

# Modelli Probabilistici

Prima parte

Leo Pasquazzi\*

Dipartimento di Statistica e Metodi Quantitativi

Università degli Studi di Milano - Bicocca

17 maggio 2019

## 1 Introduzione

Un **modello probabilistico** è una descrizione di un **esperimento casuale**, ovvero di un processo che si

---

\*Questa dispensa riassume gli argomenti trattati dall'autore durante le lezioni dell'insegnamento di "Statistica per la Finanza" del corso di laurea in Economia delle Banche, delle Assicurazioni e degli Intermediari Finanziari tenute presso l'Università degli Studi di Milano - Bicocca durante l'anno accademico 2017/18.

svolge nel mondo reale il cui esito è incerto. Modelli probabilistici sono spesso utili per prendere decisioni in situazioni di incertezza, ovvero quando non si è sicuri su come si evolveranno eventi futuri.

Secondo l'impostazione assiomatica di Kolmogorov<sup>1</sup>, un modello probabilistico è composto da tre elementi:

- 1) Da uno **spazio campionario**  $\Omega$ , ovvero da un insieme che contiene tutti gli **eventi elementari** che potrebbero verificarsi al termine di un esperimento casuale. Lo spazio campionario  $\Omega$  deve essere **esaustivo** nel senso che ad ogni possibile esito finale dell'esperimento casuale di riferimento deve corrispondere un evento elementare  $\omega \in \Omega$ , e gli eventi elementari  $\omega \in \Omega$  devono essere **mutuamente esclusivi** (ad ogni esito possibile deve corrispondere *esattamente un* evento elementare  $\omega \in \Omega$ ).

---

<sup>1</sup>La teoria moderna della probabilità si fonda sull'impostazione assiomatica proposta dal matematico russo Andrey Kolmogorov nella sua monografia "Grundbegriffe der Wahrscheinlichkeitsrechnung" pubblicata nel 1933.

- 2) Da uno spazio degli eventi  $\mathcal{A}$ , ovvero da una classe (o collezione) di sottoinsiemi di  $\Omega$  che rappresentano eventi. Più avanti si vedrà che lo spazio degli eventi  $\mathcal{A}$  di un modello probabilistico deve soddisfare determinate condizioni che lo qualificano come una  $\sigma$ -algebra.
  
- 3) Da una funzione di probabilità  $P : \mathcal{A} \mapsto [0, 1]$  che ad ogni evento  $A \in \mathcal{A}$  associa una "probabilità"  $P(A)$  il cui valore è compreso nell'intervallo  $[0, 1]$ . Queste probabilità esprimono delle misure per la verosimiglianza degli eventi inclusi nella classe  $\mathcal{A}$  e quindi, oltre a dover essere definite in modo *realistico*, devono anche soddisfare talune *condizioni di coerenza*. Le condizioni di coerenza in questione sono note come *assiomi di Kolmogorov* e verranno esposte più avanti. Esse assicurano che le probabilità  $P(A)$  degli eventi possano essere interpretate come siamo abituati a farlo. Gli assiomi di Kolmogorov assicurano, per esempio, che il valore  $P(A) = 0$  sia riservato ad eventi che secondo

il modello non possono verificarsi, che il valore  $P(A) = 1$  sia riservato ad eventi che secondo il modello si verificheranno con assoluta certezza, e che valori intermedi di  $P(A)$  siano riservati ad eventi che secondo il modello potrebbero verificarsi o meno. In particolare, assicurando che le probabilità siano *additive*<sup>2</sup>, gli assiomi di Kolmogorov assicurano che un modello probabilistico associ probabilità più elevate a eventi che forniscono una descrizione più generale dell'esito finale dell'esperimento casuale di riferimento.

In quanto segue, per indicare un modello probabilistico con spazio campionario  $\Omega$ , spazio degli eventi  $\mathcal{A}$  e funzione di probabilità  $P : \mathcal{A} \mapsto [0, 1]$  si scriverà  $(\Omega, \mathcal{A}, P)$ .

Prima di procedere con i dettagli tecnici della definizione di un modello probabilistico  $(\Omega, \mathcal{A}, P)$ , conviene rimarcare che **qualunque modello probabili-**

---

<sup>2</sup>L'esatta definizione di questo concetto verrà fornita più avanti.

stico è sempre solo una creazione delle mente umana e come tale può fornire una descrizione più o meno "realistica" di un esperimento casuale e quindi può essere più o meno utile per formulare delle previsioni sull'esito finale dell'esperimento casuale a cui si riferisce.

## 2 Lo spazio degli eventi

Come già accennato nella sezione introduttiva, gli eventi che possono verificarsi al termine di un esperimento casuale possono essere visualizzati come sottoinsiemi dello spazio campionario  $\Omega$ , e lo spazio degli eventi  $\mathcal{A}$  è una collezione di tali sottoinsiemi. In particolare, lo spazio degli eventi  $\mathcal{A}$  contiene tutti gli eventi  $A \subseteq \Omega$  ai quali la funzione di probabilità  $P : \mathcal{A} \mapsto [0, 1]$  assegna una probabilità. Siccome questo significa che il valore di  $P(A)$  non è definito se un evento  $A$  non appartiene allo spazio degli eventi  $\mathcal{A}$ , nelle applicazioni si cerca solitamente di definire un modello probabilistico  $(\Omega, \mathcal{A}, P)$  con spazio degli eventi  $\mathcal{A}$  possibilmente ampio. Quando possi-

bile, si cercherà di definire un modello probabilistico con  $\mathcal{A} = \mathcal{P}(\Omega)$ , dove  $\mathcal{P}(\Omega)$  indica il cosiddetto **insieme delle parti di  $\Omega$** , ovvero l'insieme di tutti i sottoinsiemi di  $\Omega$ .

Tuttavia, bisogna considerare anche il rovescio della medaglia: **la scelta di uno spazio di eventi  $\mathcal{A}$  troppo ampio comporta maggiori difficoltà nella fase di assegnazione della probabilità**. Infatti, una volta scelto uno spazio di eventi  $\mathcal{A}$  che contiene tutti gli eventi  $A \subseteq \Omega$  che si ritengono rilevanti, bisogna anche procedere alla definizione di una funzione di probabilità  $P : \mathcal{A} \mapsto [0, 1]$  che associa delle probabilità agli eventi  $A \in \mathcal{A}$ . Quest'ultima deve essere definita in modo **coerente** e **realistico**. Come già accennato in precedenza, per funzione di probabilità "coerente" si intende una funzione di probabilità  $P : \mathcal{A} \mapsto [0, 1]$  che rispetta gli assiomi di Kolmogorov che verranno elencati più avanti. Ma oltre a dover soddisfare gli assiomi di Kolmogorov, una funzione di probabilità deve essere anche "realistica", ovvero deve essere definita in modo tale che i valori delle probabilità  $P(A)$  rispecchino la verosimiglianza degli eventi

$A \in \mathcal{A}$ . Non è difficile intuire che *la definizione di una funzione di probabilità "coerente" e "realistica" sia un'operazione tanto più complicata quanto più ampio è lo spazio degli eventi  $\mathcal{A}$ .*

Le precedenti considerazioni suggeriscono in ogni caso di scegliere uno spazio di eventi  $\mathcal{A}$  abbastanza ampio da essere "chiuso" rispetto ad alcune operazioni fondamentali che spesso si applicano agli eventi. Le operazioni fondamentali in questione comprendono:

- a) L'operazione di **negazione**, che ad ogni evento  $A$  associa il cosiddetto **evento complementare**  $\bar{A}$  definito come l'insieme di tutti gli eventi elementari  $\omega \in \Omega$  che non appartengono all'evento  $A$ .

Si osservi che l'evento complementare  $\bar{A}$  si verifica se e solo se **non** si verifica l'evento  $A$ .

- b) L'operazione di **unione**, che ad ogni coppia di eventi  $A_1$  e  $A_2$  associa la loro **unione**  $A_1 \cup A_2$  definita come l'insieme di tutti gli eventi

elementari  $\omega \in \Omega$  che appartengono ad almeno uno dei due eventi  $A_1$  oppure  $A_2$ .

Si osservi che l'evento unione  $A_1 \cup A_2$  si verifica se e solo se si verifica **almeno uno** dei due eventi  $A_1$  **oppure**  $A_2$ .

- c) L'operazione di **intersezione**, che ad ogni coppia di eventi  $A_1$  e  $A_2$  associa la loro **intersezione**  $A_1 \cap A_2$  definita come l'insieme di tutti gli eventi elementari  $\omega \in \Omega$  che appartengono contemporaneamente sia ad  $A_1$  che ad  $A_2$ .

Si osservi che l'evento intersezione  $A_1 \cap A_2$  si verifica se e solo se si verificano **contemporaneamente** sia l'evento  $A_1$  che l'evento  $A_2$ .

Se  $A_1 \cap A_2 \neq \emptyset$ , allora si dice che  $A_1$  e  $A_2$  sono due **eventi compatibili**. Altrimenti si dice che  $A_1$  e  $A_2$  sono **eventi incompatibili** o **mutuamente esclusivi**.

Si osservi che le **operazioni di unione e intersezione** possono essere applicate anche a qualsiasi **collezione finita** di eventi, e anche a collezioni **infinita**.



te numerabili e collezioni infinite non numerabili di eventi.<sup>3</sup> Infatti, se  $\mathcal{C}$  è una collezione di eventi, allora l'unione e l'intersezione della collezione di eventi  $\mathcal{C}$  sono rispettivamente definite come

$$\bigcup_{A \in \mathcal{C}} A = \{\omega \in \Omega : \omega \in A \text{ per almeno un evento } A \in \mathcal{C}\}$$

e

$$\bigcap_{A \in \mathcal{C}} A = \{\omega \in \Omega : \omega \in A \text{ per ogni } A \in \mathcal{C}\}$$

---

<sup>3</sup>Si dice che un insieme (o una collezione, classe, famiglia) è "numerabile" se contiene un numero finito di elementi oppure se esiste una corrispondenza biunivoca tra i numeri naturali e gli elementi dell'insieme. Nel secondo caso si dice anche che l'insieme (la collezione, la classe o la famiglia) è "infinito numerabile" oppure che contiene "un'infinità numerabile" di elementi. *Tutti* gli elementi di un insieme numerabile possono sempre essere disposti in una "sequenza" o "successione" scrivendo  $x_1, x_2, \dots$ . Ovviamente, si possono anche definire sequenze o successioni (numerabili) che ripetono più volte lo stesso elemento. Se si vuole rimarcare che una sequenza è finita, si può indicare la sequenza aggiungendo un elemento "finale":  $x_1, x_2, \dots, x_n$ . Chiaramente, si possono anche definire delle sequenze  $x_1, x_2, \dots$  con elementi presi da un insieme infinito non numerabile. Tuttavia, in questo caso non sarà possibile definire una sequenza numerabile che contenga *tutti* gli elementi dell'insieme (infinito) non numerabile.

Avendo introdotto le definizioni di unione e intersezione riferite a collezioni di eventi, vale la pena osservare che queste due definizioni sono tra di loro collegate attraverso le due **leggi di De Morgan** secondo le quali

$$\overline{\left(\bigcup_{A \in \mathcal{C}} A\right)} = \bigcap_{A \in \mathcal{C}} \bar{A}$$

e

$$\overline{\left(\bigcap_{A \in \mathcal{C}} A\right)} = \bigcup_{A \in \mathcal{C}} \bar{A}.$$

Infine, prima di concludere questa sezione conviene introdurre alcune convenzioni sulla simbologia legata a collezioni di eventi. A tal fine si osservi che se  $\mathcal{C}$  è una collezione di eventi numerabile (finita o infinita), gli eventi appartenenti alla collezione potranno essere posti in corrispondenza biunivoca con un segmento iniziale dei numeri naturali (se la collezione è finita) oppure con tutti i numeri naturali (se la collezione è infinita). Nel primo caso gli eventi

della collezione possono quindi essere indicati con

$$A_1, A_2, \dots, A_n,$$

mentre nel secondo caso possono essere indicati con

$$A_1, A_2, \dots$$

senza un evento "finale". L'evento finale viene anche omesso se si vogliono indicare gli eventi di una collezione numerabile senza specificare se quest'ultima sia finita o infinita. Usando questo tipo di notazione per indicare gli eventi di una collezione numerabile, si potranno indicare l'unione e l'intersezione della collezione scrivendo

$$\bigcup_{i=1}^n A_i \quad \text{ovvero} \quad \bigcap_{i=1}^n A_i \quad (1)$$

nel caso di una collezione finita, e scrivendo

$$\bigcup_{i=1}^{\infty} A_i \quad \text{ovvero} \quad \bigcap_{i=1}^{\infty} A_i \quad (2)$$

nel caso di una collezione infinita. Quando non viene specificato se la collezione è finita oppure infinita, l'unione e l'intersezione della collezione possono essere indicate con

$$\bigcup_i A_i \quad \text{ovvero} \quad \bigcap_i A_i. \quad (3)$$

Le notazioni introdotte nella (1), (2) e nella (3) verranno anche utilizzate per indicare l'unione e l'intersezione riferite alla collezione di tutti gli eventi  $A_i$  che appartengono ad una *successione* di eventi (si noti che in una successione alcuni eventi potrebbero essere ripetuti più volte).

## 2.1 Algebre e $\sigma$ -algebre

Come già accennato in precedenza, lo spazio degli eventi  $\mathcal{A}$  di un modello probabilistico  $(\Omega, \mathcal{A}, P)$  deve sempre soddisfare talune condizioni che lo qualificano come una cosiddetta  $\sigma$ -algebra. In questa sezione verrà fornita la definizione di  $\sigma$ -algebra. Prima di procedere alla definizione di  $\sigma$ -algebra conviene tut-

tavia definire che cosa si intende per un'algebra di eventi.

**Definizione 2.1.** Una classe di eventi  $\mathcal{A}$  viene chiamata algebra (su  $\Omega$ ) se e solo se soddisfa i seguenti requisiti:

- a)  $\Omega \in \mathcal{A}$ ;
- b)  $\mathcal{A}$  è "chiusa" rispetto all'operazione di negazione nel senso che se  $A \in \mathcal{A}$ , allora  $\bar{A} \in \mathcal{A}$ ;
- c)  $\mathcal{A}$  è "chiusa" rispetto all'operazione di unione applicata a due eventi nel senso che se  $A_1 \in \mathcal{A}$  e  $A_2 \in \mathcal{A}$ , allora  $A_1 \cup A_2 \in \mathcal{A}$ .

□

Proprietà delle algebre di eventi:

- i) Se  $\mathcal{A}$  è un algebra, allora  $\emptyset \in \mathcal{A}$ .

**Dimostrazione.** Secondo il requisito a) delle algebre,  $\Omega \in \mathcal{A}$ , e usando il requisito b) delle

algebre si vede dunque che

$$\bar{\Omega} = \emptyset \in \mathcal{A}.$$

□

Siccome al termine di un esperimento casuale deve verificarsi uno degli eventi elementari  $\omega \in \Omega$  e siccome l'insieme vuoto  $\emptyset$  non contiene alcun evento elementare, l'insieme vuoto  $\emptyset$  viene chiamato **evento impossibile** mentre lo spazio campionario  $\Omega$  viene chiamato **evento certo**.

- ii) Le algebre sono "chiuse" rispetto all'operazione di unione applicata a collezioni finite (ma non necessariamente infinite) di eventi.

**Dimostrazione.** Se  $\mathcal{A}$  è un'algebra di eventi e se

$$A_1, A_2, \dots, A_n \in \mathcal{A},$$

allora il requisito c) delle algebre implica che

$$B_2 = A_1 \cup A_2 \in \mathcal{A},$$

$$B_3 = \bigcup_{i=1}^3 A_i = B_2 \cup A_3 \in \mathcal{A}$$

e più in generale che

$$B_n = \bigcup_{i=1}^n A_i = B_{n-1} \cup A_n \in \mathcal{A}$$

per ogni  $n$  finito. Questo dimostra che qualsiasi algebra è "chiusa" rispetto all'operazione di unione applicata ad una collezione finita di eventi.

Tuttavia, se

$$A_1, A_2, \dots \in \mathcal{A},$$

è una successione infinita (numerabile) di eventi appartenenti ad un'algebra  $\mathcal{A}$ , allora il requisito c) delle algebre non dice nulla riguardo l'appartenenza all'algebra  $\mathcal{A}$  dell'evento

$$B_\infty = \bigcup_{i=1}^{\infty} A_i$$

( $B_\infty \in \mathcal{A}$  può essere vero ma potrebbe anche essere falso; infatti, come sarebbero definiti gli eventi  $B_{\infty-1}$  e  $A_\infty$  la cui unione dovrebbe dare  $B_\infty$ ?).  $\square$

- iii) Le algebre sono "chiuse" rispetto all'operazione di intersezione applicata a collezioni finite (ma non necessariamente infinite) di eventi.

**Dimostrazione.** Se  $\mathcal{A}$  è un'algebra e se

$$A_1, A_2, \dots, A_n \in \mathcal{A},$$

allora per il requisito b) delle algebre dovrà essere vero che

$$\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n \in \mathcal{A}.$$

Dalla precedente proprietà ii) delle algebre discende quindi che

$$\bigcup_{i=1}^n \bar{A}_i \in \mathcal{A}$$



per ogni  $n$  finito, e siccome per le leggi di De Morgan si ha

$$\bigcup_{i=1}^n \bar{A}_i = \overline{\left( \bigcap_{i=1}^n A_i \right)},$$

si può applicare ancora una volta il requisito b) delle algebre per concludere che

$$\bigcap_{i=1}^n A_i = \overline{\left( \bigcup_{i=1}^n \bar{A}_i \right)} \in \mathcal{A}.$$

□

Come si è visto, le algebre sono classi di eventi molto "ampie" in quanto sono "chiuse" rispetto all'operazione di negazione e rispetto alle operazioni di unione e intersezione applicate a collezioni finite di eventi. Tuttavia, è stato anche mostrato che un'algebra non è necessariamente "chiusa" rispetto alle operazioni di unione e intersezione applicate a collezioni infinite di eventi. Per la definizione di un modello

probabilistico conviene quindi considerare classi di eventi più ampie, ovvero le cosiddette  $\sigma$ -algebre.

**Definizione 2.2.** Una classe di eventi  $\mathcal{A}$  viene chiamata  $\sigma$ -algebra (su  $\Omega$ ) se e solo se soddisfa i requisiti di un'algebra e anche il seguente requisito:

c\*)  $\mathcal{A}$  è "chiusa" rispetto all'operazione di unione applicata a collezioni infinite numerabili di eventi, nel senso che se  $\mathcal{A}$  contiene tutti gli eventi

$$A_1, A_2, \dots$$

appartenenti ad una collezione infinita numerabile, allora  $\mathcal{A}$  contiene anche la loro unione  $\bigcup_{i=1}^{\infty} A_i$ .

□

Proprietà delle  $\sigma$ -algebre di eventi:

- i) Ogni  $\sigma$ -algebra è anche un'algebra (ma non è vero il viceversa).
- ii) Le  $\sigma$ -algebre sono anche "chiusa" rispetto all'operazione di intersezione applicata a collezioni

**infinite numerabili di eventi.** La dimostrazione è analoga a quella della proprietà iii) delle algebre: basta porre  $n = \infty$  e sostituire il richiamo alla proprietà ii) delle algebre con un appello al requisito  $c^*$ ) delle  $\sigma$ -algebre (la legge di De Morgan rimane valida anche se  $n = \infty$ ).

