Probabilità e incertezza di misura

G. D'Agostini
Dipartimento di Fisica, Università "La Sapienza", Roma

30 ottobre 2001

Indice

Parte I

Dal concetto di probabilità ai problemi di probabilità inversa

Capitolo 1

Incertezza e probabilità

Se noi non fossimo ignoranti non ci sarebbe probabilità, ci potrebbero essere solo certezze. Ma la nostra ignoranza non può essere assoluta, altrimenti non ci sarebbe più probabilità. Così i problemi di probabilità possono essere classificati a seconda della maggiore o minore profondità della nostra ignoranza.

(H. Poincaré)

1.1 Determinismo e probabilismo nei metodi di indagine scientifica

La meccanica classica insegna che se si conoscono le proprietà di un corpo (massa, forma, etc.), le sue condizioni iniziali di moto (posizione, velocità etc.) e le condizioni esterne (campi di forze, condizioni al contorno, etc.) è possibile determinarne in modo esatto il suo comportamento negli istanti successivi. Ciò nonostante, è sufficiente analizzare il semplice esperimento del lancio di una moneta o di un dado per capire che non è affatto facile prevedere il risultato, cioè l'occorrenza rispettivamente di testa/croce o di un certo numero impresso nella faccia superiore. Un caso simile si verifica quando si tenta di trattare un sistema costituito da un gas di cui si voglia descrivere l'evoluzione partendo dai moti delle singole molecole.

Questo non implica che le leggi della meccanica classica siano necessariamente violate. Anche supponendo che il tipo di processo sia tale che, istante per istante, gli oggetti obbediscano alle leggi di Newton in ogni interazione con l'ambiente, il numero di parametri che determinano il moto rende proibitivo sia calcolare esattamente il risultato a partire dalle condizioni iniziali che cercare di riprodurre le stesse condizioni ad ogni lancio.

Diventa allora più opportuno *cambiare metodo* scientifico. Si preferisce rinunciare a predire l'esatto esito dell'esperimento e limitarsi ad esprimere affermazioni sulla plausibilità di ciascuno dei risultati possibili. In altri termini, sebbene si seguiti ad assumere che le leggi "di base" siano di tipo deterministico, il nostro *stato di incertezza* riguardo i dettagli del moto ci impedisce di arrivare a conclusioni *certe*.

Il grado di incertezza può essere ridotto se le leggi della fisica che governano il processo sono ben note e se le condizioni sperimentali sono sotto controllo. Questo significa che alcuni esiti (o intervalli di possibili valori) saranno ritenuti molto plausibili, mentre altri saranno considerati inverosimili, se non addirittura impossibili. In altri termini, si parla genericamente della *probabilità* di un possibile risultato di un certo esperimento.

Ci sono poi fenomeni in cui è la stessa meccanica classica a non essere adeguata alla descrizione del processo elementare. Questo succede sulle scale atomiche e inferiori. In questo caso sono le stesse leggi fondamentali che assumono natura aleatoria; si perde il carattere deterministico "almeno in linea di principio" della meccanica classica e si deve utilizzare il linguaggio della *meccanica quantistica*. È da notare comunque che, sebbene le due situazioni siano sostanzialmente differenti dal punto di vista fisico, esse sono simili dal punto di vista *conoscitivo*: in entrambe siamo in stato di incertezza rispetto ai possibili esiti, anche se nel caso quantistico c'è la convinzione che, pur partendo da un ben preciso stato di preparazione del sistema iniziale, l'evoluzione sia intrinsecamente probabilistica. Le diversità delle problematiche si riflette sia sul modo di intendere le leggi fisiche che sui metodi usati per valutare la probabilità dei possibili esiti. Identico sarà invece il nostro atteggiamento psicologico nei riguardi del risultato: un valore "molto elevato" della probabilità incoraggia le nostre aspettative su un certo esito, indipendentemente che si tratti del decadimento di una particella subnucleare o di lanci di dadi[1].

1.2 Incertezze in Fisica e nelle altre scienze naturali

Il concetto di probabilità - basti per ora il significato intuitivo che si attribuisce al termine - non interviene soltanto nel considerare i possibili esiti di un esperimento. Un aspetto ancora più importante è quello che riguarda le conclusioni scientifiche che seguono dalle osservazioni sperimentali, vale a dire quali ipotesi sono supportate o escluse dai dati sperimentali. Infatti, anche se comunemente si parla di "certezze" scientifiche, gli addetti ai lavori sanno bene che di certezze dimostrate con lo stesso rigore di un teorema di matematica ce ne sono ben poche, anzi, ad essere precisi, non ce n'è nessuna.

Cerchiamo di capire quali sono le ragioni di incertezza nella scienza. La figura 1.1 schematizza l'attività del fisico o degli altri ricercatori. Dai dati sperimentali si cerca di determinare il valore di una certa grandezza o di stabilire quale teoria descriva meglio i fenomeni osservati. In realtà entrambi i processi possono essere visti come due aspetti dello stesso problema: come passare dalle osservazioni alle ipotesi. Infatti i due problemi possono essere riformulati nei seguenti modi:

[1] ❄Questo approccio è compatibile sia con la corrente di pensiero che crede che la la meccanica quantistica esaurisca la descrizione del mondo fisico sia con quella che protende per le famose "variabili nascoste", secondo la visione di Einstein. Questo concetto è stato espresso molto chiaramente già nel 1748 da Hume, il quale, pur credendo che le leggi della natura dovessero essere deterministiche, affermava: *"Per quanto non vi sia nel mondo qualche cosa come il Caso, la nostra ignoranza della causa reale di ogni evento ha la stessa influenza sull'intelletto e genera una simile specie di credenza"*

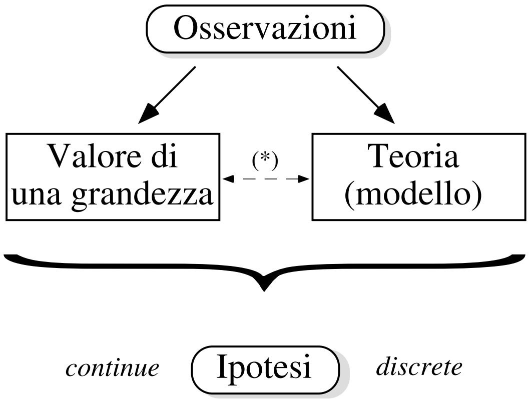

Figura 1.1: Dalle osservazioni alle ipotesi. La relazione fra valore delle grandezza e teoria sta ad indicare che in genere le grandezze hanno significato soltanto all'interno di una teoria o un modello.

A quali valori sono (più) compatibili con la definizione della grandezza oggetto della misura, avendo letto certi numeri sugli strumenti (e subordinatamente a tutte le conoscenze sugli strumenti e della grandezza in questione)?

B quale teoria è (più) compatibile con i fenomeni osservati (e subordinatamente alla credibilità della teoria basata su argomenti formali, estetici e di semplicità[2])?

La sola differenza fra i due processi di apprendimento è che, mentre nel secondo caso si ha a che fare generalmente con un piccolo numero di ipotesi, nel primo caso il numero di ipotesi è virtualmente infinito (le grandezze assumono i valori numerici con continuità, almeno in linea di principio).

Il motivo per cui non si arriva mai alle condizioni ideali di certezza, ovvero tali che soltanto una delle tante (o infinite) ipotesi sia da ritenersi *vera* e tutte le altre *false*, può essere compreso analizzando lo schema che segue.

A: Per quanto riguarda la determinazione del valore di una grandezza si dice comunemente che l'incertezza sia dovuta ad inevitabili *errori di misura* (fluttuazioni della risposta dovute a 'rumore', imperfetta calibrazione degli strumenti, effetti ambientali non perfettamente controllati, etc.);

B: Quando si tratta di una teoria possiamo distinguere due casi:

(B_1) La legge è probabilistica, ovvero *"le osservazioni non sono una mera conseguenza logica della teoria"*. Un classico esempio è quello della genetica. Un esempio più semplice è quello del lancio di una moneta. Anticipando un risultato del calcolo delle probabilità - peraltro molto noto - si ha che, se la moneta è regolare, le due sequenze di testa (T) e croce (C)

$$\texttt{TTTTTTTTTTTTTTTTTTTTTTTTT} \tag{1.1}$$

[2]Si vedrà come queste postille di subordinazione delle conclusioni scientifiche a conoscenze e "pregiudizi" a priori giocano un ruolo fondamentale nei processi di misura e di accettazione di teorie da parte della comunità scientifica.

$$\texttt{TTCCCCCTTTTCTTCCTTTCTCTTC} \qquad (1.2)$$

sono ritenute ugualmente probabili. Quindi sarà impossibile arrivare a conclusioni certe sulla regolarità di una moneta ignota pur avendo osservato una sequenza di lunghezza arbitraria[3];

(B_2) La legge è determinististica. Questa classificazione è valida solo in principio. Infatti, in tutti i casi, *"le osservazioni dipendono anche da molti altri fattori esterni alla teoria"*, siano essi condizioni iniziali e ambientali, errori sperimentali, e così via. Tutte le incertezze su questi fattori rendono la relazione teoria-osservazione anche in questo caso di tipo probabilistico.

1.3 Limiti all'accuratezza delle misure - un esempio

Come esempio dell'impossibilità di arrivare a "misure perfette", consideriamo il semplice caso in cui si voglia determinare la lunghezza di un certo oggetto. Potrebbe essere ad esempio il lato di un foglio di metallo dalla forma rettangolare. La definizione operativa di misura implica il confronto fra l'oggetto da misurare e un campione del metro. Da tale confronto potrà risultare che la lunghezza l dell'oggetto sia inferiore al metro, vale a dire $0 < l < 1\,\mathrm{m}$. Supponiamo di ottenere, dal confronto dell'oggetto con i vari sottomultipli del metro, la seguente successione di risultati:

$$\begin{array}{rcl} 0.2\,\mathrm{m} & < l < & 0.3\,\mathrm{m}\,, \\ 0.24\,\mathrm{m} & < l < & 0.25\,\mathrm{m}\,, \\ 0.247\,\mathrm{m} & < l < & 0.248\,\mathrm{m}\,, \\ 0.2473\,\mathrm{m} & < l < & 0.2474\,\mathrm{m}\,, \end{array}$$

avendo assunto di poter interpolare, eventualmente con l'ausilio di qualche strumento ottico o meccanico, fra tacche contigue distanziate di un millimetro.

Siamo interessati a capire fino a che punto possiamo andare avanti con questo procedimento. È facile convincersi che non potremo mai giungere a determinare la lunghezza di interesse come il numero reale "elemento di separazione fra due classi contigue".

A mano a mano che tentiamo di determinare meglio la quantità di interesse incontriamo nuovi problemi: inizialmente sarà la rugosità delle superfici; poi l'effetto della dilatazione termica; poi ancora la lunghezza d'onda finita della luce con cui si illumina l'oggetto; si arriva alla fine - anche assumendo di poter utilizzare un "microscopio ideale" per il confronto - a problemi legati alla natura non continua della materia: dove finisce il foglio e dove comincia il non-foglio? Ma prima ancora di arrivare a questi limiti concettuali sorge il dubbio se veramente il nostro strumento di misura sia "lungo un metro", ovvero si dovrà affrontare il problema della riproducibilità e costanza del campione di misura. Non è difficile convincersi che il meglio a cui si potrà giungere è affermare che la lunghezza di interesse è compresa fra due valori

$$l_{min} < l < l_{max}\,.$$

[3] Ma dopo l'osservazione della prima sequenza mostrata aumenta il sospetto che si tratti di una moneta con due teste, qualora ci siano delle buone ragioni per far sorgere un simile dubbio. Il concetto di probabilità servirà a quantificare il grado di tale sospetto.

A questo punto sorgono immediate delle domande:

- Quale significato dobbiamo attribuire all'espressione "essere compresa"?

- Se si progetta un esperimento molto più preciso (ovvero dal quale ci attende un intervallo $[l_{min}, l_{max}]$ molto più piccolo) quale risultato ci si dovrà aspettare?

- Se un secondo esperimento trova un intervallo

  $$l'_{min} < l < l'_{max} .$$

  diverso da quello del primo esperimento ed eventualmente in disaccordo da questo (ad esempio $l'_{max} < l_{min}$) come ci si deve regolare? A quale dei due credere? Si possono eventualmente combinare le informazioni di entrambi gli esperimenti per aumentare il nostro grado di conoscenza sulla grandezza fisica di interesse?

- Supponendo di aver effettuato anche una misura di tempo, il cui risultato è anche in questo caso $t_{min} < t < t_{max}$, e di voler utilizzare la combinazione delle due misure per stimare la velocità media v di un oggetto che ha percorso nel tempo t la distanza l, come valutare l'intervallo entro cui è compreso il valore della velocità?

Concludiamo questa discussione sull'incertezza di misura, mostrando in tabella 1.3 come si sia evoluta nel tempo la conoscenza della velocità c della luce nel vuoto da Galileo ai nostri giorni.

Si può notare la diminuzione nel corso degli anni dell'errore, inteso come la differenza fra il risultato della misura e quello "vero". Esso è indicato con e nella tabella. (Si ricorda che attualmente il valore della velocità della luce è assunto essere esatto, in quanto esso può essere riprodotto meglio di quanto non sia possibile riprodurre il metro. Quindi è la distanza ad essere grandezza derivata da velocità e tempo.)

Si nota nella tabella che, partire dalla metà del diciannovesimo secolo,[4] le misure di velocità della luce sono accompagnate da una stima quantitativa dell'incertezza (indicata con u nella tabella 1.3 e non meglio definita per il momento se non qualitativamente come "intervallo entro cui si crede ragionevolmente si trovi il valore della grandezza"). Il rapporto e/u, ovvero dell'errore di misura in unità di incertezza stimata, fornisce un'idea della bontà di stima dell'incertezza stessa. La tabella mostra come il valore attuale della velocità della luce differisce al più di qualche unità di u dai valori misurati.

1.4 Imparare dagli esperimenti: il problema dell'induzione

Ogni misura ha lo scopo di accrescere la conoscenza di chi la esegue e di chi ha interesse a quella specifica conoscenza. Questi possono essere una certa

[4]L'interesse alla stima quantitativa dell'incertezza di misura può essere fatto risalire al lavoro di Laplace e Gauss, all'inizio del 1800.

Anno	Sperimentatore (*metodo*)	c (km/s)	u (km/s)	e (km/s)	e/u
$\approx$ 1600	Galileo (*misure manuali*)	∞?	-	-	-
1676	Roemer (*satelliti di Giove*)	214000	-	-86000	-
1729	Bradley (*aberrazione*) (*posizioni stellari*)	304000	-	+4000	-
1849	Fizeau (*ruota dentata*)	315300	-	+15300	-
1862	Foucault (*specchio ruotante*)	298000	500	-1800	-3.6
1879	Michelson (*specchio ruotante*)	299910	50	+118	+2.4
1906	Rosa & Dorsey ($c = 1/\sqrt{\mu_\circ \epsilon_\circ}$)	299781	10	-11	-1.1
1927	Michelson (*specchio ruotante*)	299798	4	+5.5	+1.4
1950	Essen (*cavità a microonde*)	299792.5	3.0	+0.04	+0.01
1950	Bergstrand (*geodimetro*)	299793.1	0.25	+0.64	+2.6
1958	Froome (*interferometro*) (*a microonde*)	299792.5	0.1	+0.04	+0.4
1965	Kolibuyev (*geodimetro*)	299792.60	0.06	+0.14	+2.3
1972	Bay et al. (*da* $c = \lambda\nu$*; laser*)	299792.462	0.018	+0.004	+0.2
1973	Evenson et al. (*da* $c = \lambda\nu$*; laser*)	299792.4574	0.0012	-0.0006	-0.5
1974	Blaney et al. (*da* $c = \lambda\nu$*; laser*)	299792.4590	0.0008	+0.0010	+1.25
1983	B.I.P.M. (*assunto esatto*)	299792.458	0	0	-

Tabella 1.1: Determinazioni della velocità della luce: u ed e rappresentano rispettivamente l'incertezza dichiarata dallo sperimentatore e la differenza ("errore") rispetto al valore nominale di 299˙792˙458 m/s assunto esatto dal Bureau Interational des Poids et Mésures.

comunità scientifica, un medico che ha prescritto una certa analisi o un commerciante che deve acquistare un prodotto. È anche chiaro che la necessità stessa di eseguire misure indica che ci si trovava in uno stato di incertezza su qualcosa di interesse. Questo “qualcosa” può essere una costante fisica o una teoria sull’origine dell’universo; lo stato di salute di un paziente; la composizione chimica di un nuovo prodotto. In tutti i casi la misura ha lo scopo di *modificare* un certo stato di conoscenza.

Si sarebbe tentati di dire addirittura “acquisire”, anziché “modificare”, lo stato di conoscenza, come ad indicare che la conoscenza si possa creare dal nulla nell’atto della misura. Non è difficile convincersi che nella maggior parte dei casi si tratta invece soltanto di un aggiornamento alla luce di fatti nuovi e di un certo raziocinio. Prendiamo ad esempio la misura della temperatura di una stanza effettuata con un termometro digitale - tanto per escludere contribuiti soggettivi alla lettura dello strumento - e supponiamo di ottenere 21.7 °C. Anche se si potrà dubitare del decimo di grado, indubbiamente la misura è servita a restringere l’intervallo di temperature ritenute plausibili prima della misura - quelle compatibili con la sensazione di “ambiente confortevole”. In base alla conoscenza del termometro, o dei termometri in generale, ci saranno valori di temperatura in un certo intervallo intorno a 21.7°C ai quali crediamo di più e valori al di fuori ai quali crediamo di meno.

È però altresì chiaro che se il termometro avesse indicato, a parità di sensazione fisiologica, 17.3 °C si sarebbe tentati a ritenere che il termometro non funzioni bene. Non si avrebbero invece dubbi sul suo malfunzionamento se esso avesse indicato 2.5 °C!

I tre casi corrispondono a tre diversi gradi di aggiornamento della della conoscenza. Nell’ultimo, in particolare, l’aggiornamento[5] è nullo.

Questi esempi indicano che lo stesso concetto di probabilità con cui classificavamo le previsioni dei risultati di un esperimento intervengono nella valutazione dei possibili valori di una grandezza fisica o di ipotesi scientifiche alternative. Lo scopo delle misure è - ripetiamo - quello di modificare tali probabilità.

Il processo di apprendimento dalle osservazioni sperimentali è chiamato dai filosofi *induzione*. Probabilmente a molti lettori sarà anche noto che in filosofia esiste l’irrisolto “problema dell’induzione” dovuto alla critica di Hume a tale processo. Essa può essere sintetizzata affermando che l’induzione non è giustificata, nel senso che è impossibile dimostrare che da certe osservazioni possano seguire *necessariamente* determinate conclusioni scientifiche. L’approccio probabilistico che abbiamo appena abbozzato, e che risulterà più chiaro nel seguito, sembra essere l’unica via d’uscita a tale critica.

1.5 ❄Limiti del metodo di falsificazione

Molto spesso si pensa che l’unico metodo scientifico valido sia quello della falsificazione. Non ci sono dubbi che, se una teoria non è in grado di descrivere i risultati di un esperimento, essa vada scartata o modificata. Ma poiché non è

[5] Ma anche in questo caso si è imparato qualcosa, cioè che il termometro non funziona...

possibile dimostrare la certezza di una teoria diventa impossibile decidere fra tutte le (infinite) ipotesi non falsificate.

Il metodo probabilistico permette invece di fornire una scala di credibilità a tutte le ipotesi considerate (o rapporti di credibilità fra ogni coppia di ipotesi).

Anche se questi ragionamenti sembrano più adatti a speculazioni di filosofia della scienza, essi giocano un ruolo fondamentale nella teoria dell'incertezza di misura che sarà sviluppata nel seguito. In particolare, il criterio di falsificazione sarà recuperato come semplice sottocaso.

1.6 Decisioni in condizioni di incertezza

L'esperienza quotidiana insegna come il concetto di probabilità non sia confinato soltanto al ragionamento scientifico. Il termine probabilità è largamento usato - talvolta anche a sproposito - nel linguaggio quotidiano, in quello dei politici, nel mondo finanziario e dai mass media. Si parla per esempio di:

- probabilità che una macchina venga rubata;
- probabilità che un televisore si rompa prima dello scadere della garanzia;
- probabilità che una squadra vinca un incontro di calcio;
- probabilità che si risolva una crisi politica o un conflitto;
- probabilità che un titolo azionario aumenti di almeno il $10\,\%$ entro i prossimi sei mesi;
- probabilità che un nuovo prodotto riscuota il successo del mercato.

Il motivo del continuo uso del termine probabilità è dovuto al fatto che nella vita quotidiana e nelle situazioni professionali sono rare le occasioni in cui si agisce in stato di assoluta certezza (sotto molti aspetti è anche meglio: si immagini che noia nel caso opposto!). Ciò nonostante bisogna prendere in continuazione decisioni in stato di incertezza. Da quelle banali, come uscire o no con l'ombrello o quale marca di pelati preferire al supermercato, a quelle "serie", come accettare un posto di lavoro in un'altra città, sposarsi, o sottoporsi ad un intervento chirurgico.

Tutte queste decisioni, in mancanza della certezza sul loro esito, sono prese in base a considerazioni utilitaristiche (in senso lato), soppesando costi e benefici con le probabilità dei vari esiti.

1.7 Concetto di probabilità

Come si vede dagli esempi appena incontrati, la categoria del *probabile* viene introdotta nelle argomentazioni quando viene meno la categoria del *certo*. L'origine stessa della parola (dal latino *probàbilis*, da *probàre*, provare) sta ad indicare "degno di approvazione", "verosimile", "accettabile", "credibile", "ammissibile in base ad argomenti abbastanza sicuri". Essa contrasta con *probatus* - provato - riferito ad affermazioni per le quali è accertato il contenuto di verità (*vero o falso*).

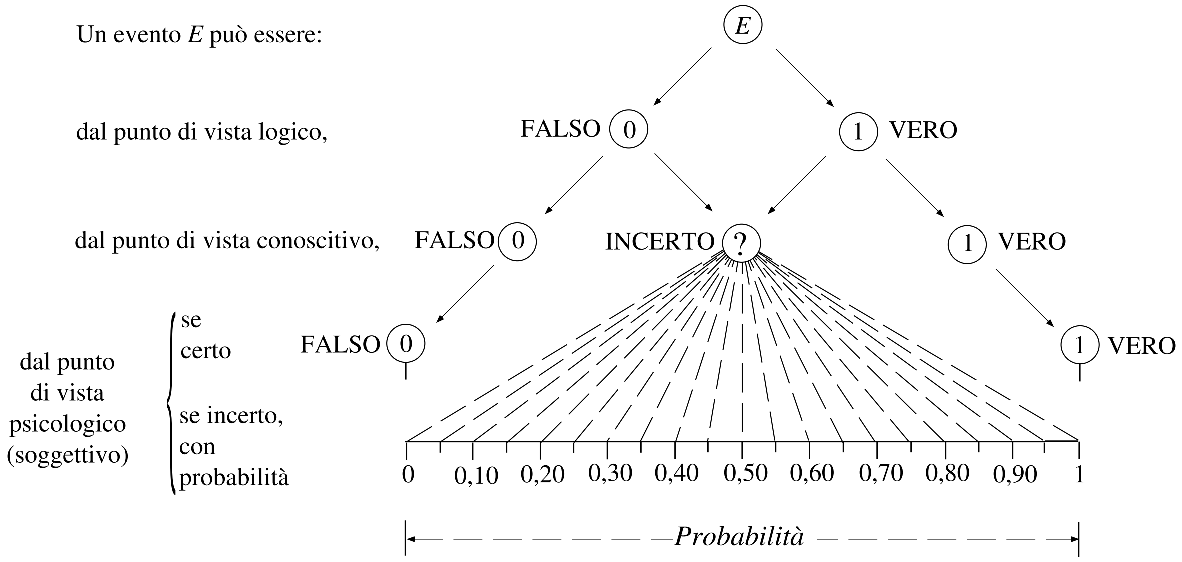

Figura 1.2: Vero, falso, probabile.

Di fronte a diverse affermazioni si può parlare di alcune *più* probabili e di altre *meno* probabili, a seconda della loro plausibilità. Il termine *probabilità* viene quindi usato come misura del grado di plausibilità di una affermazione, ovvero del “verificarsi di un certo evento”.

È da notare come anche nel linguaggio parlato si cerchi di fissare una “scala” al livello di verosimiglianza, confrontando la propria valutazione di probabilità con quella di altri eventi sui quali ci sia unanime consenso nel ritenerli più o meno possibili (si pensi, ad esempio, ad espressioni del tipo “è come vincere un terno al lotto”).

Il diagramma di figura 1.2 sintetizza molto bene lo schema logico che porta al concetto di probabilità:

- innanzitutto c’è da premettere che con il termine *evento* intenderemo qualsiasi affermazione - o proposizione - della quale sia verificabile il contenuto di verità, almeno in linea di principio. L’affermazione può riferirsi indifferentemente al passato, al presente o al futuro, ad esempio:

  – “pioveva a Roma il giorno della battaglia di Waterloo”;

  – “in quest’istante un treno merci sta transitando in un certo tratto della ferrovia Roma-Firenze”;

  – “nel prossimo lancio di un dado uscirà la faccia con il numero 5”;

  – “avendo letto sulla <u>mia</u> bilancia 10.05 g, la massa dell’oggetto risulterà essere compresa fra 10.03 e 10.07 g qualora essa sia misurata con una bilancia di altissima precisione e perfettamente calibrata”.

- Se l'evento è ben definito esso può essere, da un *punto di vista logico*, vero o falso[6]: "pioveva o non pioveva", "transita o non transita", "esce il 5 o un altro numero", "$10.03 \leq m \leq 10.07\,\mathrm{g}$ o no".

- In tutti gli esempi riportati siamo, *dal punto di vista conoscitivo*, in condizioni di incertezza. A ciascuno di questi eventi possiamo attribuire un certo livello di probabilità a seconda delle conoscenze che si hanno sull'evento. Ad esempio, se invece di pioggia a Roma si fosse interessati a tale evento a Milano o a Palermo, la valutazione di probabilità sarebbe stata diversa. Lo stesso vale se si viene a sapere che la battaglia di Waterloo è avvenuta a luglio. Nel caso della misura, la valutazione dipende dalla conoscenza della bilancia: se, invece della "mia bilancia di laboratorio", si trattasse di una bilancia di provenienza ignota, la valutazione cambierebbe.

- L'assegnazione di valutazione della probabilità può essere effettuata in vari modi ma, indipendentemente da essi, se il numero è alto si crede molto che l'evento possa verificarsi, mentre se il numero è molto piccolo si ritiene l'evento "pressoché" impossibile.

- Da queste considerazioni segue la seguente definizione del <u>concetto</u> di probabilità:

  > *la probabilità è la misura del grado di fiducia che un evento si verifichi.*

  Ricordiamo che l'espressione "si verifichi" sta per "si verifica il contenuto di verità dell'affermazione espressa dall'evento" e non dipende dal fatto che l'avvenimento debba ancora accadere (vedi esempio di Warteloo).

La definizione adottata non è nient'altro che una rivalutazione del concetto intuitivo di probabilità. Per passare dal concetto ad una teoria della probabilità è comunque necessario:

- *quantificare in un numero* il livello di probabilità delle diverse affermazioni a cui si è interessati, ovvero definire delle regole per la *valutazione* della probabilità;

- stabilire una serie di regole che questi numeri devono soddisfare, ovvero costruire uno schema formale a base della teoria.

Questo è quanto sarà trattato nel seguito, parzialmente in questo capitolo e completato nel resto del testo.

[6]❄Questa schematizzazione segue lo schema tradizionale della logica dicotomica, di origine aristotelica e universalmente diffusa, in cui un evento può essere soltanto vero o falso. La descrizione del mondo reale è più complicata e un evento può essere in parte vero in parte falso. Si pensi ad affermazioni del tipo "persona alta", "ragazza carina", "uomo sportivo", o anche "piove a Roma". Citando un esempio classico, "quale sasso è responsabile della transizione da non-mucchio a mucchio?".

Da alcuni decenni si è sviluppata la cosiddetta *fuzzy logic* - logica sfumata - che si avvicina meglio di quella tradizionale al modo di classificazione del cervello umano. Nel seguito non ci interesseremo di questo tipo di logica, sia per il carattere introduttivo del corso, sia perché la maggior parte degli eventi dei quali ci occuperemo si prestano abbastanza bene ad una classificazione di tipo dicotomico

1.8 Semplici valutazioni di probabilità

Se si chiede a qualsiasi persona quanto vale la probabilità dell'esito di eventi elementari, come una certa faccia nel lancio di una moneta o di un dado, si otterrà essenzialmente la risposta "giusta" di 1/2 e 1/6 rispettivamente. Anche se si rivolge la domanda, opportunamente formulata, a un bambino o una persona di scarsa cultura[7], cioè che non sia in grado di esprimere il risultato sotto forma di frazione o di numero razionale, si ottiene sostanzialmente una risposta equivalente: "non c'è ragione di ritenere un esito più probabile degli altri"; "tutte le faccie hanno la stessa possibilità di uscire". Cioè si esprime un giudizio di equiprobabilità - di indifferenza - sugli esiti.

