

## Riassunto delle cose principali

### 1. Variabili Aleatorie

$(\Omega, \mathcal{F})$  spazio di eventi con  $\mathcal{F}$   $\sigma$ -algebra su  $\Omega$

$\mathcal{O}$  famiglia degli insiemi aperti di  $\mathbb{R}$

$\sigma(\mathcal{O}) \equiv \mathcal{B}$  sigma algebra di Borel (in  $\mathbb{R}$ ), la più piccola che contiene  $\mathcal{O}$

Definizione:  $X: \Omega \rightarrow \mathbb{R}$  è una v.a. se  $X^{-1}(A) \in \mathcal{F}$  per ogni  $A \in \mathcal{B}$

Ricorda che  $X^{-1}(A) = \{\omega: X(\omega) \in A\}$ , che scriviamo anche come  $\{X \in A\}$ .

$\mathcal{J}$  famiglia degli intervalli  $(-\infty, b]$  con  $b \in \mathbb{R}$  (molto più “piccola” di  $\mathcal{O}$ )

Teorema:  $X^{-1}(A) \in \mathcal{F}$  per ogni  $A \in \mathcal{B} \iff X^{-1}(J) \in \mathcal{F}$  per ogni  $J \in \mathcal{J}$

**Corollario:**  $X$  è una v.a. se  $X^{-1}(J) \in \mathcal{F}$  per ogni  $J \in \mathcal{J}$  (def. di partenza del libro)

Ricorda che questo vuol dire  $\{X \leq b\} \equiv \{\omega \in \Omega: -\infty < X(\omega) \leq b\} \in \mathcal{F} \forall b \in \mathbb{R}$

Definizione:  $X$  è una v.a. *discreta* se  $X(\Omega)$  è finito o numerabile,  $X(\Omega) = \{x_1, \dots, x_n, \dots\}$

In tal caso basta che  $\{X = x_i\} \equiv \{\omega: X(\omega) = x_i\} \in \mathcal{F}$  per ogni  $x_i$ .

Dimostrazione:  $\{X \in A\} = \cup_{x_i \in A} \{X = x_i\}$  per ogni  $A \in \mathcal{B}$ .

### 2. Distribuzione di $X$

Dato uno spazio di probabilità  $(\Omega, \mathcal{F}, P)$ ,

una v.a.  $X: \Omega \rightarrow \mathbb{R}$  genera la distribuzione su  $\mathcal{B}$  (in  $\mathbb{R}$ !) definita da

$$P_X(A) := P(X \in A), \quad A \in \mathcal{B}.$$

Sia  $X$  **discreta**.

Poni  $p(x_i) = P_X(\{x_i\})$  - per essere precisi dovrebbe essere  $p_X(x_i)$

Da  $\{X \in A\} = \cup_{x_i \in A} \{X = x_i\}$  (unione disgiunta) otteniamo subito

$$P_X(A) = \sum_{x_i \in A} p(x_i) \quad A \in \mathcal{B}.$$

## Dettagli

Considera uno spazio di eventi  $(\Omega, \mathcal{F})$ , con  $\mathcal{F}$   $\sigma$ -algebra su  $\Omega$  (gli eventi sono gli insiemi in  $\mathcal{F}$ ). La famiglia di *tutti* i sottoinsiemi di  $\Omega$  può essere “troppo grande” (teorema di Vitali, 1905) quindi in generale  $\mathcal{F}$  sarà strettamente più piccolo dell’insieme potenza di  $\Omega$ .<sup>2</sup>

<sup>1</sup>S Modica 2023. Ref. Georgii, *Stochastics*, capitolo 1 e poco altro.

<sup>2</sup>Nel caso particolare di  $\Omega = \{\omega_1, \dots, \omega_n, \dots\}$  *numerabile* il problema non si pone; in quel caso come abbiamo visto una  $P$  su tutto  $\mathcal{P}(\Omega)$  resta definita allorché prediamo una sequenza  $p_n = P(\{\omega_n\})$  con  $p_n \geq 0, \sum_n p_n = 1$ , quindi possiamo prendere  $\mathcal{F} = \mathcal{P}(\Omega)$ . Per comodità useremo “numerabile” per dire “finito o numerabile”.