Alcune osservazioni sulle algebre e  $\sigma$ -algebre:

- A) Il requisito  $c^*$ ) delle  $\sigma$ -algebre non implica che esse siano anche "chiuse" rispetto alle operazioni di unione e intersezione applicate a collezioni infinite non numerabili di eventi: l'unione e/o l'intersezione di una collezione non numerabile di eventi appartenenti ad una data  $\sigma$ -algebra **potrebbe** (condizionale!!!) anche non appartenere alla  $\sigma$ -algebra.
- B) **L'insieme delle parti di  $\Omega$** , ovvero la classe di eventi che contiene tutti i sottoinsiemi di  $\Omega$  e che in seguito verrà indicata con  $\mathcal{P}(\Omega)$ , è ovviamente una  $\sigma$ -algebra ed è sempre, per qualsiasi spazio campionario  $\Omega$ , la  $\sigma$ -algebra più

”ampia” tra tutte quelle che si possono definire. D'altra parte, non è difficile rendersi conto che la classe

$$\mathcal{A}_0 = \{\emptyset, \Omega\}$$

è sempre, per qualsiasi spazio campionario  $\Omega$ , la  $\sigma$ -algebra più ”piccola” tra tutte quelle che si possono definire.

- C) Ogni algebra  $\mathcal{A}$  che contiene solo un numero finito di eventi è allo stesso tempo anche una  $\sigma$ -algebra. Infatti, se un'algebra  $\mathcal{A}$  contiene solo un numero finito di eventi, allora non sarà possibile definire delle collezioni infinite di eventi appartenenti ad  $\mathcal{A}$ , ed il requisito  $c^*$ ) delle  $\sigma$ -algebra sarà quindi vacuamente soddisfatto.
- D) Con uno spazio campionario  $\Omega$  finito si possono definire solo algebre e  $\sigma$ -algebre finite perché se  $\Omega$  è finito, allora sarà finita anche la classe di eventi  $\mathcal{P}(\Omega)$ . Infatti, se si indica con  $\#\Omega$  il numero complessivo eventi elementari  $\omega \in \Omega$ , allora  $\mathcal{P}(\Omega)$  conterrà esattamente  $2^{\#\Omega} < \infty$

eventi. Infatti, per ciascuno degli  $\#\Omega$  eventi elementari  $\omega \in \Omega$  ci sono 2 possibilità per decidere se inserirlo in un dato sottoinsieme  $A \subseteq \Omega$  o meno. Il numero complessivo di sottoinsiemi di  $\Omega$  che si possono formare sarà pertanto dato da

$$2 \times 2 \times 2 \cdots (\#\Omega \text{ volte}) \cdots \times 2 = 2^{\#\Omega}$$

(si osservi che questo conteggio include anche l'insieme vuoto  $\emptyset$  e l'insieme  $\Omega$ ).

- E) Dalle due precedenti osservazioni si evince che su uno spazio campionario  $\Omega$  finito, si possono definire solo algebre finite, e che ogni algebra sarà quindi allo stesso tempo anche una  $\sigma$ -algebra.

I seguenti esempi aiuteranno a comprendere meglio le definizioni di algebra e  $\sigma$ -algebra e la differenza che intercorre tra queste due tipologie di classi di eventi.

**Esempio 2.1.** Si consideri un esperimento casuale che consiste nel lancio di una moneta. In questo caso lo spazio campionario può essere definito come

$$\Omega = \{T, C\},$$

dove l'evento elementare "T" indica l'esito testa e l'evento elementare "C" indica l'esito croce. Trattandosi di uno spazio campionario finito, si può concludere che ogni algebra di eventi definita su  $\Omega$  è allo stesso tempo anche una  $\sigma$ -algebra (e viceversa).

La più "ampia" algebra (o  $\sigma$ -algebra) che può essere costruita su questo spazio campionario  $\Omega$  è data da

$$\mathcal{P}(\Omega) = \{\Omega, \emptyset, \{T\}, \{C\}\}.$$

Non è difficile rendersi conto che su  $\Omega$  esiste solo un'altra classe di eventi che possiede i requisiti di un'algebra (o di una  $\sigma$ -algebra). Si tratta della classe

$$\mathcal{A}_0 = \{\Omega, \emptyset\}.$$

Non esistono dunque algebre (o  $\sigma$ -algebre) "interme-

die" tra  $\mathcal{A}_0$  e  $\mathcal{P}(\Omega)$ . □

**Esempio 2.2.** Si consideri un esperimento casuale che consiste nel lancio di un dado a sei facce. In questo caso lo spazio campionario potrà essere definito come

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Siccome lo spazio campionario  $\Omega$  è finito, si può concludere che anche in questo esempio tutte le algebre definite su  $\Omega$  sono allo stesso tempo anche delle  $\sigma$ -algebre (e viceversa). L'insieme delle parti  $\mathcal{P}(\Omega)$  (la  $\sigma$ -algebra più "ampia" che si può costruire) contiene  $2^6 = 64$  sottoinsiemi di  $\Omega$ . Ovviamente

$$\mathcal{A}_0 = \{\Omega, \emptyset\}$$

è ancora la classe di eventi più "povera" tra tutte quelle che soddisfano i requisiti di un algebra (e di una  $\sigma$ -algebra).

Tuttavia, rispetto all'esempio sul lancio della moneta, nel caso del lancio di un dado si possono definire delle algebre (che sono allo stesso tempo anche  $\sigma$ -algebre) "intermedie" tra  $\mathcal{A}_0$  e  $\mathcal{P}(\Omega)$ . Un esempio

è fornito dall'algebra (o  $\sigma$ -algebra)

$$\mathcal{A}_1 = \{\emptyset, \{1, 3, 5\}, \{2, 4, 6\}, \Omega\}$$

che permette di distinguere tra punteggi pari e dispari.  $\square$

Siccome ogni  $\sigma$ -algebra è anche un'algebra (proprietà i) delle  $\sigma$ -algebre) e siccome ogni algebra definita su uno spazio campionario  $\Omega$  finito è allo stesso tempo anche una  $\sigma$ -algebra (osservazione E) sulle algebre e  $\sigma$ -algebre), occorre fare riferimento ad un esperimento casuale con spazio campionario infinito per ottenere un esempio di una classe di eventi  $\mathcal{A}$  che sia un'algebra ma non una  $\sigma$ -algebra.

**Esempio 2.3.** Si consideri

$$\Omega = \mathbb{N} = \{0, 1, 2, \dots\}$$

come spazio campionario e la classe di eventi  $\mathcal{A}$  che contiene tutti gli eventi  $A \subseteq \Omega$  tali che  $A$  oppure  $\bar{A}$  è un insieme finito. Formalmente,

$$\mathcal{A} = \{A \subseteq \Omega : A \text{ oppure } \bar{A} \text{ è un insieme finito}\}.$$



Sotto si dimostrerà che  $\mathcal{A}$  soddisfa i requisiti di un'algebra ma non quelli di una  $\sigma$ -algebra.

**Dimostrazione.** Questa dimostrazione non è richiesta all'esame.

- a) Siccome  $\overline{\Omega} = \emptyset$  e siccome  $\emptyset$  è un insieme finito (contiene zero elementi), si può concludere che  $\Omega \in \mathcal{A}$ . Il requisito a) delle algebre è quindi soddisfatto.
- b) Per verificare se la classe di eventi  $\mathcal{A}$  soddisfa anche il requisito b) delle algebre bisogna verificare se la classe  $\mathcal{A}$  è "chiusa" rispetto all'operazione di negazione. Si consideri dunque un evento  $A \in \mathcal{A}$ . Se  $A$  è un insieme finito, allora  $\overline{A} \in \mathcal{A}$  per la definizione di  $\mathcal{A}$ . D'altra parte,  $A$  non è un insieme finito, allora, sempre per la definizione della classe  $\mathcal{A}$ ,  $A \in \mathcal{A}$  implica che  $\overline{A}$  sia un insieme finito e quindi che  $\overline{A} \in \mathcal{A}$ . Questo dimostra che la classe di eventi  $\mathcal{A}$  soddisfa anche il requisito b) delle algebre.
- c) A questo punto non resta che verificare che la classe di eventi  $\mathcal{A}$  soddisfa anche il requisito c) delle algebre. Bisogna dunque dimostrare che la classe  $\mathcal{A}$  sia "chiusa" rispetto all'operazione di unione applicata a qualunque coppia di eventi  $A_1, A_2 \in \mathcal{A}$ . Si considerino dunque due eventi  $A_1$  e  $A_2$  appartenenti alla classe  $\mathcal{A}$ . Si dimostrerà che  $(A_1 \cup A_2) \in \mathcal{A}$ . A tal fine si distingueranno i seguenti tre casi:
  - c1) Primo caso:  $A_1$  e  $A_2$  sono entrambi insiemi finiti. In questo caso sarà un insieme finito anche l'unione  $A_1 \cup A_2$  e dunque si avrà  $(A_1 \cup A_2) \in \mathcal{A}$  come volevasi dimostrare.
  - c2) Secondo caso:  $A_1$  e  $A_2$  sono entrambi insiemi infiniti. In questo caso, per la definizione di  $\mathcal{A}$ , dovranno essere finiti entrambi gli insiemi complementi  $\overline{A}_1$  e

$\overline{A_2}$  e dunque sarà un insieme finito anche l'intersezione  $\overline{A_1} \cap \overline{A_2}$ . Dalla definizione della classe  $\mathcal{A}$  discende quindi che  $(\overline{A_1} \cap \overline{A_2}) \in \mathcal{A}$  e visto che la classe  $\mathcal{A}$  è chiusa rispetto all'operazione di negazione (al punto b) è stato dimostrato che  $\mathcal{A}$  soddisfa il requisito b) delle algebre), ne consegue che  $\overline{(\overline{A_1} \cap \overline{A_2})} \in \mathcal{A}$ . Ma dalle leggi di De Morgan discende che

$$\overline{(\overline{A_1} \cap \overline{A_2})} = A_1 \cup A_2$$

e quindi si conclude che  $(A_1 \cup A_2) \in \mathcal{A}$  come volevasi dimostrare.

- b3) Terzo caso:  $A_1$  è un insieme finito e  $A_2$  è un insieme infinito (o viceversa). In questo caso,  $A_1 \cup A_2$  sarà un insieme infinito e quindi si avrà  $(A_1 \cup A_2) \in \mathcal{A}$  se e solo se  $\overline{(A_1 \cup A_2)}$  è un insieme finito. Per dimostrare che ciò è vero, basta osservare che per le leggi di De Morgan si ha

$$\overline{(A_1 \cup A_2)} = \overline{A_1} \cap \overline{A_2}$$

e che il secondo membro è sicuramente un sottoinsieme di  $\overline{A_2}$ . Per completare la dimostrazione sarà quindi sufficiente dimostrare che  $\overline{A_2}$  è un insieme finito. Ma questo è un'immediata conseguenza delle ipotesi che  $A_2 \in \mathcal{A}$  e che  $A_2$  è un insieme infinito. Anche nel terzo caso risulta dunque dimostrato che  $(A_1 \cup A_2) \in \mathcal{A}$ .

I precedenti ragionamenti dimostrano che la classe di eventi  $\mathcal{A}$  soddisfa i requisiti di un'algebra di eventi. Tuttavia,  $\mathcal{A}$  non è una  $\sigma$ -algebra di eventi perché non soddisfa il requisito c\*) delle  $\sigma$ -algebre. Per rendersene conto, basta considerare la successione di insiemi

$$A_0 = \{0\}, \quad A_1 = \{2\}, \quad A_2 = \{4\}, \quad \dots, A_i = \{2i\}, \quad \dots$$

dove il generico insieme  $A_i$  della successione contiene il numero pari  $2i$ . Si osservi che tutti gli eventi  $A_i$  della successione contengono solo un elemento di  $\Omega = \mathbb{N}$  e che quindi

$$A_i \in \mathcal{A} \quad \text{per ogni } i = 0, 1, 2, \dots$$

Tuttavia,

$$\bigcup_{i=0}^{\infty} A_i = \{0, 2, 4, 6, \dots\}$$

è un insieme infinito perché contiene tutti i numero pari. D'altra parte, anche

$$\overline{\left( \bigcup_{i=0}^{\infty} A_i \right)} = \{1, 3, 5, 7, \dots\}$$

è un insieme infinito perché contiene tutti i numeri dispari. Ne consegue che

$$\bigcup_{i=0}^{\infty} A_i \notin \mathcal{A}$$

e quindi si può concludere che  $\mathcal{A}$  non soddisfa il requisito c\*) delle  $\sigma$ -algebre. □

□

La classe di eventi  $\mathcal{A}$  considerata nell'esempio precedente è stata definita in modo un po' artificiale per fornire un'esempio semplice di una classe di eventi che sia un'algebra ma non una  $\sigma$ -algebra. Nei seguenti esempi verranno invece definite delle alge-

bre e  $\sigma$ -algebre che per moltissime applicazioni sono di fondamentale importanza.

**Esempio 2.4.** Si consideri un esperimento casuale il cui esito finale può essere descritto con un numero reale. In questo caso si potrà considerare come spazio campionario  $\Omega$  l'insieme dei numeri reali  $\mathbb{R}$ . Come è noto,  $\Omega = \mathbb{R}$  contiene un'*infinità non numerabile* di numeri reali. Su uno spazio campionario così ricco di eventi elementari si possono definire tantissime classi di eventi che soddisfano i requisiti di un'algebra e/o quelli di una  $\sigma$ -algebra.

In ambito probabilistico (e non solo), sono di fondamentale importanza le due seguenti classi di eventi:

- 1) La classe  $\mathcal{B}^*$  che contiene l'evento impossibile  $\emptyset$  e tutti gli intervalli e le unioni di collezioni finite di intervalli (per intervalli si intendono intervalli aperti, chiusi, semiaperti, limitati e illimitati). Questa classe di eventi contiene ovviamente un'*infinità non numerabile* di eventi. Si può dimostrare che la classe  $\mathcal{B}^*$  soddisfa i

requisiti delle algebre ma non il requisito  $c^*$ ) delle  $\sigma$ -algebre.

- 2) La cosiddetta  **$\sigma$ -algebra di Borel  $\mathcal{B}$**  che può essere definita come la più piccola  $\sigma$ -algebra che include la classe di eventi  $\mathcal{B}^*$ . Per questo motivo la  $\sigma$ -algebra di Borel può essere anche indicata con  $\sigma(\mathcal{B}^*)$ . Ovviamente, anche  $\mathcal{B}$  contiene un'infinità non numerabile di eventi. Tuttavia, si può dimostrare che la  $\sigma$ -algebra di Borel  $\mathcal{B}$  è *strettamente* inclusa nell'insieme delle parti  $\mathcal{P}(\mathbb{R})$  (ovvero che esistono sottoinsiemi di  $\mathbb{R}$  che non appartengono a  $\mathcal{B}$ ). Ciò nonostante,  $\mathcal{B}$  è una classe di eventi molto ampia che contiene tutti gli eventi che potrebbero avere una qualche rilevanza pratica nelle applicazioni. In molti testi, la  $\sigma$ -algebra di Borel (su  $\mathbb{R}$ ) viene definita come la più piccola  $\sigma$ -algebra che contiene tutti i sottoinsiemi aperti di  $\mathbb{R}$ . Si può dimostrare che questa definizione è equivalente alla definizione data in precedenza.

□

**Esempio 2.5.** Si consideri un esperimento casuale il cui esito può essere descritto con un vettore di  $k$  numeri reali  $(x_1, x_2, \dots, x_k) \in \mathbb{R}^k$ . Anche sullo spazio campionario  $\Omega = \mathbb{R}^k$  si possono definire una miriade di classi di eventi che soddisfano i requisiti delle algebre e/o delle  $\sigma$ -algebre. In ambito probabilistico (e non solo), sono di fondamentale importanza le seguenti due classi di eventi:

- 1) La classe  $\mathcal{B}_k^*$  che contiene l'evento impossibile  $\emptyset$ , tutti i "rettangoli"  $k$ -dimensionali e tutte le unioni di collezioni finite di "rettangoli"  $k$ -dimensionali. Per "rettangoli"  $k$ -dimensionali si intendono sottoinsiemi di  $\mathbb{R}^k$  che possono essere espressi come prodotti cartesiani di  $k$  intervalli di qualsiasi tipo (aperti, semiaperti, chiusi, limitati e illimitati). I "rettangoli"  $k$ -dimensionali possono essere indicati con

$$\bigtimes_{i=1}^k I_i = \left\{ (x_1, x_2, \dots, x_k) \in \mathbb{R}^k : \right. \\ \left. x_i \in I_i \text{ per ogni } i = 1, 2, \dots, k \right\},$$

dove  $I_i$  denota un intervallo (aperto, semiaperto, chiuso; limitato o illimitato). Si noti che nel caso  $k = 1$  i "rettangoli"  $k$ -dimensionali sono semplicemente degli intervalli. Nel caso  $k = 1$  si ha dunque  $\mathcal{B}_k^* = \mathcal{B}^*$ , dove  $\mathcal{B}^*$  è l'algebra di eventi introdotta nell'esempio precedente. Ovviamente, anche la classe  $\mathcal{B}_k^*$  contiene un'infinità non numerabile di eventi, e si può dimostrare che questa classe di eventi soddisfa i requisiti di un'algebra ma non quelli di una  $\sigma$ -algebra.

- 2) La  $\sigma$ -algebra di Borel su  $\mathbb{R}^k$ , che in seguito verrà indicata con  $\mathcal{B}_k = \sigma(\mathcal{B}_k^*)$  essendo definita come la più piccola  $\sigma$ -algebra che include l'algebra di eventi  $\mathcal{B}_k^*$ . Si osservi che  $\mathcal{B}_k$  è una classe di eventi molto ampia che contiene tutti gli eventi che potrebbero avere una qualche rilevanza pratica nelle applicazioni. Infatti, anche  $\mathcal{B}_k$  contiene un'infinità non numerabile di eventi, e anche nel caso in cui  $k > 1$  si può dimostrare che la  $\sigma$ -algebra di Borel  $\mathcal{B}_k$

è *strettamente* inclusa nell'insieme delle parti  $\mathcal{P}(\mathbb{R}^k)$ . Infine, vale la pena notare che anche per  $k > 1$  la  $\sigma$ -algebra di Borel  $\mathcal{B}_k$  potrebbe essere definita come la più piccola  $\sigma$ -algebra che contiene tutti i sottoinsiemi aperti di  $\mathbb{R}^k$  (in altre parole: la più piccola  $\sigma$ -algebra che contiene tutti i sottoinsiemi aperti di  $\mathbb{R}^k$  coincide con la più piccola  $\sigma$ -algebra che include l'algebra  $\mathcal{B}_k^*$ ).

