò

$$\xi_1 + \bar{M}_n \circ \mathbf{S}\xi \geq \bar{S}_k \circ \mathbf{S}\xi + \xi_1 = \bar{S}_{k+1} \circ \xi \ . \tag{4.32}$$

Quindi,

$$\xi_1 \geq \xi_1 - \bar{M}_n \circ \mathbf{S}\xi = \bar{S}_1 \circ \xi - \bar{M}_n \circ \mathbf{S}\xi \ . \tag{4.33}$$

Allora

$$\xi_1 \geq \max\left\{\bar{S}_1 \circ \xi, .., \bar{S}_n \circ \xi\right\} - \bar{M}_n \circ \mathbf{S}\xi \tag{4.34}$$

e

$$\mathbb{E}\left[\xi_1 \mathbf{1}_{\{\omega\in\Omega: M_n(\omega)>0\}}\right] \geq \mathbb{E}\left[\left(\max\left\{\bar{S}_1 \circ \xi, .., \bar{S}_n \circ \xi\right\} - \bar{M}_n \circ \mathbf{S}\xi\right) \mathbf{1}_{\{\omega\in\Omega: M_n(\omega)>0\}}\right] \ . \tag{4.35}$$

Ma

$$\max\left\{\bar{S}_1 \circ \xi, .., \bar{S}_n \circ \xi\right\} \mathbf{1}_{\{\omega\in\Omega: M_n(\omega)>0\}} = \bar{M}_n \circ \xi \ . \tag{4.36}$$

Quindi,

$$\begin{aligned} \mathbb{E}\left[\xi_1 \mathbf{1}_{\{\omega\in\Omega: M_n(\omega)>0\}}\right] &\geq \mathbb{E}\left[\left(\bar{M}_n \circ \xi - \bar{M}_n \circ \mathbf{S}\xi\right) \mathbf{1}_{\{\omega\in\Omega: M_n(\omega)>0\}}\right] \\ &\geq \mathbb{E}\left[\left(\bar{M}_n \circ \xi - \bar{M}_n \circ \mathbf{S}\xi\right)\right] = 0 \end{aligned} \tag{4.37}$$

poiché dato che ξ è stazionaria $\xi \overset{d}{=} \mathbf{S}\xi$. ■

Teorema 226 *(Ergodico per successioni di v.a. stazionarie) Sia $\xi = \{\xi_i\}_{i\geq 1}$ una successione di v.a. definite sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ ergodica rispetto a $\mathbb{P}$, tale che $\xi_1 \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Allora, la successione di v.a. $\left\{\frac{S_n}{n}\right\}_{n\geq 1}$ tale che, $\forall n \geq 1, S_n := \sum_{k=1}^n \xi_k$, converge a $\mathbb{E}[\xi_1]$ $\mathbb{P}$-q.c..*

Dimostrazione: Siano $\overline{\eta} := \overline{\lim}_n \left(\frac{S_n}{n} - \mathbb{E}[\xi_1|\mathcal{I}]\right)$ e $\underline{\eta} := \underline{\lim}_n \left(\frac{S_n}{n} - \mathbb{E}[\xi_1|\mathcal{I}]\right)$. Poiché $\overline{\eta}$ è invariante, ovvero è $\mathcal{I}$ misurabile, $\forall \varepsilon > 0, A_\varepsilon := \{\omega \in \Omega : \overline{\eta}(\omega) > \varepsilon\}$ è invariante. Definiamo $\forall i \geq 1$

$$\xi_i^* := (\xi_i - \mathbb{E}[\xi_1|\mathcal{I}] - \varepsilon) \mathbf{1}_{A_\varepsilon} \tag{4.38}$$

e poniamo

$$S_k^* := \sum_{l=1}^k \xi_l^* \ , \qquad M_n^* := \max\{0, S_1^*, .., S_n^*\} \ . \tag{4.39}$$

Allora, poiché $M_n^* \leq M_{n+1}^*$,

$$\{\omega \in \Omega : M_n^*(\omega) > 0\} \subseteq \left\{\omega \in \Omega : M_{n+1}^*(\omega) > 0\right\} \ , \tag{4.40}$$

dunque,

$$\begin{aligned}\lim_{n\to\infty} \{\omega \in \Omega : M_n^*(\omega) > 0\} &= \left\{\omega \in \Omega : \sup_{n\geq 1} S_n^*(\omega) > 0\right\} \\ &= \left\{\omega \in \Omega : \sup_{n\geq 1} \frac{S_n^*(\omega)}{n} > 0\right\} = \left\{\omega \in \Omega : \sup_{n\geq 1} \frac{S_n(\omega)}{n} - \mathbb{E}[\xi_1|\mathcal{I}] > \varepsilon\right\} \cap A_\varepsilon\end{aligned}$$

ma, poiché

$$\sup_{n\geq 1} \frac{S_n(\omega)}{n} \geq \inf_{k\geq 1} \sup_{n\geq k} \frac{S_n(\omega)}{n} = \overline{\eta} + \mathbb{E}[\xi_1|\mathcal{I}] \,, \tag{4.41}$$

allora

$$\left\{\omega \in \Omega : \sup_{n\geq 1} \frac{S_n(\omega)}{n} - \mathbb{E}[\xi_1|\mathcal{I}] > \varepsilon\right\} \supseteq A_\varepsilon \tag{4.42}$$

perciò

$$\lim_{n\to\infty} \{\omega \in \Omega : M_n^*(\omega) > 0\} = A_\varepsilon \ . \tag{4.43}$$

Inoltre, poiché $\xi_1 \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, anche $\xi_1^* \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ dato che, $\mathbb{E}[|\xi_1^*|] \leq \mathbb{E}[|\xi_1|] + \varepsilon$. Pertanto, per il Teorema ergodico massimale, $\forall n \geq 1, \mathbb{E}\left[\xi_1^* \mathbf{1}_{\{\omega\in\Omega : M_n^*(\omega)>0\}}\right] \geq 0$. Dunque, poiché $A_\varepsilon \in \mathcal{I}$,

$$\begin{aligned}0 &\leq \lim_{n\to\infty} \mathbb{E}\left[\xi_1^* \mathbf{1}_{\{\omega\in\Omega : M_n^*(\omega)>0\}}\right] = \mathbb{E}[\xi_1^* \mathbf{1}_{A_\varepsilon}] = \mathbb{E}[(\xi_1 - \mathbb{E}[\xi_1|\mathcal{I}] - \varepsilon) \mathbf{1}_{A_\varepsilon}] \qquad (4.44)\\ &= \mathbb{E}[\xi_1 \mathbf{1}_{A_\varepsilon} - \mathbb{E}[\xi_1 \mathbf{1}_{A_\varepsilon}|\mathcal{I}]] - \varepsilon\mathbb{P}(A_\varepsilon) = -\varepsilon\mathbb{P}(A_\varepsilon) \ ,\end{aligned}$$

ovvero $\forall \varepsilon > 0, \mathbb{P}(A_\varepsilon) = 0$, il che implica $\overline{\eta} \leq 0$ $\mathbb{P}$-q.c..

Considerando al posto di ξ, la successione $\{-\xi_i\}_{i\geq 1}$ si ha che

$$\overline{\lim_n} \left(\left(-\frac{S_n}{n}\right) + \mathbb{E}[\xi_1|\mathcal{I}]\right) = -\underline{\eta} \tag{4.45}$$

da cui segue $-\underline{\eta} \leq 0$ $\mathbb{P}$-q.c., cioè $\underline{\eta} \geq 0$ $\mathbb{P}$-q.c.. Quindi, $0 \leq \underline{\eta} \leq \overline{\eta} \leq 0$ $\mathbb{P}$-q.c., ovvero

$$\lim_{n\to\infty} \frac{S_n}{n} = \mathbb{E}[\xi_1|\mathcal{I}] \ \ \mathbb{P}\text{-q.c.,} \tag{4.46}$$

ma poiché la successione è ergodica, $\mathbb{E}[\xi_1|\mathcal{I}] = \mathbb{E}[\xi_1]$ $\mathbb{P}$-q.c.. ■

4.3.1 Introduzione alla teoria ergodica

Definizione 227 *Sia $(\Omega, \mathcal{F})$ uno spazio misurabile. Un'applicazione $T : \Omega \longrightarrow \Omega$ è detta* misurabile *se, $\forall A \in \mathcal{F}$,*

$$T^{-1}A := \{\omega \in \Omega : T\omega \in A\} \in \mathcal{F} \ . \tag{4.47}$$

Definizione 228 *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità. Un'applicazione $T : \Omega \longrightarrow \Omega$ è detta* preservare la misura *se è misurabile e se $\forall A \in \mathcal{F}$,*

$$\mathbb{P}(A) = \mathbb{P}\left(T^{-1}A\right) \ , \tag{4.48}$$

ovvero se per ogni v.a. η definita su $(\Omega, \mathcal{F}, \mathbb{P})$ si ha $\mathbb{E}[\eta] = \mathbb{E}[\eta \circ T]$.

Proposizione 229 *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità, η una v.a. su di esso definita e $T : \Omega \longrightarrow \Omega$ un'applicazione che preserva la misura. Ponendo per $k = 1, \xi_1 := \eta$ e, $\forall k \geq 2, \xi_k := \eta \circ T^{k-1}$, la successione di v.a. $\xi = \{\xi_k\}_{k \geq 1}$ è stazionaria.*

Dimostrazione: $\forall B \in \mathcal{C}$,

$$A_{\mathbf{S}^{-1}B} = \{\omega \in \Omega : \mathbf{S}\xi(\omega) \in B\} = \{\omega \in \Omega : \xi(T\omega) \in B\} \ . \tag{4.49}$$

Pertanto, $\omega \in A_{\mathbf{S}^{-1}B} \Longleftrightarrow T\omega \in A_B$, ovvero $A_{\mathbf{S}^{-1}B} = T^{-1}A_B$. Ma siccome T preserva la misura $\mathbb{P}(A_{\mathbf{S}^{-1}B}) = \mathbb{P}(T^{-1}A_B) = \mathbb{P}(A_B)$. Ripetendo questo argomento per $A_{\mathbf{S}^{-k}B}$ con $k \geq 2$ si ha la tesi. ■

Proposizione 230 *Sia ζ una successione di v.a. stazionaria definita su uno spazio di probabilità $(\Xi, \mathcal{X}, \mathbb{Q})$. Allora si può costruire uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, una v.a.$\eta$ su di esso definita e un'applicazione $T : \Omega \longrightarrow \Omega$ che preserva la misura tale che la successione di v.a. ξ, definita su $(\Omega, \mathcal{F}, \mathbb{P})$ come nella proposizione precedente coincide con ζ in distribuzione.*

