Mauro Piccioni

PROBABILITÀ DI BASE

www.aracneeditrice.it
info@aracneeditrice.it

via Raffaele Garofalo, 133/A–B
00173 Roma
(06) 93781065

ISBN 978–88–548–3064–6

I edizione: febbraio 2010

Indice

1 Gli elementi fondamentali del modello **1**

1.1 Spazi di probabilità ed eventi 1
1.1.1 Esperimenti aleatori e spazi dei campioni 1
1.1.2 Insiemi e operazioni sugli insiemi 3
1.1.3 Eventi, attributi e funzioni indicatrici 8
1.1.4 Esercizi 12
1.2 Le funzioni di massa e di probabilità 13
1.2.1 Spazi finiti 13
1.2.2 Proprietà elementari delle probabilità 16
1.2.3 Somme infinite e spazi numerabili 19
1.2.4 Definizione di probabilità su spazi al più numerabili . . . 22
1.2.5 Esercizi 24
1.3 Approfondimenti 25
1.3.1 Dimostrazione delle proposizioni 1.8 e 1.9 25
1.3.2 Sulla continuità della probabilità 26

2 Trasformazioni, variabili aleatorie e leggi **31**
2.1 Leggi indotte da una trasformazione 31
2.1.1 Trasformazioni aleatorie 31
2.1.2 Alcune variabili aleatorie e le loro leggi 38
2.1.3 Esercizi 41
2.2 Funzioni di ripartizione 42
2.2.1 Proprietà delle funzioni di ripartizione 42
2.2.2 Funzioni di ripartizione di funzioni di v.a. 45
2.2.3 Test di significatività e funzioni di sopravvivenza 49
2.2.4 Esercizi 53
2.3 Approfondimenti 54
2.3.1 Prova delle Proposizione 2.4 54
2.3.2 Funzioni di ripartizione a gradino e non 56

3 La struttura della leggi congiunte **59**
3.1 Indipendenza e prodotto di probabilità 59
3.1.1 Leggi congiunte e marginali 59
3.1.2 Indipendenza tra due eventi e prodotto di due spazi di probabilità 61
3.1.3 Esercizi 69
3.2 Probabilità condizionata 70
3.2.1 Definizione di probabilità condizionata e prime proprietà . 70
3.2.2 Correlazione tra eventi 72
3.2.3 La probabilità condizionata come misura di probabilità . 77
3.2.4 Costruzione di una legge congiunta da una marginale e una famiglia di condizionate 78
3.2.5 Esercizi 87
3.3 Approfondimenti 90
3.3.1 Tabelle a doppia entrata di massima correlazione 90

4 Indipendenza e calcolo combinatorio **91**
4.1 Mutua indipendenza e probabilità prodotto 91
4.1.1 Prodotto di spazi di probabilità e trasformazioni indipendenti 91
4.1.2 Mutua indipendenza di una famiglia finita di eventi . . . 95
4.1.3 Esercizi 102
4.2 Oggetti combinatorici fondamentali 103
4.2.1 Disposizioni e loro conteggio 103
4.2.2 Combinazioni e numeri di occupazione 106
4.2.3 Coefficienti multinomiali 109
4.2.4 Le statistiche di Maxwell-Boltzmann, Fermi-Dirac e Bose-Einstein 113
4.2.5 Esercizi 114
4.3 Approfondimenti 117
4.3.1 Il numero delle combinazioni con ripetizione 117

5 Schemi d'urna e legge dei grandi numeri **119**
5.1 Schemi di estrazione da un'urna 119
5.1.1 Schema di Bernoulli e legge binomiale 119
5.1.2 Legge ipergeometrica ed estrazioni senza reimmissione . . 122
5.1.3 Formula delle probabilità composte e applicazioni: leggi di Polya 127
5.1.4 Esercizi 135
5.2 Schemi di estrazione da urne con più colori 137
5.2.1 Estrazioni con reimmissione: legge multinomiale 137
5.2.2 Estrazioni senza reimmissione 140
5.2.3 Esercizi 143
5.3 Il significato della legge dei grandi numeri 144
5.4 Approfondimenti 146
5.4.1 Simmetria della legge ipergeometrica 146

5.4.2 Urne con più colori: estrazioni con rinforzo e legge di Bose-Einstein . . . 147

6 Indipendenza condizionata e alberi infiniti 149
6.1 Indipendenza condizionata . . . 149
6.1.1 Indipendenza condizionata tra eventi . . . 149
6.1.2 Indipendenza tra due trasformazioni condizionale a una terza . . . 157
6.1.3 Lo schema di Bernoulli condizionato . . . 162
6.1.4 Esercizi . . . 164
6.2 Alberi di profondità infinita e tempi di attesa . . . 166
6.2.1 Il tempo di primo successo in alberi finiti . . . 166
6.2.2 Il tempo di primo successo in alberi infiniti . . . 168
6.2.3 Legge congiunta di tempi d'attesa . . . 173
6.2.4 Esercizi . . . 177
6.3 Approfondimenti . . . 179
6.3.1 Dimostrazione della Proposizione 6.10 . . . 179

7 La dipendenza markoviana 181
7.1 Dipendenza markoviana in generale . . . 181
7.1.1 Proprietà markoviana . . . 181
7.2 Catene di Markov . . . 186
7.2.1 Definizione ed esempi . . . 186
7.2.2 Leggi stazionarie e di equilibrio . . . 190
7.2.3 Catene irriducibili catene ergodiche e teorema di Markov 197
7.2.4 Problemi di assorbimento . . . 201
7.2.5 Esercizi . . . 210
7.3 Approfondimenti . . . 213
7.3.1 Dimostrazione della Proposizione 7.8 . . . 213
7.3.2 Dimostrazione della Proposizione 7.11 . . . 214
7.3.3 Dimostrazione della Proposizione 7.12 (2). . . . 215