□

Gli ultimi due esempi illustrano un metodo che viene spesso utilizzato per definire delle  $\sigma$ -algebre. Si noti che in entrambi gli esempi si è partiti da una data classe di eventi ristretta  $\mathcal{A}^*$  (nei due esempi in questione si è partiti da un'algebra, ma questo fatto è irrilevante per le considerazioni che seguono), e a partire dagli eventi contenuti in  $\mathcal{A}^*$  è stata definita una  $\sigma$ -algebra  $\mathcal{A} = \sigma(\mathcal{A}^*)$  stabilendo che  $\mathcal{A}$  fosse la più piccola  $\sigma$ -algebra che contiene tutti gli eventi appartenenti alla classe di partenza  $\mathcal{A}^*$ . Siccome la  $\sigma$ -algebra  $\mathcal{A} = \sigma(\mathcal{A}^*)$  viene “creata” a



partire dagli eventi contenuti nella classe  $\mathcal{A}^*$ , la  $\sigma$ -algebra  $\mathcal{A} = \sigma(\mathcal{A}^*)$  viene spesso anche chiamata “ $\sigma$ -algebra generata dalla classe di eventi  $\mathcal{A}^*$ ”. Il seguente esempio illustra che la  $\sigma$ -algebra generata da una data classe di eventi  $\mathcal{A}^*$  potrebbe contenere molti più eventi della classe di partenza.

**Esempio 2.6.** Si considerino due eventi  $A, B \subset \Omega$ , e si assuma che

$$A \cap B, \quad \bar{A} \cap B, \quad A \cap \bar{B}, \quad \bar{A} \cap \bar{B} \quad (4)$$

non siano insiemi vuoti. Tracciando un diagramma di Venn non è difficile rendersi conto che gli eventi  $A$  e  $B$  suddividono lo spazio campionario  $\Omega$  in quattro parti ciascuna delle quali può essere identificata con una delle quattro intersezioni nella (4). Usando il linguaggio della teoria degli insiemi si può quindi dire che le quattro intersezioni nella (4) formano una **partizione di  $\Omega$** .

Si consideri ora la classe di eventi

$$\mathcal{A}^* = \{A, B\}$$

e la più piccola  $\sigma$ -algebra  $\mathcal{A} = \sigma(\mathcal{A}^*)$  che contiene tutti gli eventi appartenenti alla classe  $\mathcal{A}^*$ , ovvero la più piccola  $\sigma$ -algebra che contiene i due eventi  $A$  e  $B$ . Si può dimostrare che *la  $\sigma$ -algebra  $\mathcal{A} = \sigma(\mathcal{A}^*)$  coincide con la classe  $\mathcal{C}$  che contiene gli eventi che si ottengono applicando l'operazione di unione a tutte le possibili collezioni di eventi che si possono formare con le quattro intersezioni nella (4) (compresa la collezione vuota che darà luogo all'evento impossibile  $\emptyset$ )*. Siccome a partire dalle quattro intersezioni nella (4) si possono formare  $2^4 = 16$  collezioni di eventi,<sup>4</sup> la classe di eventi  $\mathcal{C}$  dovrà contenere esattamente  $2^4 = 16$  eventi. Gli eventi appartenenti alla classe  $\mathcal{C}$  sono qui di seguito elencati:

C0) L'unico evento che si ottiene applicando l'operazione di unione all'unica collezione che non contiene *nessuno* degli eventi nella (4), ovvero

---

<sup>4</sup>Infatti, per ciascuna delle 4 intersezioni nella (4) si può decidere in due modi se includerla o meno in una data collezione. Il numero di collezioni possibili sarà pertanto dato da  $2 \times 2 \times 2 \times 2 = 2^4 = 16$ .

l'evento

$$\emptyset.$$

- C1) I 4 eventi che si ottengono applicando l'operazione di unione alle quattro collezioni che contengono *uno* solo degli eventi nella (4), ovvero gli eventi

$$A \cap B, \quad \bar{A} \cap B, \quad A \cap \bar{B}, \quad \bar{A} \cap \bar{B};$$

- C2) i 6 eventi che si ottengono applicando l'operazione di unione alle 6 collezioni che contengono *due* degli eventi nella (4), ovvero gli eventi

$$A = (A \cap B) \cup (A \cap \bar{B}), \quad B = (A \cap B) \cup (\bar{A} \cap B),$$

$$\bar{A} = (\bar{A} \cap B) \cup (\bar{A} \cap \bar{B}), \quad \bar{B} = (A \cap \bar{B}) \cup (\bar{A} \cap \bar{B});$$

$$(A \cap B) \cup (\bar{A} \cap \bar{B}), \quad (A \cap \bar{B}) \cup (\bar{A} \cap B).$$

- C3) i 4 eventi che si ottengono applicando l'operazione di unione alle 4 collezioni che contengono *tre* degli eventi nella (4), ovvero gli eventi

$$A \cup B = (A \cap B) \cup (A \cap \bar{B}) \cup (\bar{A} \cap B),$$

$$\begin{aligned}
A \cup \bar{B} &= (A \cap B) \cup (A \cap \bar{B}) \cup (\bar{A} \cap \bar{B}), \\
\bar{A} \cup B &= (A \cap B) \cup (\bar{A} \cap B) \cup (\bar{A} \cap \bar{B}), \\
\bar{A} \cup \bar{B} &= (\bar{A} \cap B) \cup (A \cap \bar{B}) \cup (\bar{A} \cap \bar{B});
\end{aligned}$$

C4) l'unico evento che si ottiene applicando l'operazione di unione all'unica collezione che contiene *tutti* gli eventi nella (4), ovvero l'evento

$$\Omega = (A \cap B) \cup (A \cap \bar{B}) \cup (\bar{A} \cap B) \cup (\bar{A} \cap \bar{B}).$$

Secondo il precedente elenco, la classe di eventi  $\mathcal{C}$  è quindi data da

$$\begin{aligned}
\mathcal{C} = \{ &\emptyset, \\
&A \cap B, A \cap \bar{B}, \bar{A} \cap B, \bar{A} \cap \bar{B}, \\
&A, B, \bar{A}, \bar{B}, (A \cap B) \cup (\bar{A} \cap \bar{B}), (A \cap \bar{B}) \cup (\bar{A} \cap B), \quad (5) \\
&A \cap B, A \cap \bar{B}, \bar{A} \cap B, \bar{A} \cup \bar{B}, \\
&\Omega \}.
\end{aligned}$$

Rimane da dimostrare che  $\sigma(\mathcal{A}^*) = \mathcal{C}$  come affermato in precedenza.

**Dimostrazione** (dell'affermazione  $\sigma(\mathcal{A}^*) = \mathcal{C}$ ). Questa dimostrazione non è richiesta all'esame. La dimostrazione che segue è un po' lunga

perché chiarisce in modo dettagliato ogni singolo passaggio logico. La dimostrazione è suddivisa in due passi: al primo passo si dimostrerà che  $\mathcal{C}$  è una  $\sigma$ -algebra che contiene gli eventi  $A$  e  $B$ ; al secondo passo la dimostrazione verrà completata dimostrando che qualsiasi  $\sigma$ -algebra che contiene gli eventi  $A$  e  $B$  deve necessariamente includere la classe di eventi  $\mathcal{C}$ .

Si consideri il primo passo della dimostrazione. Dall'elenco degli eventi contenuti nella classe  $\mathcal{C}$  si desume immediatamente che gli eventi  $A$  e  $B$  appartengono alla classe  $\mathcal{C}$ . Per completare il primo passo della dimostrazione rimane quindi da dimostrare che la classe  $\mathcal{C}$  soddisfa i requisiti delle  $\sigma$ -algebre. A tal fine si tenga presente che  $\mathcal{C}$  contiene un numero finito di eventi, e per l'osservazione C) sulle algebre e  $\sigma$ -algebre, sarà dunque sufficiente verificare che la classe  $\mathcal{C}$  soddisfi i requisiti a), b) e c) delle algebre. Siccome  $\Omega \in \mathcal{C}$ , il requisito a) delle algebre è soddisfatto. Tenendo presente che il complemento di una unione applicata ad una collezione di intersezioni nella (4) coincide con l'unione applicata alla collezione complementare,<sup>5</sup> si vede che la classe  $\mathcal{C}$  soddisfa anche il requisito b) delle algebre. Infine, non è difficile rendersi conto che la classe  $\mathcal{C}$  soddisfa anche il requisito c) delle algebre visto che un'unione di due unioni di collezioni coincide con l'unione applicata all'unione delle due collezioni.<sup>6</sup>

Avendo dimostrato che la classe  $\mathcal{C}$  è una  $\sigma$ -algebra che contiene gli eventi  $A$  e  $B$ , non resta che dimostrare che qualsiasi  $\sigma$ -algebra che contiene gli eventi  $A$  e  $B$  deve necessariamente includere la classe  $\mathcal{C}$ .

---

<sup>5</sup>Questa affermazione è apparentemente complicata ma esprime un concetto assai semplice. Si consiglia di riflettere con calma sul significato della frase e di aiutarsi con un diagramma di Venn

<sup>6</sup>Anche quest'ultima affermazione è apparentemente complicata ma esprime un concetto assai semplice.

A tal fine sono sufficienti le seguenti osservazioni:

- i) Tutte le  $\sigma$ -algebre (non solo quelle contengono gli eventi  $A$  e  $B$ ) devono contenere l'evento impossibile  $\emptyset$  perché qualsiasi  $\sigma$ -algebra è anche un'algebra (proprietà i) delle  $\sigma$ -algebre) e pertanto deve contenere l'evento impossibile  $\emptyset$  (proprietà i) delle algebre). Ne consegue che qualsiasi  $\sigma$ -algebra che contiene gli eventi  $A$  e  $B$  contiene anche l'evento impossibile  $\emptyset$ .
- ii) Qualsiasi  $\sigma$ -algebra che contiene gli eventi  $A$  e  $B$  deve anche contenere i corrispondenti eventi complementari  $\overline{A}$  e  $\overline{B}$  perché qualsiasi  $\sigma$ -algebra è anche un'algebra (proprietà i) delle  $\sigma$ -algebre) e pertanto deve essere "chiusa" rispetto a operazioni di negazione (requisito b) delle algebre).
- iii) Si considerino ora le quattro intersezioni nella (4). Anche esse devono essere contenute in ogni  $\sigma$ -algebra che contiene gli eventi  $A$  e  $B$  siccome tali  $\sigma$ -algebre devono contenere gli eventi  $\overline{A}$  e  $\overline{B}$  (questo fatto è già stato dimostrato al precedente punto ii)) e siccome, per la proprietà iii) delle algebre, le  $\sigma$ -algebre devono essere "chiusa" rispetto a operazioni di intersezione applicate a collezioni finite di insiemi (si ricordi che ogni  $\sigma$ -algebra è anche un'algebra).
- iv) Si considerino infine le unioni applicate a collezioni di eventi che si possono formare con le intersezioni nella (4). Anche tali unioni devono essere contenute in qualsiasi  $\sigma$ -algebra che contiene gli eventi  $A$  e  $B$  perché, come si è appena visto al punto iii), tali  $\sigma$ -algebre devono contenere le quattro intersezioni nella (4), e perché, per la proprietà ii) delle algebre, le  $\sigma$ -algebre devono essere chiuse rispetto a operazioni di unione applicate collezioni finite di insiemi (si ricordi ancora che ogni  $\sigma$ -algebra è anche un'algebra). Siccome, per definizione, tutti gli eventi appartenenti alla classe  $\mathcal{C}$  devono poter essere rappresentati

come unioni applicate a collezioni di intersezioni nella (4), questo dimostra che qualsiasi  $\sigma$ -algebra che contiene gli eventi  $A$  e  $B$  deve includere la classe  $\mathcal{C}$ . Anche il secondo passo della dimostrazione è quindi completo.

□

□

### 3 Gli assiomi di Kolmogorov

Avendo stabilito le condizioni che lo spazio degli eventi  $\mathcal{A}$  di un modello probabilistico  $(\Omega, \mathcal{A}, P)$  deve soddisfare, non resta che elencare anche le condizioni che devono essere soddisfatte dalla funzione di probabilità  $P : \mathcal{A} \mapsto [0, 1]$ . Come già accennato, queste ultime sono note come "assiomi di Kolmogorov". Esse assicurano che le probabilità associate agli eventi  $A \in \mathcal{A}$  siano tra di loro "coerenti". Tuttavia, come verrà illustrato in molti esempi che seguono, in generale ci sono infiniti modi per assegnare delle probabilità "coerenti" agli eventi  $A$  appartenenti ad una data  $\sigma$ -algebra  $\mathcal{A}$ , ed il problema principale è quello di definire le probabilità  $P(A)$  in modo

”realistico”, ovvero in modo tale che esse possano essere effettivamente considerate come misure per la verosimiglianza degli eventi  $A \in \mathcal{A}$ .

**Definizione 3.1** (Assiomi di Kolmogorov). Secondo l'impostazione di Kolmogorov, le probabilità assegnate agli eventi appartenenti ad un'algebra  $\mathcal{A}$  sono ”coerenti” se e solo se soddisfano i seguenti assiomi:

K1)  $P(A) \geq 0$  per ogni evento  $A \in \mathcal{A}$

K2)  $P(\Omega) = 1$

K3) (Additività numerabile) Se

$$A_1, A_2, \dots$$

è una successione infinita numerabile di eventi *incompatibili* appartenenti all'algebra  $\mathcal{A}$ , e se

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{A},$$



allora

$$P \left( \bigcup_{n=1}^{\infty} A_n \right) = \sum_{n=1}^{\infty} P(A_n)$$

□

Osservazioni sugli assiomi di Kolmogorov:

- i) Siccome ogni  $\sigma$ -algebra è anche un'algebra ma non è vero il viceversa, gli assiomi di Kolmogorov forniscono delle condizioni di "coerenza" anche per funzioni di probabilità  $P : \mathcal{A} \mapsto [0, 1]$  definite su classi di eventi  $\mathcal{A}$  che soddisfano solo i requisiti delle algebre ma non il requisito c\*) delle  $\sigma$ -algebre.
- ii) La precedente osservazione chiarisce perché nell'assioma K3 è necessario imporre l'ipotesi che l'unione  $\bigcup_{n=1}^{\infty} A_n$  appartenga all'algebra  $\mathcal{A}$ . Infatti, come già visto in precedenza, un'algebra  $\mathcal{A}$  non è necessariamente "chiusa" rispetto all'operazione di unione applicata ad una collezione infinita numerabile di eventi.

- iii) Questa osservazione e quella successiva non sono richieste all'esame. Alla luce della precedente osservazione e alla luce del fatto che lo spazio degli eventi  $\mathcal{A}$  di un modello probabilistico è sempre una  $\sigma$ -algebra, ci si potrebbe chiedere perché gli assiomi di Kolmogorov definiscono che cosa si intende per una funzione di probabilità "coerente" anche nel caso in cui  $\mathcal{A}$  è un'algebra ma non una  $\sigma$ -algebra. La risposta (un po' tecnica e un po' lunga) a questa domanda è la seguente: in molte applicazioni si considerano spazi di eventi  $\mathcal{A}$  che sono definiti come la più piccola  $\sigma$ -algebra che contiene una data algebra di eventi  $\mathcal{A}^*$  (si vedano gli esempi dove sono state introdotte le  $\sigma$ -algre di Borel su  $\mathbb{R}$  e su  $\mathbb{R}^k$ ). In questi casi lo spazio degli eventi  $\mathcal{A}$  viene indicato con il simbolo  $\sigma(\mathcal{A}^*)$ . Siccome la  $\sigma$ -algebra  $\sigma(\mathcal{A}^*)$  è spesso notevolmente più ampia dell'algebra  $\mathcal{A}^*$ , sarà sicuramente più semplice attribuire delle probabilità "coerenti" solo agli eventi  $A \in \mathcal{A}^*$  piuttosto che a tutti gli eventi  $A \in \sigma(\mathcal{A}^*)$ . Tuttavia, una volta definita una funzione di probabilità  $P^* : \mathcal{A}^* \mapsto [0, 1]$  che soddisfa gli assiomi di Kolmogorov, si potrà applicare un noto **teorema di estensione** che assicura l'esistenza di un'unica funzione di probabilità  $P : \sigma(\mathcal{A}^*) \mapsto [0, 1]$  che soddisfa gli assiomi di Kolmogorov e che è un'*estensione* della funzione di probabilità  $P^*(\cdot)$  nel senso che  $P(A) = P^*(A)$  per ogni evento  $A \in \mathcal{A}^*$ . Per definire, nel rispetto degli assiomi di Kolmogorov, una funzione di probabilità sulla  $\sigma$ -algebra  $\sigma(\mathcal{A}^*)$  è quindi sufficiente definire, nel rispetto degli assiomi di Kolmogorov, una funzione di probabilità sull'algebra  $\mathcal{A}^*$ .
- iv) Dal punto di vista teorico, potrebbe sorgere il dubbio che il valore della sommatoria nell'assioma K3 possa dipendere dall'*ordine* degli eventi  $A_1, A_2, \dots$ . Si noti tuttavia che questo non è possibile perché l'assioma K1 impone che le probabilità  $P(A_i)$  non

possono essere negative.

## 4 Le leggi del calcolo delle probabilità

Le cosiddette **leggi del calcolo delle probabilità** esprimono relazioni tra le probabilità di determinati eventi e sono spesso utili per ricavare, partendo dalle probabilità di taluni eventi, le probabilità di altri eventi. Tutte le leggi del calcolo delle probabilità sono *conseguenze* degli assiomi di Kolmogorov. In altre parole, esse non sono altro che **teoremi** che possono essere dimostrati a partire dagli assiomi di Kolmogorov. Quindi, anche se alcune leggi del calcolo delle probabilità possono sembrare "ovvie", per accertarsi che esse siano effettivamente "vere" bisogna comunque dimostrarle partendo dagli assiomi di Kolmogorov. Per comprendere l'enunciato di alcune leggi del calcolo delle probabilità e per capire l'intuizione alla base di alcune dimostrazioni, sarà spesso utile visualizzare gli eventi mediante diagrammi di Venn e

interpretare le probabilità associate agli eventi come se fossero delle "masse".

Il seguente elenco contiene alcune leggi del calcolo delle probabilità di fondamentale importanza. Come già accennato, esse devono valere per qualsiasi funzione di probabilità  $P : \mathcal{A} \mapsto [0, 1]$  definita nel rispetto degli assiomi di Kolmogorov.

L1)  $P(\emptyset) = 0$ .

