Capitolo 1

Introduzione al calcolo delle probabilità: concetti e risultati preliminari

1.1 Il ragionamento probabilistico

Il ragionamento probabilistico è un tipo di ragionamento di tipo induttivo, anzichè deduttivo come nelle usuali discipline matematiche. Il problema nasce dal fatto che spesso si è in condizioni di conoscenza parziale, cioè si è in uno stato di incertezza circa il fatto che un certo evento si sia verificato o possa verificarsi. In problemi di carattere medico, ad esempio, la diagnosi della malattia non segue in maniera univoca dal complesso dei sintomi evidenziati; oppure in problemi di carattere finanziario, la scelta di investire o meno in un certo titolo azionario non scaturisce necessariamente dall'andamento del titolo nel periodo immediatamente precedente.

La probabilità ha lo scopo di misurare la nostra incertezza in merito a fenomeni di particolare interesse chiamati eventi. Ovviamente si pone il problema di ridurre l'incertezza: a tal scopo è necessario raccogliere informazioni (dati statistici, ”campione”) ed analizzarle utilizzando i metodi propri della statistica; ciò potrà condurre poi ad aggiornare la valutazione di probabilità iniziale.

La valutazione della incertezza è un problema delicato. Ci sono casi in cui questo processo è abbastanza immediato e semplice:

1. Criterio classico o combinatorico. Consideriamo ad esempio un'urna contenente n palline di cui k bianche e le restanti $n-k$ nere; se si considera l'estrazione di una pallina, allora gli eventi possibili sono due: ”pallina bianca” oppure ”pallina nera”. Tali eventi sono incompatibili, nel senso che il risultato di un'estrazione non può essere una pallina che sia contemporaneamente bianca e nera (formalizzaremo adeguatamente più avanti questi concetti, per ora lasciamoli ad un livello più intuitivo); in questo caso la nostra incertezza circa l'estrazione di una pallina bianca può essere valutata considerando il rapporto p fra il numero di palline bianche k ed il numero totale di palline contenute nell'urna n:

$$p = \frac{k}{n}\,;$$

2. Criterio sperimentale o frequentista. Consideriamo un fenomeno ”ripetibile” espresso mediante una successione di eventi ”analoghi” $E_1, E_2, \ldots$ del tipo successo-insuccesso,

pertanto gli eventi $E_1, \ldots, E_i, \ldots$ possono essere considerati come i risultati di prove di uno stesso esperimento ripetuto più volte nelle stesse condizioni (anche qui, formalizzaremo adeguatamente più avanti questi concetti, per ora lasciamoli ad un livello più intuitivo). Supponiamo di osservare m eventi $E_1, \ldots, E_m$ (con m opportunamente grande), allora la nostra incertezza circa il verificarsi di un successo al tempo $m + 1$, cioè la probabilità dell'evento E_{m+1}, viene valutata sulla base del dato "storico" considerando la frequenza relativa del numero di "successi" nelle m prove precedenti, cioè il rapporto fra il numero di successi ottenuto nelle m prove precedenti ed il numero m di prove effettuate.

Ci sono però moltissimi altri casi – e sono i più interessanti nell'ambito dell'economia così come in quello delle scienze applicate o della vita reale – in qui la valutazione dell'incertezza non rientra in nessuno dei due casi precedenti: si pensi alla valutazione del rischio di un investimento finanziario, all'analisi delle acque di scarico di un'industria inquinante, alla valutazione dell'incertezza in merito all'esito di una richiesta di finanziamento per dato progetto, o anche alla possibilità di vittoria dell'Inter nella prossima partita. Abbiamo pertanto bisogno di un modello più generale di probabilità che consenta di poter essere utilizzato in maniera generale e che possa tener conto del complesso di informazione a nostra disposizione circa l'evento di interesse. Ciò introdurrà al concetto di probabilità come grado di fiducia.

Esempio 1.1 Analizzando le acque di scarico di un'industria fortemente inquinata, si è osservato che una frazione $\theta = 0.7$ di campioni di tali acque contiene una sostanza fortemente nociva e che tale sostanza di campioni inquinati si riduce a $\theta_D = 0.4$ se l'industria mette in funzione l'impianto di depurazione di cui è dotata (ma che ha un costo di esercizio molto elevato). Le autorità sanitarie effettuano senza preavviso un controllo delle acque di scarico, analizzando n campioni e trovando la sostanza nociva in h di essi. Ci si chiede qual era la probabilità che l'impianto di depurazione fosse in funzione al momento del prelievo?

E' evidente che in questo caso, i due criteri combinatorio e frequentista non sono adeguati.

1.2 Eventi

Per avviare la nostra discussione dobbiamo preliminarmente delimitare il campo dei fatti (o delle situazioni o delle asserzioni) nel quale si possa o si debba parlare di *incertezza* e quindi di probabilità. Nell'ambito del linguaggio probabilistico, tali fatti vengono chiamati *eventi*.

Per (o *evento* si intende un'*affermazione verificabile*, cioè qualcosa che può assumere solo due valori: *vero* o *falso*; non è detto però che si sappia con certezza la risposta, ed allora i casi sono (almeno provvisoriamente) tre: *certo*, *impossibile*, *incerto*, a seconda che sappiamo la risposta (SI o NO) o si rimanga in uno stato di incertezza (NON SO). La probabilità è una valutazione numerica dell'incertezza mediante una quantità compresa fra zero e uno, in particolare tale quantità assume valore "zero" nel caso di eventi impossibili ed assume valore "uno" nel caso di eventi certi. Se indichiamo con X il "risultato del lancio di un dado", allora X è una quantità aleatoria che sappiamo poter assumere un valore intero fra 1 e 6. L'incertezza circa gli eventi "esce il numero 4", cioè $(X = 4)$, oppure "esce un numero pari", cioè $(X \in \{2, 4, 6\})$, viene valutata calcolando la probabilità di tali eventi.

Definizione 1.2 Un *evento* è un fatto fisico o concettuale descritto da una proposizione che ammette due soli valori logici: VERO (V) o FALSO (F).

Si noti che uno stesso evento può essere descritto mediante proposizioni diverse.

Esempio 1.3 Se dalla produzione di una fabbrica di lampadine si prelevano tre pezzi, la proposizione:

$$A = \text{“almeno una delle tre lampadine dura più di 250 ore”}$$

rappresenta un evento. Fino a quando non venga effettuata una verifica, si è in condizioni di incertezza.

Esempio 1.4 Le proposizioni

$$A = \text{“Il prezzo di riferimento del titolo Alleanza al 13 febbraio 2004 è superiore a 9.24 euro”}$$
$$B = \text{“Il prezzo di riferimento del titolo Alleanza al 5 marzo 2004 è superiore a 9.55 euro”}$$

rappresentano degli eventi. La verità di tale evento può essere verificata solo al momento opportuno. Si noti che la verità o falsità di un evento non riguarda necessariamente eventi futuri, ma anche eventi già accaduti su cui non si ha completa informazione. ♣

Due eventi di particolare importanza sono *evento certo* Ω e l'*evento impossibile* $\emptyset$. Nel caso più frequente, si è in uno stato di incertezza e si parla di *evento possibile*.

In maniera equivalente, lo stato logico di un evento A può essere assegnato mediante la *funzione indicatrice* o *indicatore* di A.

Definizione 1.5 Si definisce *indicatore* di un evento A – e si indica con $|A|$ o con I_A – la quantità:

$$|A| := \begin{cases} 1 & \text{se } A \text{ è vero} \\ 0 & \text{se } A \text{ è falso} \end{cases}$$

In particolare risulta $|\Omega| = 1$ e $|\emptyset| = 0$.

Introduciamo le seguenti operazioni logiche fra due eventi A e B.

Definizione 1.6 Siano A e B due eventi. Si definisce

1. evento *somma logica* $A \vee B$ l'evento vero se e solo se almeno uno dei due eventi è vero. Le tabelle della verità (di cui la seconda è data in termini degli indicatori) sono:

A	B	$A \vee B$
V	V	V
V	F	V
F	V	V
F	F	F

$\|A\|$	$\|B\|$	$\|A \vee B\|$
1	1	1
1	0	1
0	1	1
0	0	0

In particolare, due eventi A e B si dicono *necessari* se $A \vee B = \Omega$.

2. evento *prodotto logico* $A \wedge B$ (indicato anche con AB) è l'evento vero se e solo se entrambi gli eventi sono veri. Le tabelle della verità sono:

A	B	$A \wedge B$
V	V	V
V	F	F
F	V	F
F	F	F

$\lvert A\rvert$	$\lvert B\rvert$	$\lvert A \wedge B\rvert$
1	1	1
1	0	0
0	1	0
0	0	0

In particolare due eventi A e B vengono detti *incompatibili* se $A \wedge B = \emptyset$.

3. evento *contrario* o *complementare* A^c o $\overline{A}$ è l'evento vero se e solo se l'evento A è falso. Le tabella della verità sono (la seconda in termini degli indicatori):

A	A^c
V	F
F	V

$\lvert A\rvert$	$\lvert A^c\rvert$
1	0
0	1

4. evento *differenza* A meno B, e si scrive $A \setminus B$ definito da $A \setminus B = A \wedge B^c$. Si noti che in generale gli eventi $A \setminus B$ e $B \setminus A$ sono diversi.

Ovviamente, per qualsiasi evento A risulta: $A \vee A^c = \Omega$ e $A \wedge A^c = \emptyset$.

Definizione 1.7 Siano A e B due eventi. Diremo che A *implica* B, e scriveremo $A \Rightarrow B$ se B risulta vero ogni volta che A è vero.

Definizione 1.8 Siano A e B due eventi. Diremo che A e B sono *logicamente indipendenti* se nessuno dei quattro eventi: $A \wedge B$, $A \wedge B^c$, $A^c \wedge B$, $A^c \wedge B^c$ è impossibile.

Esempio 1.9 Dalla produzione giornaliera di un dato macchinario si prelevano 4 pezzi e denotiamo con E_i l'evento "l'i-esimo pezzo funziona correttamente", con $i = 1, 2, 3, 4$. Sia:

$$A = \text{“almeno uno dei 4 articoli è difettoso”},$$
$$B = \text{“almeno due dei 4 articoli sono difettosi”}.$$

Ovviamente si ha:

$$\begin{aligned} A = \{&E_1^c E_2 E_3 E_4 \vee E_1 E_2^c E_3 E_4 \vee E_1 E_2 E_3^c E_4 \vee E_1 E_2 E_3 E_4^c \vee E_1^c E_2^c E_3 E_4 \vee E_1^c E_2 E_3^c E_4 \vee \\ &E_1^c E_2 E_3 E_4^c \vee E_1 E_2^c E_3^c E_4 \vee E_1 E_2^c E_3 E_4^c \vee E_1 E_2 E_3^c E_4^c \vee E_1^c E_2^c E_3^c E_4 \vee E_1^c E_2^c E_3 E_4^c \vee \\ &E_1^c E_2 E_3^c E_4^c \vee E_1 E_2^c E_3^c E_4^c\} \end{aligned}$$

Le proposizioni che individuano gli eventi A^c e B^c si enunciano:

$$A^c = \text{“nessuno dei 4 articoli è difettoso”},$$
$$B^c = \text{“al più uno dei 4 articoli è difettoso”}.$$

e risulta, ad esempio, $A = E_1 E_2 E_3 E_4$. Calcolare gli eventi B e B^c in termini degli eventi E_i. ♣

Diagrammi di Venn. Gli eventi possono essere descritti mediante insiemi. Considerato un insieme Ω, che costituisce l'evento certo, gli eventi sono allora sottoinsiemi di Ω. Gli eventi possono quindi essere rappresentati graficamente mediante i cosidetti *diagrammi di Venn.* I concetti della logica degli eventi si possono allora far corrispondere a quelli della *teoria degli insiemi* secondo il seguente schema:

Eventi	*Simbologia*	**Insiemi**	*Simbologia*
certo	Ω	ambiente	Ω
impossibile	$\emptyset$	vuoto	$\emptyset$
implicazione	$A \Rightarrow B$	inclusione	$A \subseteq B$
contrario	$\overline{A}$	complementare	A^c
somma logica	$A \vee B$	unione	$A \cup B$
prodotto logico	$A \wedge B$	intersezione	$A \cap B$
incompatibili	$A \wedge B = \emptyset$	disgiunti	$A \cap B = \emptyset$
classe completa	$\bigvee_i A_i = \Omega$	partizione	$\bigcup_i A_i = \Omega$

Nel seguito indicheremo con lo stesso simbolo A sia l'evento che l'insieme che lo rappresenta; inoltre per uniformarci alla notazione utilizzata usualmente in letteratura, utilizzeremo la notazione insiemistica.

Definizione 1.10 Assegnato l'evento certo Ω, diremo che A è un evento di Ω se A implica Ω ($A \Rightarrow \Omega$) o, equivalentemente, se A è contenuto in Ω ($A \subseteq \Omega$).

Sia Ω l'evento certo. Per semplicità con il termine *insieme* o *famiglia di eventi di* Ω intenderemo una qualunque famiglia $\mathcal{A} = \{A_1, \ldots, A_n\}$ di eventi di Ω.

Definizione 1.11 Sia $\mathcal{A} = \{A_1, A_2, \ldots\}$ un insieme di eventi. Un evento E si dice *elementare* rispetto ad $\mathcal{A}$ – oppure *atomo* di $\mathcal{A}$ – se per ogni $A \in \mathcal{A}$ risulta

$$A \cap E = E \qquad \text{oppure} \qquad A \cap E = \emptyset\ .$$

Tale definizione stabilisce che se E è un evento elementare, allora per ogni evento A di $\mathcal{A}$ si possono avere solo due casi: E è incluso in A, oppure E è incompatibile con A. Usualmente gli eventi elementari vengono denotati col simbolo ω. L'insieme $\{\omega\}$ viene chiamato *singoletto* di Ω.

Esercizio 1.12 Una compagnia di assicurazioni è interessata alla distribuzione per età dei coniugi: se x è l'età del marito, y quella della moglie, ad ogni osservazione si può far corrispondere un punto del primo quadrante ($x > 0, y > 0$) nel piano (x, y). Determinare gli insiemi del piano che rappresentano gli eventi:

$$\begin{aligned} A &= \text{“il marito ha più di 40 anni”}\ , \\ B &= \text{“il marito è più anziano della moglie”}\ , \\ C &= \text{“la moglie ha meno di 42 anni”}\ . \end{aligned}$$

Esercizio 1.13 Uno studente sostiene tre esami in una sessione di esami. Consideriamo i seguenti eventi:

$$E_1 = \text{“lo studente supera il primo esame”} ,$$
$$E_2 = \text{“lo studente supera il secondo esame”} ,$$
$$E_3 = \text{“lo studente supera il terzo esame”} .$$

Calcolare i seguenti eventi: i) Lo studente supera tutti e tre gli esami; ii) lo studente supera due esami; iii) lo studente supera almeno due esami; iv)lo studente supera un esame; v) lo studente supera il primo esame; vi) lo studente supera al più due esami; vii) lo studente non supera alcun esame.

