Università degli Studi di Roma "La Sapienza"
Anno Accademico 2009-2010

Facoltà di Scienze Matematiche Fisiche e Naturali
Corso di Laurea Specialistica in Matematica per le Applicazioni

Alcuni appunti per il corso di

CALCOLO DELLE PROBABILITÀ 3

Giovanna Nappo

A.A. 2009/2010

versione del 1.11.2009

Indice

Introduzione **ii**

Programma provvisorio del corso **iii**

1 Richiami su spazi di probabilità **1**
- 1.1 Esempi di spazi di probabilità . . . 1
- 1.2 Variabili aleatorie . . . 4
- 1.3 Distribuzioni di variabili aleatorie . . . 6
- 1.4 Valori attesi . . . 10
  - 1.4.1 Variabili aleatorie in spazi misurabili . . . 12
- 1.5 Misura indotta e Cambio di variabile . . . 12
- 1.6 Variabili gaussiane . . . 16

2 Costruzione di variabili aleatori in $(0, 1)$ **20**
- 2.1 Teorema di rappresentazione di Skorohod . . . 20
- 2.2 Costruzione di una succesione di variabili aleatorie indipendenti . . . 23
- 2.3 Convergenza per variabili aleatorie . . . 25

Bibliografia **28**

3 ESERCIZI PROPOSTI **29**
- 3.1 Esercizi di tipo analitico . . . 29
- 3.2 Esercizi sulla convergenza in distribuzione . . . 30
- 3.3 Esercizi con le funzioni caratteristiche . . . 30
- 3.4 Esercizi sulla condizione di Lindeberg . . . 32
- 3.5 Esercizi sulla legge dei grandi numeri . . . 33
- 3.6 Esercizio riassuntivo . . . 33

Introduzione

Lo scopo di questo corso è quello di riesaminare con gli strumenti più sofisticati (come ad esempio la teoria della misura) alcuni dei concetti principali del Calcolo delle Probabilità, già affrontati nei corsi di Calcolo delle Probabilità 1 e 2.

Argomento principale saranno i vari tipi di convergenza per variabili aletaorie e le generalizzazioni della Legge dei Grandi Numeri e del Teorema Centrale del Limite.

Come è noto, la Legge (debole) dei Grandi Numeri ed il Teorema Centrale del Limite hanno come soggetto principale una successione $\{X_n\}_{n\geq 1}$ di variabili aleatorie indipendenti, con funzione di distribuzione data[1].

Nei corsi elementari abbiamo dato per scontato che una tale successione di variabili aleatorie esistesse: in questo corso vedremo come questo fatto sia vero. Il problema va diviso in due parti:

• **primo problema:** data una funzione di ripartizione F (ossia una funzione a valori in $[0, 1]$, non decrescente, continua da destra e normalizzata) esiste sempre uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e una variabile aleatoria X, tale che $F_X(x) = F(x)$, dove $F_X(x) := \mathbb{P}(X \leq x)$?

• **secondo problema:** data una successione di funzioni di distribuzione $\{F_n\}_{n\geq 1}$ esiste uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e una successione di variabili aleatorie $\{X\}_{n\geq 1}$, indipendenti e tali che $F_{X_n}(x) = F_n(x)$?

Ci sono due risposte possibili a questa domanda, una dovuta a Kolmogorov e una dovuta a Skorohod: noi accenneremo alla risposta di Kolmogorov, mentre vedremo la risposta di Skorohod in cui lo spazio di probabilità è semplicemente $\big((0, 1), \mathcal{B}(0, 1), \lambda|_{(0,1)}\big)$, dove $\mathcal{B}(0, 1)$ è la sigma-algebra dei boreliani e $\lambda|_{(0,1)}$ è la misura di Lebesgue ristretta all'intervallo $(0, 1)$.

Questi appunti (parzialmente basati su appunti scritti per altri corsi) non sono in nessun modo completi. Le lezioni sono basate principalmente sui testi di Billingsley [1] e di Koch [2].

ATTENZIONE: le notazioni potrebbero differire da quelle usate a lezione.

GLI APPUNTI NON SONO ANCORA COMPLETATI(ad esempio, nella sezione sui vari tipi di convergenza, mancano le definizioni), E NON SONO STATI ANCORA CORRETTI.

[1]Nel caso più semplice la funzione di distribuzione F_{X_n} è addirittura sempre la stessa, ma, come vedremo questa condizione non è necessaria.

PROGRAMMA PROVVISORIO DEL CORSO

CALCOLO DELLE PROBABILITÀ 3 I Semestre , A. A. 2009/10

docente: Giovanna Nappo,

(ufficio n.108, tel. 49913262, e-mail: nappo@mat.uniroma1.it)

Prerequisiti: Nozioni di base di Probabilità, acquisibili attraverso i corsi di Calcolo delle Probabilità 1 e 2. In particolare si presuppone che lo studente conosca le nozioni di distribuzione congiunta e le distribuzioni classiche. È inoltre consigliato avere familiarità con i concetti di base di teoria della misura (tali nozioni sono acquisibili nel corso di Analisi Reale).

Obiettivi: Studio delle relazioni tra teoria della misura (misure finite) e modelli probabilistici. Studio delle possibili descrizioni e costruzioni per variabili aleatorie. Studio di alcuni tipi di convergenza per variabili aleatorie (quasi certa, in probabilità e in distribuzione). Acquisizione delle tecniche fondamentali di convergenza e dei risultati fondamentali del Calcolo delle Probabilità (Legge dei Grandi Numeri e Teorema Centrale del Limite)

TESTI CONSIGLIATI:

- **P. Billingsley**, Probability and Measure, Wiley 1984.
- **G. Koch**, La matematica del probabile, Aracne, 1997.

ALTRI TESTI CONSIGLIATI:

- **D. Williams**, Probability with martingales, Cambridge University Press, 1991;
- **L. Breiman**, Probability, Addison Wesley, 1968;
- **Y.S. Chow, H. Teicher**, Probability Theory, Springer Verlag, 1988;
- **B. De Finetti**, Teoria delle Probabilità, Einaudi, 1970;
- **W. Feller**, An introduction to probability theory and its applications (Vol 1 e 2), Wiley & Sons, 1970

Programma:

A. **MISURE DI PROBABILITÀ**:

σ-algebre di eventi, σ-additività e continuità delle misure di probabilità. sigma-algebre generate, λ-sistemi e π-sistemi di eventi, lemma $\pi - \lambda$ di Dynkin, limiti superiore ed inferiore per successioni di eventi. Indipendenza stocastica fra sigma-algebre, lemmi di Borel-Cantelli, sigma-algebra coda di una successione di eventi, Legge 0-1 di Kolmogoroff. Misure di probabilità sulla retta, funzioni di distribuzione associate.

B. **VARIABILI ALEATORIE (COME FUNZIONI MISURABILI)**:

σ-algebra generata da una funzione misurabile (e da un vettore aleatorio), misura indotta da una funzione misurabile (e da un vettore aleatorio).

Misura di probabilità indotta da una variabile aleatoria reale X (legge di una variabile aleatoria). Funzioni di distribuzione reali e spazi canonici: (a) i reali $\mathbb{R}$ e i boreliani di $\mathbb{R}$ con la misura indotta da X; (b) lo spazio $(0, 1)$ con i boreliani di $(0, 1)$ e la misura di Lebesgue ristretta a $(0, 1)$ e costruzione di Skorohod su $(0, 1)$.

Integrazione di funzioni misurabili (cenni). Valori attesi e proprietà fondamentali.

Indipendenza stocastica per variabili aleatorie e misure prodotto, costruzione di una successione di variabili aleatorie indipendenti sullo spazio canonico $(0, 1)$.

???Applicazioni probabilistiche del Teorema di Fubini.

Legge 0-1 di Kolmogoroff per variabili aleatorie. Enunciati dei teoremi di convergenza monotona e dominata. Disuguaglianza di Markov.

C. **CONVERGENZA PER SUCCESSIONI DI VARIABILI ALEATORIE**:

Definizioni di convergenza quasi certa, in probabilità ed in legge (o in distribuzione) per successioni di variabili aleatorie. Relative proprietà e relazioni. Convergenza debole per successioni di misure di probabilità, caratterizzazioni (Teoremi di Helly) e relazioni con la convergenza in legge. Teorema di Scheffé (convergenza delle densità di probabilità)[solo enunciato]. Successioni *tight* (trattenute o strette) di misure di probabilità e Teorema di Prohorov. Uniforme integrabilità e convergenza debole.

D. **FUNZIONI CARATTERISTICHE E TEOREMA CENTRALE DEL LIMITE**:
Definizione di funzione caratteristica di una distribuzione di probabilità sulla retta e relative proprietà. Teorema di Bochner [solo enunciato]. Calcolo della funzione caratteristica in casi notevoli; distribuzioni simmetriche. Relazione tra funzione caratteristica e momenti della distribuzione. Cenno al problema dei momenti: controesempio della legge lognormale. Teorema di inversione e caso in cui la funzione caratteristica appartiene a $L^1(\mathbb{R})$ (è integrabile sui reali rispetto alla misura di Lebesgue).
Teorema di continuità. Teorema Centrale del Limite per successioni di variabili aleatorie indipendenti: Teorema di Lindeberg-Levy, Teorema di Lindeberg con discussione della condizione di Lindeberg (la generalizzazione al caso di insiemi triangolari è facoltativa) Teorema di Lyapunov, Teorema di Berry-Esseen (solo enunciato).
Applicazioni del Teorema Centrale del Limite (approssimazione della legge della somma di variabili aleatorie indipendenti identicamente distribuite, relazione con la formula di Stirling)

E. **LEGGI FORTI DEI GRANDI NUMERI**:
Legge forte con esistenza e limitatezza del momento quarto (di Cantelli). Disuguaglianza di Kolmogoroff. Criterio sufficiente di Kolmogoroff per la legge forte. Legge forte per variabili aleatorie indipendenti identicamente distribuite (i.i.d.) con momento primo (Teorema di Kinchin). Estensione al caso di variabili aleatorie i.i.d., ma con momento primo della parte positiva non finito.

Riferimenti dettagliati per gli argomenti in A e B:

[Billinsgley]:
1: The Unit Interval,
2: Spaces, Classes of Sets, Probability Measures,
3: Uniqueness and the $\lambda - \pi$ Theorem (si consiglia la lettura di tutto)
4: tutto,
5: se ne consiglia la lettura,
10:, 11: e 12: i contenuti si considerano noti, e se ne consiglia la lettura,
20: Random variables and Vectors, Subfields, Distributions, Independence, Sequences of Random Variables, Convolution,
21: Expected Values and Distributions, Moments, Inequalities, Independence and Expected Values,
22: Kolmogorov's 0-1 law.

[Koch]:
Cap. 4, appendice II,
Cap. 5: 5.1, 5.2 e 5.3 (si consiglia la lettura anche dei rimanenti paragrafi),
Cap. 6: Teorema 6.87,
Cap. 8: 8.1 e 8.2 (si consiglia la lettura di 8.3),
Cap. 9: 9.1 e 9.2,
Cap.13: 13.2.

Riferimenti dettagliati per gli argomenti in C, D ed E:

[Billinsgsley]:
20: Convergence in Probability,
22: Kolmogorov's Inequality, The strong Law of Large Numbers,
25: tutto,
26: tutto,
27: Identically Distributed Summands, The Lindeberg and Lyapunov Theorems.

[Koch]:
Cap.11,
Cap.12: 12.1, 12.2.,
Cap.13: 13.1 (fino a pag. 503), Prop. 13.34, 13.3
(in particolare i Teoremi 13.49, 13.50, 13.51, 13.52, 13.59, 13.60)

Capitolo 1

Richiami su spazi di probabilità

1.1 Esempi di spazi di probabilità

Come dovrebbe essere noto uno spazio di probabilità è una terna $(\Omega, \mathcal{F}, \mathbb{P})$, dove

$\mathcal{F}$ è una ***σ-algebra***, ovvero $\mathcal{F}$ è una famiglia di sottoinsiemi di Ω, cioè $\mathcal{F}$ è un sottoinsieme di $\mathcal{P}(\Omega)$, tale che

$$\Omega \in \mathcal{F}; \tag{1.1}$$
$$\text{se } A \in \mathcal{F}, \text{ allora } A^c \in \mathcal{F}; \tag{1.2}$$
$$\text{se } A_n \in \mathcal{F}, n \in \mathbb{N}, \text{ allora } \cup_{n\in\mathbb{N}} A_n \in \mathcal{F}; \tag{1.3}$$

$\mathbb{P}$ è una ***misura di probabilità***, ovvero

$$\mathbb{P} : \mathcal{F} \mapsto [0,1]; \quad A \mapsto \mathbb{P}(A)$$

con le proprietà che

$$\mathbb{P}(\Omega) = 1; \tag{1.4}$$
$$\text{se } A_n \in \mathcal{F}, n \in \mathbb{N}, \text{ con } A_n \cap A_m = \emptyset \text{ per } n \neq m, \tag{1.5}$$
$$\text{allora } \mathbb{P}\big(\bigcup_{n\in\mathbb{N}} A_n\big) = \sum_{n\in\mathbb{N}} \mathbb{P}\big(A_n\big).$$

La σ-algebra $\mathcal{F}$ rappresenta l'informazione disponibile, ovvero gli eventi appartenenti a $\mathcal{F}$ sono gli unici eventi di cui abbiamo la possibilità di sapere se si sono verificati oppure no.

Oltre alla misura di probabilità $\mathbb{P}$, per tutti gli eventi $A \in \mathcal{F}$ con $\mathbb{P}(A) > 0$, si possono definire le ***probabilità condizionate***[1] ***all'evento*** A, che rappresentano la valutazione della probabilità nel caso in cui si verificasse l'evento A:

$$\mathbb{P}(\cdot|A)\,\mathcal{F} : \to [0,1] \tag{1.6}$$
$$E \mapsto \mathbb{P}(E|A) := \frac{\mathbb{P}(E \cap A)}{\mathbb{P}(A)} \tag{1.7}$$

Vediamo ora alcuni esempi elementari di spazi di probabilità:

[1] È facile verificare che la funzione $\mathbb{P}(\cdot|A)$ definita in (1.6) è una probabilità, cioè soddisfa gli assiomi delle probabilità. Per mettere in evidenza tale fatto va detto che Kolmogorov aveva adottato la notazione $\mathbb{P}_A(\cdot)$, ovvero $\mathbb{P}_A(E)$ invece di $\mathbb{P}(E|A)$, anche per mettere meglio in evidenza questa proprietà.

Esempio 1.1. *Qualunque sia Ω, la σ-algebra banale $\mathcal{F} = \{\emptyset, \Omega\}$ è una σ-algebra, e necessariamente $\mathbb{P}(\Omega) = 1$ e $\mathbb{P}(\emptyset) = 0$.*

Esempio 1.2. *Qualunque sia Ω, preso un sottoinsieme proprio A di Ω la σ-algebra $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$ è una σ-algebra, e necessariamente $\mathbb{P}(\Omega) = 1$, $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(A) = p$, $\mathbb{P}(A^c) = 1 - p$, per un $p \in [0, 1]$.*

Esempio 1.3. *Qualunque sia Ω, sia $\{H_m,\, m = 1, 2, \dots, N\}$ una* **partizione finita** *di Ω, cioè se gli eventi sono* **incompatibili***:*

$$H_n \cap H_m = \emptyset \text{ per } n \neq m,\ n, m \in \{1, 2, \dots, N\}$$

ed **esaustivi***:*

$$\bigcup_{m=1}^{N} H_m = \Omega,$$

allora la famiglia $\mathcal{M} = \{A = \bigcup_{m \in I} H_m,$ al variare di $I \subseteq \{1, 2, \dots, N\}\}$, (con la convenzione che $\bigcup_{m \in \emptyset} H_m = \emptyset$) è una σ-algebra. Inoltre se $p_1, p_2, \dots, p_N$ sono numeri non negativi, a somma 1, ovvero

$$p_m \geq 0,\ m = 1, 2, \dots, N, \qquad \sum_{m=1}^{N} p_m = 1,$$

allora $\mathbb{P}: \mathcal{M} \mapsto [0, 1];\ A \mapsto \mathbb{P}(A)$, con

$$\mathbb{P}(A) = \sum_{m \in I} p_m, \qquad \text{per } A = \bigcup_{m \in I} H_m, \tag{1.8}$$

definisce una probabilità su $(\Omega, \mathcal{M})$.