È importante capire che *la  $\mathcal{F}$  rappresenta l'informazione che abbiamo su  $\Omega$* , nel senso che osservato un qualunque  $\omega$ , per ogni  $E \in \mathcal{F}$  sappiamo se  $\omega \in E$  o  $\omega \in E^c$ ; in altre parole gli eventi in  $\mathcal{F}$  sono quelli che sappiamo se si sono verificati o no; e a quegli eventi possiamo assegnare probabilità. Per esempio se  $\Omega = [0, 1)$  ed  $\mathcal{F} = \{\emptyset, \Omega, [0, 1/2), [1/2, 1)\}$ , se si verifica un dato  $\omega$  sappiamo solo se  $\omega < 1/2$  oppure  $\omega \geq 1/2$ . Definendo una probabilità su  $\mathcal{F}$  - cioè assegnando una probabilità ad ogni evento in  $\mathcal{F}$  - otteniamo uno spazio di probabilità  $(\Omega, \mathcal{F}, P)$ .

## Variabili aleatorie a valori reali

Sia  $X: \Omega \rightarrow \mathbb{R}$  una funzione. Torniamo all'esempio  $\Omega = [0, 1)$  con  $\mathcal{F} = \{\emptyset, \Omega, [0, 1/2), [1/2, 1)\}$ , e considera la funzione

$$X(\omega) = \lfloor 3\omega \rfloor = \begin{cases} 0 & 0 \leq \omega < 1/3 \\ 1 & 1/3 \leq \omega < 2/3 \\ 2 & 2/3 \leq \omega < 1. \end{cases}$$

Nota che con l'informazione  $\mathcal{F}$  in nostro possesso non siamo in grado di osservare  $X$ ; per esempio se  $\omega \in [0, 1/2)$  non sappiamo se  $X = 0$  o  $X = 1$ , e analogamente se  $\omega \in [1/2, 1)$  non siamo in grado di dire se  $X = 1$  oppure  $X = 2$ . La restrizione che imporremo sulle funzioni  $X: \Omega \rightarrow \mathbb{R}$  è volta a escludere questi casi. Vogliamo che per insiemi di interesse  $A \subseteq \mathbb{R}$  sia  $X^{-1}(A) \equiv \{\omega: X(\omega) \in A\} \in \mathcal{F}$ ; in tal modo saremo in grado non solo di osservare se  $X(\omega) \in A$  o meno, ma anche di assegnare probabilità agli eventi  $X^{-1}(A)$ . Andiamo avanti sui dettagli.

Una funzione  $X: \Omega \rightarrow \mathbb{R}$  può rappresentare per esempio un guadagno, nel qual caso una cosa che ci interessa sapere è con che probabilità  $X(\omega) \geq -100$ ; o può essere un errore di misurazione, e vogliamo sapere con che probabilità  $-\epsilon < X(\omega) < \epsilon$ . In generale vogliamo poter assegnare probabilità ad eventi del tipo  $\{\omega: X(\omega) \in A\} \equiv X^{-1}(A)$  con  $A \subseteq \mathbb{R}$ , cioè vogliamo che  $X^{-1}(A) \in \mathcal{F}$ . Per *tutti* gli insiemi  $A$  di  $\mathbb{R}$ ? *No*, ma almeno per una famiglia di insiemi di  $\mathbb{R}$  "ricca abbastanza" da includere tutti gli insiemi di interesse - che contenga per esempio tutti gli insiemi aperti, chiusi, e loro unioni, intersezioni, complementi ecc. È universalmente accettato che una famiglia appropriata di insiemi è *la  $\sigma$ -algebra di Borel*, che è la più piccola  $\sigma$ -algebra in  $\mathbb{R}$  che contiene la famiglia  $\mathcal{O}$  degli insiemi aperti di  $\mathbb{R}$ ; si indica con  $\mathcal{B}(\mathbb{R})$  che abbrevieremo con  $\mathcal{B}$ . Poiché per ogni insieme di Borel vogliamo poter calcolare la probabilità  $P(X^{-1}(A))$  che  $X$  finisca in  $A$ , una variabile aleatoria (a valori reali) la definiamo così:

**Definizione 1.** Una variabile aleatoria su  $(\Omega, \mathcal{F})$  a valori reali è una funzione  $X: \Omega \rightarrow \mathbb{R}$  tale che  $X^{-1}(A) \in \mathcal{F}$  per ogni  $A \in \mathcal{B}$ .