8 Valori attesi e applicazioni 217
8.1 Valore attesi su spazi finiti o numerabili . . . 217
8.1.1 Valore atteso come estensione lineare di una probabilità . 217
8.1.2 Il principio di inclusione-esclusione . . . 222
8.1.3 Variabili non negative con infiniti valori . . . 224
8.1.4 Somme infinite di funzioni di segno qualunque e definizione generale di valore atteso . . . 229
8.1.5 Esercizi . . . 234
8.2 I momenti, la varianza e la covarianza . . . 235
8.2.1 I momenti di una legge . . . 235
8.2.2 La varianza come indice di dispersione . . . 239
8.2.3 Lo spazio delle variabili aleatorie di quadrato integrabile . 244
8.2.4 Esercizi . . . 253
8.3 Approfondimenti . . . 255

8.3.1 Dimostrazione della Proposizione 8.9 255
8.3.2 Legge dei grandi numeri per variabili aleatorie non integrabili 256
8.3.3 Approssimazione di variabili aleatorie e regressione 257

9 Funzioni generatrici e limite centrale 261
9.1 Valori attesi condizionati 261
9.1.1 Esercizi 269
9.2 La funzione generatrice delle probabilità 271
9.2.1 Definizione e prime proprietà 271
9.2.2 Convoluzione e funzioni generatrici delle probabilità . . . 272
9.2.3 Momenti e funzioni generatrici delle probabilità 276
9.2.4 Applicazioni della funzione generatrice delle probabilità . 278
9.2.5 Esercizi 285
9.3 Legge gaussiana e teorema del limite centrale 286
9.3.1 Convergenza in legge 286
9.3.2 Dalla legge dei grandi numeri al teorema del limite centrale 287
9.3.3 Applicazioni statistiche del teorema del limite centrale . . 292
9.3.4 Esercizi 295
9.4 Approfondimenti 297
9.4.1 Valore atteso condizionato come predittore ottimo 297
9.4.2 Prova della Proposizione 9.13 300
9.4.3 Prova della Proposizione 9.15 301

10 Spazi di probabilità continui 303
10.1 Schema di Bernoulli infinito e legge di Lebesgue 303
10.1.1 Dallo schema di Bernoulli alla legge di Lebesgue 310
10.1.2 Funzioni di ripartizione di v.a. assolutamente continue . . 314
10.2 Approfondimenti 318

Capitolo 1

Gli elementi fondamentali del modello

In questo capitolo presentiamo gli ingredienti fondamentali dei *modelli probabilistici* e accenniamo alle molteplici situazioni in cui risultano appropriati. Da un punto di vista matematico un modello probabilistico non è altro che uno *spazio di probabilità.* Nella accezione ristretta che considereremo durante la maggior parte del corso, uno spazio di probabilità consta di un insieme ambiente, lo *spazio dei campioni,* e di una funzione definita sui sottoinsiemi di questo spazio, che chiamiamo *eventi.* Prima di specificare le naturali proprietà che richiediamo a tale funzione, che chiameremo *probabilità*, richiamiamo alcuni elementi della *teoria elementare degli insiemi* che costituiranno il nostro linguaggio di base. La necessità di trattare probabilità su spazi dei campioni numerabili renderà necessaria una digressione sulle somme infinite e il rapporto con la nozione più comune di serie. Dopo aver derivato le più immediate conseguenze della definizione di probabilità, tra le quali si segnala per la sua utilità la *formula delle probabilità totali*, concluderemo il capitolo approfondendo il ruolo dell'assioma di *additività numerabile della probabilità.*

1.1 Spazi di probabilità ed eventi

1.1.1 Esperimenti aleatori e spazi dei campioni

Uno spazio di probabilità rappresenta un modello matematico di un *esperimento aleatorio*, quindi di una prova il cui risultato è incerto, cioè non prevedibile esattamente. Questa incertezza può derivare dal fatto che l'esperimento non ha ancora avuto luogo, oppure dal fatto che esso si è già svolto ma manca un'informazione sul suo esito. Le più semplici situazioni di questo tipo sono legate ai giochi quali:

- il lancio di uno o più dadi;

- le estrazioni dei numeri della tombola;
- la distribuzione delle carte in una mano di bridge;
- l'uscita dei segni di una schedina del Totocalcio.

Per chiarire i concetti che via via introdurremo faremo spesso ricorso a questi giochi. Tuttavia sappiamo bene che anche la nostra vita quotidiana ci offre continuamente situazioni incerte, quando:

- stiamo per recarci in un ufficio postale. Quante persone troveremo in fila davanti a noi? Riusciremo ad essere serviti prima dell'ora in cui abbiamo fissato un appuntamento?
- aspettiamo i mezzi pubblici ad una fermata. Quali autobus passeranno e soprattutto quando passeranno?
- acquistiamo dei titoli finanziari. Le azioni che ci vengono proposte frutteranno più dei titoli di stato? O quanto meno dell'interesse del conto corrente?

Ma i settori di attività in cui si è interessati all'esito di esperimenti aleatori sono molteplici. Ecco una breve lista.