**Dimostrazione.** Questa dimostrazione non è richiesta all'esame. Si consideri una successione infinita (numerabile) di eventi definita ponendo

$$A_1 = \Omega$$

e

$$A_n = \emptyset \quad \text{per } n = 2, 3, \dots$$

Ovviamente, gli eventi di questa successione sono incompatibili e appartengono a qualsiasi algebra di eventi  $\mathcal{A}$  (requisito a) delle algebre e proprietà i) delle algebre). Siccome

$$\bigcup_{n=1}^{\infty} A_n = \Omega,$$

anche l'unione apparterrà a qualsiasi algebra  $\mathcal{A}$ . Dall'assioma

K3 discende quindi che

$$\begin{aligned} P(\Omega) &= P\left(\bigcup_{n=1}^{\infty} A_n\right) \\ &= \sum_{n=1}^{\infty} P(A_n) \\ &= P(\Omega) + \sum_{n=1}^{\infty} P(\emptyset) \end{aligned}$$

da cui si evince che

$$\sum_{n=1}^{\infty} P(\emptyset) = 0.$$

Per l'assioma K1 si può dunque concludere che

$$P(\emptyset) = 0.$$

□

L2) (Additività finita) Se

$$A_1, A_2, \dots, A_n$$

sono  $n$  eventi *incompatibili* appartenenti all'algebra  $\mathcal{A}$  ( $n$  è un numero naturale positivo

finito), allora

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i).$$

**Dimostrazione.** Partendo con gli eventi

$$A_1, A_2, \dots, A_n,$$

si definisca una successione (infinita) di eventi ponendo

$$A_{n+i} = \emptyset \quad \text{per ogni } i = 1, 2, \dots$$

In questo modo,

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^n A_i,$$

e dall'assioma K3 discende quindi che

$$\begin{aligned}
 P\left(\bigcup_{i=1}^n A_i\right) &= P\left(\bigcup_{i=1}^{\infty} A_i\right) \\
 &= \sum_{i=1}^{\infty} P(A_i) \\
 &= \sum_{i=1}^n P(A_i) + \sum_{i=n+1}^{\infty} P(A_i) \\
 &= \sum_{i=1}^n P(A_i) + \sum_{i=n+1}^{\infty} P(\emptyset) \\
 [[\text{legge L1}]] &= \sum_{i=1}^n P(A_i) + \sum_{i=n+1}^{\infty} 0 \\
 &= \sum_{i=1}^n P(A_i)
 \end{aligned}$$

□

L3) Se  $A \in \mathcal{A}$ , allora  $P(\overline{A}) = 1 - P(A)$ .

**Dimostrazione.** Si osservi in primo luogo che  $\overline{A} \in \mathcal{A}$  per il requisito b) delle algebre. Siccome

$$\Omega = A \cup \overline{A}$$

appartiene a  $\mathcal{A}$  per il requisito a) delle algebre, e siccome  $A$  e  $\overline{A}$  sono incompatibili, si

può applicare la legge L2 e concludere che

$$P(\Omega) = P(A) + P(\bar{A}).$$

Ne consegue che

$$P(\bar{A}) = P(\Omega) - P(A),$$

e siccome  $P(\Omega) = 1$  (assioma K2), si ottiene

$$P(\bar{A}) = 1 - P(A)$$

come volevasi dimostrare.  $\square$

L4)  $0 \leq P(A) \leq 1$  per ogni  $A \in \mathcal{A}$ .

**Dimostrazione.** Dalla legge L3 segue che

$$P(A) + P(\bar{A}) = 1,$$

e dato che  $P(A)$  e  $P(\bar{A})$  non possono essere negative (assioma K1), si può concludere che

$$0 \leq P(A) \leq 1$$

come volevasi dimostrare.  $\square$



L5) Se  $A_1, A_2 \in \mathcal{A}$  e se  $A_1 \subseteq A_2$ , allora

$$P(A_1) \leq P(A_2).$$

**Dimostrazione.** Siccome

$$\begin{aligned} A_2 &= (A_1 \cap A_2) \cup (\bar{A}_1 \cap A_2) \\ &= A_1 \cup (\bar{A}_1 \cap A_2) \end{aligned}$$

e siccome gli eventi  $A_1$  e  $\bar{A}_1 \cap A_2$  sono incompatibili e appartengono entrambi a  $\mathcal{A}$ , si può concludere che (legge L2)

$$P(A_1) + P(\bar{A}_1 \cap A_2) = P(A_2).$$

Tenendo presente che le probabilità non possono essere negative (assioma K1), si vede dunque che

$$P(A_1) \leq P(A_2)$$

come volevasi dimostrare.  $\square$

L6) Se  $A_1, A_2 \in \mathcal{A}$  ( $A_1$  e  $A_2$  possono essere sia

compatibili che incompatibili), allora

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2).$$

**Dimostrazione.** Si osservi che

$$A_1 = (A_1 \cap \bar{A}_2) \cup (A_1 \cap A_2)$$

e che gli eventi determinati dalle due intersezioni al secondo membro sono incompatibili e che appartengono entrambi a  $\mathcal{A}$ . Applicando la legge L2 si ottiene quindi

$$P(A_1) = P(A_1 \cap \bar{A}_2) + P(A_1 \cap A_2). \quad (6)$$

Ragionando allo stesso modo con l'evento  $A_2$  si ottiene

$$P(A_2) = P(\bar{A}_1 \cap A_2) + P(A_1 \cap A_2). \quad (7)$$

Usando ora il fatto che

$$A_1 \cup A_2 = (A_1 \cap \bar{A}_2) \cup (A_1 \cap A_2) \cup (\bar{A}_1 \cap A_2),$$

ed il fatto che gli eventi determinati dalle tre intersezioni al secondo membro sono incompatibili, si vede che (ancora legge L2)

$$\begin{aligned} P(A_1 \cup A_2) &= P(A_1 \cap \bar{A}_2) + \\ &\quad + P(A_1 \cap A_2) + \\ &\quad + P(\bar{A}_1 \cap A_2). \end{aligned}$$

Combinando le equazioni (6) e (7) con quest'ultima equazione si ottiene infine

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

come volevasi dimostrare.  $\square$

L7) La probabilità dell'unione di tre eventi  $A_1$ ,  $A_2$  e  $A_3$  (possibilmente compatibili) appartenenti a  $\mathcal{A}$  è data da

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= \\ &= P(A_1) + P(A_2) + P(A_3) + \\ &\quad - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + \\ &\quad + P(A_1 \cap A_2 \cap A_3) \end{aligned}$$

Si omette la dimostrazione.

Più in generale si può dimostrare che la probabilità dell'unione di un numero finito  $n$  di eventi appartenenti all'algebra  $\mathcal{A}$  (di cui alcuni o anche tutti possono essere compatibili) è data da

- (1)  $(-1)^{1-1}$  =più la somma delle probabilità dei singoli eventi
- (2)  $(-1)^{2-1}$  =meno la somma delle probabilità di tutte le "intersezioni a due"
- (3)  $(-1)^{3-1}$  =più la somma delle probabilità di tutte le "intersezioni a tre"
- (4)  $(-1)^{4-1}$  =meno . . . ,
- ( $\vdots$ ) . . .
- ( $n$ ) più/meno (a seconda del segno di  $(-1)^{n-1}$ ) la probabilità dell'intersezione di tutti gli  $n$  eventi.

L8) Disuguaglianza di Boole per due eventi  $A_1$  e  $A_2$  appartenenti a  $\mathcal{A}$ :

$$P(A_1 \cup A_2) \leq P(A_1) + P(A_2).$$

**Dimostrazione.** La dimostrazione segue immediatamente dalla legge L6 tenendo presente che  $P(A_1 \cap A_2) \geq 0$  per l'assioma K1.  $\square$

L9) Disuguaglianza di Boole per  $n$  eventi appartenenti a  $\mathcal{A}$ :

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i).$$

**Dimostrazione.** Questa dimostrazione non è richiesta all'esame. Secondo la legge L8 la disuguaglianza di Boole è vera per  $n = 2$ . Per dimostrare che vale per qualsiasi numero finito  $n$  di eventi appartenenti a  $\mathcal{A}$ , si procede per induzione. Si assuma dunque (ipotesi di induzione) che la disuguaglianza di Boole sia vera per  $n$  eventi e si considerino  $n + 1$  eventi  $A_i$ . Siccome

$$\bigcup_{i=1}^{n+1} A_i = \left(\bigcup_{i=1}^n A_i\right) \cup A_{n+1}$$

la legge L8 implica che

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} A_i\right) &= P\left(\left(\bigcup_{i=1}^n A_i\right) \cup A_{n+1}\right) \\ &\leq P\left(\bigcup_{i=1}^n A_i\right) + P(A_{n+1}). \end{aligned}$$

Usando ora l'ipotesi di induzione secondo la quale

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i),$$

si vede che

$$P\left(\bigcup_{i=1}^{n+1} A_i\right) \leq \sum_{i=1}^{n+1} P(A_i).$$

Questo dimostra il passo induttivo. La disuguaglianza di Boole è dunque valida per qualsiasi numero finito  $n$  di eventi.  $\square$

L10) Disuguaglianza di Bonferroni per  $n$  eventi appartenenti a  $\mathcal{A}$ :

$$P\left(\bigcap_{i=1}^n A_i\right) \geq 1 - \sum_{i=1}^n P(\bar{A}_i).$$

**Dimostrazione.** Applicando la disuguaglianza di Boole agli eventi

$$\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n$$

si ottiene

$$P\left(\bigcup_{i=1}^n \bar{A}_i\right) \leq \sum_{i=1}^n P(\bar{A}_i). \quad (8)$$

Tenendo presente che secondo le leggi di De Morgan

$$\bigcup_{i=1}^n \bar{A}_i = \overline{\left(\bigcap_{i=1}^n A_i\right)},$$

si vede che secondo la legge L3 si deve avere

$$P\left(\bigcup_{i=1}^n \bar{A}_i\right) = 1 - P\left(\bigcap_{i=1}^n A_i\right).$$

Sostituendo questo risultato nella (8) si ottiene

$$1 - P\left(\bigcap_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(\bar{A}_i)$$

e quest'ultima disuguaglianza può essere riscritta come

$$P\left(\bigcap_{i=1}^n A_i\right) \geq 1 - \sum_{i=1}^n P(\bar{A}_i).$$

□

L11) Se

$$A_1 \subseteq A_2 \subseteq \dots$$

è una successione infinita e "non decrescente"  
di eventi appartenenti a  $\mathcal{A}$ , e se anche  $\bigcup_{i=1}^{\infty} A_i$

appartiene a  $\mathcal{A}$ , allora

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} P(A_n).$$

Si omette la dimostrazione.

L12) Se

$$A_1 \supseteq A_2 \supseteq \dots$$

è una successione infinita e "non crescente" di eventi appartenenti a  $\mathcal{A}$ , e se anche  $\bigcap_{i=1}^{\infty} A_i$  appartiene a  $\mathcal{A}$ , allora

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} P(A_n).$$

Si omette la dimostrazione.



## 5 Esempi

**Esempio 5.1.** Secondo un modello probabilistico  $(\Omega, \mathcal{A}, P)$ , si ha

$$P(A) = 0,7, \quad P(B) = 0,5, \quad P(A \cap \bar{B}) = 0,3.$$

Si determini il valore di  $P(A \cup B)$ .

Soluzione:

Secondo la legge L6 si ha

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \quad (9)$$

Per determinare il valore di  $P(A \cup B)$  bisogna quindi determinare il valore di  $P(A \cap B)$ . Tenendo presente che per l'additività finita (legge L2) si ha

$$P(A) = P(A \cap B) + P(A \cap \bar{B}),$$

si vede che

$$\begin{aligned} P(A \cap B) &= P(A) - P(A \cap \bar{B}) \\ &= 0,7 - 0,3 = 0,4, \end{aligned}$$

e sostituendo questo risultato nella (9) si ottiene

$$P(A \cup B) = 0,7 + 0,5 - 0,4 = 0,8.$$

Il precedente ragionamento che ha condotto al valore di  $P(A \cup B)$  è sicuramente rigoroso dal punto di vista formale, ma lascia un po' a desiderare dal punto di vista intuitivo. Infatti, ci si potrebbe chiedere che cosa abbia suggerito di ricorrere all'uso dell'additività finita (legge L2) per ricavare il valore di  $P(A \cap B)$ . Come in molti esercizi sul calcolo delle probabilità, si potrebbe rispondere che l'intuizione deriva da un diagramma di Venn. In alternativa, l'intuizione che ha condotto alla soluzione presentata può essere motivata facendo riferimento ad una **tabella a doppia entrata**. Infatti, per trovare il valore di  $P(A \cap B)$ , e dunque anche il valore di  $P(A \cup B)$  che è richiesto nella consegna dell'esercizio, si potrebbe anche partire dalla seguente tabella a doppia entrata:

evento	$B$	$\bar{B}$	Totale
$A$	$P(A \cap B) = ???$	$P(A \cap \bar{B}) = 0,3$	$P(A) = 0,7$
$\bar{A}$	$P(\bar{A} \cap B) = ???$	$P(\bar{A} \cap \bar{B}) = ???$	$P(\bar{A}) = ???$
Totale	$P(B) = 0,5$	$P(\bar{B}) = ???$	$P(\Omega) = ???$

Siccome per l'assioma K2 si deve avere  $P(\Omega) = 1$ , e siccome per l'additività finita (legge L2) le somme per riga devono coincidere con i valori riportati nella colonna marginale, e le somme per colonna devono coincidere con i valori riportati nella riga marginale, la tabella potrà essere completata solo in un unico modo:

evento	$B$	$\bar{B}$	$P(\cdot)$
$A$	0,4	0,3	0,7
$\bar{A}$	0,1	0,2	0,3
$P(\cdot)$	0,5	0,5	1

Una volta completata la tabella, il valore di  $P(A \cup B)$  potrà essere ricavato applicando la legge L6 (come è stato fatto in precedenza), oppure applicando ancora l'additività finita (legge L2). Per trovare il valore di  $P(A \cup B)$  usando l'additività finita basta osservare che

$$A \cup B = (A \cap B) \cup (A \cap \bar{B}) \cup (\bar{A} \cap B),$$

e che i tre eventi tra parentesi al secondo membro sono incompatibili. Applicando l'additività finita

(legge L2) si vede dunque che

$$\begin{aligned} P(A \cup B) &= P(A \cap B) + P(A \cap \bar{B}) + P(\bar{A} \cap B) \\ &= 0,4 + 0,3 + 0,1 = 0,8, \end{aligned}$$

come già ottenuto in precedenza. □

**Esempio 5.2.** Secondo un servizio meteo, per ciascuno dei prossimi tre giorni presi singolarmente la probabilità di pioggia è pari a 0,3, la probabilità che pioverà in entrambi i prossimi due giorni è pari a 0,2, mentre la probabilità che pioverà in tutti e tre i giorni considerati è pari a 0,1.

- a) Qual è la probabilità che pioverà in almeno uno dei prossimi due giorni?
- b) Qual è la probabilità che non pioverà in nessuno dei prossimi due giorni?
- c) Qual è la probabilità che non pioverà in almeno uno dei prossimi tre giorni?

- d) Qual è la probabilità che pioverà in entrambi i prossimi due giorni e che il giorno successivo non pioverà?
- e) Qual è la probabilità che tra tre giorni pioverà ma che non pioverà in tutti e tre i giorni considerati?

Soluzione:

Nella soluzione si indicheranno con  $A$ ,  $B$  e  $C$  rispettivamente l'evento secondo il quale domani piove, dopodomani piove e infine l'evento secondo il quale tra tre giorni piove. Dal testo dell'esercizio si desume che

$$P(A) = P(B) = P(C) = 0,3,$$

$$P(A \cap B) = 0,2, \quad P(A \cap B \cap C) = 0,1.$$

- a) L'evento secondo il quale pioverà in almeno uno dei prossimi due giorni è dato da

$$A \cup B.$$

Usando la legge L6 si ottiene quindi

$$P(A \cup B) = 0,3 + 0,3 - 0,2 = 0,4.$$

- b) L'evento secondo il quale non poverà in nessuno dei prossimi due giorni è dato da

$$\bar{A} \cap \bar{B}.$$

Secondo le leggi di De Morgan questo evento può essere espresso come

$$\bar{A} \cap \bar{B} = \overline{(A \cup B)}.$$

Usando la legge L3 si ottiene quindi

$$\begin{aligned} P(\bar{A} \cap \bar{B}) &= 1 - P(A \cup B) \\ &= 1 - 0,4 = 0,6. \end{aligned}$$

- c) L'evento secondo il quale non poverà in almeno uno dei prossimi tre giorni è dato da

$$\bar{A} \cup \bar{B} \cup \bar{C}.$$

Secondo le leggi di De Morgan questo evento

può essere espresso come

$$\overline{A} \cup \overline{B} \cup \overline{C} = \overline{(A \cap B \cap C)}.$$

Usando la legge L3 si ottiene quindi

$$\begin{aligned} P(\overline{A} \cup \overline{B} \cup \overline{C}) &= 1 - P(A \cap B \cap C) \\ &= 1 - 0,1 = 0,9. \end{aligned}$$

- d) L'evento secondo il pioverà in entrambi i prossimi due giorni e il giorno successivo non pioverà è dato da

$$A \cap B \cap \overline{C}.$$

Usando un diagramma di Venn non è difficile rendersi conto che

$$A \cap B = (A \cap B \cap C) \cup (A \cap B \cap \overline{C})$$

e che i due eventi tra parentesi al secondo membro sono incompatibili. Per l'additività finita (legge L2) ne consegue che

$$P(A \cap B) = P(A \cap B \cap C) + P(A \cap B \cap \overline{C})$$

da cui si desume che

$$P(A \cap B \cap \overline{C}) = P(A \cap B) - P(A \cap B \cap C).$$

La probabilità richiesta sarà quindi data da

$$P(A \cap B \cap \overline{C}) = 0,2 - 0,1 = 0,1.$$

e) La soluzione viene lasciata per esercizio. □

Nelle soluzioni dei due precedenti esercizi è stato dato per scontato che i valori ipotizzati per le probabilità degli eventi coinvolti fossero tra di loro "coerenti". In altre parole, non è stato verificato se esiste un modello probabilistico  $(\Omega, \mathcal{A}, P)$  con una  $\sigma$ -algebra  $\mathcal{A}$  che contiene tutti gli eventi considerati, e con una funzione di probabilità  $P : \mathcal{A} \mapsto [0, 1]$  che (i) agli eventi considerati assegni le probabilità ipotizzate nel testo del esercizio e che (ii) soddisfi anche gli assiomi di Kolmogorov. Questa omissione potrebbe avere serie conseguenze. Infatti, se le probabilità assegnate nel testo di un esercizio non sono tra di loro "coerenti" nel senso appena descritto, allora, facendo riferimento a tali probabilità e agli



assiomi di Kolmogorov (ed alle leggi del calcolo delle probabilità che ne conseguono), sarà possibile dimostrare qualsiasi affermazione.<sup>7</sup> Quindi non ha senso considerare problemi con probabilità assegnate in modo "incoerente". Il prossimo esempio illustra una situazione dove non è immediatamente ovvio che le probabilità assegnate a determinati eventi sono tra di loro "incoerenti".

**Esempio 5.3.** Secondo l'opinione di un *trader*, la probabilità che al termine di una seduta di scambi risulterà aumentata la quotazione del titolo  $X$  è pari a 0,9, mentre per il titolo  $Y$  la probabilità di un aumento della quotazione ammonta a 0,8. Inoltre, secondo l'opinione del *trader*, la probabilità che aumentino le quotazioni di entrambi i titoli  $X$  e  $Y$  è pari a 0,65.

Si determini la probabilità che la quotazione di almeno uno dei due titoli aumenti.

---

<sup>7</sup>Questa affermazione è una conseguenza del cosiddetto "teorema dello pseudo-Scoto" secondo il quale da un enunciato contraddittorio consegue logicamente qualsiasi altro enunciato.