È opportuno osservare che in generale dato un qualunque insieme  $E$  e una famiglia di insiemi  $\mathcal{G}$  in  $E$ , la più piccola  $\sigma$ -algebra in  $E$  contenente  $\mathcal{G}$  non è altro che l'intersezione di tutte le  $\sigma$ -algebre di  $E$  che contengono  $\mathcal{G}$ . Si indica con  $\sigma(\mathcal{G})$  e si dice *generata* da  $\mathcal{G}$ .<sup>3</sup>  $\sigma(\mathcal{G})$  contiene tutte le unioni e intersezioni numerabili e complementi di elementi di  $\mathcal{G}$ , unioni intersezioni e complementi di questi, ecc. ecc. Nel nostro caso  $\mathcal{B} = \sigma(\mathcal{O})$ .

Può essere istruttivo verificare che  $\mathcal{B}$  è generata anche dalla famiglia  $\mathcal{C}$  degli insiemi chiusi (dettagli in nota).<sup>4</sup> La cosa interessante è che la stessa  $\mathcal{B}$  è generata da una famiglia molto più piccola della famiglia di tutti i chiusi: la famiglia degli intervalli  $(-\infty, b]$  con  $b \in \mathbb{R}$ :

**Proposizione 1.**  $\mathcal{B} = \sigma(\mathcal{J})$  con  $\mathcal{J}$  famiglia degli intervalli  $(-\infty, b]$  con  $b \in \mathbb{R}$ .<sup>5</sup>

*Dimostrazione.* Dimostriamo (i)  $\mathcal{B} = \sigma(\mathcal{I})$  con  $\mathcal{I}$  famiglia degli intervalli  $(a, b)$ ; e (ii)  $\sigma(\mathcal{I}) = \sigma(\mathcal{J})$ . (i) Da  $\mathcal{I} \subseteq \mathcal{B}$  segue che  $\mathcal{B}$  è una  $\sigma$ -algebra che contiene  $\mathcal{I}$ , quindi  $\sigma(\mathcal{I}) \subseteq \mathcal{B}$ ; d'altra parte (ricorda da Analisi) ogni aperto in  $\mathbb{R}$  è unione numerabile di intervalli aperti  $(a, b)$ , quindi  $\mathcal{O} \subseteq \sigma(\mathcal{I})$ ; e da questo segue  $\mathcal{B} = \sigma(\mathcal{O}) \subseteq \sigma(\mathcal{I})$ . Quindi  $\mathcal{B} = \sigma(\mathcal{I})$ . (ii) Possiamo scrivere  $(-\infty, b] = \bigcup_n \bigcap_m (a_n, b + 1/m)$  con  $a_n \rightarrow -\infty$ , quindi  $\mathcal{J} \subseteq \sigma(\mathcal{I})$  il che implica  $\sigma(\mathcal{J}) \subseteq \sigma(\mathcal{I})$ ; d'altra parte  $(a, b] = (-\infty, b] \setminus (-\infty, a] \in \sigma(\mathcal{J})$ , e quindi anche  $(a, b) = \bigcup_n (a, b - 1/n] \in \sigma(\mathcal{J})$ ; ma  $\mathcal{I} \subseteq \sigma(\mathcal{J})$  implica  $\sigma(\mathcal{I}) \subseteq \sigma(\mathcal{J})$ .  $\square$

Un altro fatto rilevante è che affinché sia  $X^{-1}(A) \in \mathcal{F}$  per ogni  $A \in \mathcal{B} = \sigma(\mathcal{J})$  basta che  $X^{-1}(J) \in \mathcal{F}$  per ogni  $J \in \mathcal{J}$ . Questo vale in generale:

**Proposizione 2.** Dati due spazi di eventi  $(\Omega, \mathcal{F})$  ed  $(E, \mathcal{E})$ , sia  $X : \Omega \rightarrow E$ , e supponi che  $\mathcal{E} = \sigma(\mathcal{G})$ . Se  $X^{-1}(A) \in \mathcal{F}$  per ogni  $A \in \mathcal{G}$  allora  $X^{-1}(A) \in \mathcal{F}$  per ogni  $A \in \sigma(\mathcal{G}) = \mathcal{E}$ .