Esercizio 1.14 Si consideri il lancio di due dadi. Calcolare gli eventi tali che la somma dei due dadi sia uguale a 3, a 5, ed a 14. Qual è quel numero che corrisponde al maggior numero di coppie di risultati?

Esercizio 1.15 Un concessionario di automobili offre autoveicoli con le seguenti opzioni: i) con e senza cambio automatico; ii) con e senza climatizzatore; iii) con una o due scelte di impianto stereo; iv) tre diversi colori esterni.
Qual è l'insieme Ω? Quali sono gli eventi elementari di Ω? Quanti sono?

1.2.1 Proprietà delle operazioni fra eventi

Si può dimostrare che le operazioni fra eventi soddisfano le seguenti proprietà:

Proprietà	*Unione*	*Intersezione*
Commutativa	$A \cup B = B \cup A$	$A \cap B = B \cap A$
Associativa	$(A \cup B) \cup C = A \cup (B \cup C)$	$(A \cap B) \cap C = A \cap (B \cap C)$
Idempotenza	$A \cup A = A$	$A \cap A = A$
Distributiva	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Formule di De Morgan. Un'importante relazione fra gli eventi è descritta dalle *Formule di De Morgan*. Nel caso di due eventi A e B si ha:

$$(A \cup B)^c = A^c \cap B^c \tag{1.1}$$
$$(A \cap B)^c = A^c \cup B^c \ . \tag{1.2}$$

E' possibile fare una verifica di tali relazioni in termini di tabelle di verità, rispettivamente per la (1.1) e la (1.2):

A	B	$A \cup B$	$(A \cup B)^c$	A^c	B^c	$A^c \cap B^c$
V	V	V	F	F	F	F
V	F	V	F	V	F	F
F	V	V	F	V	F	F
F	F	F	V	V	V	V

A	B	$A \cap B$	$(A \cap B)^c$	A^c	B^c	$A^c \cup B^c$
V	V	V	F	F	F	F
V	F	F	V	V	F	V
F	V	F	V	V	F	V
F	F	F	V	V	V	V

Provare per esercizio a fare la verifica utilizzando i diagrammi di Venn.

Più in generale per un insieme al più numerabile di eventi $A_1, A_2, \ldots, A_n, \ldots$ si ha:

$$\bigcup_{n=1}^{\infty} A_n^c = \left(\bigcap_{n=1}^{\infty} A_n \right)^c \tag{1.3}$$

$$\bigcap_{n=1}^{\infty} A_n^c = \left(\bigcup_{n=1}^{\infty} A_n \right)^c \tag{1.4}$$

1.2.2 Famiglie di eventi di particolare importanza.

Introduciamo ora alcune famiglie di insiemi che verranno considerate frequentemente in seguito.

Definizione 1.16 Sia Ω un insieme. Una famiglia $\mathcal{A} = \{A_1, A_2, \ldots\}$ al più numerabile di sottoinsiemi di Ω si chiama *partizione* di Ω se possiede le seguenti proprietà:

1. Se $A_i \cap A_j = \emptyset$ per ogni $A_i, A_j \in \mathcal{A}$ con $i \neq j$ (gli eventi di $\mathcal{A}$ sono a due a due incompatibili)

2. $\displaystyle\bigcup_{i=1}^{\infty} A_i = \Omega$ (l'unione di tutti gli eventi di $\mathcal{A}$ costituisce l'evento certo).

In particolare si parla di *partizione finita* nel caso in cui $\mathcal{A}$ contiene un numero finito di eventi e di *partizione numerabile* nel caso in cui $\mathcal{A}$ è una famiglia numerabile di eventi, cioè una famiglia costituita da un numero infinito numerabile di eventi. Quindi una partizione è una classe completa di eventi a due a due incompatibili in cui uno ed uno solo è vero. Semplici esempi di partizione possono essere costruiti considerando i risultati ottenuti lanciando un dado $\Omega = \{1, 2, 3, 4, 5, 6\}$. Possiamo avere $\mathcal{A} = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}$(risultati elementari) oppure $\mathcal{A} = \{\{1, 3, 5\}, \{2, 4, 6\}\}$ (risultati dispari e risultati pari), $\mathcal{A} = \{\{1, 2, 3\}, \{4, 5, 6\}\}$ (risultati inferiori a 4 e risultati maggiori o uguali a 4), e così via.

Esempio 1.17 Sia $\mathcal{U}$ un'urna contenente $n = 7$ palline, ciascuna delle quali può essere bianca o nera. La composizione H dell'urna è un evento fra i seguenti possibili:

$$\begin{array}{llll} H_0 = \{7B\} & H_1 = \{6B + 1N\} & H_2 = \{5B + 2N\} & H_3 = \{4B + 3N\} \\ H_4 = \{3B + 4N\} & H_5 = \{2B + 5N\} & H_6 = \{1B + 6N\} & H_7 = \{7N\}. \end{array}$$

Pertanto l'insieme $\{H_0, H_1, \ldots, H_7\}$ costituisce una partizione sull'insieme delle possibili composizioni dell'urna $\mathcal{U}$. ♣

In termini di indicatori, se una famiglia $\mathcal{A}$ costituisce una partizione, allora risulta:

$$\sum_{i=1}^{n} |A_i| = |A_1| + |A_2| + \cdots + |A_n| = 1 \ ,$$

o, più in generale:

$$\sum_{i=1}^{\infty} |A_i| = |A_1| + |A_2| + \cdots + |A_n| + \cdots = 1 \ ,$$

a seconda che si tratti, rispettivamente, di una partizione finita o infinita.

Definizione 1.18 Sia Ω un insieme. Una famiglia finita $\mathcal{A} = \{A_1, \ldots, A_n\}$ di sottoinsiemi di Ω si chiama *algebra* se gode delle seguenti due proprietà:

1. Se $A \in \mathcal{A}$ allora $A^c \in \mathcal{A}$;
2. Se $A, B \in \mathcal{A}$ allora $A \cup B \in \mathcal{A}$.

Si dimostra che, se $\mathcal{A}$ è un'algebra di parti di Ω, allora $\Omega, \emptyset \in \mathcal{A}$ e per ogni $A, B \in \mathcal{A}$ si ha $A \cap B \in \mathcal{A}$. Infatti:

- $\underline{\Omega \in \mathcal{A}}$: Dalla 1., se $A \in \mathcal{A}$ allora anche $A^c \in \mathcal{A}$; dalla 2. segue quindi che $A \cup A^c = \Omega \in \mathcal{A}$.
- $\underline{\emptyset \in \mathcal{A}}$: Se $\Omega \in \mathcal{A}$, per la 1. si ha che $\Omega^c = \emptyset \in \mathcal{A}$.
- $\underline{A \cap B \in \mathcal{A}}$: siano $A, B \in A$, allora per la 1. si ha che $A^c, B^c \in \mathcal{A}$ e per la 2. $A^c \cup B^c \in \mathcal{A}$. Per la formula di De Morgan si ha che $A^c \cup B^c = (A \cap B)^c \in \mathcal{A}$ e quindi ancora per la 1. segue che $A \cap B \in A$. □

Definizione 1.19 Sia Ω un insieme. Una famiglia $\mathcal{A} = \{A_1, A_2, \ldots, A_n, \ldots\}$ di parti di Ω si chiama *σ-algebra* se:

1. $\emptyset, \Omega \in \mathcal{A}$
2. Se $A \in \mathcal{A}$ allora $A^c \in \mathcal{A}$;

3. Se $A_1, A_2, \ldots \in \mathcal{A}$ allora si ha:

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{A} \qquad \text{e} \qquad \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}\,.$$

Si noti in effetti che nella definizione precedente la proprietà 1. è ridondante e, per le formule di De Morgan, anche la seconda condizione della proprietà 3.

L'individuazione di Ω può essere a volte problematica in quanto, come in ogni astrazione matematica di situazioni reali, vi è un certo grado di arbitrarietà insito nella scelta dell'insieme Ω di tutti i possibili risultati di un esperimento. Tale scelta di solito implica opportune semplificazioni.

1.3 Le diverse concezioni della probabilità

La probabilità sorge nel XVII secolo con la pubblicazione dell'*Ars Conjectandi* di James Bernoulli dal tentativo di valutare e gestire quantitativamente l'incertezza a partire da studio di problemi sui giochi d'azzardo. Per quanto riguarda l'aspetto più propriamente matematico, le idee sulla probabilità non sono fortemente discordanti. Il nucleo della discordia concerne invece il significato della nozione di probabilità: vi sono concezioni secondo cui, per elevare a rango di scientifico la probabilità si deve depurarla di tutti gli elementi soggettivi che la riguardano e la caratterizzano; dall'altro vi sono quelle concezioni che non solo non ritengono tali elementi come causa di disturbo, ma ne fanno il punto di partenza per la definizione e la conseguente teoria matematica. Le concezioni del primo tipo (*oggettive*) ritengono la probabilità come elemento del mondo fisico, esistente fuori di ciascuno di noi; le concezioni del secondo tipo (*soggettive*) ritengono che la probabilità, esprimendo in ogni caso un'opinione personale, venga ad esistenza solo nel momento in cui il giudizio del soggetto viene a contatto di date circostanze concrete, incerte.

1.3.1 La concezione classica

Il primo tentativo di una formalizzazione matematicamente rigorosa della teoria della probabilità è dovuto a P.-S. de Laplace. Assumiamo che un evento A sia costituito dall'unione di un certo numero di eventi elementari. Nella sua opera monumentale[1], viene data la definizione *classica* della probabilità di un evento A come proporzione del numero di risultati favorevoli (eventi elementari) all'evento A ed il numero (finito) di tutti i possibili risultati (eventi elementari) connessi all'esperimento in questione, dove tali risultati vengono supposti *egualmente possibili*.

Definizione 1.20 (classica) Sia $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ l'insieme degli eventi elementari (possibili risultati) concernenti un certo esperimento $\mathcal{E}$, supposti ugualmente possibili; siano $\omega_{i_1}, \omega_{i_2}$, ω_{i_p} i p risultati favorevoli ad un certo evento A, cioè $A = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega i_p\}$. Si definisce

[1] P.-S. de Laplace, *Théorie analytique des probabilités,* Courcier, Paris, 1814.

probabilità dell'evento A la quantità:

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

dove il simbolo $\#$ denota "numero di elementi contenuti in".

Ovviamente, se l'evento A è impossibile, allora $\#A = 0$ e quindi $P(A) = 0$.

Si noti che, in accordo a tale definizione, il calcolo delle probabilità si riduce a problemi di conteggio di eventi, sulla base di regole del calcolo combinatorio. Tale definizione è inadeguata in quanto circolare e restrittiva; essa tuttavia coglie un aspetto importante legato al fatto che esistono situazioni in cui, per esempio in base a considerazioni di tipo geometrico, si può ammettere che gli eventi elementari siano equiprobabili.

1.3.2 La concezione frequentista

La concezione frequentista proposta da von Mises[2] è quella cui aderiscono molti testi scientifici, specie nell'ambito delle scienze sperimentali. Sia $A_1, A_2, \ldots, A_n$ una successione di n prove ripetute ed indipendenti (ad esempio, lanci di una moneta) in cui un certo risultato A (ad esempio, "Testa") può o meno verificarsi. Pertanto A_1 risulta vero se l'evento A si è verificato nella prima prova (cioè se esce "testa" nel primo lancio), A_2 risulta vero se l'evento A si è verificato nella seconda prova (cioè se esce "testa" nel secondo lancio), e così via. Indichiamo con f_n la frazione di successi sulle n prove, cioè la frazione del numero di volte in cui si è verificato l'evento A nelle n prove:

$$f_n = \frac{|A_1| + |A_2| + \cdots + |A_n|}{n} .$$

In altre parole, f_n è la frequenza relativa del numero di successi nelle prime n prove.

Definizione 1.21 (frequentista) Si definisce probabilità p dell'evento A il limite della frequenza relativa f_n "a lungo andare", cioè al crescere indefinito del numero n delle prove:

$$P(A) := \lim_{n\to\infty} f_n . \tag{1.5}$$

Secondo la teoria frequentista, infatti, non ha senso parlare di probabilità a proposito di un singolo evento, ma solo a proposito di una collezione di eventi dello stesso tipo.

Anche questa definizione è criticabile in quanto attiene solo ad (un numero infinito di) eventi ripetibili nelle stesse condizioni; inoltre essa non risulta adeguata dal punto di vista matematico. In base alla definizione di limite di successione, si dice che $\{a_n\}_n$ converge a l se:

$$\forall\varepsilon > 0 \quad \exists\,\nu \in \mathbb{N} \mid \text{per ogni } n > \nu \text{ si ha } \; |a_n - l| < \varepsilon . \tag{1.6}$$

In tal caso, notiamo che la definizione (1.5) non rispetta la (1.6) in quanto, trattandosi di eventi aleatori, non è possibile trovare l'indice ν tale che risulti $|f_n - p| < \varepsilon$ per ogni $n > \nu$.

[2] R. von Mises (1930), Veroyatnost i statistika (Probability and statistics), Moscow-Leningrad, (trad. inglese: Academic Press, New York, 1964).

La definizione di von Mises coglie tuttavia un aspetto interessante rispetto a quella di Laplace in quanto qui la valutazione dell'incertezza deriva dall'osservazione ripetuta di un certo fenomeno; tralascia però il fatto che la nostra conoscenza spesso non è l'esito finale di una serie di prove ripetute (per cui non si può dire nulla prima di aver fatto un numero grande di esperimenti), ma al contrario il risultato dei vari esperimenti aggiornano e modificano via via una conoscenza precedente (o una ignoranza iniziale). Al crescere del numero di esperimenti, la nostra conoscenza dipenderà sempre più dai risultati ottenuti piuttosto che dall'ipotesi iniziale.

1.3.3 La teoria assiomatica

L'impostazione assiomatica è dovuta a Kolmogorov[3]. La probabilità è un particolare tipo di misura definita su un'algebra o su una σ-algebra.

Definizione 1.22 (assiomatica) Sia Ω un insieme, $\mathcal{A}$ una σ-algebra di parti di Ω. Si definisce *probabilità* una funzione $P : \mathcal{A} \to \mathbb{R}^+$ tale che:

1. $P(A) \geq 0$ per ogni $A \in \mathcal{A}$;
2. $P(\Omega) = 1$;
3. Se $\{A_n\}$ è una famiglia di elementi di $\mathcal{A}$ a due a due disgiunti, allora risulta:

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)\,. \tag{1.7}$$

La (1.7) è nota come *proprietà di σ-additività*. Si noti che nel caso di famiglie finite di eventi, la proprietà di σ-additività si riduce alla semplice proprietà di additività:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)\,. \tag{1.8}$$

Definizione 1.23 Si definisce *spazio di probabilità* una terna $(\Omega, \mathcal{A}, P)$ dove:
Ω è un insieme;
$\mathcal{A}$ è una σ-algebra di parti di Ω;
P è una probabilità su $\mathcal{A}$.

Gli spazi di probabilità sono modelli matematici di fenomeni reali non deterministici. Questa modellizzazione viene effettuata basandosi su considerazioni empiriche e soggettive, tenendo conto della natura del problema. Ciò vuol dire che, in generale, dato un fenomeno aleatorio, non esiste un unico spazio di probabilità che lo descriva; è anzi possibile che persone diverse scelgano modelli basati su spazi di probabilità differenti.