Che tutti siano in grado di effettuare in modo intuitivo valutazioni di probabilità in situazioni semplici può essere facilmente verificato ponendo opportune domande. Per esempio, consideriamo il classico dado e interessiamoci agli eventi E_1 ="6", E_2 ="numero pari", E_3 =">1". Si proponga un gioco nella quale chi indovina vince un premio, oppure una semplice scommessa alla pari, ovvero in cui se si vince si riceve il doppio di quanto si è puntato. Tutti scommetteranno sull'esito che "riterranno più probabile" e anche un bambino sceglierà l'evento E_3, perché "è quello che uscirà più facilmente". Non è difficile convincersi che il ragionamento seguito a livello intuitivo è quello espresso in modo molto chiaro dal filosofo Hume:

> *"There is certainly a probability, which arises from a superiority of chances on any side; and according as this superiority increases, and surpasses the opposite chances, the probability receives a proportionable increase, and begets still a higher degree of belief or assent to that side, in which we discover the superiority. If a dye were marked with one figure or number of spots on four sides, and with another figure or number of spots on the two remaining sides, it would be more probable, that the former would turn up than the latter; though, if it had a thousand sides marked in the same manner, and only one side different, the probability would be much higher, and our belief or expectation of the event more steady and secure. This process of the thought or reasoning may seem trivial and obvious; but to those who consider it more narrowly, it may, perhaps, afford matter for curious speculation."*

Un altro caso di valutazione che porta a giudizi concordi è quando si mostrano risultati di un esperimento, che in principio può essere complicato a piacere, ma che può essere schematizzato in un certo modo standard che illustreremo fra poco. Ad esempio si può usare un contatore di radioattività e osservare il numero di conteggi registrato in un piccolo intervallo di tempo. Si effettuano un certo numero di osservazioni, diciamo 100, nelle quali 56 volte si

[7]❄Si badi bene che le cose cambiano molto se si prendono in considerazione eventi più complicati e che in genere richiedono dei conti anche ad esperti. Così pure - come è ben noto - l'intuizione di molti può essere tratta in inganno quando si chiede quale faccia esce dopo che si sono verificate 5 teste. Quindi bisogna stare attenti a non fare del "populismo probabilistico". Come diceva Bruno de Finetti, per conoscere il risultato di una certa operazione aritmetica non lo si chiede alla gente e poi si prende un risultato medio, ma ci si preoccuperà di insegnare loro la matematica.

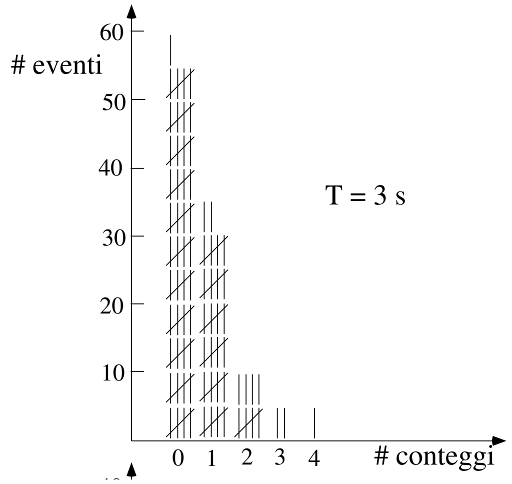

Figura 1.3: Istogramma del risultato di cento misure di conteggio. Il simbolo "#" sta per "numero". Il termine "evento" è qui usato - come avviene usualmente - nel senso di "occorrenza" o "numero di volte". Quale valore credete si verificherà in una ipotetica 101-ma misuare effettuate nelle stesse condizioni?

sono verificati zero conteggi, 32 volte un conteggio, 9 volte due, 2 volte tre e 1 volta quattro (vedi figura 1.3). Se prima di effettuare la 101^{ma} misura si chiede a coloro che hanno assistito all'esperimento quale numero si verificherà, eventualmente provocando la risposta con una scommessa alla pari, la risposta sarà scontata. Così pure emergerà spontaneamente che: "la probabiltà dello zero è del 56 %", "lo zero ha circa il doppio di possibilità di uscire rispetto all'uno", e così via. Addirittura non ci sarà nessuno disposto a scommettere che non uscirà "mai" un numero superiore al quattro, anche se affermerà che gli "sembra poco probabile".

Se però si sposta il contatore in un'altra stanza, oppure lo si cambia con uno della stessa marca e modello, o ci si mette intorno uno schermo di piombo o una sostanza misteriosa, o semplicemente lo si spegne e poi lo si riaccende, nessuno risponderà più con la stessa sicurezza: "proviamolo un po' e poi vediamo", proporranno.

Infatti il giudizio che faceva ritenere più probabili quanto era accaduto più frequentemente nel passato si basa su considerazioni del tipo: "il rivelatore sembra essere andato a regime"; "perché mai la radioattività dovrebbe cambiare nel giro di qualche minuto?"; "non mi sembra di aver osservato variazioni nel modo di fluttuare dei numeri fra i diversi valori che di volta in volta si presentavano", "credo che nei prossimi minuti il sistema si comporti come negli ultimi minuti". Anche questo processo dell'intelletto di trasferire nella probabilità di eventi futuri la frequenza di eventi passati è stato considerato da Hume:

> *"Being determined by custom to transfer the past to the future, in all our inferences; where the past has been entirely regular and uniform, we expect the event with the greatest assurance, and leave no room for any contrary supposition. But where different effects have been found to follow from causes, which are to appea-*

> *rance exactly similar, all these various effects must occur to the mind in transferring the past to the future, and enter into our consideration, when we determine the probability of the event. Though we give the preference to that which has been found most usual, and believe that this effect will exist, we must not overlook the other effects, but must assign to each of them a particular weight and authority, in proportion as we have found it to be more or less frequent.*"

Si noti inoltre che, mentre si è disposti ad affermare che la probabilità di zero conteggi alla 101^{ma} osservazione sia del 56 % (benché accompagnato da un doveroso "circa" prudenziale), si è molto più cauti a dire che la probabilità di quattro sia 1 % o che quella di numeri superiori sia nulla. In questo caso il "circa" di prima diventerà un più eloquente "all'incirca" accompagnato da vistose ondulazioni della mano. E anche lo "0 %" diventerà un "sembra al più dell'ordine del percento".

Possiamo quindi dire che il ragionamento che portava a questo secondo modo di valutare la probabilità sia legato a un giudizio sull'uniformità nel tempo delle prove (processo fisico e strumentazione) passate e future, confortato da un grande numero di risultati.

1.9 Ricapitolando

- Le osservazioni sperimentali non permettono di arrivare a conclusioni certe sulla validità di teorie scientifiche o sul valore di grandezze fisiche.

- A maggior ragione, siamo incerti sulle previsioni di eventi futuri, o comunque incerti, in quanto questa incertezza dipende dall'incertezza sulla teoria e sui suoi parametri, più le incertezze su fattori di influenza e di rumore difficilmente (o assolutamente) incontrollabili.

- Il solo paradigma della falsificazione è inadeguato a trattare le implicazioni derivanti dalle osservazioni, in quanto tutte le teorie non falsificate sono trattate alla stessa stregua. Questo approccio è in contraddizione con l'analisi storica che mostra come le comunità scientifiche abbiano sempre preferito seguire la via ritenuta più plausibile (più probabile), senza nessun argomento di necessità logica e senza attendere la falsificazione di tutte le infinite ipotesi possibili.

- La mente umana, per supplire alla mancanza di certezza pur senza considerare allo stesso modo tutto ciò che è possibile, ha sviluppato il concetto di probabilità, come misura del grado di credibilità di un evento incerto.

- In molti casi è possibile farsi intuitivamente un'idea quantitativa del livello di probabilità esprimendo giudizi di indifferenza (equiprobabilità) rispetto a più casi elementari possibili, oppure credendo che il futuro scorra allo stesso modo del passato e ciò che si è verificato più frequentemente nel passato accadrà più probabilmente nel futuro.

- Nel caso di valutazione di probabilità dalle frequenze viene spontaneo non credere esattamente al valore di frequenza, specialmente se ottenuto con un piccolo numero di prove, ma si tende a “smussare” le osservazioni cercando delle regolarità fra le frequenze osservate.

1.10 Problemi

1. Analizzare il valore di verità delle seguenti frasi pronunciate a Roma mercoledì 8 ottobre 1997: a) domani è giovedì ; b) "domani arriva la Befana"; c) "domani piove"; d) "domani nevica".

2. Analizzare il valore di verità della frase "nel 1969 un uomo arriva sulla luna" pronunciata nel: a) 1920; b) 1961; c) 1997;

3. Si considerino i seguenti eventi, relativi alla ruota di Bari del gioco del lotto:

   - E_1 = "il primo estratto del 1975 è pari";
   - E_2 = "il primo estratto di sabato prossimo (rispetto al giorno in cui stai leggendo questo testo) è pari";
   - E_3 = "la somma dei cinque numeri estratti il 10 maggio 1997 è inferiore a 15";
   - E_4 = "i cinque estratti di sabato prossimo sono tutti dispari";
   - E_5 = "i cinque estratti del 10 maggio 1997 sono tutti pari";
   - E_6 = "se i numeri estratti sono tutti pari la loro somma è maggiore di 27;
   - E_7 = "il 10 maggio 1997 è uscito il terno 1-2-3";
   - E_8 = "il 10 maggio 1997 è uscito il terno 12-45-60;
   - E_9 = "il 10 maggio 1997 è uscito l'ambo 37-45";

   (a) A quali eventi è applicabile il concetto di probabilità?

   (b) È più probabile E_7 o E_8?

   (c) Supponiamo che E_9 sia vero. In tal caso sarebbe più probabile E_7 o E_8?

4. Una scatola contiene 8 palline bianche e 2 nere, mentre un'altra ne contiene 2 bianche e 8 nere. Le due scatole sono indistinguibili. Si estrae una pallina da una scatola scelta a caso e, senza guardarla, la si ripone nell'altra scatola. Successivamente si estrae una pallina da quest'ultima scatola. Quanto vale la probabilità che la pallina sia bianca?

5. Una persona possiede due monete false, di cui una ha due teste e l'altra due croci. Ne estrae a caso dalla tasca una e la lancia: quanto vale la probabilità che esca testa?

6. *** Aggiungere altre semplici valutazioni di probabilità... ***

Capitolo 2

Valutazioni e interpretazioni della probabilità

Sebbene mi sentissi sicuro che la parità non fosse violata, c'era ancora una possibilità che lo sarebbe stata, e era importante investigare. 'Scommetteresti cento dollari contro uno che la parità non è violata?' chiese. 'No, ma cinquanta sì.'

(R.P. Feynman)

2.1 Primi interessi in stime quantitative di probabilità

L'inizio della teoria delle probabilità, chiamata all'epoca la "dottrina della sorte", avviene nel XVII secolo, come risposta a due classi di problemi, legate rispettivamente ai giochi d'azzardo e alle assicurazioni.

Nel primo caso si trattava di valutare la probabilità di vincere scommettendo sul verificarsi di un certo evento, ad esempio la faccia con su inciso il numero 6 nel lancio di un dado.

Nel secondo caso si rendeva necessaria per le banche la stima della probabilità di morte di un individuo di una certa età, ovvero la probabilità che egli potesse sopravvivere un determinato numero di anni dalla stipula del contratto.

Questi due diversi contesti hanno dato luogo a due diversi metodi per valutare la probabilità, o, come si usa ancora dire, a due "definizioni" di probabilità viste talvolta in contrapposizione fra di loro: la *"matematica"* e la *"sperimentale"* (o *"empirica"*); la *"classica"* e la *"frequentista"*; quella *"a priori"* e quella *"a posteriori"*. Il termine definizione è racchiuso fra virgolette, ad indicare che, nell'approccio che andremo a seguire, si tratta in realtà di "metodi di valutazione della probabilità", mentre la definizione è - ripetiamo - semplicemente quella intuitiva di misura della credibilità di un evento.

Nei prossimi paragrafi saranno introdotti questi due metodi di valutazione, ancora oggi fra i più usati nei casi relativamente semplici ai quali essi si applicano. Seguirà poi una discussione sia sul fatto che sostanzialmente essi vorrebbero far riferimento alla stessa grandezza, sia sulla loro pretesa oggettività. Infine verrà introdotta in modo più preciso la definizione soggettiva della probabilità mediante la quale è possibile risolvere le difficoltà insite nelle altre due definizioni.

2.2 Valutazione combinatoria

Assumiamo che l'esito di un possibile esperimento, ovvero l'evento di interesse, consista nella realizzazione di una tra più *modalità elementari* che si escludono a vicenda, ove con "elementare" si vuole indicare:

a) l'impossibilità di classificare ulteriormente gli eventi in base a qualche caratteristica che hanno alcuni e non altri;

b) che tali eventi devono essere tutti quelli possibili;

c) che essi sono fra di loro a due a due incompatibili, ovvero che il verificarsi di uno escluda tutti gli altri.

Esempi di questi eventi sono le 6 faccie di un dado, i 90 numeri del lotto, e così via. Questi eventi sono spesso chiamati "casi possibili" e li indicheremo con il simbolo e_i. Ad esempio, per il lancio del dado, $e_1 = \{1\}$, $e_2 = \{2\}, \ldots,$ $e_6 = \{6\}$.

In generale più modalità elementari possono contribuire allo stesso evento. Ad esempio l'evento "numero pari nel lancio del dado" è costituito dai tre eventi elementari $e_2 = \{2\}$, $e_4 = \{4\}$ e $e_6 = \{6\}$. Quenti eventi elementari vengono chiamati *casi favorevoli* e gli altri, di conseguenza, *casi contrari*.

Se il meccanismo di estrazione è talmente simmetrico rispetto a ciascuno dei casi possibili, nessuna modalità è da ritenersi più probabile delle altre. Questa osservazione è stata elevata a principio da Laplace. A tale principio fu successivamente dato il nome di *Principio di Ragione non Sufficiente*[1] o *Principio di Indifferenza*.

Per passare dalla equiprobabilità dei casi possibili alla valutazione numerica della probabilità rimane da definire una scala per i valori di probabilità. Per semplicità viene scelto il valore 0 per un evento impossibile e il valore 1 per un evento certo (anche se nella vita pratica si preferiscono le percentuali e si considera la probabilità compresa fra 0 e 100 %).

Se n è il numero di casi (ugualmente) possibili e $P(e_i)$ la probabilità di ciascun evento elementare, un evento generico E_j che ha n_j eventi favorevoli ha probabilità n_j volte quella elementare. Poiché l'evento certo, di probabilità 1, è costituito da n eventi elementari, ciascuno di questi ha probabilità $P(e_i) = 1/n$, mentre per il generico evento vale $P(E_j) = n_j/n$.

Quindi, la probabilità di ciascuna delle modalità elementari, sotto condizione di equiprobabilità, è pari all'inverso del numero di casi possibili, mentre un generico evento ha probabilità proporzionale al numero delle modalità elementari nelle quali può manifestarsi. Da ciò segue la così detta *"definizione" classica di Laplace* :

[1] Come contrapposizione al principio di ragion sufficiente di Leibnitz, secondo il quale "nulla accade senza che vi sia ragione perché accada proprio così invece che altrimenti" (una versione più raffinata del popolare "non muove foglia che Dio non voglia").

> *il valore della probabilità di un dato evento è pari al rapporto tra il numero dei casi favorevoli all'evento e il numero dei casi possibili, se questi sono egualmente probabili*:

$$p = \frac{\text{numero dei casi favorevoli}}{\text{numero dei casi egualmente probabili}} \,. \tag{2.1}$$

Come si intuisce, questa "definizione" è utile per assegnare un numero al livello di probabilità, ma sicuramente non definisce la probabilità come concetto, in quanto essa si basa sul concetto primitivo di equiprobabilità. È altresì ovvio che tale definizione è soddisfatta in circostanze elementari facilmente schematizzabili, ma non è affatto adeguata a descrivere fenomeni complessi, che sono quelli di maggiore interesse per le applicazioni della vita reale. Ma, e soprattutto, è da notare come il giudizio di equiprobabilità sia lasciato a discrezione di chi valuta la probabilità, cioè esso è in larga misura "soggettivo".

A tale riguardo è istruttiva questa riflessione di Poincaré[2]:

> *. . . Siamo costretti a definire il probabile dal probabile. Come possiamo sapere se due casi sono ugualmente probabili? Sarà per convenzione? Se inseriamo all'inizio di ogni problema una convenzione esplicita, bene! Allora non dobbiamo far altro che applicare le regole dell'aritmetica e dell'algebra e completare il calcolo, quando il nostro risultato non può essere chiamato in questione. Ma se vogliamo fare la minima applicazione di questo risultato dobbiamo provare che la nostra convenzione è legittima e ci troveremo in presenza della difficoltà di fondo che pensavamo di aver evitato.*

2.3 Probabilità e frequenza

Come accennato, il primo tentativo di valutare la probabilità fuori dall'ambito dei giochi d'azzardo fu motivato dal calcolo delle pensioni. L'ammontare del vitalizio da corrispondere all'assicurato dipende infatti, oltre che dal capitale accumulato e da altri fattori strettamente economici, dalla probabilità di morte dell'assicurato in funzione della sua età. Infatti, a seconda dei casi, la compagnia guadagna molto in caso di decesso prematuro, mentre perde molto nel caso di straordinaria longevità. Una valutazione realistica *** attenzione alle implicazioni di oggettività*** della probabilità è resa necessaria onde evitare perdite economiche nei casi di eccessive sottostime o sovrastime. Nel primo caso la società farà bancarotta a causa dei vitalizi troppo onerosi che si è impegnata a pagare, nel secondo sarà battuta dalla concorrenza che, disponendo di valutazioni più realistiche, potrà offrire ai clienti premi più convenienti.

Considerando la sopravvivenza di ciascuna persona da un anno all'altro, ci sono due modalità elementari, ma esse sono per fortuna non equiprobabili e quindi la "definizione" di Laplace è inapplicabile.

Il problema fu risolto compilando delle tabelle di mortalità per le varie età e stimando la probabilità dalla frequenza (ovvero da quante volte quel tipo di evento si è verificato nel passato).

[2] Henry Poincaré, "Scienza e Ipotesi"

Per essere più concreti - e meno macabri - interessiamoci ad esempio alla valutazione della probabilità di sopravvivenza nell'arco di un anno delle persone di una certa classe di età. Assumiamo che il 1° di gennaio di un certo anno la popolazione sia costituita da n individui di una certa età e che il 1° gennaio dell'anno successivo siano ancora in vita n_f di quegli n. Questo era ritenuto equivalente ad aver compiuto n esperimenti - o *prove* - dei quali n_f hanno dato esito favorevole e $(n - n_f)$ esito sfavorevole. Ciascuna prova è inoltre considerata indipendente dall'altra e la probabilità di esito favorevole è ritenuta costante (è facile intuire come questa schematizzazione sia particolarmente rozza e che non potrà mai portare ad una valutazione della probabilità di sopravvivenza del singolo individuo, ma al più ad una probabilità relativa all'individuo tipo).

Nel valutare la probabilità di sopravvivenza dei <u>nuovi</u> individui che hanno ora raggiunto tale età si effettua il seguente ragionamento: in generale, se la valutazione di probabilità è corretta, l' esito a cui è assegnata probabilità più elevata è quello dei due che si ritiene possa accadere più facilmente. In particolare, se uno dei due fosse impossibile non accadrebbe mai, mentre se fosse certo si verificherebbe con sicurezza. Supponendo che il processo si ripeta nel futuro con le stesse condizioni con cui era avvenuto nel passato allora il ragionamento precedente può essere invertito ("l'esito più frequentemente accaduto è quello che "aveva" maggiore probabilità di accadere") ed esteso al futuro [3]. La probabilità che i nuovi individui di quella classe di età sopravvivano è quindi ritenuta proporzionale al numero di eventi favorevoli (F) dell'anno precedente: $P(F) \propto n_f$.

Poiché l'evento certo ("qualcosa deve accadere"), proporzionale a $n_f + (n - n_f) = n$, viene convenzionalmente posto uguale ad 1, si ottiene che il fattore di proporzionalità vale $1/n$. Ne segue che $P(F) = n_f/n$. Ovvero che *la frequenza relativa di successo nel passato viene adottata a misura della probabilità di successo nel futuro*.

Estendendo il semplice problema dell'esperimento con due esiti ad esperimenti a molti esiti, come potrebbe essere ad esempio quello di n misure nelle quali il risultato si può presentare in tante modalità diverse (sia E_i la i-ma modalità), otteniamo

$$p(E_i) \approx \frac{n(E_i)}{n}\,, \tag{2.2}$$

[3]❄Questo processo dell'intelletto umano è descritto molto bene da Hume nel capitolo "Probabilità" del suo "Saggio sull'intelletto umano": *"Essendo costretti dalla consuetudine a trasferire il passato al futuro in tutte le nostre inferenze, quando il passato si è manifestato del tutto regolare e uniforme ci aspettiamo un evento con la massima sicurezza e non lasciamo posto a qualche altra supposizione contraria. ... Sebbene diamo la preferenza a quello che è stato trovato più usuale e crediamo che questo effetto si verificherà, non dobbiamo trascurare gli altri effetti, ma dobbiamo assegnare a ciascuno di essi un particolare peso e autorità in proporzione a come lo abbiamo trovato più o meno frequente".* Francamente, mi sembra fra le cose più sensate dette a giustificazione dell'uso della frequenza di eventi passati come valutazione della probabilità di eventi futuri. Questa semplice constatazione è senz'altro più convincente dei tentativi di far discendere tale prassi dalla "Legge empirica del Caso", dal paradossale uso del teorema di Bernoulli a tale scopo (vedi nota nel seguito), o dai macchinosi tentativi di matematizzazione attraverso i "collettivi" di von Mises.

ovvero la probabilità viene valutata come

$$p \approx \frac{\text{numero di prove con esito favorevole}}{\text{numero totale delle prove}}, \tag{2.3}$$

o, più concisamente,

$$p \quad \approx \quad \text{frequenza relativa di successo}\,.$$

Questo metodo di valutazione delle probabilità è affetto da due limitazioni:

- innanzitutto il segno “$\approx$” al posto di “$=$” sta a ricordare che questa valutazione soffre di grosse incertezze se il numero di prove non è molto grande. Ad esempio, nel caso estremo in cui sia stata effettuata una sola prova, la frequenza relativa vale 0 o 1;

- esso presuppone che gli esperimenti (passati, presenti e futuri) possano essere ripetuti nelle stesse condizioni, ovvero in modo indipendente l’uno dall’altro e con la probabilità di ciascuna delle modalità costante nella singola prova.

La prima limitazione può essere superata se si dispone di un grande numero di prove. La seconda invece è insita nella natura delle cose ed è compito di chi valuta la probabilità giudicare, di volta in volta, se l’ipotesi di equiprobabilità delle prove possa essere accettata entro i limiti di accuratezza con cui è necessario stimare la probabilità.

Quindi anche in questo caso spetta a chi valuta la probabilità decidere, con un ampio margine di arbitrarietà, se l’ipotesi di “stesse condizioni” è verificata e se le prove si svolgeranno nel futuro come si erano svolte nel passato. Ad esempio, se in una popolazione si sviluppa improvvisamente un’epidemia o se, al contrario, migliorate condizioni di vita e assenza di eventi bellici diminuiscono la mortalità non ci si può più basare sulle tabelle di sopravvivenza degli anni precedenti.

È da notare inoltre come la richiesta di un grande numero di prove possa essere in contrasto con la non - o scarsa - ripetibilità del fenomeno sotto studio. In genere sia lo strumento che la grandezza da misurare[4] possono subire modificazioni nel corso del tempo. Che senso avrebbe, ad esempio, affermare che al 95 % di probabilità una navicella spaziale si trova in un certo istante in una certa regione di spazio se si credesse che il solo modo di intendere la probabilità sia quello basato sulla frequenza?

2.4 Legge empirica del caso e “definizione” frequentista

Nel passato ha avuto grande importanza la *constatazione empirica* che, per i semplici casi in cui si sa calcolare la probabilità mediante ragionamenti di simmetria (“definizione” di Laplace), si nota che la frequenza relativa “è in genere

[4]La determinazione dei valori delle costanti fondamentali è soltanto uno dei tanti aspetti della problematica della misura.

i	N^i	N_T^i	N_T^i/N_C^i	$f(N_T^i)$	$\lvert N_T^i - N_C^i\rvert$
1	10	3	0.42857	0.3	4
2	100	53	1.12766	0.53	6
3	1000	484	0.93798	0.484	24
4	10000	4956	0.98255	0.4956	88
5	100000	49983	0.99932	0.49983	34
6	1000000	500475	1.00190	0.500475	950
7	10000000	5000790	1.00032	0.5000790	1580

Tabella 2.1: Risultati della simulazione al computer di estrazioni con probabilità a priori di successo pari a 1/2. T e C stanno per "testa" e "croce" per analogia con il lancio di monete perfette. N^i, N_T^i e $f_{N_T^i}$ indicano il numero di "lanci", il numero di "teste" e la loro frequenza relativa. È anche riportato il valore assoluto della differenza fra numero di teste e "numero" di "croci".

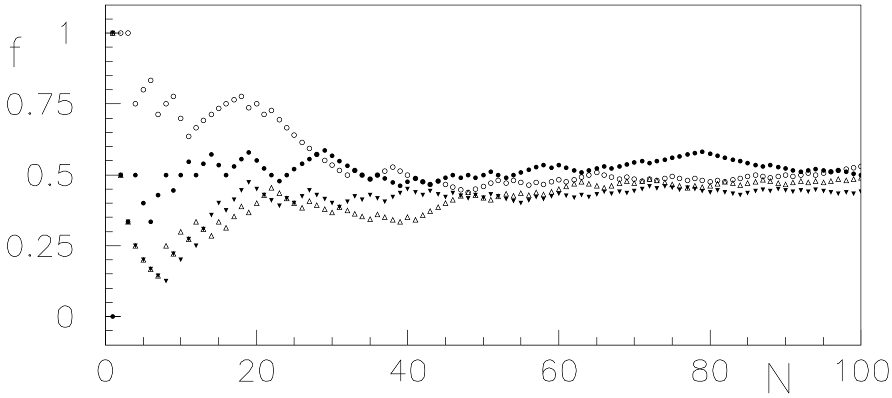

Figura 2.1: Frequenza relativa dell'evento "Testa" in funzione del numero di eventi in un esperimento simulato del lancio di una moneta. Sono riportate quattro sequenze indipendenti.

prossima" al valore di probabilità calcolato *a priori*. In effetti, sperimentalmente risulta che questo è "praticamente vero" quando il numero di prove è molto grande. In altri termini, maggiore è il numero di prove e "meno frequentemente succede" di osservare grandi scarti della frequenza dal valore di probabilità, ovvero

$$\text{“}\lim_{N\to\infty}\text{”}\,\frac{N_E}{N} = p \tag{2.4}$$

Esempi (simulati al computer) di frequenza relativa in funzione del numero dei lanci di una moneta sono mostrati nella tabella 2.1 e nella figura 2.1.

Questo tipo di osservazioni sperimentali ha condotto ad enunciare la così detta *legge empirica del caso*:

> *In una serie di prove ripetute un gran numero di volte nelle stesse condizioni, ciascuno degli eventi possibili si manifesta con una*

frequenza relativa che è circa uguale alla sua probabilità. L'approssimazione migliora con il numero delle prove.

Questo fatto prettamente empirico ha suggerito di ribaltare la (2.4) e "definire" la probabilità in termini di frequenza relativa misurata in una successione arbitrariamente grande di prove:

$$p = \text{“}\lim_{N\to\infty}\text{''}\,\frac{N_E}{N}\,. \tag{2.5}$$

(il simbolo di limite è fra virgolette perché non ha niente a che vedere con i limiti intesi in senso matematico). L'interesse del collegamento fornito dalla legge empirica del caso[5] consiste nel fatto di poter calcolare la probabilità anche in quelle circostanze in cui è difficile fare l'inventario di tutti i casi possibili ed equiprobabili su cui si basa la definizione classica di Laplace, con la convinzione che si stia sostanzialmente determinando la "stessa" quantità.

Come fatto notare a proposito della (2.4), anche la (2.5) non va intesa come un limite nel senso matematico. Non c'è infatti nessuna garanzia che data una grandezza piccola ϵ esista un $N_\circ$ arbitrariamente grande tale che se $N > N_\circ$ ci sia l'assoluta certezza che la frequenza relativa differisca da p meno di ϵ.

2.5 Interpretazione oggettivista e soggettivista della probabilità

Dal punto di vista storico le due "definizioni" di probabilità appena incontrate, legate fra loro dalla "legge empirica del caso" hanno indotto molti a

- confondere il concetto di probabilità con il suo metodo di valutazione, come si può verificare in vari libri di testo e voci di enciclopedia;
- ritenere che il valore di probabilità sia oggettivo, cioè che sia insito nella natura delle cose e non dipenda da chi lo valuta;
- credere che si possa parlare di probabilità solo in questi due casi, molto particolari e riduttivi, rispetto alla complessità del mondo reale. Sarebbero quindi esclusi da argomentazioni probabilistiche tutti quegli eventi (la stragrande maggioranza di quelli di interesse pratico e scientifico) per i quali è impossibile eseguire l'inventario dei casi possibili e di quelli favorevoli o per i quali non è possibile ripetere "infinite volte" l'esperimento nelle identiche condizioni.

[5] ❄In molti testi il ruolo di anello di congiunzione fra le due "definizioni" è affidato al Teorema di Bernoulli, che vedremo nel seguito. Anticipiamo che esso afferma che, al *"crescere del numero di prove, diventa piccola a piacere la probabilità che la frequenza relativa differisca dalla probabilità dell'evento favorevole in ciascuna delle prove"*. Preferiamo basarci invece sulle osservazioni empiriche che hanno condotto alla legge empirica del caso sia perché ci sembra più corrispondente al processo storico che per la palese illogicità dell'uso di questo teorema per questo scopo in quanto, come afferma de Finetti (1970), "non si sfugge al dilemma che la stessa cosa non si può assumere prima per definizione e poi dimostrare come teorema, né alla contraddizione di una definizione che assumerebbe una cosa certa mentre il teorema afferma che è soltanto molto probabile". Torneremo su queste osservazioni nel momento appropriato.