Dimostrazione: Poniamo $\Omega := \mathbb{R}^{\mathbb{N}}, \mathcal{F} := \mathcal{C}, \mathbb{P} := \mathbb{Q} \circ \zeta^{-1}, T := \mathbf{S}$ e η tale che $\mathbb{R}^{\mathbb{N}} \ni \mathbf{x} \longmapsto \eta(\mathbf{x}) := x_1 \in \mathbb{R}$. Allora $\forall n \geq 1$ e ogni $B \in \mathcal{B}(\mathbb{R}^n)$,

$$\begin{aligned} T^{-1}C(B) &= \{\omega \in \Omega : \mathbf{S}\omega \in C(B)\} = \{\omega \in \Omega : (\omega_2, .., \omega_{n+1}) \in B\} \\ &= \left\{\omega \in \Omega : \left((\eta \circ T)(\omega), .., \left(\eta \circ T^{n+1}\right)(\omega)\right) \in B\right\} \\ &= \left\{\omega \in \Omega : \left(\xi_2(\omega), .., \xi_{n+1}(\omega)\right) \in B\right\} \ . \end{aligned} \tag{4.50}$$

Poiché ζ è stazionaria si ha

$$\begin{aligned} \mathbb{P}(C(B)) &= \mathbb{Q}\{x \in \Xi : (\zeta_1(x), .., \zeta_n(x)) \in B\} \\ &= \mathbb{Q}\left\{x \in \Xi : \left(\zeta_2(x), .., \zeta_{n+1}(x)\right) \in B\right\} = \mathbb{P}\left(T^{-1}C(B)\right) \ , \end{aligned} \tag{4.51}$$

ovvero T preserva la misura. Inoltre, siccome $\forall n \geq 1$ e ogni $B \in \mathcal{B}(\mathbb{R}^n)$,

$$\begin{aligned} \mathbb{P}(C(B)) = \mathbb{P}\{\omega \in \Omega : (\xi_1(\omega), .., \xi_n(\omega)) \in B\} = \\ \mathbb{Q}\{x \in \Xi : (\zeta_1(x), .., \zeta_n(x)) \in B\} \ , \end{aligned} \tag{4.52}$$

$\zeta \stackrel{d}{=} \xi$. ■

Esempio 1 Un esempio di trasformazione che preserva la misura è il flusso di fase hamiltoniano che, per il Teorema di Liouville, conserva la misura di Lebesgue nello spazio delle fasi.

Teorema 231 *(di ricorrenza di Poincaré) Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità,* $T : \Omega \longrightarrow \Omega$ *un'applicazione che preserva la misura e* $A \in \mathcal{F}$*. Allora,* $\mathbb{P}$*-q.c. se* $\omega \in A, T^n\omega \in A$ *per infiniti valori di* $n \geq 1$.

Dimostrazione: Sia

$$N_0(A) := \{\omega \in A : T^n\omega \notin A, \forall n \geq 1\} = \bigcap_{n\geq 1} A \backslash \{\omega \in A : T^n\omega \in A\} \tag{4.53}$$
$$= A \backslash \bigcup_{n\geq 1} \{\omega \in \Omega : \omega \in A,\ T^n\omega \in A\}$$

l'insieme degli elementi di A la cui immagine sotto le trasformazioni della famiglia $\{T^n\}_{n\geq 1}$ non appartiene ad A. Siccome $\forall n \geq 1, N_0(A) \cap T^{-n}N_0(A) = \varnothing$, allora, $\forall m \geq 1$,

$$T^{-m}N_0(A) \cap T^{-(m+n)}N_0(A) = \varnothing \ . \tag{4.54}$$

Siccome T preserva la misura $\{T^{-n}N_0(A)\}_{n\geq 1}$ è una successione d'insiemi disgiunti di uguale misura

$$\sum_{n\geq 1} \mathbb{P}(N_0(A)) = \mathbb{P}(N_0(A)) + \sum_{n\geq 1} \mathbb{P}\left(T^{-n}N_0(A)\right) \leq \mathbb{P}(\Omega) = 1 \ , \tag{4.55}$$

ovvero $\mathbb{P}(N_0(A)) = 0$. Quindi, per $\mathbb{P}$-quasi ogni $\omega \in A, T^n\omega \in A$ per almeno un $n \geq 1$. Questo argomento si può ripetere sostituendo T^kA ad $A, \forall k \geq 1$. Pertanto, poiché posto $\forall k \geq 1, N_k(A) := N_0\left(T^kA\right), \mathbb{P}(N_k(A)) = 0$, quindi $\forall \omega \in A \backslash N(A)$, dove $N(A) := \bigcup_{k\geq 0} N_k(A), \forall m \geq 1, \exists n_m \geq 1$ tale che $(T^m)^{n_m}\omega \in A$. Dunque, $T^n\omega \in A$ per infiniti valori di $n \geq 1$. ■

Corollario 232 *Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità,* $T : \Omega \longrightarrow \Omega$ *un'applicazione che preserva la misura e* η *una v.a. non negativa. Allora,* $\mathbb{P}$*-q.c. su* $\{\omega \in \Omega : \eta(\omega) > 0\}$,

$$\eta(\omega) + \sum_{k\geq 1} \eta\left(T^k\omega\right) = \infty \ . \tag{4.56}$$

Dimostrazione: Sia $A_n := \left\{\omega \in \Omega : \eta(\omega) \geq \frac{1}{n}\right\}$. Allora, per il teorema precedente, $\mathbb{P}$-q.c. su A_n,

$$\eta(\omega) + \sum_{k\geq 1} \eta\left(T^k\omega\right) = \infty \ , \tag{4.57}$$

quindi anche su $\bigcup_{n\geq 1} A_n = \{\omega \in \Omega : \eta(\omega) > 0\}$. ■

Definizione 233 *Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità e* $T : \Omega \longrightarrow \Omega$ *un'applicazione misurabile. Un evento* $A \in \mathcal{F}$ *è detto*

- invariante *per* T *se* $T^{-1}A = A$;
- quasi invariante *per* T *se* $\mathbb{P}(A \triangle T^{-1}A) = 0$.

Chiaramente sia la collezione degli eventi invarianti $\mathcal{I}_T$ che quella degli eventi quasi invarianti $\mathcal{I}_T^*$ sono subσalgebre di $\mathcal{F}$. In particolare, $\mathcal{I}_T \subseteq \mathcal{I}_T^* \subseteq \mathcal{F}$.

Lemma 234 *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità e $T : \Omega \longrightarrow \Omega$ un'applicazione che preserva la misura. Se $A \in \mathcal{I}_T^*$ allora $\exists B \in \mathcal{I}_T$ tale che $\mathbb{P}(A \triangle B) = 0$.*

Dimostrazione: Sia $B := \overline{\lim}_n T^{-n}A = \bigcap_{n\geq 1} \bigcup_{k\geq n} T^{-k}A$. Allora, $T^{-1}B = B$ quindi $B \in \mathcal{I}_T$. Inoltre,

$$A \triangle B = \left(A \triangle T^{-1}A\right) \cup \left(\bigcup_{k\geq 1} \left(T^{-k}A \triangle T^{-(k+1)}A\right)\right) , \tag{4.58}$$

ma T preserva la misura, quindi

$$\mathbb{P}\left(\left(T^{-k}A \triangle T^{-(k+1)}A\right)\right) = \mathbb{P}\left(A \triangle T^{-1}A\right) = 0 \tag{4.59}$$

da cui segue che $\mathbb{P}(A \triangle B) = 0$. ■

Definizione 235 *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità e $T : \Omega \longrightarrow \Omega$ un'applicazione misurabile. Una v.a. η è detta* invariante *per T se è $\mathcal{I}_T$-misurabile, ovvero $\eta^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{I}_T$, o, equivalentemente $\eta = \eta \circ T$. Allo stesso modo η si dirà quasi invariante per T se è $\mathcal{I}_T^*$-misurabile ($\eta^{-1}(\mathcal{B}(\mathbb{R})) \subseteq \mathcal{I}_T^*$).*

Definizione 236 *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità e $T : \Omega \longrightarrow \Omega$ un'applicazione misurabile. Un evento $A \in \mathcal{I}_T$ è detto* metricamente indecomponibile *se non si può rappresentare come unione disgiunta di una coppia d'eventi invarianti di propabilità positiva, ovvero $\nexists A_1, A_2 \in \mathcal{I}_T$ tali che $A_1 \cap A_2 = \varnothing; \mathbb{P}(A_1), \mathbb{P}(A_2) > 0$ e $A = A_1 \bigvee A_2$.*

Definizione 237 *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità. Un'applicazione $T : \Omega \longrightarrow \Omega$ che preserva la misura è detta* ergodica *o* metricamente transitiva *se $\forall A \in \mathcal{I}_T, \mathbb{P}(A)$ può assumere soltanto i valori zero e uno.*

Proposizione 238 *Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità e $T : \Omega \longrightarrow \Omega$ un'applicazione ergodica. Allora ogni evento invariante di probabilità positiva è metricamente indecomponibile se e solo se ha probabilità uno.*

Dimostrazione:

$\Longrightarrow$ T è ergodica, quindi per definizione ogni evento invariante di probabilità positiva ha probabilità uno.

$\Longleftarrow$ Sia $A \in \mathcal{I}_T$ tale che $\mathbb{P}(A) = 1$. Se A non fosse metricamente indecomponibile esisterebbero $A_1, A_2 \in \mathcal{I}_T$ disgiunti tali che $\mathbb{P}(A_1), \mathbb{P}(A_2) > 0$ e $A = A_1 \bigvee A_2$. Ma siccome T è ergodica

$$\mathbb{P}(A_1) = \mathbb{P}(A_2) = 1 = \mathbb{P}(A) \ ; \tag{4.60}$$

perciò si avrebbe che

$$1 = \mathbb{P}(A) = \mathbb{P}(A_1) + \mathbb{P}(A_2) = 2 \ . \tag{4.61}$$

■

Lemma 239 *Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità e* $T : \Omega \longrightarrow \Omega$ *un'applicazione che preserva la misura. Allora,* T *è ergodica se e solo se ogni evento quasi invariante ha probabilità zero o uno.*

Dimostrazione:

$\Longrightarrow$ Se T preserva la misura, dal Lemma 234 segue che se $A \in \mathcal{I}_T^*$ allora $\exists B \in \mathcal{I}_T$ tale che $\mathbb{P}(A \triangle B) = 0$. Inoltre, se T è ergodica, la probabilità di un qualsiasi evento invariante, quindi anche di B, può assumere soltanto i valori zero e uno, dunque lo stesso vale per la probabilità di A.