Esempio 1.4. *Le proprietà dell'esempio precedente valgono anche nel caso di una* **partizione numerabile** *$\{H_m,\, m \in \mathbb{N}\}$ con i dovuti cambiamenti: cioè, se*

$$H_n \cap H_m = \emptyset \text{ per } n \neq m,\ n, m \in \mathbb{N},$$
$$\bigcup_{m \in \mathbb{N}} H_m = \Omega,$$

allora la famiglia

$$\mathcal{F} = \{A = \bigcup_{m \in I} H_m, \text{ al variare di } I \subseteq \mathbb{N}\},$$

(con la convenzione che $\bigcup_{m \in \emptyset} H_m = \emptyset$), è una σ-algebra[2].
Inoltre se $p_1, p_2, \dots, p_m, \dots$ sono numeri non negativi, somma 1, ovvero

$$p_m \geq 0,\ m \in \mathbb{N}, \qquad \sum_{m \in \mathbb{N}} p_m = 1,$$

[2] La verifica è banale:

$$\Omega = \bigcup_{m \in \mathbb{N}} H_m, \text{ ovvero } I = \mathbb{N}$$
$$\text{se } A = \bigcup_{m \in I} H_m, \text{ allora } A^c = \bigcup_{m \in I^c} H_m$$
$$\text{se } A_n = \bigcup_{m \in I_n} H_m,\ n \geq 1, \text{ allora } \bigcup_{n=1}^{\infty} A_n = \bigcup_{m \in I} H_m, \text{ per } I = \cup_{n=1}^{\infty} I_n.$$

allora $\mathbb{P}: \mathcal{F} \mapsto [0,1];\ A \mapsto \mathbb{P}(A)$, *con*

$$\mathbb{P}(A) = \sum_{m\in I} p_m, \qquad \text{per } A = \bigcup_{m\in I} H_m, \tag{1.9}$$

definisce una probabilità su $(\Omega, \mathcal{F})$.

La verifica di quest'ultima proprietà è banale[3].

Elenchiamo adesso alcune ***proprietà e notazioni relative alle σ-algebre***:

1 ***l'intersezione di σ-algebre è una σ-algebra***
Sia $\{\mathcal{G}_\alpha, \alpha \in \Lambda\}$ una famiglia di σ-algebre, allora $\mathcal{F} := \bigcap_{\alpha\in\Lambda} \mathcal{G}_\alpha$ è una σ-algebra[4].

2 ***l'unione di σ-algebre non è (in generale) una σ-algebra***
Basta mostrare con un controesempio che l'unione di due σ-algebre non è una σ-algebra: ad esempio se $\mathcal{G}_i = \{\emptyset, A_i, A_i^c, \Omega\}$, con $A_1 \cap A_2 \neq \emptyset, A_1, A_2$, allora $\mathcal{G}_1 \cup \mathcal{G}_2 = \{\emptyset, A_1, A_2, A_1^c, A_2^c, \Omega\}$ non è una σ-algebra.

3 ***la σ-algebra generata da una collezione di eventi***
Sia $\mathcal{K}$ un sottoinsieme di $\mathcal{P}(\Omega)$, l'insieme delle parti di Ω, allora

$$\sigma(\mathcal{K}) := \bigcap_{\mathcal{G}:\mathcal{K}\subseteq\mathcal{G}}$$

è la σ-algebra[5] generata da $\mathcal{K}$.

In particolare quindi la σ-algebra $\mathcal{M}$, generata dalla partizione $\{H_m;\ m \in \mathbb{N}\}$ come nell'Esempio 1.4, coincide con $\sigma(\{H_m;\ m \in \mathbb{N}\})$, in quanto, come già visto $\mathcal{M}$ è una σ-algebra, e inoltre ogni σ-algebra che contenga $\{H_m;\ m \in \mathbb{N}\}$, deve necessariamente contenere tutte le unioni del tipo $\cup_{m\in I} H_m$.

4 ***la σ-algebra generata da una collezione di σ-algebre***
Nel caso in cui $\mathcal{K} = \bigcup_{\alpha\in\Lambda} \mathcal{G}_\alpha$, dove $\mathcal{G}_\alpha$ sono σ-algebre, allora si pone

$$\bigvee_{\alpha\in\Lambda} \mathcal{G}_\alpha := \sigma\big(\bigcup_{\alpha\in\Lambda} \mathcal{G}_\alpha\big).$$

In particolare se $\mathcal{M} = \sigma(\{H_m;\ m \in \mathbb{N}\})$ e $\mathcal{N} = \sigma\big(\{K_\ell;\ \ell \in \mathbb{N}\}\big)$, allora

$$\mathcal{M} \vee \mathcal{N} = \sigma(\{H_m \cap K_\ell;\ m \in \mathbb{N}, \ell \in \mathbb{N}\}) = \big\{E = \bigcup_{(m,\ell)\in J} H_m \cap K_\ell;\ \text{con } J \subseteq \mathbb{N}\times\mathbb{N}\big\}.$$

[3]La funzione $\mathbb{P}: \mathcal{M} \mapsto [0,1]$ definita in (1.9) è una probabilità, infatti

$$\mathbb{P}(\Omega) = \sum_{m\in\mathbb{N}} p_m = 1,$$

$$\text{se } A_n = \bigcup_{m\in I_n} H_m \in \mathcal{M}, n\in\mathbb{N}, \text{ con } A_n \cap A_{n'} = \emptyset \text{ per } n \neq n',$$

$$\text{allora } \bigcup_{n\in\mathbb{N}} A_n = \bigcup_{m\in I} H_m \text{ con } I = \bigcup_{n\in\mathbb{N}} I_n, \text{ e con } I_n \cap I_{n'} = \emptyset \text{ per } n \neq n',$$

$$\text{e quindi } \mathbb{P}\big(\bigcup_{n\in\mathbb{N}} A_n\big) = \mathbb{P}(A) = \sum_{\ell\in I} p_\ell = \sum_{n\in\mathbb{N}}\sum_{m\in I_n} p_m = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n),$$

[4]La verifica è banale:

$\Omega \in \mathcal{F}$, in quanto $\Omega \in \mathcal{G}_\alpha$, per ogni $\alpha \in \Lambda$;

se $A \in \mathcal{F}$, cioè se $A \in \mathcal{G}_\alpha$, per ogni $\alpha \in \Lambda$, allora $A^c \in \mathcal{G}_\alpha$, per ogni $\alpha \in \Lambda$, e quindi $A^c \in \mathcal{F}$;

se $A_n \in \mathcal{F}, n \in \mathbb{N}$ cioè se $A_n \in \mathcal{G}_\alpha$, per ogni $\alpha \in \Lambda, n \in \mathbb{N}$ allora $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{G}_\alpha$, per ogni $\alpha \in \Lambda$, e quindi $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$;

[5]Il fatto che $\bigcap_{\mathcal{G}:\mathcal{K}\subseteq\mathcal{G}}$ sia una σ-algebra, deriva dalla proprietà che l'intersezione di σ-algebre è una σ-algebra.

5 ***la σ-algebra dei Boreliani*** Nel caso in cui $\mathcal{K} = \mathcal{A}$, la famiglia degli aperti di $\mathbb{R}^k$, allora

$$\mathcal{B}(\mathbb{R}^k) := \sigma(\mathcal{A})$$

è detta σ-algebra dei boreliani, o σ-algebra di Borel, ed ogni elemento di $\mathcal{I}$ di $\mathcal{B}(\mathbb{R}^k)$ è detto ***boreliano***.

1.2 Variabili aleatorie

Definizione 1.1. *Dato uno spazio di probabilità* $(\Omega, \mathcal{F}, \mathbb{P})$[6]*, una* **variabile aleatoria reale** X *è una funzione* $\mathcal{F}$-**misurabile***, ovvero una funzione*

$$X : \Omega \mapsto \mathbb{R}; \quad \omega \mapsto X(\omega),$$

tale che la controimmagine di ogni aperto $\mathcal{O} \in \mathcal{A}$ *sia un elemento di* $\mathcal{F}$[7]*, cioè tale che*

$$X^{-1}(\mathcal{O}) := \{\omega \text{ tali che } X(\omega) \in \mathcal{O}\} \in \mathcal{F}, \text{ per ogni aperto } \mathcal{O} \in \mathcal{A}.$$

Si dice anche che X *è una* **variabile aleatoria** $\mathcal{F}$**-misurabile**.
Una definizione analoga vale nel caso di variabili aleatorie multidimensionali

$$\boldsymbol{X} : \Omega \mapsto \mathbb{R}^k; \quad \omega \mapsto \boldsymbol{X}(\omega) = \big(X_1(\omega), \dots, X_k(\omega)\big),$$

basta infatti sostituire $\mathbb{R}$ *con* $\mathbb{R}^k$.

Vediamo alcuni esempi di variabili aleatorie $\mathcal{F}$-misurabili, al variare della σ-algebra $\mathcal{F}$.

Esempio 1.5. *Se* $\mathcal{F} = \{\emptyset, \Omega\}$*, allora le uniche variabili aleatorie reali* X $\mathcal{F}$*-misurabili sono le costanti:*
Se $X : \Omega \mapsto \mathbb{R}$; $\omega \mapsto X(\omega) = c$*, allora* $X^{-1}(\mathcal{O})$ *è l'evento impossibile(=insieme vuoto* $\emptyset$*), se* $c \notin \mathcal{O}$*, oppure è l'insieme certo(=*Ω*), se* $c \in \mathcal{O}$*.*
Viceversa se $X : \Omega \mapsto \mathbb{R}$; $\omega \mapsto X(\omega)$ *non è costante allora* X *assume almeno due valori* c_1 *e* c_2 *distinti (cioè esistono* ω_i *tale che* $X(\omega_i) = c_i$*, per* $i = 1, 2$*, con* $c_1 \neq c_2$*). Quindi se* $c_1 \in \mathcal{O}$*, ma* $c_2 \notin \mathcal{O}$*, allora* $\omega_1 \in X^{-1}(\mathcal{O})$*, mentre* $\omega_2 \notin X^{-1}(\mathcal{O})$*, ovvero* $\emptyset \subset X^{-1}(\mathcal{O}) \subset \Omega$ *(dove le inclusioni sono in senso stretto), e quindi* X *non è* $\mathcal{F}$*-misurabile.*

Si noti che l'esempio precedente mostra anche che tutte le variabili aleatorie costanti sono misurabili rispetto a qualunque σ-algebra ($\{\emptyset, \Omega\} \subseteq \mathcal{F}$, per ogni σ-algebra $\mathcal{F}$).

Esempio 1.6. *Sia* $\{H_m,\, m \in \mathbb{N}\}$ *una partizione numerabile, e sia* $\mathcal{M}$ *come nell'esempio 1.4. Allora* $X : \Omega \mapsto \mathbb{R}$; $\omega \mapsto X(\omega)$ *è* $\mathcal{M}$*-misurabile, se e solo se esiste una successione di costanti* $\{c_m,\, m \in \mathbb{N}\}$[8]*, tale che*

$$X(\omega) = \sum_{m \in \mathbb{N}} c_m \mathbb{I}_{H_m}(\omega). \tag{1.10}$$

Se X *è definita come in (1.10) allora* X *è* $\mathcal{M}$*-misurabile, infatti per ogni aperto* $\mathcal{O}$,

$$X^{-1}(\mathcal{O}) = \bigcup_{m: c_m \in \mathcal{O}} H_m,$$

ovvero $X^{-1}(\mathcal{O}) = \bigcup_{m \in I} H_m \in \mathcal{M}$*, per* $I = \{m : c_m \in \mathcal{O}\}$*.*
Viceversa se X *è* $\mathcal{M}$*-misurabile, cioè, per ogni aperto* $\mathcal{O}$*, esiste un* $I \subseteq \mathbb{N}$ *tale che*

$$X^{-1}(\mathcal{O}) = \bigcup_{m \in I} H_m,$$

---

[6]In realtà basta che ci sia uno spazio ***probabilizzabile***, ovvero basta solo la coppia $(\Omega, \mathcal{F})$, mentre non è necessario specificare la misura di probabilità $\mathbb{P}$.

[7]Si noti l'analogia con la definizione di funzione continua $f : \mathbb{R}^k \mapsto \mathbb{R}^d$, come una funzione tale che le controimmagini di aperti sono aperti.

[8]Si noti che non si assume che i valori di $\{c_m\}$ siano tutti distinti, ad esempio nel caso della successione costante, cioè $c_m = c$ per ogni $m \in \mathbb{N}$, si trova una variabile aleatoria costante.

allora qualunque sia $c \in \mathbb{R}$, preso $\mathcal{O}^n$ l'intervallo aperto $(c - 1/n, c + 1/n)$ si ha che

$$X^{-1}(\{c\}) = X^{-1}\big(\bigcap_n \mathcal{O}^n\big) = \bigcap_n X^{-1}(\mathcal{O}^n) = \bigcap_n \bigcup_{m \in I^n} H_m = \bigcup_{m \in \bigcap_n I^n} H_m \in \mathcal{M},$$

Esempio 1.7. *Sia $X : \Omega \mapsto \mathbb{R}$; $\omega \mapsto X(\omega)$, una funzione discreta, ovvero tale che l'immagine $X(\Omega) = \{x \in \mathbb{R}$, tali che esiste un ω con $X(\omega) = x\}$ di X sia un insieme numerabile (finito o infinito), cioè $X(\Omega) = \{x_m, m \in \mathbb{N}\}$, con $x_n \neq x_m$ per $n \neq m$. Allora*

$$X(\omega) = \sum_{m \in \mathbb{N}} x_m \mathbb{I}_{H_m}(\omega), \tag{1.11}$$

dove

$$H_m = X^{-1}(\{x_m\}) = \{\omega \text{ tali che } X(\omega) = x_m\}.$$

Si noti che $\{H_m,\, m \in \mathbb{N}\}$ forma una partizione numerabile.

*Inoltre la funzione X è una variabile aleatoria **$\mathcal{F}$-misurabile**, se e solo se*

$$H_m = X^{-1}(\{x_m\}) \in \mathcal{F}, \text{ per ogni } m \in \mathbb{N},$$

come è immediato da (1.11), osservando che, come nel caso precedente,

$$X^{-1}(\mathcal{O}) = \bigcup_{m: x_m \in \mathcal{O}} H_m.$$

*Infine la variabile aleatoria X si dice **semplice** o **elementare**, se l'insieme $X(\Omega)$ è un insieme finito.*

Si può dimostrare che

1 se X è una variabile aleatoria $\mathcal{F}$-misurabile, allora la controimmagine $X^{-1}(\mathcal{I}) \in \mathcal{F}$, per ogni boreliano $\mathcal{I} \in \mathcal{B}(\mathbb{R})$,

2 la variabile aleatoria $\boldsymbol{X}$ è $\mathcal{F}$-misurabile, se e solo se ciascuna componente X_i è $\mathcal{F}$-misurabile[9], per ogni $i = 1, \ldots, k$. In particolare $X^{-1}(\{x\}) \in \mathcal{F}$, per ogni $x \in \mathbb{R}$, in quanto $\{x\} = \bigcap_n (x - 1/n, x + 1/n)$.

Connessa con la precedente Definizione 1.1 è la seguente definizione:

Definizione 1.2. *Sia data una funzione $\boldsymbol{X} : \Omega \mapsto \mathbb{R}^k$; $\omega \mapsto \boldsymbol{X}(\omega) = \big(X_1(\omega), \ldots, X_k(\omega)\big)$. Si dice **$\boldsymbol{\sigma}$-algebra generata da $\boldsymbol{X}$**, la σ-algebra*

$$\sigma(\boldsymbol{X}) = \bigcap_{\mathcal{G} \in \mathcal{R}^{\boldsymbol{X}}} \mathcal{G}$$

dove $\mathcal{R}^{\boldsymbol{X}}$ è la famiglia delle σ-algebre, per le quali $\boldsymbol{X}$ è $\mathcal{G}$-misurabile[10].

Si dimostra che

3 La σ-algebra generata da $\boldsymbol{X}$, si può caratterizzare come:

$$\sigma(\boldsymbol{X}) = \{A = \boldsymbol{X}^{-1}(\mathcal{I}), \text{ per } \mathcal{I} \in \mathcal{B}(\mathbb{R}^k)\},$$

4 la funzione $\boldsymbol{X}$ è $\mathcal{F}$-misurabile, se e solo se $\sigma(\boldsymbol{X}) \subseteq \mathcal{F}$,

5 le variabili aleatorie $\sigma(\boldsymbol{X})$-misurabili a valori in $\mathbb{R}^d$ sono tutte e sole le variabili aleatorie $\boldsymbol{Z}$ per le quali esiste una funzione g boreliana[11] tale che

$$Z = g(\boldsymbol{X}).$$

---

[9] Dimostriamo solo la necessità, che è immediata: basta prendere $\mathcal{O} = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{i-1 \text{ volte}} \times \mathcal{O}_i \times \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{k-i \text{ volte}}$.

[10] La famiglia $\mathcal{R}^{\boldsymbol{X}}$ non è vuota, in quanto contiene almeno $\mathcal{G} = \mathcal{P}(\Omega)$, l'insieme delle parti di Ω.