Abbiamo invece visto come in quelle stesse definizioni ci sia una fondamentale componente soggettiva nel decidere caso per caso se le clausole di equiprobabilità sono soddisfatte.

Inoltre non è difficile convincersi che la probabilità dipende dallo *stato di conoscenza* (o *stato di informazione*) di chi la valuta:

- la probabilità 1/6 di ciascuno dei risultati del lancio di un dado assume la regolarità del dado. Ciò nonostante un dado reale avrà necessariamente delle asimmetrie (ad esempio dovute al diverso numero di puntini incisi sulle facce[6]) e qualcuno può credere a valori diversi da 1/6 se ha una conoscenza meno vaga di quel dado[7];

- la situazione in cui il dado è stato già lanciato, ma riparato dalla vista di chi deve stimare la probabilità, non cambia la valutazione. La probabilità (del dado regolare) è sempre 1/6 anche se dal punto di vista fisico l'esito è già determinato.

- Se inoltre si riesce a sbirciare e a leggere il valore di una faccia laterale la situazione cambia immediatamente: la probabilità della faccia vista e del suo complemento[8] a 7 si annullano e quelle delle altre facce aumentano ad 1/4.

- Se si prende un sacchetto con i 90 numeri della tombola e lo si fora da sotto, chiaramente la probabilità che esca un numero che è depositato sul fondo è chiaramente maggiore di quella dei numeri che stanno sopra, ma la nostra *ignoranza* sulla posizione dei numeri fa sì che non possiamo far altro che assegnare la stessa probabilità a tutti i numeri.

2.6 Concetto di probabilità condizionata

Se la valutazione di probabilità dipende dallo stato di informazione, non ha senso parlare di una "probabilità assoluta". Si può soltanto parlare di *probabilità condizionata* ad una certa informazione. Indichiamola genericamente con $P(E\,|\,I)$, letta "probabilità di E dato lo stato di informazione I", o "probabilità di E data I". Quando si parla di $P(E)$ senza aggiungere altro si fa riferimento a circostanze convenzionali, oppure - implicitamente - allo stato di informazione di chi la valuta.

Facciamo alcuni esempi:

- quando si dice che la probabilità della faccia di un dado sia 1/6 si sta assumendo che dado e lancio siano perfettamente regolari;

[6] Qualcuno ha stimato che l'effetto dei forellini sulle facce dei dadi sia tale che le probabilità dei valori da 1 a 6 siano da ritenersi,rispettivamente, 0.155, 0.159, 0.164, 0.169, 0.174, 0.179 (vedi Shafer, 1976).

[7] Per esempio, da misure effettuate da studenti di laboratorio, risulta che il contributo maggiore alla non equiprobabilità delle facce dei dadi è dovuto al fatto che i dadi commerciali non sono cubi perfetti. Ne segue che le faccie opposte che sono più vicine fra loro tendono ad uscire più frequentemente.

[8] Si ricorda che nei dadi la somma delle facce opposte dà 7.

- quando si dice che la probabilità di Testa nel lancio di una moneta sia 1/2 si assume che la moneta sia regolare, che il lancio sia “fatto a caso” (non è irragionevole pensare ad un prestigiatore in grado di controllare il lancio) e che inoltre la moneta non cada verticale e non vada in un tombino (né che sia catturata al volo da un amico).

Quindi la probabilità va valutata entro lo *spazio di possibilità di interesse*. Ad esempio ci si può interessare alla probalità della faccia “6” del dado limitatamente al verificarsi di un numero pari ovvero $P(\text{“}6'' \,|\, \text{“pari”}) = 1/3$, oppure alla probabilità che la somma ottenuta in due lanci sia pari a 12 nelle stesse condizioni, cioè $P(\text{“somma} = 12'' \,|\, \text{“entrambi pari}'') = 1/9$. Si noti che, più propriamente, bisognerebbe scrivere $P(\text{“}6'' \,|\, \text{“pari” \& “dado regolare”})$ e $P(\text{“somma} = 12'' \,|\, \text{“entrambi pari” \& “dado regolare”})$, anche se nel seguito eviteremo di scrivere la condizione “ovvia”[9].

Ragionamenti analoghi possono essere seguiti quando si voglia valutare la probabilità dalle frequenze (sempre nell'ipotesi che le clausole insite in tale valutazione siano soddisfatte). Si può quindi stimare la probabilità di furto di una macchina sotto la condizione che sia immatricolata a Milano, che il proprietario sia un venditore ambulante, che sia di una certa marca, o che le tre condizioni siano verificate simultaneamente (ammesso di trovare abbastanza macchine che corrispondano a tali requisiti per farne delle “statistiche”).

Quindi nei due semplici casi di valutazione della probabilità dal giudizio di equiprobabilità e dalle frequenze relative del passato si ottiene:

$$P(E\,|\,C) = \frac{\text{numero di casi favorevoli a } E \text{ e } C}{\text{numero di casi (equiprobabili) favorevoli a } C} \tag{2.6}$$

e

$$P(E\,|\,C) \approx \frac{\text{numero di prove favorevoli a } E \text{ e } C}{\text{numero di prove favorevoli a } C}\,. \tag{2.7}$$

Altre volte invece - nella maggior parte dei casi pratici della vita - lo stato di informazione cambia la valutazione, ma in modo più difficile da schematizzare: la probabilità di vittoria di un pilota di F1 dipende dal circuito e dalle condizioni ambientali (a parità di macchine); tutti sono d'accordo che la probabilità di vittoria di una squadra di calcio generalmente diminuisce se cinque titolari sono squalificati; la probabilità che un titolo azionario diminuisca di valore può dipendere dall'improvviso esonero dell'amministratore delegato (oppure può aumentare di quotazione se il dirigente allontanato era ritenuto incompetente).

Altri casi più prettamente scientifici nei quali non è possibile applicare le schematizzazioni di sopra sono quelli relativi alla probabilità di ipotesi scientifiche (teorie o valori di grandezze) alla luce di (“subordinate a”) osservazioni sperimentali. Ad esempio, quanto vali sono possibili valori di massa di un oggetto, se abbiamo letto m_1 su una bilancia? E come cambiano le mie credenze se pongo lo stesso oggetto su un'altra bilancia e leggo $x_2 \neq x_1$? Anticipiamo che in molti casi di rilievo relativi alla ricerca e alla teoria dell'incertezza di

[9] La come è ben noto a chi ha avuto cattive esperienze con compagnie di assicurazioni, quello che sembrava ovvio è contraddetto da una delle clausole scritte a caratteri microscopici nel contratto di assicurazioni. I contratti di assicurazioni, infatti, non sono altro che scommesse su eventi condizionati

misura la modifica della probabilità può essere effettata in modo logicamente consistente mediante un importante teorema del calcolo delle probabilità che vedremo a tempo debito.

Nel seguito, al fine di alleggerire la notazione, spesso scriveremo $P(E|I)$ semplicemente come $P(E)$, assumendo comunque che ci siano delle condizioni note a chi valuta la probabilità.

2.7 Eventi di probabilità nulla

Abbiamo detto che assegnamo probabilità nulla ad un evento che riteniamo impossibile. Ma non è vero il contrario, anche se ciò può sorprendere. Questo accade quando gli eventi presi in considerazione rappresentano una descrizione talmente particolareggiata di un accadimento che lo spazio di tutte le configurazioni possibili contiene un numero infinito di possibilità (anzi ∞^n, per parlare in termini matematici, con n abbastanza grande).

Quella che sembrerebbe un'affermazione sorprendente può essere verificata osservando la realtà che ci circonda e riflettendo, ad esempio, a quale sarebbe stata la probabilità - se avessi dovuto valutarla la settimana scorsa, o anche qualche minuto fa - che tu stia leggendo questa frase in questo esatto momento, mentre c'è una certa condizione meteorologica, con le persone che ti circondano vestite in un certo modo, che si scambiano certe frasi, la radio che trasmette una certa canzone o una precisa notizia di cronaca, etc.

Non servono teoremi matematici per convincersi che tutti gli eventi che accadono "avevano" probabilità nulla. Ma non per questo sono impossibili, altrimenti nulla accadrebbe. L'infinita varietà di eventi di probabilità nulla fa sì che la loro "somma" dia sempre l'evento certo ("qualcosa accade").

La possibilità di valutazione della probabilità in termini finiti è dovuta al fatto che nelle nostre schematizzazioni non siamo in genere interessati a definire un evento con tutti i dettagli. Analizzando una situazione di traffico a un incrocio, si può essere interessati al numero di macchine, allo stato dei semafori, ma non alle singole targhe delle autovetture e al colore dei capelli degli autisti, o alla velocità istantanea e posizione di ogni veicolo e pedone.

Ciò nonostante, anche quando si riflette su eventi di probabilità nulla, è possibile ritenere alcuni più credibili di altri. Ad esempio se si esaminano più foto dello stesso incrocio, alcune - anche con configurazioni di veicoli e persone completamente diverse fra loro - sembrano "più normali" e sono percepite come "ugualmente verosimili". Quindi dovendo fare considerazioni su situazioni analoghe future, queste verrebbero classificate con lo stesso grado di fiducia, sebbene tutte con probabilità uguale a zero. Ad altre situazioni, per esempio con un incidente oppure con cinque macchine d'epoca, si può dare un grado di fiducia 100 o 1000 volte inferiore, pur con la stessa probabilità nulla.

Anche se per il momento questo discorso delle probabilità nulle di eventi possibili non sembra produttivo ai fini di una valutazione numerica delle probabilità, esso sarà ripreso quando si considereranno le variabili continue. Inoltre, poiché la probabilità è sempre da valutare come probabilità condizionata, gli eventi di interesse sono spesso implicitamente subordinati a eventi di probabilità nulla. Ad esempio, anche se il caso sembra banale, considerando una qualsiasi configurazione di traffico (evento a probabilità nulla), ci si può interessare all'evento "le macchine di fabbrica italiana sono almeno il 60 %".

Un ultimo interesse educativo del considerare questo tipo di eventi è legata alla esorcizzazione dello stupore che hanno molti di fronte a fatti che sembrano *straordinari* ("estremamente poco probabili") soltanto per coincidenze fortuite o perché sono stati considerati in dettaglio per la prima volta. Questa meraviglia fa sì che si rischi di attribuire loro significati scientifici (. . . o paranormali) che essi generalmente non hanno.

2.8 Probabilità e scommesse eque

Avendo chiarito che una probabilità è da ritenersi sempre come probabilità condizionata ci resta ancora da trovare una regola generale e operativa per valutarla, ovvero dare una "definizione" che vada oltre il concetto intuitivo e che non sia limitata ai casi stereotipati degli eventi equiprobabili o delle prove effettuate in condizioni di equiprobabilità.

Un modo di superare l'impasse è quello di far ricorso al concetto di scommessa, percepibile a livello intuitivo da tutte le persone razionali. Finora, quando si è fatto uso del termine scommessa al fine di provocare un giudizio di maggiore probabilità fra più eventi alternativi, si è sempre indicata, o lasciata immaginare, una scommessa alla pari con l'intento di vincere. Ad esempio, considerando il lancio di un dado, dovendo scommettere pro o contro il "6", si è d'accordo che conviene puntare contro. Il solo problema è . . . trovare il pollo che accetti la scommessa.

Se invece si propone che chi gioca sul "6" punta mille lire e chi contro 99 mila lire si capisce come il problema si rovesci.

In entrambi i casi le scommesse hanno un verso privilegiato. Mentre nel primo caso nessuno era disposto a puntare 1000 lire per vincerne 2000, ora tutti sono favorevoli a giocare la stessa cifra per vincere 100 mila. Chiaramente il segreto non sta nell'importo superiore, perché nessuno punterebbe 50 mila sul sei per vincerne 100 mila. Così nessuno gioca alla pari contro la vittoria in casa della capolista del campionato di calcio nella partita contro l'ultima in classifica, ma potrebbe accettare la scommessa per vincere 10, 100 o 1000 volte la puntata.

Questo ci insegna che la puntata massima A che si è disposti a scommettere per ricevere una somma S se l'evento E si verifica e niente se non si verifica (eventualmente con le dovute condizioni) è proporzionale a quanto si crede che l'evento possa accadere, ossia

$$A \propto P(E)\,,$$

e all'importo S che si può vincere

$$A \propto S\,.$$

Quindi

$$A = P(E) \cdot S\,. \tag{2.8}$$

Se si ritiene che l'evento sia praticamente impossibile non si è disposti a puntare niente. Se si è assolutamente sicuri si arriva a puntare S (con la certezza di riprendersi poi i soldi).

Questa relazione potrebbe essere usata per valutare la probabilità se ci si convince che A non rappresenta soltanto la quota massima da pagare, bensì quella *giusta*. Infatti se fosse possibile giocare una puntata A' inferiore ad A il gioco sarebbe conveniente, se invece A' fosse superiore ad A la scommessa sarebbe inaccettabile, a meno di non invertirne il verso (diventando così conveniente).

La puntata A rappresenta quindi la puntata giusta che lascia il giocatore in dubbio in quale direzione giocare. Una scommessa di questo tipo è chiamata *equa* o *coerente*. Essa può essere usata per valutare $P(E)$.

> *La scommessa di una puntata A per vincere S se l'evento E si verifica e perdere A nel caso contrario, tale che essa risulti accettabile da una persona razionale in entrambe le direzioni, implica la valutazione di probabilità*
>
> $$P(E) = \frac{A}{S}\,. \tag{2.9}$$

La (2.9) impone dei vincoli su $P(E)$. Siccome nessuna persona razionale è disposto a scommettere $A > S$ o ad accettare una scommessa di $A < 0$ (si provi ad immaginare le conseguenze di quest'ultimo caso!) ne segue che

$$0 \leq P(E) \leq 1\,. \tag{2.10}$$

Questa è la prima delle regole che la probabilità deve soddisfare, che avevamo già incontrato precedentemente come scala arbitraria e che ora diventa una "scala naturale".

Un modo di immaginare una valutazione di probabilità in base alla scommessa coerente è che, una volta che una persona abbia valutato $P(E)$ e stabilito una quota A in base allla vincita totale S, l'altra persona scelga se puntare su E o contro di esso. Si tratta dello stesso principio che si applica nella suddivisione di una eredità o di una semplice torta fra due persone: la migliore garanzia di equità è che uno dei due faccia le parti e l'altro scelga. Con questa regola anche un bambino dividerebbe la torta in due parti praticamente uguali: ogni volta che si accorge che l'altro potrebbe guadagnarci scegliendo la parte visibilmente più grande riaggiusta la traiettoria del coltello.

2.9 ○ Probabilità e quote di scommessa

Talvolta la credibilità di un evento è espressa dal rapporto delle puntate che si ritiene equo scommettere pro e contro l'evento, ad esempio "1 a 1" o "100 a 1'. Mostriamo ora come sia possibile, dalla conoscenza dei rapporti delle puntate, risalire al valore di probabilità stimato. Chiamiamo B la puntata che si è disposti a scommettere su $\overline{E}$, opposto ad E, cioè quello che si verifica se non si verifica E (e viceversa). Chiaramente

$$B = P(\overline{E}) \cdot S\,. \tag{2.11}$$

Sommando membro a membro (2.8) e (2.11) si ha che

$$A + B \;\;=\;\; \big(P(E) + P(\overline{E})\big) \cdot S\,,$$

ricordandosi che la posta S è data dalla somma delle puntate A e B, si ottiene

$$P(E) + P(\overline{E}) = 1\,. \qquad (2.12)$$

Questa è, al pari della (2.10) una condizione che deve valere sempre per qualsiasi evento.

È importante notare che, anche se non è stato implicitamente indicato, la (2.12) è valida solo se le probabilità si riferiscono allo stesso stato di informazione:

$$P(E\,|\,I) + P(\overline{E}\,|\,I) = 1\,, \qquad (2.13)$$

mentre non è in genere vero per le generiche $P(E\,|\,I_1)$ e $P(\overline{E}\,|\,I_2)$.

Dal rapporto delle (2.8) e (2.11) segue inoltre che

$$\frac{A}{B} = \frac{P(E)}{P(\overline{E})}\,, \qquad (2.14)$$

da cui

$$P(E) = \frac{A}{A+B}\,. \qquad (2.15)$$

Se ad esempio uno scommettitore valuta un rapporto di puntate equo di "2 a 1" vuol dire che egli stima una probabilità di 2/3. Anche se questo modo di esprimersi non è molto familiare in Italia, questi rapporti ("odds") sono il modo naturale di esprimere la propria fiducia in un evento nei paesi anglosassoni. Anche le previsioni metereologiche sono talvolta accompagnate da stime di attendibilità del tipo "20 a 1 che domani nevica".

2.10 Definizione soggettiva di probabilità

Abbiamo visto nei paragrafi precedenti come la somma S che si può vincere se l'evento si verifica gioca un ruolo di scala inessenziale nella logica della scommessa stessa. Essa si può quindi porre uguale a 1, ottenendo un valore della puntata numericamente uguale alla probabilità ($A = P(E)\cdot 1$). Ne segue quindi la formulazione ufficiale di *probabilità soggettiva*[10]:

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

Come già accennato precedentemente, questa definizione fornisce un *significato* al numero p, indipendentemente da come esso sia valutato, ovvero se p è prossimo a 1 si è molto sicuri del verificarsi dell'evento, mentre se è prossimo a zero se ne è molto scettici.

In principio sono possibili infiniti modi per calcolarsi il valore di p, ivi compresi conti complicati di meccanica quantistica, simulazioni al computer, etc. Quelli che prenderemo in considerazione in questo testo sono essenzialmente i seguenti:

[10] R. Scozzafava, "Probabilità soggettiva - significato, valutazione, applicazioni", Masson,1997

- valutazioni combinatoriali e basate sulla frequenza, qualora si ritenga che le condizioni per la loro applicazione siano realizzate;
- valutazioni puramente intuitive nei casi unici e irripetibili della vita pratica, compresi quelli delle misure;
- aggiornamento della valutazione iniziale della probabilità - di tipo intuitivo - mediante l'uso del teorema di Bayes, che sarà illustrato nel capitolo **??**.

2.11 ❄ La "definizione ISO"

Nel 1993 l'Organizzazione Internazionale per la Standardizzazione (ISO) ha pubblicato una guida (*"Guide to the expression of uncertainty in measurement"*) per stabilire regole generali per valutare e esprimere l'incertezza della misura applicabili su un vasto spettro di misure. Essa è stata preparata da un gruppo di lavoro di esperti nominati dal Bureau International des Poids et Mesures (BIPM), l'International Electrotechnical Commission (IEC), l'ISO e l'International Organization of Legal Metrology (OIML) e supportata anche dall'International Federation of Clinical Chemistry (IFCC), l'International Union of Pure and Applied Chemistry (IUPAC) e l'International Union of Pure and Applied Physics (IUPAP).

Analizzeremo in dettaglio le raccomandazioni suggerite in tale guida quando svilupperemo la teoria probabilistica delle incertezze di misura. Vedremo allora il ruolo cruciale che avrà l'interpretazione soggettivista della probabilità, ruolo espressamente riconosciuto nella guida ISO (citazione in inglese per evitare distorsioni di traduzione):

> *. . . In contrast to this frequency-based point of view of probability, an equally valid viewpoint is that probability is a measure of the* degree of belief *that an event will occur. For example, suppose one has a chance of winning a small sum of money D and one is a rational bettor. One's degree of belief in event A occurring is $p = 0.5$ if one is indifferent to this two betting choices: (1) receiving D if event A occurs but nothing if it does not occur; (2) receiving D if event A does not occur but nothing if it does occur. Recommandation INC-1 (1980) upon which this Guide rests implicitly adopts such a viewpoint of probability since it views expressions such as equation (E.6) as the appropriate way to calculate the combined standard uncertainty of a result of a measurement.*

Aggiungiamo soltanto alcune note esplicative:

- "equally valid" può essere riduttivo, alla luce di quanto detto nei precedenti paragrafi, ma si tenga conto che l'approccio frequentista è ancora dominante in molti settori, semplicemente per propagazione di consuetudini;
- "small sum" è in relazione alla non uguale percezione del valore dei soldi da parte di diverse persone, problema illustrato fra breve;

- la raccomandazione INC-1 del Bureau International des Poids et Mesures sarà discussa nel capitolo **??**;
- l'equazione (E.6) mostra una propagazione di incertezze dovute a errori statistici e sistematici, del tipo di quelle che vedremo nel capitolo **??**. Anticipiamo che tali espressioni hanno senso soltanto se la probabilità è intesa in modo soggettivo.

2.12 ❄Note sul termine "soggettivo"

Indubbiamente l'aggettivo "soggettivo" spaventa molti, soprattutto coloro che intendono fare scienza inseguendo il sano ideale di "oggettività", così come questo termine è percepito dal senso comune. È quindi opportuno fare alcune precisazioni sull'uso di tale termine in questo contesto.

- Innanzitutto ripetiamo che "soggettivo" sta ad indicare che la valutazione di probabilità dipende dallo stato di informazione dell'individuo - soggetto - che la esegue.
- A volte si ha l'impressione che esista una probabilità oggettiva astratta, ma ci si rende conto che, in genere, questa risponde soltanto a domande generiche, a problemi stereotipati o a casi da libro di testo. Tutte le valutazioni della vita pratica richiedono delle decisioni personali e meditate e delle quali ci si deve assumere la responsabilità.

  Si provi, ad esempio, a rispondere alla seguente domanda: "quanto vale la probabilità che una molecola di gas N_2 che esce da un recipiente ad una certa temperatura abbia una velocità compresa in certo intervallo?" La risposta sembra tranquilla e assolutamente oggettiva: "si prende la distribuzione di Maxwell $f(v)$, si fanno dei conti e si ottiene un numero". Ma questo è un esercizio di Fisica Generale, non la risposta ad un quesito reale che invece potrebbe suonare così: "ti dò un recipiente contenente N_2 e un rivelatore (o i soldi per comprarli). Tu installi l'apparato (o lo fai fare ad una persona di tua fiducia) e poi rispondi alla domanda: quanto vale la probabilità che la prima molecola abbia una velocità in quell'intervallo?". Le cose cambiano drammaticamente: si studia bene il problema, ci si informa dove è stato acquistato il gas, si controlla la temperatura, eventualmente si effettuano stime preliminari e poi - molto probabilmente - si dà un intervallo di valori per $P(\ldots \leq v \leq \ldots)$.
- Soggettivo non significa arbitrario: il ruolo normativo della scommessa coerente obbliga a tener conto di tutte le informazioni a disposizione. Addirittura la conoscenza che altri valutano diversamente la probabilità dello stesso evento è una informazione che andrebbe utilizzata. Ma non tanto per mediare le due stime quanto piuttosto per capire meglio il problema e comportarsi alla fine nel modo più coerente possibile come se veramente si dovesse scommettere del denaro su o contro quell'evento.
- La così detta "oggettività", come è percepita da chi è al di fuori della ricerca scientifica, viene recuperata quando una comunità di esseri razionali condivide lo stesso stato di informazione. Ma anche in questo caso, si dovrebbe parlare più propriamente di *"intersoggettività"*.

- A volte si incontrano delle persone che, riluttanti ad accettare l'interpretazione soggettiva della probabilità ma rendendosi conto dei paradossi delle altre presunte definizioni, affermano di disinteressarsi dei "cavilli filosofici" sulla probabilità, "purché la sappia usare e valutare", inconsapevoli che "rifiutare di fare della filosofia è un atteggiamento che presuppone una certa filosofia". Se tali persone sono disposte ad accettare che quanto più alto è il valore della probabilità tanto maggiore sono le aspettative sul verificarsi dell'evento, e se sono pronte ad assumersi delle responsabilità sulle affermazioni di probabilità, non c'è nessun problema: questa posizione apparentemente agnostica potrebbe venir utilizzata come definizione alternativa del concetto di probabilità.

2.13 ❄Ruolo virtuale della scommessa, valore dei soldi e ordini di grandezza non intuitivamente percepibili

Qualcuno può avere difficoltà ad accettare il ruolo della scommessa nelle valutazioni di probabilità per ragioni assolutamente legittime e quindi sono necessari alcuni chiarimenti, oltre a quelli appena dati.

Tipiche obiezioni possono essere:

a) con chi bisogna scommettere?

b) devo sempre fare scommesse coerenti?

c) anche se ritenessi che la probabilità sia del 99.999 % non accetterei mai di scommettere 99˙999˙000 lire contro 1000 lire;

d) intuitivamente posso ritenere che la probabilità sia molto bassa ($\approx$ 1 %), bassa ($\approx$ 10-20 %), media ($\approx$ 40-60 %), alta ($\approx$ 80-90 %) o molto alta ($\approx$ 99 %), ma come si può decidere se è 43 o 44 %? O se è 99.9 invece di 99.9999 %?

Cerchiamo di rispondere brevemente.

a) La scommessa coerente deve essere intesa come operazionale, sebbene a livello ipotetico. Questo caso non è unico nella scienza.

   - Anche la definizione di sostanza velenosa è definita come "letale <u>se</u> ingerita", ma non si richiede una prova per ogni sostanza definita tale.
   - Un campo elettrico è definito dalla forza per unità di carica, ma sicuramente questo non è il modo di valutare il campo elettrico in prossimità di un elettrone.

   Il ruolo *normativo* della scommessa coerente dovrebbe obbligare chi valuta la probabilità ad essere onesto con sé stesso e con la comunità scientifica. Ad esempio, se uno sperimentatore dichiara che il valore di una grandezza è compreso in un certo intervallo con probabilità del 68 %[11]

[11] Vedremo che in molti casi questo sarà il modo standard di presentare i risultati

dovrebbe essere disposto ad accettare una scommessa di circa 2 a 1 in favore di tale evento, e 1 a 2 contro. Se egli si sente pronto a scommettere pro ma è riluttante a scommettere contro vuol dire che crede in tale risultato molto di più del 68 % dichiarato.

b) Le scommesse coerenti servono solo per valutare la probabilità. Poi nella vita (a parte le tombolate natalizie con i familiari) si fanno scommesse non eque con l'intento di vincere. Però è opportuno stimare al meglio la probabilità per cercare di puntare meno della posta equa se si gioca pro, o far puntare di più se si sta dalla parte del banco.

   Questo è quello che fanno gli allibratori professionisti, chi organizza lotto e lotterie nazionali, chi propone polizze assicurative, ma anche chi gioca semplicemente a scacchi o a scopone con gli amici. Che invece le valutazioni di chi segue i "sistemi per vincere al lotto" siano sbagliate è "dimostrato" sia dai guadagni che lo Stato fa su tali scommesse sia dal dato di fatto che i furbi, quando possono, si mettono a gestire lotterie clandestine e non di certo a giocare...

c) Se si propone ad una persona razionale di giocare ai dadi alla pari $10\,000$ o $100\,000$ lire puntando contro il "6", questa riterrà l'offerta assolutamente vantaggiosa e non esiterà a giocare (a meno che non sospetti qualche trucco, vista l'enorme differenza dal rapporto equo delle puntate). Se però si propone una puntata che corrisponde a qualche anno di stipendio, o alla casa, pur mantenendo la scommessa alla pari, la maggior parte delle persone riterranno la scommessa inaccettabile. La probabilità di 1/6 di "perdere tutto" vince la tentazione di provare a raddoppiare con probabilità 5/6. Questo è dovuto al fatto che "il valore dei soldi" non è lo stesso per tutte le persone e sicuramente non è percepito in modo lineare[12]. In risposta a questa obiezione vale quanto detto a proposito della lettura in chiave ipotetica della scommessa.

   È da notare comunque che scommesse di questo tipo sono comunemente accettate da gestori di lotterie e da assicurazioni (anche se per ovvi motivi con "puntate" non eque) abbastanza coperti finanziariamente da poter affrontare il rischio di momentanee "sovrafluttuazioni di sfortuna".

d) In effetti è vero che la probabilità soggettiva valutata in modo semplicemente intuitivo è rozza, e non potrebbe essere altrimenti. Ma vedremo dei metodi che permetteranno, a partire da stime approssimative, di riaggiornare la probabilità ottenendo valori spesso ben precisi. Questi metodi stanno ai calibri come la percezione intuitiva di probabilità sta a quella di spazio. La stessa analogia vale per valori di probabilità molti prossimi a uno o a zero. Situazioni analoghe nel dominio dello spazio sono le distanze astronomiche o nucleari.

   A questo proposito hanno un ruolo importante i semplici problemi delle urne con palline bianche e nere, oppure con numeri casuali distribuiti

[12] Questo è la ragione per cui facciamo polizze assicurative con le quali scarichiamo grandi rischi di piccolissima probabilità a chi è più coperto finanziariamente (che siccome non fa di certo scommesse eque ci guadagna nel cautelare i clienti da improbabili rovine).

uniformemente fra 0 e 1. Essi rappresentano delle scale "oggettive" di probabilità (leggi: "sulle quali non possiamo non essere tutti d'accordo") con le quali confrontare valutazioni puramente intuitive. Quindi una probabilità del 70 % dovrebbe dare circa la stessa sicurezza di quando si spera di poter estrarre una pallina bianca da un'urna che contiene 70 palline bianche e 30 palline nere. Se questo non è vero vuol dire che 70 % non corrisponde all'effettiva credibilità dell'evento.

2.14 ⥁ Speranza matematica e previsione di vincita

Il concetto della scommessa reversibile può essere ribadito pensando che se si è incerti se puntare pro o contro l'evento E - almeno dal punto di vista ipotetico - significa che A e $P(E) \cdot S$ vengono considerati di pari valore e quindi interscambiabili: si può versare la cifra "sicura" A con la speranza di poter ottenere S con probabilità $P(E)$ (se E si verifica); indifferentemente si può accettare la cifra sicura A con il rischio di pagare S con probabilità $P(E)$.