$\Longleftarrow$ Segue dal fatto che $\mathcal{I}_T \subseteq \mathcal{I}_T^*$.

■

Teorema 240 *Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità e* $T : \Omega \longrightarrow \Omega$ *un'applicazione che preserva la misura. Le seguenti affermazioni sono equivalenti:*

1. T *è ergodica.*
2. *Ogni v.a.* η *quasi invariante per* T *è costante* $\mathbb{P}$*-q.c..*
3. *Ogni v.a.* η *invariante per* T *è costante* $\mathbb{P}$*-q.c..*

Dimostrazione:

1) $\Longrightarrow$ 2) Se η è quasi invariante, allora $\forall c \in \mathbb{R}, A_c := \{\omega \in \Omega : \eta(\omega) \leq c\} \in \mathcal{I}_T^*$, e siccome T è ergodica $\mathbb{P}(A_c)$ può assumere soltanto i valori zero e uno. Posto $\kappa := \sup\{c \in \mathbb{R} : \mathbb{P}(A_c) = 0\}$, siccome $A_c \uparrow \Omega$ per $c \to \infty$ e $A_c \downarrow \varnothing$ per $c \to -\infty$ allora $|\kappa| < \infty$. Dunque,

$$\mathbb{P}\{\omega \in \Omega : \eta(\omega) < \kappa\} = \mathbb{P}\left(\bigcup_{n \geq 1}\left\{\omega \in \Omega : \eta(\omega) \leq \kappa - \frac{1}{n}\right\}\right) = 0 \ . \tag{4.62}$$

Allo stesso modo si ha che $\mathbb{P}\{\omega \in \Omega : \eta(\omega) > \kappa\} = 0$, dunque $\mathbb{P}\{\omega \in \Omega : \eta(\omega) = \kappa\} = 1$.

2) $\Longrightarrow$ 3) Se η è invariante è anche $\mathcal{I}_T^*$-misurabile perché $\mathcal{I}_T \subseteq \mathcal{I}_T^*$.

3) $\Longrightarrow$ 1) Dato $A \in \mathcal{I}_T$, allora $\mathbf{1}_A$ è una v.a. invariante per T che per ipotesi è costante $\mathbb{P}$-q.c.. Ma poiché $\mathbf{1}_A \in \{0, 1\}, \mathbb{P}(A)$ può assumero soltanto i valori zero e uno e dunque T è ergodica.

■

Definizione 241 *Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità. Un'applicazione* $T : \Omega \longrightarrow \Omega$ *che preserva la misura è detta* mixing *se* $\forall A, B \in \mathcal{F}$

$$\lim_{n \to \infty} \mathbb{P}\left(A \cap T^{-n} B\right) = \mathbb{P}(A)\, \mathbb{P}(B) \ . \tag{4.63}$$

Teorema 242 *Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità. Ogni applicazione* $T : \Omega \longrightarrow \Omega$ *mixing è ergodica.*

Dimostrazione: $\forall A \in \mathcal{F}, B \in \mathcal{I}_T$ poiché T è mixing, $\forall n \geq 1$, si ha

$$\mathbb{P}\left(A \cap T^{-n}B\right) = \mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B) \tag{4.64}$$

che, ponendo $A = B$, implica $\mathbb{P}(B) = \mathbb{P}^2(B)$, ovvero che ogni insieme invariante ha probabilità o zero o uno. ■

Teorema 243 *(di Birkhoff-Khinchin) Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità,* $T : \Omega \longrightarrow \Omega$ *un'applicazione che preserva la misura e* $\eta \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. *Allora,*

$$\lim_{n\to\infty} \frac{1}{n}\left[\eta + \sum_{k=1}^{n-1} \eta \circ T^k\right] = \mathbb{E}[\eta|\mathcal{I}_T] \;\; \mathbb{P}\text{-}q.c.. \tag{4.65}$$

Se inoltre T *è ergodica* $\mathbb{E}[\eta|\mathcal{I}_T] = \mathbb{E}[\eta] \;\; \mathbb{P}$*-q.c..*

Dimostrazione: Ponendo per $k = 1, \xi_1 := \eta$ e, $\forall k \geq 2, \xi_k := \eta \circ T^{k-1}$, per la Proposizione 229, la successione di v.a. $\xi = \{\xi_k\}_{k\geq 1}$ è stazionaria, quindi la tesi segue dal Teorema ergodico per successioni di v.a. stazionarie. Inoltre, se T è ergodica, dato che $\mathbb{E}[\eta|\mathcal{I}_T]$ è $\mathcal{I}_T$-misurabile è invariante e quindi, per il Teorema 240 $\mathbb{P}$-q.c. costante, ovvero $\mathbb{E}[\eta|\mathcal{I}_T] = \mathbb{E}[\eta]$ $\mathbb{P}$-q.c.. ■

Corollario 244 *Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità. Un'applicazione* $T : \Omega \longrightarrow \Omega$ *che preserva la misura è ergodica se e solo se* $\forall A, B \in \mathcal{F}$

$$\lim_{n\to\infty} \frac{1}{n}\left[\mathbb{P}(A \cap B) + \sum_{k=1}^{n-1} \mathbb{P}\left(A \cap T^{-k}B\right)\right] = \mathbb{P}(A)\,\mathbb{P}(B) \;\; . \tag{4.66}$$

Dimostrazione:

$\Longrightarrow$ Se T è ergodica, per il Teorema di Birkhoff-Khinchin $\forall B \in \mathcal{F}$

$$\lim_{n\to\infty} \frac{1}{n}\left[\mathbf{1}_B + \sum_{k=1}^{n-1} \mathbf{1}_B \circ T^k\right] = \lim_{n\to\infty} \frac{1}{n}\left[\mathbf{1}_B + \sum_{k=1}^{n-1} \mathbf{1}_{T^{-k}B}\right] = \mathbb{P}(B) \;\; \mathbb{P}\text{-q.c..} \tag{4.67}$$

Quindi, per il Teorema di Lebesgue della convergenza dominata,

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{n}\left[\mathbb{P}(A \cap B) + \sum_{k=1}^{n-1} \mathbb{P}\left(A \cap T^{-k}B\right)\right] \qquad (4.68)\\ &= \lim_{n\to\infty} \frac{1}{n}\left[\mathbb{E}[\mathbf{1}_{A\cap B}] + \sum_{k=1}^{n-1} \mathbb{E}[\mathbf{1}_{A\cap T^{-k}B}]\right] \\ &= \mathbb{E}\left[\mathbf{1}_A \lim_{n\to\infty} \frac{1}{n}\left(\mathbf{1}_B + \sum_{k=1}^{n-1} \mathbf{1}_{T^{-k}B}\right)\right] = \mathbb{P}(A)\,\mathbb{P}(B) \;\; . \end{aligned}$$

$\Longleftarrow$ Posto $A = B \in \mathcal{I}_T, \forall k \geq 1$, si ha $A \cap T^{-k}B = A \cap B = B$. Quindi, la (4.66) implica $\mathbb{P}(B) = \mathbb{P}^2(B)$, cioè che ogni insieme invariante ha probabilità o zero o uno.

■

Teorema 245 *Sia* $(\Omega, \mathcal{F}, \mathbb{P})$ *uno spazio di probabilità,* $T : \Omega \longrightarrow \Omega$ *un'applicazione che preserva la misura e* $\eta \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. *Allora,*

$$\lim_{n\to\infty} \mathbb{E}\left|\frac{1}{n}\left[\eta + \sum_{k=1}^{n-1} \eta \circ T^k\right] - \mathbb{E}[\eta|\mathcal{I}_T]\right| = 0 \ . \tag{4.69}$$

Se inoltre T *è ergodica* $\mathbb{E}[\eta|\mathcal{I}_T] = \mathbb{E}[\eta]$ $\mathbb{P}$*-q.c..*

Dimostrazione: Siccome $\eta \in L^1(\Omega, \mathcal{F}, \mathbb{P}), \forall \varepsilon > 0, \exists M_\varepsilon > 0$ e una v.a. $\eta_\varepsilon \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ tale che $|\eta_\varepsilon| \leq M_\varepsilon < \infty$ e $\mathbb{E}|\eta - \eta_\varepsilon| \leq \varepsilon$. Allora, ponendo, per $k = 1, \xi_1 := \eta, \forall k \geq 2$, $\xi_k := \eta \circ T^{k-1}$ e dunque $\forall n \geq 1, S_n := \sum_{k=1}^n \xi_k$ nonché, definendo allo stesso modo S_n^ε, si ha

$$\begin{aligned} \mathbb{E}\left|\frac{S_n}{n} - \mathbb{E}[\eta|\mathcal{I}_T]\right| &\leq \mathbb{E}\left|\frac{1}{n}\left[\eta - \eta_\varepsilon + \sum_{k=2}^{n}\left(\eta \circ T^{k-1} - \eta_\varepsilon \circ T^{k-1}\right)\right]\right| + \\ &+ \mathbb{E}\left|\frac{S_n^\varepsilon}{n} - \mathbb{E}[\eta^\varepsilon|\mathcal{I}_T]\right| + \mathbb{E}\left|\mathbb{E}[\eta|\mathcal{I}_T] - \mathbb{E}[\eta^\varepsilon|\mathcal{I}_T]\right| \ . \end{aligned} \tag{4.70}$$

Per il Teorema di Birkhoff-Khinchin $\lim_{n\to\infty} \frac{S_n^\varepsilon}{n} = \mathbb{E}[\eta^\varepsilon|\mathcal{I}_T]$ $\mathbb{P}$-q.c. ma, dato che η_ε è limitata, per il Teorema di Lebesgue della convergenza dominata

$$\lim_{n\to\infty} \mathbb{E}\left|\frac{S_n^\varepsilon}{n} - \mathbb{E}[\eta^\varepsilon|\mathcal{I}_T]\right| = 0 \ . \tag{4.71}$$

Inoltre, $\left|\eta \circ T^{k-1} - \eta_\varepsilon \circ T^{k-1}\right| \leq \varepsilon$ come pure

$$\left|\mathbb{E}[\eta|\mathcal{I}_T] - \mathbb{E}[\eta^\varepsilon|\mathcal{I}_T]\right| = \left|\mathbb{E}[(\eta - \eta^\varepsilon)|\mathcal{I}_T]\right| \leq \mathbb{E}[|\eta - \eta^\varepsilon||\mathcal{I}_T] \leq \varepsilon \ . \tag{4.72}$$

Pertanto, $\forall \varepsilon > 0, \mathbb{E}\left|\frac{S_n}{n} - \mathbb{E}[\eta|\mathcal{I}_T]\right| \leq 2\varepsilon$. Se inoltre T è ergodica, $\mathbb{E}[\eta|\mathcal{I}_T] = \mathbb{E}[\eta]$ $\mathbb{P}$-q.c.. ■

Capitolo 5

Moto browniano

5.1 Variabili aleatorie gaussiane

Definizione 246 *Una v.a.* ξ *definita su* $(\Omega, \mathcal{F}, \mathbb{P})$ *ha distribuzione* gaussiana *o* normale *di parametri* $m \in \mathbb{R}$, *tale che* $|m| < \infty$, *e* $\sigma^2 > 0$, *se questa è assolutamente continua rispetto alla misura di Lebesgue su* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *ed ha densità*

$$f_{\xi}(x) := \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}} . \tag{5.1}$$

Una tale v.a. è pertanto detta gaussiana *oppure* normale.

