

## Capitolo 2

# Variabili aleatorie discrete

### 2.1 Variabili aleatorie e loro distribuzioni

In molti degli esempi del capitolo precedente, abbiamo calcolato delle probabilità di eventi che si potevano esprimere in termini di una funzione dell'esito di un esperimento aleatorio.

**Esempio 2.1.1** Riprendiamo qui l'esempio 1.5.4, in cui lo spazio campionario è

$$\Omega = \{\omega = (\omega_1, \omega_2, \dots, \omega_N) : \omega_i \in \{0, 1\}\} = \{0, 1\}^N,$$

e la probabilità  $P$  è data da

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

Abbiamo calcolato la probabilità dell'evento

$$A = \left\{ \omega \in \Omega : \sum_{i=1}^N \omega_i = n \right\},$$

dove  $0 \leq n \leq N$ , e dell'evento

$$B = \{\omega \in \Omega : \omega_k = 0 \ \forall k < n, \omega_n = 1\}.$$

Posto  $X(\omega) = \sum_{i=1}^N \omega_i$ , e

$$Y(\omega) = \begin{cases} \min\{k : \omega_k = 1\} & \text{se } \{k : \omega_k = 1\} \neq \emptyset \\ N+1 & \text{altrimenti,} \end{cases}$$

possiamo riscrivere

$$A = \{\omega \in \Omega : X(\omega) = n\}$$

e

$$B = \{\omega \in \Omega : Y(\omega) = n\}.$$

**Definizione 2.1.1** Sia  $(\Omega, \mathcal{A}, P)$  uno spazio di probabilità ed  $E$  un insieme dotato della  $\sigma$ -algebra  $\mathcal{E}$ . Una funzione  $X : \Omega \rightarrow E$  si dice **variabile aleatoria** a valori in  $E$  se per ogni  $A \in \mathcal{E}$  si ha che

$$X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\} = \{X \in A\} \in \mathcal{A}$$

Introduciamo un po' di nomenclatura e di notazioni. Con riferimento alla definizione precedente, se  $E = \mathbb{R}$  diciamo che  $X$  è una variabile aleatoria **reale**. Se  $E = \mathbb{R}^n$  diciamo che  $X$  è una variabile aleatoria **vettoriale** o **vettore aleatorio** di dimensione  $n$ . Se  $E = \mathbb{C}$ , diciamo che  $X$  è una variabile aleatoria **complessa**.

Notiamo che nel calcolo delle probabilità si tende ad omettere dalle notazioni, quando possibile, la dipendenza dallo spazio probabilizzato (e quindi ad omettere di scrivere esplicitamente la dipendenza da  $\omega$ ). Molto spesso quindi si scrive  $\{X \in A\}$  e  $P\{X \in A\}$  in luogo di  $\{\omega \in \Omega \mid X(\omega) \in A\}$  (o di  $X^{-1}(A)$ ) e  $P(\{X \in A\})$  rispettivamente.

In molte applicazioni possiamo definire funzioni di variabili aleatorie: è interessante allora sapere quando queste sono a loro volta delle variabili aleatorie secondo la Definizione 2.1.1.

**Definizione 2.1.2** Sia  $f : E \rightarrow F$ , dove  $(E, \mathcal{E})$  ed  $(F, \mathcal{F})$  sono spazi misurabili. Diciamo che  $f$  è **misurabile** se per ogni  $A \in \mathcal{F}$  abbiamo  $f^{-1}(A) \in \mathcal{E}$ .

Il concetto di funzione misurabile è quindi perfettamente analogo a quello di variabile aleatoria, con l'unica differenza che lo spazio di partenza non è  $\Omega$  ma un qualunque insieme  $E$  dotato di una  $\sigma$ -algebra  $\mathcal{E}$ .

**Lemma 2.1.3** Sia  $X$  una variabile aleatoria a valori in  $(E, \mathcal{E})$ , e  $f : (E, \mathcal{E}) \rightarrow (F, \mathcal{F})$  misurabile. Allora l'applicazione  $f(X) : \Omega \rightarrow F$  è una variabile aleatoria a valori in  $(F, \mathcal{F})$ .

**Dimostrazione.** Bisogna dimostrare che per ogni  $A \in \mathcal{F}$  abbiamo

$$f(X)^{-1}(A) = \{f(X) \in A\} = \{X \in f^{-1}(A)\} \in \mathcal{E}$$

Ma questo è vero poichè  $f^{-1}(A) \in \mathcal{E}$  ed  $X$  è una variabile aleatoria a valori in  $(E, \mathcal{E})$ . ■

## 2.2 Legge di una variabile aleatoria

In molte applicazioni pratiche, si è interessati non tanto a come è definita una variabile aleatoria  $X$ , quanto alla probabilità di eventi del tipo  $\{X \in A\}$ , con  $A \in \mathcal{E}$ . Ci viene allora in aiuto la seguente definizione.

**Definizione 2.2.1** Sia  $X$  una variabile aleatoria discreta a valori in  $E$ . La funzione  $P_X : \mathcal{E} \rightarrow [0, 1]$ , definita da

$$P_X(A) := P\{X \in A\}$$

si dice **distribuzione**, o **legge**, della variabile aleatoria  $X$ . Se non c'è rischio di ambiguità (cioè se su  $(\Omega, \mathcal{A})$  è definita solo la probabilità  $P$ ), si usa anche denotare la legge di  $X$  con  $\mu_X$ .

**Proposizione 2.2.2** La distribuzione  $\mu_X$  di una variabile aleatoria  $X$  a valori in  $(E, \mathcal{E})$  è una probabilità su  $(E, \mathcal{E})$ .

**Dimostrazione.** Bisogna verificare che  $\mu_X$  soddisfa gli assiomi P1 e P2 della Definizione 1.1.2:

1.  $\mu_X(E) = P\{X \in E\} = 1$ ;

2. se  $(A_n)_{n \geq 0}$  è una successione di insiemi disgiunti contenuta in  $\mathcal{E}$ , allora gli eventi  $\{X \in A_n\}$  sono disgiunti (verificarlo!), e ne segue che

$$\mu_X \left( \bigcup_n A_n \right) = P \left\{ X \in \bigcup_n A_n \right\} = P \left( \bigcup_n \{X \in A_n\} \right) = \sum_n P\{X \in A_n\} = \sum_n \mu_X(A_n)$$

■

La legge di  $X$  è dunque la probabilità che ad ogni elemento  $A \in \mathcal{E}$  associa “la probabilità che  $X$  cada in  $A$ ”. Per indicare che la variabile aleatoria  $X$  ha legge  $\mu_X$ , si usa la scrittura  $X \sim \mu_X$ .

Vediamo ora il più semplice esempio di legge di variabili aleatorie.