Siccome non sempre le scommesse sono eque, è interessante chiamare *speranza matematica (di vincita)*, o *valore atteso (di vincita)*, la somma che si spera di vincere per la probabilità di vincita. Chiamiamo questa quantità $M(= P(E) \cdot S)$, per distinguerla dalla generica posta A che chi organizza il gioco può richiedere di pagare. Se A è uguale a M il gioco è equo. Quindi il rapporto A/M caratterizza l'equità del gioco. Facciamo degli esempi.

- Roulette[13], giocando al "rosso e nero" (ad esempio E= "rosso"):

$$\begin{aligned} P(E) &= \frac{18}{37} \\ S &= 2\,A \\ \frac{M}{A} &= \frac{36}{37} = 0.973\ (97.3\,\%)\,. \end{aligned}$$

  Il gioco è evidentemente non equo, ma non troppo sbilanciato.

- Roulette, puntando su un numero (es. E = "esce il 24"):

$$\begin{aligned} P(E) &= \frac{1}{37} \\ S &= 36\,A \\ \frac{M}{A} &= \frac{36}{37} = 0.973\ (97.3\,\%)\,. \end{aligned}$$

  In effetti tutte le vincite della roulette sono calcolate in modo tale da mantenere un rapporto costante fra puntata e speranza matematica.

[13]Si ricorda che alla roulette ci sono 37 numeri, 1-36 più lo zero. Dei numeri da 1 a 36 la metà sono rossi e la metà neri (senza un ordine semplice). Lo zero è un numero particolare considerato né pari né dispari, né rosso né nero. Se si gioca puntando sui pari o sui dispari, o sul colore si vince una somma totale che è il doppio della puntata, ovvero la vincita netta è pari alla puntata stessa. Se si punta su un numero singolo si vince globalmente 36 volte la puntata. Sono anche possibili molte altre combinazioni di gioco sulle quali non ci possiamo soffermare. Si ricorda inoltre che esistono al mondo altre versioni di roulette.

- Lotto, puntando <u>su un numero</u> su una ruota (es. E = "esce il 25 sulla ruota di Bari"):

$$\begin{aligned} P(E) &= \frac{5}{90} = \frac{1}{18} \\ S &= 10.9\,A \\ \frac{M}{A} &= \frac{10.9}{18} = 0.605\ (60.5\,\%)\,. \end{aligned}$$

  Il banco del lotto vince mediamente quasi il 40 % delle puntate[14], contro il 2.7 di quello della roulette.

Come si vede, la conoscenza della speranza matematica permette di giudicare l'equità di un gioco d'azzardo, cioè caratterizza il gioco, anche senza che se ne conoscano in dettaglio le regole. Si vede, ad esempio, che il gioco della roulette è molto più equo del lotto. Anzi si potrebbe dire per inciso che sembra anche "troppo equo", nel senso che la speranza matematica differisce soltanto del 2.7 % dalla puntata. Presumibilmente bisogna tener conto che, anche se il gioco è quasi equo dal punto di vista probabilistico, non lo è dal punto di vista psicologico, e infatti quasi sempre i giocatori seguitano a giocare anche dopo che hanno vinto, e spesso finiscono per perdere tutta la cifra che erano disposti a giocare e sono costretti a smettere (esiste un classico problema del calcolo delle probabilità detto della "rovina del giocatore").

2.15 ❄Previsione di guadagno e decisioni

La quantità $M = P(E) \cdot S$ può essere riscritta nel seguente modo

$$\begin{aligned} M &= P(E) \cdot S + (1 - P(E)) \cdot 0 \\ &= P(E) \cdot S + P(\overline{E}) \cdot 0\,, \end{aligned}$$

mettendo in luce che essa è ottenuta dalla somma degli importi che si possono vincere se si verifica un certo evento, ciascuno moltiplicato per sua probabilità di avvenire.

Se si eseguono contemporaneamente tante scommesse su tanti eventi E_1, E_2,..., E_n, la speranza matematica è la somma delle speranza matematiche di ciascuna delle scommesse. Includendo anche l'evento $E_\circ$ ("non si verifica nessuno degli altri n eventi"), con vincita nulla ($S_\circ$), otteniamo

$$M = \sum_{i=0}^{n} P(E_i) \cdot S_i\,. \tag{2.16}$$

Si noti che non è richiesto che gli E_i siano incompatibili fra di loro. Ad esempio, giocando alla roulette si può puntare contemporaneamente sul "2", sul "3", sui pari e sui rossi. Se le puntate sono rispettivamente 10, 10, 100 e 200 mila lire la speranza matematica sarà pari (in migliaia di lire) a

$$M = \frac{1}{37} \times 360 + \frac{1}{37} \times 360 + \frac{18}{37} \times 200 + \frac{18}{37} \times 400\,.$$

[14] L'estratto semplice è pagato 11.232 volte la puntata, meno il 3 %.

Per l'ovvio significato che appare in questi esempi la speranza matematica è anche chiamata - sono questi i termini attualmente preferiti - *valore atteso*, o *previsione*, di vincita.

Questi termini potrebbero generare confusione se non li si pensa accompagnati dall'aggettivo "probabilistico", e si intendesse che vogliano dire "predizioni" in termini di certezza. Non si deve nemmeno intendere che essi si riferiscano agli importi che effettivamente si possono vincere (0, S_1, S_2, etc). L'esempio della roulette rende abbastanza bene l'idea. Le possibili vincite sono 0 (non si realizza nessuna delle puntate), 200 mila (un pari che non sia il "2" e nemmeno "rosso"), 560 o 960 mila se esce il "2" (nelle due ipotesi che questo numero sia definito nero o rosso), e così via.

Quindi l'attesa, o previsione, probabilistica dà l'idea di una sorta di vincita media, e su questo concetto ritorneremo quando si parlerà delle variabili casuali. Quello che è importante è che, in scommesse eque, la previsione di vincita deve essere uguale alla somma delle poste.

Estendiamo il concetto di scommessa a situazioni più complicate in cui si gioca contemporaneamente pro e contro diversi eventi. Ovvero per alcuni eventi S_i può essere positivo e per altri può essere negativo. Questo è quello che si verifica tutti i giorni quando possono accadere degli eventi in grado di produrre vantaggi o svantaggi a seconda che si verifichino oppure no: affrontare un viaggio per procurarsi un lavoro comporta una spesa che potrebbe essere compensata dall'assunzione; viaggiare sull'autobus senza biglietto ha un vantaggio immediato che però può trasformarsi in una perdita se passa il controllore. Anche la roulette può essere vista in questo modo considerando la puntata A una perdita quando non si verifica nessuno degli eventi sui quali si è scommesso.

In questi casi, al posto della previsione di vincita (calcolata senza tener conto della posta pagata) è preferibile parlare della previsione di *guadagno*, ottenuta con la stessa formula di M e indicata con $\mathbb{P}(G)$ e calcolata in due modi alternativi ma equivalenti:

- gli eventi E_i devono essere tali per cui <u>almeno uno</u> deve accadere (quando c'è un evento e il suo contrario questo è automatico); le vincite vanno considerate come *vincite nette* S_N, ovvero già detratte delle puntate; le puntate sono considerate negative e associate agli eventi opposti a quelli sui quali si è scommesso;

- alternativamente si possono considerare le vincite lorde (come in M) e sottarre la somma delle puntate (A_{tot}) (con probabilità 1 in quanto esse si considerano versate in anticipo);

Otteniamo quindi:

$$\mathbb{P}(G) = \sum_i P(E_i) \cdot S_{N_i} \tag{2.17}$$

o

$$\mathbb{P}(G) = \sum_i P(E_i) \cdot S_i - A_{tot}\,. \tag{2.18}$$

Ne segue che

- le scommesse eque, o coerenti, hanno una previsione (probabilistica) di guadagno nulla e le *decisioni* su quale verso della scommessa preferire sono *indifferenti*;

- le *decisioni vantaggiose* hanno una previsione di guadagno positiva;

Come esempio numerico riprendiamo il caso della roulette puntando 1000 lire su un solo numero. I due modi per calcolare la previsione di guadagno sono:

$$\begin{aligned}
\mathrm{I\!P}(G) &= \frac{1}{37}\times(36\times 1000-1000)+\frac{36}{37}\times(-1000) = -\frac{1}{37}\times 1000 \\
\mathrm{I\!P}(G) &= \frac{1}{37}\times(36\times 1000)+\frac{36}{37}\times 0-1000 = -\frac{1}{37}\times 1000\,.
\end{aligned}$$

Come esempio di decisione, consideriamo la situazione di un automobilista che deve scegliere fra pagare il biglietto del parchimetro, di costo B, o rischiare una multa di importo C se arriva il vigile, evento la cui probabilità $P(V\,|\,I)$ è subordinata allo stato di informazione su città, quartiere, ora, traffico, situazione meteorologica, etc. La previsione di guadagno è

$$\mathrm{I\!P}(G) \;=\; P(\overline{V}\,|\,I)\cdot B + P(V\,|\,I)\cdot(-C+B) = B - C\cdot P(V\,|\,I)\,.$$

Quindi la multa è veramente scoraggiante se

$$C \;>\; \frac{B}{P(V\,|\,I)}\,.$$

2.16 ❄Decisioni vantaggiose e etica della ricerca

Terminiamo questa discussione sulle decisioni con alcune doverose delucidazioni sul percorso prettamente pecuniario sul quale sembra che ci stiamo indirizzando. Come detto precedentemente, il "valore" dei soldi è abbastanza soggettivo e quindi si può scegliere di scambiarli con beni (che per definizione non hanno un valore fisso, ma solo uno di mercato), piaceri, emozioni, etc. Quindi nella scelta delle decisioni i soldi, che ci sono serviti a schematizzare rozzamente il problema, possono essere sostituiti da altre scale di valori.

Effettivamente queste considerazioni entrano in gioco continuamente. In particolare, basta farsi due rapidi conti sulle probabilità di multa o di altre sanzioni per capire che sarebbe preferibile non pagare parchimetri, biglietti dell'autobus e tasse (gli evasori lo sanno perfettamente), o addirittura compiere atti criminali. Per fortuna il senso civico, o semplicemente un ragionamento utilitaristico più lungimirante, legato alla consapevolezza che questi piccoli benefici individuali si possono ripercuotere negativamente nel futuro, fanno sì che la convivenza civile sia ancora possibile e produttiva.

Tornando all'argomento più vicino a questo testo e al fatto che anche nella valutazione delle incertezza di misura (e quindi anche nell'eventuale annuncio di grandi risultati scientifici) è molto importante attenersi a regole di etica, citiamo quanto riportato dall'Organizzazione Internazionale per la Standardizzazione (ISO) a conclusione della già citata "Guida all'espressione dell'incertezza di misura".

> *Although this Guide provides a framework for assessing uncertainty, it cannot substitute for critical thinking, intellectual honesty, and professional skill. The evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement. The quality and utility of the uncertainty quoted for the result of a measurement therefore ultimately depend on the understanding, critical analysis, and integrity of those who contribute to the assignment of its value.*

Questo invito è in perfetto accordo con la concezione della probabilità soggettiva, accompagnata dalla regola normativa della scommessa coerente.

2.17 ❄Regola di penalizzazione - il bastone e la carota

C'è ancora un punto interessante da discutere, legato ai guadagni e alle stime puramente soggettive della probabilità. Nei casi in cui non si possa fare uso né di ragionamenti di simmetria né di dati sperimentali per riaggiornare le nostre probabilità soggettive secondo lo schema che mostreremo nel capitolo **??**, bisogna affidarsi all' *"intuito"*. L'esperienza mostra come tali valutazioni di probabilità non siano affatto arbitrarie e, in effetti, ci sono esperti che tendono ad "indovinare" le previsioni meglio della gente comune.

Il concetto stesso di "esperto" implica una solida esperienza nel proprio settore, sia esso calcio, poker, finanza, politica, marketing, etc. Una delle caratteristiche di queste persone è quella di avere, oltre ad una conoscenza approfondita del problema, una notevole abilità a imparare dall'esperienza, sia in caso di pronostici "indovinati" che "sbagliati". È chiaro che, essendo i pronostici di tipo probabilistico (nessun consulente è disposto a mettere la mano sul fuoco che quello che riterrà molto probabile si verificherà veramente) il concetto di indovinare o meno è legato a quanto spesso non accade un evento giudicato molto probabile, o quanto spesso ne accade uno poco probabile. Quindi una possibile misura della qualità della previsione potrebbe essere la differenza fra la probabilità dell'evento e il suo valore logico (0 o 1) una volta che esso passa dallo stato di incertezza a quello di certezza. Quindi si può ricompensare l'esperto in base alla sua abilità.

Il metodo più conosciuto, e anche effettivamente applicato in ambienti in cui è importante aguzzare l'ingegno degli esperti è quello della *regola della penalizzazione*. Al posto di ricompense si applicano sempre delle penali, proporzionali al quadrato della differenza fra probabilità e *indicatore* (*di verità*) $|E|$ dell'evento dopo che è passato dallo stato di incertezza a quello di certezza ($|E|$ vale 1 o 0 a seconda che l'evento si verifichi o meno):

$$\text{Penale} = S \cdot (P(E) - |E|)^2 \,. \tag{2.19}$$

S è un fattore di scala tale da dissuadere le persone dal bluffare. La scelta della dipendenza quadratica dallo scarto è dovuto al fatto che essa è quella più semplice che obbliga ad esprimere la valutazione di probabilità della quale si è veramente convinti. Questo aspetto sarà mostrato fra i problemi. Concludiamo con delle osservazioni su tale procedura:

- c'è sempre una penale, a meno che non si verifichi un evento pronosticato con probabilità 1 (o non se ne verifichi uno pronosticato di probabilità 0). La ricompensa va vista come vincere una sfida di minima somma di penali su un lungo periodo o come un importo iniziale al quale vanno sottratte le penali.

- la penalizzazione non va assolutamente intesa come una "valutazione a posteriori" (cosa significherebbe?) della probabilità, ma semplicemente come l'equivalente del classico bastone e carota dei domatori;

- si può dimostrare (vedi fra i problemi) che non conviene mai assegnare un valore di probabilità diverso da quello che si pensa. Ad esempio, assegnando una probabilità 0.5 all'evento testa di una moneta si ha un valore atteso di penale di $0.25\,S$. Nel caso si dichiarasse una probabilità di 0.6 il valore atteso salirebbe a $0.26\,S$.

2.18 Ricapitolando

- La nascita del calcolo delle probabilità è legata ai problemi relativi a giochi d'azzardo e assicurazioni sulla vita, dai quali hanno avuto luogo rispettivamente le valutazioni combinatorie e frequentiste:

$$\begin{aligned} p &= \frac{\text{numero dei casi favorevoli}}{\text{numero dei casi (egualmente) probabili}}, \\ p &\approx \frac{\text{numero di prove con esito favorevole}}{\text{numero di prove totali eseguite nelle stesse condizioni}}. \end{aligned}$$

- La constatazione empirica del fatto che quando è possibile valutare la probabilità con ragionamenti di simmetria si trovano frequenze relative di successo prossime a tale valore, con approssimazione che migliora con l'aumentare del numero di prove, ha condotto alla definizione di probabilità come il limite della frequenza (per $N \to \infty$) e alla concezione oggettivista della probabilità.

- In realtà le "definizioni" combinatorie e frequentiste richiedono la conoscenza del concetto primitivo di probabilità. Inoltre le valutazioni dipendono da condizioni che devono essere verificate di volta in volta dal soggetto che valuta la probabilità. Questa constatazione conduce alla così detta concezione soggettivista della probabilità e al fatto che la probabilità sia sempre condizionata.

- La scommessa coerente (reversibile), seppure da intendere in modo ipotetico, ha il fondamentale ruolo normativo di poter definire la scommessa anche nei casi pratici della vita e della ricerca scientifica che non rientrano nella schematizzazione dei casi ugualmente possibili o degli infiniti esperimenti eseguiti nelle stesse condizioni. Quando le condizioni sulle quali tali "definizioni" si basano sono ritenute valide, queste vengono recuperate, purché interpretate nella logica della scommessa ("se dai conti mi viene un valore alto, ma in realtà non mi sento sufficientemente sicuro sul verificarsi dell'evento, c'è qualcosa non va con la valutazione").

- Se la scommessa coerente suggerisce puntate pro e contro un evento rispettivamente di A a B vuol dire che il valore di probabilità è stimato essere

$$p = \frac{A}{A+B}\,.$$

- Le scommesse coerenti sono anche dette eque. In esse la posta deve essere uguale al valore atteso di vincita (probabilità $\times$ possibile vincita), detto anche speranza matematica.

- La previsione di guadagno è un estensione della previsione di vincita, e si ottiene sommando tutti i possibili ricavi e perdite, ciascuno moltiplicato con la propria probabilità di accadere. Questo concetto sta alla base della teoria delle decisioni.

2.19 Problemi

Nota: alcuni di questi problemi possono richiedere elementi di calcolo combinatorio facilmente ricavabili intuitivamente.

1. Si lanciano due monete e si è interessati ai seguenti eventi $E_1 = TT$, $E_2 = CC$ e $E_3 = TC$, ove T e C stanno per testa e croce. Quanto vale la probabilità dei tre eventi?

2. Quanto vale la probabilità che, lanciando due dadi regolari, esca un doppio 6?

3. Quanto vale la probabilità che, lanciando due dadi regolari, esca 4 in un dado e 5 nell'altro?

4. Il telecomando di un cancello elettrico ha un codice selezionabile mediante combinazione di dieci microinterruttori, ciascuno dei quali puo' assumere due posizioni. Quanto vale la probabilità che una persona che possiede un telecomando in grado di inviare impulsi radio alla frequenza corretta ma che non conosce il codice segreto riesca ad aprire il cancello al primo tentativo?

5. Una valigia ha due chiusure, ciascuna delle quali ha una serratura a codice decimale di tre cifre. Un'altra ha una sola serratura a codice decimale da sei cifre. Quale delle due valigie è più sicura?

6. Giocando all'enalotto[15] è più probabile che esca una colonna con tutti i simboli uguali (ad esempio 111111111111) oppure la colonna 11X21X1X12X2?

7. La poesia "Cielo e mare" di Giuseppe Ungaretti, che si esprime in

   M'illumino d'immenso

   è indubbiamente una delle più brevi della letteratura italiana. Quanto vale la probabilità di ottenere il testo della composizione battendo a caso su una tastiera di computer da 106 tasti? (Considerare equivalenti le lettere minuscole e maiuscole.)

8. Gettando tre dadi si può ottenere il numero 11 con le seguenti 6 triplette: (6,3,2), (6,4,1), (5,5,1), (5,4,2), (5,3,3) e (4,4,3). Anche il numero 12 può ottenere ottenuto con sei diverse triplette: (6,5,1), (6,4,2), (6,3,3), (5,5,2), (5,4,3) e (4,4,4). Quali dei due numeri ha maggiore probabilità di verificarsi?

[15]Si ricorda che la colonna vincente dell'enalotto è ottenuta da 12 estrazione di numeri del lotto, assegnando i simboli "1", "X" e "2" a seconda che i numeri usciti siano compresi rispettivamente fra 1 e 30, fra 31 e 60 e fra 61 e 90.

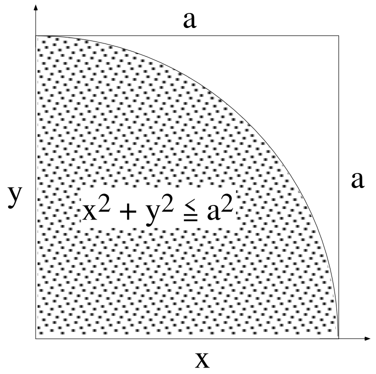

Figura 2.2: Quanto vale la probabilità che scegliendo a caso un punto nel quadrato di lato a esso cada dentro il quarto di cerchio disegnato?

9. Supponiamo di avere un quadrato di lato a ed un cerchio di raggio $r = a$ il cui centro coincide con uno dei vertici del quadrato (figura 2.2). Supponiamo di generare dei punti distribuiti uniformemente all'interno del quadrato. Quanto vale la probabilità che un punto cada all'interno del quarto di cerchio contenuto nel quadrato?

10. In una citta il numero di giovani che effettua la visita medica per il servizio militare in quattro anni successivi sono: 5467, 5511, 5437 e 5481. Di questi, coloro che non risultano abili al servizio sono, rispettivamente nei vari anni, 394, 386, 360 e 386. Stimare la probabilità che il cinquantesimo giovane visitato l'anno successivo risulti abile.

11. Viene registrata, per sei anni successivi, l'incidenza di una certa malattia fra la popolazione di un certo paese, ottenendo i seguenti valori: $0.50\,\%$, $0.85\,\%$, $0.80\,\%$, $1.31\,\%$, $1.26\,\%$, $1.72\,\%$. Quanto vale la probabilità che una persona sia affetta da quella malattia durante l'anno successivo?

12. Una persona si reca in una località di mare il 4 aprile di un certo anno e trova un tempo splendido che, a detta dei locali, dura ormai da oltre una settimana ed è previsto mantenersi per tutta la settimana successiva, come mostrato anche dalle foto da satellite. La sera, mentre si gode un eccezionale tramonto con degli amici, partecipa ad una discussione sulle previsioni del tempo. Alcuni affermano infatti che il giorno dopo potrebbe piovere in quanto negli ultimi 10 anni è sempre piovuto in occasione del 5 aprile. In base a queste informazioni, quanto vale la probabilità che piova il giorno dopo?

13. Consideriamo il quarto pronostico di un concorso di totocalcio, quanto vale la probabilità che si verifichi "1", "X" o "2"?

14. Rispondere qualitativamente al problema precedente sapendo che tale pronostico è relativo al girone di ritorno, in cui la prima squadra in classifica ospita l'ultima.

15. Come varia la valutazione della probabilità del problema precedente se si venisse a conoscenza che da 10 giornate si è verificato sempre "1" al quarto pronostico?

16. Ammettiamo di sapere che negli ultimi anni la percentuale di studenti che supera l'esame relativo a questo corso sia stata del $63.5\,\%$ e che essa si sia mantenuta costante e indipendente dalle sessioni. Se Tu (nel senso di te e non di un'altra persona) dovessi sostenere tale esame Domani (nel senso di domani …), in quanto stimi la probabilità di essere promosso?

17. Due semplici casi per i quali è facile dare una risposta banale, più complicato darne una "seria". Un nuovo farmaco viene somministrato a sei persone affette da una certa malattia e tutte guariscono: qual'è l'efficienza di tale farmaco per curare la malattia in oggetto? Rispondere alla stessa domanda nel caso che un altro farmaco venga somministrato a tre persone affette da un'altra malattia e nessuna guarisca.

18. Secondo la teoria dell'ereditarietà di Mendel, il carattere è determinato da geni *alleli* (geni omologhi con effetti differenti). Nel caso dei piselli, per esempio, i geni che determinano il colore dei semi si presentano in due alleli di cui uno *dominante* e l'altro *recessivo*. Se nella coppia di geni è presente almeno un gene dominante la pianta ha i semi gialli, altrimenti verdi. Indichiamo con G e g i due alleli, dove la notazione g ricorda che il carattere "seme verde" (ovvero "seme non giallo") è recessivo. In ogni processo di riproduzione ciascuno dei genitori trasmette, con pari probabilità, uno dei suoi due geni.

    (a) Se piselli omozigoti (ovvero con alleli identici) del tipo $(G,\,G)$ sono incrociati con piselli del tipo $(g,\,g)$, quanto vale la probabilità di ottenere discendenti aventi i genotipi $(G,\,G)$, $(G,\,g)$ e $(g,\,g)$. Quanto vale la probabilità che essi presentino il carattere "seme verde"?

    (b) Supponiamo di lasciare incrociare fra di loro gli individui della prima generazione. Quanto vale la probabilità che nella seconda generazione si ottengano piselli dai semi gialli e verdi?

    (c) Se i piselli ottenuti nella prima generazione sono incrociati con piselli dai semi verdi, quanto vale la probabilità di ottenere piselli dai semi verdi?

    (d) Se i piselli gialli della seconda generazione sono incrociati con piselli verdi, quanto vale la probabilità di ottenere piselli dai semi verdi?

    (e) Supponiamo di lasciar incrociare fra loro i soli piselli dai semi gialli ottenuti nella seconda generazione e di continuare la selezione nelle generazioni successive. Quante altre generazioni sono necessarie per ottenere piante con almeno il $99.9\,\%$ della specie dai semi gialli?

19. Continuando sul tema dell'esercizio precedente: non tutte le coppie di alleli seguono lo schema dominante/recessivo. Per esempio nella razza di bovini *short horn* il colore del manto è determinato da una coppia di alleli: la combinazione $(R,\,R)$ produce individui rossi, la combinazione $(r,\,r)$ individui bianchi e la combinazione $(R,\,r)$ individui chiazzati, detti *roani*.

    - Se si incrocia un toro bianco con una mucca rossa, quanto vale la probabilità di ottenere un vitello bianco?
    - Quanto vale la probabilità di ottenere un roano dall'incrocio di due roani?
    - Supponiamo di ottenere dall'incrocio fra un toro e una mucca roani un certo numero di vitelli. Successivamente, ad ogni nuova generazione, facciamo in modo far accoppiare fra di loro soltanto animali dello stesso colore. Quanto vale la percentuale di roani alla prima, seconda, terza e quarta generazione?

20. Un circolo sportivo ha 1360 soci adulti, di cui: 400 uomini sposati e 350 celibi; 250 donne sposate e 360 nubili. Quanto vale la probabilità che, scelta a caso la cartellina di un socio, questo sia donna? Quanto vale la probabilità che sia donna se si sapesse che la persona in questione è sposata? Quanto vale invece la probabilità che la persona sia sposata sapendo che è un uomo?

21. Tre signore ("bionda", "mora" e "rossa") si incontrano in treno e, conversando, si scambiano delle informazioni sulla loro vita privata. Dopo un po' entra nello stesso scompartimento anche

un matematico. Questi si siede, ascolta distrattamente la conversazione fra le signore e, dopo un po', scende alla fermata successiva a quella nella quale era salito. Mentre si allontana ripensa un'attimo al fatto che, da quanto ha sentito, tutte le signore hanno due figli e che, inoltre: la bionda ha almeno un figlio maschio; il figlio maggiore della mora è maschio. Ne conclude che la probabilità che le signore abbiano entrambi i figli maschi è diversa nei tre casi. Ha ragione?

22. Un liceo ha 1000 studenti, di cui 600 sono maschi e 400 femmine. L' 80% delle ragazze fuma, mentre l'80% dei ragazzi non fuma. Quanto vale la probabilità che uno studente scelto a caso risulti maschio e fumatore? Incontrando lungo il corridoio uno studente, qual'è la probabilità che sia un fumatore? Sapendo invece che uno studente fuma, quanto vale la probabilità che sia un ragazzo?

23. Un modo per far scegliere al caso colui che deve fare qualcosa, molto più usato del lancio della monetina, è il "pari o dispari" giocato con le dita di una mano. Un ragazzo si accorge che sia lui che i suoi amici raramente tirano lo zero, mentre i numeri da 1 a 5 si presentano all'incirca con la stessa frequenza. Stanti queste ipotesi, se il ragazzo tira a caso un numero compreso fra 1 e 5, è conveniente che egli punti sul pari o sul dispari? È possibile sviluppare una strategia che aumenti ancora di più la probabilità di vincita?

24. Durante una trasmissione televisiva due concorrenti hanno la possibilità di vincere un premio. Il conduttore mostra loro tre scatole, spiega che in una di esse c'è l'assegno e che le altre sono vuote. Quindi li invita a scegliere ciascuno una scatola. Successivamente il primo concorrente apre la sua scatola e questa risulta vuota. Il conduttore offre allora all'altro concorrente la possibilità di scambiare la scatola che egli ha scelto con la terza rimasta. L'offerta è vantaggiosa, svantaggiosa o indifferente?

25. Variazione sul tema. Durante una trasmissione televisiva un concorrente ha la possibilità di vincere un premio. Il conduttore gli mostra tre scatole, gli spiega che in una di esse c'è l'assegno e che le altre sono vuote e quindi lo invita a scegliere una scatola. Dopo che il concorrente ha scelto la scatola, ma non l'ha ancora aperta, il conduttore dichiara di conoscere in quale scatola è il premio. Dichiara inoltre che aprirà, fra le due scatole non scelte, una scatola che è sicuramente vuota. In effetti la scatola che apre non contiene il premio. A questo punto il conduttore offre al concorrente la possibilità di scambiare la scatola che egli ha scelto con la terza rimasta.

    L'offerta è vantaggiosa, svantaggiosa o indifferente? (Confrontare con il problema precedente e dare almeno una risposta intuitiva.)

26. Una persona ritiene che il prezzo equo di un biglietto che dà diritto ad un premio singolo di un milione di lire sia di 500 lire. Quanti biglietti sono stati venduti?

27. Un allibratore accetta scommesse su un incontro di pugilato fra A e B. Nel caso di vittoria di A egli è disposto a pagare 1.25 volte la puntata. Se vince B pagherà invece 3.5 volte la puntata. Quale pugile viene ritenuto favorito? La scommessa sono eque? Quanto dovrebbero essere le vincite in caso di gioco equo, se l'allibratore ritenesse che le probabilità di vittoria sono 3/4 per A e 1/4 per B?

28. Una persona dichiara che la probabilità che una moneta regolare e lanciata a caso dia testa è pari a 0.8, in quanto la valutazione soggettiva della probabilità gli consente di dire qualsiasi numero compreso fra 0 e 1. Cosa gli si può proporre per mettere alla prova la sua convinzione? E se avesse dichiarato 0.3?