[11] Una funzione $g : \mathbb{R}^k \mapsto \mathbb{R}^d$, si dice boreliana se è una funzione tale che le controimmagini di aperti sono boreliani.
Ovviamente le funzioni continue sono boreliane. Sono boreliane anche le funzioni continue a tratti, o meglio ancora costanti a tratti.
Per chi non avesse familiarità con i concetti di misurabilità può pensare a queste funzioni, o a funzioni che siano limite puntuale di funzioni di uno dei due tipi precedenti.

Esempio 1.8. *Sia X una funzione semplice, come in Esempio 1.7, allora*

$$\sigma(X) = \sigma(\{H_m,\, m \in \mathbb{N}\}) = \{A = \bigcup_{m\in I} H_m;\ I \subseteq \mathbb{N}\},$$

dove $H_m = X^{-1}(\{x_m\})$.

Inoltre tutte e sole le variabili aleatorie $\sigma(X)$-misurabili sono le funzioni

$$Z : \Omega \mapsto \mathbb{R};\ \omega \mapsto Z(\omega) := \sum_m c_m \mathbb{I}_{H_m},$$

come discende immediatamente dall'Esempio 1.6. Di conseguenza se $g : \mathbb{R} \mapsto \mathbb{R}$ tale che $g(x_m) = c_m$, per ogni $m \in \mathbb{N}$, allora

$$Z(\omega) := \sum_m c_m \mathbb{I}_{H_m} = Z(\omega) = \sum_m g(x_m) \mathbb{I}_{X^{-1}(\{x_m\})}(\omega) = \sum_m g(x_m) \mathbb{I}_{\{x_m\}}\big(X(\omega)\big) = g(X(\omega)).$$

Terminiamo questa sezione, ricordando che le operazioni di massimo, minimo, somma, prodotto, di due funzioni misurabili, danno luogo a funzioni misurabili: quindi se X ed Y sono variabili aleatorie $\mathcal{F}$-misurabili, lo sono anche $X \vee Y = \max(X,Y)$, $X \wedge Y = \min(X,Y)$, $X+Y$, XY. In particolare sono variabili aleatorie $X^+ := X \vee 0$ e $X^- := (-X) \vee 0$.

1.3 Distribuzioni di variabili aleatorie

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

$$X : \Omega \mapsto \mathbb{R}^k;\ \omega \mapsto X(\omega)$$

una variabile aleatoria a valori in $\mathbb{R}^k$. Tramite X è possibile definire una misura di probabilità $\mathbf{P}_X$ sullo spazio misurabile $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ nel seguente modo:

$$\mathbf{P}_X\ : \mathcal{B}(\mathbb{R}^k) \mapsto [0,1] \quad \mathcal{I} \mapsto \mathbf{P}_X(\mathcal{I}) := \mathbb{P}\big(X \in \mathcal{I}\big).$$

È facile verificare che effettivamente $\mathbf{P}_X$ definisce una probabilità sui boreliani $\mathcal{B}(\mathbb{R}^k)$. La misura di probabilità così definita è detta ***misura di probabilità indotta da*** X, o ***distribuzione di*** X.

A volte, per indicare la misura di probabilità indotta, si usa il simbolo $\mathbb{P}X^{-1}$, che nasce dall fatto che $\mathbf{P}_X(\mathcal{I}) := \mathbb{P}\big(X \in \mathcal{I}\big) = \mathbb{P}\big(X^{-1}(\mathcal{I})\big)$. Nel seguito, a volte useremo anche il simbolo μ_X per indicare la distribuzione di probabilità di X.

Come è noto, associata alla variabile aleatoria X c'è anche la funzione di distribuzione[12]

$$F_X(x) := \mathbb{P}(\omega \in \Omega :\ X(\omega) \le x) = \mathbf{P}_X\big((-\infty, x]\big), \qquad x \in \mathbb{R}^k. \tag{1.12}$$

La funzione di distribuzione gode di alcune proprietà caratterizzanti[13]:

Proprietà delle funzioni di distribuzione

0 $F_X(x) \in [0,1]$

1 La funzione F_X è continua dall'alto[14], nel senso che, per ogni $x \in \mathbb{R}^k$ si ha

$$\lim_{y \searrow x} F_X(y_1, \cdots, y_i, \cdots, y_k) = F(x_1, \cdots, x_i, \cdots, x_k),$$

dove $y \searrow x$ significa $y_i \to x_i^+$, per ogni $i = 1, \cdots, k$.

[12]Si ricordi che, per $k \ge 1$, l'evento e l'insieme nella (1.12) sono rispettivamente

$$\{\omega \in \Omega :\ X(\omega) \le x\} = \{\omega \in \Omega :\ X_1(\omega) \le x_1, \cdots, X_k(\omega) \le x_k\} \quad \text{e} \quad (-\infty, x] = (-\infty, x_1] \times \cdots \times (-\infty, x_k].$$

[13]Si veda la sezione 2.1

[14]Nel caso $k = 1$ la proprietà **1** corrisponde alla continuità da destra.

2 La funzione $F_X(x)$ è monotona non decrescente.

3 Siano $a = (a_1, \cdots, a_k)$ e $b = (b_1, \cdots, b_k)$, si definisca

$$\Delta(a, b) = \{x \in \mathbb{R}^k : \forall i = 1, \cdots, k, \text{ si ha } x_i = a_i \text{ oppure } x_i = b_i\},$$

e si definisca $n_a(x)$ il numero di i tali che $x_i = a_i$, per $x \in \Delta(a, b)$.
Se $a_i \leq b_i$, per ogni $i = 1, \cdots, k$, allora[15]

$$\sum_{x \in \Delta(a,b)} (-1)^{n_a(x)} F_X(x) \geq 0.$$

4 Per ogni $x \in \mathbb{R}^k$ e per ogni $i = 1, \cdots, k$ si ha che

$$\lim_{y_i \to -\infty} F_X(x_1, \cdots, x_{i-1}, y_i, x_{i+1}, \cdots, x_k) = 0.$$

Inoltre

$$\lim_{|x| \to +\infty} F_X(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_k) = 1.$$

È importante sottolineare che la funzione di distribuzione F_X individua la misura di probabilità indotta $\mathbf{P}_X$ sulla famiglia (di boreliani)

$$(-\infty, b] := \{x \in \mathbb{R}^k \,:\, x_i \leq b_i\} \qquad \text{con } b = (b_1, \cdots, b_k) \in \mathbb{R}^k.$$

Questa famiglia ha la proprietà di essere chiusa rispetto all'intersezione finita:

$$(-\infty, b] \cap (-\infty, b'] = (-\infty, b \wedge b'], \qquad \text{dove}\, b \wedge b' := (b_1 \wedge b_1', \cdots, b_k \wedge b_k').$$

Ciò è sufficiente a individuare la misura di probabilità indotta, grazie a un risultato molto utile di teoria della misura:

Lemma 1.1 (Lemma di Dynkin, Billingsley 1984 [1])**.** *Sia $\mathcal{A}$ una famiglia di eventi che genera la σ-algebra $\mathcal{G}$ e che è chiusa rispetto alla intersezione finita (cioè: A, $B \in \mathcal{A}$ implica $A \cap B \in \mathcal{A}$). Se due misure di probabilità ν e μ coincidono su $\mathcal{A}$, allora le due misure coincidono su $\mathcal{G} = \sigma(\mathcal{A})$.*

Definizione 1.3 (variabili aleatorie con densità discreta)**.** *Si dice che una variabile aleatoria elementare X ha densità discreta*

$$\begin{pmatrix} x_1 & x_2 & \cdots & \cdots & x_m \\ p_1 & p_2 & \cdots & \cdots & p_m \end{pmatrix}$$

dove $x_1, x_2, \cdots x_m$ sono elementi di $\mathbb{R}^k$ e $p_1, p_2, \cdots p_m$ sono numeri reali tali che

$$p_j \geq 0 \quad \textit{per ogni } j = 1, 2, \cdots, m, \qquad \sum_{j=1}^{m} p_j = 1,$$

[15] Nel caso $k = 1$, la proprietà **3** corrisponde alla proprietà di monotonia **2**:

$$\text{se } a \leq b \quad \text{allora} \quad F_X(a) \leq F_X(b).$$

Nel caso $k = 2$, invece la proprietà **3** diviene:

$$\text{se } a_1 \leq b_1 \text{ e } a_2 \leq b_2 \quad \text{allora} \quad F_X(b_1, b_2) - F_X(a_1, b_2) - F_X(b_1, a_2) + F_X(a_1, a_2) \geq 0.$$

Per $k \geq 2$ la proprietà **3** non si riduce alla proprietà di monotonia **2**, come mostra il seguente controesempio:

$$F(x_1, x_2) =] \begin{cases} 0 & \text{se } x < 0, \quad \text{oppure se } x + y < 1, \quad \text{oppure se } y < 0. \\ 1 & \text{se } x \geq 0,\ y \geq 0, \text{ e } x + y \geq 1 \end{cases}$$

Si vede facilmente che F è una funzione monotona. Tuttavia F non soddisfa la proprietà $\not\models$, infatti

$$F(1, 1) - F(1, 0) - F(0, 1) + F(0, 0) = 1 - 1 - 1 + 0 = -1.$$

se, per ogni boreliano $\mathcal{I}$, vale

$$\mathbf{P}_X(\mathcal{I}) := \mathbb{P}(X \in \mathcal{I}) = \sum_{\substack{j=1 \\ x_j \in \mathcal{I}}}^{m} p_j.$$

In particolare quindi il significato di p_j è chiaro, essendo

$$\mathbb{P}(X = x_j) = p_j.$$

La definizione è analoga nel caso di variabili aleatorie discrete, la cui distribuzione viene caratterizzata attraverso una densità discreta su un insieme numerabile $\{x_k,\ k \geq 1\}$

$$\begin{pmatrix} x_1 & x_2 & \cdots & \cdots & x_m & x_{m+1} & \cdots \\ p_1 & p_2 & \cdots & \cdots & p_m & p_{m+1} & \cdots \end{pmatrix}$$

Esempio 1.9 (variabili aleatorie con distribuzione binomiale)**.** *Ogni variabile aleatoria X per la quale*

$$\mathbf{P}_X(\mathcal{I}) := \sum_{\substack{h=0 \\ h \in \mathcal{I}}}^{n} \binom{n}{h} p^h \, (1-p)^{n-h}$$

viene detta una variabile aleatoria binomiale di parametri n e p e si scrive in breve $X \sim Bin(n,p)$.

Definizione 1.4 (variabili con densità)**.** *Sisupponga di avere una funzione $f : \mathbb{R}^k \mapsto \mathbb{R}$ con le proprietà:*

$$f(x) \geq 0 \quad \textit{per ogni } x \in \mathbb{R}^k, \qquad \int_{\mathbb{R}^k} f(x)\,dx = 1,$$

si dice che X ha distribuzione con **densità (di probabilità)** *f se accade che, per ogni boreliano $\mathcal{I} \in \mathcal{B}(\mathbb{R}^k)$,*

$$\mathbf{P}_X(\mathcal{I}) := \int_{\mathcal{I}} f(x)\,dx.$$

Esempio 1.10 (distribuzione gaussiana)**.** *Come caso particolare si consideri il caso della variabile aleatoria unidimensionale con densità*

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

dove μ è un numero reale e σ è un numero (strettamente) positivo. Una variabile aleatoria con questa distribuzione è detta **gaussiana** *o* **normale** *di valore atteso (o valore medio) μ e varianza σ^2. Brevemente si indica $X \sim N(\mu, \sigma^2)$. Se $\mu = 0$ e $\sigma^2 = 1$ si dice che X è una variabile gaussiana (o normale)* **standard**.

Vediamo ora dei semplici esempi di calcolo della distribuzione indotta.

Esempio 1.11 (una variabile aleatoria binomiale)**.** *Sia*

$$\Omega = \{0,1\}^N = \{\omega = (\omega_1, \omega_2, \ldots, \omega_N),\ \textit{con } \omega_i \in \{0,1\},\ \textit{per } i = 1, 2, \ldots, N\},$$

sia

$$\mathcal{F} = \mathcal{P}(\Omega),$$

l'insieme delle parti di Ω, sia la probabilità definita attraverso la relazione

$$\mathbb{P}(\{\omega\}) := p^{\sum_{i=1}^{N} \omega_i} \, (1-p)^{N - \sum_{i=1}^{N} \omega_i},$$

dove p è un numero fissato con la condizione che $p \in (0,1)$. Sia infine X la variabile aleatoria definita da

$$X(\omega) := \sum_{i=1}^{N} \omega_i.$$

Si vede facilmente che

• 1 *la variabile aleatoria* X *assume solo i valori* $\{0, 1, \ldots N\}$,

• 2 *per* $h \in \{0, 1, \ldots N\}$ *si ha*[16]

$$\mathbf{P}_X(h) := \mathbb{P}(X = h) = \binom{N}{h} p^h\,(1-p)^{N-h},$$

• 3 *per ogni boreliano* $\mathcal{I}$

$$\mathbf{P}_X(\mathcal{I}) := \mathbb{P}(X \in \mathcal{I}) = \sum_{\substack{h=0\\ h\in\mathcal{I}}}^{N} \binom{N}{h} p^h\,(1-p)^{N-h}$$

Esempio 1.12 (Variabili esponenziali)**.** *Sia* $\Omega = (0,1)$ $\mathcal{F} = \mathcal{B}(0,1)$ *e* $\mathbb{P}$ *la misura di Lebesgue su* $(0,1)$*. Sia* $\lambda > 0$ *e* $X(\omega) := -\log(1-\omega)/\lambda$*. Allora*

$$F_X(x) = \mathbb{P}(X \le x) = mis\{\omega \in (0,1)\,:\, -\log(1-\omega) \le \lambda\,x\} = mis\{\omega \in (0,1)\,:\, \omega \le 1 - e^{-\lambda\,x}\},$$

e quindi

$$F_X(x) = \begin{cases} 0 & per\ x \le 0,\\ 1 - e^{-\lambda\,x} & per\ x > 0. \end{cases}$$

Per il Lemma di Dynkin (Lemma 1.1) sappiamo che la funzione di distribuzione individua univocamente la distribuzione di X*. È quindi facile convincersi che, tale distribuzione coincide con la distribuzione*

$$\nu_\lambda(dx) = \mathbf{1}_{(0,\infty)}(x)\,\lambda\,e^{-\lambda\,x}\,dx,$$

che è nota come la distribuzione esponenziale di parametro λ*.*
Sempre nello stesso spazio si può definire la variabile aleatoria

$$Y(\omega) = -\frac{\log(\omega)}{\mu},$$

dove μ *è una costante strettamente positiva. È facile vedere che* Y *ha distribuzione esponenziale, di parametro* μ*.*

Esempio 1.13. *Sempre nello stesso ambito dell'esempio precedente, ci si può chiedere quale sia la distribuzione congiunta di* X *e* Y*, ossia la distribuzione del vettore aleatorio* (X,Y)*.*

Chiaramente, si ha $X(\omega)$, $Y(\omega) > 0$ *e inoltre*

$$Y(\omega) = -\frac{\log\big(1 - e^{-\lambda\,X(\omega)}\big)}{\mu},$$

come si ottiene subito da

$$\omega = 1 - e^{-\lambda\,X(\omega)}.$$

Di conseguenza, se $G := \{(x,y)\,:\, x > 0, y > 0, ey = -\frac{\log\left(1-e^{-\lambda\,x}\right)}{\mu}\}$*, è facile convincersi che*

$$\mathbf{P}_{X,Y}(\mathcal{I}) = \nu_\lambda\big(\pi_x(G\cap\mathcal{I})\big)\Big(= \nu_\mu\big(\pi_y(G\cap\mathcal{I})\big)\Big)$$

dove π_x *e* π_y *sono le proiezione sull'asse* x *e sull'asse* y*, rispettivamente.*

[16] L'evento $A_h := \{X = h\}$ è rappresentato dall'insieme, di cardinalità $\binom{N}{h}$, i cui elementi $\omega = (\omega_1, \omega_2, \ldots, \omega_N)$ hanno la proprietà che $\sum_{i=1}^N \omega_i = h$. La probabilità di ciascuno di questi ω vale quindi

$$\mathbb{P}(\omega) = p^{\sum_{i=1}^N \omega_i}\,(1-p)^{N-\sum_{i=1}^N \omega_i} = p^h\,(1-p)^{N-h}$$

e la probabilità dell'insieme vale

$$\mathbb{P}(X = h) = \mathbb{P}(A_h) = \sum_{\omega\in A_h} \mathbb{P}(\omega) = \sum_{\omega\in A_h} p^h\,(1-p)^{N-h} = |A_h| p^h\,(1-p)^{N-h} = \binom{N}{h} p^h\,(1-p)^{N-h}$$

Esempio 1.14 (trasformazione di Box-Müller). *Sia $\Omega = (0,1) \times (0,1)$, con la misura di Lebesgue sui boreliani. Siano*

$$X(\omega_1, \omega_2) := \sqrt{-2\,\log \omega_1}\,\cos(2\,\pi\,\omega_2);$$
$$Y(\omega_1, \omega_2) := \sqrt{-2\,\log \omega_1}\,\sin(2\,\pi\,\omega_2);$$

Si può dimostrare che la distribuzione congiunta di (X,Y) ammette densità si probabilità

$$p_{X,Y}(x,y) = \frac{1}{2\,\pi}\,e^{-\frac{x^2+y^2}{2}} = \frac{1}{\sqrt{2\,\pi}}\,e^{-\frac{x^2}{2}}\,\frac{1}{\sqrt{2\,\pi}}\,e^{-\frac{x^2}{2}}$$

Tale densità caratterizza le variabili aleatorie gaussiane com media nulla e matrice di covarianza l'identità (si veda l'Appendice 1.6).