**Esempio 2.2.1 (variabile aleatoria di Bernoulli)** Sia dato un evento  $A$ , tale che  $P(A) = p \in (0, 1)$ , e consideriamo la cosiddetta **funzione indicatrice** di  $A$ , definita su  $\Omega$  come

$$X(\omega) = \begin{cases} 1 & \text{se } \omega \in A \\ 0 & \text{se } \omega \in A^c \end{cases}$$

Sappiamo che la funzione  $X$  assume valori su  $E = \{0, 1\}$ . Innanzitutto verifichiamo che è una variabile aleatoria secondo la Definizione 2.1.1: se poniamo su  $E$  la  $\sigma$ -algebra  $\mathcal{E} := \mathcal{P}(E)$ , verifichiamo che la controimmagine di ogni sottoinsieme di  $E$  sia un evento in  $\mathcal{A}$ :

$$\begin{aligned} X^{-1}(\{0\}) &= \{X = 0\} = A^c \in \mathcal{A}, \\ X^{-1}(\{1\}) &= \{X = 1\} = A \in \mathcal{A}, \\ X^{-1}(E) &= \{X \in E\} = \Omega \in \mathcal{A}, \\ X^{-1}(\emptyset) &= \{X \in \emptyset\} = \emptyset \in \mathcal{A}. \end{aligned}$$

Questo implica che  $X$  è una variabile aleatoria: la sua legge è data da

$$\begin{aligned} \mu_X(\{0\}) &= P\{X = 0\} = P(A^c) = 1 - p, \\ \mu_X(\{1\}) &= P\{X = 1\} = P(A) = p, \\ \mu_X(E) &= P\{X \in E\} = P(\Omega) = 1, \\ \mu_X(\emptyset) &= P\{X \in \emptyset\} = P(\emptyset) = 0. \end{aligned}$$

(le ultime due quantità sono anche una conseguenza della Proposizione 2.2.2). Una variabile aleatoria con questo tipo di legge viene chiamata **variabile aleatoria di Bernoulli** di parametro  $p$ , e la sua legge viene indicata con  $Be(p)$ ; come visto sopra, in simboli si può indicare che  $X$  ha legge di Bernoulli di parametro  $p$  con  $X \sim Be(p)$ . Il nome variabile aleatoria di Bernoulli viene utilizzato di solito quando non interessa descrivere l’evento  $A$ , mentre quando interessa mettere enfasi su  $A$  si preferisce chiamarla funzione indicatrice di  $A$ , come fatto all’inizio, utilizzando il simbolo  $\mathbf{1}_A$  in luogo di  $X$ .

Finiamo la sezione dando questa definizione. Due variabili aleatorie  $X$  e  $Y$  si dicono **identicamente distribuite** (abbreviato con i.d.) se hanno la stessa legge, cioè se  $\mu_X = \mu_Y$ . Chiaramente se  $X$  e  $Y$  sono uguali allora sono identicamente distribuite, ma il viceversa non è vero.

**Esempio 2.2.2** Prendiamo una variabile  $X$  di Bernoulli di parametro  $1/2$  e definiamo  $Y = 1 - X$ ; allora  $X$  e  $Y$  chiaramente non sono uguali. Tuttavia esse sono identicamente distribuite, poiché hanno entrambe legge  $Be(1/2)$ . Infatti anche  $Y$  può assumere solo i valori  $0$  e  $1$ , e  $P\{Y = 1\} = P\{1 - X = 1\} = P\{X = 0\} = 1 - 1/2 = 1/2$ .

Il concetto di variabili aleatorie identicamente distribuite si estende a funzioni di variabili aleatorie in questo modo.

**Lemma 2.2.3** *Siano  $X$  e  $Y$  variabili aleatorie identicamente distribuite a valori in  $(E, \mathcal{E})$ , e  $f : (E, \mathcal{E}) \rightarrow (F, \mathcal{F})$  misurabile. Allora le variabili aleatorie  $f(X)$  ed  $f(Y)$  sono a loro volta identicamente distribuite su  $(F, \mathcal{F})$ .*

**Dimostrazione.** Bisogna dimostrare che per ogni  $A \in \mathcal{F}$  abbiamo

$$\mu_{f(X)}(A) = P\{f(X) \in A\} = P\{X \in f^{-1}(A)\} = P\{Y \in f^{-1}(A)\} = P\{f(Y) \in A\} = \mu_{f(Y)}(A)$$

e quindi  $\mu_{f(X)} = \mu_{f(Y)}$ . ■

Per conoscere la legge  $\mu_X$  di una variabile aleatoria  $X$  bisognerebbe in principio conoscere tutti i valori di  $\mu_X(A)$  per ogni sottoinsieme  $A \in \mathcal{E}$ , cosa che in principio può essere complessa. Ci sono casi in cui questo non è necessario, poichè si possono ricavare tutti questi valori da un numero inferiore di mattoncini costituenti. Il caso più semplice verrà trattato nel resto di questo capitolo, ed è costituito dalle cosiddette variabili aleatorie discrete. Un altro caso significativo verrà trattato nel Capitolo 3 e sarà costituito dalle cosiddette variabili aleatorie continue.

## 2.3 Variabili aleatorie discrete e densità discreta

**Definizione 2.3.1** Sia  $X$  una variabile aleatoria a valori in uno spazio misurabile  $(E, \mathcal{E})$ . Diciamo che  $X$  è una variabile aleatoria **discreta** se  $E$  è discreto.

**Osservazione 2.3.1** Ci sono casi in cui  $E$  non è discreto, ma possiamo restringerci a questo caso definendo un opportuno sottoinsieme di  $E$ . Difatti, se  $\{X = x\} \in \mathcal{A}$  per ogni  $x \in E$ , definiamo

$$E_0 := \{x \in E \mid \mu_X(\{x\}) = P\{X = x\} > 0\}$$

Allora  $E_0$  è numerabile in virtù della Proposizione 2.3.2 nel seguito, e se  $\mu_X(E_0) = 1$  possiamo restringere  $X$  al solo insieme  $E_0$  ed ottenere tutti i risultati seguenti. Vediamo ora perchè  $E_0$  è numerabile.

**Proposizione 2.3.2** *Sia  $(x_i)_{i \in I}$  una successione di reali non negativi, con  $I$  insieme di indici qualsiasi. Definiamo la loro somma come*

$$\sum_{i \in I} x_i := \sup \left\{ \sum_{j \in J} x_j \mid J \subseteq I \text{ finito} \right\}$$

e definiamo

$$I_0 := \{i \in I \mid x_i > 0\}$$

Se  $\sum_{i \in I} x_i =: M < +\infty$ , allora  $I_0$  è al più numerabile.