29. Il montepremi di una lotteria di Capodanno è stato il seguente: sei premi principali rispettivamente da 6, 3, 2.5, 2, 1.6 e 1.2 miliardi di lire; 102 premi da 250 milioni; 252 premi da 50 milioni. Sapendo che sono stati venduti 30 milioni di biglietti a 5000 Lire l'uno, trovare il rapporto fra il prezzo del biglietto e la speranza matematica di vincita.

30. Un ragazzo, al riaprirsi delle scuole, effettua un rapido calcolo sui viaggi effettuati con gli autobus urbani nell'ultimo anno e stima di aver effettuato circa 1500 corse, durante le quali ha incontrato solo tre volte il controllore. Sapendo che il costo dell'abbonamento è di 50000 lire al mese, mentre l'eventuale multa costa 60000 lire, troverà più conveniente comprare l'abbonamento o utilizzare i soldi che riceve dalla famiglia per divertirsi?

31. Fra i problemi legati ai giochi d'azzardo dai quali ha tratto origine la teoria del calcolo delle probabilità c'è il caso (sottoposto dal Cavalier de Méré a Pascal nel 1654) della suddivisione della posta nel caso di gioco interrotto prima del termine.
    Supponiamo che si tratti di un gioco in cui valga il numero di partite vinte indipendentemente dal punteggio raggiunto in ciascuna di esse, che non si possano verificare delle patte e che al momento dell'interruzione ai due giocatori (di pari abilità

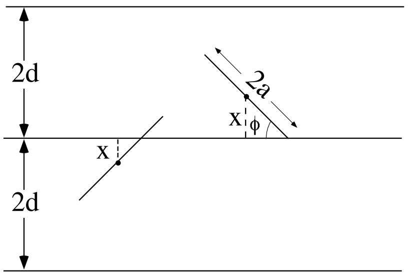

Figura 2.3: Esperimento dell'ago di Buffon.

e identificati con le lettere A e B) manchino rispettivamente 2 e 3 partite per la vittoria. Quale è il criterio per suddividere la posta in palio fra i giocatori?

32. Un altro classico problema è illustrato in figura **??** ("ago di Buffon"). Su un piano orizzontale vengono disegnate delle linee parallele distanti fra loro $2d$. Poi si lascia cadere (a caso) un ago di lunghezza $2a < 2d$. Trovare la probabilità che l'ago tocchi una delle linee.

33. Si sa che una scatola contiene sicuramente due palline, ma esse possono essere entrambe dello stesso colore, oppure di colori diversi. Quanto vale la probabilità che esse siano entrambe dello stesso colore?

34. Ad una persona vengono presentate due buste e gli viene comunicato che ciascuna di esse contiene un assegno. La cifra dei due assegni è ignota, ma sicuramente uno dei due assegni è di un importo 10 volte l'altro. La persona può scegliere a caso una delle buste, aprirla, constatare il contenuto e dopo ha il diritto di decidere se cambiare l'assegno visto con quello dell'altra busta. Ammettiamo, per fissare le idee, che la busta contenga 100˙000 lire. Qual'è la decisione più conveniente? Discutere gli aspetti paradossali del problema.

35. Si immagini che l'offerta descritta nel problema precedente venga effettivamente rivolta ad una persona e che questa sia in serio dubbio se accettare o rifiutare lo scambio. Che probabilità sta - inconsciamente - attribuendo all'eventualità che l'altra busta contenga un assegno da un milione?

36. Sempre sul problema delle due buste: immaginiamo che questa scelta sia proposta ad uno studente dai suoi amici, con la variante che il rapporto fra il valore dei due assegni è di 1000. Lui apre la busta e trova 10000 lire. Cosa deciderà?

37. Si valuti il valore atteso della penalizzazione (con S unitario per semplicità) e lo si confronti con quanto si otterrebbe se una persona pensa che la probabilità vale p ma dichiara $p + \epsilon$.

38. Risolvere lo stesso quesito del problema precedente immaginando una regola di penalizzazione che dipenda dal modulo della differenza fra $P(E)$ e $|E|$.

39. I problemi 9 e 32 non presentano grande difficoltà di comprensione, ma ad essere rigorosi non dovrebbero comparire in questo capitolo. Quale valutazione di probabilità viene usata?

Capitolo 3

Elementi di calcolo combinatorio

Molti dei problemi classici di calcolo delle probabilità si riducono al calcolo dei casi favorevoli e di quelli possibili, supposte le condizioni di equiprobabilità di questi ultimi. Quando le situazioni diventano complicate e i ragionamenti intuitivi non bastano più, la possibilità di risolvere tali problemi è legata all'abilità di eseguire operazioni di calcolo combinatorio.

In questo testo tale problema è abbastanza marginale e quindi questo capitolo può essere saltato da chi non ha interesse a certe valutazioni senza perdere nulla della sostanza di quello che segue. Molti dei problemi che richiederebbero il calcolo combinatorio possono essere affrontati usando il concetto di probabilità condizionata e la formula delle probabilità composte che incontreremo nel prossimo capitolo.

3.1 Problemi elementari tipici

Immaginiamo di avere n elementi *distinguibili*, che possono essere lettere dell'alfabeto ($n = 26$), cifre del sistema decimale ($n = 10$), simboli del totocalcio ("1", "X", "2"), palline numerate (per esempio le 15 palle da biliardo) o un gruppo di persone (una classe di 28 studenti). Supponiamo di voler calcolare:

1. il numero di parole di 4 lettere (considerando anche quelle impronunciabili tipo "qwzg"...); il numero di possibili colonne del totocalcio (ciascuna colonna ha 13 risultati);

2. il numero di parole di 4 lettere ottenute con la condizione che ciascuna lettera non compaia più di una volta; il numero di squadre di calcio che può essere formato con gli studenti, tenendo conto anche degli 11 possibili ruoli diversi che essi possono ricoprire;

3. il numero di partite di biliardo che possono essere giocate se all'inizio di ogni partita si cambia l'ordine delle palle all'interno del triangolo con cui vengono posizionate; il numero di file che possono essere formate dagli studenti prima che ciascuno abbia ricoperto tutte le posizioni della fila (dal numero 1 al numero 28 nell'esempio citato);

4. il numero di squadre di pallavolo (6 giocatori) che possono essere formate con gli studenti della classe, tenendo conto che la differenza di ruolo è inessenziale; il numero di strette di mano che possono essere scambiate fra gli studenti.

Ciascuno di questi gruppi di problemi corrisponde ad un diverso concetto del calcolo combinatorio.

3.2 Disposizioni e combinazioni

3.2.1 Regola fondamentale del calcolo combinatorio

Tutte le regole che seguono si basano su quella di base secondo cui, se un certo oggetto (nel senso matematico) si forma mediante k scelte successive con: n_1 possibilità per la prima scelta; n_2 per la seconda (con le nuove condizioni dovute all'aver effettuato la prima scelta); n_3 per la terza e così via, il numero totale di oggetti è il prodotto

$$n_1\, n_2\, \cdots\, n_k. \tag{3.1}$$

3.2.2 Numero di r-disposizioni di n oggetti

Supponiamo di avere r caselle in ciascuna delle quali può essere collocato uno degli n oggetti, indipendentemente da cosa è stato messo nelle altre. Questa schematizzazione risponde ai problemi del punto 1.

Per calcolare il numero desiderato si ragiona nel modo seguente: ci sono n possibilità per la prima casella e a ciascuna di queste possono essere associate n possibilità per la seconda; a ciascuna coppia di possibilità per le prime due se ne possono associare n per la terza, così via. Quindi:

$$\text{numero di } r\text{-disposizioni di } n \text{ oggetti} = n^r\,. \tag{3.2}$$

Queste disposizioni vengono anche chiamate "disposizioni di n oggetti a gruppi di r *con possibili ripetizioni*".

Si noti che non ci sono limiti superiori al numero intero r, presupponendo che ci sia una disponibilità infinita di ciascuno degli oggetti, come succede per le lettere dell'alfabeto.

Per i due problemi del punto 1 abbiamo: $456\,976$ parole e $1\,594\,323$ colonne.

3.2.3 Numero di r-disposizioni semplici di n oggetti

In alcuni problemi lo stesso oggetto non può essere messo contemporaneamente in più caselle, come succede quando si pensa a disposizioni di persone o di altri oggetti dei quali ne esiste uno per tipo, oppure quando tale condizione è richiesta nel problema. Quindi si ha un primo vincolo dovuto al fatto che il numero di caselle non può essere superiore al numero di oggetti:

$$r \;\leq\; n\,.$$

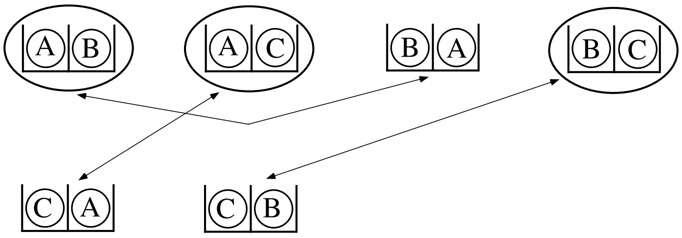

Figura 3.1: Numero di disposizioni semplici (senza ripetizione) di tre oggetti presi due a due. Quelle cerchiate corrispondono ad una scelta di disposizioni dalle quali le altre differiscono per l'ordine degli oggetti (come indicato dalle frecce). Il numero di disposizioni cerchiate corrisponde quindi al numeri di combinazioni di n oggetti presi r a r.

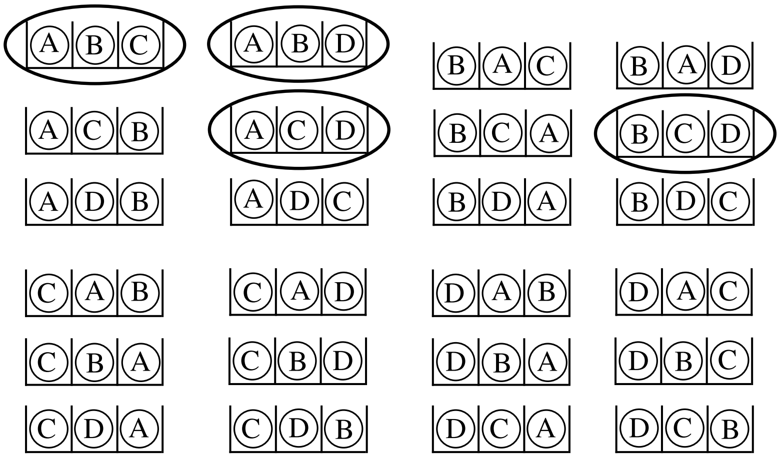

Figura 3.2: Come figura 3.1 per $n = 4$ e $r = 3$.

In questo caso si hanno n possibilità per la prima casella, a ciascuna di queste possono essere associate $n-1$ possibilità per la seconda e così via, fino a $n-r+1$ per la casella numero r (vedi figura 3.1 per $n=3$ e $r=2$ e figura 3.2 per $n=4$ e $r=3$).

$$
\begin{aligned}
(n)_r &= n \cdot (n-1) \cdot (n-2) \; \cdots \; (n-r+1) &(3.3)\\
&= n \cdot (n-1) \cdot (n-2) \; \cdots \; (n-r+1) \cdot \frac{(n-r)!}{(n-r)!} \\
&= \frac{n!}{(n-r)!} \qquad (r \le n) \,. &(3.4)
\end{aligned}
$$

Si è introdotto il simbolo $(n)_r$ per indicare il numero di r-disposizioni semplici (o "*senza possibili ripetizioni*") ed è stato moltiplicato numeratore e denominatore per $(n-r)!$ per semplificare il calcolo. Per riottenere la (3.3) dalla (3.4) anche nel caso $r=n$ (ovvero $(n)_n = n!$) si adotta la convenzione

$$0! = 1\,.$$

Per i problemi del punto 2 otteniamo: $358^{\cdot}800$ parole e $\approx 857^{\cdot}181$ miliardi di squadre.

È da notare come anche l'operazione elementare di conteggio possa essere considerata come una r-disposizione con $r=1$.

3.2.4 Numero di permutazioni di n oggetti

Il numero di permutazioni di n oggetti non è altro che il numero di ordinamenti possibili di n oggetti. In termini del numero di disposizioni, esso può essere visto come il numero di disposizioni ordinate di n oggetti in $r=n$ caselle:

$$
\begin{aligned}
\text{numero di permutazioni di } n \text{ oggetti } [= (n)_n] &= n \cdot (n-1) \; \cdots \; \cdot 1 \\
&= n!\,. &(3.5)
\end{aligned}
$$

Per i due problemi del punto 2 abbiamo: $\approx 1.308 \times 10^{12}$ partite e $\approx 3.05 \times 10^{29}$ file.

Un esempio di permutazioni per $n=3$ palline è mostrato in figura 3.3.

3.2.5 Combinazioni

In alcuni problemi non si ha interesse a distinguere fra le $(n)_r$ disposizioni semplici tutte quelle che contengono gli stessi oggetti ma che differiscono soltanto per l'ordine. Si parla allora di *"combinazioni di n oggetti r a r"*, indicato con

$$\binom{n}{r}\,,$$

e letto *"n sopra r"*. Esso si ottiene da $(n)_r$ dividendolo per il numero di ordinamenti diversi degli r elementi:

$$\binom{n}{r} = \frac{(n)_r}{r!} = \frac{n!}{(n-r)!\,r!} \qquad (r \le n)\,. \tag{3.6}$$

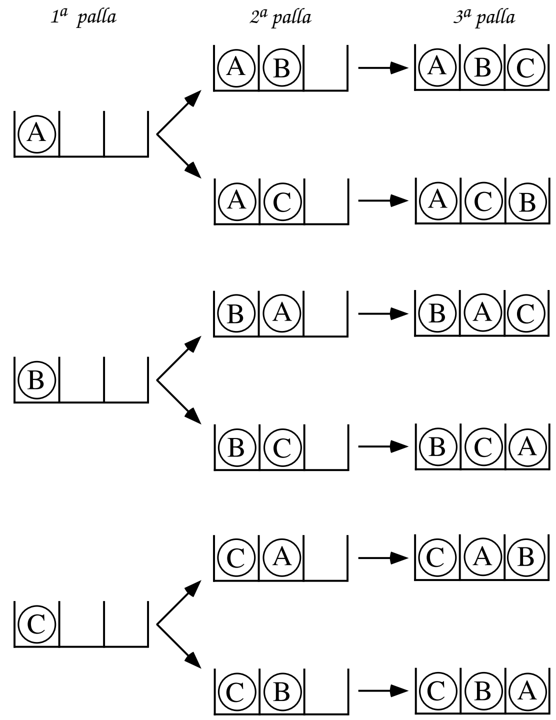

Figura 3.3: Numero di permutazioni di tre oggetti.

Infatti, utilizzando la regola fondamentale del calcolo combinatorio e chiamando provvisoriamente n_x il numero incognito che si vuole calcolare, abbiamo che il numero di possibili scelte di r elementi fra n senza ripetizione è pari al numero di combinazioni di n oggetti r a r, moltiplicato per il numero di ordinamenti degli r elementi di ciascuna combinazione:

$$(n)_r \quad = \quad n_x \cdot r!\,,$$

da cui segue la (3.6). La quantità "n sopra r" gode della proprietà di simmetria

$$\binom{n}{r} = \binom{n}{n-r}, \tag{3.7}$$

ovvero è equivalente considerare gli r oggetti che si prendono o gli $n-r$ che si tralasciano. Vale anche la relazione

$$\binom{n}{n} = \binom{n}{0} = 1\,, \tag{3.8}$$

la quale significa che c'è un solo modo di prendere o di tralasciare tutti di n oggetti se non ci si interessa dell'ordine (formalmente anche "0 sopra 0" vale 1, ma non ha alcun significato applicativo).

Esempi di combinazioni sono mostrati nelle figure 3.1 e 3.2.

Negli esempi del punto 4 otteniamo: $376^{\cdot}740$ squadre e 378 ($= 28 \times 27/2$) strette di mano.

3.2.6 Coefficienti binomiali

Consideriamo lo sviluppo della potenza n-ma del binomio $(a+b)^n$. Scriviamolo in dettaglio per i casi più semplici:

$$
\begin{aligned}
(a+b)^2 &= (a+b)\cdot(a+b) = a^2+2ab+b^2 \\
(a+b)^3 &= (a+b)\cdot(a+b)\cdot(a+b) = a^3+3a^2b+3ab^2+b^3 \\
(a+b)^4 &= (a+b)\cdot(a+b)\cdot(a+b)\cdot(a+b) \\
&= a^4+4a^3b+6a^2b^2+4ab^3+b^4\,.
\end{aligned}
$$

Da questi esempi si capisce già la relazione fra coefficienti dello sviluppo e il concetto di combinazione appena incontrato:

- le potenze di ordine n compaiono una sola volta, in quanto ottenute moltiplicando fra di loro tutti i primi o secondi termini del binomio, e questo può succedere una sola volta;

- le potenze di ordine $n-1$ di a (tanto per fissare le idee) compaiono quando vengono moltiplicate fra di loro $(n-1)$ volte a e 1 volta b; questo può accadere n volte, ovvero il numero di possibili scelte di b sugli n fattori $(a+b)$;

- le potenze $n-2$ di a derivano dalle possibili coppie di b, e così via.

Per questo motivo le espressioni

$$\binom{n}{r}$$

sono anche chiamate *coefficienti binomiali*. L'espressione generale dello sviluppo di un binomio diventa

$$
\begin{aligned}
(a+b)^n &= a^n+\binom{n}{1}a^{n-1}b+\binom{n}{2}a^{n-2}b^2+\cdots \\
&\quad +\binom{n}{n-1}ab^{n-1}++b^n \\
&= \sum_{i=0}^{n}\binom{n}{i}a^{n-i}b^i \qquad (3.9)\\
&= \sum_{i=0}^{n}\frac{n!}{(n-i)!i!}a^ib^{n-i}\,. \qquad (3.10)
\end{aligned}
$$

Si noti il caso particolare di $a=b=1$. Per tali valori $(a+b)^n=2^n$ e la sommatoria diventa:

$$\sum_{i=0}^{n}\binom{n}{i}=2^n\,. \qquad (3.11)$$

La ben nota rappresentazione a triangolo dei coefficienti binomiali (chiamato "di Pascal" o "di Tartaglia") è mostrata in tabella 3.1.

n	Coefficienti binomiali	$\sum_{i=0}^{n}\binom{n}{i}$
0	1	1
1	1 1	2
2	1 2 1	4
3	1 3 3 1	8
4	1 4 6 4 1	16
5	1 5 10 10 5 1	32
6	1 6 15 20 15 6 1	64
7	1 7 21 35 35 21 7 1	128
8	1 8 28 56 70 56 28 8 1	256

Tabella 3.1: Rappresentazione a triangolo dei coefficienti binomiali.

3.2.7 Note su nomenclatura e simbologia

La terminologia del calcolo combinatorio tende a confondere molti (non solo studenti!) e bisogna fare attenzione a capire di cosa si sta parlando, per evitare il rischio di malintesi. Infatti, all'ambiguità intrinseca di termini non intuitivi (combinazioni, permutazioni e disposizioni tendono a confondersi) si aggiunge quella conseguente all'uso, nei vari testi, di nomi e simboli diversi per le stesse quantità:

- A volte le disposizioni semplici sono anche chiamate genericamente *permutazioni* e indicate con ${}_nP_r$:
$$(n)_r \iff {}_nP_r\,;$$
- Le combinazioni sono anche indicate con i simboli ${}_nC_r$ e $C_{n,r}$:
$$\binom{n}{r} \iff {}_nC_r \iff C_{n,r}\,.$$

3.3 Note sul calcolo dei grandi numeri

Nelle applicazioni pratiche può succedere che sia impossibile calcolare i valori di disposizioni, permutazioni e combinazioni usando direttamente le formule riportate. Ad esempio, se si vuole calcolare il numero di quartetti che si possono costituire da un gruppo di 200 persone, le calcolatrici vanno in "overflow" se si tenta di calcolare
$$\frac{200!}{196!4!}\,.$$
.

Lo stesso succede se si tenta di valutare 26^{200} (il numero di frasi diverse di 200 lettere a partire da un alfabeto di 26 lettere). A seconda dei casi si possono adottare diverse tecniche per risolvere il problema.

- Il metodo più semplice, soprattutto quando si ha interesse al risultato esatto (un intero), consiste nel risalire dalle formule compatte a quelle originali. In genere significa tornare dalla (3.4) alla (3.3) e riscrivere, ad esempio,

$$\frac{200!}{196!4!} = \frac{200 \times 199 \times 198 \times 197}{4!} = 64'684'950\,.$$

  Un altro trucco per evitare overflow anche nelle stesse moltiplicazioni, consiste nel fare la divisione dopo qualche moltiplicazione, oppure, soprattutto se si vuole arrivare a numeri interi esatti, conviene espandere anche il fattoriale del denominatore ed effettuare le opportune semplificazioni.

  Questo metodo funziona quando i grandi fattoriali si semplificano e i risultati rientrano nella capacità della calcolatrice.

- Se si hanno numeri diversi da 10 elevati a potenze molto grandi, come nel nostro esempio di 26^{200}, conviene utilizzare le proprietà dei logaritmi[1]. Infatti, valendo

$$\log b^a \;=\; a \log b\,,$$

  ne segue

$$b^a \;=\; 10^{a \log b}\,.$$

  Quindi nel nostro esempio avremmo

$$26^{200} \;=\; 10^{282.9946696} = 10^{282+0.9946696} = 9.878013 \times 10^{282}\,.$$

  Logaritmi importanti che andrebbero ricordati sono quelli in base dieci di 2 e di e, almeno nella loro forma approssimata: $\log 2 = 0.30$ e $\log e = 0.43 \approx 0.4$. Essi permettono infatti di stimare l'ordine di grandezza di potenze (calcolabili o no con la calcolatrice) che possono capitare di frequente. Ad esempio

$$\begin{aligned} 2^{64} &\approx 10^{19} \\ e^{-20} &\approx 10^{-9} \div 10^{-8} \end{aligned}$$

- Per i grandi fattoriali si usa invece la formula approssimata di Stirling:

$$n! \approx \sqrt{2\pi n}\, n^n\, e^{-n}\,, \tag{3.12}$$

  in genere nella sua forma logaritmica poiché quando $n!$ dà overflow sulla calcolatrice a maggior ragione diventa incalcolabile n^n:

$$\log \frac{n!}{\sqrt{2\pi}} \approx \left(n + \frac{1}{2}\right) \log n - n \log e\,. \tag{3.13}$$

[1] Il simbolo log indica il logaritmo decimale (a volte si incontra $\log_{10}$). Il logaritmo naturale verrà indicato con ln.

Ad esempio, provando l'approssimazione con il maggiore fattoriale calcolabile con la maggior parte delle calcolatrici tascabili, abbiamo (con tutte le cifre indicate)

$$\log \frac{69!}{\sqrt{2\pi}} \approxeq 97.83369255 = 97 + 0.83369255\,,$$

da cui

$$69! = \sqrt{2\pi} \times 10^{0.83369255} \times 10^{97} = 1.70915907 \times 10^{98}\,,$$

da confrontare con $1.711224523 \times 10^{98}$ calcolato direttamente. Quindi l'approssimazione è accurata a circa una parte su mille già per $n \gtrsim 70$.

3.4 Ordinamenti, occupazioni ed estrazioni

Gli esempi con i quali sono stati introdotti i concetti di r-disposizioni (con ripetizioni o non) e delle combinazioni erano legati ai possibili "ordinamenti" - per utilizzare un termine generico che comprenda disposizioni, permutazioni e combinazioni.

È interessante notare come si possa arrivare agli stessi concetti combinatori partendo da altri punti di vista: quello dei numeri di occupazione di oggetti in scatole e quello del numero di estrazioni.

r-disposizioni di n oggetti. A seconda dei problemi si può pensare a

- disposizioni di n oggetti a gruppi di r con ripetizioni (secondo l'esempio delle n lettere dell'alfabeto per formare parole di r lettere);
- modi di collocare r palline numerate in n scatole diverse, permettendo a più palline di occupare la stessa scatola;
- estrazioni ordinate e con reintroduzione di r palline da un'urna che contiene n palline distinte.

Sebbene tutti e tre i casi si riconducano alla stessa formula (n^r) in un caso n indica gli "oggetti" (lettere) e r le caselle (posizioni all'interno della parola); nel secondo gli "oggetti" sono le r palline e n stanno per le posizioni ove tali oggetti vanno collocati; nel terzo, infine, n sono gli "oggetti" nell'urna e r sono quelli estratti e poi rimessi dentro.

Un modo di avvicinare i primi due punti di vista è di pensare agli r indicatori di un display che possono "andare" in n stati diversi. (L'ambiguità che crea una certa difficoltà psicologica ha origine nella possibilità di immaginare sia le lettere "andare" (apparire) nelle posizioni dell'indicatore, che gli indicatori "andare" (commutare) nei diversi stati possibili. La figura 3.4 mostra i due punti di vista per $r = 3$ e $n = 10$. Un esempio più dettagliato, con $n = 3$ e $r = 2$ è mostrato in figura 3.5.

r-disposizioni semplici. Analogamente al caso precedente si può pensare a

- disposizioni di n oggetti a gruppi di r, con la condizione che ogni oggetto sia considerato al più una volta;

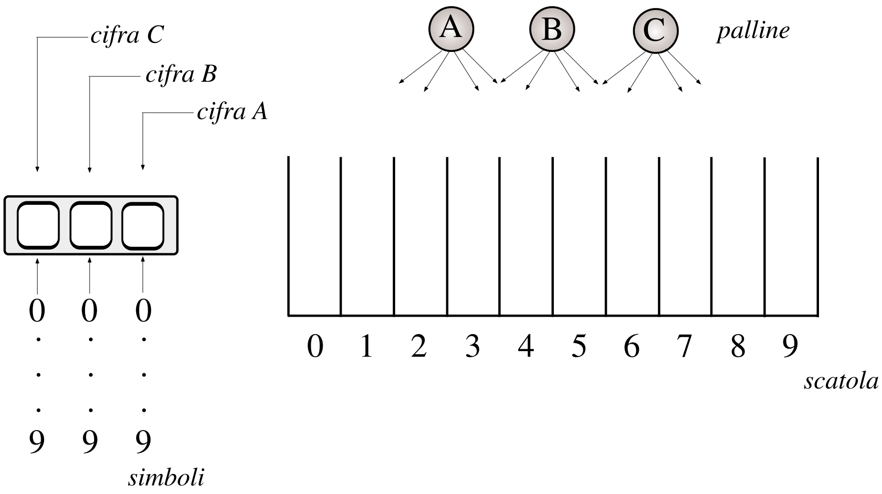

Figura 3.4: Due diversi punti di vista del calcolo combinatorio: ordinamenti e occupazione.

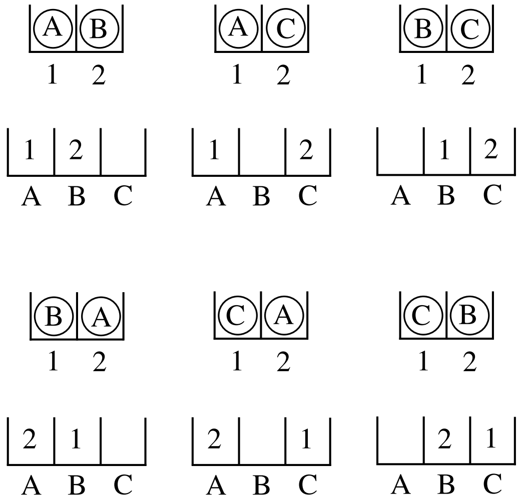

Figura 3.5: Esempio di r-disposizioni con ripetizioni di n oggetti, con $r = 2$ e $n = 3$ (già mostrato in figura 3.1) visto sia in termini di ordinamenti che di occupazioni.

- modi di collocare r palline numerate in n scatole diverse, con la condizione che ogni scatola contenga al più una pallina;
- estrazioni ordinate senza reintroduzione di r palline da un'urna che contiene n palline distinte.

Combinazioni. Anche in questo caso possiamo avere:

- disposizioni di n oggetti a gruppi di r, con la condizione che ogni oggetto sia considerato al più una volta;
- modi di collocare r palline indistinguibili in n scatole diverse, con la condizione che ogni scatola contenga al più una pallina; le diverse configurazioni sono distinte soltanto dallo stato di occupazione della scatola;
- estrazioni senza reintroduzione di r palline da un'urna che contiene n palline distinte.

3.5 Alcuni esempi classici

Mostriamo ora alcuni esempi classici di applicazioni del calcolo combinatorio:

1. **P.** Massimo numero intero senza segno rappresentabile con una parola di un computer a 32 bit.

   R. È equivalente al numero di parole di 32 lettere che si possono formare da un alfabeto di 2 lettere (0, 1): $2^{32} = 4294967295$.

2. **P.** Una moneta regolare viene lanciata 50 volte. È più probabile la sequenza con 50 teste o la sequenza con testa e croce alternate?

   R. Si tratta di r-disposizione di $n = 2$ oggetti ("T", "C") con $r = 50$. Il loro numero è $2^{50} \approx 1.13 \cdot 10^{15}$. Assumendo l'indipendenza dei lanci ogni sequenza ha la stessa probabilità ($\approx 8.88 \cdot 10^{-16}$).

3. **P.** Perché molti sono convinti che la sequenza che ha il $50\,\%$ di teste e il $50\,\%$ di croci sia più probabile di quella con tutte teste?

   R. Si confonde la probabilità della singola sequenza con la probabilità di una qualsiasi sequenza che abbia metà teste e metà croci. Il numero totale di queste è dato da $\binom{50}{25} = 1.26 \cdot 10^{14}$. Queste sono in effetti l'$11.2\,\%$ del totale, ma purtroppo per vincere bisogna indovinare quella giusta.