A volte, invece di definire lo spazio di probabilità e la variabile aleatoria X ed infine trovare la distribuzione di X, si può dare direttamente la distribuzione di X. Questo è il caso delle variabili aleatorie che vengono caratterizzate solo attraverso la densità discreta o con densità (di probabilità).

Più in generale, le distribuzioni si possono specificare solo attraverso la funzione di distribuzione.

Quando si specifica una variabile aleatoria attraverso la sua distribuzione, e ancor di più se invece si specifica solo una funzione che goda delle proprietà delle funzioni di distribuzione (si veda pag. 6), rimane il dubbio che una tale variabile aleatoria esista, ovvero che esista uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e una variabile aleatoria X. A questo problema risponde il teorema di Skorohod (vedere Appendice 2.1).

1.4 Valori attesi

In questa sezione ricordiamo come si può definire il valore atteso per variabili aleatorie generali, a partire dalla sua definizione per variabili aleatorie semplici. Per maggiori approfondimenti si rimanda, ad esempio, al libro di Billingsley [1] o a quello di Williams [3].

Definizione 1.5 (Valore atteso per variabili semplici). *Sia X una variabile aleatoria in $(\Omega, \mathcal{F}, \mathbb{P})$, non negativa e semplice, cioè come in Esempio 1.7,*

$$X(\omega) = \sum_{m\in\mathbb{N}} x_m \mathbb{I}_{H_m}(\omega), \quad \text{con } H_m \in \mathcal{F} \text{ per ogni } m \in \mathbb{N},$$

allora si definisce

$$\mathbb{E}[X] = \sum_{m\in\mathbb{N}} x_m \mathbb{P}(H_m).$$

Osservazione 1.1. *Ogni variabile aleatoria X in $(\Omega, \mathcal{F}, \mathbb{P})$, non negativa, ammette una successione di variabili aleatorie X_n, semplici e non negative, tali che*

$$0 \le X_n(\omega) \le X_{n+1}(\omega), \quad \text{e tali che } \lim_{n\to\infty} X_n(\omega) = X(\omega).$$

Infatti[17] *basta prendere*

$$X_n(\omega) = \sum_{m=0}^{n2^n-1} \frac{m}{2^n} \mathbb{I}_{H_m^{(n)}}(\omega) + n\mathbb{I}_{H_{n2^n}^{(n)}}(\omega) = \sum_{m=0}^{n2^n-1} \frac{m}{2^n} \mathbf{1}_{[\frac{m}{2^n}, \frac{m+1}{2^n})}(X(\omega)) + n\mathbf{1}_{[n,\infty)}(X(\omega)), \tag{1.13}$$

[17] La monotonia della successione delle variabili aleatorie X_n è evidente:
• se $X_n(\omega) = m/2^n$, con $m < n2^n$, allora i soli casi possibili sono

$$X_{n+1}(\omega) = (2m)/2^{n+1} = m/2^n = X_n(\omega),$$

oppure

$$X_{n+1}(\omega) = (2m+1)/2^{n+1} = m/2^n + 1/2^{n+1} > X_n(\omega);$$

• se $X_n(\omega) = n$ allora $X_{n+1}(\omega)$ può assumere un valore compreso tra n ed $n+1$.

Per la convergenza basta osservare che, qualunque sia ω, pur di prendere n sufficientemente grande e in modo che $X(\omega) < n$, si ha che

$$0 \le X(\omega) - X_n(\omega) \le 1/2^n.$$

dove si è posto

$$H_m^{(n)} = X^{-1}\left(\left[\tfrac{m}{2^n}, \tfrac{m+1}{2^n}\right)\right) \in \mathcal{F} \text{ per } 0 \le m \le n2^n - 1, \quad H_{n2^n}^{(n)} = X^{-1}\big([n,\infty)\big),$$

e, per $A \in \mathcal{F}$,

$$\mathbb{I}_A(\omega) = 1 \quad se \quad \omega \in A \qquad e \qquad \mathbb{I}_A(\omega) = 0 \quad se \quad \omega \notin A,$$

ed infine, per $a < b$ *numeri reali,*

$$\mathbf{1}_{[a,b)}(x) = 1 \quad se \quad x \in [a,b) \qquad e \qquad \mathbf{1}_{[a,b)}(x) = 0 \quad se \quad x \notin [a,b).$$

È infine interessante notare che, posto $\lfloor x \rfloor$ *la parte intera inferiore*[18] *di* x, *si può riscrivere nel seguente modo*

$$X_n(\omega) = \frac{\lfloor 2^n X(\omega) \rfloor}{2^n} \wedge n.$$

Definizione 1.6 (Valore atteso per variabili nonnegative). *Sia* X *una variabile aleatoria in* $(\Omega, \mathcal{F}, \mathbb{P})$, *non negativa, si definisce*

$$\mathbb{E}[X] = \sup\Big\{ \sum_{i=1}^{n} \inf_{\omega \in A_i} X(\omega)\, \mathbb{P}(A_i) \text{ al variare tra le partizioni dell'evento certo } A_1, \cdots, A_n \Big\}.$$

Si dimostra che

$$\mathbb{E}[X] = \lim_{n\to\infty} \mathbb{E}[X_n],$$

dove $\{X_n;\ n \in \mathbb{N}\}$ *è la successione monotona definita come in (1.13) dell'Osservazione precedente. Il limite esiste ed è monotono, per la proprietà di monotonia del valore atteso, sulle variabili aleatorie semplici. Si noti bene che tale limite può valere anche* $+\infty$, *nel qual caso si dice che la variabile* X *ha valore atteso infinito.*

Osservazione 1.2. ***Ovviamente se* $\{\widehat{X}_n;\ n \in \mathbb{N}\}$ *è un'altra successione di variabili aleatorie semplici che converge monotonamente ad* X, *anche la successione dei valori attesi* $\mathbb{E}[\widehat{X}_n]$ *è una successione che converge monotonamente. Si può dimostrare che il limite non dipende dalla successione scelta*[19] *ed in particolare coincide con il limite considerato nella precedente Definizione 1.6.*

Arriviamo ora alla definizione generale del valore atteso:

Definizione 1.7 (Valore atteso per variabili generali). *Sia* X *una variabile aleatoria in* $(\Omega, \mathcal{F}, \mathbb{P})$, *Siano* $X^+ := X \vee 0$ *e* $X^- := (-X) \vee 0$, *le variabili aleatorie non negative, definite alla fine della sezione precedente. Si noti che* $X = X^+ - X^-$ *e che invece* $|X| = X^+ + X^-$. *Si definisce allora, se ha senso*[20]

$$\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-].$$

[18] La parte intera inferiore $\lfloor x \rfloor$ di x è quel numero intero k tale che $k \le x < k+1$.

[19] **Per ottenere l'unicità del limite basta dimostrare che se $\{Y_n;\ n \in \mathbb{N}\}$ e $\{Z_n;\ n \in \mathbb{N}\}$ sono due successioni di variabili aleatorie semplici che convergono monotonamente ad X, allora per ogni k si ha

$$\mathbb{E}[Y_k] \le \lim_{n\to\infty} \mathbb{E}[Z_n],$$

da cui si deduce immediatamente che $\lim_{k\to\infty} \mathbb{E}[Y_k] \le \lim_{n\to\infty} \mathbb{E}[Z_n]$e quindi l'uguaglianza, scambiando il ruolo delle due successioni.
Si fissi quindi k e si consideri che, per ipotesi Y_k è semplice e che quindi si può scrivere $Y_k = \sum_{i=1}^{\ell} y_i \mathbb{I}_{A_i}$, dove $A_i = \{Y_k = y_i\}$ (ovviamente ℓ ed y_i dipendono da k, ma tralasciamo l'indice k per comodità di notazione e perché è inessenziale). Sia ora $\varepsilon > 0$ e $B_i^{(n)} = A_i \cap \{Z_n > y_i - \varepsilon\}$. Essendo $\{Z_n;\ n \in \mathbb{N}\}$ una successione monotona si ottiene che $B_i^{(n)} \subset B_i^{(n+1)}$. Inoltre $Y_k(\omega) \le X(\omega)$ e $Z_n(\omega) \nearrow X(\omega)$ e quindi se $\omega \in A_i$, ossia se $Y_k(\omega) = y_i$, allora per un n sufficientemente grande deve valere $Z_n(\omega) > y_i - \varepsilon$ e quindi $\bigcup_n B_i^{(n)} = A_i$. Per la continuità della probabilità deve valere allora che $\mathbb{P}(B_i^{(n)}) \nearrow \mathbb{P}(A_i)$. Ovviamente si ha

$$\mathbb{E}[Z_n] \ge \sum_{i=1}^{\ell} (y_i - \varepsilon)\mathbb{P}(B_i^{(n)}) \qquad \text{e quindi} \qquad \lim_{n\to\infty} \mathbb{E}[Z_n] \ge \sum_{i=1}^{\ell} (y_i - \varepsilon)\mathbb{P}(A_i).$$

Per l'arbitrarietà di ε si ha allora $\lim_{n\to\infty} \mathbb{E}[Z_n] \ge \sum_{i=1}^{\ell} y_i \mathbb{P}(A_i) = \mathbb{E}[Y_k]$.

[20] Si considera che la somma $\mathbb{E}[X^+] - \mathbb{E}[X^-]$ ha senso

1 se $\mathbb{E}[X^+] < \infty$, $\mathbb{E}[X^-] < \infty$, nel qual caso $\mathbb{E}[X] \in \mathbb{R}$ e inoltre si ha anche $\mathbb{E}[|X|] = \mathbb{E}[X^+] + \mathbb{E}[X^-] < \infty$;

2 se $\mathbb{E}[X^+] < \infty$, $\mathbb{E}[X^-] = \infty$, nel qual caso $\mathbb{E}[X] = -\infty$;

3 se $\mathbb{E}[X^+] = \infty$, $\mathbb{E}[X^-] < \infty$, nel qual caso $\mathbb{E}[X] = +\infty$;

Il caso che rimane escluso è quindi il caso in cui $\mathbb{E}[X^+] = \infty$, $\mathbb{E}[X^-] = \infty$, del resto si avrebbe la forma indeterminata $\infty - \infty$.

Se invece di usare la probabilità $\mathbb{P}$ si usa la probabilità condizionata ad un evento A, ovvero $\mathbb{P}(\cdot|A)$, allora si parla di ***valore atteso di*** X ***condizionato all'evento*** A e si usa la notazione

$$\mathbb{E}[X|A].$$

Ciò significa che, nel caso di una variabile aleatoria semplice

$$X(\omega) = \sum_{m\in\mathbb{N}} x_m \mathbb{I}_{H_m}(\omega), \quad \text{con } H_m \in \mathcal{F} \text{ per ogni } m \in \mathbb{N},$$

si ha

$$\mathbb{E}[X|A] = \sum_{m\in\mathbb{N}} x_m \mathbb{P}(H_m|A).$$

**Terminiamo questa sezione ricordando che la definizione di valore atteso di una variabile aleatoria X corrisponde alla definizione dell'integrale della funzione misurabile X rispetto alla misura $\mathbb{P}$ e che per il valore atteso valgono i due famosi risultati di passaggio al limite sotto il segno di integrale:
Teorema della convergenza monotona: se X_n sono variabili aletaorie limitate dal basso e che convergono monotonamente ad X ($\mathbb{P} - q.c.$) allora la successione dei valori attesi $\mathbb{E}[X_n]$ converge monotonamente a $\mathbb{E}[X]$.

Teorema della convergenza dominata: se X_n sono variabili aletaorie che convergono ad X $\mathbb{P} - q.c.$ e se Y è una variabile aleatoria tale che $|X_n| \leq Y$, con $\mathbb{E}[Y] < \infty$, allora la successione dei valori attesi $\mathbb{E}[X_n]$ converge a $\mathbb{E}[X]$.

1.4.1 Variabili aleatorie in spazi misurabili

QUESTA SEZIONE SI PUO' SALTARE Oltre a definire le variabili aleatorie reali o vettoriali si possono definire in modo naturale anche variabili aleatorie a valori in spazi misurabili.

Definizione 1.8 (variabile aleatoria (o ente alealorio) a valori in $(S, \mathcal{S})$)**.** *Siano $(\Omega, \mathcal{F})$ e $(S, \mathcal{S})$ due spazi misurabili. Una variabile aleatoria a valori in S è una funzione misurabile*

$$X: \ (\Omega, \mathcal{F}) \to (S, \mathcal{S}); \ \omega \mapsto X(\omega).$$

In altre parole una funzione da Ω in S e tale che per ogni $B \in \mathcal{S}$, la sua controimmagine tramite X appartiene a $\mathcal{F}$, ossia l'insieme $X^{-1}(B) \in \mathcal{F}$.

Se S è uno **spazio metrico** *(o più in generale uno spazio topologico, allora la sigma-algebra $\mathcal{S}$ coincide con la sigma-algebra dei boreliani, ossia la sigma-algebra generata dagli aperti.*

Esempi tipici nascono quando si vogliono trattare i processi aleatori come funzioni aleatorie, ed in particolare a funzioni aleatorie continue. In tale caso si può prendere, ad esempio, lo spazio delle funzioni continue su $[0, T]$ a valori reali. Prendendo poi come sigma-algebra la sigma-algebra dei boreliani, allora si può affermare che funzioni come il massimo o il minimo, sono variabili aleatorie.

Come si vede, nella definizione di variabile aleatoria non abbiamo neanche nominato la misura di probabilità su $(\Omega, \mathcal{F})$.

1.5 Misura indotta e Cambio di variabile

QUESTO ARGOMENTO E' SVOLTO QUI IN MODO PIU' APPROFONDITO CHE A LEZIONE.

Negli Esempi 1.12, 1.13, 1.14 abbiamo trovato le distribuzioni di alcune variabili aleatorie a valori reali o vettoriali. In termini astratti quello che abbiamo fatto è caratterizzare la misura indotta.

Definizione 1.9 (Misura indotta)**.** *Siano $(A_1, \mathcal{A}_1)$ e $(A_2, \mathcal{A}_2)$ due spazi con le rispettive sigma-algebre, e sia $\psi: \ A_1 \to A_2, \ a_1 \mapsto \psi(a_1)$ una funzione misurabile (cioè per ogni $B_2 \in \mathcal{A}_2$ si ha che la controimmagine $\psi^{-1}(B_2) \in \mathcal{A}_1$). Supponiamo che su $(A_1, \mathcal{A}_1)$ sia definita una misura μ_1. Allora si definisce misura indotta (da ψ) la misura*

$$\mu_2(B_2) := \mu_1\big(\psi^{-1}(B_2)\big), \qquad B_2 \in \mathcal{A}_2.$$

Ovviamente perché la precedente definizione sia ben posta bisogna verificare che effettivamente definisca una misura (questo è un semplice esercizio ed è lasciato al lettore).

Tornando agli Esempi precedentemente citati ed in particolare agli Esempi 1.12, 1.13, in entrambi $(A_1, \mathcal{A}_\infty) = \big((0,1), \mathcal{B}(0,1)\big)$ e $\mu_1 = \mathbb{P}$, la misura di Lebesgue ristretta a $(0,1)$, mentre $A_2 = \mathbb{R}$ nel primo esempio e invece $A_2 = \mathbb{R}^2$, nel secondo esempio. Inoltre nel primo esempio sono state considerate due funzioni $\psi_1(\omega) = X(\omega) = -\frac{\log(1-\omega)}{\lambda}$ e $\psi_2(\omega) = Y(\omega) = -\frac{\log(\omega)}{\lambda}$, mentre nel secondo esempio è stata considerata la funzione $\psi := (\psi_1, \psi_2)$. Nell'Esempio 1.14, invece $(A_1, \mathcal{A}_\infty) = \big((0,1) \times (0,1), \mathcal{B}((0,1) \times (0,1)\big)$ e $A_2 = \mathbb{R}^2$ e μ_1 è la misura di Lebesgue ristretta a $(0,1) \times (0,1)$. Infine la funzione ψ è definita da $\psi(\omega_1, \omega_2) = \big(\sqrt{-2\,\log \omega_1}\, \cos(2\,\pi\, \omega_2), \sqrt{-2\,\log \omega_1}\, \sin(2\,\pi\, \omega_2)\big)$.