Dalla precedente definizione segue che se ξ è una v.a. gaussiana definita su $(\Omega, \mathcal{F}, \mathbb{P})$ di parametri m e σ^2, si ha che

$$\mathbb{E}(\xi) = m , \tag{5.2}$$

$$\mathbb{E}\left((\xi - \mathbb{E}(\xi))^2\right) = \mathbb{E}\left(\xi^2\right) - \mathbb{E}^2(\xi) = \sigma^2 . \tag{5.3}$$

Nel caso $m = 0$ e $\sigma^2 = 1$, ξ è detta avere distribuzione *normale standard* o, equivalentemente, essere *normale standard*. In tal caso, dalla definizione precedente si ha che f_{ξ} è simmetrica per cui, $\forall x > 0$,

$$\mathbb{P}\{\omega \in \Omega : \xi(\omega) > x\} = \mathbb{P}\{\omega \in \Omega : \xi(\omega) < -x\} . \tag{5.4}$$

Inoltre, se η è una v.a. gaussiana definita su $(\Omega, \mathcal{F}, \mathbb{P})$ di parametri m e σ^2, $\xi := \frac{\eta - m}{\sigma}$ è una v.a. di distribuzione normale standard. Infatti, ponendo per una qualsiasi v.a. ζ e $\forall x \in \mathbb{R}$,

$$F_{\zeta}(x) := \mathbb{P}\{\omega \in \Omega : \zeta(\omega) \leq x\} , \tag{5.5}$$

si ha

$$\mathbb{P}\left\{\omega \in \Omega : \frac{\eta(\omega) - m}{\sigma} \leq x\right\} = \mathbb{P}\{\omega \in \Omega : \eta(\omega) \leq \sigma x + m\} = \int_{-\infty}^{\sigma x + m} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-m)^2}{2\sigma^2}} dy \tag{5.6}$$

per cui

$$f_{\xi}(x) = \frac{d}{dx} F_{\xi}(x) = \frac{d}{dx} F_{\eta}(x) = \sigma \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} . \tag{5.7}$$

Osservazione 247 *Se ξ è una v.a. normale standard definita su $(\Omega, \mathcal{F}, \mathbb{P})$, allora, dall'espressione della funzione caratteristica di ξ,*

$$\varphi_{\xi}(t) := \mathbb{E}\left(e^{it\xi}\right) = e^{itm - \frac{1}{2}t^2\sigma^2}, \tag{5.8}$$

segue che, nel limite di $\sigma \downarrow 0$, ξ converge in distribuzione alla v.a. degenere costantemente uguale ad m.

Proposizione 248 (stima delle code) *Se ξ è una v.a. normale standard definita su $(\Omega, \mathcal{F}, \mathbb{P})$, allora, $\forall x > 0$, posto $n(x) := f_{\xi}(x)$, si ha*

$$\frac{n(x)}{x} - \frac{n(x)}{x^3} < \mathbb{P}\{\omega \in \Omega : \xi(\omega) > x\} < \frac{n(x)}{x}. \tag{5.9}$$

Dimostrazione: $\forall x > 0$, poiché $n'(x) = -xn(x)$,

$$\begin{aligned} \mathbb{P}\{\omega \in \Omega : \xi(\omega) > x\} &= \int_x^{+\infty} dyn(y) = \int_x^{+\infty} dyy\frac{n(y)}{y} = \int_x^{+\infty} dy\left(-\frac{n'(y)}{y}\right) = \\ &= \left[-\frac{n(y)}{y}\right]_x^{+\infty} + \int_x^{+\infty} dy\left(-\frac{n(y)}{y^2}\right) = \frac{n(x)}{x} - \int_x^{+\infty} dy\frac{n(y)}{y^2} \\ &< \frac{n(x)}{x}. \end{aligned} \tag{5.10}$$

D'altronde, dall'ultima uguaglianza segue

$$\begin{aligned} \mathbb{P}\{\omega \in \Omega : \xi(\omega) > x\} &= \frac{n(x)}{x} - \int_x^{+\infty} dy\frac{n(y)}{y^2} = \frac{n(x)}{x} - \int_x^{+\infty} dyy\frac{n(y)}{y^3} \\ &= \frac{n(x)}{x} - \int_x^{+\infty} dy\left(-\frac{n'(y)}{y^3}\right) \\ &= \frac{n(x)}{x} - \left(\left[-\frac{n(y)}{y^3}\right]_x^{+\infty} - 3\int_x^{+\infty} dy\frac{n'(y)}{y^4}\right) \\ &= \frac{n(x)}{x} - \frac{n(x)}{x^3} + 3\int_x^{+\infty} dy\frac{n'(y)}{y^4} > \frac{n(x)}{x} - \frac{n(x)}{x^3}. \end{aligned} \tag{5.11}$$

■

Osservazione 249 *Dato un vettore aleatorio (ξ, η) su $(\Omega, \mathcal{F}, \mathbb{P})$ le cui componenti appartengono a $L^2(\mathcal{F})$, sia $Var(\xi)$ la varianza della componente ξ, $Var(\eta)$ quella della componente η, $Cov(\xi, \eta)$ la covarianza delle due componenti e*

$$\rho := \frac{Cov(\xi, \eta)}{\sqrt{Var(\xi)}\sqrt{Var(\eta)}} \tag{5.12}$$

il loro coefficiente di correlazione.

Nel caso in cui ξ e η sono variabili aleatorie gaussiane la curva di regressione è lineare. Infatti,

$$f_{(\xi,\eta)}(x,y) = \frac{\exp\left\{-\frac{\left[\frac{\left(x-\mu_\xi\right)^2}{\sigma_\xi^2}-2\rho\frac{\left(x-\mu_\xi\right)(y-\mu_\eta)}{\sigma_\xi\sigma_\eta}+\frac{(y-\mu_\eta)^2}{\sigma_\eta^2}\right]}{2(1-\rho^2)}\right\}}{2\pi\sigma_\xi\sigma_\eta\sqrt{1-\rho^2}}, \tag{5.13}$$

$$\begin{aligned} f_{\eta|\xi}(y|x) = \frac{f_{(\xi,\eta)}(x,y)}{f_\xi(x)} &= \frac{\exp\left\{-\frac{\left[\frac{\left(x-\mu_\xi\right)^2}{\sigma_\xi^2}\left(1-\left(1-\rho^2\right)\right)-2\rho\frac{\left(x-\mu_\xi\right)(y-\mu_\eta)}{\sigma_\xi\sigma_\eta}+\frac{(y-\mu_\eta)^2}{\sigma_\eta^2}\right]}{2(1-\rho^2)}\right\}}{\sigma_\eta\sqrt{2\pi\left(1-\rho^2\right)}} \\ &= \frac{\exp\left\{-\frac{\left[\frac{\left(x-\mu_\xi\right)^2}{\sigma_\xi^2}\rho^2-2\rho\frac{\left(x-\mu_\xi\right)(y-\mu_\eta)}{\sigma_\xi\sigma_\eta}+\frac{(y-\mu_\eta)^2}{\sigma_\eta^2}\right]}{2(1-\rho^2)}\right\}}{\sigma_\eta\sqrt{2\pi\left(1-\rho^2\right)}} \\ &= \frac{\exp\left\{-\frac{\left[y-\left(\mu_\eta+\frac{\sigma_\eta}{\sigma_\xi}\rho\left(x-\mu_\xi\right)\right)\right]^2}{2\sigma_\eta^2(1-\rho^2)}\right\}}{\sigma_\eta\sqrt{2\pi\left(1-\rho^2\right)}}. \end{aligned} \tag{5.14}$$

Pertanto,

$$\begin{aligned} \varphi^*(\xi) = \mathbb{E}[\eta|\xi] &= \mu_\eta + \frac{\sigma_\eta}{\sigma_\xi}\rho\left(\xi-\mu_\xi\right) \\ &= \frac{Cov(\xi,\eta)}{Var(\xi)}\xi + \left(\mathbb{E}[\eta] - \frac{Cov(\xi,\eta)}{Var(\xi)}\mathbb{E}[\xi]\right) \end{aligned} \tag{5.15}$$

e

$$\begin{aligned} \|\eta-\mathbb{E}[\eta|\xi]\|_{L^2}^2 = \mathbb{E}\left[(\eta-\mathbb{E}[\eta|\xi])^2\right] &= \sigma_\eta^2\left(1-\rho^2\right) \\ &= Var(\eta) - \frac{Cov^2(\xi,\eta)}{Var(\xi)}. \end{aligned} \tag{5.16}$$

Dunque, nel caso in cui il vettore (ξ,η) è gausiano la curva di regressione è una retta, cioè $\mathbb{E}[\eta|\xi]$ è una funzione lineare di ξ.

Che $\mathbb{E}[\eta|\xi]$ è una funzione lineare di ξ, vale anche nel caso più generale in cui ξ e η sono vettori aleatori gaussiani, come si vedrà nell'Osservazione 262.