4. **P.** Calcolare la probabilità di estratto semplice, ambo, terna, quaderna e cinquina su una ruota del lotto (indipendentemente dall'ordine di estrazione).

   R. Il lotto ha $N = 90$ "palline numerate" (evitiamo di chiamarle "numeri" per evitare confusione) delle quali ne vengono sorteggiate $n = 5$. Il numero di gruppi di r elementi (con $r = 1, 2, 3, \ldots$ per singolo estratto, ambo, terna, etc.) che si possono formare con le N palline è:

   $$\binom{90}{r}.$$

r	5	10	15	20	23	30	41	46	57	70
P (%)	2.7	11.7	25.3	41.1	50.7	70.6	90.3	94.8	99.0	99.9

Tabella 3.2: Probabilità che in un gruppo di r persone almeno due abbiano il compleanno lo stesso giorno. Il calcolo è basato sull'equiprobabilità delle nascite e trascurando l'effetto degli anni bisestili

Di questi gruppi possibili, quelli favorevoli per vincere sono quelli che è possibile formare con le 5 palline estratte. Ne segue

$$P(r) = \frac{\binom{5}{r}}{\binom{90}{r}} = \frac{5 \times \cdots \times (5-r+1)}{90 \times \cdots \times (90-r+1)}\,, \tag{3.14}$$

da cui segue la formula ricorsiva

$$P(r) \;=\; P(r-1)\cdot\frac{6-r}{91-r} \qquad r = 2, 3, \ldots, 5\,,$$

con $P(r=1) = 1/18$.

È interessante mostrare anche un altro modo di ragionare per ottenere lo stesso risultato: ci sono

$$\binom{90}{5}$$

possibili cinquine. Di queste ce ne sono

$$\binom{90-r}{5-r}$$

che contengono le r palline della scommessa (pari al numero di combinazioni delle rimanenti $90-r$ palline nelle $5-r$ posizioni che non interessano ai fini della scommessa). Ne segue

$$P(r) = \frac{\binom{90-r}{5-r}}{\binom{90}{5}}\,, \tag{3.15}$$

che ovviamente dà gli stessi risultati della (3.14).

5. **P.** Quanto vale la probabilità che in un gruppo di r persone ce ne siano almeno due che hanno il compleanno lo stesso giorno? (Si assuma una distribuzione uniforme nelle nascite nei 365 giorni dell'anno e si trascurino per semplicità i bisestili.)

   R. Conviene partire dalla probabilità che tutte le r persone abbiano compleanni diversi, ovvero che r giorni estratti fra i 365 giorni dell'anno siano tutti differenti. Essa, data l'ipotesi di equiprobabilità, si riduce al calcolo del numero dei casi possibili (pari alle r-disposizioni di n oggetti) e a quello dei casi favorevoli (pari alle r-disposizioni semplici di n

oggetti):

$$
\begin{aligned}
P(E_r) = 1 - P(\overline{E}_r) &= 1 - \frac{(n)_r}{n^r} \qquad (3.16)\\
&= 1 - \frac{365 \times 364 \times \cdots \times (365 - r + 1)}{365^r}\\
&= 1 - \frac{365}{365} \times \frac{364}{365} \times \cdots \times \frac{365 - r + 1}{365}\,.
\end{aligned}
$$

In tabella 3.2 è riportata $P(E_r)$ per alcuni valori di r. Si tenga conto che in realtà le nascite non sono distribuite uniformemente, ma si verificano più spesso in alcuni periodi dell'anno. Questo effetto tende a far aumentare la probabilità di coincidenze di compleanno (si immagini se, ad esempio, il 90 % delle nascite si verificasse concentrate in un solo mese e il restante 10 % negli altri mesi dell'anno).

3.6 Ricapitolando

- Il calcolo combinatorio si basa sulla regola fondamentale del prodotto dei numeri di possibilità delle scelte successive.
- I concetti principali del calcolo combinatorio sono
  - r-disposizioni di n oggetti (con ripetizioni):
    $$n^r\,;$$
  - r-disposizioni semplici di n oggetti (senza ripetizione), chiamato anche "permutazioni di n elementi r a r":
    $$(n)_r \;=\; {}_nP_r = \frac{n!}{(n-r)!}\,;$$
  - permutazioni di n oggetti:
    $$n!\,;$$
  - numero di combinazioni di n oggetti presi r a r:
    $$\binom{n}{r} \;=\; {}_nC_r = C_{n,r} = \frac{n!}{(n-r)!r!}\,.$$
- Si può arrivare agli stessi concetti del calcolo combinatorio da diversi "punti di vista": quello degli ordinamenti, quello delle occupazioni e quello delle estrazioni. A seconda del punto di vista può cambiare il ruolo fra "oggetti fisici" e "posizioni".
- Questo capitolo costituisce una specie di parentesi e non è essenziale per la comprensione del resto del testo.

3.7 Problemi

1. Un gruppo di 25 persone partecipa ad un veglione di capodanno. Allo scoccare della mezzanotte si stappano le bottiglie di champagne e tutti si scambiano gli auguri facendo tintillare i bicchieri. Quanti tintinnii si sentiranno in totale? (Sorvolando sul fatto essi che saranno coperti dal fragore dei botti...)
2. Il menù di una trattoria prevede 5 antipasti, 8 primi, 6 secondi, 3 contorni, e 3 dolci. Una persona decide di provare tutti i possibili pranzi che la trattoria può offrire. Quante volte vi si dovrà recare?
3. Quanto vale la probabilità che estraendo a caso 4 lettere dell'alfabeto italiano si ottenga una parola formata dalla sequenza consonante-vocale--consonante-vocale (tipo "mano", "tiro", etc). Si escludano, per ovvi motivi, le lettere h e q.
4. Nel gioco del Superenalotto ogni scommessa consiste nel pronosticare 6 numeri. Nel caso che tutti i sei numeri si verifichino si raggiunge la vincita massima. Il prezzo di ciascuna scommessa è di 800 lire. È anche possibile giocare indicare sulla scheda più di 6 numeri. Quanto costa una scheda in cui sono segnati 10 numeri?
5. Formulare i problemi 1, 3, 4, 6 e 7 del capitolo 2 nei termini del calcolo combinatorio.
6. Viene proposto un sistema di targhe costituito da un gruppo di 2 lettere seguito da uno di 4 cifre decimali (ad esempio AB3456). Quante macchine sarà possibile registrare?
7. Sul problema precedente: cosa cambia se le 4 cifre e le 2 lettere possono comparire in ogni posizione (ovvero anche 3A456B)?
8. Formulare il problema 20 del successivo capitolo 4 nei termini del calcolo combinatorio (l'anomalia di riferirsi ad un esercizio del capitolo successivo è giustificata dal carattere di complementarità di questo capitolo).
9. Calcolare le probabilità di fare estratto singolo ("ambata"), ambo, terna, quaterna e cinquina su una ruota. Sapendo che il premio pagato è rispettivamente 11.232, 250, 4250, 80˙000 e 1˙000˙000 di volte la puntata (a cui va sottratto il 3 %), si calcoli il rapporto fra speranza matematica e puntata per le diverse combinazioni di gioco.
10. Verificare che la formula (3.14) e (3.15) sono equivalenti.
11. Quanto vale la probabilità di fare 13 al totocalcio con una sola colonna?

Capitolo 4

Regole della probabilità

Abbiamo visto come valutare la probabilità di eventi semplici tramite ragionamenti di simmetria (combinatori), fiducia nella regolarità dei processi considerati (frequentisti) o semplicemente intuitivi. Abbiamo cercato di mostrare come il concetto di probabilità vada tenuto separato dal metodo di valutazione e come l'approccio soggettivista sia quello più ampio, nel quale gli altri rientrano come sottocasi particolari.

A volte ci sono degli eventi che possono essere espressi come espressioni logiche di eventi semplici. È possibile calcolare in modo "automatico" la probabilità di questi eventi se si conoscono le proprietà della probabilità delle combinazioni logiche di base. "Automatico" sta a significare che se le valutazioni di probabilità degli eventi di partenza veramente rappresentano il nostro grado di fiducia sull'accadere di essi, il risultato avrà l'analogo significato rispetto all'evento di interesse. Così anche valutazioni intuitive di eventi semplici e necessariamente grossolane (del tipo 10-20 %) possono condurre a valutazioni numeriche di probabilità di eventi complessi in quelle regioni molto prossime a 0 o a 1, dove l'intuizione confonderebbe gli ordini di grandezza della distanza da tali limiti di certezza.

4.1 Probabilità della somma logica di due eventi incompatibili

Se consideriamo un qualsiasi evento e il suo opposto, non ha molto senso scommettere sul fatto che si verifichi uno qualsiasi dei due, in quanto l'occorrenza di uno oppure dell'altro forma la certezza. Se questi casi sono contemplati è solo come caso limite.

Prendiamo ora due eventi E_1 e E_2 fra loro incompatibili (cioè che non possono essere entrambi veri) anche se non necessariamente opposti.

Prima di proseguire introduciamo una notazione sintetica che indichi l'evento composto "si verifica E_1 o si verifica E_2". La congiunzione "o" non va intesa nel senso di "aut ... aut", bensì in quello di "sia che si verifichi E_1 sia che si verifichi E_2". Questa operazione logica è chiamata *somma logica*, anche nota come "OR":

$$\text{somma logica (OR) di } E_1 \text{ e } E_2: \quad E_1 \cup E_2 \,.$$

Cerchiamo di valutare la probabilità di $E_1 \cup E_2$ sapendo che a E_1 e E_2 sono state assegnate probabilità $P(E_1)$ e $P(E_2)$. Ragioniamo con la logica della scommessa coerente per vincere un importo unitario (vedi paragrafo 2.10)

Le due scommesse separate su E_1 e E_2 equivalgono ad una sola scommessa con puntata $P(E_1) + P(E_2)$ sull'evento $E_1 \cup E_2$. Infatti:

- se non si verifica né E_1 né E_2, e quindi non si verifica $E_1 \cup E_2$, si perde $P(E_1) + P(E_2)$;
- se si verifica uno qualsiasi dei due si vince 1;
- non si possono verificare simultaneamente entrambi (con eventuale vittoria doppia) in quanto essi sono incompatibili.

Ne segue che

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) \quad \text{(se } E_1 \text{ e } E_2 \text{ incompatibili)}\,. \tag{4.1}$$

Questa è la terza proprietà fondamentale della probabilità, dopo quella che limita $P(E)$ fra 0 e 1 e quella che afferma che la probabilità dell'evento certo è uguale a 1.

4.2 Eventi e insiemi

Prima di procedere con le altre proprietà della probabilità è conveniente introdurre il formalismo con cui sono indicate le operazioni logiche di eventi. Tali operazioni verranno a mano a mano esemplificate sul caso degli esiti del lancio del dado.

Evento: come detto, rappresenta qualsiasi affermazione. Nel caso dei dadi possiamo avere ad esempio: "2", "pari", "$\leq$ 3", "1, 2, 5", etc. Questi eventi vengono indicati generalmente con lettere maiuscole: E, A, B, etc. A volte si usa H, a indicare "ipotesi", in quanto gli eventi sono associati ad ipotesi ("nell'ipotesi che esce il sei vinco").

A volte si indicano con lettere minuscole gli *eventi elementari*, volendo con essi indicare descrizioni di avvenimenti che non possono essere classificati ulteriormente in base a caratteristiche che hanno alcuni e non altri. Quindi nel caso del dado si avrebbe: e_1 = "1", e_2 = "2".

(Si noti che nei casi reali non ha senso parlare di eventi elementari, in quanto, dato un certo avvenimento, le caratterizzazioni possono essere virtualmente infinite. Ad esempio l'evento "la squadra X vince" può essere caratterizzata da "se gioca in casa", "con almeno due goal di scarto", "nonostante il portiere espulso", etc.)

Un modo generale di indicare gli eventi del lancio del dado è di elencare i casi elementari che li costituiscono, come ad esempio

- A = "pari" = $\{2, 4, 6\}$;
- B = "$\leq$ 3" = $\{1, 2, 3\}$;

Eventi	Insiemi	simbolo
evento	insieme	E
evento certo	ambiente (spazio campionario)	Ω
evento impossibile	insieme vuoto	$\emptyset$
implicazione	inclusione (sottoinsieme)	$E_1 \subseteq E_2$
uguaglianza	uguaglianza	se $\left\{ \begin{array}{l} E_1 \subseteq E_2 \\ \text{e} \\ E_2 \subseteq E_1 \end{array} \right. \Rightarrow E_1 = E_2$
evento opposto (o complementare)	insieme complementare	$\overline{E}$ $E \cup \overline{E} = \Omega$ $\overline{\Omega} = \emptyset, \overline{\emptyset} = \Omega$
prodotto logico	intersezione	$E_1 \cap E_2$
somma logica	unione	$E_1 \cup E_2$
eventi incompatibili	insiemi disgiunti	$E_1 \cap E_2 = \emptyset$
classe completa	partizione finita	$\left\{ \begin{array}{l} E_i \cap E_j = \emptyset \; \forall i \neq j \\ \bigcup_i E_i = \Omega \end{array} \right.$

Tabella 4.1: Corrispondenza fra eventi ed insiemi.

- C = "6" = $e_6 = \{6\}$.

Evento certo: è indicato con Ω. Nel caso del dado

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

L'evento certo dipende dal problema. Ad esempio nel lancio di una moneta l'evento certo può essere {"testa", "croce"}, o {"testa", "croce", "dritta"}. Anche se queste precisazioni sembrano cavillose e sono sottintese nella maggior parte dei problemi didattici, esse sono invece importanti nella vita reale (si pensi alle classiche postille dei contratti assicurativi, intere pagine scritte a caratteri microscopici che vengono fatte firmare esplicitamente, oltre a quanto sottoscritto nell'atto principale - "non si preoccupi, firmi tranquillo: niente d'importante"...).

Evento impossibile: sono tutte le affermazioni incompatibili con gli esiti del dado: "esce sia 3 che 4 allo stesso colpo"; "esce $\pi/4$", etc.

Implicazione, indicata con $E_1 \subseteq E_2$ ("E_1 implica E_2"). Indica che se E_1 è vero anche E_2 è necessariamente vero. Esempio:

- se $E_1 = \{2\}$ e E_2 = "pari", ne segue che $E_1 \subseteq E_2$;
- se $E_1 = \{1,\ 3,\ 5\}$ e E_2 = "dispari", ne segue che $E_1 \subseteq E_2$ e $E_2 \subseteq E_1$. Quando E_1 implica E_2 e anche E_2 implica E_1 vuol dire che i due eventi sono **uguali**: $E_1 = E_2$.

Evento opposto ad E, indicato con $\overline{E}$ (si incontra anche il simbolo E^c). Se E è vero segue che $\overline{E}$ è falso e viceversa, come "pari" e "dispari". È chiamato anche evento complementare.

Prodotto logico, indicato con $E_1 \cap E_2$ (si incontra anche $E_1 \wedge E_2$). Esso è vero se sono veri sia E_1 che E_2. Esso è anche indicato con il simbolo "AND", con l'"e" commerciale "&" o, anche semplicemente, con il classico simbolo di prodotto "·". Esempi:

- se $E_1 = \{1, 3, 5\}$ e $E_2 = \{1\}$ segue che $E_1 \cap E_2 = \{1\}$;
- se E_1 = "≤ 4" e E_2 = "≥ 3" segue che $E_1 \cap E_2 = \{3, 4\}$;
- se $E_1 = \{1\}$ e E_2 = "pari" segue che il prodotto logico è un evento impossibile: $E_1 \cap E_2 = \emptyset$;
- se consideriamo due dadi e chiamiamo E_1 = "≤ 2" al primo dado e E_2 = "pari" al secondo dado, il prodotto logico $E_1 \cap E_2$ è dato dalle coppie $\{(1,2),(1,4),(1,6),(2,2),(2,4),(2,6)\}$, ove le coppie ordinate di valori si riferiscono ai risultati dei due dadi.

Dalla definizione segue che il prodotto logico di E_1 e E_2 implica sia E_1 che E_2: $(E_1 \cap E_2) \subseteq E1$; $(E_1 \cap E_2) \subseteq E2$.

Eventi incompatibili: quando non possono essere veri entrambi, ovvero se il loro prodotto logico è un evento impossibile ($E_1 \cap E_2 = \emptyset$). Un esempio è riportato nel punto precedente. Un evento e il suo opposto sono sempre incompatibili: $E \cap \overline{E} = \emptyset$.

Somma logica: $E_1 \cup E_2$ (si incontra anche $E_1 \vee E_2$). È un evento vero se è vero almeno uno dei due eventi. È anche indicata con il simbolo "OR". Esempi

- se $E_1 = \{1\}$ e E_2 = "pari", ne segue che $E_1 \cup E_2 = \{1, 2, 4, 6\}$.
- se consideriamo un evento e il suo opposto, la loro unione logica dà la certezza: $E \cup \overline{E} = \Omega$.

Dalla definizione segue che ciascuno degli eventi E_1 e E_2 implica la loro somma logica: $E_1 \subseteq (E_1 \cup E_2)$; $E_2 \subseteq (E_1 \cup E_2)$.

Classe completa: è composta da eventi tali che essi siano a due a due mutuamente esclusivi e tali che la loro somma logica costituisca l'evento certo (quest'ultima proprietà è anche espressa dicendo che sono *esaustivi*).

$$\begin{cases} E_i \cap E_j = \emptyset & \forall i \neq j \\ \bigcup_{i=1}^{n} E_i = \Omega \end{cases} \tag{4.2}$$

La notazione $\bigcup_{i=1}^{n} E_i$ indica la somma logica degli n eventi della classe (è il corrispondente del simbolo di sommatoria dell'aritmetica).

involuzione	$\overline{\overline{A}} = A$
commutatività	$A \cup B = B \cup A$ $A \cap B = B \cap A$
associatività	$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$
distributività	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
distributività numerabile	$A \cap (B_1 \cup B_2 \cup \ldots) = (A \cap B_1) \cup (A \cap B_2) \cup \ldots$ $A \cup (B_1 \cap B_2 \cap \ldots) = (A \cup B_1) \cap (A \cup B_2) \cap \ldots$
idempotenze	$A \cup A = A$ $A \cap A = A$
assorbimento	$A \cup (A \cap B) = A$ $A \cap (A \cup B) = A$
assorbimento da Ω e da $\emptyset$	$A \cup \Omega = \Omega$ $A \cap \emptyset = \emptyset$
identità	$A \cup \emptyset = A$ $A \cap \Omega = A$
legge di contraddizione	$A \cap \overline{A} = \emptyset$
legge del terzo escluso	$A \cup \overline{A} = \Omega$
leggi di De Morgan	$\overline{A \cap B} = \overline{A} \cup \overline{B}$ $\overline{A \cup B} = \overline{A} \cap \overline{B}$

Tabella 4.2: Proprietà fondamentali delle operazioni fra insiemi (ed eventi). Esse possono essere dimostrate facilmente utilizzando i diagrammi di Venn.

- E_1 = "pari" e E_2 = "dispari" rispettano questa condizione; lo stesso vale per E_1 = "≤ 2", E_2 = "3", E_3 = "> 3";
- E_1 = "< 4" e E_2 = "≥ 3" non formano una classe completa, e nemmeno E_1 = "< 4" e E_2 = "> 4" (violano rispettivamente la prima e la seconda condizione della (4.2);
- non è necessario che la classe completa sia composta dagli "eventi elementari", anche perché questi eventi non sono in genere definibili nei problemi reali.

È spesso molto pratico mettere in relazione gli eventi con gli insiemi. Essi infatti obbediscono a identiche proprietà formali se si fanno delle opportune analogie fra definizioni sugli eventi e definizioni sugli insiemi, come mostrato in tabella 4.1. Per gli uni e per gli altri valgono ad esempio le proprietà riportate in tabella 4.2. Inoltre un importante vantaggio della corrispondenza fra

Simbolo booleano	Evento	A	B 0	1
AND	$A \cap B$	0	0	0
		1	0	1
OR	$A \cup B$	0	0	1
		1	1	1
NAND	$\overline{A \cap B}$ $(= \overline{A} \cup \overline{B})$	0	1	1
		1	1	0
NOR	$\overline{A \cup B}$ $(= \overline{A} \cap \overline{B})$	0	1	0
		1	0	0
XOR	$(A \cap \overline{B}) \cup (\overline{A} \cap B)$	0	0	1
		1	1	0

Tabella 4.3: Simboli dell'algebra booleana e dell'algebra degli insiemi con tabelline della verità (0 e 1 stanno rispettivamente per falso e vero).

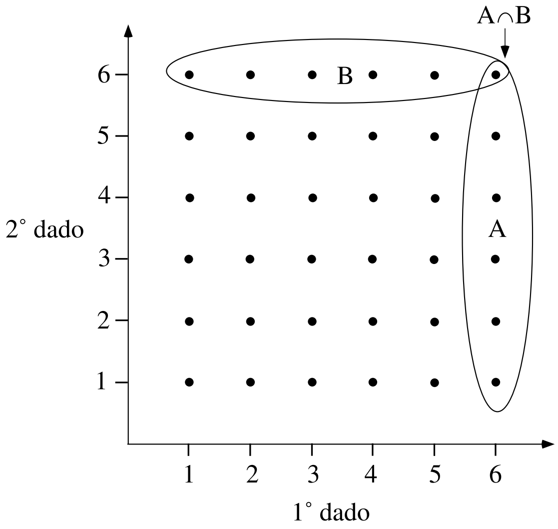

Figura 4.1: Spazio campionario ed esempi di eventi nel caso di lancio di due dadi: A="il primo dado dà 6"; B = "il secondo dado dà 6"; $A \cap B$ = "entrambi i dadi danno 6".

eventi ed insiemi è che si può utilizzare la visualizzazione grafica dei diagrammi di Venn. È allora conveniente disegnare le aree con cui si rappresentano gli insiemi proporzionali alle probabilità dei corrispettivi eventi. Il semplice caso dello spazio campionario legato al lancio di due dadi è mostrato in figura 4.1 (per non appesantire la figura si è evitato di disegnare Ω intorno ai 36 punti). Alcune proprietà di eventi e insiemi sono raffigurate in figura 4.2. Come completamento di quanto detto sull'algebra degli eventi riportiamo pure in tabella 4.3 i simboli usualmente utilizzati nell'algebra booleana con il corrispondente dell'algebra degli eventi e degli insiemi. Per ciascuna operazione è riportata anche la cosiddetta tavola della verità, di interpretazione immediata.

4.3 probabilità come misura

*** si può mettere qui una nota sulla probabilità come misura, come interpretazione geometrica dei diagrammi di Venn?

*** vedi BdF "probabilità come misura e probabilità come massa"

4.4 Evento condizionato

Abbiamo già introdotto il concetto di probabilità condizionata, intuitivo alla stessa stregua della probabilità in generale. Anzi, abbiamo cercato di mostrare come in realtà la valutazione di probabilità vada sempre intesa come subordinata a qualche informazione, ipotesi, dati sperimentali o pregiudizi estetici (fondamentali questi ultimi nella fisica teorica).

Chiariamo ora meglio il concetto di evento condizionato. $E \mid H$ è una qualsiasi affermazione rispetto alla quale siamo in stato di incertezza, ma che può essere vera o falsa *nell'ipotesi che H sia vera*. Nell'ipotesi che H sia falsa l'evento $E \mid H$ perde di significato. Può essere schematizzata quindi con:

$$\mathbf{E} \mid \mathbf{H} \rightarrow \begin{cases} \textbf{Vero}\text{: se, essendo vero } H\text{, è vero } E; \\ \textbf{Falso}\text{: se, essendo vero } H\text{, è falso } E; \\ \textbf{Indeterminato}\text{: nessun valore logico se } H \text{ è falso;} \end{cases} \tag{4.3}$$

In termini di scommessa il terzo caso corrisponde ad annullare la scommessa[1]. Ad esempio, dovendo fare una scommessa su una partita di calcio si può porre la condizione "se non piove". Anche lanciando i dadi in giochi di società, si usa talvolta la condizione "se il dado cade dal tavolo l'esito non è valido e il dado va rilanciato".

Bisogna fare attenzione a non confondere $E \cap H$ e $E \mid H$. Consideriamo l'esempio del lancio dei due dadi di figura 4.1. Prendiamo l'evento condizionato "6 al secondo dado condizionato dal verificarsi di 6 al primo dado". Sebbene dal punto di vista geometrico, considerando il verificarsi del doppio

[1] Ad esempio, il testo della legge che regola il gioco del lotto specifica: "Quando le matrici rivelano incompletezza di dati o le scommesse sono state accettate in violazione delle disposizioni dell'articolo 3 o i dati non sono pervenuti al centro di elaborazione, le scommesse si considerano non avvenute e il giocatore escluso dalla partecipazione all'estrazione ha diritto al rimborso dell'importo della scommessa previa esibizione dello scontrino al raccoglitore".

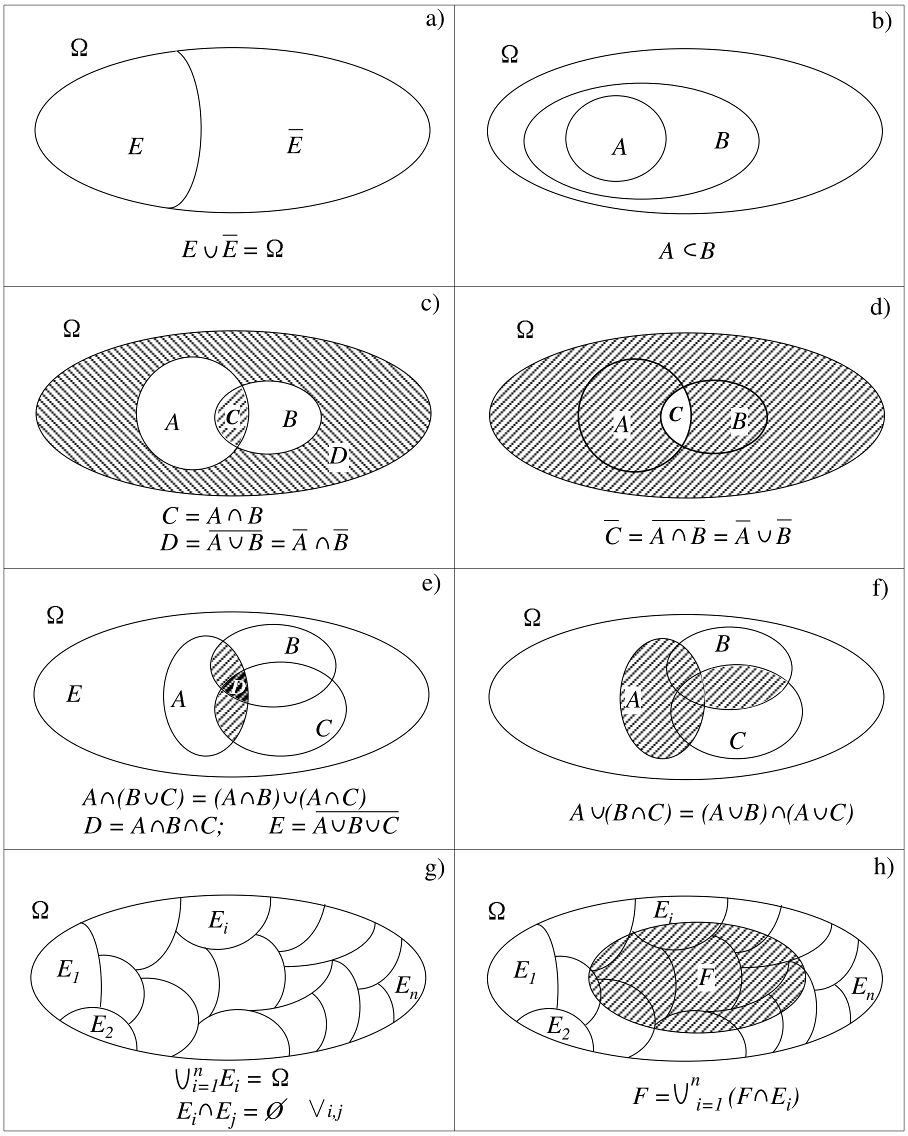

Figura 4.2: Proprietà degli insiemi e diagrammi di Venn. a) insieme complementare; b) sottoinsieme (proprio); c-f) proprietà dell'unione e dell'intersezione; g) partizione finita; h) decomposizione di un insieme F nelle proprie intersezioni con una partizione finita

6, $B \cap A$ e $B \,|\, A$ *sembrerebbero* la stessa cosa, cambia l'ambiente entro cui tale evento è considerato. Infatti $A \cap B$ è l'intersezione di A e di B riferita allo spazio campionario Ω costituito dalle 36 possibilità, mentre $B \,|\, A$ è sempre l'intersezione dei due, ma riferita allo spazio campionario ridotto Ω^* costituito da A.

Quindi, anche se $E \cap H$ e $E \,|\, H$ sono veri simultaneamente (diciamo che sono "uguali dal punto di vista fisico"), essi differiscono quando siamo in condizione di incertezza, in quanto siamo interessati al verificarsi di E solo nell'ipotesi che H sia vero. Ciò si riflette sulla valutazione della probabilità. In particolare si intuisce come $P(E \,|\, H)$ sia maggiore o uguale a $P(E \cap H)$, in quanto l'evento è contemplato all'interno di una classe di ipotesi più ristretta.

Un caso particolare immediato che mostra la diversa valutazione di probabilità nei due casi è

$$P(H \,|\, H) = 1, \tag{4.4}$$

valida per qualsiasi evento H, qualunque sia il suo valore di probabilità (anche nullo, nel senso chiarito nel paragrafo 2.7).