Più in generale, nel caso di variabili aleatorie X a valori in $(S; \mathcal{S})$, se nello spazio misurabile $(\Omega, \mathcal{F})$ è definita una misura di probabilità $\mathbb{P}$, si definisce legge di X o distribuzione di X, la probabilità $\mathbf{P}_X : \mathcal{S} \to [0,1]$ definita come la misura indotta da $(\Omega, \mathcal{F}, \mathbb{P})$ tramite X:

$$\mathbf{P}_X(B) := \mathbb{P}(X \in B), \quad B \in \mathcal{S}.$$

Quello che più ci interessa qui è la formula del cambio di variabile negli integrali, che, nell'ambito del calcolo delle probabilità, corrisponde alla possibilità di calcolare i valori attesi di funzioni di variabili aleatorie X a valori[21] in $(S, \mathcal{S})$ sia come integrali sullo spazio $(\Omega, \mathcal{F}, \mathbb{P})$ che come integrali sullo spazio $(S, \mathcal{S}, \mathbf{P}_X)$.

In tale caso si ottiene che i valori attesi di $f(X)$, per f funzioni misurabili e limitate, si possono calcolare sia come integrali sulo spazio degli eventi Ω

$$\mathbb{E}[f(X)] = \int_\Omega f(X(\omega))\, \mathbb{P}(d\omega),$$

sia come integrale sullo spazio degli stati S

$$\mathbb{E}[f(X)] = \int_S f(x)\, \mathbf{P}_X(dx).$$

Riportiamo qui la dimostrazione nell'ambito astratto della Definizione 1.9 di misura indotta.

Lemma 1.2 (Cambio di variabile)**.** *Sia $f \in \mathcal{M}_b(A_2)$, ossia una funzione misurabile da $\big(A_2, \mathcal{A}_2\big)$ in $\big(\mathbb{R}, \mathcal{B}(\mathbb{R})\big)$ e limitata. Allora*

$$\int_{A_2} f(a_2)\mu_2(da_2) = \int_{A_1} f(\psi(a_1))\mu_1(da_1) \tag{1.14}$$

Dimostrazione. Iniziamo con il mostrare che, per definizione di μ_2, (1.14) è valida per $f = \mathbb{I}_{B_2}$, per ogni $B_2 \in \mathcal{A}_2$: da una parte

$$\int_{A_2} \mathbb{I}_{B_2}(a_2)\mu_2(da_2) = \mu_2(B_2) := \mu_1\big(\psi^{-1}(B_2)\big),$$

dall'altra, tenuto conto che $\mathbb{I}_{B_2}(\psi(a_1)) = \mathbb{I}_{\psi_{-1}(B_2)}(a_1)$, in quanto $\psi(a_1) \in B_2$ se e solo se $a_1 \in \psi^{-1}(B_2)$,

$$\int_{A_1} \mathbb{I}_{B_2}(\psi(a_1))\mu_1(da_1) = \int_{A_1} \mathbb{I}_{\psi^{-1}(B_2)}(a_1)\mu_1(da_1) = \mu_1\big(\psi^{-1}(B_2)\big).$$

La dimostrazione segue poi con una tecnica che è standard nell'ambito della teoria della misura.

Sia $\mathcal{H}$ l'insieme delle funzioni f per cui è valida l'uguaglianza (1.14).

L'insieme $\mathcal{H}$ verifica le seguenti proprietà:

(i) **linearità**, ovvero se f, $g \in \mathcal{H}$, allora, per ogni a,$b \in \mathbb{R}$ la funzione $a\,f + b\,g \in \mathcal{H}$

[21] Nel caso di variabili aleatorie vettoriali lo spazio $(S, \mathcal{S})$ coincide con $\big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\big)$. Ma la formula vale anche per variabil aleatorie a valori in spazi più generali, come ad esempio gli spazi metrici, prendendo come sigma-algebra la sigma-algebra dei boreliani, ossia la sigma-algebra generata dagli aperti (in altre parole la più piccola sigma algebra contenente gli aperti). Come già detto esempi di tale genere si incontrano quando ci si interessa di processi aleatori, pensati come variabili aletorie a valori in uno spazio di funzioni, ad esempio lo spazio delle funzioni continue su un intervallo $[0, T]$, con la metrica della norma uniforme.

(come segue dalla proprietà di linearità per gli integrali rispetto a μ_1)
(ii) la funzione **1**, cioè la funzione costante uguale ad 1, appartiene a $\mathcal{H}$
(come segue dall'osservazione iniziale e notando che $\mathbf{1} = \mathbb{I}_{A_2}$,)
(iii) **monotonia**, ovvero se $f_n \in \mathcal{H}$ e $f_n \nearrow f$, $f \in \mathcal{M}_2(A_2)$ allora $f \in \mathcal{H}$
(come segue dalla proprietà della convergenza monotona degli integrali rispetto a μ_1)
(iv) per ogni $B_2 \in \mathcal{A}_2$, la funzione $\mathbb{I}_{B_2} \in \mathcal{H}$
(come segue immediatamente dalla osservazione iniziale; si noti inoltre che in realtà la (ii) segue da questa proprietà)

Le precedenti proprietà assicurano che $\mathcal{H}$ è una classe monotona. Basta allora applicare il teorema delle classi monotone, che per comodità del lettore riportiamo di seguito. $\square$

Teorema 1.3 (Teorema delle classi monotone). *Sia $(\Omega, \mathcal{F})$ uno spazio misurabile e sia $\mathcal{H}$ un insieme di funzioni reali misurabili e limitate, con le seguenti proprietà:*
(i) $\mathcal{H}$ è uno spazio vettoriale,
(ii) $\mathcal{H}$ contiene la funzione costante **1**,
(iii) $f_n \in \mathcal{H}$, $f_n \nearrow f$, f limitata implicano $f \in \mathcal{H}$
cioè $\mathcal{H}$ è una classe monotona.

Se inoltre $\mathcal{H}$ soddisfa anche la seguente proprietà

(iv) $\mathcal{H}$ contiene le funzioni del tipo $\mathbb{I}_A$ per ogni $A \in \mathcal{A}$, dove $\mathcal{A} \subset \mathcal{F}$ è un π-sistema, cioè è chiuso per intersezione finita,

allora $\mathcal{H}$ contiene tutte le funzioni limitate e $\sigma(\mathcal{A})$-misurabili.

Il precedente Teorema 1.2 si applica anche quando vogliamo calcolare la distribuzione di una trasformazione di una variabile aleatoria: ad esempio, se Z è una variabile aleatoria con distribuzione $\mathbf{P}_Y$ ed $Z = \varphi(Y)$, allora $\mathbf{P}_Z(B) = \mathbf{P}_Y\big(\varphi^{-1}(B)\big)$, come è immediato verificare.

Nel caso di variabili aleatorie multivariate, e per funzioni φ sufficientemente regolari, si possono ottenere formule esplicite, utilizzando noti risultati di analisi: ad esempio, se Y ammette densità f_Y e φ è invertibile[22] e con derivate continue, allora anche Z ammette densità e si ha

$$f_Z(z) = f_Y(\varphi^{-1}(z)) \left| det \left(\frac{\partial \varphi^{-1}(z)}{\partial z} \right) \right| = f_Y(\varphi^{-1}(z)) \frac{1}{\left| det \left(\frac{\partial \varphi(y)}{\partial y} \right) \right|_{y=\varphi^{-1}(z)}}.$$

Particolarmente semplice è il caso di trasformazioni lineari (o affini) in cui lo Jacobiano è il determinante della matrice. Ad esempio se $Z = AY$, con A invertibile, allora $\varphi^{-1}(z) = A^{-1}z$ e la formula precedente diviene

$$f_Z(z) = f_Y(A^{-1}(z)) \frac{1}{|det(A)|}.$$

Esempio 1.15. *Un esempio di trasformazione che incontreremo spesso nel seguito è il caso in cui $Y = (Y_1, Y_2, \cdots, Y_m)$ e*

$$\begin{aligned} Z_1 &= Y_1, \\ Z_2 &= Y_1 + Y_2, \\ &\cdots \\ Z_m &= Y_1 + Y_2 + \cdots + Y_m, \end{aligned} \qquad \textit{ossia } z = \varphi(y) = A\,y, \textit{ con } \qquad \begin{aligned} z_1 &= y_1, \\ z_2 &= y_1 + y_2, \\ &\cdots \\ z_m &= y_1 + y_2 + \cdots + y_m. \end{aligned}$$

Allora la matrice A è la matrice triangolare $A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}$ *con determinante uguale ad 1.*

[22] In realtà basta che esista un aperto $\mathcal{O}$, tale che la densità $f_Y(y) = 0$ per $y \notin A$ e tale che φ sia invertibile da $\mathcal{O}$ a $\varphi(\mathcal{O}$,

La trasformazione inversa è

$$
\begin{aligned}
y_1 &= z_1,\\
y_2 &= z_2 - z_1,\\
&\cdots\\
y_m &= z_m - z_{m-1},
\end{aligned}
\qquad \textit{ossia } y = \varphi^{-1}(z) = A^{-1}y \textit{ dove } A^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0\\ -1 & 1 & 0 & \cdots & 0\\ 0 & -1 & 1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}
$$

per cui, se Y ammette densità di probabilità,

$$f_Z(z_1, z_2, \cdots, z_m) = f_Y(z_1, z_2 - z_1, \cdots, z_m - z_{m-1}).$$

Il caso $m = 2$ è particolarmente interessante in quanto permette di ricavare la densità della somma di due variabili aleatorie, semplicemente calcolando la densità marginale di $Z_2 = Y_1 + Y_2$: per $z \in \mathbb{R}$

$$f_{Y_1+Y_2}(x)\Big(= f_{Z_2}(x)\Big) = \int_{\mathbb{R}} f_{Z_1,Z_2}(x, x')\, dx' = \int_{\mathbb{R}} f_{Y_1,Y_2}(x, x' - x)\, dx'.$$

1.6 Variabili gaussiane

QUESTO ARGOMENTO DOVREBBE ESSERE GIA' STATO SVOLTO IN ALTRO CORSO, ALMENO IN PARTE...

Cominciamo con il definire una variabile aleatoria gaussiana standard unidimensionale:

Definizione 1.10. *Si dice che una variabile aleatoria reale Z è* ***gaussiana*** *di valore atteso μ e varianza σ^2, se ammette densità*

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}.$$

In questo caso si usa la notazione $Z \sim N(\mu, \sigma^2)$. Se $\mu = 0$ e $\sigma^2 = 1$ allora si dice che Z segue una ***legge normale o gaussiana standard****.*

Caso $n-$dimensionale: iniziamo con il caso di un vettore (colonna) aleatorio

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_k \\ \cdots \\ Y_n \end{pmatrix}$$

a componenti indipendenti e tutte gaussiane standard, ovvero il caso in cui

$$\begin{aligned} f_Y(\boldsymbol{y}) &= \prod_{i=1}^{n} f_{Y_i}(y_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}y_i^2\right\} \\ &= \frac{1}{(\sqrt{2\pi})^n} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} y_i^2\right\} = \frac{1}{(2\pi)^{n/2}} \exp\left\{-\frac{1}{2}\boldsymbol{y}'\boldsymbol{y}\right\}. \end{aligned}$$

dove l'apice indica l'operazione di trasposizione, ovvero $\boldsymbol{y'}$ è il vettore riga $(y_1, y_2, \cdots, y_n)$.

È immediato verificare che $\mathbb{E}(Y_i) = 0$, $Var(Y_i) = 1$ e che $Cov(Y_i, Y_j) = 0$, per $i \neq j$.

Sia ora A una matrice non singolare e sia $\boldsymbol{m}$ un vettore (colonna). Definiamo ora $Z = AY + \boldsymbol{m}$ e cerchiamo la sua densità. Sappiamo dai risultati generali che se Y ammette densità e $Z = \varphi(Y)$ con φ invertibile e con derivate continue, allora anche Z ammette densità:

$$f_Z(z) = f_Y(\varphi^{-1}(z)) \left| det\left(\frac{\partial \varphi^{-1}(z)}{\partial z}\right)\right| = f_Y(\varphi^{-1}(z)) \frac{1}{\left|det\left(\frac{\partial \varphi(y)}{\partial y}\right)\right|_{y=\varphi^{-1}(z)}}$$

di conseguenza, poiché nel nostro caso $\varphi(y) = Ay + m$ e $\varphi^{-1}(z) = A^{-1}(z - m)$

$$f_Z(z) = \frac{1}{\sqrt{2\pi}^n} \exp\left\{-\frac{1}{2}\left(A^{-1}(z-m)\right)' A^{-1}(z-m)\right\} \frac{1}{|det(A)|}.$$

Essendo

$$\begin{aligned} &(A^{-1}(z-m))'A^{-1}(z-m) = (z-m)'(A^{-1})'A^{-1}(z-m) \\ &= (z-m)'(A')^{-1}A^{-1}(z-m) = (z-m)'(AA')^{-1}(z-m) \end{aligned}$$

si ottiene

$$f_Z(z) = \frac{1}{(2\pi)^{n/2}} \frac{1}{|det(A)|} \exp\left\{-\frac{1}{2}(z-m)'\left(AA'\right)^{-1}(z-m)\right\}.$$

La precedente espressione si basa sulle seguenti proprietà:

(i) $\boxed{(A')^{-1} = (A^{-1})'}$

in quanto

$$A'z = w \Leftrightarrow z = (A')^{-1}w$$

e inoltre

$$A'z = w \Leftrightarrow (z'A)' = w \Leftrightarrow z'A = w' \Leftrightarrow z' = w'A^{-1}$$

$$\Leftrightarrow z = \left(w'A^{-1}\right)' \Leftrightarrow z = \left(A^{-1}\right)' w.$$

(ii) $\boxed{(AA')^{-1} = (A')^{-1} A^{-1}}$

in quanto

$$(AA')^{-1} z = w \Leftrightarrow z = AA'w \Leftrightarrow A^{-1}z = A'w \Leftrightarrow (A')^{-1} A^{-1}z = w.$$

È interessante notare che sia il vettore $\boldsymbol{m}$ che la matrice $AA' = A'A$ hanno una interpretazione probabilistica:

$$\mathbb{E}(Z_i) = \mathbb{E}(\sum_{k=1}^{n} a_{i,k}Y_k) + m_i = \sum_{k=1}^{n} a_{i,k}\mathbb{E}(Y_k) + m_i = m_i$$

$$Cov(Z_i, Z_j) = \mathbb{E}[(Z_i - m_i)(Z_j - m_j)] = \mathbb{E}[\sum_{k=1}^{n} a_{i,k}Y_k \sum_{h=1}^{n} a_{j,h}Y_h] = \sum_{k=1}^{n}\sum_{h=1}^{n} a_{i,k}a_{j,h}\mathbb{E}[Y_kY_h]$$

e quindi

$$Cov(Z_i, Z_j) = \sum_{k=1}^{n} a_{i,k}a_{j,k}\mathbb{E}[Y_kY_k] + \sum_{k=1}^{n}\sum_{h\neq k}^{1,n} a_{i,k}a_{j,h}\mathbb{E}[Y_kY_h] = \sum_{k=1}^{n} a_{i,k}a_{j,k} = (AA')_{i,j}$$

Si osservi che se $\boldsymbol{Z} = (Z_1, ..., Z_n)$ è un vettore gaussiano allora $(Z_1...Z_k)$ e $(Z_{k+1}, ..., Z_n)$ sono indipendenti, se e solo se $Cov(Z_i, Z_h) = 0$ per ogni $i = 1, \cdots, k$ e $h = k+1, \cdots, n$. In tale caso allora è ovvio che il vettore $(Z_1...Z_k)$ è un vettore gaussiano[23]

Terminiamo questo paragrafo con il ricordare quanto valgono i ***momenti di una variabile aleatoria gaussiana***. Sia Z una variabile aleatoria $\mathcal{N}(0, \sigma^2)$. Per quanto visto prima possiamo considerare $Z = \sigma Y$ con Y una variabile aleatoria $\mathcal{N}(0, 1)$. Da questa osservazione segue subito che

$$\mathbb{E}[Z^k] = \sigma^k \mathbb{E}[Y^k]$$

e

$$\mathbb{E}[|Z|^k] = |\sigma|^k \mathbb{E}[|Y|^k].$$

[23]Per ottenere lo stesso risultato nel caso generale, ovvero che se $\boldsymbol{Z} = (Z_1, ..., Z_n)$ è un vettore gaussiano allora $(Z_1...Z_k)$ è un vettore gaussiano, si può procedere nel seguente modo. Innanzitutto basta considerare il caso in cui i valori attesi sono nulli senza ledere in generalità. Inoltre si può pensare che $\boldsymbol{Z} = A\boldsymbol{Y}$. Se la matrice $A' = (a'_{ij})$ è definita in modo che $a'_{ij} = a_{ij}$ qualunque siano $i = 1, ...k$ e $j = 1, ..., n$, e il vettore aleatorio $\boldsymbol{Z}'$ è definito da

$$\boldsymbol{Z}' = A'\boldsymbol{Y},$$

allora, chiaramente,

$$Z'_i = (A'\boldsymbol{Y})_i = Z_i = (A\boldsymbol{Y})_i, \quad \text{per } i = 1, ...k.$$

Se inoltre a'_{hj} per $h = k+1, ...n$ e $j = 1, ..., n$ sono presi in modo che il vettore $(Z'_1, \cdots, Z'_k) = (Z_1, \cdot, Z_k)$ sia indipendente dal vettore $(Z'_{k+1}, \cdots, Z'_n)$, ovvero in modo che

$$0 = E[Z_iZ'_h] = Cov(Z_i, Z'_h) = \sum_{\ell=1}^{n} a_{i,\ell}a'_{h,\ell}$$

per $i = 1, ...k$ e $h = k+1, ...n$, allora si ottiene il risultato voluto.
Nel caso in cui la matrice A sia non singolare ciò è sempre possibile perché i vettori $\boldsymbol{a}_{(i)} = (a_{i1}, a_{i2}, \cdots, a_{in})$ sono linearmente indipendenti e quindi basta trovare $n-k$ vettori $\boldsymbol{a}'_{(h)} = (a'_{i1}, a'_{h2}, \cdots, a'_{hn})$ ortogonali allo spazio vettoriale k-dimensionale $span(\boldsymbol{a}_{(i)}, i = 1, \cdot, k)$.