- Nella sperimentazione dei farmaci un gruppo di pazienti viene trattato con il farmaco di interesse, mentre altri non vengono trattati e costituiscono il gruppo di controllo. Anche se il farmaco è inefficace, può accadere che il gruppo trattato mostri risultati migliori rispetto al gruppo di controllo. Valutare con che probabilità questo può verificarsi serve ad assicurare che il risultato dell'esperimento sia *significativo* dell'efficacia del farmaco.
- Per dimensionare una rete di comunicazione (sia essa una connessione tra computer oppure una rete stradale) occorre, sulla base di caratteristiche ipotizzate per il traffico, rendere la probabilità che la rete si blocchi sufficientemente piccola.
- Un'azienda di assicurazione riceve dei premi fissi dai suoi assicurati, a fronte dei rimborsi a loro dovuti per l'eventuale verificarsi degli eventi contro i quali essi sono assicurati. L'ammontare dei premi richiesti deve rendere la probabilità di insolvenza dell'azienda sufficientemente piccola.

Il primo elemento di uno spazio di probabilità è l'elenco dei possibili **esiti** dell'esperimento, che vanno elencati in un insieme Ω che chiameremo **spazio dei campioni.** Gli esiti di questo elenco devono essere *esaustivi* e *incompatibili*, nel senso che l'esperimento deve necessariamente dar luogo ad uno ed uno solo tra di essi.

Una prima classificazione della complessità degli spazi di probabilità, fondamentale per il seguito, va fatta in base alla loro **cardinalità.** Per un **insieme**

finito A la sua cardinalità $|A|$ è il numero degli elementi dell'insieme. Per determinare la cardinalità di **insiemi infiniti** si utilizzano degli insiemi di riferimento. In particolare, diciamo che un insieme ha cardinalità numerabile quando può essere messo in corrispondenza biunivoca con l'insieme $\mathbf{N}$ dei naturali, e che ha la cardinalità del continuo quando può essere messo in corrispondenza biunivoca con l'insieme $\mathbf{R}$ dei numeri reali.

- Uno *spazio dei campioni* costituito da un numero *finito* di esiti è appropriato per descrivere esempi quali l'estrazione di una o più palline da un'urna. Ad un esperimento di questo tipo si riconduce l'estrazione di una o più unità *campione da una popolazione finita*, che è alla base delle procedure di *sondaggio*. Ad un esperimento di questo tipo possono ovviamente essere assimilati il lancio di un dado, che corrisponde all'estrazione da un'urna con 6 palline, la scelta di una carta da una mazzo di carte italiane, che corrisponde all'estrazione da un'urna con 40 palline, e così via.

- Uno *spazio dei campioni* costituito da un'infinità *numerabile* di elementi è appropriato a descrivere, ad esempio, il numero di persone in fila in un certo istante davanti ad uno sportello bancario. Se non altro per motivi di spazio questo numero non potrà essere enorme. Ma ci troveremmo in difficoltà dovendo specificare una soglia massima che non può essere oltrepassata, e inoltre vorremmo che la scelta di questa soglia non influenzi i risultati ottenuti. Esperimenti di *conteggio* di questo tipo possono avere per oggetto il numero di componenti di una famiglia scelta a caso sull'elenco telefonico, il numero di gol in una certa partita di calcio, il numero di auto che transitano attraverso un casello autostradale in un fissato lasso di tempo, e così via.

- Uno *spazio dei campioni* costituito da un'infinità *continua* di elementi è appropriato per descrivere, ad esempio, il *tempo di attesa* di un autobus ad una fermata oppure la durata di una chiamata ad un centralino telefonico. E' chiaro che un orologio digitale può rilevare questi tempi solo con un'approssimazione dipendente dal numero delle cifre mostrate dallo schermo. Anche in questo caso non vogliamo che i risultati possano dipendere dalla precisione specificata per lo strumento di misura: e troviamo quindi conveniente ipotizzare di poter rilevare esattamente questi tempi. Esperimenti in cui si rilevano misure di lunghezza, peso, temperatura, hanno ovviamente caratteristiche simili.

1.1.2 Insiemi e operazioni sugli insiemi

Una volta elencati nello spazio dei campioni tutti i possibili esiti di un esperimento, il nostro interesse per il risultato dell'esperimento stesso è sempre legato al verificarsi o meno di certi specifici *eventi*. Ecco alcuni esempi di eventi legati agli esperimenti aleatori precedentemente descritti.

- Nell'estrazione della tombola siamo interessati che esca un numero appartenente ad una nostra cartella.

- Entrando in una stazione di servizio siamo interessati al fatto che i clienti non siano in numero superiore al numero delle pompe di benzina.
- Se aspettiamo l'autobus ci interessa che il tempo di attesa non superi un certo tempo massimo da noi tollerabile.

Dato che una volta che l'esito dell'esperimento è noto dobbiamo essere in grado di determinare se un evento cui siamo interessati si è verificato o no (altrimenti lo spazio dei campioni non è adeguato per i nostri scopi), un evento deve individuare univocamente un sottoinsieme dello spazio dei campioni. Per questo motivo nel seguito *chiameremo **evento** un qualsiasi sottoinsieme dello spazio dei campioni.* Se $A \subset \Omega$ è un evento e $\omega \in \Omega$ uno degli esiti possibili dell'esperimento, quando $\omega \in A$ diciamo che ω *realizza* A.

Per prima cosa è quindi richiamare le operazioni che possiamo effettuare sugli eventi.

Definizione 1.1 *La **complementazione** di un evento $A \subset \Omega$ è definita da*

$$A^c = \{\omega : \omega \notin A\}\,.$$

*L'**unione** di due eventi A e $B \subset \Omega$ è definita da*

$$A \cup B = \{\omega : \omega \in A \text{ oppure } \omega \in B\}\,,$$

*mentre la loro **intersezione** è definita da*

$$A \cap B = \{\omega : \omega \in A \text{ e } \omega \in B\}\,.$$

Ricordiamo che la relazione di inclusione $A \subset B$ tra eventi A e B di Ω vale quando ogni elemento di A è un elemento di B. Quindi per dimostrare un'uguaglianza $A = B$ tra eventi di Ω bisogna prima mostrare che $A \subset B$ e successivamente che $B \subset A$. E' facile verificare che l'inclusione può essere espressa equivalentemente con l'ausilio delle operazioni di unione o di intersezione.