5.2 Vettori aleatori gaussiani

5.2.1 Vettori aleatori a componenti in $L^2(\Omega, \mathcal{F}, \mathbb{P})$

Definizione 250 *Dato un vettore aleatorio* $\xi := (\xi_1, .., \xi_n), n \geq 2$, *definito su uno spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$ *si definisce* matrice di varianza-covarianza *l'operatore lineare* $C(\xi) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *i cui elementi di matrice sono le covarianze delle componenti di* ξ, *ovvero*

$$C_{i,j}(\xi) := \mathbb{E}\left[(\xi_i - \mathbb{E}[\xi_i])\left(\xi_j - \mathbb{E}\left[\xi_j\right]\right)\right] \ , \ i, j = 1, .., n \ . \tag{5.17}$$

$C(\xi)$ ha le seguenti proprietà:

- è simmetrica, come evidentemente risulta dalla (5.17);
- è definita non negativa. Infatti $\forall (\lambda_1, .., \lambda_n) \in \mathbb{R}^n$

$$\begin{aligned}(C(\xi)\lambda, \lambda) &= \sum_{i,j=1}^{n} \mathbb{E}\left[(\xi_i - \mathbb{E}[\xi_i])\left(\xi_j - \mathbb{E}\left[\xi_j\right]\right)\right]\lambda_i\lambda_j \qquad (5.18)\\ &= \mathbb{E}\left[\left(\sum_{i=1}^{n}(\xi_i - \mathbb{E}[\xi_i])\lambda_i\right)^2\right] \geq 0 \ .\end{aligned}$$

Pertanto, esiste una matrice ortogonale O tale che $C = ODO^t$ dove D è diagonale. Inoltre, poiché C è non negativa tutti gli elementi di D sono non negativi. Perciò, ponendo $\sqrt{D} := diag\left[\sqrt{c_1}, .., \sqrt{c_n}\right]$, dove i numeri c_i sono gli autovalori di C, e $B := O\sqrt{D}$, si ha $C = BB^t$.

Definizione 251 *Sia* $\xi = (\xi_1, ..\xi_n) \in \mathbb{R}^n, n \geq 2$, *un vettore aleatorio definito su uno spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$, *tale che,* $\forall i = 1, .., n, \xi_i \in L^2(\mathcal{F})$. *Le componenti di* ξ *si dicono* linearmente indipendenti *se non esiste* $a = (a_1, .., a_n) \in \mathbb{R}^n$ *diverso dal vettore nullo tale che*

$$\sum_{i=1}^{n} a_i\xi_i = 0 \ \mathbb{P} - q.c.. \tag{5.19}$$

Chiaramente il sottospazio lineare *span* $(\xi_1, .., \xi_n)$ generato dalle componenti di ξ è un sottospazio lineare di $L^2\left(\mathcal{F}^{\xi}\right)$, quindi di $L^2(\mathcal{F})$. Inoltre, se ξ è un vettore aleatorio definito su $(\Omega, \mathcal{F}, \mathbb{P})$ a componenti in $L^2(\mathcal{F})$ linearmete indipendenti, la matrice di varianza-covarianza $C(\xi) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ di ξ ha rango massimo, allora il vettore aleatorio $\eta := B^{-1}\xi = \left(O\sqrt{D}\right)^{-1}\xi = \sqrt{D^{-1}}O^t\xi$, dove le matrici B, O sono quelle introdotte nella dimostrazione della proposizione precedente e

$\sqrt{D^{-1}} := diag\left[\frac{1}{\sqrt{c_1}}, .., \frac{1}{\sqrt{c_n}}\right]$, ha componenti ortogonali in $L^2(\mathcal{F})$. Infatti,

$$
\begin{aligned}
Cov\left[\eta_i\eta_j\right] &= \mathbb{E}\left[\left(\frac{1}{\sqrt{c_i}}\sum_{k=1}^{n} O_{i,k}^t\left(\xi_k - \mathbb{E}\left[\xi_k\right]\right)\right)\left(\frac{1}{\sqrt{c_j}}\sum_{l=1}^{n} O_{j,l}^t\left(\xi_l - \mathbb{E}\left[\xi_l\right]\right)\right)\right] \\
&= \frac{1}{\sqrt{c_i}}\frac{1}{\sqrt{c_j}}\sum_{k,l=1}^{n} O_{i,k}^t O_{j,l}^t Cov\left[\xi_k\xi_l\right] \\
&= \frac{1}{\sqrt{c_i}}\frac{1}{\sqrt{c_j}}\sum_{k,l=1}^{n} O_{i,k}^t C_{k,l}\left(\xi\right) O_{l,j} = \frac{c_j\delta_{i,j}}{\sqrt{c_i c_j}} = \delta_{i,j}
\end{aligned} \tag{5.20}
$$

e pertanto le componenti del vettore $\zeta := (\zeta_1, .., \zeta_n)$ tale che $\forall i = 1, ...n, \zeta_i := \eta_i - \mathbb{E}[\eta_i]$, formano una base ortonormale di $span(\xi_1, .., \xi_n)$.

Dato un vettore aleatorio ξ definito su $(\Omega, \mathcal{F}, \mathbb{P})$ a componenti in $L^2(\mathcal{F})$, se $\det C(\xi) = 0$ sia $r < n$ il rango di $C(\xi)$. Allora esistono $n - r$ vettori linearmente indipendenti $\lambda^{(1)}, .., \lambda^{(n-r)} \in \mathbb{R}^n$, tali che $\forall i = 1, .., n, \lambda^{(i)} = \left(\lambda_1^{(i)}, .., \lambda_n^{(i)}\right)$, per cui

$$
\begin{aligned}
\mathbb{E}\left[\left(\sum_{k=1}^{n}\left(\xi_k - \mathbb{E}\left[\xi_k\right]\right)\lambda_k^{(i)}\right)^2\right] &= \sum_{k,l=1}^{n} Cov\left(\xi_k, \xi_l\right)\lambda_k^{(i)}\lambda_l^{(i)} = \sum_{k,l=1}^{n} C_{k,l}\left(\xi\right)\lambda_k^{(i)}\lambda_l^{(i)} \\
&= \left(C\left(\xi\right)\lambda^{(i)}, \lambda^{(i)}\right) = 0 \ ,
\end{aligned} \tag{5.21}
$$

ovvero

$$
\sum_{k=1}^{n}\left(\xi_k - \mathbb{E}\left[\xi_k\right]\right)\lambda_k^{(i)} = 0 \ \mathbb{P} - q.c. \ , \ \forall i = 1, .., n - r \ , \tag{5.22}
$$

dunque il vettore aleatorio $\bar{\xi} = \left(\bar{\xi}_1, .., \bar{\xi}_n\right)$ tale che $\forall i = 1, .., n, \bar{\xi}_i := \xi_i - \mathbb{E}[\xi_i]$, non ha componenti linearmente indipendenti. In particolare, se il rango di $C(\xi)$ è $r < n, r$ componenti di $\bar{\xi}$ saranno linearmente indipendenti.

5.2.2 Vettori aleatori con distribuzione gaussiana

Definizione 252 *Un v.a.* $\xi = (\xi_1, .., \xi_n)$ *su* $(\Omega, \mathcal{F}, \mathbb{P})$ *è detto* gaussiano *o* normalmete distribuito, *se ha distribuzione assolutamente continua rispetto alla misura di Lebesgue su* $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ *di densità*

$$
f_\xi\left(x\right) := \sqrt{\frac{\det A}{\left(2\pi\right)^n}} e^{-\frac{1}{2}\left(A(x-m),(x-m)\right)} = \frac{e^{-\frac{1}{2}\left(A(x-m),(x-m)\right)}}{\sqrt{\left(2\pi\right)^n \det A^{-1}}} \ , \tag{5.23}
$$

dove:

1. $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *è una matrice simmetrica definita non negativa, ovvero tale che:*

   (a) $A^t = A$ *ovvero* $a_{i,j} = a_{j,i} \quad \forall i, j = 1, .., n$;

(b) $(Ax,x) \geq 0 \quad \forall x \in \mathbb{R}^n$;

2. $m \in \mathbb{R}^n$ tale che, $\forall i = 1,..,n$, $|m_i| < \infty$.

Osservazione 253 *Poiché la matrice A che compare nell'espressione della densità di probabilità di un vettore gaussiano è simmetrica, $\forall x, y \in \mathbb{R}^n$, $(Ax,y) = (x,Ay) = (Ay,x)$. Quindi,*

$$\begin{aligned}(A(x-m),(x-m)) &= (Ax,x) - (Ax,m) - (Am,x) + (Am,m) \\ &= (Ax,x) + (Am,m) - 2(Am,x) \ .\end{aligned} \tag{5.24}$$

Perciò, la densità di probabilità di ξ risulta

$$f_\xi(x) = \frac{e^{-\frac{1}{2}(Am,m)}}{\sqrt{(2\pi)^n \det A^{-1}}} e^{-\frac{1}{2}(Ax,x)+(Am,x)} \ . \tag{5.25}$$

Pertanto, se un v.a. η definito su $(\Omega, \mathcal{F}, \mathbb{P})$ a valori in $\mathbb{R}^n$ ha distribuzione a.c. rispetto alla misura di Lebesgue in $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ di densità

$$f_\eta(x) = Ke^{-\frac{1}{2}(Ax,x)+(b,x)} \ , \tag{5.26}$$

con K costante di normalizzazione, tale che:

1. $A : \mathbb{R}^n \longmapsto \mathbb{R}^n$ *è una matrice simmetrica definita non negativa;*
2. $b \in \mathbb{R}^n$;

   allora, η è gaussiano.

Osservazione 254 *Siccome una matrice simmetrica è diagonalizzabile, esiste una matrice O ortogonale, ovvero tale che $O^t = O^{-1}$, tale che $O^tAO = D := diag[a_1,..,a_n]$ dove, $\forall i = 1,..,n$, a_i è un autovalore di A. Nel caso in cui A sia definita non negativa, i suoi autovalori sono non negativi ed in generale possono anche avere molteplicità algebrica maggiore di uno. Restano allora definite le matrici $\sqrt{D} := diag[\sqrt{a_1},..,\sqrt{a_n}]$ e $B := \sqrt{D}O^t$ per cui $A = B^tB$. Inoltre, poiché in generale alcuni autovalori di A possono essere nulli, ovvero $KerA$ è isometricamente isomorfo a $\mathbb{R}^k$, $0 \leq k \leq n$, $B : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-k}$, il che è equivalente ad affermare che, indicando con Π_k il proiettore ortoganale su $KerA$ e con $I_n : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ la matrice identità, se $\Pi_{n-k} := I_n - \Pi_k$,*

$$A = (\Pi_k + \Pi_{n-k}) A (\Pi_k + \Pi_{n-k}) = \Pi_{n-k}A\Pi_{n-k} = \Pi_{n-k}B^tB\Pi_{n-k} = \Pi_{n-k}ODO^t\Pi_{n-k} \ . \tag{5.27}$$