*Dimostrazione.* La famiglia  $\mathcal{A} := \{A \subseteq E : X^{-1}(A) \in \mathcal{F}\}$  è una  $\sigma$ -algebra (verifica!) che contiene  $\mathcal{G}$  e quindi anche  $\sigma(\mathcal{G})$ .  $\square$

Il punto di partenza del libro (Baldi capitolo 2) è il seguente

**Corollario.** Considera i due spazi di eventi  $(\Omega, \mathcal{F})$  ed  $(\mathbb{R}, \mathcal{B})$  e la funzione  $X : \Omega \rightarrow \mathbb{R}$ .  $X$  è una variabile aleatoria (cioè  $X^{-1}(A) \in \mathcal{F}$  per ogni  $A \in \mathcal{B}$ ) se per ogni  $b \in \mathbb{R}$  si ha  $X^{-1}((-\infty, b]) \equiv \{\omega \in \Omega : -\infty < X(\omega) \leq b\} \in \mathcal{F}$ .

## Variabili aleatorie discrete

Una funzione  $X : \Omega \rightarrow \mathbb{R}$  si dice *discreta* se la sua immagine  $X(\Omega) = \{X(\omega) : \omega \in \Omega\}$  è numerabile, cioè si ha  $X(\Omega) = \{x_1, \dots, x_i, \dots\}$ . Poiché qualunque singoletto è in  $\mathcal{B}$  (perché è chiuso), perché  $X$  sia una v.a.<sup>6</sup> deve essere  $E_i := X^{-1}\{x_i\} \in \mathcal{F}$  per ogni  $n$ . E questo

<sup>3</sup>Questa intersezione non è vuota perché contiene l'insieme potenza di  $E$ ; è facile verificare che è una  $\sigma$ -algebra; e per costruzione è la più piccola che contiene  $\mathcal{G}$ .

<sup>4</sup>Un chiuso è complemento di un aperto quindi  $\mathcal{C} \subseteq \sigma(\mathcal{O})$  che implica  $\sigma(\mathcal{C}) \subseteq \sigma(\mathcal{O})$ ; e analogamente poiché un aperto è complemento di un chiuso  $\mathcal{O} \subseteq \sigma(\mathcal{C})$  che implica  $\sigma(\mathcal{O}) \subseteq \sigma(\mathcal{C})$ .

<sup>5</sup>in realtà basta  $b \in \mathbb{Q}$ , ma non importa

<sup>6</sup>abbreviazione di variabile aleatoria

basta, perché per  $A \in \mathcal{B}$ ,  $X(\omega) \in A$  se e solo se  $X(\omega) = x_i$  per qualche  $x_i \in A$ ; quindi  $X^{-1}(A) = \cup\{E_i: x_i \in A\}$  (unione numerabile) che è in  $\mathcal{F}$  se gli  $E_i$  lo sono.

Nota che, denotando con  $\mathbf{1}_E$  l'indicatore di  $E \subseteq \Omega$  (cioè  $\mathbf{1}_E(\omega) = 1$  se  $\omega \in E$  e zero altrimenti), la famiglia  $\{E_i\}$  così definita è una partizione di  $\Omega$  ed  $X = \sum_i x_i \mathbf{1}_{E_i}$  - la funzione che prende valore  $x_i$  in  $E_i$ . Arriviamo quindi alla

**Definizione 2.** Una  $X: \Omega \rightarrow \mathbb{R}$  è una *variabile aleatoria discreta* se  $X = \sum_i x_i \mathbf{1}_{E_i}$  dove  $\{E_i\}$  è una partizione di  $\Omega$  con  $E_i \in \mathcal{F} \forall i \geq 1$ .

Queste sono le v.a. considerate nel capitolo 2. Possiamo guardare agli esempi lì contenuti. Intanto facciamo un esempio.