4.5 Regole di base della probabilità - assiomi

Abbiamo incontrato alcune regole a cui la valutazione della probabilità deve soddisfare, derivate dal concetto di scommessa coerente: la probabilità deve essere compresa fra zero e 1; vale 1 per l'evento certo e 0 per quello impossibile; vale la regola di somma per probabilità di eventi incompatibili.

Da queste regole, mediante le proprietà formali degli eventi e degli insiemi è possibile derivare altre proprietà cui la probabilità deve soddisfare. È possibile dimostrare facilmente che queste regole di base sono soddisfatte automaticamente anche dalle valutazioni combinatorie e frequentiste.

Esiste un approccio molto formale alla probabilità in cui le tre regole di base sono assunte come assiomi e le proprietà che ne seguono sono ricavate come teoremi. In questo approccio però la probabilità *non è definita come concetto*, così come anche l'evento è soltanto un oggetto matematico. In questa teoria la probabilità è semplicemente un numero reale che soddisfi i tre assiomi dati dalle regole[2]:

Regola 1 (*positività*): $0 \le P(E) \le 1$;

Regola 2 (*certezza*): $P(\Omega) = 1$, $P(\emptyset) = 0$;

Regola 3 (*unione*): $P(E_1 \cup E_2) = P(E_1) + P(E_2)$, se $E_1 \cap E_2 = \emptyset$.

In particolare, la Regola 3 può essere estesa ad un numero[3] n di eventi:

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i) \quad \text{se } E_i \cap E_j = \emptyset \quad \forall i \neq j\,. \tag{4.5}$$

[2] ❄Al fine di smitizzare il pur storicamente importante approccio assiomatico preferiamo parlare semplicemente di "regole", che ovviamente corrispondono agli assiomi. Gli assiomi 1 e 2 sono spesso presentati nelle forme come $P(E) \ge 0$ e $P(\Omega) = 1$, le quali ovviamente lasciano invariate tutte le proprietà che ne discendono.

[3] ❄L'estensione dell'unione a infiniti eventi è delicata e controversa, ma inessenziale per la nostra trattazione.

Questa relazione è nota con il nome di teorema delle probabilità totali.

Da queste regole di base seguono alcune relazioni importanti che devono essere sempre soddisfatte dalle valutazioni della probabilità:

Proprietà 1 : $P(E) = 1 - P(\overline{E})$.

Proprietà 2 : $P(B) = P(B \cap A) + P(B \cap \overline{A})$. Questa proprietà può essere estesa ad una classe completa di eventi, ottenendo

$$P(B) = P\left(\bigcup_{i=1}^{n}(B \cap E_i)\right) = \sum_{i=1}^{n} P(B \cap E_i) \qquad (4.6)$$

come mostrato anche in figura 4.2, riquadro h).

Proprietà 3 : Se l'evento A implica l'evento B, cioè $A \subseteq B$, allora $P(A) \leq P(B)$. In particolare ne segue che

$$\begin{array}{rcl} P(A \cap B) & \leq & P(A) \\ P(A \cap B) & \leq & P(B) \\ P(A \cup B) & \geq & P(A) \\ P(A \cup B) & \geq & P(B) \end{array}$$

Proprietà 4 : probabilità della somma logica nel caso generale:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)\,. \qquad (4.7)$$

Ne segue che

$$P(A \cup B) \leq P(A) + P(B)\,. \qquad (4.8)$$

Nel caso di tre eventi la proprietà diventa:

$$\begin{array}{rcl} P(A \cup B \cup C) & = & P(A) + P(B) + P(C) \\ & & -P(A \cap B) - P(A \cap C) - P(B \cap C) \\ & & +P(A \cap B \cap C)\,, \end{array} \qquad (4.9)$$

estendibile a n eventi come

$$\begin{array}{rcl} P\left(\bigcup_{i=1}^{n} E_i\right) & = & \sum_{i=1}^{n} P(E_i) - \sum_{i<j} P(E_i \cap E_j) \\ & & + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) - \cdots \\ & & -(-1)^n P(E_1 \cap E_2 \cap \cdots \cap E_n)\,. \end{array} \qquad (4.10)$$

Questa formula è chiamata *principio di inclusione-esclusione* a causa dell'alternanza dei segni.

Queste proprietà sono abbastanza intuitive e si possono dimostrare facilmente utilizzando i diagrammi di Venn. La 4, in particolare, è molto importante e, nella soluzione dei problemi, dovrebbe essere presa in considerazione prima della Regola 3, che può essere vista come suo sottocaso valido quando $P(A \cap B) = 0$. Il motivo per cui si sottrae $P(A \cap B)$ a $(P(A) + P(B))$ è dovuto al fatto che, per dirlo alla buona, altrimenti l'elemento $P(A \cap B)$ verrebbe contato due volte. Si può vedere in un semplice caso legato al gioco delle carte:

- Si consideri un mazzo di carte da gioco italiane con 4 semi e 10 valori per seme. Si vuole calcolare la probabilità di estrarre una Coppe (C) o un Asso (A). Supponendo l'equiprobabilità (carte ben mischiate), le probabilità di una Coppe o di un Asso sono rispettivamente $P(C) = \frac{1}{4}$, $P(A) = \frac{1}{10}$. È chiaro che se si sommano semplicemente queste probabilità l'Asso di Coppe viene contato due volte ed è per questo che bisogna sottrarre dalla somma la sua probabilità ($P(A \cap C) = \frac{1}{40}$), ottenendo come risultato $P(A \cup B) = \frac{13}{40}$.

4.5.1 ❄Dimostrazioni delle proprietà della probabilità

Come esercizio sull'algebra degli eventi dimostriamo le proprietà descritte nel paragrafo precedente (con la stessa numerazione progressiva).

1. E ed $\overline{E}$ sono incompatibili e la loro unione vale Ω. Ne segue:

$$\begin{aligned} P(E \cup \overline{E}) &= P(\Omega) = 1 \\ P(E) + P(\overline{E}) &= 1\,. \end{aligned}$$

2. Un qualsiasi evento B può essere scritto in generale come

$$B = B \cap \Omega = B \cap (A \cup \overline{A}) = (B \cap A) \cup \left(B \cap \overline{A}\right)\,, \tag{4.11}$$

in quanto esso può essere vero sia quando è vero anche A che quando A è falso.

Poiché B è stato suddiviso in due eventi fra loro incompatibili si ha:

$$P(B) = P(B \cap A) + P(B \cap \overline{A}). \tag{4.12}$$

3. Si riparte dalla (4.11) osservando che, se A implica B, quando è vero A è vero anche $A \cap B$. Si ottiene in questo caso:

$$B = A \cup (B \cap \overline{A}) \tag{4.13}$$

(provare a visualizzare con un diagramma di Venn: B à costituito da tutti i punti di A più quelli non di A ma che appartengono a B). Essendo A e $(B \cap \overline{A})$ eventi incompatibili si ha

$$P(B) = P(A) + P(B \cap \overline{A}) \geq P(A)\,, \tag{4.14}$$

in quanto $P(B \cap \overline{A})$ è non negativa per la prima regola della probabilità.

4. L'unione logica di due eventi A e B può essere espressa come unione logica di tre eventi fra loro incompatibili:

$$A \cup B = (A \cap \overline{B}) \cup (\overline{A} \cap B) \cup (A \cap B)\,, \tag{4.15}$$

ossia $A \cup B$ è vero quando è vero soltanto uno di essi oppure sono veri entrambi. Tenendo conto anche della proprietà 2 applicata a ciascuno dei due eventi relativamente all'altro, ossia

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

si ottiene

$$\begin{aligned} P(A \cup B) &= P(A \cap \overline{B}) + P(\overline{A} \cap B) + P(A \cap B) \\ P(A) &= P(A \cap B) + P(A \cap \overline{B}) \\ P(B) &= P(B \cap A) + P(B \cap \overline{A})\,. \end{aligned}$$

Sottraendo membro a membro le ultime due equazioni dalla prima si arriva alla relazione da provare.

4.6 Relazione fra probabilità condizionata e congiunta

Abbiamo già parlato del fatto che l'evento condizionato sia sostanzialmente lo stesso evento fisico del prodotto logico con il condizionante, ma con una diversa condizione di incertezza in quanto subentra l'ipotesi che il condizionante sia vero. La condizione aggiuntiva (rispetto all'ambiente Ω) fa sì che in genere la probabilità di $E|H$ sia maggiore di quella di $E \cap H$ ogni qual volta che $H \subset \Omega$ ("sottoinsieme proprio") e ovviamente $P(E \cap H) \neq 0$. In effetti, ipotizzando che il "fattore di amplificazione" sia proprio $1/P(H)$ - se è diverso da zero - si riottengono facilmente i casi limite di $H = \Omega$ o $E = H$. In effetti in tutti i casi vale la relazione

$$P(E \cap H) = P(H) \cdot P(E \,|\, H)\,, \tag{4.16}$$

da cui

$$P(E \,|\, H) = \frac{P(E \cap H)}{P(H)} \qquad (P(H) \neq 0)\ . \tag{4.17}$$

Ovviamente i ruoli di E e di H possono essere scambiati e la (4.16) può essere scritta più in generale

$$P(E \cap H) = P(H) \cdot P(E \,|\, H) = P(E) \cdot P(H \,|\, E)\,, \tag{4.18}$$

Queste relazioni sono molto importanti per le applicazioni, ma necessitano delle delucidazioni sul loro esatto significato, al fine di capire quando è lecito utilizzarle. Soprattutto può sorgere un po' di confusione in quanto a volte si incontra la (4.17) indicata come *definizione* della probabilità condizionata, altre volte si incontra la (4.16) come *teorema delle probabilità composte*. Vediamo come le formule appaiono nei diversi approcci:

- nella valutazione combinatoria e frequentista la (4.16) è una diretta conseguenza della definizione. Ad esempio, se consideriamo la valutazione combinatoria, indicando con $\#\Omega$ il numero[4] dei casi totali di Ω (la sua *numerosità*), $\#H$ il numero di casi favorevoli ad H e $\#EH$ quelli favorevoli a E e H simultaneamente si ottiene (ricordandosi della (2.6)):

$$P(E\,|\,H) = \frac{\#EH}{\#H} = \frac{\frac{\#EH}{\#\Omega}}{\frac{\#H}{\#\Omega}} = \frac{P(E\cap H)}{P(H)}\,. \tag{4.19}$$

  Una analoga dimostrazione si ottiene per la valutazione delle probabilità dalle frequenze relative.

- Anche nell'impostazione soggettivista la (4.16) è un teorema, essendo essa condizione necessaria e sufficiente per una scommessa coerente sull'evento $P(E\,|\,H)$. La dimostrazione è molto semplice. Si immagini di fare una scommessa coerente per vincere un importo unitario sull'evento condizionato $E\,|\,H$ (vedi paragrafo 4.4): si paga (con certezza!) la cifra $P(E\,|\,H)$ per vincere 1 se si verificano sia E che H (con probabilità $P(E\cap H)$); nel caso che H non si verifichi (con probabilità $P(\overline{H})$) la scommessa viene invalidata e si recupera la puntata $P(E\,|\,H)$. Quindi la previsione di guadagno è:

$$\begin{aligned}\mathrm{I\!P}(G) &= 1\cdot(-P(E\,|\,H)) + P(E\cap H)\cdot 1 + P(\overline{H})\cdot P(E\,|\,H)\\ &= -P(E\,|\,H) + P(E\cap H) + (1-P(H))\cdot P(E\,|\,H)\\ &= P(E\cap H) - P(E\,|\,H)\cdot P(H)\,. \end{aligned}\tag{4.20}$$

  Essendo la scommessa coerente e quindi nulla la previsione di guadagno, segue la (4.16).

- Soltanto nell'approccio assiomatico la (4.17) è presentata come *definizione* della probabilità condizionata (anche se poi a volte la (4.18) viene chiamata nello stesso testo "teorema"...). Ora se questa formula fosse veramente la definizione della probabilità condizionata sarebbe necessario valutare ogni volta $P(H)$ e $P(E\cap H)$ per poi ottenere $P(E\,|\,H)$. Questo punto di vista è chiaramente non accettabile e fortunatamente non è seguito in pratica (nemmeno dalla maggior parte di coloro che pensano che la (4.17) sia una definizione!).

  Si pensi a dei fisici che lavorano al progetto di un nuovo rivelatore per la nuova generazione di acceleratori e siano interessati alla probabilità che un certo algoritmo riesca a identificare un dato decadimento di una nuova particella elementare, ipotizzando una certa massa, certi modi di decadimento, dei parametri del rivelatore, etc. Non è assolutamente pensabile valutare questa probabilità condizionata dalla probabilità dei condizionanti (che la particella esista sul serio, che abbia quella massa, etc) e dalla congiunta dei condizionanti con l'evento "il decadimento è riconosciuto dall'algoritmo"). In pratica si fissano i valori dei condizionanti nei programmi di simulazione e si valutano le probabilità dalle frequenze.

[4] Il simbolo "#" sta per "numero".

> Se questi esempi possono sembrare un po' particolari, si provi a calcolare secondo la (4.17) la probabilità che, avendo estratto dalla tombola 40 numeri senza che sia uscito il "47", questo esca al 41^{mo} tentativo (la risposta "al volo" è 1/50).

Nella maggior parte dei problemi pratici e scientifici è molto più pratico (e in molti casi il solo metodo praticabile) assegnare direttamente $P(E\,|\,H)$ (ad esempio come descritto nel paragrafo 2.6) ed eventualmente da questa, se $P(H)$ è diverso da zero, calcolare $P(E \cap H)$ mediante la (4.16).

Facciamo un esempio banale ma che mostra a quali difficoltà si può andare incontro prendendo alla lettera la (4.17). Si consideri di lanciare quattro volte una moneta regolare e di voler calcolare "la probabilità che al quarto lancio sia uscita Testa se anche al secondo era uscita Testa". Chiaramente la risposta diretta di 1/2 è quella corretta. La si confronti con il seguente procedimento:

- ci sono $2^4 = 16$ casi possibili, dei quali $2^3 = 8$ contengono Testa al 2^o colpo e $2^2 = 4$ contengono Testa congiuntamente al 2^o e al 4^0;

- quindi $P(T_2) = 8/16 = 1/2$, $P(T_2 \cap T_4) = 4/16 = 1/4$, da cui

$$P(T_4\,|\,T_2) = \frac{P(T_2 \cap T_4)}{P(T_2)} = \frac{1/4}{1/2} = \frac{1}{2}\,.$$

4.7 Condizionamento da eventi di probabilità nulla

Il fatto che la formula (4.17) non possa definire la probabilità condizionata, ma ne offra soltanto una valutazione se sono noti gli altri due termini, può essere compreso, oltre che dai ragionamenti di buon senso fatti precedentemente, dal considerare gli eventi di probabilità nulla nel senso chiarito nel paragrafo 2.7.

Innanzitutto è chiaro che la probabilità di un evento condizionato con sé stesso deve essere sempre uguale ad uno, indipendentemente dal valore di $P(E)$, in quanto $P(E\,|\,E)$ rappresenta *"la probabilità di E nell'ipotesi che E sia vero"*:

$$P(E\,|\,E) = 1 \quad .$$

Inoltre, se consideriamo due eventi A e B di probabilità nulla, ma ai quali attribuiamo lo stesso grado di fiducia (si considerino ad esempio le condizioni "normali" di traffico descritte nel paragrafo 2.7) deve valere:

$$P(A) = P(B) = P(A \cup B) = 0$$

ma

$$\begin{aligned} P(A \cup B\,|\,A \cup B) &= 1 \\ P(A\,|\,A \cup B) + P(B\,|\,A \cup B) &= 1 \end{aligned}$$

da cui

$$\begin{aligned} P(A\,|\,A \cup B) &= \frac{1}{2} \\ P(B\,|\,A \cup B) &= \frac{1}{2}\,, \end{aligned}$$

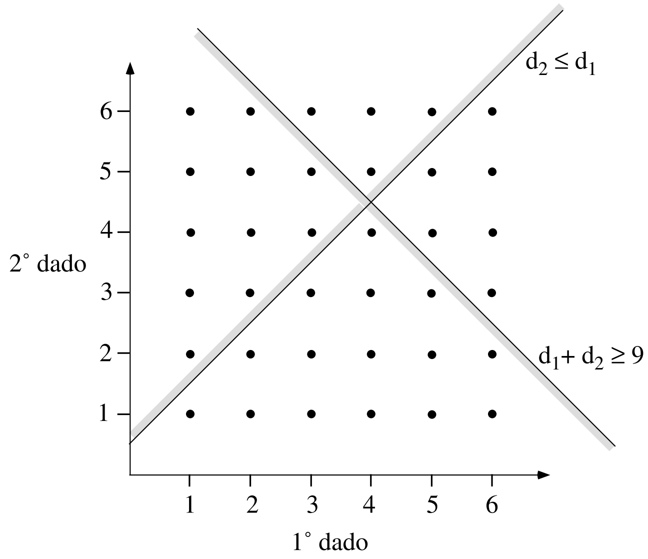

Figura 4.3: Spazio campionario ed esempi di eventi nel caso di lancio di due dadi. Sono indicati i condizionamenti "la somma è maggiore o uguale a 9" e "il secondo valore è non superiore al primo".

sebbene queste probabilità non siano ottenibili dalla (4.17).

Gli esempi appena riportati possono sembrare più delle curiosità che dei casi di interesse scientifico. È invece facile convincersi, ad esempio, che ogni volta che si valuta la probabilità di qualche evento mediante simulazioni al computer che dipendono da parametri rispetto ai quali si è in condizione di incertezza si sta usando esattamente il concetto di condizionamento da eventi di probabilità nulla.

4.8 Indipendenza stocastica (o in probabilità)

Immaginiamo l'esperimento del lancio di due dadi regolari distinguibili e di indicare con d_1 e d_2 i rispettivi valori. Interessiamoci alla probabilità dei seguenti eventi relativi al 2^0 dado "$d_2 = 1$" "$d_2 = 3$" e "$d_2 = 6$". Per simmetria possiamo solo assegnare probabilità 1/6 a ciascuno degli eventi. Consideriamo ora di voler valutare la probabilità nell'ipotesi che i seguenti eventi siano alternativamente veri: "$d_1 = 5$"; "$d_1 + d_2 \geq 9$"; "$d_2 \leq d_1$". (Una persona potrebbe sbirciare il risultato e darci queste informazioni.)

Calcolando direttamente[5] le probabilità condizionate facendo uso della

[5] I masochisti possono anche usare la 4.17, che in questo caso è applicabile…

figura 4.3 otteniamo per i tre condizionamenti:

$$\text{“}d_1 = 5\text{''} \quad : \quad \begin{cases} P(d_2 = 1 \,|\, d_1 = 5) = 1/6 & [= P(d_2 = 1)] \\ P(d_2 = 3 \,|\, d_1 = 5) = 1/6 & [= P(d_2 = 3)] \\ P(d_2 = 6 \,|\, d_1 = 5) = 1/6 & [= P(d_2 = 6)] \end{cases}$$

$$\text{“}d_1 + d_2 \geq 9\text{''} \quad : \quad \begin{cases} P(d_2 = 1 \,|\, d_1 + d_2 \geq 9) = 0 & [< P(d_2 = 1)] \\ P(d_2 = 3 \,|\, d_1 + d_2 \geq 9) = 1/10 & [< P(d_2 = 3)] \\ P(d_2 = 6 \,|\, d_1 + d_2 \geq 9) = 4/10 & [> P(d_2 = 6)] \end{cases}$$

$$\text{“}d_2 \leq d_1\text{''} \quad : \quad \begin{cases} P(d_2 = 1 \,|\, d_2 \leq d_1) = 6/21 & [> P(d_2 = 1)] \\ P(d_2 = 3 \,|\, d_2 \leq d_1) = 4/21 & [> P(d_2 = 3)] \\ P(d_2 = 6 \,|\, d_2 \leq d_1) = 1/21 & [< P(d_2 = 6)] \end{cases}$$

Si vede quindi come il nuovo stato di informazione possa aumentare, diminuire o inalterare la credibilità dell'evento in considerazione. Anche chi non sa - o non ha tempo di - fare i conti in dettaglio intuisce come, se la somma dei due risultati è maggiore o uguale a 9, sia più facile che si verifichino valori alti per il secondo dado, e sicuramente non inferiori a 3. Se l'ipotesi che un evento H sia vero cambia la probabilità di un altro evento (E) diciamo che i due eventi sono *correlati*:

- se $P(E\,|\,H) > P(E) \longrightarrow E$ e H sono correlati *positivamente*;
- se $P(E\,|\,H) < P(E) \longrightarrow E$ e H sono correlati *negativamente*.

Altrimenti

- se $P(E\,|\,H) = P(E)$ i due eventi sono *indipendenti*, o più propriamente *indipendenti in probabilità* (o *stocasticamente indipendenti*).

In questo caso il teorema delle probabilità composte diventa:

$$P(E \cap H) = P(E) \cdot P(H)\,, \tag{4.21}$$

nota come *regola della moltiplicazione della probabilità di eventi indipendenti*. Questa relazione può essere un modo alternativo alla condizione $P(E\,|\,H) = P(E)$ per definire l'indipendenza stocastica. Si noti come l'uso della (4.21) sia duplice. Ci sono dei casi in cui è possibile valutare direttamente i tre termini della relazione e verificare l'indipendenza, oppure *assumere* l'indipendenza e utilizzare tale formula per calcolare la probabilità congiunta a partire dalle probabilità dei due eventi (o in genere del terzo incognito).

4.9 Altre proprietà della probabilità condizionata

Vediamo altre relazioni importanti sulle probabilità ottenute dal teorema delle probabilità composte e dall'algebra degli eventi.

4.9.1 Legge della moltiplicazione

Le relazioni (4.16) e (4.21) possono essere estese al caso di n eventi, ottenendo la *legge della moltiplicazione delle probabilità*, dall'espressione più generale

$$\begin{aligned} P(E_1 \cap E_2 \cap \cdots \cap E_n) &= P(E_1) \cdot P(E_2\,|\,E_1) \cdot P(E_3\,|\,E_1 \cap E_2) \cdots \\ &\quad \cdot P(E_n\,|\,E_1 \cap E_2 \cap \cdots \cap E_{n-1})\,. \end{aligned} \tag{4.22}$$

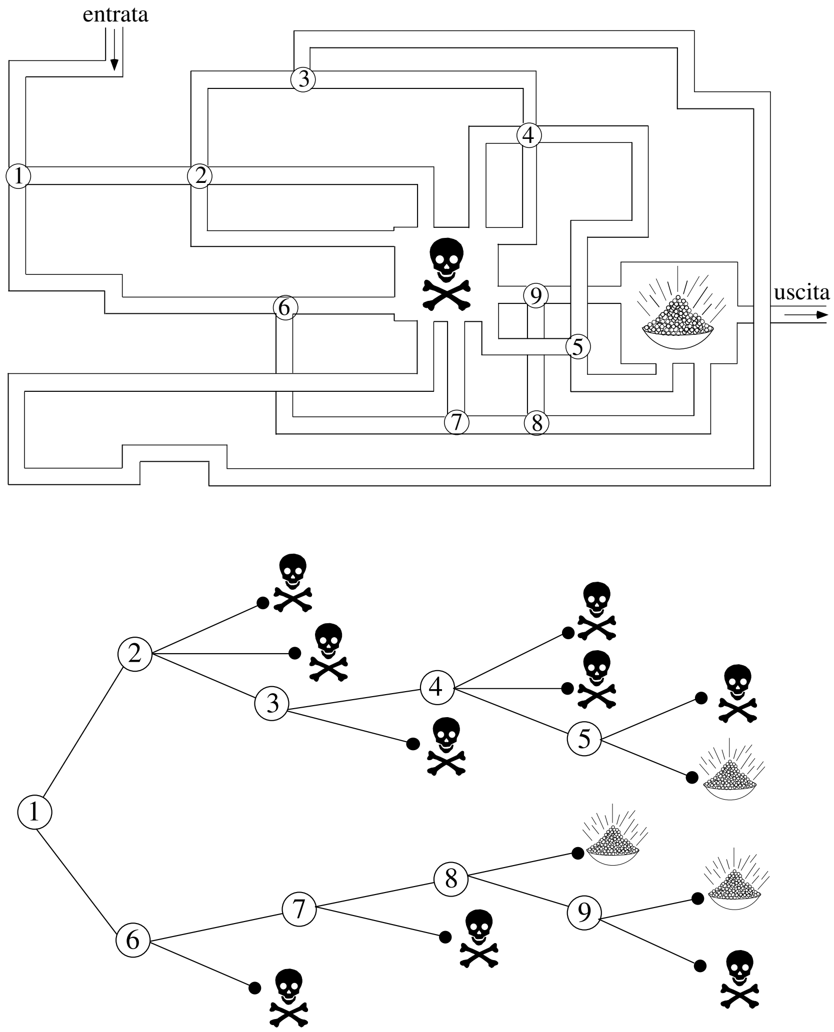

Figura 4.4: Esempio di probabilità condizionata.

In analogia al semplice caso di due eventi, quando si verifica che

$$P(E_1 \cap E_2 \cap \cdots \cap E_n) = P(E_1) \cdot P(E_2) \cdot \cdots \cdot P(E_n) \qquad (4.23)$$

gli n eventi fra di loro indipendenti. Si noti come l'indipendenza fra ciascuna delle coppie di eventi E_i e E_j non implica la 4.23.

Come esempio di applicazione della (4.22) consideriamo il labirinto della figura 4.4. La probabilità di arrivare al tesoro (T) passando prima per il bivio 5 vale (associamo alla biforcazione numero i l'evento E_i).

$$\begin{aligned} P(T \,|\, E_1 \cap E_2 \cap E_3 \cap E_4 \cap E_5) &= P(1) \cdot P(E_2 \,|\, E_1) \cdot P(E_3 \,|\, E_1 \cap E_2) \cdots \\ &= 1 \cdot \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{72} \,. \qquad (4.24) \end{aligned}$$

4.9.2 Legge delle alternative

Ricordiamo la formula che esprime la probabilità di un certo evento E come somma delle probabilità di tutti i prodotti logici dell'evento E con ciascuno degli eventi di una classe completa (vedi 4.6):

$$P(E) = \sum_{i=1}^{n} P(E \cap H_i) \qquad \text{se} \left\{ \begin{array}{l} H_i \cap H_j = \emptyset \quad \forall i \neq j \\ \bigcup_{i=1}^{n} H_i = \Omega \end{array} \right. \tag{4.25}$$

Applicando a ciascuno dei prodotti logici il teorema delle probabilità composte otteniamo

$$P(E) = \sum_i P(H_i) \cdot P(E \,|\, H_i)\,, \tag{4.26}$$

conosciuta come *legge delle alternativa* o *formula di disintegrazione* (si incontra anche la denominazione *formula delle probabilità totali*, ma non va confusa con l'omonima (4.5)!).

Anche se la relazione è una semplice conseguenza della (4.6) e del teorema delle probabilità composte, essa è molto importante sia per le applicazioni che per una migliore "visualizzazione" della probabilità condizionata. Infatti:

- riscrivendo la 4.26 come

  $$P(E) = \frac{\sum_i P(H_i) \cdot P(E \,|\, H_i)}{\sum_i P(H_i)} \tag{4.27}$$

  (in quanto $\sum_i P(H_i) = 1$) si mette in luce come la probabilità di un qualsiasi evento sia pari alla *media delle sue probabilità condizionate dagli eventi di una classe completa, pesate con le probabilità dei condizionanti*;

- un altro punto di vista interessante, già usato nell'esempio del labirinto di figura 4.4, è quello dei percorsi alternativi con cui si può "arrivare" all'evento E. (vedi figura 4.5). Ne seguono due regole pratiche:

  1. la probabilità di seguire un certo percorso è pari al prodotto delle probabilità lungo il percorso (rilettura del teorema delle probabilità composte: il dato "percorso" i-mo è equivalente al verificarsi sia di E che di H_i);
  2. la probabilità di un evento è pari alla somma della probabilità dei percorsi che conducono a tale evento (rilettura della legge delle alternative).

4.10 Indipendenza logica e indipendenza stocastica

4.11 Ricapitolando

- Qualunque sia il modo per valutare la probabilità, questa deve soddisfare le tre regole di base:

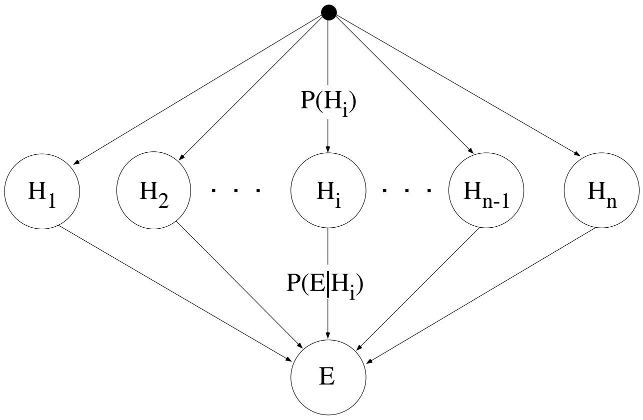

Figura 4.5: Legge delle alternative (o formula di disintegrazione) nel punto di vista dei percorsi alternativi che conducono ad un certo evento.

1. $0 \leq P(E) \leq 1$,
2. $P(\Omega) = 1$ e $P(\emptyset) = 0$,
3. $P(E_1 \cup E_2) = P(E_1) + P(E_2)$ se $E_1 \cap E_2 = \emptyset$,

ove Ω indica l'evento certo, $\emptyset$ l'evento impossibile, e gli operatori $\cup$ e $\cap$ indicano rispettivamente la somma logica e il prodotto logico.

- Le regole di base sono anche note come "assiomi", secondo l'approccio assiomatico introdotto da Kolmogorov. In questo approccio la probabilità è semplicemente un numero reale che soddisfi gli assiomi, disinteressandosi di come essa vada valutata.