[3] A. J. Kolmogorov (1933), *Grundbegriffe der Wahrsheinlichkeitsrechung*, Springer Verlag, Berlin (trad. inglese: Chlesea, New York, 1950)

1.3.4 La concezione soggettiva

Secondo l'approccio soggettivo, dovuto principalmente a de Finetti[4], la probabilità è una misura del *grado di fiducia* che un soggetto coerente ripone nel verificarsi di un certo evento A. In tale contesto, il termine "soggettivo" non va inteso come "mancante di oggettività" ma, al contrario, come risultante dal soppesare – con il massimo sforzo di oggettività – ogni elemento di giudizio di cui il soggetto dispone. Pertanto in questo constesto, si parla usualmente di *valutazione di probabilità* anzichè di *stima* come nel caso delle concezioni oggettive.

Un modo di valutare la probabilità di un evento è quello del cosidetto *meccanismo della scommessa*. Dato un numero reale p si dice che si effettua una *scommessa* di quota p su un evento A se, versando una somma pS (dove $S > 0$ è una quantità convenzionale, usualmente si considera S come quantità unitaria), si riceve un *importo* S se A si verifica (scommessa *vinta*) e niente in caso contrario (perdendo la puntata pS).

Nota 1.24 Bisogna sottolineare che lo strumento della scommessa serve solo per capire il concetto di probabilità come grado di fiducia, nel senso che se si ha una fiducia maggiore nel fatto che un certo evento A sia vero o possa verificarsi, allora si è disposti a rischiare di più per guadagnare l'importo S.

Per esempio la "Guide to the Expression of Uncertainty in Measurement" pubblicata a Ginevra nel 1993 e contenente le raccomandazioni del *Comité International des Poids et Mesures* sul modo di esprimere le incertezze sperimentali afferma che "In contrasto con il punto di vista della probabilità basato sulla frequenza , un punto di vista ugualmente valido è quello che sostiene che la probabilità è una misura del grado di fiducia nel verificarsi di un certo evento A. Per esempio, il grado di fiducia in A è uguale a 0.5 se un individuo ritiene indifferente la scelta fra le due scommesse: *i*) ricevere una somma S se A si verifica e niente se A non si verifica; *ii*) ricevere S se A non si verifica e niente se A si verifica. Questa guida adotta questo punto di vista sulla probabilità come il modo appropriato per calcolare l'incertezza sul risultato di una misurazione. La valutazione dell'incertezza non è né un lavoro di routine né una procedura esclusivamente matematica; essa dipende da una conoscenza dettagliata della natura dell'evento da studiare". Si noti infine che alla stesura di questa guida hanno collaborato enti come: *International Organization for Standardization* (ISO), *Bureau International des Poids et Mesures*, *International Electrotechnical Commission*, *International Federation of Clinical Chemistry*, *International Union of Pure and Applied Chemistry*, *International Union of Pure and Applied Physics*, *International Organization of Legal Metrology*.

Una scommessa di quota p (su un evento A *possibile*) da luogo ad un guadagno:

$$\begin{aligned} G_1 &= G(A) = S - pS = (1-p)S && \text{nel caso in cui l'evento } A \text{ si verifica} \\ G_2 &= G(A^c) = -pS && \text{nel caso in cui l'evento } A \text{ non si verifica.} \end{aligned}$$

Il guadagno globale può essere scritto sinteticamente

$$G = (|A| - p)S \tag{1.9}$$

[4]B. de Finetti (1937). La prévision: ses lois logiques, ses sources subjectives, *Annales de l'Institut H. Poincaré*, **7**, 1-68.

dove $G = G_1$ se l'evento A è vero (e quindi $|A| = 1$) e $G = G_2$ se l'evento A è falso (e quindi risulta $|A| = 0$).

Definizione 1.25 Una scommessa si dice *coerente* se *non* esiste alcuna somma $S \neq 0$ tale che le quantità G_1 e G_2 corrispondenti ai due possibili esiti della scommessa (cioè il verificarsi di E o del suo contrario) siano dello stesso segno.

La condizione precedente è del tutto naturale ed ha lo scopo di evitare che sia garantito *a priori* – cioè indipendentemente dall'esito della scommessa – un guadagno (o una perdita). In termini analitici, una scommessa di quota p su un evento A possibile si dice *coerente* se non esiste alcuna somma S diversa da zero ($S \neq 0$) tale che la quantità $G = (|A| - p)S$ assuma lo stesso segno in corrispondenza dei due valori logici A e A^c, cioè se risulta:

$$G_1 G_2 = -pS \cdot (1-p)S = -p(1-p)S^2 \leq 0 \qquad \text{per ogni } S \neq 0 \, . \tag{1.10}$$

In altre parole una scommessa è coerente se non da luogo a guadagni o perdite certe.

Proposizione 1.26 Una scommessa di quota p su un evento qualunque A è coerente se e solo se la quantità p è compresa fra 0 e 1.

Dimostrazione. La tesi segue immediatamente dalla (1.10). Infatti $G_1 G_2 = -p(1-p)S^2$ risulta ≤ 0 se e solo se si ha $0 \leq p \leq 1$. □

In particolare la condizione di coerenza implica che:

- se $A = \emptyset$ allora $p = 0$;
- se $A = \Omega$ allora $p = 1$.

Definizione 1.27 (Definizione soggettiva) Si chiama *probabilità* di un evento A la misura del "grado di fiducia" in A espressa da un numero reale $p = P(A)$ tale che una scommessa di quota p su A sia coerente.

Risulta quindi $0 \leq P(A) \leq 1$ ed in particolare $P(\emptyset) = 0$ e $P(\Omega) = 1$.

Il meccanismo della scommessa come strumento di valutazione della probabilità è generale e può essere utilizzato in qualunque situazione e per qualunque tipo di evento, senza presupporre l'esistenza di una particolare famiglie di eventi $\mathcal{A}$, cui l'evento di interesse A dovrebbe appartenere. La definizione di probabilità nell'approccio soggettivo è di tipo ipotetico-operativo (non è detto che debba scommettere necessariamente) e trova corrispettivi in definizioni nell'ambito di altre discipline quali la fisica, la chimica, l'economia.

Si noti che la definizione soggettiva fornisce un *significato* del numero p indipendentemente dal modo in cui viene valutato.

L'assegnazione di probabilità ad una famiglia $\mathcal{A}$ di eventi in modo tale che la condizione di coerenza sia soddisfatta contemporaneamente per ciascuno di essi origina il concetto di *distribuzione di probabilità*.

Definizione 1.28 Sia $\mathcal{A}$ una famiglia di eventi di Ω. Diremo che sulla famiglia $\mathcal{A}$ è assegnata una *distribuzione di probabilità,* se ad ogni $A \in \mathcal{A}$ viene associato un numero reale $p = P(A)$ ad ogni $A \in \mathcal{A}$, tale che:

1. per ogni $A \in \mathcal{A}$ risulta $0 \leq P(A) \leq 1$;
2. $P(\emptyset) = 0$, $P(\Omega) = 1$;
3. per ogni partizione $A_1, A_2, \ldots$ di Ω contenuta in $\mathcal{A}$ risulta:

$$\sum_{i=1}^{\infty} P(A_i) = P(A_1) + P(A_2) + \cdots + P(A_n) + \cdots = P(\Omega) = 1 \; . \tag{1.11}$$

Si noti che le prime due condizioni fanno riferimento ad eventi singoli, mentre la terza si riferisce alla famiglia $\mathcal{A}$. Si noti in particolare che, nell'approccio assiomatico, la terza condizione costituisce un caso particolare delle (1.7) o (1.8) a seconda che si consideri una partizione di Ω costituita da un numero infinito o finito di eventi.

In particolare la terza proprietà (1.11) segue immediatamente dalla proprietà di σ-additività (1.7) qualora l'insieme $\{A_n\}$ costituisce una partizione dell'evento certo Ω. Nell'impostazione soggettiva la proprietà (1.11) scaturisce dalla proprietà di coerenza solo quando si consideri una partizione finita $\{A_1, A_2, \ldots, A_n\}$ di Ω. Infatti, se $\{A_1, A_2, \ldots, A_n\}$ è una partizione di Ω, allora considerata una qualunque combinazione di scommesse su $A_1, A_2, \ldots, A_n$ di quote rispettivamente $p_1, p_2, \ldots, p_n$, analogamente alla (1.9) il guadagno può esprimersi come segue:

$$G = \sum_{i=1}^{n} (|A_i| - p_i) S_i \; . \tag{1.12}$$

Posto $S_1 = S_2 = \cdots = S_n = 1$ e tenendo conto che – essendo $\{A_1, A_2, \ldots, A_n\}$ una partizione di Ω – solo uno degli eventi deve necessariamente risultare vero e tutti gli altri falsi, segue:

$$G = 1 - (p_1 + p_2 + \cdots + p_n).$$

Poichè il guadagno totale deve essere uguale a zero (per il principio di coerenza), ne segue che $p_1 + p_2 + \cdots + p_n = 1$, cioè la (1.11) per una partizione finita di Ω.

Nel caso di partizioni infinite $\{A_1, A_2, \ldots, A_n, \ldots, \}$ di Ω, la (1.11) non può essere ricavata dal principio di coerenza e bisogna comunque ammettere la proprietà di σ-additività.

Nota 1.29 Nel seguito assumeremo sempre che $\mathcal{A}$ sia un'algebra o una σ-algebra di parti di Ω e, in quest'ultimo caso, che sia sempre soddisfatta l'ipotesi di σ-additività (1.7). Faremo quindi sempre riferimento ad uno spazio di probabilità $(\Omega, \mathcal{A}, P)$.

1.4 Primi risultati del calcolo delle probabilità

Dalle tre proprietà della probabilità viste sopra discendono alcune regole per il calcolo della probabilità di vari eventi.

Teorema 1.30 (probabilità dell'evento contrario) Sia A un evento e sia A^c il suo contrario. Allora risulta:

$$P(A^c) = 1 - P(A) \; .$$

Dimostrazione. Poichè A^c è l'evento contrario di A, allora risulta ovviamente $A \cup A^c = \Omega$. Inoltre, per lo stesso motivo, A e A^c sono incompatibili. Pertanto segue:

$$P(\Omega) = P(A \cup A^c) = P(A) + P(A^c) \ .$$

Poichè $P(\Omega) = 1$, allora si ha $P(A) + P(A^c) = 1$, da cui $P(A^c) = 1 - P(A)$. □

In particolare, da questo risultato segue che $P(\emptyset) = 1 - P(\Omega) = 0$.

Teorema 1.31 Siano A, B due eventi e sia $A \setminus B$ l'evento differenza "A meno B". Allora risulta:

$$P(A \setminus B) = P(A) - P(A \cap B) \ .$$

Dimostrazione. Preliminarmente ricordiamo che si ha $A \setminus B \equiv (A \cap B^c)$. Per la proprietà distributiva dell'intersezione risulta:

$$A = A \cap \Omega = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c) \ . \tag{1.13}$$

I due eventi $(A \cap B)$ e $(A \cap B^c)$ sono inoltre disgiunti, pertanto dalla (1.13) segue:

$$P(A) = P(A \cap B) + P(A \cap B^c) \ . \tag{1.14}$$

da cui la tesi: $P(A \cap B^c) = P(A) - P(A \cap B)$. □
Si noti in particolare che, se $B \subseteq A$, allora $P(A \cap B) = P(B)$ e quindi $P(A \setminus B) = P(A) - P(B)$.

Teorema 1.32 (delle probabilità totali) Siano A, B due eventi. Allora risulta:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \ . \tag{1.15}$$

Dimostrazione. Poichè risulta:

$$A \cup B = \Omega \cap (A \cup B) = (A \cup A^c) \cap (A \cup B) = A \cup (A^c \cap B)$$

e quindi, essendo gli eventi a secondo membro a due a due incompatibili, segue:

$$P(A \cup B) = P[A \cup (A^c \cap B)] = P(A) + P(A^c \cap B) \ . \tag{1.16}$$

Dalla (1.14), scambiando gli eventi A e B, segue $P(A^c \cap B) = P(B) - P(A \cap B)$ e quindi sostituendo nella (1.16), si ha la tesi.

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

□

Si noti che, dal teorema delle probabilità totali, si ottiene immediatamente la relazione:

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

Dal teorema delle probabilità totali segue anche $P(A \cup B) \leq P(A) + P(B)$; più in generale si ha:

$$P\left(\bigcup_n A_n\right) \leq \sum_n P(A_n) \ .$$

In particolare, se A e B sono eventi incompatibili, dalla (1.15), segue:

$$P(A \cup B) = P(A) + P(B) \ . \tag{1.17}$$

Più in generale, se $A_1, \ldots, A_n$ è un insieme di eventi a due a due incompatibili, allora dalla (1.17) segue immediatamente:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) \ . \tag{1.18}$$

In particolare, se $A_1, \ldots, A_n$ costituisce una partizione dell'evento certo, cioè $\cup_{i=1}^n A_i = \Omega$, allora si ritrova la terza proprietà (1.11) delle distribuzioni di probabilità.

Esercizio 1.33 La percentuale degli studenti che hanno superato gli esami A, B e C sono stati le seguenti: A: 50%, B:40%, C:30%; A e B: 35%, A e C: 25%, B e C: 20%; infine il 15% degli studenti ha superato tutti e tre gli esami.

Soluzione. La probabilità che uno studente abbia superato almeno un esame è data da:

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) \\ &- P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \\ &= 0.5 + 0.4 + 0.3 - 0.35 - 0.25 - 0.2 + 0.15 = 0.55 \ . \end{aligned}$$

Calcolare la probabilità che uno studente abbia superato almeno due esami. ♣

Teorema 1.34 (proprietà di monotonia) Siano due eventi A, B con $B \subseteq A$. Allora risulta $P(B) \leq P(A)$.