In tal caso, la densità di probabilità di un vettore gaussiano ξ associata alla matrice A può essere definita come limite in distribuzione della densità di probabilità di un vettore gaussiano $\xi^{(\varepsilon)}$, la cui densità di probabilità è identica a quella di ξ salvo che la matrice associata è $A^{(\varepsilon)} := A + \varepsilon^{-1}\Pi_k$, nel limite di $\varepsilon \downarrow 0$. Più precisamente, posto

$$\begin{aligned} y &:= O^t(x-m) = O^t((\Pi_{n-k} + \Pi_k)(x-m)) \\ &= O^t\Pi_{n-k}(x-m) + O^t\Pi_k(x-m) = u + v, \end{aligned} \tag{5.28}$$

$$\Pi_{n-k}x = Ou + \Pi_{n-k}m, \quad \Pi_k x = Ov + \Pi_k m, \quad x = Oy + m = Ou + Ov + m \tag{5.29}$$

si ha

$$\begin{aligned}\left(A^{(\varepsilon)}(x-m),(x-m)\right) &= \left(\left(\Pi_{n-k}A\Pi_{n-k}+\varepsilon^{-1}\Pi_k\right)(x-m),(x-m)\right) \\ &= \left(\Pi_{n-k}A\Pi_{n-k}(x-m),(x-m)\right)+\left(\Pi_k\varepsilon^{-1}I_n\Pi_k(x-m),(x-m)\right) \\ &= \left(\Pi_{n-k}ODO^t\Pi_{n-k}(x-m),(x-m)\right)+\left(\Pi_k\varepsilon^{-1}OI_nO^t\Pi_k(x-m),(x-m)\right) \\ &= ((D)u,u)+\left(\varepsilon^{-1}I_nv,v\right)\end{aligned} \tag{5.30}$$

Inoltre, poiché O *è ortogonale* $|\det O|=1$ *e* $\det A=\det D$,

$$\begin{aligned}\det\left(A^{(\varepsilon)}\right) &= \det\left(\Pi_{n-k}A\Pi_{n-k}+\varepsilon^{-1}\Pi_k\right)=\det\left(\Pi_{n-k}A\Pi_{n-k}\right)\det\left(\Pi_k\varepsilon^{-1}I_n\Pi_k\right) \\ &= \frac{1}{\varepsilon^k}\det D=\frac{1}{\varepsilon^k}\prod_{i=1}^{n-k}a_i\end{aligned} \tag{5.31}$$

allora, $\forall\phi\in C_b(\mathbb{R}^n,\mathbb{R})$,

$$\lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^n}dxf_{\xi^{(\varepsilon)}}(x)\phi(x)=\lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^n}dx\frac{e^{-\frac{1}{2}\left(A^{(\varepsilon)}(x-m),(x-m)\right)}}{\sqrt{(2\pi)^n\det\left(A^{(\varepsilon)}\right)^{-1}}}\phi(x)= \tag{5.32}$$

$$\lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^{n-k}}du\left(\prod_{i=1}^{n-k}\sqrt{\frac{a_i}{2\pi}}e^{-\frac{1}{2}a_iu_i^2}\right)\int_{\mathbb{R}^k}dv\left(\prod_{i=1}^{k}\frac{1}{\sqrt{2\pi\varepsilon}}e^{-\frac{v_i^2}{2\varepsilon}}\right)\phi(Ou+Ov+m)$$

$$=\int_{\mathbb{R}^{n-k}}du\left(\prod_{i=1}^{n-k}\sqrt{\frac{a_i}{2\pi}}e^{-\frac{1}{2}a_iu_i^2}\right)\phi(Ou+m)\ .$$

Proposizione 255 *Sia* ξ *un v.a. gaussiano definito su* $(\Omega,\mathcal{F},\mathbb{P})$. *Allora, se la densità di probabilità di* ξ *è*

$$f_\xi(x)=\frac{e^{-\frac{1}{2}(A(x-m),(x-m))}}{\sqrt{(2\pi)^n\det A^{-1}}}\ , \tag{5.33}$$

la sua funzione caratteristica risulta

$$\varphi_\xi(t):=\mathbb{E}\left(e^{i(t,\xi)}\right)=e^{i(t,m)-\frac{1}{2}\left(A^{-1}t,t\right)}\ . \tag{5.34}$$

Dimostrazione: Nel caso in cui A è definita positiva, si ha:

$$\int_{\mathbb{R}^n}dx\frac{e^{-\frac{1}{2}(A(x-m),(x-m))}}{\sqrt{(2\pi)^n\det A^{-1}}}e^{i(t,x)}=\int_{\mathbb{R}^n}dy\frac{e^{-\frac{1}{2}(Dy,y)}}{\sqrt{(2\pi)^n(\det D)^{-1}}}e^{i(t,Oy+m)}$$

$$e^{i(t,m)}\int_{\mathbb{R}^n}dy\frac{e^{-\frac{1}{2}(Dy,y)}}{\sqrt{(2\pi)^n(\det D)^{-1}}}e^{i\left(O^tt,y\right)}=e^{i(t,m)}\prod_{i=1}^{n}\sqrt{\frac{a_i}{2\pi}}\int_{\mathbb{R}}dy_ie^{-\frac{a_iy_i^2}{2}+i\left(O^tt\right)_iy_i}$$

$$=e^{i(t,m)}\prod_{i=1}^{n}e^{-\frac{\left(O^tt\right)_i^2}{2a_i}}=e^{i(t,m)}e^{-\frac{1}{2}\left(D^{-1}O^tt,O^tt\right)}=e^{i(t,m)-\frac{1}{2}\left(OD^{-1}O^tt,t\right)}=e^{i(t,m)-\frac{1}{2}\left(A^{-1}t,t\right)}\ .$$

Se A possiede un nucleo non banale, dall'osservazione precedente, ponendo

$$\phi(Ou+m)=e^{i(t,Ou+m)}=e^{i(t,m)+i(t,Ou)}\,, \tag{5.35}$$

che è ovviamente una funzione limitata su $\mathbb{R}^n$, si ottiene

$$\begin{aligned}\int_{\mathbb{R}^{n-k}} du\left(\prod_{i=1}^{n-k}\sqrt{\frac{a_i}{2\pi}}e^{-\frac{1}{2}a_iu_i^2}\right)\phi(Ou+m) &= e^{i(t,m)}\prod_{i=1}^{n-k}\sqrt{\frac{a_i}{2\pi}}\int_{\mathbb{R}}du_ie^{-\frac{a_iu_i^2}{2}+i\left(O^tt\right)_iu_i} \\ &= e^{i(t,m)-\frac{1}{2}\left(A^{-1}t,t\right)}\,.\end{aligned} \tag{5.36}$$

■

Osservazione 256 *Poiché per il teorema di unicità ogni distribuzione di probabilità risulta univocamente determinata dalla sua funzione caratteristica, è possibile definire che un vettore aleatorio n-dimensionale, e dunque anche una variabile aleatoria, è gaussiano se la sua funzione caratteristica $\varphi(t)=e^{i(t,m)-\frac{1}{2}(Ct,t)}$, dove $m\in\mathbb{R}^n$ è tale che $\forall i=1,..,n$, $|m_i|<\infty$ e $C:\mathbb{R}^n\longrightarrow\mathbb{R}^n$, è una matrice simmetrica definita non negativa.*

Inoltre, ciò permette di definire la misura di probabilità associata ad un vettore gaussiano quando C abbia nucleo non banale. Infatti, ponendo $C^{(\varepsilon)}=C+\varepsilon I_n$ si ha che $\forall t\in\mathbb{R}^n$,

$$\lim_{\varepsilon\downarrow 0}\varphi^{(\varepsilon)}(t)=\lim_{\varepsilon\downarrow 0}e^{i(t,m)-\frac{1}{2}\left(C^{(\varepsilon)}t,t\right)}=\varphi(t)\ . \tag{5.37}$$

Proposizione 257 *Sia ξ un v.a. gaussiano definito su $(\Omega,\mathcal{F},\mathbb{P})$. Allora, se la densità di probabilità di ξ è*

$$f_\xi(x)=\frac{e^{-\frac{1}{2}(A(x-m),(x-m))}}{\sqrt{(2\pi)^n\det A^{-1}}}=\frac{e^{-\frac{1}{2}\left(C^{-1}(x-m),(x-m)\right)}}{\sqrt{(2\pi)^n\det C}}\,, \tag{5.38}$$

$$\mathbb{E}(\xi)=m\in\mathbb{R}^n\,,\quad \mathbb{E}\left((\xi-m)_i(\xi-m)_j\right)=C_{i,j}=\left(A^{-1}\right)_{i,j}\quad i,j=1,..,n\,. \tag{5.39}$$

Pertanto m è il vettore corrispondente ai valori attesi delle componenti di ξ e C è la matrice di varanza-covarianza.

Dimostrazione: Nel caso in cui C, e dunque A, è definita positiva, $\forall i,j=1,..,n$, si ha:

$$\begin{aligned}\mathbb{E}(\xi_i) &= \int_{\mathbb{R}^n}dx\frac{e^{-\frac{1}{2}\left(C^{-1}(x-m),(x-m)\right)}}{\sqrt{(2\pi)^n\det C}}x_i=\int_{\mathbb{R}^n}dy\frac{e^{-\frac{1}{2}(Dy,y)}}{\sqrt{(2\pi)^n(\det D)^{-1}}}(Oy+m)_i \\ &= \prod_{i=1}^n\sqrt{\frac{a_i}{2\pi}}\int_{\mathbb{R}}dy_ie^{-\frac{a_iy_i^2}{2}}(Oy)_i+m_i\,.\end{aligned} \tag{5.40}$$

Ma, essendo $(Oy)_i$ una funzione dispari di y_i, $\forall i=1,..,n$,

$$\int_{\mathbb{R}}dy_ie^{-\frac{a_iy_i^2}{2}}(Oy)_i=0\,, \tag{5.41}$$

percui $\mathbb{E}(\xi_i) = m_i$.