**Esempio.** Sia  $\Omega = \{1, \dots, 6\}^2$  - lancio di due dadi,  $\omega = (\omega_1, \omega_2)$  con  $\omega_i \in \{1, \dots, 6\}$ . Scriviamo la  $X = \text{"somma dei due numeri"}$ . Ovviamente  $X(\omega) = \omega_1 + \omega_2$ . Nella forma "canonica" è  $X = \sum_{i=2}^{12} i \cdot \mathbf{1}_{E_i}$  con  $E_i = \{\omega: \omega_1 + \omega_2 = i\}$ . Un'altra, chiamiamola  $Y$  è "numero di volte che esce 6". Indicando con  $\lfloor x \rfloor$  il massimo intero minore o uguale ad  $x \in \mathbb{R}$  (es.  $\lfloor -2.5 \rfloor = -3$ ,  $\lfloor 2.5 \rfloor = 2$ ) abbiamo  $Y(\omega) = \lfloor \omega_1/6 \rfloor + \lfloor \omega_2/6 \rfloor$ . In forma canonica, ponendo  $S_1 = \{6\} \times \{1, \dots, 6\}$  (6 al primo lancio) ed  $S_2 = \{1, \dots, 6\} \times \{6\}$  (6 al secondo) abbiamo  $Y = \mathbf{1}_{S_1 \setminus S_2} + \mathbf{1}_{S_2 \setminus S_1} + 2 \cdot \mathbf{1}_{S_1 \cap S_2} + 0 \cdot \mathbf{1}_{(S_1 \cup S_2)^c}$ .

## Variabili a valori in $\mathbb{R}^n$

Come vedremo, su uno spazio  $(\Omega, \mathcal{F}, P)$  possono essere definite più v.a. che conviene esaminare *congiuntamente*. Per fare un esempio banale considera il modello del lancio di moneta con probabilità  $p$  di testa ripetuto  $n$  volte, cioè  $\Omega = \{0, 1\}^n$ ,  $P\{\omega\} = p^{\sum_{i=1}^n \omega_i} (1-p)^{n-\sum_{i=1}^n \omega_i}$  (qui  $\mathcal{F} = \mathcal{P}(\Omega)$  perché  $\Omega$  è finito). Sia  $X_1(\omega) = \sum_{i=1}^n \omega_i$  ed  $X_2(\omega) = n - \sum_{i=1}^n \omega_i$  (numero di successi e di insuccessi negli  $n$  lanci), e poniamo  $X = (X_1, X_2)$ . Sappiamo allora che  $P\{\omega: X(\omega) = (k, n-k)\} = \binom{n}{k} p^k (1-p)^{n-k}$  per  $k = 0, \dots, n$ .

L'idea della condizione che una v.a. deve soddisfare è la stessa che nel caso di valori in  $\mathbb{R}$ : devo essere in grado di calcolare  $P(X^{-1}(A))$  per una famiglia ricca abbastanza di insiemi di  $\mathbb{R}^n$ . Come avrai forse indovinato, questa sarà la famiglia  $\mathcal{B}(\mathbb{R}^n) := \sigma(\mathcal{O})$ , la famiglia degli insiemi di Borel in  $\mathbb{R}^n$  che è quella generata dagli aperti  $\mathcal{O}$  di  $\mathbb{R}^n$ . Analogamente a quanto fatto prima la abbreviamo con  $\mathcal{B}^n$ . Ripetendo la dimostrazione del caso unidimensionale, usando il fatto che un aperto di  $\mathbb{R}^n$  è unione numerabile di rettangoli aperti  $\Pi_{i=1}^n(a_i, b_i)$  (dove  $a_i, b_i \in \mathbb{R}$  se vogliamo anche razionali), si dimostra che  $\mathcal{B}^n$  è generata dalla famiglia  $\mathcal{J}$  dei rettangoli del tipo  $\Pi_{i=1}^n(-\infty, b_i]$ :  $\mathcal{B}^n = \sigma(\mathcal{J})$ . E la definizione di v.a. è la stessa di prima: dati  $(\Omega, \mathcal{F}, P)$  ed  $(\mathbb{R}^n, \mathcal{B}^n)$ ,  $X: \Omega \rightarrow \mathbb{R}^n$  è una v.a. se  $X^{-1}(A) \in \mathcal{F} \forall A \in \mathcal{B}^n$ .