Dimostrazione. In base al Teorema 1.31, si ha $P(A \backslash B) = P(A \cap B^c) = P(A) - P(A \cap B) \geq 0$. Poichè $B \subseteq A$, allora $P(A \cap B) = P(B)$ e quindi $P(A \setminus B) = P(A) - P(B) \geq 0$ da cui $P(B) \leq P(A)$. □

Esercizio 1.35 Dati n eventi arbitrari $A_1, A_2, \ldots, A_n$, dimostrare per induzione la seguente disuguaglianza (di Bonferroni):

$$P(A_1, A_2, \ldots, A_n) \geq \sum_{i=1}^{n} P(A_i) - (n-1) = 1 - \sum_{i=1}^{n} P(A_i^c) \ .$$

Soluzione. Dimostriamo la disuguaglianza dapprima nel caso $n = 2$; dal teorema delle probabilità totali, si ha:

$$P(A_1, A_2) = P(A_1) + P(A_2) - P(A_1 \cup A_2) \geq P(A_1) + P(A_2) - 1 \ .$$

Inoltre possiamo scrivere (aggiungendo e togliendo 1):

$$P(A_1) + P(A_2) - 1 - 1 + 1 = 1 - [1 - P(A_1) + 1 - P(A_2)] = 1 - P(A_1^c) - P(A_2^c) \ .$$

Ora assumiamo valida la disuguaglianza per $n = k - 1$ e dimostriamola per $n = k$. Essendo:

$$\begin{aligned} P(A_1, \ldots, A_{k-1}, A_k) &= P(A_1, \ldots, A_{k-1}) + P(A_k) - P[(A_1, \ldots, A_{k-1}) \cup A_k] \\ &\geq P(A_1, \ldots, A_{k-1}) + P(A_k) - 1 \end{aligned} \tag{1.19}$$

e, poichè la disuguaglianza si assume vera per $n = k - 1$, risulta:

$$P(A_1, \ldots, A_{k-1}) \geq \sum_{i=1}^{k-1} P(A_i) - (k - 1 - 1)$$

e quindi sostituendo nella (1.19), segue:

$$P(A_1, \ldots, A_{k-1}, A_k) \geq \sum_{i=1}^{k-1} P(A_i) - (k - 2) + P(A_k) - 1 = \sum_{i=1}^{k} P(A_i) - (k - 1) \ .$$

La seconda relazione segue come nel caso precedente:

$$\begin{aligned} \sum_{i=1}^{k} P(A_i) - (k - 1) &= \sum_{i=1}^{k} P(A_i) - (k - 1) - 1 + 1 \ = \ 1 - \sum_{i=1}^{k} [1 - P(A_i)] \\ &= 1 - \sum_{i=1}^{k} P(A_i^c) \ . \end{aligned}$$

□

Esercizio 1.36 Dati tre eventi incompatibili $A, B, C \subset \Omega$ determinare l'insieme delle coppie (α, β) tali che le valutazioni di probabilità $P(A) = \alpha$, $P(B) = \beta$ e $P(C) = 0.2$ siano coerenti.

Soluzione. Essendo gli eventi incompatibili, in base alle ipotesi, risulta:

$$0 \leq P\{A \cup B \cup C\} = P(A) + P(B) + P(C) = \alpha + \beta + 0.2 \leq 1 \ .$$

Pertanto l'insieme I delle valutazioni ammissibili (α, β) è dato da:

$$I = \{(\alpha, \beta) \ : \ 0 \leq \alpha + \beta \leq 0.8\} \ .$$

♣

Esercizio 1.37 Si supponga che un dado sia costruito in modo tale che la probabilità di avere un numero sia proporzionale al numero stesso. Calcolare:

1. La probabilità di ogni singolo numero;
2. la probabilità di avere un numero pari e quella di avere un numero superiore a 4.

Soluzione. Indichiamo con A_i l'evento "esce il numero i", per $i = 1, \dots, 6$. Denotato con p la probabilità che esca la faccia A_1, cioè $P(A_1) = p$, segue:

$$P(A_2) = 2p, \qquad P(A_3) = 3p, \qquad P(A_4) = 4p, \qquad P(A_5) = 5p, \qquad P(A_6) = 6p.$$

Essendo gli eventi $A_1, \dots, A_6$ disgiunti, per la terza condizione della definizione di distribuzione di probabilità, deve risultare:

$$P(A_1 \cup \dots \cup A_6) = P(A_1) + \dots + P(A_6) = 1$$

cioè

$$p + 2p + 3p + 4p + 5p + 6p = 21p = 1$$

da cui $p = 1/21$. Segue pertanto:

$$P(A_2) = 2/21, \quad P(A_3) = 3/21, \quad P(A_4) = 4/21, \quad P(A_5) = 5/21, \quad P(A_6) = 6/21$$

e quindi $P(A_2 \cup A_4 \cup A_6) = 12/21$ e $P(A_5 \cup A_6) = 11/21$. ♣

Esercizio 1.38 Dato un esperimento tale che: $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$ e $P(A \cap B) = \frac{1}{4}$, calcolare:

$$\begin{array}{ccc} P(A \cup B) & P(A^c \cap B) & P(A^c \cup B) \\ P(A \cap B^c) & P(A^c \cap B^c) & P(A^c \cup B^c) \end{array}$$

Soluzione. Da applicazioni delle relazioni trovate in precedenza, segue:

$$\begin{aligned}
P(A \cup B) &= P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{4} = \frac{7}{12} \\
P(A^c \cap B) &= P(B \setminus A) = P(B) - P(A \cap B) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \\
P(A^c \cup B) &= P(A^c) + P(B) - P(A^c \cap B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{12} = \frac{3}{4} \\
P(A \cap B^c) &= P(A \setminus B) = P(A) - P(A \cap B) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4} \\
P(A^c \cap B^c) &= P((A \cup B)^c) = 1 - P(A \cup B) = \frac{5}{12} \\
P(A^c \cup B^c) &= P((A \cap B)^c) = 1 - P(A \cap B) = \frac{3}{4}
\end{aligned}$$

♣

1.5 Valutazioni combinatorie della probabilità

Sia Ω un insieme finito di n eventi elementari, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. Essendo gli eventi ω_i ($i = 1, \ldots, n$) a due a due disgiunti, risulta ovviamente:

$$\begin{aligned} P(\Omega) &= P(\{\omega_1\} \cup \{\omega_2\} \cup \cdots \cup \{\omega_n\}) \\ &= P(\{\omega_1\}) + P(\{\omega_2\}) + \cdots + P(\{\omega_n\}) = 1 \; . \end{aligned} \tag{1.20}$$

Se esistono ragioni di "simmetria" che inducono a ritenere tali eventi *equiprobabili*, cioè:

$$P(\{\omega_i\}) = p \qquad\qquad i = 1, \ldots, n$$

allora, dalla (1.20), si ottiene:

$$P(\{\omega_1\}) + P(\{\omega_2\}) + \cdots + P(\{\omega_n\}) = \underbrace{p + p + \cdots + p}_{n} = np = 1$$

e quindi

$$p = \frac{1}{n} \; .$$

Pertano alla probabilità p di ciascuno di essi, deve necessariamente attribuirsi, in base alla (1.11), il valore $p = 1/n$. Ne segue che se un evento A è l'unione di h degli eventi ω_i:

$$A = \{\omega_{i_1}\} \cup \{\omega_{i_2}\} \cup \cdots \cup \{\omega_{i_h}\}$$

allora la sua probabilità è:

$$\begin{aligned} P(A) &= P(\{\omega_{i_1}\} \cup \{\omega_{i_2}\} \cup \cdots \cup \{\omega_{i_h}\}) = P(\{\omega_{i_1}\}) + P(\{\omega_{i_2}\}) + \cdots + P(\{\omega_{i_h}\}) \\ &= \underbrace{\frac{1}{n} + \cdots + \frac{1}{n}}_{h} = \frac{h}{n} \; , \end{aligned}$$

cioè è uguale al rapporto fra il "numero di casi favorevoli" ed il "numero di casi possibili".

In questo modo, una valutazione *quantitativa* della probabilità si ottiene in base ad una semplice considerazione *qualitativa*: il giudizio di equiprobabilità degli eventi $A_1, A_2, \ldots, A_n$.

Il calcolo del numero di casi favorevoli e di quello dei casi possibili in certi casi è quasi immediato, attraverso ragionamenti di carattere elementare. In molte altre circostanze, invece, il calcolo di tali quantità richiede nozioni di calcolo combinatorio. Nel seguito forniamo pertanto alcune elementari nozioni del calcolo combinatorio.

Definizione 1.39 Sia $S = \{a_1, \ldots, a_n\}$ un insieme di n elementi distinti. Si dice *disposizione di classe k* estratta da S *con ripetizione* una qualunque k-upla ordinata:

$$(a_{i_1}, a_{i_2}, \ldots, a_{i_k})$$

di elementi non necessariamente distinti di S, dove $a_{i_1}, \ldots, a_{i_k} \in S$.

Pertanto due disposizioni di classe k estratte da uno stesso insieme S sono uguali quando in essa figurano gli *stessi elementi* di S nello *stesso ordine*.

Esempio 1.40 Una colonna della schedina del Totocalcio è una disposizione di classe k estratta da $S = \{1, X, 2\}$ (e quindi $n = 3$) con $k = 13$. ♣

Teorema 1.41 Il numero delle disposizioni di classe k da un insieme di n elementi è n^k.

Dimostrazione. Per formare una di queste disposizioni, ciascuno dei k elementi che la compongono può essere scelto in n modi diversi. Perciò si hanno complessivamente $n \cdot n \cdots n = n^k$ modi diversi di costruire tali disposizioni. □

Consideriamo ora una particolare disposizione di classe k da S con la condizione che i suoi elementi (estratti da S) siano tutti *distinti*. In questo caso si parla di *disposizioni semplici* (mentre nel caso precedente, come già evidenziato, si parla di disposizioni con ripetizione).

Teorema 1.42 Il numero delle disposizioni semplici di classe k da un insieme di n elementi (con $k \leq n$) è uguale a:

$$(n)_k := n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!} \ . \tag{1.21}$$

Dimostrazione. A differenza del caso precedente, nel formare una di queste disposizioni, si può scegliere in n modi diversi soltanto il primo dei k elementi che la compongono. Per la scelta del secondo elemento si hanno solamente $n - 1$ possibilità; e così via. Per il k-esimo si hanno a disposizione $n - (k - 1) = n - k + 1$ possibilità. □

Nota 1.43 La scrittura $n!$ si legge "n fattoriale" e denota il prodotto dei primi n numeri naturali:

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

In particolare si pone $0! = 1$.

Esercizio 1.44 Un ufficio dispone di 10 sportelli che forniscono servizi diversi, e arrivano contemporaneamente 7 clienti che si distribuiscono fra i vari sportelli. Definita "coda" un insieme di almeno due clienti allo stesso sportello, quante sono le possibili distribuzioni dei clienti che non danno luogo al formarsi di code? Considerando poi equiprobabili tutte le possibili distribuzioni dei clienti fra i vari sportelli, calcolare la probabilità p che non si formino code.

Soluzione. Il problema viene affrontato utilizzando gli strumenti del calcolo combinatorio. In base alla definizione data, affinchè non vi siano code, è necessario che i 7 clienti si distribuiscano fra 7 dei 10 sportelli presenti nell'ufficio. Pertanto il numero di distribuzioni dei clienti fra i 10 sportelli in maniera tale che non vi siano code è dato dal numero di disposizioni semplici (cioé senza ripetizione) di classe 7 da un insieme di 10 elementi, cioè:

$$(10)_7 = \frac{10!}{(10-7)!} = 604800\,.$$

Assumendo successivamente l'equidistribuzione fra tutte queste distribuzioni, la probabilità che non vi siano code viene valutata in base al rapporto fra in numero di casi favorevoli (cioè il numero di distribuzioni che non comportano code) ed il numero di distribuzioni possibili, che è dato dal numero di tutte le disposizioni (con possibili ripetizione) di classe 7 da un insieme di 10 elementi, che sono 10^7. Pertanto la probabilità richiesta è data da:

$$\frac{604800}{10^7} = 0.0605\,. \tag{1.22}$$

♣

Definizione 1.45 Dato un insieme S di n elementi, si chiama *permutazione* una disposizione semplice di classe n.

Corollario 1.46 Il numero delle permutazioni di un insieme di n elementi è:

$$(n)_n = \frac{n!}{0!} \;=\; n!\,. \tag{1.23}$$

Esempio 1.47 Una parola di lunghezza k è una qualsiasi successione ordinata di lettere dell'alfabeto. Se l'alfabeto ha n lettere, il numero di parole di k lettere è n^k. In particolare, quelle senza lettere ripetute sono $(n)_k$. Gli anagrammi di parole con n lettere distinte (per esempio, con $n = 4$, "rima", "mari", "rami", "armi") sono particolari permutazioni. ♣

Dato un insieme S di n elementi, finora abbiamo considerato gruppi ordinati di elementi estratti (una o più volte) da S. Consideriamo ora gruppi non ordinati di elementi di S, non necessariamente distinti.

Definizione 1.48 Sia $S = \{a_1, \ldots, a_n\}$ un insieme di n elementi distinti. Si chiama *combinazione di classe k* estratta da S un qualunque gruppo

$$\{a_1, a_2, \ldots, a_k\}$$

non ordinato di k elementi estratti da S. Diremo che un dato elemento di S ha *molteplicità* m se esso figura m volte nella combinazione.

Pertanto due combinazioni di classe k estratte da uno stesso insieme S sono uguali quando in essa figurano gli stessi elementi di S con la stessa molteplicità.

Teorema 1.49 Il numero delle combinazioni semplici di classe k estratte da un insieme di n elementi (con $k \leq n$) è

$$\binom{n}{k} = \frac{(n)_k}{(k)_k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!} \tag{1.24}$$

Dimostrazione. Il numero delle disposizioni semplici di classe k di un insieme di n elementi è $(n)_k$: ciascuna di esse costituisce una k-upla ordinata, e considerando tutte le sue permutazioni si ottengono ancora delle disposizioni semplici di classe k, in numero di $k!$ e formate con gli stessi elementi. Esse individuano quindi lo stesso insieme, cioè la stessa combinazione semplice di classe k. Cioè, se consideriamo 2 disposizioni costituite dagli stessi elementi, allora esse differiranno solo per l'ordinamento. In altre parole, ad ogni combinazione corrispondono $k!$ disposizioni semplici, e quindi il numero complessivo $(n)_k$ di queste ultime si ottiene moltiplicando $k!$ per il numero delle combinazioni, e da ciò segue la (1.24). □

Teorema 1.50 Sia assegnata un'urna contenente N palline, di cui r bianche e $N - r$ nere. Se ne estraggano n (simultaneamente o successivamente, ma senza restituzione), supponendo equiprobabili tutte le combinazioni semplici di classe n delle N palline. Allora la probabilità che nella n-upla estratta vi siano k palline bianche è:

$$\frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}} \,. \tag{1.25}$$

Dimostrazione. Una n-upla "favorevole" all'evento "k palline bianche fra le n estratte" si ottiene mettendo k palline, scelte fra le r bianche, insieme ad $n - k$ palline scelte fra le $N - r$ nere. Ogni possibile scelta del primo tipo può essere accoppiata con una qualunque scelta del secondo tipo, e quindi la probabilità richiesta è la (1.25). □

Considerata una popolazione di n elementi, il numero in cui la popolazione può essere suddivisa in k sottoinsiemi contenenti rispettivamente $r_1, r_2, \ldots, r_k$ elementi, con $r_1 + r_2 + \cdots + r_k = n$ e $0 \leq r_i \leq n$ è dato da

$$\binom{n}{r_1, r_2, \ldots, r_k} := \frac{n!}{r_1! r_2! \cdots r_k!} \,.$$

Tale quantità viene chiamate *coefficiente multinomiale*.

Esercizio 1.51 Calcoliamo in quanti modi in cui quattro giocatori di bridge possono ricevere ciascuno 13 carte delle 52 del mazzo. In termini di palline, il gioco del bridge consiste di 52 palline che devono essere poste in 4 scatole tali che in ogni scatola vi siano esattamente 13 palline. Si ha quindi:

$$\binom{52}{13, 13, 13, 13} = \frac{52!}{13!13!13!13!} \approx 2.08 \cdot 10^{48}.$$

Un altro esempio è dato dal numero di anagrammi della parola "matematica" che risulta:

$$\binom{10}{3, 1, 1, 2, 1, 2} = \frac{10!}{3!1!1!2!1!} = 151200 \,.$$

♣

Esercizio 1.52 Gli archivi di un ufficio sono stati memorizzati in 30 dischi da 100 Mbyte, contenenti ognuno 100 file. Un programma dovrà accedere a 28 di questi file (tutti diversi). Qual è la probabilità che esso non debba usare il primo disco?