D'altronde,

$$\mathbb{E}\left((\xi - m)_i (\xi - m)_j\right) = \int_{\mathbb{R}^n} dx \frac{e^{-\frac{1}{2}\left(C^{-1}(x-m),(x-m)\right)}}{\sqrt{(2\pi)^n \det C}} (x-m)_i (x-m)_j \tag{5.42}$$
$$= \int_{\mathbb{R}^n} dy \frac{e^{-\frac{1}{2}(Dy,y)}}{\sqrt{(2\pi)^n (\det D)^{-1}}} (Oy)_i (Oy)_j$$
$$= \left(\prod_{k=1}^n \sqrt{\frac{a_k}{2\pi}}\right) \int_{\mathbb{R}^n} dy e^{-\sum_{k=1}^n \frac{a_k y_k^2}{2}} \sum_{m,l=1}^n O_{i,m} y_m O_{j,l} y_l \,.$$

Ma, poiché se $l \neq m$, $O_{i,m} y_m O_{j,l} y_l$ è una funzione dispari di y_l e y_m

$$\int_{\mathbb{R}} dy_l e^{-\frac{a_l y_l^2}{2}} \int_{\mathbb{R}} dy_m e^{-\frac{a_k y_m^2}{2}} O_{i,m} y_m O_{j,l} y_l = 0 \,. \tag{5.43}$$

Perciò,

$$\sum_{m,l=1}^n \int_{\mathbb{R}^n} dy e^{-\sum_{k=1}^n \frac{a_k y_k^2}{2}} O_{i,m} y_m O_{j,l} y_l = \sum_{m,l=1}^n \int_{\mathbb{R}^n} dy e^{-\sum_{k=1}^n \frac{a_k y_k^2}{2}} O_{i,m} y_m O_{j,l} y_l \delta_{l,m} \tag{5.44}$$
$$= \sum_{l=1}^n \int_{\mathbb{R}^n} dy e^{-\sum_{k=1}^n \frac{a_k y_k^2}{2}} O_{i,l} O_{j,l} y_l^2 = \left(\prod_{k=1}^n \sqrt{\frac{2\pi}{a_k}}\right) \sum_{l=1}^n \frac{O_{i,l} O_{j,l}}{a_l}$$
$$= \left(\prod_{k=1}^n \sqrt{\frac{2\pi}{a_k}}\right) \sum_{l=1}^n \frac{O_{i,l} O_{l,j}^t}{a_l} = \left(\prod_{k=1}^n \sqrt{\frac{2\pi}{a_k}}\right) \left(O D^{-1} O^t\right)_{i,j} \,.$$

Quindi,

$$\mathbb{E}\left((\xi - m)_i (\xi - m)_j\right) = \left(O D^{-1} O^t\right)_{i,j} = \left(A^{-1}\right)_{i,j} = C_{i,j} \,. \tag{5.45}$$

Nel caso in cui C abbia un nucleo non banale la tesi segue ponendo nella (5.32) $\forall i, j = 1, .., n$, $\phi(x) = x_i$ e $\phi(x) = (x-m)_i (x-m)_j$. ■

Osservazione 258 *Sia ξ un v.a. gaussiano definito su $(\Omega, \mathcal{F}, \mathbb{P})$ e $\varphi_\xi(t) = e^{i(t,m) - \frac{1}{2}(Ct,t)}$ la sua funzione caratteristica, analogamente al caso delle variabili aleatorie, il valore atteso della componente i-sima e il generico elemento di matrice della matrice di varanza-covarianza potevano ottenersi derivando ponendo $t = 0$ nelle espressioni di $\frac{\partial}{\partial t_i} \varphi_\xi(t)$ e $\frac{\partial^2}{\partial t_i \partial t_j} \varphi_\xi(t)$.*

Teorema 259
1. *Le componenti di un vettore aleatorio gaussiano sono scorrelate se e solo se sono variabili aleatorie indipendenti.*
2. *Un vettore aleatorio $\xi = (\xi_1, .., \xi_n) \in \mathbb{R}^n$ definito su $(\Omega, \mathcal{F}, \mathbb{P})$ è gaussiano se e solo se $\forall \lambda = (\lambda_1, .., \lambda_n) \in \mathbb{R}^n$ la variabile aleatoria (λ, ξ) è gaussiana.*

Dimostrazione:

1. Se le componenti di un vettore aleatorio gaussiano sono scorrelate allora, poiché la sua matrice di varianza-covarianza risulta essere diagonale, la funzione caratteristica di tale vettore si scrive come prodotto delle funzioni caratteristiche delle sue componenti e pertanto queste sono stocasticamente indipendenti. D'altra parte se le componenti di un vettore aleatorio sono indipendenti, sono anche scorrelate.

2. Se ξ è un v.a. gaussiano, di valore atteso m e matrice di varianza-covarianza C, $\forall\lambda \in \mathbb{R}^n$, $t \in \mathbb{R}$,
$$\varphi_{(\lambda,\xi)}(t) = \mathbb{E}\left(e^{it(\lambda,\xi)}\right) = \mathbb{E}\left(e^{i(t\lambda,\xi)}\right) = \varphi_\xi(t\lambda) = e^{i(t\lambda,m)-\frac{1}{2}(Ct\lambda,t\lambda)} = e^{it(\lambda,m)-\frac{t^2}{2}(C\lambda,\lambda)}, \quad (5.46)$$
che è la funzione caratteristica di una v.a. gaussiana di valore atteso (λ, m) e varianza $(C\lambda,\lambda)$. D'altra parte, se, $\forall\lambda \in \mathbb{R}^n$, (λ,ξ) è una v.a. gaussiana posto $t = 1$ nella pecedente espressione della sua funzione caratteristica si ha che
$$\varphi_{(\lambda,\xi)}(1) = \mathbb{E}\left(e^{i(\lambda,\xi)}\right) = e^{i(\lambda,m)-\frac{1}{2}(C\lambda,\lambda)} = \varphi_\xi(\lambda) . \quad (5.47)$$
Poiché ciò è vero $\forall\lambda \in \mathbb{R}^n$, ξ è un v.a. gaussiano.

■

Proposizione 260 *Una matrice $C : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ è la matrice di varianza-covarianza di un vettore aleatorio $\xi := (\xi_1, .., \xi_n) \in \mathbb{R}^n$ definito su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$, tale che, $\forall i = 1, .., n, \xi_i \in L^2(\mathcal{F})$, se e solo se C è simmetrica e definita non negativa.*

Dimostrazione:

$\Longrightarrow$ Segue dalla sua definizione che la matrice di varianza-covarianza $C(\xi)$ del vettore aleatorio $\xi := (\xi_1, .., \xi_n)$, è simmetrica e definita non negativa.

$\Longleftarrow$ Se C è una matrice simmetrica esiste una matrice ortogonale O tale che $C = ODO^t$ dove D è diagonale. Inoltre, se C è definita non negativa tutti gli elementi di D sono non negativi. Perciò, ponendo $\sqrt{D} := diag\left[\sqrt{c_1}, .., \sqrt{c_n}\right]$, dove i numeri c_i sono gli autovalori di C, e $B := O\sqrt{D}$, si ha $C = BB^t$. Quindi, se $\eta := (\eta_1, .., \eta_n)$ è un vettore aleatorio le cui componenti sono tra loro indipendenti e normali standard (ovvero $\forall i = 1, .., n, \eta_i \overset{d}{=} N(0, 1)$), posto $\xi := B\eta$, vettore aleatorio a valori in $\mathbb{R}^d$, si ha che $\xi = (\xi_1, .., \xi_n)$ ha componenti gaussiane, poiché $\forall i = 1, ..n, \xi_i = \sum_{j=1}^n B_{i,j}\eta_j$ è gaussiano per il Teorema 259. Inoltre, $\forall i, j = 1, ..n$,
$$\begin{aligned}
\mathbb{E}\left[\xi_i\xi_j\right] &= \mathbb{E}\left[(B\eta)_i (B\eta)_j\right] = \quad (5.48)\\
&= \mathbb{E}\left[\sum_{k,l=1}^n B_{i,k}B_{j,l}\eta_k\eta_l\right] = \sum_{k,l=1}^n B_{i,k}B_{j,l}\mathbb{E}\left[\eta_k\eta_l\right]\\
&= \sum_{k,l=1}^n B_{i,k}B_{j,l}\delta_{k,l} = \sum_{k=1}^n B_{i,k}B_{j,k} = \left(BB^t\right)_{i,j}
\end{aligned}$$

e

$$\mathbb{E}[\xi_i] = \sum_{j=1}^{n} B_{i,j}\mathbb{E}[\eta_j] = 0 \ , \tag{5.49}$$

si ottiene che la matrice di varianza-covarianza di ξ è proprio BB^t, cioè C.

■

Corollario 261 *Un vettore aleatorio $\xi := (\xi_1, .., \xi_n) \in \mathbb{R}^n$ definito su uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ a componenti in $L^2(\mathcal{F})$ linearmente indipendenti è gaussiano se e solo se esistono n variabili aletorie indipendenti identicamente distribuite $\eta_1, .., \eta_n$ tali che $\eta_1 \stackrel{d}{=} N(0,1)$ e una matrice B non singolare di ordine n tale che $\forall i = 1, .., n, \xi_i - \mathbb{E}[\xi_i] = \sum_{j=1}^n B_{i,j}\eta_j$ e $C(\xi) = BB^t$.*

Dimostrazione:

$\Longrightarrow$ Se O è la matrice ortogonale che diagonalizza $C(\xi)$ e $\sqrt{D} := diag\left[\sqrt{c_1}, .., \sqrt{c_n}\right]$ dove $c_1, .., c_n$ sono gli autovalori di $C(\xi)$, posto $B := O\sqrt{D}$, il vettore aleatorio $\eta := B^{-1}(\xi - \mathbb{E}[\xi])$ ha componenti non correlate di varianza pari a 1, come segue dalla (5.20). Inoltre, per il Teorema 259 η è gaussiano, pertanto le componenti di η sono indipendenti, e siccome $\mathbb{E}[\eta] = 0$, ognuna di queste è normale standard.

$\Longleftarrow$ Sia $\eta := (\eta_1, .., \eta_n)$ e $B : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ tale che $C(\xi) = BB^t$. Posto $\bar{\xi} := B\eta$, per la (5.48) $C(\bar{\xi}) = C(\xi)$, quindi la differenza tra i vettori aleatori ξ e $\bar{\xi}$ è un vettore aleatorio degenere; ma poiché $\mathbb{E}[\bar{\xi}] = 0$,

$$\xi = \bar{\xi} + \mathbb{E}[\xi] \ , \ \mathbb{P} - q.c.. \tag{5.50}$$

■

Osservazione 262 *Sia $\xi = (\xi_1, .., \xi_n)$ un vettore aleatorio gaussiano definito sullo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ a componenti linearmente indipendenti e ζ una v.a. gaussiana. Ragionando come nell'Osservazione 90 miglior stimatore in media quadratica di ζ, date le osservazioni del vettore aleatorio ξ, è $\mathbb{E}[\zeta|\xi]$ poiché questa v.a. realizza il minimo di $\inf_{\theta \in L^2(\mathcal{F}^\xi)} \|\zeta - \theta\|^2_{L^2}$, ma $\mathbb{E}[\zeta|\xi] = \mathbb{E}[\zeta|\eta]$, dove η è il vettore aleatorio $\eta := B^{-1}(\xi - \mathbb{E}[\xi])$ costruito nel precedente corollario che ha componenti normali standard fra loro indipendenti. Calcolando la densità di probabilità del vettore aleatorio $(\zeta, \eta) \in \mathbb{R}^{n+1}$,*