Chiama  $X_i$  le coordinate di  $X: \Omega \rightarrow \mathbb{R}^n$ , così  $X = (X_1, \dots, X_n)$  con  $X_i: \Omega \rightarrow \mathbb{R}$ .<sup>7</sup> L'ultima cosa da sapere è che  $X$  è una v.a. se e solo se lo sono tutte le  $X_i$  (cioè se  $X_i^{-1}(A) \in \mathcal{F} \forall A \in \mathcal{B}$ ,

<sup>7</sup>  $X_i$  non è altro che la proiezione  $\pi_i \circ X$  dove  $\pi_i(x_1, \dots, x_n) = x_i$ .

per ogni  $i$ ). Questo è molto importante per accertarsi che nel quadro generale le cose siano “a posto”, ma dimostrarlo richiede un po’ di lavoro; mettiamo in appendice la dimostrazione per completezza.

## Distribuzione di una variabile aleatoria

Come abbiamo accennato all’inizio, data su  $(\Omega, \mathcal{F}, P)$  una v.a.  $X: \Omega \rightarrow \mathbb{R}$  siamo interessati alla probabilità con cui  $X$  sta in certi insiemi (per es. con che probabilità non è negativa). È esattamente per questo che abbiamo imposto la condizione di misurabilità  $X^{-1}A \equiv \{\omega: X(\omega) \in A\} \in \mathcal{F} \forall A \in \mathcal{B}$ : l’interesse è su  $P(X^{-1}A)$ . Questo vale sia per  $X: \Omega \rightarrow \mathbb{R}$  che per  $X: \Omega \rightarrow \mathbb{R}^n$ . Ponendo  $(\mathbb{R}^1, \mathcal{B}^1) \equiv (\mathbb{R}, \mathcal{B})$  possiamo considerare il caso generale  $(\mathbb{R}^n, \mathcal{B}^n), n \geq 1$  con  $X: \Omega \rightarrow \mathbb{R}^n$ . Possiamo alleggerire la notazione indicando l’evento  $X^{-1}A = \{\omega: X(\omega) \in A\}$  con  $\{X \in A\}$ , e scrivere  $P(X^{-1}A) = P(X \in A)$ .

Osserva che la funzione  $A \mapsto P(X^{-1}A)$  assegna probabilità ad insiemi  $A \in \mathcal{B}^n$ ; è in effetti una probabilità su  $\mathcal{B}^n$ , che si indica con  $P_X$ , e si chiama *distribuzione di  $X$* .<sup>8</sup> Dunque la distribuzione di  $X$  è la probabilità su  $\mathcal{B}^n$  definita da

$$P_X(A) := P(X \in A), \quad A \in \mathcal{B}^n.$$

Per esempio nel caso  $\Omega = \{\text{sole, pioggia}\}$  con probabilità 50-50 siamo indifferenti fra le due v.a.  $X = \text{“vendo gelati”}$  e  $Y = \text{“vendo tè”}$  che danno  $(100, -100)$  e  $(-100, 100)$  perché hanno la stessa distribuzione su  $\mathbb{R}$ , data da  $P_X(100) = P_Y(100) = 1/2$  e  $P_X(-100) = P_Y(-100) = 1/2$ .

## Distribuzione di $X: \Omega \rightarrow \mathbb{R}$ discrete

Abbiamo visto che queste sono della forma  $X = \sum_i x_i \mathbf{1}_{E_i}$  con  $\{E_i\}$  partizione di  $(\Omega, \mathcal{F})$  ed  $E_i \in \mathcal{F} \forall i$ ; hanno immagine  $X(\Omega) = \{x_1, \dots, x_n, \dots\}$ , ed  $E_i = \{\omega: X(\omega) = x_i\}$ ; poniamo  $p(x_i) := P(E_i) \equiv P_X(\{x_i\})$ . Abbiamo visto che  $X^{-1}(A) = \cup\{E_i: x_i \in A\}$  (e l’unione è disgiunta), sicché

$$P_X(A) = P(X \in A) = \sum_{x_i \in A} P(E_i) = \sum_{x_i \in A} p(x_i).$$

La  $P_X(A) = \sum_{x_i \in A} p(x_i)$  è notevole: per calcolare  $P_X(A)$ , per qualunque  $A$ , bastano i  $p(x_i)$  (cioè i valori di  $P_X$  nei punti  $x_i$ ). Questo caratterizza la semplicità delle variabili discrete (che non per niente sono anche chiamate “semplici”).