**Dimostrazione.** Definiamo  $I_n := \{i \in I \mid x_i > \frac{1}{n}\}$ . Allora  $|I_n| \leq Mn$ . Se infatti questo non fosse vero, avremmo

$$M \geq \sum_{j \in I_n} x_j > \sum_{j \in I_n} \frac{1}{n} > Mn \cdot \frac{1}{n} = M$$

ma siccome  $M < +\infty$  questo è impossibile, dunque  $|I_n| \leq Mn$ . Siccome  $I_0 = \cup_{n \geq 1} I_n$ , la sua cardinalità può essere solo finita o numerabile. ■

Sia  $X$  una variabile aleatoria discreta a valori in un insieme  $E$ , e supponiamo che  $\mathcal{E} = \mathcal{P}(E)$ .

**Definizione 2.3.3** La funzione

$$\begin{aligned} p_X : E &\rightarrow [0, 1] \\ x &\mapsto P\{X = x\} \end{aligned}$$

si dice **densità discreta** della variabile aleatoria  $X$ .

Nel seguito l'aggettivo "discreta" verrà omissso, fino a quando non introdurremo, nel prossimo capitolo, la nozione "continua" di densità.

**Osservazione 2.3.2** La densità  $p_X$  e la distribuzione  $\mu_X$  di una variabile aleatoria discreta, si determinano l'un l'altra tramite le relazioni, per  $x \in E$ ,  $A \subseteq E$

$$(2.3.1) \quad p_X(x) = \mu_X(\{x\}), \quad \mu_X(A) = \sum_{x \in A} p_X(x).$$

**Osservazione 2.3.3** Se  $X$  è una variabile aleatoria discreta a valori in  $(E, \mathcal{E})$  con  $\mathcal{E} = \mathcal{P}(E)$ , si verifica banalmente che per qualunque funzione  $f$  definita su  $E$  e a valori in qualunque spazio misurabile  $(F, \mathcal{F})$  abbiamo che  $f(X)$  è una variabile aleatoria.

## 2.4 Vettori aleatori discreti. Densità congiunta e densità marginali.

In questa sezione tratteremo vettori aleatori  $X = (X_1, \dots, X_n)$  in cui ogni componente assumerà valori in un generico insieme discreto  $E_i$ ,  $i = 1, \dots, n$ . Dire che  $X$  è un vettore aleatorio equivale a dotare l'insieme di arrivo, cioè lo spazio prodotto  $E := E_1 \times \dots \times E_n = \prod_{i=1}^n E_i$ , di una  $\sigma$ -algebra  $\mathcal{E}$  e verificare che  $X^{-1}(A) = \{X \in A\} \in \mathcal{A}$  per ogni  $A \in \mathcal{E}$ .

**Proposizione 2.4.1** Siano  $X_1, \dots, X_n$  funzioni definite sullo spazio probabilizzato  $(\Omega, \mathcal{A}, P)$  a valori negli insiemi discreti  $E_1, \dots, E_n$  rispettivamente, ciascuno dotato della  $\sigma$ -algebra  $\mathcal{E}_i = \mathcal{P}(E_i)$ . Sono equivalenti:

1. per ogni  $i = 1, \dots, n$ ,  $X_i$  è una variabile aleatoria a valori in  $(E_i, \mathcal{E}_i)$ ;
2.  $X := (X_1, \dots, X_n)$  è un vettore aleatorio a valori in  $E := E_1 \times \dots \times E_n$  rispetto alla  $\sigma$ -algebra  $\mathcal{E} := \mathcal{P}(E)$ .

**Dimostrazione.** Se  $(X_1, \dots, X_n)$  è un vettore aleatorio a valori in  $E_1 \times \dots \times E_n$ , allora per ogni  $i = 1, \dots, n$  e per ogni  $A_i \in \mathcal{E}_i$  si ha che

$$\begin{aligned} \{X_i \in A_i\} &= \{X_1 \in E_1, \dots, X_i \in A_i, \dots, X_n \in E_n\} = \\ &= \{(X_1, \dots, X_n) \in E_1 \times \dots \times E_{i-1} \times A_i \times E_{i+1} \times \dots \times E_n\} \in \mathcal{A} \end{aligned}$$

e quindi  $X_i$  è una variabile aleatoria. Viceversa, se le  $X_i$ ,  $i = 1, \dots, n$  sono variabili aleatorie, allora per ogni  $x = (x_1, \dots, x_n) \in E_1 \times \dots \times E_n$  si ha

$$\{X = x\} = \{(X_1, \dots, X_n) = (x_1, \dots, x_n)\} = \{X_1 = x_1, \dots, X_n = x_n\} = \bigcap_{i=1}^n \{X_i = x_i\} \in \mathcal{A}$$

Siccome  $(X_1, \dots, X_n)$  assume valori su  $E_1 \times \dots \times E_n$ , che è un insieme discreto, quanto appena dimostrato basta per affermare che  $(X_1, \dots, X_n)$  è una variabile aleatoria. Difatti per ogni  $A \in \mathcal{P}(E)$  si ha

$$\{X \in A\} = \bigcup_{x \in A} \{X = x\}$$

Questa è un'unione numerabile di insiemi che abbiamo appena dimostrato appartenere ad  $\mathcal{A}$ , quindi è un evento. ■

Se  $X$  è un vettore aleatorio discreto, si può definire la sua densità discreta, che verrà denotata con  $p_X(x)$ , o con  $p_{X_1, \dots, X_n}(x_1, \dots, x_n)$  ogni qual volta si voglia evidenziare la sua dipendenza dalle  $n$  variabili  $x_1, \dots, x_n$ ; in tal caso la chiameremo **densità congiunta** delle variabili  $X_1, X_2, \dots, X_n$ .

**Proposizione 2.4.2** Sia  $1 \leq k < n$ . Allora il vettore aleatorio  $(X_1, \dots, X_k)$  di dimensione  $k$  ha densità data da

$$(2.4.1) \quad p_{X_1, \dots, X_k}(x_1, \dots, x_k) = \sum_{(x_{k+1}, \dots, x_n) \in H} p_{X_1, \dots, X_n}(x_1, \dots, x_n)$$

dove  $H := E_{k+1} \times \dots \times E_n$ .

**Dimostrazione.** Fissiamo  $x_1, \dots, x_k$ ; allora:

$$\begin{aligned} p_{X_1, \dots, X_k}(x_1, \dots, x_k) &= P\{(X_1, \dots, X_k) = (x_1, \dots, x_k)\} = \\ &= P\{(X_1, \dots, X_k, \dots, X_n) \in \{(x_1, \dots, x_k)\} \times H\} = \\ &= \mu_X(\{(x_1, \dots, x_k)\} \times H) \\ &= \sum_{(x_{k+1}, \dots, x_n) \in H} p_{X_1, \dots, X_n}(x_1, \dots, x_k, \dots, x_n), \end{aligned}$$

dove nell'ultimo passaggio abbiamo usato (2.3.1). ■

**Osservazione 2.4.1** Siano  $i_1, \dots, i_k$ ,  $k$  indici distinti tra 1 e  $n$ . È chiaro che la proposizione 2.4.2 può essere riformulata per esprimere la densità di  $(X_{i_1}, \dots, X_{i_k})$  in termini di quella di  $(X_1, \dots, X_n)$ .