Soluzione:

Si indichi con  $A$  l'evento secondo il quale la quotazione del titolo  $X$  aumenta e con  $B$  l'evento secondo il quale aumenta la quotazione del titolo  $Y$ . Dai dati forniti nel testo dell'esercizio si desume che

$$P(A) = 0,9, \quad P(B) = 0,8$$

e che

$$P(A \cap B) = 0,65.$$

Applicando la legge L6 si ottiene dunque

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= 0,9 + 0,8 - 0,65 = 1,05 \end{aligned}$$

ma questo valore non è compreso tra 0 e 1 e contraddice dunque la seconda disuguaglianza nella legge L4. Siccome questa contraddizione è stata ottenuta applicando in modo rigoroso gli assiomi di Kolmogorov e le leggi del calcolo delle probabilità, si può concludere che le probabilità assegnate dal *trader* non sono tra di loro "coerenti".  $\square$

Nel precedente esempio è stato abbastanza sem-

plice scoprire che le probabilità assegnate non fossero tra loro "coerenti" ma in altri problemi ciò potrebbe essere alquanto complicato. Spesso, per verificare se le probabilità assegnate a determinati eventi sono tra loro "coerenti" si ricorre a dei cosiddetti **teoremi di estensione** che solitamente non fanno parte del programma di corsi introduttivi alla teoria della probabilità. Pertanto, **in quanto segue si darà sempre per scontato che le probabilità assegnate nei testi degli esercizi siano tra loro "coerenti", a meno che risolvendo un esercizio incidentalmente non si riscontri una contraddizione così come è avvenuto nell'esempio precedente.**

**Esempio 5.4.** Questo esempio può essere saltato. Si supponga che  $P(A) = 0,9$  e che  $P(B) = 0,8$ . Nell'esempio precedente si è visto che in questo caso è impossibile che  $P(A \cap B)$  sia pari a  $0,65$ . Qual è l'insieme dei possibili valori per  $P(A \cap B)$ ?

Soluzione:

Per determinare l'insieme dei possibili valori per la probabilità  $P(A \cap B)$  bisogna innanzitutto determinare l'insieme di tutte le possibili funzioni di probabilità  $P : \mathcal{A} \mapsto [0,1]$  che si possono definire (nel rispetto degli assiomi di Kolmogorov) sulla più piccola  $\sigma$ -algebra che contiene entrambi gli eventi  $A$  e  $B$ , ovvero sulla  $\sigma$ -algebra  $\mathcal{A} = \sigma(\{A, B\})$ . Si osservi che questa  $\sigma$ -algebra coincide con quella già vista nell'Esempio 2.6 (ovvero con la  $\sigma$ -algebra nella (5)).

Siccome si può dimostrare che tra l'insieme di tutte le possibili funzioni di probabilità  $P : \mathcal{A} \mapsto [0, 1]$  e l'insieme di tutte le possibili tabelle a doppia entrata del tipo

evento	$B$	$\bar{B}$	$P(\cdot)$
$A$	$P(A \cap B)$	$P(A \cap \bar{B})$	$P(A)$
$\bar{A}$	$P(\bar{A} \cap B)$	$P(\bar{A} \cap \bar{B})$	$P(\bar{A})$
$P(\cdot)$	$P(B)$	$P(\bar{B})$	$P(\Omega) = 1$

esiste una corrispondenza biunivoca, si può concludere che l'insieme di tutte le possibili funzioni di probabilità  $P : \mathcal{A} \mapsto [0, 1]$  con  $P(A) = 0,9$  e  $P(B) = 0,8$  è dato dall'insieme di tutte le possibili tabelle a doppia entrata che soddisfano questi due vincoli. Per determinare l'insieme di tutti possibili valori per la probabilità  $P(A \cap B)$  bisogna quindi considerare l'insieme di tutte le possibili tabelle a doppia entrata del tipo

evento	$B$	$\bar{B}$	$P(\cdot)$
$A$	$P(A \cap B) = ???$	???	0,9
$\bar{A}$	???	???	0,1
$P(\cdot)$	0,8	0,2	1

Da questa tabella si vede che il valore di  $P(A \cap B)$  determina univocamente anche i rimanenti valori mancanti nella tabella, e che l'insieme dei possibili valori di  $P(A \cap B)$  è dato dall'insieme di tutti i valori mediante i quali si può completare la tabella con valori non negativi in tutte le caselle. Per ottenere i vincoli che devono essere soddisfatti dal valore di  $P(A \cap B)$  conviene dunque completare la tabella con le formule per calcolare gli altri valori mancanti a partire dal valore di  $P(A \cap B)$ . Così facendo si ottiene:

evento	$B$	$\bar{B}$	$P(\cdot)$
$A$	$P(A \cap B)$	$0,9 - P(A \cap B)$	$0,9$
$\bar{A}$	$0,9 - P(A \cap B)$	$0,1 - [0,8 - P(A \cap B)]$ oppure $0,2 - [0,9 - P(A \cap B)]$	$0,1$
$P(\cdot)$	$0,8$	$0,2$	$1$

Da quest'ultima tabella si desume che il valore di  $P(A \cap B)$  potrebbe essere qualsiasi valore che soddisfa le seguenti disuguaglianze:

$$\begin{aligned}
 P(A \cap B) &\geq 0 \\
 0,9 - P(A \cap B) &\geq 0 \\
 0,8 - P(A \cap B) &\geq 0 \\
 0,1 - [0,8 - P(A \cap B)] &\geq 0
 \end{aligned}$$

Non è difficile verificare che tutte e quattro queste disuguaglianze sono soddisfatte se e solo se

$$0,7 \leq P(A \cap B) \leq 0,8.$$

L'insieme dei possibili valori di  $P(A \cap B)$  che sono compatibili con i valori  $P(A) = 0,8$  e  $P(B) = 0,9$  è quindi dato dall'intervallo  $[0,7; 0,8]$ .

□

## 6 Metodi per l'assegnazione delle probabilità

Una volta definito lo spazio campionario  $\Omega$  e la  $\sigma$ -algebra di eventi  $\mathcal{A}$ , si passa alla fase cruciale della definizione di un modello probabilistico: quella dell'assegnazione di valori alle probabilità  $P(A)$  degli eventi  $A \in \mathcal{A}$ . A questo punto sorge spontanea la seguente domanda: esistono dei metodi per assegnare delle probabilità che siano allo stesso tempo "coerenti" e anche "realistiche"?

La risposta a questa domanda dipende ovviamente dalle condizioni fisiche sotto le quali si svolge l'esperimento casuale di riferimento e quindi non può essere data in modo generale. In ogni caso, per definire un modello probabilistico "coerente" e "realistico" è sicuramente utile tenere presente che qualunque metodo può essere interpretato come una implementazione di uno dei seguenti tre metodi generali:

- a) Il **metodo classico**.<sup>8</sup> Questo metodo è applica-

---

<sup>8</sup>Il metodo classico dà luogo a valori delle probabilità che corrispon-

bile solo se  $\Omega$  contiene un numero finito di eventi elementari  $\omega$ , e quando le condizioni fisiche sotto le quali si svolge l'esperimento casuale di riferimento sono tali da ritenere che tutti gli eventi elementari  $\omega \in \Omega$  siano "equiprobabili". In questi casi si può considerare l'insieme delle parti  $\mathcal{P}(\Omega)$  come  $\sigma$ -algebra di eventi, e si dovrà definire la funzione di probabilità  $P : \mathcal{P}(\Omega) \mapsto [0, 1]$  ponendo

$$P(A) = \frac{\#A}{\#\Omega} \quad \text{per ogni } A \in \mathcal{P}(\Omega),$$

---

dono alla cosiddetta *definizione classica della probabilità* che spesso viene attribuita al matematico (nonché fisico e astronomo) francese Pierre Simon Laplace vissuto durante il periodo napoleonico dal 1749 al 1827. Secondo la *definizione classica della probabilità*, "la probabilità di un evento è il rapporto tra il numero dei casi favorevoli all'evento e il numero dei casi possibili, purché questi ultimi siano tutti equiprobabili". Questa definizione del concetto di probabilità ha suscitato non poche perplessità. In primo luogo per l'apparente circolarità della definizione che richiede che i casi possibili siano tutti "equiprobabili"; in secondo luogo perché non chiarisce quale sia la probabilità di un evento nel caso in cui gli eventi possibili non siano "equiprobabili"; e infine perché non chiarisce quale sia la probabilità di un evento nel caso in cui ci sono infiniti casi possibili.

dove  $\#A$  indica il numero di eventi elementari  $\omega$  che appartengono ad  $A$  (ovvero il numero di "casi favorevoli") e dove  $\#\Omega$  indica il numero di eventi elementari che appartengono a  $\Omega$  (ovvero il numero di "casi possibili"). Le probabilità assegnate secondo questo metodo sono quindi date dal rapporto tra il **numero di casi favorevoli (ovvero  $\#A$ )** ed in **numero di casi possibili (ovvero  $\#\Omega$ )**.

Si può dimostrare che il metodo classico fornisce sempre una funzione di probabilità  $P : \mathcal{P}(\Omega) \mapsto [0, 1]$  che rispetta gli assiomi di Kolmogorov.

- b) Il **metodo frequentista**.<sup>9</sup> Questo metodo può essere applicato quando l'esito dell'esperimen-

---

<sup>9</sup>Il metodo frequentista dà luogo a valori delle probabilità che corrispondono alla cosiddetta *definizione frequentista* del concetto di probabilità. Questa definizione fu proposta dal matematico (nonché ingegnere) austriaco (naturalizzato negli Stati Uniti) Richard von Mises vissuto dal 1883 fino al 1953. Secondo la definizione frequentista, "la probabilità di un evento è il limite cui tende la frequenza relativa dell'evento al crescere del numero degli esperimenti". Tuttavia, anche la *definizione frequentista* della probabilità pone dei seri problemi applicativi. Infatti, la definizione frequentista non chiarisce quante volte



to casuale di riferimento è già stato osservato in un elevato numero di repliche dell'esperimento, e se si ritiene che le repliche si siano svolte sotto condizioni simili a quelle sotto le quali si svolgerà l'esperimento casuale di interesse. Anche con il metodo frequentista si può considerare l'insieme delle parti  $\mathcal{P}(\Omega)$  come  $\sigma$ -algebra di eventi, e si definirà la funzione di probabilità  $P : \mathcal{P}(\Omega) \mapsto [0, 1]$  ponendo

$$P(A) = \frac{\text{numero di volte che si è verificato l'evento } A}{\text{numero di repliche osservate}}$$

per ogni  $A \in \mathcal{P}(\Omega)$ .

Non è difficile dimostrare che anche il metodo frequentista fornisce sempre una funzione

---

bisogna ripetere un dato esperimento casuale per poter ritenere che la frequenza relativa di un dato evento sia ragionevolmente prossima al suo limite. Inoltre, la definizione frequentista non può essere applicata per assegnare delle probabilità nei casi in cui un esperimenti casuale non può essere ripetuto oppure può essere ripetuto solo un numero ridotto di volte.

di probabilità  $P : \mathcal{P}(\Omega) \mapsto [0, 1]$  che rispetta gli assiomi di Kolmogorov.

- b) Il **metodo soggettivo**.<sup>10</sup> Secondo questo metodo, un individuo deve assegnare le probabilità usando tutte le informazioni di cui dispone sull'esperimento casuale in questione, ed eventualmente anche su esperimenti casuali simili. In mancanza di informazioni certe l'individuo potrà anche ricorrere al proprio intuito.

Si osservi che due individui che applicano il metodo soggettivo potrebbero benissimo attribuire probabilità diverse ad uno stesso evento. In questo caso i due individui faranno semplicemente riferimento a due modelli probabili-

---

<sup>10</sup>Come origine del metodo soggettivo vengono solitamente citati i lavori di Frank Plumpton Ramsey (matematico, logico, statistico ed economista inglese vissuto dall'anno 1903 fino all'anno 1930) e di Bruno de Finetti (matematico e statistico italiano vissuto dall'anno 1906 fino all'anno 1985). In "Fondamenti logici del ragionamento probabilistico" (contributo pubblicato nel "Bollettino dell'Unione Matematica Italiana", anno IX N. 5, dicembre 1930, edito dalla Casa Editrice Nicola Zanichelli, Bologna) de Finetti scrive che la probabilità deve essere considerata come il "grado di fiducia sentito da un dato individuo nell'avverarsi di un dato evento".

stici  $(\Omega, \mathcal{A}, P)$  diversi. Si osservi inoltre che un individuo potrebbe incorrere nell'errore di non valutare in modo "coerente" le probabilità degli eventi inclusi in una data  $\sigma$ -algebra di eventi  $\mathcal{A}$  e che il metodo soggettivo potrebbe dunque condurre ad una funzione di probabilità  $P : \mathcal{A} \mapsto [0, 1]$  che non rispetta gli assiomi di Kolmogorov. Quando si applica il metodo soggettivo bisogna quindi sempre verificare che le probabilità siano assegnate in modo "coerente".

**Esempio 6.1** (Assegnazione delle probabilità secondo il metodo classico). Si consideri un esperimento casuale che consiste nel lancio di un dado. In questo caso lo spazio campionario potrà essere definito come

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Come si è visto in precedenza, se si ritiene che il dado sia regolare, le probabilità dovrebbero essere assegnate secondo il metodo classico. La probabilità assegnata ad un evento  $A \in \mathcal{P}(\Omega)$  dovrebbe quindi

essere definita come il rapporto tra il numero di casi favorevoli all'evento  $A$ , ed il numero di casi possibili che è dato da  $\#\Omega = 6$ . Per esempio, la probabilità dell'evento  $\{1, 4, 5, 6\}$  dovrebbe essere definita ponendo

$$P(\{1, 4, 5, 6\}) = \frac{4}{6} = \frac{2}{3}.$$

Come già osservato in precedenza, il metodo classico assicura che le condizioni poste dagli assiomi di Kolmogorov siano automaticamente soddisfatte.  $\square$

**Esempio 6.2** (Assegnazione delle probabilità mediante il metodo frequentista e quello soggettivo). Tizio vuole scommettere sull'esito di una partita di calcio tra le squadre  $A$  e  $B$ . Negli ultimi 20 incontri

- (A) la squadra  $A$  ha vinto 12 volte;
- (X) la partita è terminata con un pareggio 6 volte;
- (B) la squadra  $B$  ha vinto solo 2 volte.

Pertanto, secondo il metodo frequentista, la probabilità che vinca la squadra  $A$  dovrebbe essere definita

ponendo  $P(A) = 12/20 = 0,6$ , la probabilità che la partita termini con un pareggio dovrebbe essere definita ponendo  $P(X) = 6/20 = 0,3$ , e la probabilità che invece vinca la squadra  $B$  è dovrebbe essere definita ponendo  $P(B) = 2/20 = 0,1$ . A partire da queste probabilità si possono ricavare le probabilità di tutti gli altri eventi che si possono formare mediante operazioni di unione e intersezione applicate agli eventi  $A$ ,  $X$  e  $B$ .

Si supponga ora che Tizio ritenga che le 20 partite di cui conosce l'esito finale non siano repliche "sotto le medesime condizioni" della partita che ancora si dovrà giocare. Pertanto, secondo Tizio, il metodo frequentista non può essere applicato. Infatti, siccome molti giocatori della squadra  $A$  sono infortunati, Tizio ritiene che la probabilità che la squadra  $B$  vincerà sia pari a 0,30 e non a 0,10 come suggerito dal metodo frequentista, e che la probabilità di un pareggio sia pari a 0,40 piuttosto che a 0,30 (se Tizio è "coerente" dovrebbe quindi assegnare probabilità 0,30 anche alla vittoria della squadra  $A$ ). Si noti che in questo esempio Tizio ha usato le probabilità

suggerite dal metodo frequentista come riferimento iniziale, e a partire da esse ha definito delle nuove probabilità seguendo il proprio intuito. Così facendo Tizio rischia comunque di incorrere nell'errore di assegnare dei valori non "coerenti" alle probabilità degli eventi (assegnando per esempio il valore 0,5 a  $P(A)$ ).  $\square$

## 7 Calcolo combinatorio

Con qualche approssimazione si può affermare che il calcolo combinatorio è la branca della matematica che si occupa di problemi di conteggio. I metodi del calcolo combinatorio sono spesso utili per determinare il numero di elementi contenuti in un dato insieme come per esempio in uno spazio campionario  $\Omega$  oppure in un evento  $A \subset \Omega$ . Per questo motivo, i metodi e le tecniche del calcolo combinatorio trovano largo impiego nel calcolo delle probabilità quando si ritiene che le probabilità debbano essere assegnate in base al metodo classico.

## 7.1 Esperimenti casuali suddivisi in passi

Si consideri un esperimento casuale che si svolge in  $k \geq 1$  passi, con

$n_1$  esiti possibili al primo passo,

$n_2$  esiti possibili al secondo passo, . . .

. . .

$n_k$  esiti possibili all'ultimo ( $k$ -esimo) passo.

Non è difficile rendersi conto che il numero complessivo di eventi elementari  $\omega$  che possono verificarsi in un esperimento casuale di questo tipo sia dato da

$$n_1 \times n_2 \times \cdots \times n_k. \quad (10)$$

Infatti, gli esiti di un tale esperimento possono essere rappresentati mediante sequenze di lunghezza  $k$  i cui elementi descrivono gli esiti dei singoli passi dell'esperimento. Siccome il primo elemento della sequenza può essere scelto in  $n_1$  modi diversi, e può essere combinato con ciascuno degli  $n_2$  modi possibili per scegliere il secondo elemento della sequenza e così via, il numero totale di sequenze possibili dovrà

essere dato dal prodotto nella (10).

Per visualizzare il funzionamento di un esperimento suddiviso in passi è spesso utile costruire un **diagramma ad albero** come illustrato nel seguente esempio.

**Esempio 7.1.** Si consideri un esperimento casuale che consiste nel

- 1) lancio di una moneta,
- 2) seguito dall'estrazione di una pallina da un'urna che contiene una pallina gialla, una verde ed una rossa,
- 3) che a sua volta è seguita dal lancio di un tetraedro con facce numerate da 1 a 4.

Con riferimento all'esperimento casuale in questione:

- a) Si costruisca un diagramma ad albero che descrive gli esiti dei vari passi dell'esperimento e si elenchino gli eventi elementari che descrivono tutti i possibili esiti finali dell'esperimento. Quanti sono?



- b) Si definisca un opportuno modello probabilistico  $(\Omega, \mathcal{A}, P)$ . Secondo quale metodo si dovrebbero assegnare le probabilità?