Alleggerendo la notazione possiamo scrivere  $E_i = \{X = x_i\}$ , così  $p(x_i) = P(X = x_i)$ . A questo punto siamo riagganciati al libro.

---

<sup>8</sup>Verifichiamo che  $P_X$  è una probabilità: che è in  $[0, 1]$  è evidente, e altrettanto lo è che  $P_X(\mathbb{R}^n) = 1$  perché  $X^{-1}(\mathbb{R}^n) = \Omega$ . Infine, se  $\{A_i\}$  è una famiglia disgiunta in  $\mathcal{B}^n$  la famiglia  $\{X^{-1}A_i\}$  è disgiunta in  $\mathcal{F}$ , sicché  $P_X(\cup A_i) = P(X^{-1}(\cup A_i)) = P(\cup X^{-1}A_i) = \sum P(X^{-1}A_i) = \sum P_X(A_i)$ .

## Appendice

Vogliamo dimostrare che dati  $(\Omega, \mathcal{F}, P)$  ed  $(\mathbb{R}^n, \mathcal{B}^n)$ ,  $X: \Omega \rightarrow \mathbb{R}^n$  è una v.a. se e solo se lo sono le sue coordinate  $X_i: \Omega \rightarrow \mathbb{R}$  cioè se  $X_i^{-1}(A) \in \mathcal{F} \forall A \in \mathcal{B}$ , per ogni  $i$ . Per farlo conviene definire su  $\mathbb{R}^n$  un'altra  $\sigma$ -algebra che chiameremo  $\mathcal{B}^{\otimes n}$ ; dimostrare l'equivalenza voluta usando  $\mathcal{B}^{\otimes n}$ ; e infine dimostrare che  $\mathcal{B}^{\otimes n} = \mathcal{B}^n$ .

La definizione di  $\mathcal{B}^{\otimes n}$  è la (1.19) in Georgii, la riportiamo per comodità. Indica con  $\pi_i$  la proiezione di  $x \in \mathbb{R}^n$  sulla  $i$ -esima coordinata,  $\pi_i(x) = x_i$ . La definizione è  $\mathcal{B}^{\otimes n} = \sigma\{\pi_i^{-1}(A): A \in \mathcal{B}, i = 1, \dots, n\}$  - la più piccola  $\sigma$ -algebra in  $\mathbb{R}^n$  che fa sì che le proiezioni da  $\mathbb{R}^n$  ad  $\mathbb{R}$  siano v.a. La dimostrazione di  $\mathcal{B}^{\otimes n} = \mathcal{B}^n$  è in Georgii (non è difficile).

Usiamo dunque  $(\mathbb{R}^n, \mathcal{B}^{\otimes n})$ . Se  $X: \Omega \rightarrow \mathbb{R}^n$  è una v.a. rispetto a questo spazio anche  $X_i = \pi_i \circ X$  sarà v.a. perché per  $A \in \mathcal{B}$  abbiamo  $X_i^{-1}(A) = X^{-1}(\pi_i^{-1}(A)) \in \mathcal{F}$  perché  $\pi_i^{-1}(A) \in \mathcal{B}^{\otimes n}$ . Assumi viceversa che le  $X_i$  siano v.a. Riflettendo un minuto sulle definizioni se vede che  $\mathcal{B}^{\otimes n}$  è generata dalla famiglia  $\mathcal{D} = \{A_1 \times \dots \times A_n : A_i \in \mathcal{B}\} = \{\cap_{i=1}^n \pi_i^{-1}(A_i) : A_i \in \mathcal{B}\}$ ; quindi per la Proposizione 2 basta provare che  $X^{-1}(D) \in \mathcal{F}$  per ogni  $D \in \mathcal{D}$ . Ma  $X^{-1}(\cap_{i=1}^n \pi_i^{-1}(A_i)) = \cap_{i=1}^n X_i^{-1}(A_i) \in \mathcal{F}$  come segue subito dall'ipotesi.