Nel caso in cui  $k = 1$ , la Proposizione 2.4.2 consente di esprimere la densità di una componente di un vettore aleatorio in termini della densità congiunta. Le densità delle componenti vengono chiamate **densità marginali**. Come mostriamo nei due seguenti esempi, non è possibile ricostruire la densità congiunta a partire dalla sola conoscenza delle densità marginali.

**Esempio 2.4.2** Da un'urna contenente 6 palline numerate da 1 a 6 si effettuano due estrazioni *con rimpiazzo*, cioè rimettendo nell'urna la pallina estratta subito dopo ogni estrazione. Indichiamo con  $X = (X_1, X_2)$  il risultato delle due estrazioni, e calcoliamone densità congiunta e densità marginali.

Supponendo che ogni risultato sia equiprobabile, la legge congiunta darà uguale probabilità ad ogni elemento dello spazio di arrivo  $\{1, \dots, 6\}^2$ , e quindi  $\mu_X$  sarà la legge uniforme su  $\{1, \dots, 6\}^2$ : questo significa che per ogni  $(x_1, x_2) \in \{1, \dots, 6\}^2$  abbiamo

$$p_X(x_1, x_2) = P\{(X_1, X_2) = (x_1, x_2)\} = \mu_X(\{(x_1, x_2)\}) = \frac{1}{36}$$

Utilizzando la formula 2.4.1 abbiamo poi

$$p_{X_1}(x_1) = \sum_{x_2=1}^6 p_X(x_1, x_2) = 6 \cdot \frac{1}{36} = \frac{1}{6},$$

$$p_{X_2}(x_2) = \sum_{x_1=1}^6 p_X(x_1, x_2) = 6 \cdot \frac{1}{36} = \frac{1}{6},$$

quindi anche le due leggi marginali sono probabilità uniformi su  $\{1, \dots, 6\}$ .

**Esempio 2.4.3** Dalla stessa urna del precedente esempio si effettuano ora due estrazioni *senza rimpiazzo*, cioè non rimettendo nell'urna la pallina estratta dopo ogni estrazione. Indichiamo come prima con  $X = (X_1, X_2)$  il risultato delle due estrazioni, e calcoliamone densità congiunta e densità marginali.

Supponendo che ogni risultato tra quelli possibili sia equiprobabile, la legge congiunta darà uguale probabilità ad ogni elemento dell'insieme  $\{(x_1, x_2) \mid x_i = 1, \dots, 6, x_1 \neq x_2\}$ , e quindi  $\mu_X$  sarà la legge uniforme su questo insieme; tuttavia, siccome le due variabili aleatorie in principio possono assumere ogni valore in  $\{1, \dots, 6\}$ , l'insieme naturale su cui  $X = (X_1, X_2)$  assume valori è  $\{1, \dots, 6\}^2$ : questo significa che per ogni  $(x_1, x_2) \in \{1, \dots, 6\}^2$  abbiamo

$$p_X(x_1, x_2) = P\{(X_1, X_2) = (x_1, x_2)\} = \mu_X(\{(x_1, x_2)\}) = \begin{cases} \frac{1}{30} & \text{se } x_1 \neq x_2, \\ 0 & \text{se } x_1 = x_2. \end{cases}$$

Utilizzando la formula 2.4.1 abbiamo poi

$$p_{X_1}(x_1) = \sum_{x_2=1}^6 p_X(x_1, x_2) = 5 \cdot \frac{1}{30} + 1 \cdot 0 = \frac{1}{6},$$

$$p_{X_2}(x_2) = \sum_{x_1=1}^6 p_X(x_1, x_2) = 5 \cdot \frac{1}{30} + 1 \cdot 0 = \frac{1}{6},$$

Anche stavolta le due leggi marginali sono probabilità uniformi su  $\{1, \dots, 6\}$ , ma notiamo che la legge congiunta non è più la legge marginale su  $\{1, \dots, 6\}^2$ .

Nei due esempi abbiamo quindi due vettori aleatori diversi con una legge congiunta diversa ma uguali leggi marginali: questo significa che, come abbiamo detto prima, non è possibile ricostruire la legge congiunta di un vettore aleatorio a partire dalla sola conoscenza delle leggi marginali. Questo sarà però possibile in un caso particolare, che sarà l'argomento della prossima sezione.

## 2.5 Indipendenza di variabili aleatorie

Nel caso di eventi indipendenti, abbiamo visto che riusciamo a dare una definizione generale per una qualsiasi famiglia (finita, numerabile o di cardinalità più grande) di eventi. Vediamo ora che la stessa cosa vale per le variabili aleatorie.

**Definizione 2.5.1** Sia  $I$  un insieme qualsiasi di indici, e  $\{X_i : i \in I\}$  una famiglia di variabili aleatorie a valori negli spazi  $(E_i, \mathcal{E}_i)$ ,  $i \in I$ . Si dice che le variabili aleatorie di tale famiglia sono **indipendenti** se, per ogni  $J \subseteq I$  finito e per ogni scelta di  $A_j \in \mathcal{E}_j$ ,  $j \in J$ , si ha

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

Il semplice confronto tra la Definizione 2.5.1 e quella di indipendenza tra eventi ci dà la seguente proprietà.

**Proposizione 2.5.2** Siano  $\{X_i : i \in I\}$  variabili aleatorie come nella definizione 2.5.1. Le seguenti affermazioni sono equivalenti:

- i. le  $(X_i)_{i \in I}$  sono indipendenti;
- ii. per ogni scelta di  $A_i \in \mathcal{E}_i$ ,  $i \in I$ , gli eventi  $\{X_i \in A_i\}$ ,  $i \in I$ , sono indipendenti.

Poichè la distribuzione congiunta di  $n$  variabili aleatorie discrete è completamente determinata dalla loro densità congiunta, non è sorprendente che l'indipendenza si possa caratterizzare in termini della densità congiunta.

**Proposizione 2.5.3** Siano  $X_1, X_2, \dots, X_n$  variabili aleatorie discrete definite sullo stesso spazio di probabilità  $(\Omega, \mathcal{A}, P)$ , a valori rispettivamente negli insiemi  $E_1, E_2, \dots, E_n$ . Denotiamo con  $p_{X_1, \dots, X_n}$  la loro densità congiunta, e con  $p_{X_i}$  le densità marginali. Allora  $X_1, X_2, \dots, X_n$  sono indipendenti se e solo se

$$(2.5.1) \quad p_{X_1, \dots, X_n}(x_1, \dots, x_n) = \prod_{i=1}^n p_{X_i}(x_i)$$

per ogni  $(x_1, \dots, x_n) \in E_1 \times \dots \times E_n$ .