Soluzione. La situazione in esame si modellizza con uno schema successo-insuccesso senza reimmissione: sono 20 prove ripetute in ognuna delle quali viene scelto un file fra i 3000 possibili. Consiedereremo "successo" la scelta di un file appartenente al primo disco (sono 100) e insuccesso la scelta di uno degli altri dischi (2900 files). Non c'è restituzione perchè un file, una volta letto, non verrà più richiamato in seguito. La probabilità richiesta è dunque

$$p = \frac{\binom{100}{0}\binom{2900}{28}}{\binom{3000}{28}} = 0.385 \; .$$

♣

Esercizio 1.53 Una scatola contiene m palline numerate progressivamente, con $m \geq 5$.

1. Vengono estratte in blocco 5 palline. Qual è la probabilità α che venga estratta la pallina numero 1?

2. Vengono estratte 5 palline ogni volta con reimmissione. Qual è la probabilità β che venga estratta la pallina numero 1?

Soluzione. Indicato con A l'evento di interesse, calcoliamo la probabilità dell'evento complementare, che è più semplice. Infatti l'evento "non esce la pallina n. 1" è composto dalle cinquine che non contengono 1 e sono formate senza reimmissione, essendo l'estrazione fatta in blocco. In quanti modi posso formare una tale cinquina cioè un evento favorevole a A^c? La prima pallina è una qualsiasi delle $m-1$ (escluso 1); la seconda è scelta tra $m-2$ (cioè $m-1$ meno la pallina estratta); la terza fra $m-3$, la quarta fra $m-4$ e l'ultima fra $m-5$. In altre parole,il numero di esiti favorevoli all'evento A^c è uguale al numero di disposizioni semplici di $m-1$ oggetti presi a 5 a 5:

$$(m-1)_5 = (m-1)(m-2)(m-3)(m-4)(m-5)$$

Il numero di esiti possibili è dato dal numero di cinquine senza reimmissione, cioè al numero di disposizioni di m oggetti presi a 5 a 5:

$$(m)_5 = m(m-1)(m-2)(m-3)(m-4) \; .$$

Pertanto:

$$P(A) = 1 - P(A^c) = 1 - \frac{(m-1)_5}{(m)_5} = 1 - \frac{m-5}{m} = \frac{5}{m} \; .$$

Per quanto riguarda il secondo punto, è anche in questo caso più semplice calcolare la probabilità dell'evento complementare. Il numero di eventi favorevoli è – nel caso di estrazioni con reimmissione – uguale al numero di disposizioni con ripetizione di $m-1$ elementi presi a 5 a 5, che è uguale a $(m-1)^5$, mentre il numero di esiti possibili è uguale al numero di disposizioni con ripetizione di m elementi presi a 5 a 5, che è uguale a m^5. Si ha pertanto:

$$P(B) = 1 - P(B^c) = 1 - \frac{(m-1)^5}{m^5} = 1 - \left(1 - \frac{1}{m}\right)^5 .$$

♣

Esercizio 1.54 Un giocatore di poker riceve all'inizio del gioco cinque carte da un normale mazzo di 52.

1. Qual è la probabilità di ricevere almeno 2 assi?
2. Qual è la probabilità di ricevere cinque carte dello stesso seme?

Per quanto riguarda il primo punto, possiamo considerare le 52 carte del mazzo sudivise in due gruppi, uno composto dai 4 assi e l'altro dalle restanti 48 carte. Indichiamo con A_k l'evento "il giocatore riceve k assi fra le sue 5 carte", con $k = 1, 2, 3, 4$. Il problema può essere considerato come un problema di estrazioni di 5 elementi da un'urna contenente 4 elementi di un tipo e 48 dell'altro tipo. In base a quanto visto in precedenza si ha:

$$P(A_k) = \frac{\binom{4}{k}\binom{48}{5-k}}{\binom{52}{5}} .$$

da cui: $P(A_2) = 0.04$, $P(A_3) = 0.0017$ e $P(A_4) = 1.847 \cdot 10^{-5}$.

L'evento "il giocatore riceve almeno due assi" è dato da $A_2 \cup A_3 \cup A_4$ e poichè gli eventi A_2, A_3 e A_4 sono a due a due disgiunti, ne segue $P(A_2 \cup A_3 \cup A_4) = P(A_2) + P(A_3) + P(A_4) = 0,042$.

Per quanto riguarda il secondo punto, indichiamo con $A_\diamondsuit$ l'evento "il giocatore riceve 5 carte di quadri", con $A_\heartsuit$ l'evento "il giocatore riceve 5 carte di cuori", con $A_\clubsuit$ l'evento "il giocatore riceve 5 carte di picche" e con $A_\spadesuit$ l'evento "il giocatore riceve 5 carte di picche". L'evento "il giocatore riceve 5 carte dello stesso seme" è pertanto $A_\diamondsuit \cup A_\heartsuit \cup A_\clubsuit \cup A_\spadesuit$.

Assumendo che le carte siano equiprobabili, ne segue che i 4 eventi $A_\diamondsuit, A_\heartsuit, A_\clubsuit$ e $A_\spadesuit$ sono equiprobabili, cioè $P(A_\diamondsuit) = P(A_\heartsuit) = P(A_\clubsuit) = P(A_\spadesuit)$. Segue pertanto $P(A_\diamondsuit \cup A_\heartsuit \cup A_\clubsuit \cup A_\spadesuit) = P(A_\diamondsuit + A_\heartsuit + A_\clubsuit + A_\spadesuit)$.

Considerato ad esempio l'evento $A_\diamondsuit$ come riferimento, si ha: $P(A_\diamondsuit \cup A_\heartsuit \cup A_\clubsuit \cup A_\spadesuit) = 4P(A_\diamondsuit)$. Calcoliamo quindi $P(A_\diamondsuit)$. Lo schema è analogo al precedente, in questo caso l'urna di 52 elementi si considera suddivisa in un insieme di 13 elementi di un tipo (successi) e 39 elementi dell'altro (insuccessi):

$$P(A_\diamondsuit) = \frac{\binom{13}{5}\binom{39}{0}}{\binom{52}{5}} = 4.95 \cdot 10^{-4} ,$$

e quindi $P(A_\diamondsuit \cup A_\heartsuit \cup A_\clubsuit \cup A_\spadesuit) = P(A_\diamondsuit + A_\heartsuit + A_\clubsuit + A_\spadesuit) = 4P(A_\diamondsuit) = 4 \cdot 4,95 \cdot 10^{-4} = 0,198\%$. $\clubsuit$

1.6 Probabilità su eventi condizionati

Dato un evento E, ad esso corrisponde, come è noto, l'alternativa "VERO o FALSO"; spesso è necessario o interessa considerare certe situazioni solo subordinatamente all'eventuale verificarsi di una data circostanza, la quale può essere espressa a sua volta mediante un evento H. Anzi, in realtà, è questo il punto di vista che andrebbe *sempre* adottato: infatti, anche quando si lancia una moneta e si considera l'evento T(= "testa") o C (= "croce"), si dà tacitamente per scontato che sia sempre vero un altro evento:

$$H = \text{"la moneta mostra una delle due facce"} \ ,$$

cioè si identifica H con l'evento Ω (la moneta potrebbe infatti rotolare in un tombino, oppure restare "in piedi" appoggiata su una parete: anche se è difficile, non è però impossibile – né logicamente né praticamente – che ciò si verifichi).

Definizione 1.55 Siano E e H due eventi di Ω. Si definisce *evento condizionato* E dato H, oppure E condizionato a H, e si indica con $E|H$, l'evento tale che:

- $E|H$ risulta VERO se $E \cap H$ è vero
  (cioè quando essendo vero H, risulta vero anche E),
- $E|H$ risulta FALSO se $E^c \cap H$ è vero
  (cioè quando, essendo vero H, risulta falso E),
- $E|H$ risulta INDETERMINATO se H è falso
  (cioè quando H risulta falso e quindi non interessa, una volta acquisita tale informazione, sapere se E sia vero o falso).

Si noti che, a differenza dell'evento semplice (vedi Definizione 1.2), l'evento condizionato è un ente logico a 3 valori: vero, falso, indeterminato.

Sia $\mathcal{A}$ una famiglia di parti di Ω e sia $H \in \Omega$. Nel seguito denotiamo con $\mathcal{A}|H$ l'insieme $\{E|H \, : \, E \in \mathcal{A}\}$

Per chiarire il significato della condizione $H \neq \emptyset$, si noti che essa equivale (in base alla definizione di evento impossibile) che "non è noto che H sia falso": in base ad una nuova informazione, H può successivamente risultare vero o falso. Si noti inoltre che per $H = \Omega$ si ritrova il concetto di *evento* non condizionato, cioè $E|\Omega = E$. Infatti l'alternativa INDETERMINATO non può presentarsi (poichè Ω è sempre vero), e negli altri due casi la condizione "essendo vero H" è automaticamente soddisfatta. Si noti che l'evento $E|H$ si può scrivere come $(E \cap H)|H$; tale forma si chiama *forma ridotta*.

Come ad un evento E si associa in modo naturale una scommessa (vinta quando E è vero e persa quando E è falso), ad un evento condizionato $E|H$ si può associare un tipo diverso di scommessa, da considerarsi *annullata* se H non si verifica (e vinta se si verificano H e E, persa se si verifica H ma non E). Precisamente il concetto di *scommessa condizionata* su $E|H$, di quota p, generalizza il concetto di scommessa introdotto precedentemente: versando una somma pS (con $S \neq 0$ arbitrario), si riceve un importo S se si verificano H e E, niente se si verifica H ma non E (perdendo così la puntata pS), oppure viene restituita la somma pS se H non si verifica (scommessa annullata). In altre parole, se si punta la somma pS sull'evento $E|H$ si riceve:

- la somma S se $E|H$ vero;
- niente (cioè si perde la somma pS) se $E|H$ falso;
- viene restituita la somma pS se H falso.

Anche per gli eventi condizionati è possibile introdurre un concetto al quale attribuire il significato di "grado di fiducia" nel verificarsi di E condizionatamente a $H \neq \emptyset$ cioè assegnare $P(E|H)$. Più in generale, si consideri una famiglia di eventi $\mathcal{A}$ e un evento $H \neq \emptyset$. Per ogni $E \in \mathcal{A}$, affinchè una scommessa condizionata su $E|H$ sia *coerente*, richiediamo che la distribuzione di probabilità su $\mathcal{A}|H$ verifichi le tre condizioni della Definizione 1.28 di distribuzione di probabilità:

1. per ogni $A \in \mathcal{A}$ risulta $0 \leq P(A|H) \leq 1$;
2. $P(\emptyset|H) = 0$, $P(\Omega|H) = 1$;
3. per ogni partizione $A_1, A_2, \ldots$ di Ω contenuta in $\mathcal{A}$ risulta:

$$\sum_{i=1}^{\infty} P(A_i|H) = P(A_1|H) + P(A_2|H) + \cdots + P(A_n|H) + \cdots = P(\Omega|H) = 1 \ . \quad (1.26)$$

Si noti in particolare che per $H = \Omega$ si ritrovano le tre proprietà della Definizione 1.28.

Nota 1.56 Si noti che l'evento $E|H$ è un oggetto unico, pertanto $P(E|H)$ è la probabilità dell'evento "E condizionato al verificarsi dell'evento H" (e non, come a volte si dice, la probabilità dell'evento E dopo aver appreso che l'evento H è vero). La quantità $P(E|H)$ viene usualmente chiamata *probabilità condizionale* di E dato H.

Assegnati gli eventi E, H e $E|H$, le probabilità di $E \cap H$, H e $E|H$ sono legate dal seguente risultato (di cui si omette la dimostrazione).

Teorema 1.57 (delle probabilità composte) Siano E, H due eventi e si consideri l'evento condizionato $E|H$. Allora risulta:

$$P(E \cap H) = P(H)\, P(E|H) = P(E)P(H|E) \ . \quad (1.27)$$

Nell'impostazione assiomatica, assegnato uno spazio di probabilità $(\Omega, \mathcal{A}, P)$, e considerato $H \in \mathcal{A}$ tale che $P(H) > 0$, si definisce probabilità condizionata la quantità:

$$P(E|H) = \frac{P(E \cap H)}{P(H)} \ . \quad (1.28)$$

Nota 1.58 Si noti che la (1.28) viene spesso assunta quale definizione di probabilità condizionata; in altre parole si definisce la probabilità condizionata a partire dalla (1.28), nell'ipotesi $P(H) > 0$, e si ottiene (banalmente!) il teorema delle probabilità composte come conseguenza. Nell'impostazione soggettiva, l'evento $E|H$ viene assegnato alla sola condizione che $H \neq \emptyset$ e si ricava la (1.27) come conseguenza, che resta valida anche nel caso in cui risulta $P(H) = 0$.

Le tre condizioni della Definizione 1.28 possono essere dimostrate immediatamente, in base al teorema delle probabilità composte, nel caso in cui si assume $P(H) > 0$:

1. $\underline{0 \leq P(A|H) \leq 1}$. Sia $A \in \mathcal{A}$, poichè $A \cap H \subseteq H$, allora per la proprietà di monotonia segue: $0 \leq P(A \cap H) \leq P(H)$ e quindi:
$$P(A|H) = \frac{P(A \cap H)}{P(H)} \leq \frac{P(H)}{P(H)} \leq 1 \; .$$

2. $\underline{P(\emptyset|H) = 0 \text{ e } P(\Omega|H) = 1}$. Si ha infatti:
$$P(\emptyset|H) = \frac{P(\emptyset \cap H)}{P(H)} = \frac{P(\emptyset)}{P(H)} = 0 \; ,$$
$$P(\Omega|H) = \frac{P(\Omega \cap H)}{P(H)} \; = \; \frac{P(H)}{P(H)} \; = \; 1 \; .$$

3. Per ogni partizione $\{A_1, \ldots, A_n, \ldots,\}$ di Ω, con di $A_1, \ldots, A_n, \ldots, \in \mathcal{A}$, deve risultare:
$$P(A_1|H) + P(A_2|H) + \cdots + P(A_n|H) + \cdots = 1 \; .$$
Infatti, poichè $A_1, A_2, \ldots, A_n \ldots$ costituisce una partizione (cioè risulta $A_i \cap A_j = \emptyset$ per $i \neq j$ e $\bigcup_{n=1}^{\infty} A_n = \Omega$) segue:
$$\begin{aligned} P(A_1|H) + \cdots + P(A_n|H) + \cdots &= \frac{P(A_1 \cap H)}{P(H)} + \cdots + \frac{P(A_n \cap H)}{P(H)} + \cdots \\ &= \frac{P[(A_1 \cap H) \cup \cdots \cup (A_n \cap H) \cup \cdots]}{P(H)} \\ &= \frac{P[(A_1 \cup \cdots \cup A_n \cup \cdots) \cap H]}{P(H)} \\ &= \frac{P(H)}{P(H)} = 1 \; . \end{aligned}$$

Nel caso generale, la *probabilità condizionale* di H dati $A_1, \ldots, A_n$ si scrive $P(H|A_1, \ldots, A_n)$ ed è data da:
$$P(H|A_1, \ldots, A_n) \equiv P(H|A_1 \cap \ldots \cap A_n) = \frac{P(H \cap A_1 \cap \ldots \cap A_n)}{P(A_1 \cap \ldots \cap A_n)} \; . \qquad (1.29)$$
Se $P(A_1, \cdots, A_n) > 0$, in base a semplici operazioni algebriche si ottiene:
$$\begin{aligned} P(A_1, \cdots, A_n) &= P(A_1) \frac{P(A_1, A_2)}{P(A_1)} \frac{P(A_1, A_2, A_3)}{P(A_1, A_2)} \cdots \frac{P(A_1, A_2, \ldots, A_n)}{P(A_1, A_2, \ldots, A_{n-1})} \\ &= P(A_1) P(A_2|A_1) P(A_3|A_1, A_2) \cdots P(A_n|A_1, \ldots, A_{n-1}) \; . \qquad (1.30) \end{aligned}$$

1.6.1 Indipendenza stocastica di eventi

Dati due eventi A, B, in generale, si ha $P(B) \neq P(B|A)$. Tuttavia sono molte le situazioni in cui le valutazioni di probabilità sono tali che il rendere noto che l'evento A si sia verificato o meno non modifica l'incertezza sulla verità o falsità di B, cioè la probabilità dell'evento B resta inalterata indipendentemente dal fatto che l'evento A sia vero o falso. Ad esempio, supponiamo di lanciare due volte una moneta: ovviamente la probabilità che al secondo lancio si verifichi "testa" non dipende dal risultato ottenuto nel lancio precedente. In simboli:

$$P(B|A) = P(B) .$$

Da qui nasce il concetto di indipendenza stocastica.