$$f_{(\zeta,\eta)}(y, x_1, ...x_n) = \frac{\exp\left[-\frac{\left(C^{-1}(\zeta,\eta)(y-\mathbb{E}[\zeta],x_1,..,x_n),(y-\mathbb{E}[\zeta],x_1,..,x_n)\right)}{2}\right]}{\sqrt{(2\pi)^{n+1}\det C(\zeta,\eta)}} \ , \tag{5.51}$$

dove $C_{11}(\zeta, \eta) = Var(\zeta), \forall i, j = 2, .., n+1, C_{1,j}(\zeta, \eta) = Cov(\zeta, \eta_{j-1}), C_{i,j}(\zeta, \eta) = \delta_{i,j}$, se $n(x)$ denota la densità di una v.a. $N(0,1)$, si ottiene

$$f_{\zeta|\eta}(y|x_1, ..x_n) = \frac{f_{(\zeta,\eta)}(y, x_1, ...x_n)}{f_\eta(x_1, ...x_n)} = \frac{f_{(\zeta,\eta)}(y, x_1, ...x_n)}{\prod_{i=1}^{n} n(x_i)} \ , \tag{5.52}$$

perciò

$$\mathbb{E}[\zeta|\eta] = \int_{\mathbb{R}} dy y f_{\zeta|\eta}(y|x_1,..x_n) = \mathbb{E}[\zeta] + \sum_{i=1}^{n} Cov(\zeta,\eta_i)\eta_i \ , \tag{5.53}$$

il che implica che $\mathbb{E}[\zeta|\eta]$ *è una funzione lineare di* η*, ovvero appartiene a span* $(\eta_1,..,\eta_n)$ *, ma poiché per costruzione le variabili aleatorie* $\eta_1,..,\eta_n$ *formano una base ortonormale per span* $(\xi_1,..,\xi_n)$ *, ciò dimostra che, dato* $\xi = (\xi_1,..,\xi_n)$ *un vettore aleatorio gaussiano definito sullo spazio di probabilità* $(\Omega,\mathcal{F},\mathbb{P})$ *a componenti linearmente indipendenti e* ζ *una v.a. gaussiana,* $\mathbb{E}[\zeta|\xi] \in$ *span* $(\xi_1,..,\xi_n)$ *.*

Questa affermazione vale anche se al posto di una singola v.a. ζ *gaussiana si considera un vettore aleatorio gaussiano* $\zeta = (\zeta_1,..,\zeta_m)$*; in tal caso, se* $B : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *è tale che* $C(\xi) = BB^t, \forall j = 1,..,m,$

$$\begin{aligned}\mathbb{E}\left[\zeta_j|\xi\right] &= \mathbb{E}\left[\zeta_j\right] + \sum_{i=1}^{n} Cov\left(\zeta_j,\left(B^{-1}\left(\xi-\mathbb{E}[\xi]\right)\right)_i\right)\left(B^{-1}\left(\xi-\mathbb{E}[\xi]\right)\right)_i \\ &= \mathbb{E}\left[\zeta_j\right] + \sum_{i=1}^{n}\sum_{k=1}^{m} Cov\left(\zeta_j,\xi_k\right)C^{-1}_{k,i}(\xi)\left(\xi_i-\mathbb{E}[\xi_i]\right) \ .\end{aligned} \tag{5.54}$$

Proposizione 263 *Sia* $\left\{\xi^{(n)}\right\}_{n\geq 1}$ *una successione di v.a. gaussiani definiti su* $(\Omega,\mathcal{F},\mathbb{P})$ *convergente in probabilità al v.a.* ξ*. Allora, anche* ξ *è un v.a. gaussiano.*

Dimostrazione: Siano $m^{(n)} := \mathbb{E}\left(\xi^{(n)}\right)$ e $C^{(n)}$ la matrice di varianza-covarianza di $\xi^{(n)}$. Allora, poiché $\left\{\xi^{(n)}\right\}_{n\geq 1}$ converge in probabilità, $\forall\lambda\in\mathbb{R}^n$, anche la successione di v.a. $\left\{\left(\xi^{(n)},\lambda\right)\right\}_{n\geq 1}$ converge in probabilità, quindi in distribuzione, ovvero $\forall t\in\mathbb{R}$, esiste

$$\lim_{n\to\infty}\varphi_{\left(\lambda,\xi^{(n)}\right)}(t) = \lim_{n\to\infty}\varphi_{\xi^{(n)}}(t\lambda) = \lim_{n\to\infty} e^{it\left(\lambda,m^{(n)}\right)-\frac{t^2}{2}\left(C^{(n)}\lambda,\lambda\right)} = \varphi_{(\lambda,\xi)}(t) \ . \tag{5.55}$$

Dunque posto $\mu_\lambda := \mathbb{E}\left((\xi,\lambda)\right)$ e $\sigma^2_\lambda := \mathbb{E}\left(\left((\xi,\lambda)-\mathbb{E}\left((\xi,\lambda)\right)\right)^2\right)$, si ha che

$$\lim_{n\to\infty}\mathbb{E}\left(\left(\xi^{(n)},\lambda\right)\right) = \mu_\lambda\,, \quad \lim_{n\to\infty}\mathbb{E}\left(\left((\xi,\lambda)-\mathbb{E}\left((\xi,\lambda)\right)\right)^2\right) = \sigma^2_\lambda\,. \tag{5.56}$$

Pertanto,

$$\lim_{n\to\infty} e^{it\left(\lambda,m^{(n)}\right)-\frac{t^2}{2}\left(C^{(n)}\lambda,\lambda\right)} = e^{it\mu_\lambda-\frac{t^2}{2}\sigma^2_\lambda} \tag{5.57}$$

ovvero, $\forall\lambda\in\mathbb{R}^n$, (ξ,λ) è una v.a. gaussiana e dunque, per il Teorema 259 ξ è un vettore aleatorio gaussiano. ■

Osservazione 264 *Riassumendo quanto finora esposto, dall'ossevazione precedente segue che una collezione di v.a. gaussiani forma uno spazio lineare, in quanto ogni sua combinazione lineare è ancora un v.a. gaussiano. La chiusura in media quadratica di questa verietà lineare è ancora composta da v.a. gaussiani.*

5.3 Sistemi gaussiani

Definizione 265 *Sia $\mathcal{I}$ un insieme di indici. Una collezione di v.a. definite su $(\Omega, \mathcal{F}, \mathbb{P})$, $\{\xi_i\}_{i\in\mathcal{I}}$ è detto* sistema gaussiano *se, $\forall n \geq 1$ e per ogni sottoinsieme di indici della forma $\{i_k\}_{k=1}^n \subset \mathcal{I}$, il vettore aleatorio $\xi^{(n)} = \left(\xi_{i_1}, .., \xi_{i_n}\right) \in \mathbb{R}^n$ è gaussiano.*

Un sistema gaussiano, possiede le seguenti proprietà:

1. Se $\{\xi_i\}_{i\in\mathcal{I}}$ è un sistema gaussiano definito su $(\Omega, \mathcal{F}, \mathbb{P})$, ogni collezione $\{\xi_i\}_{i\in\mathcal{I}'}$, con $\mathcal{I}' \subseteq \mathcal{I}$ è ancora un sistema gaussiano;

2. Se $\{\xi_i\}_{i\in\mathcal{I}}$ è un sistema gaussiano definito su $(\Omega, \mathcal{F}, \mathbb{P})$, $\forall n \geq 1$ e per ogni collezione $\{\xi_{i_k}\}_{k=1}^n$ di v.a. gaussiane tali che $\{i_k\}_{k=1}^n \subset \mathcal{I}$, la chiusura in media quadratica della varietà lineare generata da $\{\xi_{i_k}\}_{k=1}^n$ è un sistema gaussiano.

Osservazione 266 *Se ξ_1 e η_1 sono v.a. gaussiane definite su $(\Omega, \mathcal{F}, \mathbb{P})$, il vettore aleatorio $X = (\xi, \eta)$ tale che $\xi := \xi_1$ e $\eta := |\eta_1| \mathbf{1}_{[0,+\infty)}(\xi_1) - |\eta_1| \mathbf{1}_{(-\infty,0)}(\xi_1)$ non è gaussiano. Quindi l'affermazione contraria alla 1 è falsa.*

Proposizione 267 *Se $\mathcal{I}$ è un insieme di indici, $m = \{m_i\}_{i\in\mathcal{I}} \in \mathbb{R}^{\mathcal{I}}$ e $\{C_{i,j}\}_{i,j\in\mathcal{I}}$ una collezione di numeri reali tali che: $\forall i, j \in \mathcal{I}$, $C_{i,j} = C_{j,i}$ e $\forall x \in \mathbb{R}^{\mathcal{I}}$, $\sum_{i,j\in\mathcal{I}} x_j C_{j,i} x_i$, allora esiste uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ ed un sistema gaussiano $\{\xi_i\}_{i\in\mathcal{I}}$ su di esso definito, tale che $\forall i, j \in \mathcal{I}$, $\mathbb{E}(\xi_i) = m_i$ e $\mathbb{E}\left(\xi_i \xi_j\right) = C_{i,j}$.*

Dimostrazione: Sia $\mathcal{J}$ l'insieme delle sottocollezioni di indici di $\mathcal{I}$ della forma $I^{(n)} := \{i_k\}_{k=1}^n \subset \mathcal{I}$, $\forall n \geq 1$. Sia allora $\mathbf{P}_{I^{(n)}}$ la distribuzione di probabilità su $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ associata alla funzione caratteristica

$$\mathbb{R}^n \ni t \longmapsto \varphi_{I^{(n)}}(t) = e^{i\left(t, m^{(n)}\right) - \frac{1}{2}\left(C^{(n)} t, t\right)} \in \mathbb{C} \tag{5.58}$$

del v.a. gaussiano $\xi^{(n)} = \left(\xi_{i_1}, .., \xi_{i_n}\right)$ di valore atteso $m^{(n)} = \{m_i\}_{i\in I_n}$ e matrice di varianza-covarianza $C = \{C_{i,j}\}_{i,j\in I^{(n)}}$. La famiglia $\{\mathbf{P}_{I^{(n)}}\}_{I^{(n)}\subset\mathcal{J}}$ verifica le ipotesi del teorema di Kolmogorov di estensione di una misura su $\left(\mathbb{R}^{\mathcal{I}}, \mathcal{B}\left(\mathbb{R}^{\mathcal{I}}\right)\right)$, pertanto se $\mathbb{P}$ è tale estensione, $(\Omega, \mathcal{F}, \mathbb{P}) = \left(\mathbb{R}^{\mathcal{I}}, \mathcal{B}\left(\mathbb{R}^{\mathcal{I}}\right), \mathbb{P}\right)$. ■