Soluzione:

- a) Il diagramma ad albero è dato da

$$\begin{array}{c}
 \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \begin{array}{l} \text{giallo} \\ \text{verde} \\ \text{rosso} \end{array} \begin{cases} 1 \\ 2 \\ 3 \\ 4 \end{cases} \\
 \text{testa} \\
 \dots \text{(sotto testa)} \dots \text{croce} \\
 \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \begin{array}{l} \text{giallo} \\ \text{verde} \\ \text{rosso} \end{array} \begin{cases} 1 \\ 2 \\ 3 \\ 4 \end{cases}
 \end{array}$$

Gli eventi elementari che possono verificarsi in questo esperimento casuale possono essere rappresentati attraverso sequenze del tipo

$$\omega = (\text{testa}, \text{giallo}, 1),$$

dove il primo elemento della sequenza indica il lato *testa* oppure il lato *croce* della moneta, il

secondo elemento indica uno dei colori *giallo*, *verde* oppure *rosso* delle palline presenti nell'urna, e il terzo ed ultimo elemento della sequenza è uno dei primi quattro numeri naturali positivi riportati sulle quattro facce del tetraedro. Per tener conto di tutti i possibili esiti di questo esperimento casuale si dovrà quindi fare riferimento ad uno spazio campionario  $\Omega$  che contiene  $2 \times 3 \times 4 = 24$  eventi elementari  $\omega$  (che potrebbero essere rappresentati dalle sequenze appena descritte). A ciascun evento elementare  $\omega \in \Omega$  corrisponde un percorso da sinistra verso destra lungo i nodi dell'albero.

- b) Nell'esempio in questione è ragionevole ritenere che gli eventi elementari dello spazio campionario  $\Omega$  definito al punto a) siano "equiprobabili". Quindi si dovrebbero assegnare le probabilità secondo il metodo classico. Si ricordi che tale metodo permette di assegnare una probabilità

ad ogni  $A \in \mathcal{P}(\Omega)$  ponendo

$$P(A) = \frac{\#A}{\#\Omega}.$$

Si noti che in questo modo si avrà

$$P(\{\omega\}) = \frac{1}{24} \quad \text{per ogni } \omega \in \Omega.$$

Si ricordi inoltre che quando si assegnano le probabilità in base al metodo classico si può essere certi che le condizioni poste dagli assiomi di Kolmogorov siano soddisfatte.

□

## 7.2 Disposizioni e permutazioni

Un caso particolare di esperimento suddiviso in passi consiste nell'estrazione, uno alla volta, di  $k$  oggetti da un'urna che contiene  $n$  oggetti.

- a) Se le estrazioni avvengono **senza reimmissione**, allora ci saranno

$n$  esiti possibili alla prima estrazione,  
 $n - 1$  esiti possibili alla seconda estrazione, . . .  
:  
 $n - k + 1$  esiti possibili alla  $k$ -esima (e ultima) estrazione.

Applicando quanto già visto per esperimenti che si svolgono in più passi, si vede che lo spazio campionario  $\Omega$  associato a questo tipo di esperimento casuale dovrà contenere

$$n(n - 1)(n - 2) \cdots (n - k + 1)$$

eventi elementari. Gli eventi elementari in questione vengono chiamati **disposizioni senza ripetizione** di  $n$  oggetti presi  $k$  alla volta. Si osservi che due disposizioni senza ripetizione si possono distinguersi (i) sia per la presenza di oggetti diversi, (ii) sia per l'ordine in cui si presentano gli oggetti.

Nel caso particolare in cui  $n = k$  si parlerà di **permutazioni senza ripetizione** di  $n$  oggetti. Il

numero di permutazioni senza ripetizione di  $n$  oggetti viene spesso indicato con  $n!$  (si legge  $n$  fattoriale). Ovviamente,

$$n! = n(n - 1)(n - 2) \times \cdots \times 2 \times 1.$$

Visto che qualunque permutazione contiene tutti gli oggetti presenti nell'urna, due permutazioni si possono distinguere solamente per l'ordine in cui si presentano gli stessi oggetti.

- b) Se d'altra parte le estrazioni avvengono **con reimmissione**, allora il numero di esiti possibili per ciascuna delle  $k$  estrazioni sarà sempre pari ad  $n$ , e pertanto in questo caso lo spazio campionario  $\Omega$  conterrà

$$n \times n \times \cdots \times n = n^k$$

eventi elementari che vengono chiamati **disposizioni con ripetizione** di  $n$  elementi presi  $k$  alla volta. Si noti che nel caso di estrazioni con reimmissione, il numero  $k$  di oggetti estratti potrebbe anche eccedere il numero  $n$  di oggetti

presenti nell'urna.

Due disposizioni con ripetizione possono distinguersi (i) per la presenza di oggetti diversi, (ii) per il numero di volte che contengono alcuni oggetti oppure (iii) per l'ordine in cui si presentano gli oggetti.

**Esempio 7.2.** Da un'urna che contiene  $n = 9$  palline numerate si estraggono  $k = 3$  palline senza reimmissione.

- a) Si descriva lo spazio campionario  $\Omega$  associato a questo esperimento casuale. Quanti eventi elementari contiene  $\Omega$ ?
- b) Si assegnino delle probabilità agli eventi di  $\mathcal{P}(\Omega)$ .
- c) Si risponda alle precedenti domande ipotizzando che le estrazioni avvengano con reimmissione.

Soluzione:

- a) Lo spazio campionario  $\Omega$  è costituito dall'insieme di tutte le disposizioni senza ripetizione che si possono formare con  $n = 9$  palline prendendone  $k = 3$  alla volta. Il numero totale di disposizioni di questo tipo ammonta a

$$\#\Omega = 9 \times 8 \times 7 = 504.$$

- b) Siccome lo spazio campionario  $\Omega$  è finito, e siccome è ragionevole ipotizzare che gli eventi elementari  $\omega \in \Omega$  siano “equiprobabili”, le probabilità dovranno essere assegnate in base al metodo classico. Per ogni disposizione  $\omega \in \Omega$  si avrà quindi

$$P(\{\omega\}) = \frac{1}{504} = 0,00198$$

e la probabilità di qualsiasi evento  $A \in \mathcal{P}(\Omega)$  sarà dunque data dal rapporto tra il numero  $\#A$  di casi favorevoli ed il numero  $\#\Omega = 504$  di casi possibili.

- c) Nel caso di estrazioni con reimmissione lo spa-

zio campionario  $\Omega$  conterrà

$$\#\Omega = 9 \times 9 \times 9 = 9^3 = 729$$

disposizioni con ripetizione. Secondo il metodo di assegnazione delle probabilità classico, la probabilità associata a ciascuna di tali disposizioni sarà data da

$$P(\{\omega\}) = \frac{1}{729} = 0,00137,$$

e la probabilità associata a qualsiasi altro evento  $A \in \mathcal{P}(\Omega)$  sarà data dal rapporto tra il numero  $\#A$  di casi favorevoli ed il numero  $\#\Omega = 729$  di casi possibili.

□

### 7.3 Combinazioni

Se un esperimento casuale consiste nell'estrazione **in blocco** di  $k$  oggetti da un'urna che contiene  $n$  oggetti, allora lo spazio campionario  $\Omega$  sarà composto



da eventi elementari  $\omega$  che sono **insiemi non ordinati di  $k$  oggetti diversi** scelti tra gli  $n$  oggetti presenti nell'urna. In questo caso, gli eventi elementari  $\omega$  vengono chiamati **combinazioni** di  $n$  oggetti presi  $k$  alla volta. Il numero di combinazioni diverse viene solitamente indicato con il simbolo

$$\binom{n}{k}.$$

Non è difficile dimostrare che

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

**Dimostrazione.** Ciascuna combinazione contiene esattamente  $k$  oggetti diversi e  $k$  oggetti diversi possono essere ordinati in  $k!$  modi diversi (si vedano le permutazioni). Ordinando tutte le  $\binom{n}{k}$  combinazioni diverse in tutti i  $k!$  modi possibili, si ottengono tutte le disposizioni di  $n$  oggetti presi  $k$  alla volta. Pertanto si ha

$$\binom{n}{k} \times k! = n(n-1) \cdots (n-k+1)$$

Tenendo presente che il secondo membro può essere espresso come

$$\frac{n!}{(n-k)!}$$

si vede che

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

□

**Esempio 7.3.** Da un'urna che contiene  $n = 90$  palline numerate si estraggono in blocco  $k = 6$  palline.

- a) Quanti eventi elementari contiene lo spazio campionario associato a questo esperimento casuale?
- b) Si definisca un modello probabilistico  $(\Omega, \mathcal{A}, P)$  per questo tipo di esperimento casuale. Secondo quale metodo si dovrebbero assegnare le probabilità?

Soluzione:

a) Lo spazio campionario  $\Omega$  contiene

$$\begin{aligned}\binom{90}{6} &= \frac{90!}{6!(90-6)!} \\ &= \frac{90 \times 89 \times 88 \times 87 \times 86 \times 85}{6 \times 5 \times 4 \times 3 \times 2 \times 1} \\ &= 622.614.630\end{aligned}$$

combinazioni di palline prese 6 alla volta.

b) Siccome è ragionevole ritenere che ciascuna combinazione di palline abbia la medesima probabilità di verificarsi, le probabilità dovranno essere assegnate mediante il metodo classico. Secondo tale metodo, la probabilità associata a ciascuna combinazione di palline dovrebbe essere data da

$$\begin{aligned}P(\{\omega\}) &= \frac{1}{622.614.630} \\ &= 1,606 \times 10^{-9},\end{aligned}$$

e le probabilità associate agli altri eventi  $A \in \mathcal{P}(\Omega)$  dovrebbero essere date dal rapporto tra

il numero di casi favorevoli  $\#A$  ed il numero di casi possibili  $\#\Omega = 622.614.630$ .

□

## 7.4 Esercizi

**Esempio 7.4.** Si consideri un mazzo di 32 carte contrassegnate con i valori

7, 8, 9, 10, J, Q, K, A

e con i quattro semi cuori, quadri, fiori e picche.

- A) Dal mazzo vengono estratte, una dopo l'altra, cinque carte senza riposizione.
- $A_1$ ) Qual è la probabilità che le cinque carte estratte siano tutte di cuori?
- $A_2$ ) Qual è la probabilità che le prime due carte estratte siano entrambe di cuori?
- $A_3$ ) Qual è la probabilità che l'ultima carta estratta sia di cuori?

- B) Si risponda alle domande del punto precedente ipotizzando che le estrazioni avvengano con riposizione.
- C) Dal mazzo vengono estratte in blocco cinque carte.
- $C_1$ ) Qual è la probabilità di ottenere quattro carte del valore  $A$ ?
- $C_2$ ) Qual è la probabilità di ottenere tre carte del valore  $A$  e due carte del valore  $K$ ?
- $C_3$ ) Qual è la probabilità di ottenere due carte del valore  $A$ , due carte del valore  $K$  e una carta del valore  $Q$ ?

Soluzione:

- A) Estraendo, una alla volta, cinque carte senza riposizione si ottiene una delle

$$32 \times 31 \times 30 \times 29 \times 28 = 24.165.120$$

disposizioni (senza ripetizione) possibili. E' ragionevole ritenere che ciascuna di esse abbia la

medesima probabilità di verificarsi e quindi si dovrebbero assegnare le probabilità in base al metodo classico: per ogni  $A \in \mathcal{P}(\Omega)$  si dovrebbe quindi definire il corrispondente valore di  $P(A)$  come rapporto tra il numero  $\#A$  di casi favorevoli e il numero  $\#\Omega = 24.165.120$  di casi possibili. Nelle seguenti risposte ai quesiti  $A_1$ ,  $A_2$  e  $A_3$  si determinerà il valore di questo rapporto per gli eventi descritti nei quesiti.

$A_1$ ) Siccome nel mazzo ci sono  $32/4 = 8$  carte di cuori, la probabilità di ottenere cinque carte di cuori (evento  $A_1$ ) sarà data da

$$\begin{aligned} P(A_1) &= \frac{8 \times 7 \times 6 \times 5 \times 4}{32 \times 31 \times 30 \times 29 \times 28} \\ &= \frac{6720}{24.165.120} = 0,000278. \end{aligned}$$

Infatti, la prima carta di cuori può essere scelta in 8 modi diversi e, dopo aver scelto la prima carta di cuori, rimangono 7 modi per scegliere la seconda carta di cuori, ...

$A_2$ ) La probabilità che le prime due carte estratte siano entrambe di cuori è data da

$$\begin{aligned} P(A_2) &= \frac{8 \times 7 \times 30 \times 29 \times 28}{32 \times 31 \times 30 \times 29 \times 28} \\ &= \frac{1.364.160}{24.165.120} = 0,05645 \end{aligned}$$

perché ci sono 8 modi per scegliere la prima carta di cuori tra le 8 carte di cuori nel mazzo, 7 modi per scegliere la seconda carta di cuori tra le rimanenti 7 carte di cuori, 30 modi per scegliere la terza carta (che eventualmente potrebbe anche essere di cuori) tra le 30 carte rimaste nel mazzo, 29 modi per...

$A_3$ ) La probabilità che l'ultima (ovvero la quinta) carta estratta sia di cuori è data da

$$\begin{aligned} P(A_2) &= \frac{8 \times 31 \times 30 \times 29 \times 28}{32 \times 31 \times 30 \times 29 \times 28} \\ &= \frac{6.041.280}{24.165.120} = 0,25 \end{aligned}$$

perché ci sono 8 modi per scegliere la quinta carta estratta tra le 8 carte di cuori a disposizione, 31 modi per scegliere la quarta carta estratta (che eventualmente potrebbe anche essere di cuori) tra le 31 carte rimaste a disposizione, 30 modi per scegliere la terza carta estratta . . .

Si osservi che la probabilità che l'ultima carta estratta sia di cuori è uguale alla probabilità che la prima carta estratta sia di cuori oppure alla probabilità di ottenere una carta di cuori in una qualsiasi posizione prefissata.

- B) Nel caso di estrazioni con riposizione lo spazio campionario conterrà

$$\begin{aligned} 32 \times 32 \times 32 \times 32 \times 32 &= \\ &= 32^5 = 33.554.432 \end{aligned}$$

disposizioni con ripetizione. Nelle risposte ai quesiti  $B_1$ ,  $B_2$  e  $B_3$  si assumerà che tali di-



sposizioni siano tutte “equiprobabili” e quindi si determineranno delle probabilità che corrispondono al metodo di assegnazione delle probabilità classico.

$B_1$ ) Siccome nel mazzo ci sono  $32/4 = 8$  carte di cuori, la probabilità di ottenere cinque carte di cuori (evento  $A_1$ ) sarà data da

$$\begin{aligned} P(A_1) &= \frac{8 \times 8 \times 8 \times 8 \times 8}{32 \times 32 \times 32 \times 32 \times 32} \\ &= \frac{32.768}{33.554.432} = 0,000976. \end{aligned}$$

Infatti, ci sono 8 modi per scegliere la prima carta tra quelle di cuori, 8 modi per scegliere la seconda carta tra quelle di cuori, ...

$B_2$ ) La probabilità che le prime due carte estratte siano entrambe di cuori è data da

$$\begin{aligned} P(A_2) &= \frac{8 \times 8 \times 32 \times 32 \times 32}{32 \times 32 \times 32 \times 32 \times 32} \\ &= \frac{2.097.152}{33.554.432} = 0,0625 \end{aligned}$$

perché ci sono 8 modi per scegliere la prima carta di cuori, 8 modi per scegliere la seconda carta di cuori, 32 modi per scegliere la terza carta (che eventualmente potrebbe anche essere di cuori), 32 modi per...

$B_3$ ) La probabilità che l'ultima carta estratta sia di cuori è data da

$$\begin{aligned} P(A_2) &= \frac{8 \times 32 \times 32 \times 32 \times 32}{32 \times 32 \times 32 \times 32 \times 32} \\ &= \frac{8.388.608}{33.554.432} = 0,25 \end{aligned}$$

perché ci sono 8 modi per scegliere la quinta carta estratta tra le 8 carte di cuori a disposizione, 32 modi per scegliere la quarta carta estratta (che eventualmente potrebbe anche essere di cuori), 32 modi per scegliere la terza carta estratta ...

C) Nel caso di un'estrazione in blocco di cinque

carte, lo spazio campionario conterrà

$$\binom{32}{5} = \frac{32!}{5!(32-5)!} = 201.376$$

combinazioni. Nelle risposte ai quesiti  $C_1$ ,  $C_2$  e  $C_3$  si assumerà che tali combinazioni siano tutte “equiprobabili” e quindi si determineranno delle probabilità che corrispondono al metodo di assegnazione delle probabilità classico.

$C_1$ ) La probabilità di ottenere quattro carte del valore  $A$  è data da

$$\begin{aligned} P(C_1) &= \frac{\binom{4}{4} \times \binom{28}{1}}{\binom{32}{5}} \\ &= \frac{1 \times 28}{201.376} = 0,000139, \end{aligned}$$

perché ci sono  $\binom{4}{4} = 1$  modi per scegliere 4 carte del valore  $A$  tra le 4 carte del valore  $A$  nel mazzo e  $\binom{28}{1} = 28$  modi per scegliere una carta tra le 28 carte restanti.

$C_2$ ) La probabilità di ottenere tre carte del valore  $A$  e due carte del valore  $K$  è data da

$$\begin{aligned} P(C_2) &= \frac{\binom{4}{3} \times \binom{4}{2}}{\binom{32}{5}} \\ &= \frac{4 \times 6}{201.376} = 0,000119. \end{aligned}$$

Infatti, ci sono  $\binom{4}{3} = 4$  modi per scegliere 3 carte del valore  $A$  tra le 4 carte del valore  $A$  disponibili nel mazzo, e ciascuno di questi modi può essere combinato con uno dei  $\binom{4}{2}$  modi possibili per scegliere 2 carte del valore  $K$  tra le 4 carte del valore  $K$  disponibili nel mazzo.

$C_3$ ) La probabilità di ottenere due carte del valore  $A$ , due carte del valore  $K$  ed una

carta del valore  $Q$  è data da

$$\begin{aligned} P(C_3) &= \frac{\binom{4}{2} \times \binom{4}{2} \times \binom{4}{1}}{\binom{32}{5}} \\ &= \frac{6 \times 6 \times 4}{201.376} \\ &= \frac{144}{201.376} = 0,000715, \end{aligned}$$

perché ci sono  $\binom{4}{2}$  modi per scegliere due carte del valore  $A$  tra le quattro nel mazzo,  $\binom{4}{2}$  modi per scegliere due carte del valore  $K$  tra le quattro nel mazzo e infine  $\binom{4}{1}$  modi per scegliere una carta del valore  $Q$  tra le quattro nel mazzo.

□

**Esempio 7.5.** Ad un concorso canoro per bambini la classifica finale viene redatta in modo casuale. Al concorso partecipano 6 bambini.

- a) Qual è la probabilità che il bambino Marco vinca il concorso?

- b) Qual è la probabilità che il bambino Marco finisca in una delle prime tre posizioni?
- c) Qual è la probabilità che Marco e Chiara finiscano entrambi nelle prime tre posizioni?

Soluzione:

Siccome per ipotesi tutte le

$$6! = 720$$

classifiche hanno la stessa probabilità di verificarsi, le probabilità richieste saranno date dal rapporto tra il numero di casi favorevoli ed il numero di casi possibili.

- a) La probabilità che Marco vinca il concorso (evento  $M_1$ ) sarà quindi data da

$$P(M_1) = \frac{1 \times 5 \times 4 \times 3 \times 2 \times 1}{6!} = \frac{1}{6}$$

perché dopo aver posizionato Marco in prima posizione, rimangono 5 modi per scegliere il bambino in seconda posizione;

e ciascun modo di riempire le prime due posizioni può essere combinato con 4 modi diversi (perché?) per riempire la terza posizione;

e ciascun modo di riempire le prime tre posizioni può essere combinato con 3 modi diversi (perché?) per riempire la quarta posizione; e così via. . .