Definizione 1.59 Siano A, B due eventi di Ω. Diremo che l'evento A è *stocasticamente indipendente* dall'evento B se risulta:

$$P(A|B) = P(A) ,$$

e, similmente, l'evento B è indipendente dall'evento A se risulta:

$$P(B|A) = P(B) .$$

In particolare, dalla (1.27), segue che due eventi A e B sono stocasticamente indipendenti se e solo se:

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

Ovviamente se gli eventi A e B sono stocasticamente indipendenti, lo sono anche gli eventi A^c e B, A e B^c, A^c e B^c.

A differenza dell'indipendenza logica (vedi Definizione 1.8), l'indipendenza stocastica dipende non solo dagli eventi, ma anche dalla loro distribuzione di probabilità. Al fine di comprendere meglio questo aspetto consideriamo il seguente esempio.

Esempio 1.60 Sia assegnata un'urna contenente tre palline: una bianca, una rossa ed una nera e supponiamo di effettuare una sola estrazione. Indichiamo con B l'evento "la pallina estratta è bianca", analogamente con R e N gli altri due eventi "la pallina estratta è rossa" e "la pallina estratta è nera". Ovviamente si ha $P(B) = P(N) = P(R) = \frac{1}{3}$. Consideriamo i due eventi:

$$E_1 : B \cup R \quad \text{(la pallina estratta è bianca oppure rossa)}$$
$$E_2 : B \cup N \quad \text{(la pallina estratta è bianca oppure nera);}$$

in questo caso risuta:

$$P(E_1) = \frac{2}{3} \quad P(E_2) = \frac{2}{3} \quad \text{da cui} \quad P(E_1 \cap E_2) = P(B) = \frac{1}{3} \neq P(E_1)P(E_2) = \frac{4}{9}$$

e pertanto gli eventi E_1 e E_2 non sono stocasticamente indipendenti. Supponiamo ora di aggiungere un'altra pallina gialla nella nostra urna (pertanto cambia la composizione in quanto vi è una pallina in più) ma consideriamo lo stesso esperimento di prima, in particolare ora risulta $P(B) = P(N) = P(R) = \frac{1}{4}$. Considerati gli stessi eventi di prima, si ha:

$$P(E_1) = \frac{1}{2} \quad P(E_2) = \frac{1}{2} \quad \text{da cui} \quad P(E_1 \cap E_2) = P(B) = \frac{1}{4} = P(E_1)P(E_2) \,.$$

Si vede pertanto che l'aggiunta di una pallina ha cambiato la distribuzione di probabilità per cui gli eventi E_1 e E_2 risultano ora stocasticamente indipendenti. ♣

Più in generale, n eventi $A_1, \ldots, A_n$ si dicono stocasticamente indipendenti se

$$P(A_{i_1} \cap \cdots \cdots A_{i_k}) = \prod_{j=1}^{k} P(A_{i_j}) \tag{1.31}$$

per ogni sottoinsieme $\{i_1, \ldots, i_k\}$ di $\{1, \ldots, n\}$ o, equivalentemente,

$$P(A_j | A_{i_1}, \cdots, A_{i_k}) = P(A_j)$$

per ogni scelta di indici j e $\{i_1, \ldots, i_k\}$ tale che $j \notin \{i_1, \ldots, i_k\}$.

Si noti che la (1.31) non indica un'unica relazione, ma molte in quanto essa deve essere valida <u>comunque</u> si scelga il numero k di eventi e <u>comunque</u> si scelgano i k eventi nella famiglia $\{A_1, \ldots, A_n\}$. Si tratta pertanto di una condizione molto forte. Ad esempio, se si considerano tre eventi A_1, A_2, A_3, la condizione di indipendenza stocastica (1.31) richiede che siano verificate <u>tutte</u> le relazioni:

$$\begin{array}{ll} P(A_1 \cap A_2) = P(A_1)P(A_2) & P(A_1 \cap A_3) = P(A_1)P(A_3) \\ P(A_2 \cap A_3) = P(A_2)P(A_3) & P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3) \,. \end{array}$$

Esempio 1.61 (segue Esempio 1.60) Con riferimento all'urna contenente quattro palline (vedi Esempio 1.60 precedente): bianca, rossa, nera, gialla, consideriamo ora i tre eventi:

$E_1 : B \cup R$ (la pallina estratta è bianca oppure rossa)
$E_2 : B \cup N$ (la pallina estratta è bianca oppure nera)
$E_3 : B \cup G$ (la pallina estratta è bianca oppure gialla).

Con ragionamento analogo a quello precedente, si vede che gli eventi sono a due a due stocasticamente indipendenti e quindi:

$$P(E_1 \cap E_2) = P(E_1)P(E_2) \quad P(E_1 \cap E_3) = P(E_1)P(E_3) \quad P(E_2 \cap E_3) = P(E_2)P(E_3)$$

invece non risultano a tre a tre stocasticamente indipendenti in quanto:

$$P(E_1 \cap E_2 \cap E_3) = P(B) = \frac{1}{4} \neq P(E_1)P(E_2)P(E_3) = \frac{1}{8}$$

e pertanto gli eventi non sono stocasticamente indipendenti. ♣

Esempio 1.62 (Paradosso di De Méré) E' più probabile ottenere almeno una volta "la faccia 6" lanciando quattro volte un dado, oppure ottenere almeno una volta "un doppio 6" lanciando ventiquattro volte due dadi?

Questo problema viene chiamato "paradosso di De Méré" (in realtà si era già occupato di questo problema, un secolo prima, G. Cardano). De Méré fu un cultore di probabilità del XVII secolo, il quale scrisse al matematico B. Pascal per chiedergli come mai la frequenza relativa di "successo" risultasse sistematicamente" inferiore, nel secondo caso, a quelle del primo. Egli infatti riteneva che i due risultati avessero la stessa probabilità, probabilmente faceva il seguente ragionamento (sbagliato!):

$$\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{4}{6}=\frac{2}{3}\,,$$
$$\frac{1}{36}+\frac{1}{36}+\cdots+\frac{1}{36}=\frac{24}{36}=\frac{2}{3}\,.$$

Il ragionamento corretto conduce alla spiegazione dei risultati osservati. Nel primo caso consideriamo l'evento $A_i =$"esce la faccia 6 nell'i-esimo lancio" per $i = 1, 2, 3, 4$. La probabilità richiesta è:

$$P(A_1 \cup A_2 \cup A_3 \cup A_4)\ .$$

Per la formula di De Morgan si ha $(A_1 \cup A_2 \cup A_3 \cup A_4) = (A_1^c \cap A_2^c \cap A_3^c \cap A_4^c)^c$ e quindi:

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) = P[(A_1^c \cap A_2^c \cap A_3^c \cap A_4^c)^c] = 1 - P(A_1^c \cap A_2^c \cap A_3^c \cap A_4^c)\ .$$

Essendo inoltre gli eventi indipendenti, segue:

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3 \cup A_4) &= 1 - P(A_1^c)P(A_2^c)P(A_3^c)P(A_4^c) \\ &= 1-\left(1-\frac{1}{6}\right)^4 = 1-\left(\frac{5}{6}\right)^4 \approx 0.5177\ . \end{aligned}$$

Nel secondo caso, considerato l'evento $B_i =$"esce un doppio 6 nell'i-esimo lancio" per $i = 1, \ldots, 24$, con ragionamenti analoghi, si ha

$$\begin{aligned} P(B_1 \cup B_2 \cup \cdots \cup B_{24}) &= P[(B_1^c \cap B_2^c \cap \cdots \cap B_{24}^c)^c] = 1 - P(B_1^c \cap B_2^c \cap \cdots \cap B_{24}^c) \\ &= 1 - P(B_1^c)P(B_2^c)\cdots P(B_{24}^c) = 1-\left(1-\frac{1}{36}\right)^{24} \\ &= 1-\left(\frac{35}{36}\right)^{24} \approx 0.4914\ . \end{aligned}$$

♣

Esercizio 1.63 Un'urna contiene 15 palline di cui 3 bianche, 5 nere e 7 rosse.

1. Si estrae una sola pallina: qual è la probabilità che sia rossa? oppure non nera?

2. Si estraggono due palline *con reimmissione*: qual è la probabilità di estrarre due palline bianche? una bianca e una rossa? due non nere? almeno una bianca?

3. Si estraggono due palline *senza reimmissione*: quali sono le probabilità degli eventi descritti al punto 2.?

In base ai dati del problema si ha $\#\Omega = 15$, $\#B = 3$, $\#N = 5$ e $\#R = 7$. Per quanto riguarda il primo punto si ha pertanto:

$$\begin{aligned} P(R) &= \frac{\#R}{\#\Omega} = \frac{7}{15} \\ P(N^c) &= \frac{\#\Omega - \#N}{\#\Omega} = \frac{15-5}{15} = \frac{10}{15} = \frac{2}{3} \,. \end{aligned}$$

Per quanto riguarda il secondo punto indichiamo con B_1 l'evento "estrazione di una pallina bianca alla prima estrazione", con B_2 l'evento "estrazione di una pallina bianca alla seconda estrazione" e così via. L'estrazione *con ripetizione* dà luogo ad eventi indipendenti in quanto la composizione dell'urna rimane la stessa dopo ogni estrazione. Si ha pertanto:

$$\begin{aligned} P(B_1) = P(B_2) = P(B) &= \frac{\#B}{\#\Omega} = \frac{3}{15} = \frac{1}{5} \,, \\ P(N_1) = P(N_2) = P(N) &= \frac{\#N}{\#\Omega} = \frac{5}{15} = \frac{1}{3} \,, \\ P(R_1) = P(R_2) = P(R) &= \frac{\#B}{\#\Omega} = \frac{7}{15} \,. \end{aligned}$$

Segue che:

$$\begin{aligned} P(B_1 \cap B_2) &= P(B_1)P(B_2) = [P(B)]^2 = \left(\frac{3}{15}\right)^2 = \frac{1}{25} \,, \\ P[(B_1 \cap R_2) \cup (R_1 \cap B_2)] &= P(B_1 \cap R_2) + P(R_1 \cap B_2) \\ &= P(B_1)P(R_2) + P(R_1)P(B_2) = 2 \cdot P(B)P(R) \\ &= 2 \cdot \frac{3}{15} \cdot \frac{7}{15} = \frac{42}{225} = \frac{14}{75} \,, \\ P(N_1^c \cap N_2^c) &= P(N_1^c)P(N_2^c) = [P(N^c)]^2 = \left(1 - \frac{5}{15}\right)^2 = \frac{4}{9} \,, \end{aligned}$$

e infine

$$\begin{aligned} P[(B_1 \cap B_2^c) \cup (B_1^c \cap B_2) \cup (B_1 \cap B_2)] &= P(B_1 \cap B_2^c) + P(B_1^c \cap B_2) + P(B_1 \cap B_2) \\ &= P(B_1)P(B_2^c) + P(B_1^c)P(B_2) + P(B_1)P(B_2) \\ &= 2 \cdot P(B) \cdot P(B^c) + [P(B)]^2 \\ &= 2 \cdot \frac{3}{15}\left(1 - \frac{3}{15}\right) + \left(\frac{3}{15}\right)^2 = \frac{81}{225} = \frac{9}{25} \,. \end{aligned}$$

Per quanto riguarda il terzo punto, l'estrazione *senza ripetizione* dà luogo ad eventi non indipendenti in quanto, dopo ogni estrazione, l'urna cambia la sua composizione. In questo caso si ha:

$$\begin{aligned}
P(B_1 \cap B_2) &= P(B_1)P(B_2|B_1) = \frac{3}{15} \cdot \frac{2}{14} = \frac{1}{5} \cdot \frac{1}{7} = \frac{1}{35} \,, \\
P[(B_1 \cap R_2) \cup (R_1 \cap B_2)] &= P(B_1)P(R_2|B_1) + P(R_1)P(B_2|R_1) \\
&= \frac{3}{15} \cdot \frac{7}{14} + \frac{7}{15} \cdot \frac{3}{14} = 2 \cdot \frac{21}{210} = \frac{1}{5} \,, \\
P(N_1^c \cap N_2^c) &= P(N_1^c)P(N_2^c|N_1^c) = \left(1 - \frac{5}{15}\right)\left(1 - \frac{5}{14}\right) = \frac{10}{15} \cdot \frac{9}{14} = \frac{3}{7}
\end{aligned}$$

e infine

$$\begin{aligned}
P[(B_1 \cap B_2^c) \cup (B_1^c \cap B_2) \cup (B_1 \cap B_2)] &= P(B_1)P(B_2^c|B_1) + P(B_1^c)P(B_2|B_1^c) \\
&+ P(B_1)P(B_2|B_1) \\
&= \frac{3}{15}\left(1 - \frac{2}{14}\right) + \left(1 - \frac{3}{15}\right)\frac{3}{14} + \frac{3}{15} \cdot \frac{2}{14} \\
&= \frac{3}{15} \cdot \frac{12}{14} + \frac{12}{15} \cdot \frac{3}{14} + \frac{3}{15} \cdot \frac{2}{14} = \frac{13}{35} \,.
\end{aligned}$$

♣

Esercizio 1.64 Dato $\Omega = \{1, 2, 3\}$, $A = \{1\}$, $B = \{3\}$ e $C = \{2\}$, $P(A) = P(B) = \frac{1}{3}$, calcolare

$$\begin{array}{ccc}
P(C) & P(A \cup B) & P(A^c) \\
P(B \cup C) & P(A^c \cap B^c) & P(A^c \cup B^c)
\end{array}$$

Esercizio 1.65 I componenti prodotti da una certa ditta possono presentare due tipi di difetti, con percentuali rispettivamente del 3% e del 7%. I due tipi di difettosità possono presentarsi in momenti diversi della produzione per cui si può assumere che le presenze dell'uno o dell'altro siano indipendenti fra loro.