**Dimostrazione.** Assumiamo che  $X_1, X_2, \dots, X_n$  siano indipendenti. Allora

$$p_{X_1, \dots, X_n}(x_1, \dots, x_n) = P\{X_1 = x_1, \dots, X_n = x_n\} = \prod_{i=1}^n P\{X_i = x_i\} = \prod_{i=1}^n p_{X_i}(x_i).$$

Viceversa, assumiamo che valga (2.5.1). Per semplificare le espressioni che seguono, assumiamo  $n = 2$ ; l'argomento che usiamo è però valido in generale. Siano  $A_1 \in E_1$ ,  $A_2 \in E_2$  fissati ma arbitrari. Si ha

$$\begin{aligned} P\{X_1 \in A_1, X_2 \in A_2\} &= P\{(X_1, X_2) \in A_1 \times A_2\} = \sum_{x_1 \in A_1, x_2 \in A_2} p_{X_1, X_2}(x_1, x_2) \\ &= \sum_{x_1 \in A_1, x_2 \in A_2} p_{X_1}(x_1)p_{X_2}(x_2) = \sum_{x_1 \in A_1} p_{X_1}(x_1) \sum_{x_2 \in A_2} p_{X_2}(x_2) = P\{X_1 \in A_1\}P\{X_2 \in A_2\}. \end{aligned}$$

■

**Osservazione 2.5.1** Siano  $X, Y$  due variabili aleatorie (ma l'argomento è generalizzabile a  $n$  variabili aleatorie) la cui densità congiunta si fattorizza nella forma

$$p_{X,Y}(x, y) = \alpha(x)\beta(y).$$

Allora

$$p_X(x) = \sum_y p_{X,Y}(x, y) = b\alpha(x),$$

dove  $b = \sum_y \beta(y)$ . Analogamente

$$p_Y(y) = a\beta(y),$$

con  $a = \sum_x \alpha(x)$ . Inoltre

$$1 = \sum_{x,y} p_{X,Y}(x, y) = ab.$$

Di conseguenza

$$p_{X,Y}(x, y) = \alpha(x)\beta(y) = p_X(x)p_Y(y),$$

e dunque  $X$  e  $Y$  sono indipendenti.

**Esempio 2.5.2** Si considerino gli esempi 2.4.2 e 2.4.3. Le variabili aleatorie  $X_1, X_2$  dell'Esempio 2.4.2 sono indipendenti, mentre quelle dell'Esempio 2.4.3 non sono indipendenti.

**Osservazione 2.5.3** Siano  $X_1, X_2, \dots, X_n$  variabili aleatorie indipendenti, e  $I = \{i_1, \dots, i_h\}$ ,  $J = \{j_1, \dots, j_k\}$  due sottoinsiemi non vuoti e *disgiunti* di  $\{1, 2, \dots, n\}$ . Denotiamo con  $X_I$  la variabile aleatoria

$$X_I = (X_{i_1}, \dots, X_{i_h}),$$

e, analogamente,  $X_J$ . Mostriamo che  $X_I, X_J$  sono variabili aleatorie indipendenti. Scriviamo  $x_I$  in luogo di  $(x_{i_1}, \dots, x_{i_h})$ , e analogamente  $x_J$ . Allora, tenendo conto che, per l'Osservazione 2.4.1, le variabili aleatorie  $\{X_l : l \in I \cup J\}$  sono indipendenti, si ha

$$\begin{aligned} p_{X_I, X_J}(x_I, x_J) &= p_{X_{i_1}, \dots, X_{i_h}, X_{j_1}, \dots, X_{j_k}}(x_{i_1}, \dots, x_{i_h}, x_{j_1}, \dots, x_{j_k}) \\ &= \prod_{r=1}^h p_{X_{i_r}}(x_{i_r}) \prod_{s=1}^k p_{X_{j_s}}(x_{j_s}) = p_{X_I}(x_I) p_{X_J}(x_J), \end{aligned}$$

da cui segue l'indipendenza.

Come vedremo, il risultato nell'Osservazione 2.5.3 viene usato congiuntamente a quello della seguente Proposizione, che stabilisce che l'indipendenza si conserva per trasformazioni. Il risultato che segue è enunciato e dimostrato per due variabili aleatorie, ma è facilmente estendibile a ogni  $n \geq 2$ .

**Proposizione 2.5.4** Siano  $X_1, X_2$  due variabili aleatorie definite sullo stesso spazio di probabilità  $(\Omega, \mathcal{A}, P)$ , e a valori negli spazi  $(E_1, \mathcal{E}_1)$  e  $(E_2, \mathcal{E}_2)$  rispettivamente. Siano inoltre  $(F_1, \mathcal{F}_1)$ ,  $(F_2, \mathcal{F}_2)$  due spazi, e  $f_i : E_i \rightarrow F_i$ ,  $i = 1, 2$  funzioni misurabili. Allora: se le variabili aleatorie  $X_1$  e  $X_2$  sono indipendenti, allora anche le variabili aleatorie  $f_1(X_1)$  e  $f_2(X_2)$  lo sono.

**Dimostrazione.** Basta osservare che per ogni  $A_i \in \mathcal{F}_i$ ,  $i = 1, 2$ , si ha

$$\begin{aligned} P\{f_1(X_1) \in A_1, f_2(X_2) \in A_2\} &= P\{X_1 \in f_1^{-1}(A_1), X_2 \in f_2^{-1}(A_2)\} = \\ &= P\{X_1 \in f_1^{-1}(A_1)\}P\{X_2 \in f_2^{-1}(A_2)\} = P\{f_1(X_1) \in A_1\}P\{f_2(X_2) \in A_2\} \end{aligned}$$

■

**Esempio 2.5.4** Siano  $X_1, \dots, X_n, X_{n+1}, \dots, X_{n+m}$  variabili aleatorie reali indipendenti. Allora  $X_1 + \dots + X_n$  e  $X_{n+1} + \dots + X_{n+m}$  sono indipendenti. Basta applicare la Proposizione precedente con  $I = \{1, 2, \dots, n\}$ ,  $J = \{n+1, \dots, n+m\}$ ,  $f(x_1, \dots, x_n) = x_1 + \dots + x_n$ ,  $g(x_{n+1}, \dots, x_{n+m}) = x_{n+1} + \dots + x_{n+m}$ .

Diamo infine questa definizione. Una famiglia di variabili aleatorie indipendenti  $(X_n)_n$  che siano anche identicamente distribuite si dice, appunto, una famiglia di variabili aleatorie **indipendenti ed identicamente distribuite** (abbreviato con i.i.d.). Questa condizione può a volte semplificare calcoli altrimenti complessi, come illustra il seguente esempio.

### 2.5.1 Variabili aleatorie binomiali

Una variabile aleatoria  $X$  si dice **binomiale** di parametri  $n \geq 1$  e  $p \in [0, 1]$ , e si scrive  $X \sim B(n, p)$ , se  $X$  ha la medesima distribuzione di una variabile aleatoria del tipo  $X_1 + \dots + X_n$ , dove le  $X_i$  sono i.i.d. e hanno legge  $Be(p)$ .