Proposizione 1.1 *Qualunque siano gli eventi A, $B \subset \Omega$*

$$\begin{array}{rcl} A & \subset & B \iff A \cup B = B \\ & \Updownarrow & \\ A \cap B & = & A \end{array}$$

Prova. Cominciamo a dimostrare le implicazioni $\Longrightarrow$ e $\Downarrow$. Se $A \subset B$ è immediato che l'unione di A e B non aggiunge alcun elemento a B e l'intersezione di A e B non ne toglie alcuno ad A. Per provare le implicazioni opposte si può ragionare *per assurdo*, e negare la tesi $A \subset B$. Supponiamo quindi che esista un elemento ω^* di A che non appartiene a B. Allora $\omega^* \in A \cup B$ e $\omega^* \notin A \cap B$. Otteniamo quindi la negazione delle ipotesi $A \cup B = B$ e $A \cap B = A$ rispettivamente, concludendo quindi la prova. ∎

Per semplificare le notazioni è utile introdurre anche la **differenza tra gli eventi** A e $B \subset \Omega$, definita come

$$A \setminus B = A \cap B^c$$

e la **differenza simmetrica**

$$A \triangle B = (A \setminus B) \cup (B \setminus A) = (A \cap B^c) \cup (B \cap A^c)\,.$$

Prima di enunciare alcune delle proprietà delle operazioni di unione, intersezione e complementazione ricordiamo che l'insieme vuoto $\varnothing$ è l'evento di Ω che non è realizzato da alcun elemento.

Proposizione 1.2 *Le operazioni di unione, intersezione e complementazione hanno le proprietà*

1. **Commutativa:**

$$A \cup B = B \cup A, \quad A \cap B = B \cap A.$$

2. **Distributiva una rispetto all'altra:**

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)\,,\; A \cap (B \cup C) = (A \cap B) \cup (A \cap C)\,.$$

3. $\varnothing$ **è l'identità di** $\cup$, Ω **è l'identità di** $\cap$:

$$A \cup \varnothing = A, \;\; A \cap \Omega = A.$$

4. **Partizione binaria:**

$$A \cup A^c = \Omega, \;\; A \cap A^c = \varnothing,$$

che valgono qualunque siano i sottoinsiemi A, B e C di Ω.

Prova. Le uguaglianze (1), (3) e (4) sono immediate dalla definizione. Per la prima delle (2) proviamo innanzi tutto che il membro di sinistra è contenuto in quello di destra. Dato che $B \cap C \subset B$ è chiaro che $A \cup (B \cap C) \subset A \cup B$. Analogamente $A \cup (B \cap C) \subset A \cup C$, per cui essendo $A \cup (B \cap C)$ contenuto in $A \cup B$ e $A \cup C$, è contenuto nella loro intersezione. Per provare che il membro di destra è contenuto in quello di sinistra osserviamo per prima cosa che A è contenuto in entrambi. D'altra parte, se $\omega \notin A$ appartiene al membro di destra, deve necessariamente essere $\omega \in B \cap C$, il che conclude la prova. La seconda delle (2) si prova con argomenti analoghi. Una prova alternativa particolarmente rapida è possibile utilizzando le formule di de Morgan, che tratteremo tra poco.
■

L'argomento usato nella prova della proprietà (2) può essere visualizzato facendo ricorso ad un **diagramma di Venn,** ottenuto associando a ciascun evento A, B, C un sottoinsieme del piano come nella Figura 1a). Per provare

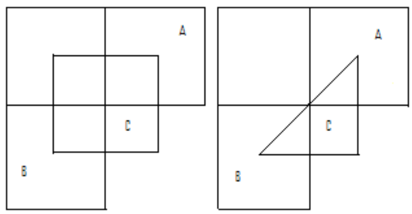

Figura 1.1: Sottoinsiemi a) in posizione generica e b) non in posizione generica.

direttamente una uguaglianza tra insiemi mediante un diagramma di Venn occorre che i sottoinsiemi associati agli eventi A, B, C,... (in assenza di vincoli tra di essi), siano in *posizione generica*. Questo significa che *tutti gli atomi della partizione da essi generata* (che definiremo in generale solo più avanti) *devono essere non vuoti*. Nel caso dei tre eventi A, B e C questo significa che $A \cap B \cap C$, $A^c \cap B \cap C$, $A \cap B^c \cap C$, $A \cap B \cap C^c$, $A^c \cap B^c \cap C$, $A^c \cap B \cap C^c$, $A \cap B^c \cap C^c$ e $A^c \cap B^c \cap C^c$ devono essere tutti non vuoti come nella Figura 1.1 a). Nella Figura 1.1 b) questo non accade perchè $A \cap B \cap C = \varnothing$. Utilizzando un diagramma di Venn di quest'ultimo tipo, ad esempio, si può ritenere che

$$(A \cap B) \, \Delta \, (A \cap C) = (A \cap B) \cup (A \cap C) \, ,$$

valga qualunque siano A, B e C, mentre è vera soltanto quando $A \cap B \cap C = \varnothing$.

Si osserva ora che, qualunque sia l'evento $A \subset \Omega$, si ha

$$(A^c)^c = A, \tag{1.1}$$

che ci torna utile per dimostrare la seguente proposizione.