1. Qual è la probabilità che un componente presenti entrambi i difetti?
2. Qual è la probabilità che un componente sia difettoso (cioè presenti almeno uno dei due difetti)?
3. Qual è la probabilità che il componente presenti il primo difetto, sapendo che è difettoso?
4. Qualè la probabilità che esso presenti uno solo dei due difetti sapendo che esso è difettoso?

Indichiamo con A e B gli eventi corrispondenti rispettivamente alla presenza del primo e del secondo difetto, cioè $P(A) = 0.03$ e $P(B) = 0,07$; inoltre gli eventi A e B sono indipendenti. La probabilità che entrambi i difetti siano presenti è:

$$P(A \cap B) = P(A)P(B) = 0.03 \cdot 0.07 = 0.0021 \, .$$

L'evento "almeno uno dei due difetti è presente" è $A \cup B$, per cui:

$$P(A \cup B) = P(A) + P(B) + P(A \cap B) = 0.03 + 0.07 - 0.0021 = 0.0979 = 9.79\% \, .$$

L'evento "un pezzo presenta il primo difetto, sapendo che presenta uno dei due difetti" è $A|(A \cup B)$, e pertanto la sua probabilità è data da:

$$P(A|A \cup B) = \frac{P[A \cap (A \cup B)]}{P(A \cup B)} = \frac{P(A)}{P(A \cup B)} = \frac{0.03}{0.0979} = 0.0306 = 30.6\% \, .$$

Infine l'evento "è presente solo un difetto , essendo noto che il pezzo è difettoso" è dato dall'evento complementare a $A \cap B|A \cup B$ e quindi, essendo $(A \cap B) \cap (A \cup B) = (A \cap B)$, segue:

$$1 - P(A \cap B|A \cup B) \;\; = \;\; 1 - \frac{P(A \cap B)}{P(A \cup B)} = 1 - \frac{0.0021}{0.0979} = 0.978 = 97.8\% \, .$$

♣

Esercizio 1.66 (Il paradosso del concorrente) Questo giochino costituisce una versione aggiornata del ben più noto paradosso del carceriere. In un quiz a premi il concorrente viene posto di fronte a tre porte, indicate con A, B e C: dietro due di esse vi è una capra dietro l'altra vi è una Ferrari. Il concorrente effetta la scelta A e successivamente il conduttore apre una porta B dietro cui c'è la capra. Quest'informazione modifica la probabilità che dietro la porta A vi sia la Ferrari?

Soluzione. Indichiamo con A l'evento "dietro la porta A c'è la Ferrari" e con A^c l'evento "dietro la porta A c'è la capra"; analogamente per gli eventi B e C. Consideriamo l'evento:

B^*: il conduttore apre la porta B dietro cui c'è la capra.

Il problema chiede di confrontare $P(A)$ e $P(A|B^*)$: conviene cambiare scelta se e solo se $P(A|B^*) > P(A)$.

Ovviamente, trattandosi di eventi equiprobabili si ha:

$$P(A) = P(B) = P(C) \;\; = \;\; \frac{1}{3} \, .$$

Passiamo al calcolo di $P(A|B^*)$. In base alla relazione della probabilità condizionata possiamo scrivere:

$$P(A|B^*) \;\; = \;\; \frac{P(B^*|A)P(A)}{P(B^*)} \, . \qquad (1.32)$$

Essendo $P(A) = 1/3$, dobbiamo calcolare $P(B^*)$ e $P(B^*|A)$. Osserviamo che, in base alle regole del gioco, l'evento B^* si verifica in uno dei due seguenti casi: *i*) se dietro la porta A c'è la capra (e quindi dietro C c'è la Ferrari): in questo caso il conduttore – dovendo indicare una porta con la capra diversa da A – non può che segnalare la stanza B; *ii*) il concorrente ha indovinato, cioè dietro la porta A c'è la Ferrari: in tal caso il presentatore, potendo scegliere se aprire la porta B o la C, sceglie la C. In simboli:

$$B^* \; = \; C \cup (A \cap B^*)$$

da cui, trattandosi di eventi disgiunti, segue:

$$P(B^*) \; = \; P(A \cap B^*) + P(C) = P(B^*|A)P(A) + P(C) \; . \tag{1.33}$$

A questo punto la questione è la valutazione di $P(B^*|A)$, cioè la probabilità che il presentatore – potendo scegliere indifferentemente fra due porte dietro le quali ci sono le capre – preferisca la porta B. Nel caso in cui egli sceglie con la stessa probabilità una delle due porte si ha:

$$P(B^*|A) \; = \; P(C^*|A) = 1/2 \; , \tag{1.34}$$

e quindi segue:

$$P(B^*) \; = \; P(B^*|A)P(A) + P(C) = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{3} = \frac{1}{2} \; .$$

Da cui sostituendo nella (1.32), si ottiene

$$P(A|B^*) \; = \; \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3} \; .$$

In altre parole, sotto l'ipotesi (1.34), la probabilità che dietro la porta A si nasconda la Ferrari non cambia e quindi è indifferente cambiare la scelta iniziale

<u>*Nota bene.*</u> Si noti che la questione sta nel fatto che NON si verifica l'evento "il conduttore apre a caso una delle altre due porte e vi trova una capra" ma l'evento "il conduttore apre una porta dietro cui c'è una capra".

In generale, non si può dire che l'informazione fornita dal presentatore aumenti la probabilità che dietro la porta A vi sia la Ferrari se non si conosce il criterio con il quale il presentatore sceglie la porta B nell'ipotesi (attenzione!) in cui il concorrente abbia scelto la porta con la Ferrari. Infatti, tenendo conto della (1.33), si ha:

$$P(A|B^*) \; = \; \frac{P(B^*|A)}{1 + P(B^*|A)} \; , \tag{1.35}$$

per cui $P(A|B^*)$ dipende dal criterio di scelta del conduttore:

1. Se $P(B^*|A) = 1/2$ allora $P(A|B^*) = P(A)$ e quindi è indifferente cambiare o meno;

2. Se $P(B^*|A) < 1/2$ allora $P(A|B^*) > P(A)$ e quindi conviene cambiare;

3. Se $P(B^*|A) > 1/2$ allora $P(A|B^*) < P(A)$ e quindi conviene mantenere la scelta iniziale.

Si noti che, dalla (1.35), si vede che $P(A|B^*)$ varia fra 0 e 1/2 in dipendenza del criterio di scelta del presentatore, in particolare risulta $P(A|B^*) = 1/2$ per $P(B^*|A) = 1$ cioè nel caso in cui il presentatore ha un comportamento deterministico per cui, qualora il concorrente avesse scelto la porta con la Ferrari, egli apre sempre la porta B. ♣

Esercizio 1.67 Un'urna contiene N palline di cui pN bianche e $qN = (1-p)N$ nere, con $0 < p < 1$. Si effettua una prima estrazione senza restituzione E_1 di cui si ignora il risultato. Qual è la probabilità che la seconda estrazione E_2 fornisca una pallina bianca?

Soluzione. Il risultato della seconda estrazione dipende dal risultato della prima estrazione E_1 che è incognito. Le probabilità che la prima estrazione E_1 fornisca rispettivamente una pallina bianca $(E_1 = b)$ o una pallina nera $(E_1 = n)$, possono essere valutate in base al rapporto fra casi favorevoli e casi possibili, ottenendo i valori

$$\begin{aligned} P(E_1 = b) &= \frac{pN}{N} = p \\ P(E_1 = n) &= \frac{(1-p)N}{N} = 1 - p \,. \end{aligned}$$

La probabilità che la seconda estrazione fornisca una pallina bianca è allora data da:

$$\begin{aligned} P(E_2 = b) &= P(\{E_1 = b, E_2 = b\} \cup \{E_1 = n, E_2 = b\}) \\ &= P(E_1 = b, E_2 = b) + P(E_1 = n, E_2 = b) \\ &= P(E_1 = b)P(E_2 = b|E_1 = b) + P(E_1 = n)P(E_2 = b|E_1 = n) \,. \end{aligned} \tag{1.36}$$

Se la prima estrazione ha fornito una pallina bianca allora, prima della seconda estrazione E_2, l'urna contiene $N-1$ palline di cui $pN-1$ bianche; in questo caso la probabilità che la seconda estrazione fornisca una pallina bianca è valutata mediante il rapporto fra il variabile di palline bianche ed il variabile totale di palline presenti nell'urna, che fornisce il valore:

$$P(E_2 = b|E_1 = b) = \frac{pN-1}{N-1} \,. \tag{1.37}$$

Analogamente, se la prima estrazione ha fornito una pallina nera allora, prima della seconda estrazione E_2, l'urna contiene $N-1$ palline di cui ancora pN bianche; in questo caso la probabilità che la seconda estrazione fornisca una pallina bianca è data da:

$$P(E_2 = b|E_1 = n) = \frac{pN}{N-1} \,. \tag{1.38}$$

Dalla (1.36), mediante le (1.37) e (1.38), segue la probabilità che la seconda estrazione fornisca una pallina bianca:

$$\begin{aligned} P(E_2) &= P(E_1 = b)P(E_2 = b|E_1 = b) + P(E_1 = n)P(E_2 = b|E_1 = n) \\ &= p\,\frac{pN-1}{N-1} + (1-p)\frac{pN}{N-1} \\ &= \frac{p^2N - p + pN - p^2N}{N-1} \\ &= p\,. \end{aligned}$$

Nota. In generale, iterando tale procedimento, si può dimostrare per induzione che se il risultato delle prime k estrazioni senza restituzioni $E_1, \ldots E_k$ non è noto, la probabilità che la $k+1$-esima estrazione E_{k+1} fornisca una pallina bianca è ancora p. Ciò vale in particolare per la N-esima estrazione E_N. ♣

Esercizio 1.68 Una persona deve colpire con una freccia un bersaglio quadrato di un metro di lato. L'evento A consiste nel colpire la metà destra del bersaglio, e l'evento B consiste nel colpire il triangolo inferiore sinistro. L'area di A, come quella di B, è uguale a 0.5. Calcolare

1. $P(A)$,
2. $P(A \cap B)$,
3. $P(A \cap A^c)$,
4. $P(A^c \cap B)$,
5. $P(A^c \cap B^c)$.

1.7 Il Teorema di Bayes

Il teorema di Bayes costituisce, in un certo senso, una diretta applicazione della probabilità condizionata che consente di poter aggiornare una valutazione di probabilità su un evento H sulla base di nuove informazioni derivanti dall'aver appreso che un certo evento E risulta vero.

Teorema 1.69 (Teorema di Bayes) Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità. Sia $\{H_1, H_2, \ldots, H_n\}$ una partizione di Ω tale che $P(H_i) > 0$ per $i = 1, \ldots, n$ e $E \in \mathcal{A}$ un qualunque altro evento di probabilità positiva, cioè $P(E) > 0$. Allora si ha, per ogni $i = 1, 2, \ldots, n$:

$$P(H_i|E) = \frac{P(E|H_i)\,P(H_i)}{\sum_{j=1}^{n} P(E|H_j)\,P(H_j)}\,. \tag{1.39}$$

Dimostrazione. In base al teorema delle probabilità composte si ha:

$$P(H_i|E) = \frac{P(H_i \cap E)}{P(E)} = \frac{P(H_i)P(E|H_i)}{P(E)} \tag{1.40}$$

Poichè gli eventi $H_1, H_2, \ldots, H_n$ costituiscono una partizione dell'evento certo Ω, risulta:

$$\begin{aligned} E &= E \cap \Omega = E \cap (H_1 \cup H_2 \cup \cdots \cup H_n) \\ &= (E \cap H_1) \cup (E \cap H_2) \cup \cdots \cup (E \cap H_n) = \bigcup_{i=1}^{n} (E \cap H_i) \ . \end{aligned} \tag{1.41}$$

Dal principio delle probabilità totali e dalla (1.27) – o direttamente dalla (1.41) – segue:

$$\begin{aligned} P(E) &= P\left(E \cap \left(\bigcup_{j=1}^{n} H_j\right)\right) = P\left((E \cap H_1) \cup \cdots \cup (E \cap H_n)\right) \\ &= \sum_{j=1}^{n} P(E \cap H_j) = \sum_{j=1}^{n} P(H_j)P(E|H_j) \ . \end{aligned} \tag{1.42}$$

La formula (1.42) viene chiamata *formula di disintegrazione* di $P(E)$ in base alla partizione $\{H_1, \ldots, H_n\}$ e costituisce un risultato spesso utile ai fini del calcolo poichè spesso la probabilità di un evento E è difficile a calcolarsi direttamente, ma risulta agevole determinare $P(E|H_i)$, cioè la probabilità dell'evento E condizionatamente a H_i. Sostituendo la (1.42) nella (1.40), si ottiene la tesi. $\square$

Il *teorema di Bayes* costituisce una delle formule più importanti ed interessanti del calcolo delle probabilità. Nella terminologia statistica $H_1, H_2, \ldots, H_n$ sono dette *cause* o *ipotesi*, le probabilità $P(H_i)$, $i = 1, 2, \ldots, n$ sono dette *probabilità iniziali* (o a priori) di H_i, $P(H_i|E)$ è detta *probabilità finale* (o a posteriori) di H_i; infine la quantità $P(E|H_i)$ viene chiamata *probabilità di verosimiglianza*.

Il teorema di Bayes può essere visto come uno strumento che permette di correggere le informazioni a priori $P(H_i)$ sulla base delle osservazioni sperimentali $P(E|H_i)$, fornendo la probabilità a posteriori $P(H_i|E)$. In questa formula, infatti, si combinano informazioni a priori e verosimiglianze, e quanto più la probabilità a posteriori $P(H_i|E)$ è diversa dalla probabilità a priori $P(H_i)$, tanto più la verosimiglianza ha modificato le informazioni a priori sulle cause H_i.