- b) Per determinare la probabilità che Marco finisca in una delle prime tre posizioni (evento  $M_1 \cup M_2 \cup M_3$ ) conviene in primo luogo osservare che gli eventi  $M_1$ ,  $M_2$  e  $M_3$  sono incompatibili, e che quindi (legge dell'additività finita L2)

$$P(M_1 \cup M_2 \cup M_3) = P(M_1) + P(M_2) + P(M_3).$$

Siccome

$$P(M_1) = P(M_2) = P(M_3)$$

(verificare), e siccome

$$P(M_1) = \frac{1}{6}$$

come visto al punto a), si può concludere che

$$P(M_1 \cup M_2 \cup M_3) = 3 \times \frac{1}{6} = \frac{1}{2}.$$

- c) Per determinare la probabilità che Marco e Chiara finiscano entrambi nelle prime tre posizioni, ovvero la probabilità dell'evento

$$A = (M_1 \cup M_2 \cup M_3) \cap (C_1 \cup C_2 \cup C_3),$$

conviene rifarsi direttamente al rapporto tra il numero di casi favorevoli ed il numero di casi possibili. In questo modo si vede che

$$P(A) = \frac{\binom{3}{2} \times 2 \times 4 \times 3 \times 2 \times 1}{6!}$$

perché ci sono  $\binom{3}{2}$  modi per scegliere le due posizioni da occupare con Marco e Chiara tra



le prime tre posizioni;

per ciascun modo di scegliere le due posizioni da occupare con Marco e Chiara ci sono 2 modi diversi per posizionare Marco e Chiara nelle due posizioni scelte (prima Marco e poi Chiara oppure viceversa);

poi rimangono 4 modi per riempire la terza posizione rimasta vuota tra le prime tre posizioni dove già ci sono Marco e Chiara;

a questo punto rimangono 3 modi per occupare la quarta posizione, 2 modi per occupare la quinta posizione e un unico modo per occupare l'ultima posizione.

□

## 8 Probabilità condizionata

### 8.1 Definizione di probabilità condizionata

Come già visto in precedenza, la probabilità  $P(A)$  di un evento  $A$  fornisce una misura per la verosimiglianza dell'evento. Tuttavia, a volte si vuole valutare la verosimiglianza di un dato evento  $A$  **tenendo conto del fatto che si è già verificato un altro evento  $B$** . Il seguente esempio mostra che in questo caso potrebbe essere "sbagliato" valutare la verosimiglianza dell'evento  $A$  usando  $P(A)$ .

**Esempio 8.1.** Un esperimento casuale consiste nel lancio di due dadi regolari di cui uno è di colore rosso mentre l'altro è di colore verde.

- a) Qual è la probabilità che la somma dei punteggi sia pari a 12?
- b) Sapendo che il punteggio del dado rosso è pari a 6, qual è la probabilità che la somma dei punteggi sia 12?

- c) Sapendo che il punteggio del dado rosso è pari a 5, qual è la probabilità che la somma dei punteggi sia 12?

Soluzione:

Nell'esperimento casuale in questione si può considerare come spazio campionario  $\Omega$  l'insieme di tutte le coppie ordinate  $\omega = (x_1, x_2)$  che si possono formare con i primi 6 numeri naturali. Il primo elemento di ciascuna coppia ordinata indica il punteggio ottenuto con il dado rosso, mentre il secondo elemento indica il punteggio ottenuto con il dado verde. Siccome non c'è motivo di ritenere che sia più verosimile osservare alcune coppie ordinate piuttosto che altre, si risponderà ai quesiti supponendo che le probabilità siano definite come previsto dal metodo classico, ovvero come rapporti tra il numero di casi favorevoli ed il numero di casi possibili. Ovviamente, il numero di casi possibili sarà dato da  $\#\Omega = 6 \times 6 = 36$ .

- a) Sia  $A$  l'evento secondo il quale la somma dei punteggi è pari a 12. Siccome  $A$  contiene solo

un evento elementare (ovvero la coppia ordinata  $\omega = (6, 6)$ ), si avrà

$$P(A) = \frac{1}{36}$$

- b) Anche se  $P(A) = \frac{1}{36}$ , intuitivamente la risposta corretta alla domanda b) sembrerebbe essere  $1/6$ .
- c) Anche se  $P(A) = \frac{1}{36}$ , intuitivamente la risposta corretta alla domanda c) sembrerebbe essere 0.

□

Per valutare la verosimiglianza di un evento  $A$ , **tenendo conto del fatto che si è già verificato un altro evento  $B$** , bisogna fare riferimento alla cosiddetta **probabilità condizionata dell'evento  $A$  dato l'evento  $B$** .

**Definizione 8.1** (Probabilità condizionata). Se il valore di  $P(B)$  non è nullo, la probabilità condizio-

nata dell'evento  $A$  dato l'evento  $B$  è definita come

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

□

**Esempio 8.2.** Si risponda ai quesiti b) e c) dell'esempio precedente.

Soluzione:

- b) Si indichi con  $B$  l'evento secondo il quale il punteggio del dado rosso è pari a 6. Siccome  $B$  contiene 6 eventi elementari (le 6 coppie ordinate con il numero 6 in prima posizione), si avrà

$$P(B) = \frac{6}{36} = \frac{1}{6}.$$

Tenendo presente che  $A \subset B$ , si vede che

$$A \cap B = A,$$

e ricordando che

$$P(A) = \frac{1}{36},$$

si ottiene

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6}$$

come ci si aspettava.

- c) Si indichi con  $C$  l'evento secondo il quale il punteggio del dado rosso è pari a 5. Siccome anche  $C$  contiene 6 eventi elementari (le 6 coppie ordinate con il numero 5 in prima posizione), si avrà

$$P(C) = \frac{6}{36} = \frac{1}{6}.$$

Siccome

$$A \cap C = \emptyset,$$

in questo caso si avrà

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{P(\emptyset)}{P(C)} = \frac{0}{\frac{1}{6}} = 0.$$

Anche questo risultato coincide con il valore che sembrava corretto dal punto di vista intuitivo (si veda la risposta al quesito c) nell'Esempio 8.1).

□

## 8.2 Probabilità condizionata e probabilità congiunta

Con il termine **probabilità congiunta di due o più eventi**  $A_1, A_2, \dots, A_k$  si intende la probabilità dell'intersezione degli eventi:

$$P(A_1 \cap A_2 \cap \dots \cap A_k).$$

A volte, per evitare confusione in ragionamenti che coinvolgono delle probabilità congiunte, le probabi-

lità  $P(A_i)$  riferite ai singoli eventi vengono chiamate **probabilità marginali**.

Si supponga ora che  $P(A_1) > 0$ , in modo tale che la probabilità condizionata

$$P(A_2|A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)}$$

sia ben definita. In questo caso si vede immediatamente che la **probabilità congiunta**  $P(A_1 \cap A_2)$  può essere espressa come

$$P(A_1 \cap A_2) = P(A_1) \times P(A_2|A_1). \quad (11)$$

Si supponga ora invece che sia positiva la probabilità congiunta  $P(A_1 \cap A_2)$ . Si osservi che in questo caso dovrà essere positiva anche la probabilità marginale  $P(A_1)$  in quanto la legge L5 impone che

$$P(A_1) \geq P(A_1 \cap A_2) > 0.$$

Come visto in precedenza, questo fatto implica che la probabilità congiunta  $P(A_1 \cap A_2)$  può essere espres-



sa come nella (11). Ma siccome con  $P(A_1 \cap A_2) > 0$ , risulterà ben definita anche la probabilità condizionata

$$P(A_3|A_1 \cap A_2) = \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)}.$$

Ne consegue che anche la probabilità congiunta  $P(A_1 \cap A_2 \cap A_3)$  potrà essere espressa mediante una formula analoga alla (11), ovvero come

$$P(A_1 \cap A_2 \cap A_3) = P(A_1 \cap A_2) \times P(A_3|A_1 \cap A_2).$$

Sostituendo al posto di  $P(A_1 \cap A_2)$  il secondo membro della (11), si vede quindi che

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \times P(A_2|A_1) \times \\ \times P(A_3|A_1 \cap A_2).$$

Iterando questi ragionamenti si può dimostrare il seguente teorema:

**Teorema 8.1.** Se

$$P(A_1 \cap A_2 \cap \cdots \cap A_{k-1}) > 0, \quad (12)$$

allora si avrà

$$\begin{aligned} P(A_1 \cap A_2 \cap \cdots \cap A_k) &= \\ &= P(A_1) \times P(A_2|A_1) \times P(A_3|A_1 \cap A_2) \times \cdots \\ &\quad \cdots \times P(A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1}), \end{aligned} \tag{13}$$

ovvero

$$P\left(\bigcap_{i=1}^k A_i\right) = P(A_1) \times \prod_{i=2}^k P\left(A_i \left| \bigcap_{j=1}^{i-1} A_j\right.\right).$$

□

La formula (13) viene spesso utilizzata per calcolare probabilità congiunte in esperimenti che si svolgono in più passi.

**Esempio 8.3.** Un'urna contiene 10 palline bianche e 5 palline rosse. Dall'urna vengono estratte, una dopo l'altra, 4 palline. Dopo ogni estrazione vengono reinserite nell'urna due palline del colore della pallina estratta.

Si calcoli la probabilità di ottenere

- (i) una pallina bianca alla prima estrazione,
- (ii) seguita da un'altra pallina bianca,
- (iii) seguita da una pallina rossa che a sua volta
- (iv) è seguita da una pallina bianca.

Soluzione:

Ovviamente, l'esperimento casuale in questione può essere visto come un esperimento che si svolge in  $k = 4$  passi. Gli esiti possibili al primo passo sono  $10 + 5 = 15$  (il numero iniziale di palline nell'urna) e siccome è ragionevole ritenere che ciascuno di tali esiti abbia la medesima probabilità di verificarsi, la probabilità di ottenere una pallina bianca alla prima estrazione (evento  $B_1$ ) dovrà essere data da

$$P(B_1) = \frac{10}{10 + 5} = \frac{10}{15}.$$

A questo punto, prima di procedere alla seconda estrazione (secondo passo), nell'urna vengono inserite due palline bianche e la probabilità condizionata

di ottenere una pallina bianca alla seconda estrazione (evento  $B_2$ ), *sapendo che alla prima estrazione si è ottenuta una pallina bianca*, sarà quindi data da

$$P(B_2|B_1) = \frac{10 - 1 + 2}{(10 - 1 + 2) + 5} = \frac{11}{16}.$$

Visto l'esito della seconda estrazione, prima di procedere alla terza estrazione verranno inserite nell'urna ancora una volta due palline bianche. La probabilità condizionata di ottenere una pallina rossa alla terza estrazione (evento  $R_3$ ), *sapendo che nelle prime due estrazioni sono state ottenute due palline bianche*, sarà quindi data da

$$P(R_3|B_1 \cap B_2) = \frac{5}{(10 - 1 + 2 - 1 + 2) + 5} = \frac{5}{17}.$$

Infine, visto l'esito della terza estrazione, prima di procedere alla quarta e ultima estrazione verranno inserite nell'urna due palline rosse. La probabilità condizionata di ottenere alla quarta estrazione una pallina bianca, *sapendo che nelle prime due estrazioni sono state ottenute due palline bianche e*

*che nella terza estrazione è stata ottenuta una pallina rossa, sarà dunque data da*

$$\begin{aligned} P(B_4|B_1 \cap B_2 \cap R_3) &= \\ &= \frac{10 - 1 + 2 - 1 + 2}{(10 - 1 + 2 - 1 + 2) + (5 - 1 + 2)} = \frac{12}{18}. \end{aligned}$$

Usando queste probabilità condizionate si vede che la probabilità di ottenere

- i) una pallina bianca alla prima estrazione,
- ii) che è seguita da un'altra pallina bianca alla seconda estrazione,
- iii) che è seguita da una pallina rossa alla terza estrazione,
- iv) che a sua volta è seguita da una pallina bianca alla quarta ed ultima estrazione

sarà data da (si usa la formula (13))

$$\begin{aligned} P(B_1 \cap B_2 \cap R_3 \cap B_4) &= \\ &= P(B_1) \times P(B_2|B_1) \times P(R_3|B_1 \cap B_2) \times \end{aligned}$$

$$\begin{aligned} & \times P(B_4|B_1 \cap B_2 \cap R_3) = \\ & = \frac{10}{15} \times \frac{11}{16} \times \frac{5}{17} \times \frac{12}{18} = 0,0899. \end{aligned}$$

Si noti che in questo tipo di esperimento casuale non si può applicare il metodo classico per assegnare *direttamente* dei valori alle probabilità degli eventi che descrivono l'esito finale delle quattro estrazioni.

□

### 8.3 Eventi indipendenti

**Definizione 8.2** (Eventi indipendenti). Due eventi  $A$  e  $B$  si dicono **indipendenti** se la loro probabilità congiunta è uguale al prodotto delle probabilità marginali:

$$P(A \cap B) = P(A) \times P(B)$$

□

Si osservi che se  $A$  e  $B$  sono due eventi indipen-

denti, allora si ha

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A) \quad \text{se } P(B) > 0$$

e

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = P(B) \quad \text{se } P(A) > 0.$$

Non è difficile dimostrare che se  $A$  e  $B$  sono una coppia di eventi indipendenti, allora lo sono anche la coppia di eventi  $A$  e  $\bar{B}$ , la coppia di eventi  $\bar{A}$  e  $B$  e la coppia di eventi  $\bar{A}$  e  $\bar{B}$ .

**Dimostrazione.** Si consideri la seguente tabella a doppia entrata:

	$B$	$\bar{B}$	Totale
$A$	$P(A \cap B)$	$P(A \cap \bar{B})$	$P(A)$
$\bar{A}$	$P(\bar{A} \cap B)$	$P(\bar{A} \cap \bar{B})$	$P(\bar{A})$
Totale	$P(B)$	$P(\bar{B})$	1

dove per la legge L2 le probabilità marginali  $P(A)$ ,  $P(\bar{A})$ ,  $P(B)$  e  $P(\bar{B})$  devono essere uguali alle somme per riga o colonna delle probabilità congiunte, e

dove per la legge L3

$$P(\overline{A}) = 1 - P(A) \quad \text{e} \quad P(\overline{B}) = 1 - P(B).$$

Se  $A$  e  $B$  sono indipendenti, ovvero se

$$P(A \cap B) = P(A) \times P(B),$$

allora, per i vincoli sulle probabilità marginali, devono valere le seguenti uguaglianze:

$$\begin{aligned} P(A \cap \overline{B}) &= P(A) - P(A \cap B) \\ &= P(A) - P(A) \times P(B) \\ &= P(A) \times (1 - P(B)) \\ &= P(A) \times P(\overline{B}) \end{aligned}$$

(questo dimostra che  $A$  e  $\overline{B}$  sono eventi indipendenti),

$$\begin{aligned} P(\overline{A} \cap B) &= P(B) - P(A \cap B) \\ &= P(B) - P(A) \times P(B) \\ &= (1 - P(A)) \times P(B) \\ &= P(\overline{A}) \times P(B) \end{aligned} \tag{14}$$



(questo dimostra che anche  $\bar{A}$  e  $B$  sono eventi indipendenti) e

$$\begin{aligned} P(\bar{A} \cap \bar{B}) &= P(\bar{A}) - P(\bar{A} \cap B) \\ \text{[[per la (14)]]} &= P(\bar{A}) - P(\bar{A}) \times P(B) \\ &= P(\bar{A}) \times (1 - P(B)) \\ &= P(\bar{A}) \times P(\bar{B}) \end{aligned}$$

(questo dimostra che anche  $\bar{A}$  e  $\bar{B}$  sono eventi indipendenti).  $\square$

**Esempio 8.4.** Un risparmiatore ha investito in due titoli azionari. La probabilità che ad un anno dall'investimento il primo titolo presenti un rendimento positivo (evento  $A$ ) è pari a 0,60. Per il secondo titolo, invece, la probabilità di ottenere ad un anno dall'investimento un rendimento positivo (evento  $B$ ) è solo pari a 0,55.

- a) Supponendo che  $P(A|B) = 0,90$ , qual è la probabilità che ad un anno dall'investimento almeno uno dei due titoli presenti un rendimento positivo?

- b) Supponendo ora invece che gli eventi  $A$  e  $B$  siano indipendenti, qual è la probabilità che ad un anno dall'investimento almeno uno dei due titoli presenti un rendimento positivo?

Soluzione:

Per la legge L6, la probabilità che ad un anno dall'investimento almeno uno dei due titolo presenti un rendimento positivo deve essere data da

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

- a) Se  $P(A|B) = 0,90$ , allora si avrà

$$\begin{aligned} P(A \cap B) &= P(A|B) \times P(B) \\ &= 0,90 \times 0,55 \\ &= 0,495, \end{aligned}$$

e quindi si otterrà

$$P(A \cup B) = 0,60 + 0,55 - 0,495 = 0,655.$$

- b) Se d'altra parte  $A$  e  $B$  sono eventi indipenden-

ti, allora si avrà

$$\begin{aligned}P(A \cap B) &= P(A) \times P(B) \\ &= 0,60 \times 0,55 \\ &= 0,33.\end{aligned}$$

In questo caso si ottiene

$$P(A \cup B) = 0.60 + 0.55 - 0,33 = 0,82.$$

□

## 8.4 Indipendenza globale di tre o più eventi

Tre eventi  $A_1, A_2, A_3$  si dicono **globalmente indipendenti** se

$$P(A_1 \cap A_2) = P(A_1) \times P(A_2)$$

$$P(A_1 \cap A_3) = P(A_1) \times P(A_3)$$

$$P(A_2 \cap A_3) = P(A_2) \times P(A_3)$$

e se

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \times P(A_2) \times P(A_3).$$

In generale,  $k \geq 2$  eventi

$$A_1, A_2, \dots, A_k$$

si dicono globalmente indipendenti se sono soddisfatte tutte le seguenti condizioni:

1) Per qualsiasi combinazione di due eventi si ha

$$P(A_{i_1} \cap A_{i_2}) = P(A_{i_1}) \times P(A_{i_2})$$

(si osservi che con  $k$  eventi si possono formare  $\binom{k}{2}$  combinazioni di eventi presi due alla volta);

2) per qualsiasi combinazione di tre eventi si ha

$$P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) = P(A_{i_1}) \times P(A_{i_2}) \times P(A_{i_3})$$

(si osservi che con  $k$  eventi si possono formare  $\binom{k}{3}$  combinazioni di eventi presi tre alla volta);

3) ...

k-1) per l'unica combinazione che comprende tutti i  $k$  eventi si ha

$$P(A_1 \cap A_2 \cap \dots \cap A_k) = P(A_1) \times P(A_2) \times \dots \times P(A_k)$$

(si osservi che con  $k$  eventi si può formare un'unica combinazione di eventi presi  $k$  alla volta: infatti,  $\binom{k}{k} = 1$  per ogni  $k = 1, 2, \dots$ ).

**Esempio 8.5.** Si consideri un gioco che consiste in tre lanci di una moneta. Sia  $A_1$  l'evento secondo il quale gli esiti dei primi due lanci sono uguali,  $A_2$  l'evento secondo il quale gli esiti del primo e del terzo lancio sono uguali, e sia  $A_3$  l'evento secondo il quale gli esiti degli ultimi due lanci sono uguali.