Proposizione 1.3 ***Formule di de Morgan*** . *Qualunque siano gli eventi A e B di Ω si ha che*

$$(A \cap B)^c = A^c \cup B^c \tag{1.2}$$

$$(A \cup B)^c = A^c \cap B^c. \tag{1.3}$$

Prova. Per provare che il membro di sinistra della (1.2) è contenuto nel membro di destra si percorrono le implicazioni che seguono da sinistra verso destra, mentre per provare la relazione opposta si percorrono in senso opposto.

$$\omega \in (A \cap B)^c \iff \omega \notin A \cap B \Longleftrightarrow \omega \in A^c \text{ o } \omega \in B^c \iff \omega \in A^c \cup B^c.$$

La seconda formula di de Morgan (1.3) si ottiene sostituendo in (1.2) A e B con A^c e B^c; in questo modo si ottiene

$$(A^c \cap B^c)^c = A \cap B.$$

Complementando entrambi i membri e tenendo presente (1.1) si ottiene immediatamente (1.3). ■

Si può dimostrare, anche se non è un risultato banale, che le proprietà (1)-(4) caratterizzano le operazioni di unione, intersezione e complementazione, nel senso che tutte le altre proprietà possono essere ricavate a partire da queste. Uno dei vantaggi di questo approccio consiste nel fatto che le proprietà (1)-(4) rimangono immutate se si scambiano $\cup$ con $\cap$ e Ω con $\varnothing$ per cui vale il seguente

Corollario 1.1 ***Principio di dualità della teoria degli insiemi.*** *Ad ogni relazione di uguaglianza valida qualunque siano gli eventi di Ω che vi compaiono, corrisponde un'altra uguaglianza ugualmente valida, che si ottiene scambiando $\cup$ con $\cap$ e Ω con $\varnothing$.*

Una proprietà immediata delle operazioni di unione e intersezione è la **proprietà associativa.**

Proposizione 1.4 *Qualunque siano gli eventi A, B e C di Ω si ha*

$$A \cup (B \cup C) = (A \cup B) \cup C = \{\omega : \omega \in A \text{ oppure } \omega \in B \text{ oppure } \omega \in C\}, \quad (1.4)$$

$$A \cap (B \cap C) = (A \cap B) \cap C = \{\omega : \omega \in A \text{ e } \omega \in B \text{ e } \omega \in C\} \quad (1.5)$$

Dalle (1.4) e (1.5) è chiaro come iterare le definizioni di unione e intersezione per famiglie finite di eventi. Ma in effetti non c'è alcun bisogno di limitare la cardinalità di una famiglia di eventi per definirne l'unione e l'intersezione. Se $\{A_n, n \in T\}$ è una famiglia qualunque di eventi di Ω definiamo

$$\bigcup_{n \in T} A_n = \{\omega : \exists i \in T \text{ con } \omega \in A_i\} \quad (1.6)$$

$$\bigcap_{n \in T} A_n = \{\omega : \forall i \in T, \, \omega \in A_i\}.$$

Nel caso in cui $A_i \cap A_j = \emptyset$ per ogni coppia di indici distinti i e $j \in T$, parliamo di **famiglia disgiunta di eventi** e corrispondentemente della loro

unione disgiunta. Si verifica facilmente che la validità delle formule di de Morgan e della proprietà distributiva non dipende dalla cardinalità delle famiglie di eventi, quindi

$$\left(\bigcup_{n\in T} A_n\right)^c = \bigcap_{n\in T} A_n^c, \quad \left(\bigcap_{n\in T} A_n\right)^c = \bigcup_{n\in T} A_n^c, \tag{1.7}$$

$$\left(\bigcup_{n\in T} A_n\right) \cap B = \bigcup_{n\in T} (A_n \cap B), \left(\bigcap_{n\in T} A_n\right) \cup B = \bigcap_{n\in T} (A_n \cup B). \tag{1.8}$$

1.1.3 Eventi, attributi e funzioni indicatrici

Solo in un esercizio molto artificiale può accadere che un evento venga descritto mediante la lista degli esiti di Ω che lo realizzano. Nelle applicazioni concrete un evento è dato descrivendo nel linguaggio comune l'**attributo** che un esito deve possedere per realizzarlo. Un attributo S è una caratteristica che ciascun esito di Ω possiede oppure no. L'evento associato ad un attributo è il sottoinsieme degli esiti di Ω che possiedono l'attributo. Deve essere sempre possibile determinare quali esiti possiedono un attributo e quali no: se scommettiamo sull'accadere di S vogliamo essere in grado di verificare se la scommessa è stata vinta o persa.

Esempio 1.1 *Nel gioco della roulette $\Omega = \{0, 1, ..., 36\}$ una delle scommesse possibili è su Pair. Gli esiti che pagano questa scommessa sono i numeri pari di Ω, escluso lo 0. Invece, se Ω è l'insieme degli studenti iscritti al secondo anno del corso di laurea in Matematica presso l'Università di Roma La Sapienza, "aver sostenuto l'esame di Calcolo delle Probabilità 1 è un attributo, mentre "secchione" non lo è, a meno di non specificare esattamente quali requisito uno studente secchione deve possedere.*

Attribuendo S ad uno specifico elemento ω dello spazio Ω si deve ottenere un'**asserzione** $S(\omega)$, cioè un'affermazione che è vera (se ω possiede l'attributo S) oppure falsa (se ω non lo possiede). Da un punto di vista matematico, quindi, un attributo definisce una funzione definita sullo spazio dei campioni Ω, a valori nell'insieme $\{Vero, Falso\}$, detta **funzione di verità** dell'attributo. L'evento associato all'attributo è dunque l'**insieme di verità**

$$\{\omega : \ S(\omega) = Vero\}$$

di questa funzione. Ovviamente attributi diversi possono dar luogo alla stessa funzione di verità su Ω, nel qual caso si dicono equivalenti su Ω. Inoltre ad ogni evento A di Ω è possibile associare l'attributo esito appartenente ad A che è ovviamente posseduto da tutti e soli gli elementi di A. In questo modo abbiamo quindi definito una **corrispondenza biunivoca tra** classi di equivalenza di **attributi ed eventi** di Ω.