Vale poi la pena di ricordare che

$$\mathbb{E}[Y^{2k+1}] = 0,$$

$$\mathbb{E}[Y^{2k}] = (2k-1)!! = (2k-1)(2k-3)\cdots 5\cdot 3\cdot 1,$$

mentre[24] infine

$$\mathbb{E}[|Y|^{2k+1}] = \sqrt{\frac{2}{\pi}}(2k)!! = \sqrt{\frac{2}{\pi}}(2k)(2k-2)\cdots 4\cdot 2 = \sqrt{\frac{2}{\pi}}2^k k!.$$

Prima di dimostrare queste tre uguaglianze si osservi che le ultime due si possono scrivere in modo sintetico come

$$\mathbb{E}[|Y|^n] = C_{((-1)^n)}\,(n-1)!! \qquad C_{(+1)} = 1 \qquad C_{(-1)} = \sqrt{\frac{2}{\pi}}.$$

La prima relazione è banale, per ragioni di simmetria, e permette di ricavare la seconda osservando che

$$\mathbb{E}[\mathrm{e}^{uY}] = \mathrm{e}^{\frac{u^2}{2}} = \sum_{h=0}^{\infty} \frac{1}{h!}\frac{u^2}{2}^h = \sum_{h=0}^{\infty} \frac{1}{h!}\frac{u^{2h}}{2^h}.$$

e d'altra parte, essendo appunto ovviamente $\mathbb{E}[Y^{2k+1}] = 0$,

$$\mathbb{E}[\mathrm{e}^{uY}] = \mathbb{E}[\sum_{k=0}^{\infty}\frac{1}{k!}u^k Y^k] = \sum_{h=0}^{\infty}\frac{1}{(2h)!}u^{2h}\mathbb{E}[Y^{2h}]$$

si deve necessariamente avere che i coefficienti delle due serie devono coincidere:

$$\frac{1}{h!}\frac{1}{2^h} = \frac{1}{(2h)!}\mathbb{E}[Y^{2h}],$$

ovvero

$$\begin{aligned}\mathbb{E}[Y^{2h}] &= \frac{(2h)!}{h!2^h} = \frac{2h(2h-1)(2h-2)(2h-3)\cdots 3\cdot 2\cdot 1}{h(h-1)\cdots 3\cdot 2\cdot 1\cdot 2^h}\\ &= \frac{(2h)!!(2h-1)!!}{h!2^h} = \frac{2^h h!\cdot(2h-1)!!}{2^h h!} = (2h-1)!!.\end{aligned}$$

Infine la terza si ricava per integrazione per parti e calcolando a mano che $\mathbb{E}[|Y|] = \sqrt{\frac{2}{\pi}}$.

Concludiamo questo paragrafo con un lemma che riguarda il comportamento asintotico della funzione di sopravvivenza di una gaussiana standard e del modulo di una gaussiana standard.

Lemma 1.4. *Sia Y una gaussiana standard, allora, posto $f_Y(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$, si ha, per $x>0$,*

$$\left(x+\frac{1}{x}\right)^{-1} f_Y(x) \le \mathbb{P}(Y > x) \le \frac{1}{x}\,f_Y(x), \qquad x>0, \tag{1.15}$$

$$\left(x+\frac{1}{x}\right)^{-1} f_{|Y|}(x) \le \mathbb{P}(|Y| > x) \le \frac{1}{x}\,f_{|Y|}(x), \qquad x>0, \tag{1.16}$$

$$\mathbb{P}(|Y|>x) \le e^{-\frac{x^2}{2}}, \qquad x>0. \tag{1.17}$$

[24] Si noti che dalle ultime due relazioni sui momenti si ottiene che

$$\mathbb{E}[|Y|^m] = (m-1)!!C_{(-1)^m}, \quad \text{con } C_{+1} = 1, C_{-1} = \sqrt{\frac{2}{\pi}}.$$

Dimostrazione. La disuguaglianza (1.16) discende immediatamente dalla prima disuguaglianza (1.15), la quale equivale a

$$\left(x+\frac{1}{x}\right)^{-1}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}}\le\mathbb{P}(Y>x)\le\frac{1}{x}\,\frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}},$$

e discende dalla seguente relazione

$$\left(w+\frac{1}{w}\right)^{-1}e^{-\frac{w^2}{2}}\le\int_w^{+\infty}e^{-\frac{z^2}{2}}\,dz\le\frac{1}{w}\,e^{-\frac{w^2}{2}},\qquad w>0. \tag{1.18}$$

La disuguaglianza destra della (1.18) discende da

$$\int_w^{+\infty}e^{-\frac{z^2}{2}}\,dz\le\frac{1}{w}\int_w^{+\infty}z\,e^{-\frac{z^2}{2}}\,dz=\frac{1}{w}\,e^{-\frac{w^2}{2}}\quad,$$

Inoltre

$$\frac{d}{dw}\frac{1}{w}\,e^{-\frac{w^2}{2}}=-\left(1+\frac{1}{w^2}\right)\,e^{-\frac{w^2}{2}}$$

e quindi

$$\frac{1}{w}\,e^{-\frac{w^2}{2}}=\int_w^{+\infty}\left(1+\frac{1}{z^2}\right)\,e^{-\frac{z^2}{2}}\,dz\le\left(1+\frac{1}{w^2}\right)\int_w^{+\infty}e^{-\frac{z^2}{2}}\,dz\quad,$$

che prova l'altra disuguaglianza nella (1.18).

Infine, per provare la disuguaglianza (1.17), basta osservare che,

$$\begin{aligned}\mathbb{P}(|Y|>x)&=2\int_x^{+\infty}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{y^2}{2}}\,dy=2\,e^{-\frac{x^2}{2}}\int_x^{+\infty}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{y^2-x^2}{2}}\,dy\\&=2\,e^{-\frac{x^2}{2}}\int_x^{+\infty}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{(y+x)(y-x)}{2}}\,dy=2\,e^{-\frac{x^2}{2}}\int_0^{+\infty}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{(z+2x)z}{2}}\,dz\qquad\text{(essendo } x>0,\text{)}\\&\le2\,e^{-\frac{x^2}{2}}\int_0^{+\infty}\frac{1}{\sqrt{2\pi}}\,e^{-\frac{z^2}{2}}\,dz=e^{-\frac{x^2}{2}}.\end{aligned}$$

□

Capitolo 2

Costruzione di variabili aleatori in $(0, 1)$

2.1 Teorema di rappresentazione di Skorohod

In questa sezione affrontiamo il problema seguente:

Data una funzione F, esiste uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ e una variabile aleatoria X, definita su questo spazio, per la quale F è la funzione di distribuzione, cioè $F = F_X$?

Chiaramente F deve soddisfare le proprietà delle funzioni di distribuzione, (ossia le proprietà **0** $-$ **4** di pagina 6). Si può dimostrare che tali proprietà sono sufficienti a individuare una misura di probabilità $\mu = \mu_F$ sui boreliani di $\mathbb{R}^k$, per la quale $F(x) = \mu(-\infty, x]$. Di conseguenza si può prendere come spazio di probabilità $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \mu_F)$ e come variabile aleatoria l'identità, ossia $X(x_1, \cdots, x_k) = (x_1, \cdots, x_k)$.

Tuttavia c'è un altro spazio in cui costruire tale variabile aleatoria, lo spazio $(0, 1)$ con la misura di Lebesgue sui boreliani di $(0, 1)$.

In questa sezione ci limitiamo al caso unidimensionale, il caso a più dimensioni (e addirittura per successioni di variabili aleatorie viene brevemente considerato nella sottosezione **??**).

Teorema 2.1 (di rappresentazione di Skorohod)**.** *Sia data una funzione F, che verifica le seguenti proprietà:*

P0 *F è a valori in* $[0, 1]$*;*

P1 *F è non decrescente;*

P2 *F è continua a destra, cioè, per ogni $t \in \mathbb{R}$,* $\lim_{\varepsilon \to 0^+} F(t + \varepsilon) = F(t)$*;*

P3 *F è normalizzata, cioè* $\lim_{t \to -\infty} F(t) = 0$*;* $\lim_{t \to +\infty} F(t) = 1$*.*

Sia $\varphi : (0, 1) \to \mathbb{R}$ definita da

$$\varphi(u) := \inf\{y : \; F(y) \geq u\}.$$

Allora φ è boreliana, inoltre la variabile aleatoria

$$\begin{array}{rcl} X : (0, 1) & \to & \mathbb{R} \\ \omega & \mapsto & X(\omega) = \varphi(\omega) \end{array}$$

definita nello spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P}) \equiv ((0, 1), \mathcal{B}(0, 1), \lambda)$ con λ la misura di Lebesgue, ha funzione di distribuzione F, ovvero $F_X(x) = F(x)$.

Dimostrazione. La dimostrazione è basata sul fatto che

$$\{u \in (0, 1) : \; \varphi(u) \leq x\} = \{u \in (0, 1) : \; u \leq F(x)\} \tag{2.1}$$

per ogni $x \in \mathbb{R}$. Dalla precedente affermazione segue infatti che:

(i) φ è misurabile rispetto a $\mathcal{B}(0,1)$;

(ii) Per ogni $x \in \mathbb{R}$ risulta

$$\begin{aligned} F_X(x) &= \mathbb{P}(\omega \in \Omega : \ X(\omega) \in (-\infty, x]) = \lambda(\{u \in (0,1) : \ \varphi(u) \leq x\}) && \text{(per definizione di } X \text{ e di } (\Omega, \mathcal{F}, P)) \\ &= \lambda(\{u \in (0,1) : \ u \leq F(x)\}) && \text{(per l'affermazione (2.1))} \\ &= \lambda\left((0,1) \cap (-\infty, F(x)]\right) = \begin{cases} \lambda\left((0, F(x)]\right), & \text{se } F(x) < 1, \\ \lambda(0,1), & \text{se } F(x) = 1, \end{cases} \end{aligned}$$

ovvero

$$F_X(x) = F(x).$$

Si tratta dunque di provare l'affermazione (2.1). Dimostriamo innanzitutto che

$$\varphi(u) = \min\{y : \ F(y) \geq u\}, \qquad \text{cioè} \qquad F(\varphi(u)) \geq u. \tag{2.2}$$

Infatti, essendo $\varphi(u) = \inf\{y : \ F(y) \geq u\}$, esiste una successione $\{y_n\}_{n=0}^{\infty}$ tale che:
(a) $F(y_n) \geq u$ per ogni n, e quindi $y_n \geq \varphi(u)$,
(b) $\{y_n\}$ tende a $\varphi(u)$ per $n \to \infty$.
Allora, poiché F è continua a destra,

$$F(\varphi(u)) = \lim_{n\to\infty} F(y_n) \geq u.$$

Possiamo ora mostrare l'uguaglianza in (2.1)

- Prima facciamo vedere che $\{u \in (0,1) : \ \varphi(u) \leq x\} \subseteq \{u \in (0,1) : \ u \leq F(x)\}$, mostrando che, per ogni u in $(0,1)$

  $$\text{se } \varphi(u) \leq x \quad \text{allora} \quad u \leq F(x).$$

  E infatti, poiché F è non decrescente, se $\varphi(u) \leq x$, allora $F(\varphi(u)) \leq F(x)$ e quindi per la (2.2)

  $$u \leq F(\varphi(u)) \leq F(x).$$

- Proviamo ora l'inclusione opposta $\{u \in (0,1) : \ \varphi(u) \leq x\} \supseteq \{u \in (0,1) : \ u \leq F(x)\}$, mostrando che, per ogni u in $(0,1)$,

  $$\text{se } u \leq F(x), \quad \text{allora} \quad \varphi(u) \leq x.$$

  Infatti, se $u \leq F(x)$, allora $x \in \{y : \ F(y) \geq u\}$ e quindi $\inf\{y : \ F(y) \geq u\} \leq x$, cioè $\varphi(u) \leq x$.

Prima di terminare la dimostrazione lasciamo al lettore il compito di osservare che fino ad ora non abbiamo (esplicitamente) usato la proprietà **P3** di normalizzazione. Tuttavia tale proprietà serve per garantire che la funzione φ sia a valori reali. □

Dall'affermazione (2.1) segue anche il seguente Corollario.

Corollario 2.2. *Sia U una v.a. uniformemente distribuita*[1] *in $(0,1)$. Allora $X := \varphi(U)$ ha distribuzione F*

[1] Ricordiamo che $U \sim$Unif$(0,1)$ è una v.a. che ha densità

$$f(t) := \begin{cases} 1, & 0 < t < 1 \\ 0, & \text{altrove} \end{cases}$$

che è la derivata della funzione

$$F(t) \equiv F_U(t) = \begin{cases} 0, & t < 0 \\ t, & 0 \leq t \leq 1 \\ 1, & t > 1 \end{cases}$$

Per una funzione con questa densità vale

$$\mathbb{P}(U \in [a,b]) = \int_a^b dt = b - a$$

quantità che dipende solo dall'ampiezza dell'intervallo, il che spiega la dizione *uniforme*

Dimostrazione. Infatti

$$\begin{aligned}
\mathbb{P}(X \leq x) &= \mathbb{P}(\varphi(U) \leq x) = \mathbb{P}(U \in \{u \in (0,1) : \varphi(u) \leq x\}) && \text{(per l'affermazione (2.1))} \\
&= \mathbb{P}(U \in \{u \in (0,1) : u \leq F(x)\}) = F(x). && \text{poiché } U \sim \text{Unif}(0,1)
\end{aligned}$$

□

Sempre la stessa proprietà (2.1) ci garantisce che la funzione $\widetilde{X}$ è crescente:

Osservazione 2.1 (Crescenza di $\widetilde{X}$)**.** *Sia* $\omega \leq \omega'$*, vogliamo mostrare che* $\widetilde{X}(\omega) = \min\{x : F(x) \geq \omega\} \leq \widetilde{X}(\omega')$*.*

La relazione (2.1) *garantisce che* $\widetilde{X}(\omega') = \min\{x : F(x) \geq \omega'\}$ *e quindi si ha che*

$$F\big(\widetilde{X}(\omega')\big) \geq \omega'.$$

Quindi (in quanto $\omega \leq \omega'$*)*

$$F\big(\widetilde{X}(\omega')\big) \geq \omega' \geq \omega,$$

e cioè $\widetilde{X}(\omega') \in \{x : F(x) \geq \omega\}$ *da cui*

$$\widetilde{X}(\omega) = \min\{x : F(x) \geq \omega\} \leq \widetilde{X}(\omega').$$

Osservazione 2.2 (spazi completi)**.** *Terminiamo questa sezione con una osservazione importante, che riguarda la possibilità di considerare spazi di probabilità completi, cioè di spazi che contengono anche gli insiemi trascurabili*

$$\mathcal{N} = \{A \subset \Omega : \forall \epsilon \exists B_\epsilon \in \mathcal{F}, \text{ tale che } A \subseteq B_\epsilon, \text{ e } \mathbb{P}(B_\epsilon) \leq \epsilon\},$$

ossia i sottoinsiemi degli insiemi di misura nulla: infatti, se A *è trascurabile allora esiste una successione di eventi* $B_{1/n}$ *tali che, per ogni* n*,* $B_{1/n}$ *contiene* A *e con probabilità minore uguale a* $1/n$*. Non si lede in generalità a supporre che* $B_{1/n}$ *sia una successione monotona. Di conseguenza* $A \subseteq B := \cap_{n=1}^{\infty} B_{1/n}$ *e* $\mathbb{P}(B) = 0$*.*

È noto che la misura di Lebesgue si può costruire (estendendo la misura definita sull'algebra delle unioni finite di intervalli) su spazi completi, e sulla σ*-algebra* $\mathcal{L}(0,1)$ *degli insiemi Lebesgue misurabili, che è appunto il completamento della* σ*-algebra* $\mathcal{B}(0,1)$ *dei boreliani. Tutte le funzioni boreliane sono ovviamente* $\mathcal{L}$*-misurabili (cioè misurabili secondo Lebesgue). Quindi la variabile aleatoria definita nel teorema di Skorohod è ancora misurabile se consideriamo* $\big((0,1), \mathcal{L}(0,1), \lambda\big)$*, invece di* $\big((0,1), \mathcal{B}(0,1), \lambda\big)$*, e quindi possiamo affermare che il teorema di Skorohod assicura che, data* F *che soddisfa le proprietà* ***P0 - P3*** *esiste uno spazio completo, dove è possibile definire una variabile aleatoria* X *con* $F_X = F$*.*

2.2 Costruzione di una succesione di variabili aleatorie indipendenti

L'affermazione che una successione di variabili aleatorie $\{X_n\,,\ n \in \mathbb{N}\}$ è una successione di v.a. indipendenti con $\mu_{X_n} = \mu_n$, è un'affermazione che riguarda le distribuzioni finito dimensionali del processo $\{X_n\,,\ n \in \mathbb{N}\}$. L'esistenza di una tale successione si potrebbe quindi dedurre dal teorema di rappresentazione di Kolmogorov, o magari da un risultato *ad hoc* la cui prova fosse la semplificazione del procedimento usato nel dimostrare tale teorema. Tuttavia l'esistenza di una tale successione tuttavia si può dedurre direttamente, pur di dare per scontato che esiste la misura di Lebesgue su $(0,1)$. Infatti su $(0,1)$ si possono definire delle variabili aleatorie indipendenti ed identicamente distribuite, a valori nell'insieme $\{0,1\}$, e che assumono il valore 0 con probabilità $1/2$ (lo stesso vale per il valore 1). A partire da questa successione di variabili aleatorie si può costruire una successione di variabili aleatorie $\{U_j\,,\ j \in \mathbb{N}\}$ indipendenti ed uniformi in $(0,1)$, come descritto qui di seguito. Infine, posto $F_n(x) = \mu_n\big((-\infty,x]\big)$, la successione cercata è data dalla successione delle v.a. $F_n^{-1}(U_n)$.