Nel caso particolare in cui $\Omega = H \cup H^c$ e quindi $H_1 = H$ e $H_2 = H^c$, con $0 < P(H) < 1$, dalla (1.42) segue immediatamente:

$$P(E) \ = \ P(H)P(E|H) + P(H^c)P(E|H^c).$$

Pertanto la (1.39) si scrive

$$P(H_i \mid E) \ = \ \frac{P(E|H_i)\, P(H_i)}{P(H)P(E|H) + P(H^c)P(E|H^c)} \qquad i = 1, 2 \ . \tag{1.43}$$

Esempio 1.70 In una fabbrica, un certo manufatto viene prodotto per il 30% da macchine di tipo H_1, per il 25% da macchine di tipo H_2 e per il rimanente 45% da macchine di tipo H_3. Le macchine H_1 producono il 2% di scarto, le H_2 l'1.2% e le H_3 l'1% di scarto. In una giornata di lavoro le tre macchine producono complessivamente 10.000 pezzi. Un pezzo scelto a caso è difettoso. Calcoliamo la probabilità che esso provenga dalle macchine H_1, H_2 o H_3. Indicato con E l'evento "il pezzo in esame è difettoso", in base ai dati risulta:

$$\begin{aligned} P(E|H_1) &= 0.02 & \qquad P(H_1) &= 0.30 \\ P(E|H_2) &= 0.012 & \qquad P(H_2) &= 0.25 \\ P(E|H_3) &= 0.01 & \qquad P(H_3) &= 0.45. \end{aligned}$$

Il problema richiede il calcolo delle quantità:

$$P(H_1|E) \qquad P(H_2|E) \qquad P(H_3|E) \ .$$

In base al teorema di Bayes, dalla (1.39), risulta: Il problema richiede il calcolo delle quantità:

$$\begin{aligned} P(H_1|E) &= \frac{P(H_1)P(E|H_1)}{P(H_1)P(E|H_1) + P(H_2)P(E|H_2) + P(H_3)P(E|H_3)} \ , \\ P(H_2|E) &= \frac{P(H_2)P(E|H_2)}{P(H_1)P(E|H_1) + P(H_2)P(E|H_2) + P(H_3)P(E|H_3)} \ , \\ P(H_3|E) &= \frac{P(H_3)P(E|H_3)}{P(H_1)P(E|H_1) + P(H_2)P(E|H_2) + P(H_3)P(E|H_3)} \ . \end{aligned}$$

In base ai dati, il denominatore di tali espressioni è dato da:

$$\begin{aligned} P(E) &= P(H_1)P(E|H_1) + P(H_2)P(E|H_2) + P(H_3)P(E|H_3) \\ &= 0.02 \cdot 0.30 + 0.012 \cdot 0.25 + 0.01 \cdot 0.45 = 0.0135 \ . \end{aligned}$$

E quindi, sostituendo i dati in tali espressioni si ottiene:

$$\begin{aligned} P(H_1|E) &= \frac{0.3 \cdot 0.02}{0.0135} = 0.4444 \ , \\ P(H_2|E) &= \frac{0.25 \cdot 0.012}{0.0135} = 0.2222 \ , \\ P(H_3|E) &= \frac{0.45 \cdot 0.01}{0.0135} = 0.3333 \ . \end{aligned}$$

(si noti che le quantità precedenti sono numeri periodici). Nella seguente tabella confrontiamo le probabilità a priori e quelle a posteriori, possiamo quindi vedere come si è modificata la probabilita su ciascuna delle cause in seguito all'osservazione:

Macchine	H_1	H_2	H_3	Somma
Prob. a priori	0.30	0.25	0.45	1
Prob. a posteriori	0.4444	0.2222	0.3333	1

Tale risultato è motivato dal fatto che la macchina H_3 è quella che ha una maggiore probabilità a priori di fornire un pezzo difettoso ed è la macchina che fornisce il maggior numero di pezzi. ♣

Esercizio 1.71 Da un lotto di N componenti, contenente al più un pezzo difettoso, si effettuano senza restituzione $n \leq N$ estrazioni. Considerati gli eventi:

$$\begin{aligned} H &: \text{"il lotto contiene un componente difettoso} \\ E_i &: \text{"l'}i\text{-esimo pezzo estratto è non difettoso" con } i = 1, \dots, n \end{aligned}$$

e posto $P(H) = p$,

1. calcolare la funzione $f(p) = P(H|E_1, \dots, E_n)$
2. determinare, in funzione di N, i valori di n tali che $f(0.7) < 0.4$.

Soluzione. In base al teorema di Bayes si ha:

$$P(H|E_1, \dots, E_n) = \frac{P(E_1, \dots, E_n|H)P(H)}{P(E_1, \dots, E_n|H)P(H) + P(E_1, \dots, E_n|H^c)P(H^c)}$$

$$= \frac{\frac{\binom{N-1}{n}\binom{1}{0}}{\binom{N}{n}} \cdot p}{\frac{\binom{N-1}{n}\binom{1}{0}}{\binom{N}{n}} \cdot p + 1 \cdot (1-p)}$$

Essendo:

$$\frac{\binom{N-1}{n}\binom{1}{0}}{\binom{N}{n}} = \frac{\frac{(N-1)!}{n!(N-1-n)!}}{\frac{N!}{n!(N-n)!}} = \frac{N-n}{N}$$

segue:

$$P(E|H_1, \dots, H_n) = \frac{\frac{N-n}{N} \cdot p}{\frac{N-n}{N} \cdot p + (1-p)} = \frac{(N-n)p}{(N-n)p + N(1-p)} = \frac{(N-n)p}{N-np}$$

e quindi

$$f(p) = \frac{(N-n)p}{N-np} \; .$$

Per quanto riguarda il secondo punto, si cercano i valori di n tali che:

$$\frac{(N-n)\cdot 0.7}{N-0.7n} < 0.4 \; .$$

Svolgendo tale disuguaglianza si trova che i valori di n che soddisfano alla relazione richiesta sono:

$$n > 0.7143N \; .$$

♣

Nota 1.72 Nelle applicazioni pratiche del teorema di Bayes, la forma più semplice (1.43) è spesso quella che risulta più utile:

$$P(E|H) \;=\; \frac{P(H|E)P(E)}{P(H|E)P(E)+P(H|E^c)P(E^c)} \; .$$

In questo contesto, è forse più espressivo mettere in evidenza il rapporto fra le probabilità condizionali all'evento E ed alla sua negazione:

$$\begin{aligned} P(E|H) \;&=\; \frac{P(H|E)P(E)}{P(H|E)P(E)+P(H|E^c)P(E^c)} \;=\; \frac{P(E)}{P(E)+(1-P(E))\dfrac{P(H|E^c)}{P(H|E)}} \\ &=\; \frac{P(E)}{P(E)+\dfrac{1-P(E)}{R}} \end{aligned}$$

dove

$$R \;=\; \frac{P(H|E)}{P(H|E^c)} \; .$$

In particolare notiamo che

$$\begin{array}{rcl} \text{se } R=1 & \text{allora} & P(E|H)=P(E) \\ \text{se } R>1 & \text{allora} & P(E|H)>P(E) \\ \text{se } R<1 & \text{allora} & P(E|H)<P(E) \end{array}$$

cioè il fatto che H si verifichi fa aumentare la fiducia nell'evento E se $R>1$; al contrario, se $R<1$, allora il fatto che H si verifichi fa diminuire la fiducia nel verificarsi di E.

Esempio 1.73 Assumiamo che l'1% degli abitanti di una città sia affetto da una certa malattia. Un nuovo test diagnostico fornisce un risultato positivo nel 97% dei casi in cui l'individuo è affetto da tale malattia, e fornisce un risultato negativo nel 95% dei casi in cui l'individuo è sano. Un individuo scelto a caso risponde positivamente al test: calcolare la probabilità che questi sia malato. Consideriamo i seguenti eventi:

M	=	l'individuo scelto è malato
$T\|M$	=	l'individuo scelto, nel caso in cui sia malato, risponde positivamente al test
$T^c\|M^c$	=	l'individuo scelto, nel caso in cui sia sano, risponde negativamente al test.

Il problema fornisce le seguenti valutazioni di probabilità:

$$\begin{aligned} P(M) &= 0.01 \\ P(T|M) &= 0.97 \\ P(T^c|M^c) &= 0.95 \ , \end{aligned}$$

e richiede di calcolare la probabilità che un individuo sia malato, nell'ipotesi in cui egli risponda positivamente al test, cioè $P(M|T)$. In base al teorema di Bayes, dalla (1.39), si ha:

$$P(M|T) \ = \ \frac{P(M)P(T|M)}{P(M)P(T|M) + P(M^c)P(T|M^c)} \tag{1.44}$$

poichè in base ai dati del problema risulta:

$$\begin{aligned} P(M^c) &= 1 - P(M) = 0.99 \\ P(T|M^c) &= 1 - P(T^c|M^c) = 0.05 \ , \end{aligned}$$

la (1.44) porge:

$$P(M|T) \ = \ \frac{0.01 \cdot 0.97}{0.01 \cdot 0.97 + 0.99 \cdot 0.05} \ = \ 0.1639 \ .$$

Supponiamo ora che il test venge ripetuto, fornendo ancora esito positivo: calcolare la probabilità che l'individuo sia malato. Indicati ora con T_1 e T_2 rispettivamente gli eventi "l'individuo risulta positivo al test nella prima analisi" e "l'individuo risulta positivo al test nella seconda analisi", si richiede la probabilità che l'individuo sia malato, nell'ipotesi in cui due analisi forniscano esito positivo, cioè $P(M|T_1 \cap T_2)$. Ancora per il teorema di Bayes si ha:

$$P(M|T_1 \cap T_2) \ = \ \frac{P(M)P(T_1 \cap T_2|M)}{P(M)P(T_1 \cap T_2|M) + P(M^c)P(T_1 \cap T_2|M^c)} \ . \tag{1.45}$$

E' ragionevole assumere che, nell'ipotesi in cui l'individuo si malato, gli esiti delle due analisi siano eventi indipendenti, cioè $P(T_1 \cap T_2|M) = P(T_1|M)P(T_2|M)$, e che forniscano lo stesso risultato, cioè $P(T_1|M) = P(T_2|M) = P(T|M)$, da cui segue $P(T_1 \cap T_2|M) = [P(T|M)]^2$; analogamente nel caso in cui l'individuo sia sano si ha $P(T_1 \cap T_2|M^c) = P(T_1|M^c)P(T_2|M^c) = [P(T|M^c)]^2$. Ne segue che la (1.45) si scrive

$$\begin{aligned} P(M|T_1 \cap T_2) &= \frac{P(M)P(T_1|M)P(T_2|M)}{P(M)P(T_1|M)P(T_2|M) + P(M^c)P(T_1|M^c)P(T_2|M^c)} \\ &= \frac{0.01 \cdot 0.97^2}{0.01 \cdot 0.97^2 + 0.99 \cdot 0.95^2} \\ &= 0.7917 \ . \end{aligned}$$

Si vede pertanto come, dopo l'esito della prima analisi, non si può ancora ragionevolmente supporre che l'individuo sia malato; al contrario dopo l'esito della seconda analisi, anch'esso positivo, si può ragionevolmente supporre che l'individuo sia malato. ♣

1.8 Altri esercizi

Esercizio 1.74 Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità e siano A, B e C tre eventi di $\mathcal{A}$ tali che: $P(A) = 2/5$, $P(B) = 6/10$, $P(B|A) = 1/2$ e $P(A \cap B \cap C) = 1/10$. Calcolare: *i*) $P(A \cup B)$, *ii*) $P(A^c \cap B^c)$, *iii*) $P(A|B^c)$, *iv*) $P(C|A \cap B)$.

Esercizio 1.75 Sapendo che $P(A|B^c) = 2P(A|B) = 2/123$ e che $P(B) = P(B^c)$, calcolare $P(A)$.

Esercizio 1.76 Siano A, B e C tre eventi stocasticamente indipendenti tale che $P(A) = 1/17$ e $P(B) = 1/5$. Calcolare $P(A \cap B|C)$ e $P(A \cup B)$.

Esercizio 1.77 Un'impresa produce strumenti per il controllo del corretto funzionamento di dispositivi elettronici. Considerati gli eventi:

H = il dispositivo in esame funziona correttamente,

E = lo strumento rileva un funzionamento corretto del dispositivo in esame,

l'azienda pubblicizza l'affidabilità dei propri prodotti con le seguenti specifiche: $P(E|H) = P(E^c|H^c) = 0.95$. Essendo noto che la percentuale dei dispositivi in esame che risultano difettosi è del 5%:

1. Calcolare $P(H|E)$;
2. Calcolare quale deve essere il valore di $P(E|H) = P(E^c|H^c) = p$ affinchè risulti $P(H|E) = 0.9$.

Esercizio 1.78 In un gruppo di n persone, qual è la probabilità che ve ne siano almeno due che festeggiano il compleanno nello stesso mese e giorno dell'anno? Si consideri un anno non bisestile (=365 giorni) e si assuma l'equiprobabilità fra tutte le possibili distribuzioni delle n persone fra i 365 giorni dell'anno. Calcolare successivamente la probabilità richiesta per $n = 10$.

Esercizio 1.79 In un negozio di generi alimentari è arrivata una partita di 100 yogurt tra i quali ve ne sono 10 guasti. Un cliente ne compra 2 e li mangia entrambi. Egli accusa dei dolori allo stomaco (causati naturalmente da almeno uno yogurt guasto). Calcolare la probabilità p che entrambe le confezioni acquistate fossero guaste.

Esercizio 1.80 In una mostra di di dipinti, ci sono 12 opere di cui 10 sono originali. Un visitatore sceglie un quadro a caso e, prima di acquistarlo, chiede l'opinione di un esperto circa l'autenticità del dipinto. L'esperto riconosce correttamente l'opera originale mediamente 9 volte su 10.

1. Assumendo che l'esperto attesta l'autenticità dell'opera, qual è la probabilità che il dipinto sia effettivamente autentico?

2. Se l'esperto valuta che il dipinto non sia autentico, allora il visitatore decide di non acquistare quel dipinto ma un altro scelto a caso. Qual è la probabilità che quest'ultimo sia autentico?

Esercizio 1.81 Si consideri una moneta regolare e si indichino con C e T rispettivamente gli eventi ”croce” e ”testa” nel generico lancio.

a) Si considerino tre lanci della moneta: si calcoli la probabilità che non si presenti mai la successione CT.

b) Si considerino quattro lanci della moneta: si calcoli la probabilità che non si verifichino le successioni CT o TCC.

Esercizio 1.82 Con riferimento ad una data estrazione del lotto, si supponga che Tizia abbia giocato i numeri $\{1, 2, 3, 4, 5\}$. Considerati gli eventi: A = ”Tizio fa esattamente un terno oppure una quaterna” e B = ”‘i numeri 1 e 2 vengono estratti”, calcolare: $P(A)$, $P(B)$ e $P(A|B)$.

Esercizio 1.83 Il signor Giacometti, per decidere se andare in vacanza per il fine settimana, ascolta le previsioni del tempo su due differenti emittenti televisive, Tele-A e Tele-B. E' noto che Tele-A è affidabile al 70% mentre l'altra lo è al 50%. Il signor Giacometti si affida usualmente il 60% delle volte alle previsioni di Tele-A e per il restante 40% segue quelle di Tele-B. Sapendo che, andando in vacanza per il fine settimana, è stato sorpreso dal maltempo, calcolare la probabilità che il signor Giacometti si sia affidato alle previsioni dell'emittente Tele-A.

Esercizio 1.84 Siano E_1, E_2 due eventi aventi probabilità rispettivamente uguale a $P(E_1) = 0.4$ e $P(E_2) = 0.8$. Stabilire se le due valutazioni $P(E_1 \cup E_2) = 0.7$ e $P(E_1 \cap E_2) = 0.3$ sono coerenti; in caso affermativo calcolare i valori ammissibili per $P(E_1 \cup E_2)$ e $P(E_1 \cap E_2)$.