Se utilizziamo la codifica binaria associando a $Vero$ il numero 1 e a $Falso$ il numero 0 traduciamo la funzione di verità di un attributo in una **funzione**

indicatrice su Ω, cioè una funzione f definita su Ω a valori in $\{0,1\}$. La funzione f indica l'evento

$$A = \{\omega : f(\omega) = 1\},$$

nel qual caso scriviamo $f = 1_A$. Ovviamente si ha anche

$$A^c = \{\omega : f(\omega) = 0\}.$$

In questo modo gli **eventi** di Ω sono posti **in corrispondenza biunivoca** anche **con** l'insieme delle **funzioni indicatrici** definite su Ω.

Queste corrispondenze biunivoche divengono isomorfismi (rispetto alle operazioni $\cup$, $\cap$, ^{c} con gli elementi neutri $\varnothing$ e Ω dotate delle proprietà (1)-(4) della Proposizione 1.2) se si definiscono delle opportune operazioni sugli attributi e sulle funzioni indicatrici. Nella Tabella 1.1 indichiamo come queste ultime operazioni sono definite. Per comodità di notazione indichiamo gli eventi con lettere maiuscole e gli attributi a questi corrispondenti con le stesse lettere, ma minuscole. Lasciamo al lettore il compito di esplicitare le definizioni delle operazioni sugli attributi, comunque già abbastanza evidenti dalla terminologia utilizzata.

eventi	funzioni indicatrici	attributi
A^c	$1 - 1_A$	$\neg a$ (negazione)
$A \cup B$	$1_A + 1_B - 1_A 1_B$	$a \vee b$ (disgiunzione)
$A \cap B$	$1_A 1_B$	$a \wedge b$ (congiunzione)
$A \triangle B$	$1_A + 1_B - 21_A 1_B$	$a^x\vee b$ (disgiunzione forte)
$A \cap B = \emptyset$	$1_A \cdot 1_B \equiv 0$	a e b incompatibili
$A \cup B = \Omega$	$1_A + 1_B \equiv 1$	a e b necessari
$A \subset B$	$1_A \equiv 1_A \cdot 1_B$	$a \Longrightarrow b$ (implicazione)
$A = B$	$1_A = 1_B$	$a \iff b$ (doppia implicazione)
$A = \Omega$	$1_A \equiv 1$	a certo
$A = \varnothing$	$1_A \equiv 0$	a impossibile

Tabella 1.1

Si noti dalla tabella che l'indicatrice dell'unione di due eventi A e B è la somma della funzione indicatrice di A e della funzione indicatrice di B se e solo se A e B sono disgiunti. Per questo motivo, e per altri che risulteranno evidenti nel seguito, può essere particolarmente conveniente rappresentare un evento come unione disgiunta.

Esempio 1.2 *Una classe importante di esempi è quella delle* ***reti aleatorie.*** *Una rete non è altro che un grafo in cui un certo insieme di vertici (ad esempio un certo insieme di città) è collegato da spigoli (ad esempio strade): si veda ad esempio la rete di Figura 1.2 a). Ciascuna strada può essere aperta o chiusa al traffico. Tra due città possono esistere più strade diverse. E' naturale per questa applicazione scegliere come spazio Ω quello delle funzioni indicatrici sull'insieme degli n spigoli del grafo, indicando con* 1 *l'apertura e con* 0 *la chiusura di ciascuna strada. Con questa convenzione l'evento* i-*esima strada aperta corrisponde* *al sottoinsieme*

$$A_i = \{\omega \in \Omega : \omega(i) = 1\} \subset \{0,1\}^n, i = 1, ..., n. \tag{1.9}$$

Fissata una città di partenza o e una di arrivo i un attributo di ovvio interesse è dato da è possibile portare a termine il viaggio da o a i. La funzione indicatrice dell'evento E associato a questo attributo è la funzione che a ciascuna delle n-ple in Ω che rappresentano gli stati possibili delle n strade associa 1 *o* 0 *a seconda che il viaggio in questione sia possibile o meno.*

Questa funzione può sempre essere scritta individuando tutti i cammini, *ovvero gli itinerari che portano da o a i (ovviamente bastano i cammini* minimali, *cioè quelli che non visitano una città più di una volta). E' possibile raggiungere i da o se e soltanto se c'è un cammino praticabile, cioè con tutte le strade previste aperte. In questo modo nella rete di Figura 1.2 a) si arriva a scrivere ad esempio*

$$E = (A_1 \cap A_2) \cup (A_4 \cap A_5) \cup (A_1 \cap A_3 \cap A_5) \cup (A_2 \cap A_3 \cap A_4) \tag{1.10}$$

Ovviamente una rappresentazione di questo tipo non è unica; ad esempio

$$E = \{A_3 \cap (A_1 \cup A_4) \cap (A_2 \cup A_5)\} \cup \{A_3^c \cap [(A_1 \cap A_2) \cup (A_4 \cap A_5)]\} \tag{1.11}$$

*si ottiene facilmente osservando che i due eventi tra parentesi graffe sono rispettivamente $E \cap A_3$ e $E \cap A_3^c$ (e quindi sono disgiunti). In particolare se o e i sono le sole città della rete e ci sono n strade alternative tra di esse (**connessione in parallelo,** Figura 1.2 b)) allora*

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

*mentre se c'è un solo itinerario possibile tra o e i, che passa per n città intermedie (**connessione in serie,** Figura 1.2 c)), si ha*

$$E = \bigcap_{i=1}^{n} A_i.$$

L'isomorfismo tra eventi e attributi ha un'importanza soprattutto terminologica, dato che in probabilità il linguaggio degli eventi tende a confondersi con quello degli attributi. Ad esempio l'evento Ω viene di solito chiamato **evento certo,** mentre l'insieme vuoto $\emptyset$ è detto **evento impossibile.** L'isomorfismo tra eventi e funzioni indicatrici è invece di grande utilità pratica perchè permette di "tradurre" operazioni e relazioni sugli insiemi in operazioni e relazioni sui valori delle funzioni indicatrici, cioè su numeri. L'esempio seguente è solo una prima illustrazione, che riceverà conferme ulteriori durante tutto il corso.