Lemma 2.3 (Successioni di v.a. indipendenti uniformi in $(0,1)$: esistenza)**.** *Nello spazio $\Omega = (0,1)$ con la misura di Lebesgue sui boreliani, è possibile avere una successione di v.a. uniformi in $(0,1)$ ed indipendenti.*

Per costruire tale successione si ricordi che scrivendo $\omega \in (0,1)$ in forma diadica

$$\omega = \sum_{1}^{\infty} W_i(\omega)\frac{1}{2^i},$$

le v.a. W_i risultano indipendenti e $\mathbb{P}(W_i = 0) = \mathbb{P}(W_i = 1) = \frac{1}{2}$. La successione U_n di v.a. uniformi ed indipendenti si può costruire, a partire dalle v.a. $\{W_i\}$, riordinandole in modo che formino una sequenza a doppio indice $\{\widetilde{W}_{i,n}\}$ così da poter definire

$$U_n(\omega) = \sum_{i=0}^{\infty} \widetilde{W}_{i,n}(\omega)\,\frac{1}{2^i}.$$

Ad esempio si può prendere $\widetilde{W}_{i,n} = W_{2^{i-1}(2n+1)}$, che corrisponde a riordinare la successione $\{W_i\}$ in questo modo:

$$\begin{array}{cccccc}
W_1 & W_3 & W_5 & W_7 & W_9 & \cdots \\
W_2 & W_6 & W_{10} & W_{14} & \cdots & \\
W_4 & W_{12} & W_{20} & W_{28} & \cdots & \\
W_8 & W_{24} & W_{40} & \cdots & & \\
\vdots & \vdots & \vdots & & &
\end{array}$$

ottenendo da $\{W_i\}$ infinite sottosuccessioni (corrispondenti alle colonne di questa matrice) in modo tale che nessuna W_i venga tralasciata né ripetuta.

****

Rimane da osservare che, per ogni n, la variabile aleatoria U_n è misurabile secondo $\widetilde{\mathcal{F}}_n := \sigma(\widetilde{W}_{i,n}, i \geq 1)$, e che tali sigma algebre sono indipendenti[2]. Quindi le variabili aleatorie U_n formano una succesione di variabili aleatorie indipendenti.

Inoltre poiché le variabili aleatorie $U_n(\omega)$ sono definite attraverso le $\widetilde{W}_{k,n}$, $k \geq 1$, esattamente come la variabile aleatoria $U(\omega) = \omega$ è costruita attraverso le $W_k(\omega)$, $k \geq 1$, e, per ogni n, la successioni di variabili aleatorie $\widetilde{W}_{k,n}$, $k \geq 1$ è una successione di variabili aleatorie indipendenti identicamente distribute (esattamente come $W_k(\omega)$, $k \geq 1$) anche la legge di U_n è la stessa di U e cioè uniforme su $(0,1)$.

Osservazione 2.3. ***Ribadiamo che la precedente costruzione è riportata affinché sia chiaro che l'affermazione che esiste una successione di v.a. indipendenti ed uniformi in $(0,1)$, è vera.*

A questo punto, data una successione di funzioni di distribuzione F_n, per costruire una successione di variabili aleatorie indipendenti Y_n e con funzione di distribuzione F_n, basta mettersi nello spazio $\big((0,1), \mathcal{B}(0,1), \lambda|_{(0,1)}\big)$, definire

$$Y_n(\omega) = \varphi_{F_n}\big(U_n(\omega)\big),$$

[2] Si confronti il Corollario 2 pag. 50 di [1] oppure, meglio il Teorema 20.2 pag. 268.

dove le U_n sono le variabili aleatorie definite sopra, e infine usare il Corollario 2.2, per ottenere che Y_n ha funzione di distribuzione $F_{Y_n} = F_n$, e infine, grazie all'osservazione che Y_n è misurabile rispetto alla sigma algebra generata da U_n, dedurre che le Y_n formano una successione di variabili aleatorie indipendenti.

2.3 Convergenza per variabili aleatorie

Esistono vari tipi di convergenza per variabili aleatorie. Noi ci occuperemo principalmente di tre tipi di convergenza:

1 CONVERGENZA QUASI CERTA
mettere definizione

2 CONVERGENZA IN PROBABILITÀ
mettere definzione

3 CONVERGENZA IN LEGGE O IN DISTRIBUZIONE (DETTA ANCHE DEBOLE)
mettere definizione

Valgono le seguenti implicazioni

$$\begin{array}{c} X_n \underset{n\to\infty}{\longrightarrow} X \quad \mathbb{P}-q.c. \\ \Downarrow \\ X_n \underset{n\to\infty}{\overset{in\,Pr}{\longrightarrow}} X \\ \Downarrow \\ X_n \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} X \end{array}$$

Nel caso in cui X è una variabile aleatoria degenere (ossia se esiste un $a \in \mathbb{R}$ tale che $\mathbb{P}(X = a) = 1$) allora la convergenza in probabilità equivale alla convergenza in legge:

$$\begin{array}{c} X_n \underset{n\to\infty}{\overset{in\,Pr}{\longrightarrow}} a \\ \Updownarrow \\ X_n \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} a \end{array}$$

Infine vale anche una specie di implicazione inversa, nel senso specificato dal teorema di immersione di Skorohod (vedi il successivo Teorema 2.4)

$$\begin{array}{c} X_n \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} X \\ \Downarrow \\ \exists\,(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}), \quad \widetilde{X}_n,\, \widetilde{X},\ \text{with}\ \widetilde{X}_n \overset{\mathcal{L}}{=} X_n,\ \widetilde{X} \overset{\mathcal{L}}{=} X \end{array}$$

e tali che

$$\widetilde{X}_n \underset{n\to\infty}{\longrightarrow} \widetilde{X} \quad \widetilde{\mathbb{P}}-q.c.$$

Teorema 2.4 (Teorema di rappresentazione di Skorohod per successioni)**.** *Se la successione di variabili aleatorie X_n converge in legge ad X, allora esiste uno spazio di probabilità $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$, e una successione di variabili aleatorie $\widetilde{X}_n$ e una variabile aleatoria $\widetilde{X}$ su tale spazio, con la stessa legge di X_n e di X, rispettivamente, e tali che*

$$\widetilde{X}_n \underset{n\to\infty}{\longrightarrow} \widetilde{X} \quad \widetilde{\mathbb{P}}-q.c.$$

Dimostrazione. Come nel teorema di rappresentazione di Skorohod (Teorema 2.1) lo spazio di probabilità è $\widetilde{\Omega} = (0, 1)$ con $\widetilde{\mathcal{F}} = \mathcal{B}(0, 1)$, i boreliani di $(0, 1)$ e $\widetilde{\mathbb{P}}$ è la misura di Lebesgue ristretta ai boreliani di $(0, 1)$. Inoltre, posto F_n ed F le funzioni di distribuzione di X_n ed X, rispettivamente, le variabili aleatorie sono

$$\widetilde{X}_n(\omega) = \varphi_{F_n}(\omega) = \inf\{y :\ F_n(y) \geq \omega\}, \qquad \widetilde{X}(\omega) = \varphi_F(\omega) = \inf\{y :\ F(y) \geq \omega\}.$$

Dalla definizione di $\widetilde{X}$ e dalla (2.1) sappiamo che

$$\widetilde{X}(\omega) \leq x \Leftrightarrow F(x) \geq \omega \quad \text{(e quindi } \widetilde{X}(\omega) > x \Leftrightarrow F(x) > \omega).$$

Analogamente possiamo affermare che

$$\widetilde{X}_n(\omega) \le x \Leftrightarrow F_n(x) \ge \omega \quad \text{(e quindi } \widetilde{X}_n(\omega) > x \Leftrightarrow F_n(x) > \omega\text{)}.$$

Questa osservazione sarà fondamentale per dimostrare le seguenti due relazioni

$$\textbf{(I)} \qquad \forall \omega \in (0,1), \qquad \underline{\lim}_{n\to\infty} \widetilde{X}_n(\omega) \ge \widetilde{X}(\omega), \tag{2.3}$$

$$\textbf{(II)} \qquad \forall \omega,\, \omega' \in (0,1), \omega < \omega' \qquad \overline{\lim}_{n\to\infty} \widetilde{X}_n(\omega) \le \widetilde{X}(\omega') \tag{2.4}$$

A loro volta le precedenti affermazioni, insieme al fatto che $\widetilde{X}$ è una funzione non descrescente (ossia crescente in senso lato) e quindi ammette al più un'infinità numerabile di punti di discontinuità, permettono di arrivare alla dimostrazione:

Dalle (2.3) e (2.4) sappiamo che, per ogni $\omega < \omega' \in (0,1)$ vale

$$\widetilde{X}(\omega) \le \underline{\lim}_{n\to\infty} \widetilde{X}_n(\omega) \le \overline{\lim}_{n\to\infty} \widetilde{X}_n(\omega) \le \widetilde{X}(\omega')$$

e quindi

$$\widetilde{X}(\omega) \le \underline{\lim}_{n\to\infty} \widetilde{X}_n(\omega) \le \overline{\lim}_{n\to\infty} \widetilde{X}_n(\omega) \le \lim_{\omega'\to\omega^+} \widetilde{X}(\omega').$$

La tesi segue allora in quanto le disuguaglianze sono tutte uguaglianze se ω è un punto di continuità per $\widetilde{X}$:

$$\widetilde{X}(\omega) \le \underline{\lim}_{n\to\infty} \widetilde{X}_n(\omega) \le \overline{\lim}_{n\to\infty} \widetilde{X}_n(\omega) \le \lim_{\omega'\to\omega^+} \widetilde{X}(\omega') = \widetilde{X}(\omega)$$

Quindi $\widetilde{\mathbb{P}} - q.c.$ esiste il limite di $\widetilde{X}_n(\omega)$ e coincide con $\widetilde{X}(\omega)$.

Per finire la dimostrazione non rimane che verificare le relazioni (2.3) e (2.4).

(I) Cominciamo con il fissare $\omega \in (0,1)$ e scegliere un $\varepsilon > 0$ e prendere[3] un x di continutità per F e tale che $\widetilde{X}(\omega) - \varepsilon < x < \widetilde{X}(\omega)$. Come abbiamo visto, la disuguaglianza $\widetilde{X}(\omega) > x$ implica che

$$F(x) > \omega.$$

Di conseguenza, poiché sappiamo che $F_n(x)$ converge a $F(x)$, per n sufficientemente grande possiamo affermare che

$$F_n(x) > \omega.$$

La precedente relazione ci assicura che

$$X_n(\omega) > x,$$

e quindi

$$\underline{\lim}_{n\to\infty} X_n(\omega) \ge x > \widetilde{X}(\omega) - \varepsilon.$$

Per l'arbitarietà di ε (mandando ε a zero) otteniamo quindi che

$$\underline{\lim}_{n\to\infty} X_n(\omega) \ge \widetilde{X}(\omega),$$

ossia proprio la (2.3).

(II) Per ottenere la (2.4), dobbiamo sempre fissare $\omega \in (0,1)$, ma inoltre dobbiamo considerare[4] anche un $\omega' \in (0,1)$ con $\omega < \omega'$. Comunque scelto $\varepsilon > 0$ possiamo prendere un y tale che $F(y) = F(y^-)$ e

$$\widetilde{X}(\omega) \le \widetilde{X}(\omega') < y < \widetilde{X}(\omega') + \varepsilon,$$

[3]Un tale x esiste di sicuro, perché il numero dei punti di discontinuità di F, che è una funzione monotona, sono al massimo un'infinità numerabile.

[4]Lo scopo è ottenere che esista un y di continuità per F e tale che $F(y) > \omega$, per poter garantire poi che, per n sufficientemente grande $\omega \le F_n(y)$ ossia che $\widetilde{X}_n(\omega) \le y$. Purtroppo il fatto di sapere che $\widetilde{X}(\omega) < y$ implica solo che $F(y) \ge y$ e ciò a sua volta non ci garantisce che $F_n(y) \ge y$, per n sufficientemente grande...

(dove la prima disuguaglianza deriva semplicemente dal fatto che $\widetilde{X}(\omega)$ è una funzione non descrescente[5].
La disuguaglianza $\widetilde{X}(\omega') < y$ implica che

$$F\big(\widetilde{X}(\omega')\big) \leq F(y).$$

Inoltre sappiamo che $F\big(\widetilde{X}(\omega')\big) \geq \omega'$, e quindi

$$\omega < \omega' \leq F\big(\widetilde{X}(\omega')\big) \leq F(y).$$

Il fatto che

$$\omega < F(y),$$

garantisce che definitivamente (ossia per n sufficientemente grande)

$$\omega < F_n(y),$$

da cui

$$\omega \leq F_n(y), \quad \leftrightarrow \widetilde{X}(\omega) \leq y.$$

Poiché sappiamo che $y < \widetilde{X}(\omega') + \varepsilon$ otteniamo quindi

$$\widetilde{X}(\omega) \leq y < \widetilde{X}(\omega') + \varepsilon.$$

A questo punto possiamo affermare che

$$\overline{\lim}_{n\to\infty} \leq \widetilde{X}(\omega') + \varepsilon,$$

e per ottenere la (2.4), basta mandare ε a zero e ω' a ω per ottenere che

$$\overline{\lim}_{n\to\infty} \leq \widetilde{X}(\omega^+).$$

La funzione $\widetilde{X}$ è non decrescente (ossia crescente in senso lato) e quindi ammette al più un'infinità numerabile di punti in cui è discontinua. Da cui si otteniene

$$\overline{\lim}_{n\to\infty} \leq \widetilde{X}(\omega),$$

per ogni $\omega \in (0,1)$, esclusi al più un'inifinità numerabile e quindi quasi ovunque rispetto alla misura di Lebesgue, che appunto è la nostra $\widetilde{\mathbb{P}}$.

$\square$

[5] Si ricordi l'Osservazione 2.1.

Bibliografia

[1] Billingsley, P. *Probability and measure*, third ed. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1995. A Wiley-Interscience Publication.

[2] Koch, G. *La matematica del probabile*, prima ed. Aracne, Roma, 1997.

[3] Williams, D. *Probability with martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Capitolo 3

ESERCIZI PROPOSTI

In questo capitolo sono raccolti alcuni esercizi da scolgere: alcuni sono di tipo analitico e sono parte integrante di alcune dimostrazioni fndamentali, alcuni sono stati parzialmente o totalmente svolti a lezione.