Si verifichi che gli eventi  $A_1$ ,  $A_2$  e  $A_3$  sono indipendenti a coppie ma che non sono globalmente indipendenti.

Soluzione:

Gli eventi elementari che possono verificarsi al termine dell'esperimento casuale in questione possono essere rappresentati con delle terne ordinate i cui elementi sono delle  $T$  (per testa) oppure delle  $C$  (per croce). Siccome ogni elemento di una terna può essere scelto in 2 modi diversi, lo spazio campionario  $\Omega$  dovrà contenere

$$\#\Omega = 2 \times 2 \times 2 = 2^3 = 8$$

eventi elementari.

Nel contesto in questione, gli eventi elementari  $\omega \in \Omega$  possono essere considerati equiprobabili e quindi si dovranno assegnare le probabilità secondo il metodo classico. Siccome il numero di casi favorevoli all'evento  $A_1$  è dato da

$$\#A_1 = 2 \times 1 \times 2 = 4$$

(infatti, il primo elemento di una terna compatibile con l'evento  $A_1$  può essere scelto in 2 modi diversi; il secondo elemento della terna potrà poi essere scelto in un unico modo in quanto dovrà coincidere con il

primo elemento della terna; il terzo e ultimo elemento della terna potrà di nuovo essere scelto in due modi diversi), la probabilità dell'evento  $A_1$  sarà data da

$$P(A_1) = \frac{\#A_1}{\#\Omega} = \frac{4}{8} = \frac{1}{2}.$$

Ragionando in modo analogo si vede che

$$\#A_2 = 2 \times 2 \times 1 = 4$$

e che

$$\#A_3 = 2 \times 2 \times 1 = 4.$$

Ne consegue che

$$P(A_1) = P(A_2) = P(A_3) = \frac{4}{8} = \frac{1}{2}. \quad (15)$$

Si consideri ora l'intersezione  $A_1 \cap A_2$ . Si osservi l'evento  $A_1 \cap A_2$  si verifica se e solo se gli esiti di tutti e tre i lanci sono uguali. Ne consegue che il numero di casi favorevoli all'evento  $A_1 \cap A_2$  sarà dato da

$$\#(A_1 \cap A_2) = 2 \times 1 \times 1 = 2$$

e quindi si avrà

$$P(A_1 \cap A_2) = \frac{\#(A_1 \cap A_2)}{\#\Omega} = \frac{2}{8} = \frac{1}{4}.$$

Siccome questo risultato coincide con il valore di

$$P(A_1) \times P(A_2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

si può concludere che  $A_1$  e  $A_2$  sono eventi indipendenti.

Si osservi ora che anche gli eventi  $A_1 \cap A_3$  e  $A_2 \cap A_3$  si verificano se e solo se gli esiti di tutti e tre i lanci sono uguali. Quindi si avrà

$$A_1 \cap A_2 = A_1 \cap A_3 = A_2 \cap A_3$$

e dunque anche

$$P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{4}.$$

Usando la (15) si vede dunque che sono indipendenti anche gli eventi  $A_1$  e  $A_3$  e gli eventi  $A_2$  e  $A_3$ . Questo dimostra che gli eventi  $A_1$ ,  $A_2$  e  $A_3$  sono indipendenti



a coppie.

Per verificare che gli eventi  $A_1$ ,  $A_2$  e  $A_3$  non sono globalmente indipendenti bisogna ora verificare che

$$P(A_1 \cap A_2 \cap A_3) \neq P(A_1) \times P(A_2) \times P(A_3).$$

A tal proposito si osservi che, per la (15), il prodotto delle probabilità marginali al secondo membro è dato da

$$P(A_1) \times P(A_2) \times P(A_3) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}. \quad (16)$$

Rimane da determinare il valore di  $P(A_1 \cap A_2 \cap A_3)$ . A tal fine si noti che anche l'evento  $A_1 \cap A_2 \cap A_3$  si verifica se e solo se gli esiti di tutti e tre i lanci sono uguali. Quindi si può concludere che

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_1 \cap A_2) = \\ &= P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{4}, \end{aligned}$$

e siccome questo risultato è diverso dal valore del prodotto delle probabilità marginali nella (16), questo dimostra che gli eventi  $A_1$ ,  $A_2$  e  $A_3$  non sono

globalmente indipendenti. □

**Esempio 8.6.** Si consideri un'urna che contiene 100 palline delle quali 60 sono contrassegnate con un punto nero, 80 hanno la superficie ruvida e 90 contengono una sorpresa. Le palline contrassegnate con un punto nero che hanno la superficie ruvida sono 48, le palline contrassegnate con un punto nero che contengono una sorpresa sono 54 e le palline con superficie ruvida che contengono una sorpresa sono 72. Le palline che presentano tutte e tre le caratteristiche sono invece solo 30.

Dall'urna viene estratta una pallina in modo casuale. Si considerino i seguenti eventi:

- l'evento  $A_1$  secondo il quale la pallina estratta è contrassegnata con un punto nero;
- l'evento  $A_2$  secondo il quale la pallina estratta ha la superficie ruvida;
- l'evento  $A_3$  secondo il quale la pallina estratta contiene una sorpresa.

Si verifichi se gli eventi  $A_1$ ,  $A_2$  e  $A_3$  sono globalmente indipendenti.

Soluzione:

Secondo le informazioni fornite,

$$P(A_1) = 0,60, \quad P(A_2) = 0,80, \quad P(A_3) = 0,90,$$

$$P(A_1 \cap A_2) = 0,48, \quad P(A_1 \cap A_3) = 0,54,$$

$$P(A_2 \cap A_3) = 0,72$$

e

$$P(A_1 \cap A_2 \cap A_3) = 0,30.$$

Dato che

$$\begin{aligned} P(A_1) \times P(A_2) &= 0,60 \times 0,80 \\ &= 0,48 \\ &= P(A_1 \cap A_2), \end{aligned}$$

$$\begin{aligned} P(A_1) \times P(A_3) &= 0,60 \times 0,90 \\ &= 0,54 \\ &= P(A_1 \cap A_3) \end{aligned}$$

e

$$\begin{aligned}P(A_2) \times P(A_3) &= 0,80 \times 0,90 \\ &= 0,72 \\ &= P(A_2 \cap A_3),\end{aligned}$$

si può concludere che gli eventi  $A_1$ ,  $A_2$  e  $A_3$  sono **indipendenti a coppie**.

Tuttavia,

$$\begin{aligned}P(A_1) \times P(A_2) \times P(A_3) &= 0,60 \times 0,80 \times 0,90 \\ &= 0,432 \\ &\neq 0,30 \\ &= P(A_1 \cap A_2 \cap A_3)\end{aligned}$$

e questo dimostra che gli eventi  $A_1$ ,  $A_2$  e  $A_3$  non sono globalmente indipendenti.  $\square$

Come verrà illustrato nel seguente esempio, eventi globalmente indipendenti sorgono solitamente nel contesto di esperimenti casuali che si svolgono in più passi quando gli esiti dei singoli passi non si influenzano a vicenda.

**Esempio 8.7.** Un esperimento casuale consiste nel lancio di una moneta, che è seguito dal lancio di un tetraedro con 4 facce numerate che a sua volta è seguito dal lancio di un dado con 6 facce numerate.

Si considerino

- l'evento  $A_1$  secondo il quale l'esito del lancio della moneta è testa;
- l'evento  $A_2$  secondo il quale l'esito del lancio del tetraedro è la faccia che riporta il numero 4;
- l'evento  $A_3$  secondo il quale l'esito del lancio del dado è la faccia che riporta il numero 6.

Si verifichi che gli eventi  $A_1$ ,  $A_2$  e  $A_3$  sono globalmente indipendenti.

Soluzione:

Gli eventi elementari che possono verificarsi al termine dell'esperimento casuale in questione possono essere rappresentati con delle terne ordinate il cui primo elemento può essere il lato "testa" oppure il lato "croce" della moneta, il cui secondo elemento può

essere uno dei primi quattro numeri naturali positivi che compaiono sulle quattro facce del tetraedro, e il cui terzo ed ultimo elemento può essere uno dei primi sei numeri naturali positivi che compaiono sulle sei facce del dado. Lo spazio campionario associato all'esperimento casuale conterrà quindi

$$\#\Omega = 2 \times 4 \times 6 = 48$$

eventi elementari.

Siccome gli eventi elementari possono essere considerati come “equiprobabili”, si dovrà procedere all'assegnazione delle probabilità mediante il metodo classico. Quindi si otterrà una funzione di probabilità  $P : \mathcal{P}(\Omega) \mapsto [0, 1]$  che associerà delle probabilità a tutti i sottoinsiemi di  $\Omega$  in base alla regola del rapporto tra il numero di casi favorevoli ed il numero di casi possibili:

$$P(A) = \frac{\#A}{\#\Omega} \quad \text{per } A \in \mathcal{P}(\Omega).$$

Le probabilità degli eventi  $A_1$ ,  $A_2$  e  $A_3$  saranno

pertanto date da (perché?)

$$P(A_1) = \frac{1 \times 4 \times 6}{48} = \frac{1}{2},$$

$$P(A_2) = \frac{2 \times 1 \times 6}{48} = \frac{1}{4},$$

e da

$$P(A_3) = \frac{2 \times 4 \times 1}{48} = \frac{1}{6}.$$

Le probabilità delle intersezioni riferite a tutte le coppie di eventi che si possono formare con gli eventi  $A_1$ ,  $A_2$  e  $A_3$  saranno invece date da (perché?)

$$P(A_1 \cap A_2) = \frac{1 \times 1 \times 6}{48} = \frac{1}{8},$$

$$P(A_1 \cap A_3) = \frac{1 \times 4 \times 1}{48} = \frac{1}{12},$$

e da

$$P(A_2 \cap A_3) = \frac{2 \times 1 \times 1}{48} = \frac{1}{24}.$$

Siccome

$$P(A_i \cap A_j) = P(A_i) \times P(A_j) \quad \text{per ogni } i \neq j$$

(la verifica è immediata), si può concludere che tutte le coppie di eventi sono indipendenti.

Per verificare se gli eventi  $A_1$ ,  $A_2$  e  $A_3$  sono globalmente indipendenti, bisogna quindi solo più verificare se

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \times P(A_2) \times P(A_3).$$

Non è difficile rendersi conto che anche quest'ultima condizione è soddisfatta. Infatti,

$$P(A_1 \cap A_2 \cap A_3) = \frac{1 \times 1 \times 1}{48} = \frac{1}{48}$$

e

$$P(A_1) \times P(A_2) \times P(A_3) = \frac{1}{2} \times \frac{1}{4} \times \frac{1}{6} = \frac{1}{48}.$$

Questo dimostra che gli eventi  $A_1$ ,  $A_2$  e  $A_3$  sono globalmente indipendenti. Procedendo in modo analogo si può verificare che nell'ambito del modello probabilistico considerato sono anche globalmente indipendenti tutte le terne di eventi  $A_1$ ,  $A_2$  e  $A_3$  dove  $A_1$  descrive solo l'esito del lancio della moneta,



$A_2$  descrive solo l'esito del lancio del tetraedro e  $A_3$  descrive solo l'esito del lancio del dado.

□

## 8.5 La formula della probabilità totale

Si considerino  $k$  eventi  $A_1, A_2, \dots, A_k$  che formano una **partizione dello spazio campionario  $\Omega$** , ovvero  $k$  eventi  $A_1, A_2, \dots, A_k$  tali che

$$A_i \cap A_j = \emptyset \quad \text{per ogni } i \neq j,$$

e tali che

$$A_1 \cup A_2 \cup \dots \cup A_k = \Omega.$$

Si consideri inoltre un altro evento  $B$ . Siccome

$$B = (B \cap A_1) \cup (B \cap A_2) \cup \dots \cup (B \cap A_k)$$

e siccome gli eventi

$$B \cap A_i \quad i = 1, 2, \dots, k$$

sono incompatibili, l'additività finita (legge L2) implica che

$$P(B) = \sum_{i=1}^k P(B \cap A_i). \quad (17)$$

Supponendo ora che

$$P(A_i) > 0 \quad \text{per ogni } i = 1, 2, \dots, k,$$

e che quindi siano ben definite le  $k$  probabilità condizionate  $P(B|A_i)$ , si vede che

$$P(B \cap A_i) = P(B|A_i) \times P(A_i).$$

Sostituendo quest'ultima espressione nella (17) si ottiene la cosiddetta **formula della probabilità totale**:

$$P(B) = \sum_{i=1}^k P(B|A_i) \times P(A_i). \quad (18)$$

Si osservi che se per qualche  $i = 1, 2, \dots, k$  la probabilità marginale  $P(A_i)$  è nulla, e per questo motivo la corrispondente probabilità condizionata  $P(B|A_i)$  non è definita, allora sarà nullo anche il corrispon-

dente termine  $P(B \cap A_i)$  nella (17), e quindi il termine  $P(B|A_i) \times P(A_i)$  nella formula della probabilità totale potrà essere omissis.

Prima di passare ad un esempio, vale la pena osservare che la formula della probabilità totale vale anche nel caso in cui  $k = \infty$ , ovvero nel caso di una partizione dello spazio campionario formata da una collezione di eventi infinita numerabile. In questo caso, il ragionamento che conduce alla formula della probabilità totale rimane identico a quello visto in precedenza. L'unica differenza consiste nel fatto che per ottenere la (17) bisogna fare appello all'assioma K3 piuttosto che all'additività finita (legge L2).

**Esempio 8.8.** Un gioco consiste nel lancio di un dado e, a seconda del punteggio ottenuto nel lancio del dado, nell'estrazione di una pallina da una delle tre urne qui di seguito descritte:

- 1) se nel lancio del dado si ottiene il punteggio 1 (evento  $A_1$ ), la pallina viene estratta da un'urna che contiene 2 palline bianche e 3 palline nere;

- 2) se nel lancio del dado si ottiene il punteggio 2 oppure il punteggio 3 (evento  $A_2$ ), la pallina viene estratta da un'urna che contiene 2 palline bianche e 4 palline nere;
- 3) in tutti gli altri casi (evento  $A_3$ ) la pallina viene estratta da un'urna che contiene 2 palline bianche e 6 palline nere.

Qual è la probabilità di ottenere una pallina bianca (evento  $B$ )?

Soluzione:

Sia  $A_1$  l'evento secondo il quale il punteggio del dado è pari a 1, sia  $A_2$  l'evento secondo il quale il punteggio del dado è pari a 2 oppure a 3, e sia  $A_3$  l'evento secondo il quale il punteggio del dado è maggiore di 3. Si indichi inoltre con  $B$  l'evento secondo il quale la pallina ottenuta nell'estrazione dall'urna è bianca.

Dalla descrizione del gioco si desume che

$$P(A_1) = \frac{1}{6}, \quad P(A_2) = \frac{2}{6} = \frac{1}{3},$$

$$P(A_3) = \frac{3}{6} = \frac{1}{2}$$

e che

$$P(B|A_1) = \frac{2}{2+3} = \frac{2}{5}, \quad P(B|A_2) = \frac{2}{2+4} = \frac{1}{3},$$

$$P(B|A_3) = \frac{2}{2+6} = \frac{1}{4}.$$

Sostituendo questi valori nella formula della probabilità totale si ottiene

$$\begin{aligned} P(B) &= \frac{2}{5} \times \frac{1}{6} + \frac{1}{3} \times \frac{1}{3} + \frac{1}{4} \times \frac{1}{2} \\ &= \frac{1}{15} + \frac{1}{9} + \frac{1}{8} = \frac{327}{1080} = 0,303. \end{aligned}$$

□

## 8.6 La formula di Bayes

Come visto nella sezione precedente, la formula della probabilità totale permette di ricavare il valore di  $P(B)$  a partire dalle probabilità marginali degli eventi  $A_1, A_2, \dots, A_k$  di una partizione del-

lo spazio campionario e dai valori delle probabilità condizionate  $P(B|A_1)$ ,  $P(B|A_2)$ ,  $\dots$ ,  $P(B|A_k)$ .

In questa sezione verrà dimostrata la cosiddetta formula di Bayes che, a partire dalle suddette quantità, permette di ricavare i valori delle probabilità condizionate

$$P(A_1|B), \quad P(A_2|B), \quad \dots, \quad P(A_k|B).$$

Siccome queste probabilità condizionate sono definite se e solo se  $P(B) > 0$ , d'ora in poi si assumerà che questa disuguaglianza sia soddisfatta. In tal caso, per la definizione di probabilità condizionata, si avrà

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)},$$

e tenendo presente che

$$P(A_i \cap B) = \begin{cases} P(B|A_i) \times P(A_i) & \text{se } P(A_i) > 0 \\ 0 & \text{altrimenti,} \end{cases}$$

si vede che

$$P(A_i|B) = \begin{cases} \frac{P(B|A_i) \times P(A_i)}{P(B)} & \text{se } P(A_i) > 0 \\ 0 & \text{altrimenti.} \end{cases}$$

Sostituendo al posto di  $P(B)$  la formula della probabilità totale (formula (18)); si conviene che i termini per i quali  $P(B|A_i)$  non è definita perché  $P(A_i)$  è nulla sono tutti nulli), si ottiene la cosiddetta **formula di Bayes**, ovvero

$$P(A_i|B) = \begin{cases} \frac{P(B|A_i) \times P(A_i)}{\sum_{j=1}^k P(B|A_j) \times P(A_j)} & \text{se } P(A_i) > 0 \\ 0 & \text{altrimenti.} \end{cases}$$

Ovviamente, anche la formula di Bayes rimane valida se  $k = \infty$ , ovvero se al posto di una partizione finita si considera una partizione infinita numerabile.

**Esempio 8.9.** Si consideri ancora l'esempio precedente. Si calcoli per ciascuno dei tre eventi  $A_1$ ,  $A_2$  e  $A_3$  la corrispondente probabilità condizionata  $P(A_i|B)$ . Si interpretino i valori ottenuti.

Soluzione:

Sostituendo nella formula di Bayes si ottiene:

$$P(A_1|B) = \frac{\frac{1}{15}}{\frac{1}{15} + \frac{1}{9} + \frac{1}{8}} = 0,220;$$

$$P(A_2|B) = \frac{\frac{1}{9}}{\frac{1}{15} + \frac{1}{9} + \frac{1}{8}} = 0,367;$$

$$P(A_3|B) = \frac{\frac{1}{8}}{\frac{1}{15} + \frac{1}{9} + \frac{1}{8}} = 0,413.$$

Dal confronto di queste tre probabilità condizionate si deduce che quando si ottiene una pallina bianca nell'estrazione dall'urna e non si conosce l'esito del lancio del dado, l'ipotesi più verosimile è un punteggio maggiore di 3. Infatti, la probabilità condizionata  $P(A_i|B)$  raggiunge il suo valore massimo quando  $i = 3$ . □



## Bibliografia

- Billingsley P. (1995), *Probability and Measure - Third Edition*, Wiley Series in Probability and Mathematical Statistics, Wiley 1995
- Zenga M. (1996), *Modello probabilistico e variabili casuali*, Giappichelli, Torino 1996