È facile calcolare la densità di  $X$ . Se le  $X_i$  rappresentano gli esiti di  $n$  prove ripetute indipendenti per ognuna delle quali l'esito positivo  $\{X_i = 1\}$  ha probabilità  $p$ , l'evento  $\{X_1 + \dots + X_n = k\}$  rappresenta la probabilità di avere ottenuto  $k$  esiti positivi. Tale probabilità è stata calcolata nell'Esempio 1.5.4:

$$p_X(k) = P\{X_1 + \dots + X_n = k\} = \binom{n}{k} p^k (1-p)^{n-k}.$$

Un'ultima, rilevante proprietà delle variabili aleatorie binomiali è la seguente.

**Proposizione 2.5.5** *Siano  $X \sim B(n, p)$  e  $Y \sim B(m, p)$  indipendenti. Allora  $X + Y \sim B(n+m, p)$ .*

**Dimostrazione.** Siano  $X_1, X_2, \dots, X_n, X_{n+1}, \dots, X_{n+m}$  variabili aleatorie di Bernoulli di parametro  $p$  indipendenti. La distribuzione congiunta di  $(X, Y)$  è uguale (perché?) alla distribuzione congiunta di  $(X_1 + \dots + X_n, X_{n+1} + \dots + X_{n+m})$ . Ne segue che  $X + Y$  ha la stessa distribuzione di  $X_1 + \dots + X_n + X_{n+1} + \dots + X_{n+m}$  che, per definizione, è una variabile aleatoria binomiale di parametri  $n + m$  e  $p$ . ■

### 2.5.2 Variabili aleatorie di Poisson

In numerose situazioni concrete, ci si trova a considerare delle variabili aleatorie binomiali i cui parametri,  $n$  e  $p$ , sono tali che  $n$  è molto grande e  $p$  è molto piccolo. In altre parole, si eseguono numerose ripetizioni di una prova che ha esito positivo con probabilità piccola. Supponiamo, ad esempio, di considerare il numero di accessi, in un fissato intervallo di tempo, ad un certo servizio (lo sportello di un ufficio pubblico, un pagina web, un centralino...). Vi sarà un numero grande di utenti, diciamo  $n$ , che ha accesso a tale servizio. Si osserva però che, tipicamente, un numero di

utenti molto minore di  $n$  accede effettivamente al servizio in un intervallo di tempo della lunghezza di quello fissato. Un ragionevole modello matematico per tale situazione, è di supporre che ogni utente abbia una probabilità piccola,  $p \ll 1$ , di accedere al servizio. In una approssimazione che è assai buona in molti casi, assumiamo che il valore di  $p$  sia uguale per tutti gli utenti, e che ogni utente si comporti in modo indipendente dagli altri. Sotto queste ipotesi, il numero  $X$  di utenti che effettivamente accede al servizio è una variabile aleatoria binomiale di parametri  $n$  e  $p$ , cioè

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Se  $n$  è molto grande, calcoli espliciti con questa densità risultano estremamente pesanti, se non impraticabili, data la difficoltà di calcolare i coefficienti binomiali per grandi valori di  $n$ . È allora interessante analizzare il comportamento asintotico di  $p_X(k)$  quando  $n \rightarrow +\infty$  e  $p \rightarrow 0$ . Per ottenere un comportamento limite non banale, è necessario che  $p$  vada a zero “proporzionalmente” a  $\frac{1}{n}$ , come illustra il risultato seguente.

**Proposizione 2.5.6** *Supponiamo che  $X_n \sim B(n, p_n)$ , con  $p_n$  successione reale tale che  $\lim_{n \rightarrow +\infty} np_n = \lambda > 0$ . Allora per ogni  $k \in \mathbb{N}$  si ha*

$$\lim_{n \rightarrow +\infty} p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

**Dimostrazione.** Fissiamo un generico  $k \in \mathbb{N}$  e calcoliamo, per ogni  $n \geq 1$ ,

$$\begin{aligned} p_{X_n}(k) &= \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{n-k} = \\ &= \frac{(np_n)^k}{k!} \frac{n(n-1) \cdots (n-k+1)}{n^k} \frac{1}{\left(1 - \frac{np_n}{n}\right)^k} \left(1 - \frac{np_n}{n}\right)^n. \end{aligned}$$

D'altra parte

$$\lim_{n \rightarrow +\infty} \frac{n(n-1) \cdots (n-k+1)}{n^k} = 1, \quad \lim_{n \rightarrow +\infty} \left(1 - \frac{np_n}{n}\right)^k = 1, \quad \text{e} \quad \lim_{n \rightarrow +\infty} \left(1 - \frac{np_n}{n}\right)^n = e^{-\lambda}.$$

Ne segue che

$$\lim_{n \rightarrow +\infty} p_{X_n}(k) = \frac{\lambda^k}{k!} \cdot 1 \cdot 1 \cdot e^{-\lambda} = e^{-\lambda} \frac{\lambda^k}{k!}.$$

■

Si noti che l'espressione appena ottenuta è la densità di una variabile aleatoria a valori in  $\mathbb{N}$ , in quanto

$$\sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = 1.$$

Tale procedura di approssimazione giustifica la definizione di una nuova classe di variabili aleatorie. Una variabile aleatoria  $X$  a valori in  $\mathbb{N}$  si dice variabile aleatoria **di Poisson** di parametro  $\lambda > 0$ , e si scrive  $X \sim Po(\lambda)$ , se

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

per ogni  $k \in \mathbb{N}$ . Tali variabili aleatorie vengono comunemente usate per modellare, tra le altre, quantità del tipo “numero di accessi ad un servizio”, come abbiamo sopra giustificato.

Il risultato che ora illustriamo afferma che la somma di due variabili aleatorie di Poisson indipendenti è ancora una variabile aleatoria di Poisson.

**Proposizione 2.5.7** *Siano  $X \sim Po(\lambda)$  e  $Y \sim Po(\mu)$  indipendenti. Allora  $X + Y \sim Po(\lambda + \mu)$ .*

Prima di dimostrare la Proposizione 2.11.4, mostriamo un risultato del tutto generale.

**Proposizione 2.5.8** *Siano  $X$  e  $Y$  due variabili aleatorie discrete reali, e sia  $p_{X,Y}$  la loro densità congiunta. Allora*

$$(2.5.2) \quad p_{X+Y}(z) = P\{X + Y = z\} = \sum_x p_{X,Y}(x, z - x) = \sum_y p_{X,Y}(z - y, y).$$

**Dimostrazione.** Usando la  $\sigma$ -additività della probabilità, si ha

$$\begin{aligned} P\{X + Y = z\} &= P\left(\bigcup_{x \in X(\Omega)} \{X = x, Y = z - x\}\right) = \sum_{x \in X(\Omega)} P\{X = x, Y = z - x\} \\ &= \sum_x p_{X,Y}(x, z - x), \end{aligned}$$

ove si è osservato che  $P\{X = x, Y = z - x\} = 0$  se  $x \notin X(\Omega) := \{x \in \mathbb{R} \mid p_X(x) > 0\}$ . L'ultima uguaglianza in (2.11.2) si ottiene scambiando i ruoli di  $X$  e  $Y$ . ■

Nel caso in cui le variabili aleatorie  $X$  e  $Y$  nella Proposizione 2.11.5 siano indipendenti, la (2.11.2) assume la forma

$$(2.5.3) \quad p_{X+Y}(z) = \sum_x p_X(x)p_Y(z - x) = \sum_y p_X(z - y)p_Y(y),$$

che ora useremo.