Esempio 1.3 *Mostriamo come si ottiene la proprietà (2) della Proposizione 1.2 (che in precedenza abbiamo omesso di provare)*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

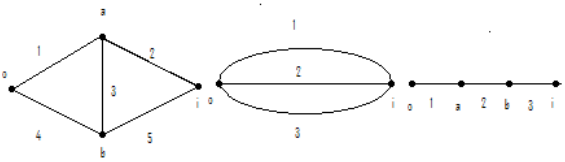

Figura 1.2: Esempi di reti aleatorie a) generica; b) in parallelo; c) in serie.

valida qualunque siano gli eventi A, B e C, verificandola mediante l'identità delle funzioni indicatrici. Per il membro di sinistra

$$1_{A\cap(B\cup C)} = 1_A 1_{B\cup C} = 1_A\left(1_B + 1_C - 1_B 1_C\right) = 1_A 1_B + 1_A 1_C - 1_A 1_B 1_C$$

mentre per quello di destra

$$1_{(A\cap B)\cup(A\cap C)} = 1_{A\cap B} + 1_{A\cap C} - 1_{A\cap B}1_{A\cap C} = 1_A 1_B + 1_A 1_C - 1_A 1_B 1_C$$

che conclude la semplice dimostrazione.

Ritorniamo infine sul problema della scelta dello spazio dei campioni Ω. E' naturale chiedersi se è possibile scegliere in modo automatico lo spazio dei campioni Ω dopo aver specificato quali sono gli attributi di interesse. Questo è possibile, ma può non risultare conveniente per la specificazione della probabilità su Ω, come vedremo nella prossima sezione. Certamente gli attributi interessanti possono però segnalare la necessità di aggiornare la scelta dello spazio dei campioni, come accade nell'esempio seguente.

Esempio 1.4 *Gli eventi sportivi offrono una gran quantità di esempi di esperimenti aleatori (per la gioia degli scommettitori!). Consideriamo ad esempio i risultati di un girone eliminatorio della Champions League di calcio, cui partecipano 4 squadre. Nelle 6 giornate in cui il girone si articola ciascuna squadra incontra tutte le altre, una volta in casa, l'altra in trasferta (si parla in questo caso di girone all'italiana). Se lo svolgimento sequenziale delle partite non interessa, possiamo limitarci a rappresentare tutti i risultati di queste in una matrice*

4×4 in cui ad ogni squadra è associata una riga e una colonna. L'elemento nella casella (i, j) della matrice rappresenta il risultato della partita tra la squadra i (che gioca in casa) e la squadra j (che gioca in trasferta), ovviamente con $i \neq j$. Dato che il risultato è una coppia di interi non negativi e che vengono giocate $4 \times 3 = 12$ partite lo spazio delle matrici di risultati può essere identificato con $\Omega = \left(\mathbf{N}^2\right)^{12}$, dopo aver ordinato in un modo qualsiasi le caselle della matrice. Supponiamo di essere interessati all'attributo la squadra 1 è qualificata per il turno successivo. Come possiamo determinare il sottoinsieme di Ω in cui questo avviene? La matrice di risultati permette di calcolare la classifica del girone, ottenuta contando 3 punti per ogni vittoria e 1 per ogni pareggio: se una squadra è tra le prime due della classifica girone si qualifica. Ovviamente può succedere che due o più squadre totalizzino gli stessi punti, nel qual caso il regolamento prevede meccanismi più o meno complicati per determinare il passaggio del turno (differenza reti, numero reti segnate, classifica avulsa, che possono comunque tutte essere determinate dalla matrice dei risultati). Ma qualsiasi meccanismo per risolvere il passaggio del turno in caso di parità punti che non privilegi a priori *una squadra rispetto ad un'altra deve prevedere l'effettuazione di un sorteggio in caso di parziale o completa simmetria della matrice dei risultati, ad esempio quando tutte le partite terminano in pareggio sullo $0-0$. Questo significa che l'attributo cui siamo interessati non è associato ad un sottoinsieme di Ω. Lo spazio dei campioni va quindi allargato includendo anche l'esito di un eventuale sorteggio se vogliamo far corrispondere all'attributo di interesse un evento. La scelta fatta per Ω è invece adeguata se vige la regola di non pagare una scommessa sulla qualificazione della squadra 1 al turno successivo se questa avviene per sorteggio.*

1.1.4 Esercizi

1. Siano A, B e C eventi di un qualche spazio dei campioni Ω. Esprimere i seguenti eventi attraverso operazioni di unione, intersezione e complementazione su A, B e C e scrivere le loro funzioni indicatrici mediante le funzioni indicatrici di A, B e C.

   a) Esattamente uno tra questi eventi si verifica.

   b) Esattamente due tra questi eventi si verificano.

   c) Al più due tra questi eventi si verificano.

   d) Almeno due tra questi eventi si verificano.