Lo studente dovrà consegnare gli esercizi e discuterne durante l'esame orale, un paio a scelta del docente.

Sarà possibile, limitatamente alla sessione della prova in itinere (novembre 2009) portare gli esercizi scritti anche dopo l'esame orale. In tal caso la discussione degli esercizi avverrà alla consegna degli esercizi.

3.1 Esercizi di tipo analitico

Esercizio 3.1. *Dimostrare che se*

$$\lim_{n\to\infty} a_n = \alpha,$$

allora

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} a_k = \alpha.$$

Esercizio 3.2. *Dimostrare che, per $\alpha > 1$*

$$\lim_{n\to\infty} \frac{\sum_{i=n}^{\infty} i^{-\alpha}}{\frac{n^{1-\alpha}}{\alpha-1}} = 1,$$

e che, per $\beta > -1$,

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} i^{\beta}}{\frac{n^{1+\beta}}{1+\beta}} = 1$$

Esercizio 3.3. *Mostrare che*

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \leq \min\left(\frac{|x|^{n+1}}{(n+1)!}, 2\,\frac{|x|^n}{n!} \right)$$

Suggerimento: *utilizzare la formula*[1]

$$e^{ix} = \sum_{k=0}^{m} \frac{(ix)^k}{k!} + \frac{i^{m+1}}{m!} \int_0^x (x-s)^m \, e^{is} \, ds, \qquad (*)$$

con $m = n$ per ottenere

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \leq \frac{|x|^{n+1}}{(n+1)!},$$

e per $m = n-1$, insieme all'osservazione che

$$\int_0^x (x-s)^m \, e^{is} \, ds = \int_0^x (x-s)^m \, ds + \int_0^x (x-s)^m \, (e^{is}-1) \, ds = \frac{x^{m+1}}{m+1} + \int_0^x (x-s)^m \, (e^{is}-1) \, ds$$

e che $|e^{is}-1| \leq 2$, per ottenere

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \leq 2 \, \frac{|x|^n}{n!}.$$

Esercizio 3.4. *Mostrare che, per $x > 0$*

$$|e^{-x} - (1-x)| \leq \frac{x^2}{2}.$$

3.2 Esercizi sulla convergenza in distribuzione

Esercizio 3.5. *Siano*

$$F_n(x) = \int_{-\infty}^{x} p_n(t) \, dt, \qquad \text{per } n \geq 1,$$

dove

$$p_n(x) = \mathbf{1}_{(0,1)}(x)\big(1 + \cos(2\pi n x)\big)$$

Dimostrare che F_n sono funzioni di distribuzioni e che

$$\lim_{n\to\infty} F_n(x) = F(x) = \int_{-\infty}^{x} p(t) \, dt,$$

dove

$$p(x) = \mathbf{1}_{(0,1)}(x),$$

e che invece

$$p_n(x) \nrightarrow p(x).$$

Discutere la relazione di questo esempio con il Teorema di Scheffé.

Esercizio 3.6. *Sia X uniforme in $(0,1)$ e siano $X_n := X^n$. Indivuduare le funzioni di distribuzione F_n delle variabili aleatorie X_n. Dimostrare che la successione di variabili aleatorie X_n converge in distribuzione e individuare la variabile aleatoria limite. La successione X_n converge anche in qualche altro senso?*

---

[1] La formula $(*)$ è sostanzialmente la formula del resto nello sviluppo di Taylor, e si può facilmente ottenere per induzione dalla formula

$$e^{ix} = 1 + i \int_0^x e^{iy_1} \, dy_1.$$

Come esempio si consideri

$$\begin{aligned} e^{ix} &= 1 + i \int_0^x e^{iy_1} \, dy_1 = 1 + i \int_0^x \left(1 + i \int_0^{y_1} e^{iy_2} \, dy_2\right) dy_1 = 1 + ix + i^2 \int_0^x \left(\int_0^{y_1} e^{iy_2} \, dy_2\right) dy_1 \\ &= 1 + ix + i^2 \int_0^x \left(\int_{y_2}^{x} e^{iy_2} \, dy_1\right) dy_2 = 1 + ix + i^2 \int_0^x (x - y_2) \, e^{iy_2} dy_2. \end{aligned}$$

Esercizio 3.7. *Sia X uniforme in $(\frac{1}{2}, \frac{3}{2})$ e siano $X_n := X^n$. Individuare le funzioni di distribuzione F_n delle variabili aleatorie X_n. La successione di variabili aleatorie X_n converge in distribuzione? e se sì, a quale variabile aleatoria? e se no, la successione X_n è una successione tight (ovvero trattenuta)?*

3.3 Esercizi con le funzioni caratteristiche

Esercizio 3.8. *(a) Per $n \geq 1$, sia X_n una variabile aleatoria esponenziale di parametro n. La successione di variabili aleatorie X_n converge in distribuzione? e se sì, a quale variabile aleatoria?*
(b) Per $n \geq 1$, sia Y_n una variabile aleatoria Gamma di parametri (n, n). La successione di variabili aleatorie X_n converge in distribuzione? e se sì, a quale variabile aleatoria?

Suggerimento: *usare le funzioni caratteristiche.*

Esercizio 3.9 (esercizio svolto a lezione). *Siano X ed Y due variabili aleatorie indipendenti, entrambe esponenziali di parametro 1.*

(a) Dimostrare che la variabile aleatoria $Z = X - Y$ ammette densità

$$p_Z(z) = \frac{1}{2}\, e^{-|z|}.$$

Suggerimento: *utilizzare il fatto che, se le variabili aleatorie sono indipendenti e ammettono densità, allora la densità della somma è la convoluzione.*
(b) Dimostrare che la funzione caratteristica di $Z = X - Y$ è

$$\varphi_Z(t) = \frac{1}{1+t^2}.$$

(c) Utilizzare la formula di inversione per le funzioni caratteristiche integrabili e dedurre che l'espressione della funzione caratteristica di una variabile aleatoria V con distribuzione di Cauchy (cioè con densità $\frac{1}{\pi}\frac{1}{1+x^2}$) è

$$\varphi_V(t) = e^{-|t|}.$$

(d) Dedurre, dalla forma della funzione caratteristica di una variabile aleatoria V con distribuzione di Cauchy, che non esiste finito $\mathbb{E}\big[|V|\big]$.

Esercizio 3.10 (parzialmente svolto). *Sia X una variabile aleatoria con densità*

$$p_X(x) = \begin{cases} 0 & \textit{per } |x| \leq 2, \\ c\, \frac{1}{x^2 \log|x|} & \textit{per } |x| > 2, \end{cases}$$

*dove $c > 0$ è la costante di normalizzazione (**N.B.** la costante c non deve essere calcolata)*

(i) Dimostrare che X non è integrabile (ossia che $\mathbb{E}[|X|] = \infty$*).*

Suggerimento: *utilizzare il fatto che la densità è una funzione pari.*

(ii) Dimostrare che $\varphi_X(t)$ ammette derivata prima in $t = 0$ e che $\varphi'_X(0) = 0$.
Soluzione: (ATTENZIONE leggermente diverso da come fatto a lezione)
utilizzando il fatto che la densità è una funzione pari il rapporto incrementale $\frac{\varphi_X(h)-1}{h}$ si può scrivere come un integrale su $(2,\infty)$ rispetto a $p_X(x)\,dx$, in cui i termini:

$$\frac{\varphi_X(h)-1}{h} = c \int_{\{x:|x|>2\}} \frac{\cos hx - 1}{h}\, \frac{1}{\log|x|}\, \frac{1}{x^2}\, dx$$

Quindi, per mostrare che

$$\frac{\varphi_X(h)-1}{h}\underset{h\to 0}{\longrightarrow} 0 \quad \Leftrightarrow \quad \left|\frac{\varphi_X(h)-1}{h}\right|\underset{h\to 0}{\longrightarrow} 0$$

e tenendo conto che

$$\left|c\int_{\{x:|x|>2\}}\frac{\cos hx-1}{h}\,\frac{1}{\log|x|}\,\frac{1}{x^2}\,dx\right|\leq c\int_{\{x:|x|>2\}}\left|\frac{\cos hx-1}{h}\right|\frac{1}{\log|x|}\,\frac{1}{x^2}\,dx=2c\int_2^\infty\frac{1-\cos|hx|}{|h|}\,\frac{1}{\log|x|}\,\frac{1}{x^2}\,dx,$$

basta mostrare che

$$\int_2^\infty\frac{1-\cos|h|x}{|h|}\,\frac{1}{\log|x|}\,\frac{1}{x^2}\,dx\underset{h\to 0}{\longrightarrow} 0.$$

Effettuando il cambio di variabile di integrazione $y=|h|x$ si ottiene che

$$\int_2^\infty\frac{1-\cos|h|x}{|h|}\,\frac{1}{\log|x|}\,\frac{1}{x^2}\,dx=\int_{2|h|}^\infty\frac{1-\cos y}{|h|}\,\frac{1}{\log(y/|h|)}\,\frac{1}{\frac{y^2}{|h|^2}}\,\frac{1}{|h|}dy=\int_{2|h|}^\infty(1-\cos y)\,\frac{1}{\log(y/|h|)}\,\frac{1}{y^2}\,dy.$$

Senza ledere in generalità possiamo assumere $|h|<1$ e che quindi, per $|y|\geq 2|h|$, si ha $\frac{1}{\log|y/h|}=\frac{1}{\log|y|+|\log|h||}\to 0$, per h che tende a zero.

Di conseguenza, posto $g_{|h|}$ la funzione

$$g_{|h|}(y)=\mathbf{1}_{\{|y|\geq 2|h|\}}(1-\cos y)\,\frac{1}{\log(y/|h|)}\,\frac{1}{y^2},\qquad y>0,$$

si ha che

$$g_{|h|}(y)\underset{h\to 0}{\longrightarrow} 0.$$

Inoltre, considerato che

$$\frac{1}{\log|y/h|}\leq\frac{1}{\log 2},\quad \textit{per } y\geq 2|h|,$$

$$0\leq 1-\cos y=2\big(\sin(y/2)\big)^2\leq 2\left(\frac{y}{2}\right)^2=\frac{1}{2}\,y^2,\quad \textit{per ogni } y$$

e

$$0\leq 1-\cos y\leq 2,\quad \textit{per ogni } y,$$

possiamo affermare che

$$g_{|h|}(y)\leq g(y)=\min\left(\frac{1}{2\log 2},\,\frac{2}{\log 2}\,\frac{1}{y^2}\right),\qquad y>0$$

Per ottenere la tesi basta infine osservare che

$$\int_0^\infty g(y)dy<\infty$$

e utilizzare il teorema della convergenza dominata su $(0,\infty)$.

Osservazione: *questo esempio mostra che, se la funzione caratteristica di una variabile aleatoria X ammette una derivata di ordine k dispari, non è detto che esista finito il momento di ordine k, cioè che può accadere che $\mathbb{E}[|X|^k]=\infty$.*

3.4 Esercizi sulla condizione di Lindeberg

Esercizio 3.11. *Siano X_n variabili aleatorie indipendenti, discrete e tali che*

$$\mathbb{P}(X_n=\pm c_n)=\frac{1}{2},$$

per una succesione c_n di numeri positivi. Dimostrare che la condizione di Lindeberg

$$\forall\varepsilon>0,\qquad \lim_{n\to\infty}\sum_{k=1}^n\frac{1}{s_n^2}\,\mathbb{E}[|X_k|^2\mathbf{1}_{\{|X_k|>\varepsilon\, s_n\}}]=0,$$

dove $s_n^2 = Var(X_1 + \cdots + X_n)$, *equivale a*

$$\lim_{n\to\infty} \frac{\max_{k=1,\ldots,n} c_k^2}{\sum_{h=1}^n c_h^2} = 0.$$

Considerare il caso $c_n = n^\alpha$, *per* $\alpha > 0$.

Esercizio 3.12. *Dimostrare che la condizione di Lindeberg non è necessaria affinché per una successione di variabili aletaorie indipendenti valga il Teorema Centrale del Limite (ossia che la somma delle prime n variabili aleatoria, opportunamente standardizzata, converga in distribuzione ad una variabile gaussiana standard)*

Suggerimento: *considerare* X_n *una successione di variabili aleatorie indipendenti, gaussiane di media zero e varianza* σ_n^2, *e scegliere* σ_n^2 *in modo che, posto*

$$s_n^2 = \sum_{k=1}^n \sigma_k^2$$

non valga la condizione

$$\lim_{n\to\infty} \frac{\max_{k=1,\ldots,n} \sigma_k^2}{s_n^2} \neq 0.$$

3.5 Esercizi sulla legge dei grandi numeri

Esercizio 3.13 (esercizio parzialmente svolto a lezione). *Siano* X_n *variabili aleatorie con* $\mathbb{P}(X_n = \pm n^\alpha) = \frac{1}{2}$.

(a) Mostrare che per $\alpha < 0$ *vale la legge dei grandi numeri (**N.B.** senza assumere l'indipendenza delle variabili aleatorie* X_n*)*

Si assuma l'ulteriore ipotesi che le variabili aleatorie X_n *siano indipendenti.*

(b) Mostrare che per $\alpha \in [0, 1/2)$ *vale la condizione del Teorema di Kolmogorov, e quindi vale la legge dei grandi numeri.*

*(c) **(da svolgere dopo aver studiato il teorema centrale del limite, ed in particolare dopo aver svolto l'esercizio 3.11)***

Mostrare che, per $\alpha > 0$, *posto* $S_n = X_1 + X_2 + \ldots X_n$ *ed* $s_n^2 = Var(S_n)$ *si ha*

$$s_n^2 \sim \frac{n^{2\alpha+1}}{2\alpha+1} \quad \textit{cioè} \quad \frac{s_n^2}{\frac{n^{2\alpha+1}}{2\alpha+1}}$$

e che vale la condizione di Lindeberg.

(d) Come conseguenza del punto (c), mostrare che, per $\alpha = 1/2$, *la successione*

$$\frac{S_n}{n} = \frac{S_n}{s_n}\frac{s_n}{n}$$

converge ad una variabile aleatoria con distribuzione gaussiana e determinarne la media e la varianza.

(e) Invece, per $\alpha > 1/2$, *mostrare che la successione*

$$\frac{S_n}{n} = \frac{S_n}{s_n}\frac{s_n}{n}$$

non converge neanche in distribuzione.

Suggerimento: *si dimostri ad esempio che, posto* $Y_n := \frac{S_n}{s_n}$ *e* $\beta_n = \frac{s_n}{n}(> 0)$ *si ha ovviamente* $\beta_n \to \infty$ *e, per ogni* $x > 0$, *si ha* $x/\beta_n \to 0$. *Dedurre quindi che*

$$\overline{\lim}_{n\to\infty}\mathbb{P}(\beta_n Y_n \leq x) \leq \frac{1}{2}, \qquad \forall\, x > 0,$$

e che quindi la successione non può convergere in distribuzione.

Esercizio 3.14. *Siano* X_n *variabili aleatorie indipendenti identicamente distribuite, tutte con distribuzione di Cauchy.*

(a) Dimostrare che $(|X_1| + |X_2| + \ldots + |X_n|)/n$ *converge quasi certamente ad infinito, quindi per le variabili aleatorie* $|X_n|$ *vale la legge dei grandi numeri.*

(b) Dimostrare che $(X_1 + X_2 + \ldots + X_n)/n$ *converge in distribuzione ed individuare la distribuzione limite.*

Suggerimento: *Utilizzare le funzioni caratteristiche.*

3.6 Esercizio riassuntivo

Esercizio 3.15. *Utilizzare la rappresentazione della variabile aleatoria $U(\omega) = \omega$ in $\Omega = (0,1)$ [come al solito si assume anche $\mathcal{F}$ uguale ai boreliani di $(0,1)$ e la probabilità $\mathbb{P}$ uguale alla misura di Lebesgue ristretta a $(0,1)$]*

$$\omega = \sum_1^\infty W_i(\omega) \frac{1}{2^i},$$

dove W_i sono variabili aleatorie indipendenti identicamente distribuite, con $\mathbb{P}(W_i = 1) = \mathbb{P}(W_i = 0) = \frac{1}{2}$, per ottenere una dimostrazione probabilistica della relazione

$$\frac{\sin t}{t} = \prod_{i=1}^\infty \cos \frac{t}{2^i}.$$

Suggerimento: *utilizzare il fatto che U ha distribuzione uniforme su $(0,1)$, e che la variabile aleatoria*

$$V(\omega) := 2U(\omega) - 1 = 2\sum_1^\infty W_i(\omega)\frac{1}{2^i} - 1 = \sum_1^\infty (2\,W_i(\omega) - 1)\,\frac{1}{2^i}$$

e che quindi V è il limite quasi certo di

$$V_n(\omega) := \sum_1^n (2\,W_i(\omega) - 1)\,\frac{1}{2^i},$$

e che le variabili aleatorie $\overline{W}_i = 2\,W_i - 1$ sono indipendenti.