**Dimostrazione della Proposizione 2.11.4.** Per la (2.11.3) si ha

$$p_{X+Y}(n) = \sum_{k=0}^n e^{-\lambda} e^{-\mu} \frac{\lambda^k}{k!} \frac{\mu^{n-k}}{(n-k)!} = e^{-(\lambda+\mu)} \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} \lambda^k \mu^{n-k} = e^{-(\lambda+\mu)} \frac{(\lambda + \mu)^n}{n!},$$

dove, nell'ultima uguaglianza si è usata la formula di Newton per la potenza di un binomio.

### 2.5.3 Variabili aleatorie geometriche

Abbiamo visto nell'Esempio 1.5.4 che, se si eseguono  $N$  prove ripetute indipendenti per ognuna delle quali l'esito positivo ha probabilità  $p$ , allora la probabilità di ottenere il primo esito positivo all' $n$ -esimo tentativo, con  $n \leq N$ , è  $p(1-p)^{n-1}$ . Per liberarsi della restrizione  $n \leq N$ , immaginiamo di eseguire una successione infinita di prove ripetute  $N = +\infty$ . C'è un problema, però: l'insieme che descrive gli esiti di una successione infinita di prove,  $\Omega = \{0, 1\}^{\mathbb{N}^*}$ , non è numerabile, e dunque non è immediato definire una misura di probabilità  $P$  su di esso. Come vedremo nel prossimo Capitolo, tale problema può essere superato. Tralasciando i dettagli tecnici è possibile definire una variabile aleatoria  $X$  tale che l'evento  $\{X = n\}$  corrisponde a "il primo esito positivo è stato ottenuto al tentativo  $n$ -esimo", ed ha probabilità  $p(1-p)^{n-1}$ . In altre parole  $X$  è l'istante in cui si verifica il primo successo.

In generale, diremo che una variabile aleatoria  $X$  a valori in  $\mathbb{N}$  è una variabile aleatoria **geometrica** di parametro  $p \in (0, 1)$ , e scriveremo

$$X \sim Ge(p),$$

se  $X$  assume valori su  $\mathbb{N}^*$  e

$$p_X(n) = p(1-p)^{n-1} \quad \forall n \geq 1.$$

Le variabili aleatorie geometriche godono di una proprietà rilevante, detta **assenza di memoria**, che si può esprimere con l'affermazione: *se nei primi  $n$  tentativi non è stato ottenuto alcun successo, la probabilità di dover attendere altri  $m$  tentativi prima del primo successo non dipende da  $n$* . La conseguente inutilità di puntare, ad esempio, sui numeri ritardatari nel gioco del Lotto è, speriamo, evidente!

**Proposizione 2.5.9** *Sia  $X \sim Ge(p)$ . Allora, per ogni  $n, m \geq 0$ ,*

$$(2.5.4) \quad P\{X > n+m \mid X > n\} = P\{X > m\}.$$

**Dimostrazione.** Anzitutto si osservi che, poichè  $\{X > n+m\} \subseteq \{X > n\}$ , si ha che

$$P\{X > n+m \mid X > n\} = \frac{P\{X > n+m, X > n\}}{P\{X > n\}} = \frac{P\{X > n+m\}}{P\{X > n\}}.$$

Dobbiamo dunque calcolare probabilità del tipo  $P\{X > n\}$ . Si ottiene:

$$P\{X > n\} = \sum_{k=n+1}^{+\infty} p(1-p)^{k-1} = p(1-p)^n \sum_{h=0}^{+\infty} (1-p)^h = (1-p)^n,$$

dove si è usato il fatto di saper calcolare la somma di una serie geometrica:

$$\sum_{h=0}^{+\infty} (1-p)^h = \frac{1}{1-(1-p)} = \frac{1}{p}.$$

Ma allora:

$$P\{X > n+m \mid X > n\} = \frac{P\{X > n+m\}}{P\{X > n\}} = \frac{(1-p)^{n+m}}{(1-p)^n} = (1-p)^m = P\{X > m\}.$$

■

La proprietà di perdita di memoria caratterizza le variabili aleatorie geometriche nella classe delle variabili aleatorie a valori naturali, come mostriamo nella seguente Proposizione.

**Proposizione 2.5.10** *Sia  $X$  una variabile aleatoria a valori in  $\mathbb{N}$ , tale che (2.11.1) è verificata per ogni  $n, m \geq 0$ . Allora o  $X$  è una variabile aleatoria Geometrica oppure  $P\{X = 0\} = 1$ .*

**Dimostrazione.** Posto  $m = 1$  in (2.11.1), si ottiene

$$P\{X > n+1\} = P\{X > n\}P\{X > 1\}$$

che, per induzione, implica

$$P\{X > n\} = P\{X > 1\}^n,$$

per ogni  $n \geq 1$ . Ma allora, osservando che  $\{X = n\} = \{X > n-1\} \setminus \{X > n\}$ ,

$$\begin{aligned} p_X(n) &= P\{X = n\} = P\{X > n-1\} - P\{X > n\} = P\{X > 1\}^n - P\{X > 1\}^{n+1} = \\ &= (1 - P\{X > 1\})P\{X > 1\}^n, \end{aligned}$$

per ogni  $n \geq 1$ . Inoltre

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

Dunque, posto  $p = 1 - P\{X > 1\}$ , si ha che

$$p_X(n) = p(1 - p)^{n-1}$$

per ogni  $n \geq 1$ , che conclude la dimostrazione. ■

## 2.6 Funzione di ripartizione

**Definizione 2.6.1** Sia  $X$  una variabile aleatoria discreta reale. La funzione  $F_X : \mathbb{R} \rightarrow [0, 1]$  definita da

$$F_X(x) = P(X \leq x)$$

si dice **funzione di ripartizione** della variabile aleatoria  $X$ .

Alcune proprietà elementari delle funzione di ripartizione sono mostrate nella seguente proposizione.