2. a) Utilizzando le funzioni indicatrici, si dimostri che l'operazione di intersezione $\cap$ è distributiva rispetto a quella di differenza simmetrica Δ.

   b) Si determini l'operazione duale di Δ, che chiamiamo per comodità ∇.

   c) Si determini la relazione di uguaglianza duale di quella provata in a).

   d) Si dimostri direttamente la relazione in c) aiutandosi con un diagramma di Venn.

3. All'inizio di un torneo di tennis i 32 giocatori vengono inseriti nel tabellone, che determina le partite che devono essere effettuate fino alla fine del torneo. Dopo le prime 16 partite i 16 perdenti sono eliminati e i vincenti si accoppiano nel modo determinato dal tabellone, e così via.

   a) Se consideriamo lo svolgimento del torneo come un esperimento aleatorio, quale spazio dei campioni va scelto?

   b) Che cardinalità ha questo spazio?

1.2 Le funzioni di massa e di probabilità

1.2.1 Spazi finiti

L'elenco degli esiti possibili di un esperimento è solo il primo elemento per specificare un modello matematico dell'esperimento stesso. E' necessario poi quantificare l'incertezza circa il risultato dell'esperimento dando *valutazioni di probabilità* a questi esiti (e più in generale ad un qualsiasi evento). Dato che riteniamo inopportuna in questa fase iniziale una discussione estesa dei vari possibili approcci riguardo al significato della probabilità, preferiamo limitarci a fare appello all'intuizione. Possiamo quindi convenire che *la probabilità misura la tendenza degli eventi a verificarsi*, collocandoli in una scala che va dall'*impossibilità* (probabilità 0) alla *certezza* (probabilità 1). La nostra discussione si focalizza invece sulle regole che in ogni caso (qualunque sia il significato cui facciamo riferimento) la probabilità deve rispettare. Cominciamo col supporre *lo spazio dei campioni* Ω *finito.*

Definizione 1.2 *Una* ***funzione di massa*** p *è una funzione definita su uno spazio dei campioni* Ω*, a valori non negativi, tale che*

$$\sum_{\omega \in \Omega} p(\omega) = 1. \tag{1.12}$$

La massa $p(\omega)$ dell'esito $\omega \in \Omega$ va interpretata come "peso" o probabilità dell'**evento elementare** o **singoletto** $\{\omega\}$, evento realizzato dall'unico esito ω. Dato che gli esiti sono incompatibili, per "pesare" un generico evento A evento dobbiamo mettere sulla bilancia tutti i pesi degli esiti che lo realizzano, che è quello che facciamo definendo la funzione

$$\mathrm{P}(A) = \sum_{\omega \in A} p(\omega), \quad \forall A \subset \Omega. \tag{1.13}$$

Proposizione 1.5 *La funzione* P *definita in (1.13) sull'insieme delle parti* $\mathcal{P}(\Omega)$ *dello spazio dei campioni* Ω *è* ***normalizzata,*** *cioè*

$$\mathrm{P}(\Omega) = 1, \tag{1.14}$$

e ***finitamente additiva,*** *cioè se* $A_1, ..., A_n$ *è una famiglia disgiunta di eventi, allora*

$$\mathrm{P}(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathrm{P}(A_i). \tag{1.15}$$

Prova. La (1.14) si ottiene ponendo $A = \Omega$ nella (1.13) e sfruttando la (1.12). La finita additività risulta dal fatto che

$$\mathrm{P}(\bigcup_{i=1}^{n} A_i) = \sum_{\omega \in \bigcup_{i=1}^{n} A_i} p(\omega) = \sum_{i=1}^{n} \sum_{\omega \in A_i} p(\omega) = \sum_{i=1}^{n} \mathrm{P}(A_i). \tag{1.16}$$

■

Chiamiamo quindi *probabilità* su Ω una qualsiasi funzione non negativa, normalizzata e finitamente additiva, definita su $\mathcal{P}(\Omega)$.

Perchè una funzione P, definita sugli eventi di uno spazio dei campioni Ω, sia finitamente additiva basta che soddisfi la (1.15) solo per $n = 2$. La dimostrazione si ottiene facilmente per ***induzione.*** Nelle **dimostrazioni per induzione** si tratta di dimostrare che una certa proprietà è vera per ogni valore di n. Si inizia a dimostrare la proprietà per il valore più piccolo di n per cui ha senso, che in questo caso è $n = 2$, in cui la proprietà è vera per ipotesi. Si assume quindi la proprietà vera per un generico n (*ipotesi induttiva)* e si dimostra per l'intero successivo $n+1$. Allora la proprietà è vera per ogni n intero. Assumiamo quindi la (1.15) vera per un generico n, e consideriamo una famiglia disgiunta $\{A_i, i = 1, ..., n+1\}$. Dato che

$$A_{n+1} \cap (\bigcup_{i=1}^{n} A_i) = \bigcup_{i=1}^{n} (A_{n+1} \cap A_i) = \emptyset$$

applicando la (1.15) ai 2 eventi $\bigcup_{i=1}^{n} A_i$ e A_{n+1} e utilizzando poi l'ipotesi induttiva si ottiene

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

che conclude la prova.

Proposizione 1.6 *Una probabilità P su di uno spazio finito* Ω *si può sempre rappresentare come in (1.13), prendendo la funzione di massa* $p(\omega) = \mathrm{P}(\{\omega\})$.

Prova. Qualunque sia la cardinalità di A, possiamo scrivere

$$A = \bigcup_{\omega \in A} \{\omega\} \tag{1.17}$$

e se A è finito possiamo dedurre dalla finita additività che

$$\mathrm{P}(A) = \mathrm{P}(\bigcup_{\omega \in A} \{\omega\}) = \sum_{\omega \in A} \mathrm{P}(\{\omega\}) = \sum_{\omega \in A} p(\omega).$$