**Proposizione 2.6.2** Sia  $F_X$  la funzione di ripartizione di una variabile aleatoria  $X$ . Allora

- i.  $F_X$  è non decrescente
- ii.  $F_X$  è continua da destra.
- iii.

$$\lim_{x \rightarrow -\infty} F_X(x) = 0.$$

iv.

$$\lim_{x \rightarrow +\infty} F_X(x) = 1.$$

**Dimostrazione.**

- i. Se  $x < y$  allora  $\{X \leq x\} \subseteq \{X \leq y\}$ . Perciò

$$F_X(x) = P\{X \leq x\} \leq P\{X \leq y\} = F_X(y).$$

- ii. Sia  $x \in \mathbb{R}$ . Basta dimostrare che se  $(x_n)$  è una successione decrescente tale che  $x_n \downarrow x$  allora

$$(2.6.1) \quad \lim_{n \rightarrow +\infty} F_X(x_n) = F(x).$$

Si osservi che

$$\{X \leq x\} = \bigcap_n \{X \leq x_n\},$$

e che la successione di eventi  $\{X \leq x_n\}$  è decrescente. Per la Proposizione 1.2.2 si ha

$$F_X(x) = P\{X \leq x\} = \lim_{n \rightarrow +\infty} P\{X \leq x_n\} = \lim_{n \rightarrow +\infty} F_X(x_n).$$

iii. È sufficiente applicare la Proposizione 1.2.2 alla famiglia decrescente di eventi  $\{X \leq x_n\}$  dove  $x_n \downarrow -\infty$ , osservando che

$$\bigcap_n \{X \leq x_n\} = \emptyset.$$

iv. È sufficiente applicare la Proposizione 1.2.2 alla famiglia crescente di eventi  $\{X \leq x_n\}$  dove  $x_n \uparrow +\infty$ , osservando che

$$\bigcup_n \{X \leq x_n\} = \Omega.$$

■

Passiamo ora a studiare la funzione di ripartizione di una variabile aleatoria discreta. Innanzitutto, per quanto visto nell'Osservazione 2.3.2, la funzione di ripartizione si può esprimere in termini della densità come segue:

$$F_X(x) = \sum_{y \leq x} p_X(y).$$

Viceversa, mostriamo che la densità è esprimibile in termini della funzione di ripartizione. In particolare, la funzione di ripartizione determina completamente la distribuzione di una variabile aleatoria discreta: questo sarà conseguenza del seguente risultato generale.

**Proposizione 2.6.3** *Se  $X$  è una variabile aleatoria reale, per ogni  $x \in \mathbb{R}$  vale la relazione*

$$p_X(x) = F_X(x) - F_X(x^-),$$

dove

$$F_X(x^-) = \lim_{y \rightarrow x^-} F_X(y).$$

**Dimostrazione.** Se  $y < x$ , si ha che

$$P\{X \in (y, x]\} = P\{X \leq x\} - P\{X \leq y\} = F_X(x) - F_X(y).$$

Allora, se  $(y_n)$  è una successione di elementi di  $(-\infty, x)$  tali che  $y_n \uparrow x$ , poichè gli eventi  $\{X \in (y_n, x]\}$  formano una successione decrescente per cui  $\bigcap_n \{X \in (y_n, x]\} = \{X = x\}$ , si ha

$$\begin{aligned} p_X(x) &= P\{X = x\} = \lim_{n \rightarrow +\infty} P\{X \in (y_n, x]\} = \lim_{n \rightarrow +\infty} [F_X(x) - F_X(y_n)] = \\ &= F_X(x) - \lim_{n \rightarrow +\infty} F_X(y_n) = F_X(x) - F_X(x^-). \end{aligned}$$

■

**Osservazione 2.6.1** Dalla Proposizione precedente segue che  $F_X$  è discontinua in  $x$  se e solo se  $p_X(x) > 0$ , e il valore di  $p_X(x)$  rappresenta l'entità del "salto" di  $F_X$  in  $x$ . Nel caso in cui  $X(\Omega) = \{x \in \mathbb{R} \mid p_X(x) > 0\}$  non abbia punti di accumulazione, dall'identità

$$F_X(x) = \sum_{y \leq x} p_X(y).$$

segue che  $F_X$  è una funzione costante a tratti.

La funzione di ripartizione è spesso utile nei calcoli con le distribuzioni di variabili aleatorie. In particolare, come ora vedremo, è usata quando si tratta di calcolare la distribuzione del massimo o del minimo di variabili aleatorie indipendenti.

### 2.6.1 Massimo e minimo di variabili aleatorie indipendenti

**Proposizione 2.6.4** *Siano  $X_1, X_2, \dots, X_n$  variabili aleatorie reali indipendenti. Definiamo:*

$$Z = \max(X_1, \dots, X_n), \quad W = \min(X_1, \dots, X_n).$$

Allora

$$(2.6.2) \quad F_Z(x) = \prod_{k=1}^n F_{X_k}(x),$$

e

$$(2.6.3) \quad F_W(x) = 1 - \prod_{k=1}^n [1 - F_{X_k}(x)].$$

**Dimostrazione.** Cominciamo col dimostrare (2.11.5). Si osservi che

$$\{Z \leq x\} = \bigcap_{k=1}^n \{X_k \leq x\},$$

usando l'indipendenza delle  $X_k$  si ha

$$F_Z(x) = P\{Z \leq x\} = \prod_{k=1}^n P\{X_k \leq x\} = \prod_{k=1}^n F_{X_k}(x).$$

Per quanto riguarda (2.11.6), si vede che

$$\{W > x\} = \bigcap_{k=1}^n \{X_k > x\}.$$

e perciò

$$1 - F_W(x) = P\{W > x\} = \prod_{k=1}^n P\{X_k > x\} = \prod_{k=1}^n [1 - P\{X_k \leq x\}] = \prod_{k=1}^n [1 - F_{X_k}(x)].$$

da cui segue (2.11.6). ■

**Esempio 2.6.2** Siano  $X_1, \dots, X_n \sim Ge(p)$  indipendenti. Determiniamo la densità di  $Z = \max(X_1, \dots, X_n)$  e  $W = \min(X_1, \dots, X_n)$ .

Come si vede dai calcoli fatti nella Proposizione 2.11.2,

$$F_{X_i}(k) = 1 - (1 - p)^{k+1}.$$

Dunque

$$F_Z(k) = [1 - (1 - p)^{k+1}]^n.$$

Essendo, per  $k \in \mathbb{N}$ ,  $F_Z(k^-) = F_Z(k - 1)$ , si ha, usando la Proposizione 2.11.8

$$p_Z(k) = F_Z(k) - F_Z(k - 1) = [1 - (1 - p)^{k+1}]^n - [1 - (1 - p)^k]^n.$$

Analogamente,

$$F_W(k) = 1 - [(1 - p)^{k+1}]^n = 1 - [(1 - p)^n]^{k+1},$$

che coincide con la funzione di ripartizione di una variabile aleatoria geometrica di parametro  $1 - (1 - p)^n$ . Poichè la funzione di ripartizione individua completamente la distribuzione, possiamo concludere che

$$W \sim Ge(1 - (1 - p)^n).$$