- Dalle regole di base e dalle operazioni logiche fra gli eventi seguono altre proprietà delle probabilità. Fra le più importanti c'è quella della probabilità dell'unione logica nel caso generale:

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

- L'uguaglianza fra l'algebra degli eventi e l'algebra degli insiemi permette di associare gli uni agli altri e, in particolare, di visualizzare le operazioni logiche mediante i diagrammi di Venn.

- Il concetto dell'evento condizionato di E rispetto ad H è legato alla valutazione della probabilità che verifichino sia E e H, ma subordinatamente all'ipotesi che H sia vera: ne segue che $P(E\,|\,H) \geq P(E \cap H)$. Più esattamente, $P(E\,|\,H)$ e $P(E \cap H)$ sono legate dal teorema delle probabilità composte:

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

- Nell'approccio assiomatico la probabilità condizionata è invece definita dalla probabilità del prodotto logico divisa la probabilità del condizionante. Questo punto di vista non è affatto naturale e provoca solo inconvenienti se preso alla lettera.

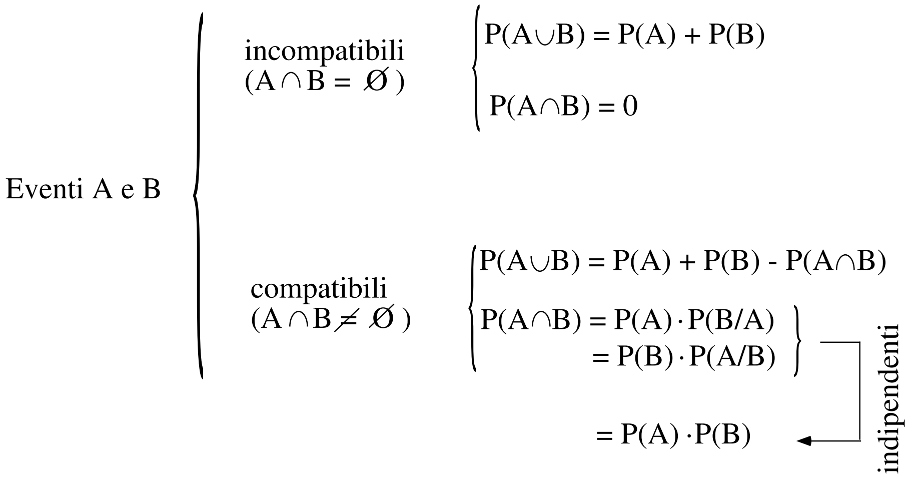

Figura 4.6: Probabilità della somma logica e del prodotto logico di eventi compatibili e non.

- Il concetto di indipendenza stocastica va distinto da quello di indipendenza logica o fisica degli eventi e consiste nel richiedere che il condizionamento non modifichi la valutazione di probabilità, ossia $P(E \,|\, H) = P(E)$, da cui:
$$P(E \cap H) = P(E) \cdot P(H)\,.$$
- Lo schema della figura 4.6 riassume le varie possibilità del calcolo della probabilità di somma e prodotto logico di eventi.
- Il teorema delle probabilità composte può essere esteso al prodotto logico di n eventi:
$$P(E_1 \cap E_2 \cap \cdots \cap E_n) \;=\; P(E_1) \cdot P(E_2 \,|\, E_1) \cdot P(E_3 \,|\, E_1 \cap E_2) \cdots\,,$$
che nel caso di eventi indipendenti si riduce al prodotto delle probabilità:
$$P(E_1 \cap E_2 \cap \cdots \cap E_n) \;=\; P(E_1) \cdot P(E_2) \cdot \cdots \cdot P(E_n)\,.$$
- Una proprietà importante è quella che permette di decomporre un evento in
$$P(E) = \sum_{i=1}^{n} P(E \cap H_i)\,,$$
dove gli n eventi H_i formano una classe completa. Applicando a ciascun termine della decomposizione il teorema delle probabilità composte si ottiene la legge delle alternative (o formula di disintegrazione)
$$P(E) = \sum_{i=1}^{n} P(H_i) \cdot P(E \,|\, H_i)\,.$$

4.12 Problemi

1. Dati i eventi A, B e C, scrivere l'espressione dei seguenti eventi:

   (a) E_1 = "si verifica soltanto A";

   (b) E_2 = "si verificano sia A che B che C";

   (c) E_3 = "non si verifica né A, né B, né C";

   (d) E_4 = "si verifica almeno uno degli eventi";

   (e) E_5 = "si verificano al più due eventi";

   (f) E_6 = "si verificano A e B, indipendentemente dall'esito di C";

   (g) E_7 = "si verifica B indipendentemente dall'esito degli altri";

   (h) E_8 = "si verifica A, ma non B, indipendentemente dall'esito di C";

   (i) E_9 = "si verificano tutti o ("OR") nessuno degli eventi"

   (j) E_{10} = "si verifica A indipendentemente dagli altri due, o ("OR") si verifica uno degli altri due"

2. Consideriamo i seguenti eventi: A = "una macchina è della marca X"; B = "una macchina è bianca". Sapendo che le probabilità di A vale 1/3, la probabilità di B vale 1/2 e che il 50 % delle macchine X sono bianche:

   (a) calcolare le probabilità dei seguenti eventi: $P(\overline{A})$; $P(\overline{B})$; $P(A \cap B)$; $P(A \cup B)$; $P(\overline{B}|A)$; $P(\overline{A} \cup B)$; $P(A \cup \overline{B})$; $P(\overline{A} \cap \overline{B})$; $P(\overline{A} \cup \overline{B})$; $P(\overline{A} \cap B)$; $P(A \cap \overline{B})$;

   (b) esprimere a parole gli eventi di cui è richiesto il calcolo della probabilità al punto a);

   (c) supponendo che il numero di macchine sia talmente elevato per cui la probabilità di osservare un certo tipo di macchina non sia influenzato dall'osservazione precedente, calcolare: la probabilità di osservare due volte consecutive una macchina bianca; la probabilità che dopo una macchina X segua una macchina non bianca;

3. Una persona crede che $P(E) = 0.7$ e $P(\overline{E}) = 0.4$. È possibile tale assegnazione di probabilità? Si immagini che, con tale convinzione, egli faccia due scommesse legate a tale evento nelle quali punta 7000 lire per ricevere 10000 se l'evento si verifica e punta 4000 lire per ricevere 10000 se l'evento non si verifica. Fareste mai simili scommesse?

4. Un allibratore improvvisato ritiene che $P(E) = 0.5$ e $P(\overline{E}) = 0.4$. Pertanto accetta una scommessa di 5000 per pagarne 10000 se l'evento si verifica, e contemporaneamente una scommessa di 4000 per pagarne 10000 se l'evento non si verifica. Sta facendo un buon affare?

5. Consideriamo due eventi arbitrari E_1 e E_2. Una persona attribuisce ad essi $P(E_1) = 0.4$, $P(E_2) = 0.5$ e $P(E_1 \cup E_2) = 0.6$. Le valutazioni sono coerenti?

6. Una persona ritiene che $P(A) = 0.6$, $P(B) = 0.7$. È possibile? Quanto deve essere il valore minimo della probabilità del prodotto logico dei due eventi?

7. Quanto vale la probabilità che, estraendo da un mazzo di carte italiane una carta, questa sia una coppe o una figura?

8. Definiti gli eventi A e B, una persona valuta $P(A) = 1/2$, $P(B) = 1/16$ e $P(A \cap B) = 1/4$. Verificare se l'assegnazione delle probabilità è coerente.

9. Una persona si dichiara convinta che una squadra di calcio ha una probabilità dell' 80 % di passare in vantaggio al primo tempo e di vincere poi l'incontro. Successivamente afferma che la probabilità che la squadra passi in vantaggio al primo tempo è del 50 %? Sono coerenti le sue affermazioni?

10. Un'urna contiene 4 palline nere e 3 rosse. Si estraggono a caso 3 palline senza rimetterle nell'urna. Quanto vale la probabilità che rispettivamente nessuna, una, due e tre palline siano rosse?

11. Rispondere alle stesse domande dell'esercizio precedente nel caso che le palline vengano reintrodotte nell'urna.

12. Un'urna contiene una pallina bianca, una rossa e una nera. Una seconda urna ne contiene una bianca, una rossa, una nera e una gialla. Si estrare una pallina da ciascuna delle urne. Identifichiamo l'eventi con l'iniziale del colore e un indice che vale 1 o due a seconda dell'urna. Determinare la probabilità dei seguenti eventi:

    (a) B_1;

    (b) R_2;

    (c) $A = B_1 \cap G$;

    (d) $E_1 = B_1 \cup R_1$;

    (e) $F_1 = B_1 \cup N_1$

    (f) $E_2 = B_2 \cup R_2$;

    (g) $F_2 = B_2 \cup N_2$;

(h) $H = B_2 \cup G$.

Verificare inoltre se sono fra di loro indipendenti:

(a) E_1 e F_1;

(b) E_2 e F_2;

(c) E_2 e H;

(d) E_2, F_2 e H.

13. Roulette russa: ogni giocatore ruota a caso il tamburo di una pistola e poi si spara. Qual'è la probabilità di sopravvivere dopo n prove, se ogni volta si fa ruotare il tamburo? (La pistola ha sei colpi)

14. Variante (col morto) della roulette russa. Sei giocatori decidono di ruotare a caso una sola volta il tamburo e di provare uno dopo l'altro. Quale dei giocatori ha maggiore probabilità di sopravvivere?

15. È più probabile che esca un 'sei" su 4 lanci di un dado o un "doppio sei" su 24 lanci di due dadi? (Questo è uno dei problemi classici della probabilità proposto dal Cavalier de Méré a Pascal.)

16. Quante volte bisogna lanciare un dado per essere sicuri al 99 % che esca una certa faccia?

17. In relazione al problema precedente: quanto vale la probabilità che la faccia prescelta si verifichi al venticinquesimo lancio se non si è verificato nei precedenti ventiquattro? Dare sia la risposta intuitiva che quella che si ottiene valutando

$$P(E_{25} \,|\, \overline{E_1 \cap \cdots \cap E_{24}})\,.$$

18. Ai tre eventi mutualmente incompatibili E_1, E_2 ed E_3 vengono assegnate probabilità $P(E_1)$, $P(E_2)$ e $P(E_3)$ (ad esempio, rispettivamente 10, 20 e 30 %). Mostrare come la probabilità di ciascuno degli eventi, subordinata al verificarsi di uno dei tre, è pari a $P(E_i)/\sum_i P(E_i)$.

19. Tre amici (Antonio, Berto e Carlo) scommettono sull'uscita del numero 51 sulla ruota di Venezia. I termini della scommessa sono i seguenti: se il numero esce alla prima settimana di attesa vince A, se esce alla seconda settimana vince B, se alla terza vince C. Se il numero non esce entro le prime tre settimane la scommessa viene invalidata. Chi ha maggiore probabilità di vincere la scommessa? Quanto devono puntare B e C se A punta 10˙000 lire, affinché la scommessa sia equa?

20. Un quotidiano[6] pubblica i pronostici per il totocalcio riportati nella seguente tabella, espressi in probabilità (in %) dei diversi segni (per curiosità è riportata anche la schedina vincente). Il giornalista si era basato, apparentemente, su sole considerazioni tecniche[7].

	1	X	2	Ris.
1	20	40	40	1
2	30	25	45	1
3	40	25	35	X
4	45	20	35	1
5	50	20	30	1
6	45	35	20	X
7	40	35	25	2
8	60	30	10	1
9	35	40	25	1
10	30	25	45	X
11	35	40	25	2
12	35	30	35	2
13	45	35	20	X

Valutare la probabilità della(e) colonna(e) ritenuta(e) più probabile(i) e meno probabile(i). Confrontare con quanto si otterrebbe se tutti i segni avessero la stessa probabilità. È più probabile la colonna con tutti 1 o la colonna "12X12X12X12X1"?

21. Ammettiamo che due amici (A e B) concordino con i pronostici del problema precedente e vogliano fare la seguente scommessa: A punta su tutti segni 1 e B punta una sequenza alternata di 1 e X, a partire da 1. Se non si verifica nessuno dei due eventi la scommessa è invalidata. Quanto deve puntare ciascuno affinché la scommessa sia equa?

22. In un cassetto sono riposti disordinatamente 10 calzini bianchi, 10 rossi e 10 neri. Quanto vale la probabilità che, alzandosi la mattina al buio e prendendo due calzini a caso, se ne estraggano due dello stesso colore? E se se ne prendono 3? Quanti calzini bisogna prendere per avere la certezza che fra di essi ve ne siano due dello stesso colore?
Quanto vale invece la probabilità che estraendo 4 calzini ce ne siano almeno due del colore del primo calzino estratto? Dopo quanti calzini si ha la certezza di avere almeno due calzini dello stesso colore del primo estratto?

23. Qual'è la probabilità che giocando a tressette (gioco italiano da 40 carte) un giocatore non riceva

[6]La Repubblica, sabato 11/9/1993.

[7]Spesso si incontrano in altre pubblicazioni anche elementi scaramantici, legati all'alternanza dei segni o ai segni precedentemente usciti.

nemmeno un “pezzo”? (A questo gioco si chiamano pezzi l’Asso, il Due e il Tre). Qual’è la probabilità che questo capiti ad almeno due giocatori? E a tre?

24. Riprendiamo il problema 4 del capitolo 1 delle due scatole di cui una contiene 8 palline bianche e 2 nere e l’altra 2 bianche e 8 nere. Si estrae una pallina da una scatola scelta a caso e, senza guardarla, la si ripone nell’altra scatola. Successivamente si estrae una pallina da quest’ultima scatola. Calcolare la probabilità che la pallina sia bianca facendo i conti dettagliati.

25. In una razza di cani il gene N che determina il mantello nero domina su quello responsabile del mantello rosso (gene n). Una cagna nera dal genotipo eterozigote (N, n) viene incrociata con un maschio nero avente lo stesso genotipo. Quanto vale la probabilità che su 5 cuccioli non ce ne sia nemmeno uno rosso? (Trascurare la possibilità di gemelli monoovulari.)

26. Una persona fa il seguente ragionamento per calcolare la probabilità che ci sia vita animale su Marte (la storiella è antecedente le missiomi spaziali): siccome non so niente, la probabilità che ci siano i felini è del 50 %; ugualmente c’è il 50 % di probabilità che ci siano rettili; e così via. Ne segue che considerando n specie animali, la probabilità che non ce ne sia nessuna è pari a 0.5^n, da cui segue che $P(\text{vita}) = 1 - 0.5^n$. Quindi la probabilità di vita tende a 1 pur di considerare un numero se il ragionamento viene esteso a molte specie specie. Questa amenità viene presentata da taluni come un paradosso derivante dall’uso della probabilità soggettiva. Dove stanno gli errori di ragionamento?

27. Una persona ritiene che la probabilità che una squadra vinca un’incontro di calcio sia dell’80 %, mentre la probabilità che essa termini il primo tempo in vantaggio sia del 60 %. Inoltre valuta nel 90 % la probabilità che la squadra vinca l’incontro qualora cominci in vantaggio il secondo tempo. Quanto vale la probabilità che essa vinca l’incontro pur non essendo in vantaggio al primo tempo?

28. Riferendosi al labirinto di figura 4.4: quanto vale la probabilità di arrivare al tesoro e uscire vivi?

29. Consideriamo il problema 33 del capitolo 2. Quanto vale la probabilità che entrambe le palline siano dello stesso colore se riteniamo che la procedura di preparazione sia al 70 % la prima descritta nella soluzione di tale problema e al 30 % la seconda?

30. Riprendiamo il problema 25 del capitolo 2: si immagini che che il concorrente non si fidi del presentatore e pensi che possa aver bluffato. È ancora conveniente cambiare scatola? Ad esempio quanto vale la probabilità che il premio sia nell’altra scatola se crede al 50 % che il presentatore possa aver bluffato?

31. Quanto vale la probabilità che il secondo estratto al lotto sia il 25? (Si consiglia di usare, anche se non è strettamente necessario, la legge delle alternative).

32. Quanto vale la probabilità che il terzo estratto sia il 16, nell’ipotesi che il secondo estratto sia il 47? E se si venisse a sapere anche che il primo estratto era il 58?

33. Ancora sui problemi 24 e 25 del capitolo 2: Risolvere i problemi in modo più formale usando la legge delle alternative.

Capitolo 5

Probabilità delle cause e meccanismo di aggiornamento delle probabilità

...questi problemi sono classificati come
probabilità delle cause e sono i più importanti
di tutti per le loro applicazioni scientifiche ...

Un effetto potrebbe essere prodotto dalla causa a
o dalla causa b. *L'effetto è appena stato osservato.*
Ci domandiamo la probabilità che sia dovuto alla causa a.
Questa è una probabilità di causa a posteriori.
Ma non la potrei calcolare, se una convenzione più o
o meno giustificata non mi dicesse in anticipo qual'è
la probabilità a priori che la causa a *entri in gioco.*

(H. Poincaré)

5.1 Inferenza probabilistica

Un modo interessante di rileggere la probabilità condizionata è di pensare il condizionante *causa* dell'evento (visto come *effetto*). Questo vale, ad esempio, se si considerano gli eventi condizionati:

- "sopravvive almeno cinque anni" | "ha subito un trapianto al fegato";

- "auto rubata" | "auto nuova e di valore";

- "l'ago della bilancia si posiziona su 1000.00 g" | "chilogrammo campione su bilancia commerciale di laboratorio";

- "si registrano x conteggi in un contatore di radioattività" | "la radioattività ambientale vale r";

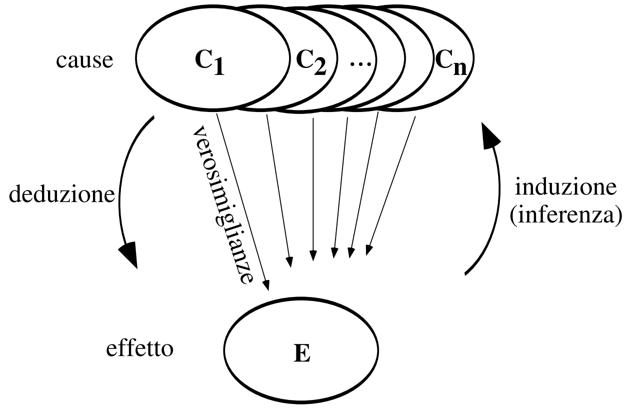

Figura 5.1: Deduzione e induzione

- "si osservano x globuli bianchi in un piccolo campione di sangue osservato al microscopio" | "il sangue contiene n globuli bianchi per unità di volume".

Il condizionante viene visto come causa che può provocare i vari effetti corrispondenti agli eventi condizionati (vedi figura 5.2):

$$P(E \,|\, H) \Longleftrightarrow P(\text{Effetto} \,|\, \text{Causa})\,. \tag{5.1}$$

La causa è da intendersi inoltre in senso lato, come *teoria* che può dar luogo a dei fenomeni osservati, ovvero *grandezza fisica* responsabile dei valori registrati dagli strumenti:

$$P(\text{fenomenologia} \,|\, \text{teoria} \,\cap\, \text{condizioni al contorno})$$

ovvero

$$P(\text{osservazioni strumentali} \,|\, \text{grandezza} \,\cap\, \text{fattori di influenza})\,.$$

Nel condizionante sono stati esplicitate anche altre cause concomitanti che sono ugualmente responsabili degli effetti osservati. *"Condizioni al contorno"* e *"fattori di influenza"* indicano genericamente tutte le condizioni iniziali o altri fattori esterni alla teoria, calibrazioni degli strumenti, variabili e disturbi ambientali, e così via.

Si intuisce che, come indicato nella citazione di Poincaré posta all'inizio del capitolo, se si riesce a trovare una regola per *invertire la probabilità* e valutare $P(H \,|\, E)$ a partire da $P(E \,|\, H)$, questa potrà essere utilizzata come una via per imparare dall'esperienza, alternativa allo scetticismo di Hume o al falsificazionismo popperiano. In questo capitolo indicheremo la possibile strada da seguire e mostreremo che i risultati che si ottengono sono sensati. Tentare di dimostrare invece la "certezza" di questo approccio sarebbe invece un paradosso autoreferenziale, anche se qualcuno osa dire che "nei ragionamenti in condizioni di incertezza l'unica cosa che oggigiorno sembra certa sia come trattare le incertezze".

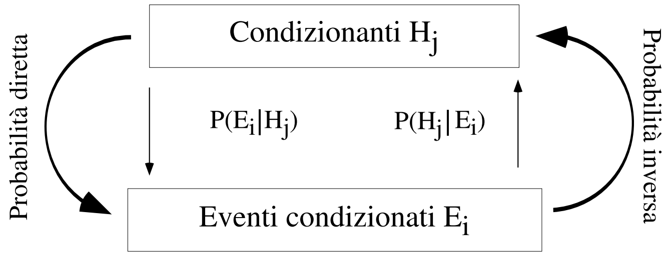

Figura 5.2: Schema di inversione della probabilità

Prima di proseguire è importante aggiungere due note, per evitare che la visione causa-effetto stravolga il concetto più generale - conoscitivo - della probabilità condizionata. Innanzitutto ricordiamo che la probabilità condizionata $P(E\,|\,H)$ è da intendersi come probabilità di E *nell'ipotesi che* H sia vera. Quindi non va intesa come se fosse calcolabile soltanto *dopo* che H risulta essere vera. In secondo luogo, l'evento condizionante è da intendersi come uno stato di informazione e non deve essere vista in funzione strettamente causale. Questo risulterà chiaro fra poco, quando, nell'inversione di probabilità l'evento che consideriamo effetto giocherà il ruolo di condizionante rispetto alla conoscenza sulla "causa fisica".

5.2 Teorema di Bayes

Consideriamo un evento E e una classe completa di ipotesi H_i, ovvero - ricordiamo - che siano esaustive e mutuamente esclusive:

$$\begin{cases} H_i \cap H_j = \emptyset \ \ (i \neq j) \\ \bigcup_{i=1}^n H_i = \Omega \end{cases}$$

Applicando la formula della probabilità composta si ottiene

$$P(E \cap H_i) = P(E) \cdot P(H_i \mid E) = P(H_i) \cdot P(E \mid H_i) \tag{5.2}$$

da cui segue

$$P(H_i\,|\,E) = \frac{P(H_i) \cdot P(E\,|\,H_i)}{P(E)}\,, \tag{5.3}$$

ovvero

$$\frac{P(H_i\,|\,E)}{P(H_i)} = \frac{P(E\,|\,H_i)}{P(E)}\,. \tag{5.4}$$

(Per la validità di queste due equazioni non è necessaria la condizione che le ipotesi H_i formino una classe completa.)

Utilizzando la formula di disintegrazione della probabilità (4.26), valida sotto ipotesi di classe completa degli H_i (e che riportiamo per comodità

$$P(E) = \sum_j P(E\,|\,H_j) \cdot P(H_j)\,,$$

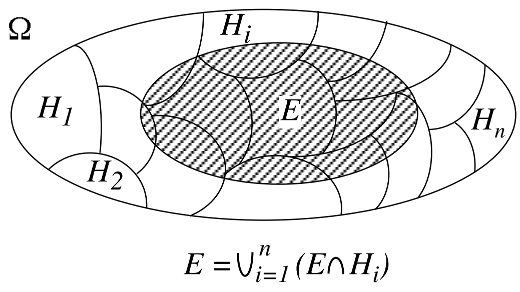

Figura 5.3: Decomposizione di un evento nell'unione dei prodotti logici con gli elementi di una partizione finita H_i.

insieme alla figura 5.3 della decomposizione dell'evento E da cui la (4.26) discende) si ottiene:

$$P(H_i \mid E) = \frac{P(E \mid H_i) \cdot P(H_i)}{\sum_j P(E \mid H_j) \cdot P(H_j)} . \tag{5.5}$$

Questa è la forma standard con cui è conosciuto il *teorema di Bayes*, sebbene (5.3) e (5.4) siano modi alternativi di scriverlo.

Se si osserva che il denominatore non è altro che un fattore di normalizzazione, ossia tale che $\sum_i P(H_i \mid E) = 1$, la formula (5.5) può essere più facilmente memorizzata come

$$P(H_i \mid E) \propto P(E \mid H_i) \cdot P(H_i) . \tag{5.6}$$

Inoltre, fattorizzando $P(H_i)$ al secondo membro della (5.5) e ricordando che in realtà le probabilità sono sempre da intendersi condizionate da una ipotesi implicita che qui indichiamo con $H_\circ$, possiamo riscrivere la formula (5.5) come[1]

$$P(H_i \mid E, H_\circ) = \alpha \cdot P(H_i \mid H_\circ) , \tag{5.10}$$

con

$$\alpha = \frac{P(E \mid H_i, H_\circ)}{\sum_i P(E \mid H_i, H_\circ) P(H_i \mid H_\circ)} . \tag{5.11}$$

[1] Il condizionamento multiplo dell'evento E dalle ipotesi $H_\circ$ e H può essere indicato con $P(E \mid H_\circ \cap H)$ o più semplicemente con $P(E \mid H_\circ, H)$. Quando $H_\circ$ è una ipotesi di base e di cui non si ha alcun interesse di effettuare un inversione di probabilità si può scrivere anche $P_{H_\circ}(E \mid H)$, $P_{H_\circ}(H)$ e $P_{H_\circ}(H \mid E)$, da cui segue che

$$P_{H_\circ}(H \mid E) \propto P_{H_\circ}(E \mid H) \cdot P_{H_\circ}(H) , \tag{5.7}$$

ovvero

$$P(H \mid E \cap H_\circ) \propto P(E \mid H \cap H_\circ) \cdot P(H \mid H_\circ) , \tag{5.8}$$

o anche

$$P(H \mid E, H_\circ) \propto P(E \mid H, H_\circ) \cdot P(H \mid H_\circ) , \tag{5.9}$$

Se si considera la sola ipotesi H_i e la sua opposta $\overline{H}_i$ la (5.5) si riduce a

$$P(H_i\,|\,E) = \frac{P(E\,|\,H_i)\cdot P(H_i)}{P(E\,|\,H_i)\cdot P(H_i) + P(E\,|\,\overline{H}_i)\cdot P(\overline{H}_i)}\,, \tag{5.12}$$

da cui è possibile riscrivere il teorema di Bayes ancora in un'altra forma:

$$P(H_i\,|\,E) = \frac{P(H_i)}{P(H_i) + P(\overline{H}_i)\cdot \frac{P(E\,|\,\overline{H}_i)}{P(E\,|\,H_i)}}\,. \tag{5.13}$$

Un'ultima forma, molto conveniente perché lega due qualsiasi ipotesi H_i e H_j incompatibili (anche se non formano una classe completa) è

$$\frac{P(H_i\,|\,E)}{P(H_j\,|\,E)} = \frac{P(E\,|\,H_i)}{P(E\,|\,H_j)}\cdot\frac{P(H_i)}{P(H_j)} \tag{5.14}$$

Questi diversi modi di scrivere essenzialmente la stessa formula riflettono l'importanza di questo semplice teorema.

5.3 Chiavi di lettura del teorema di Bayes

Dei diversi modi di scrivere il teorema di Bayes la formula (5.5) è quella più pratica da usare, anche se a volte è da preferire la (5.3) se $P(E)$ è nota direttamente. Le altre versioni offrono interessanti modi di interpretare il teorema.

- La (5.4) mostra che

  > *$P(H_i)$ è modificata dall'ipotesi che E sia vera dello stesso fattore dell quale $P(E)$ è modificata dall'ipotesi che H_i sia vera.*

  Ad esempio, se stimo che la probabilità che una squadra di calcio vinca un incontro raddoppi qualora venissi a sapere che nel primo tempo è in vantaggio, ne segue che, se so che la squadra ha vinto l'incontro, la probabilità che essa fosse già in vantaggio al primo tempo diventa il doppio di quanto non l'avessi valutata precedentemente.

- La (5.6) rappresenta il modo più semplice e intuitivo di formulare il teorema:

  > *la probabilità di H_i dato E è proporzionale alla probabilità di H_i per la probabilità di E dato H_i.*

- Infine la (5.10-5.11) mostra come la probabilità di una certa ipotesi è aggiornata quando cambia lo stato di informazione

  $\boxed{P(H_i\,|\,H_\circ)}$ (anche indicata come $P_{H_\circ}(H_i)$ o, ancora più succintamente, come $P_\circ(H_i)$) è la probabilità *iniziale*, o *a priori*, ovvero la probabilità di H_i condizionata da tutte le ipotesi preliminari $H_\circ$, con esclusione del verificarsi o meno di E;

$\boxed{P(H_i \,|\, E, H_\circ)}$ (o semplicemente $P(H_i \,|\, E)$) è la probabilitità *finale*, o *a posteriori*: è la probabilità di H_i riaggiornata alla luce dell'ipotesi che E sia vero;

$\boxed{P(E \,|\, H_i, H_\circ)}$ (o semplicemente $P(E \,|\, H_i)$) è la *verosimiglianza*, ovvero una misura di quanto è verosimile E alla luce di H_i (o, in termini di causa ed effetto, di quanto facilmente H_i possa produrre E). Si noti come anche la verosimiglianza sia valutata subordinatamente allo stato di conoscenza $H_\circ$.

- Infine la (5.14) mostra il riaggiornamento del rapporto delle probabilità (e quindi delle quote di scommessa) subordinatamente all'ipotesi E:

  il rapporto delle probabilità viene modificato dal rapporto delle verosimiglianze

  $$\frac{P(E \,|\, H_i)}{P(E \,|\, \overline{H}_i)}\,, \tag{5.15}$$

  denominato *fattore di Bayes*.

  Questo modo di vedere il teorema è molto pratico quando non ha interesse, o non è pratico, considerare tutte le ipotesi che formano una classe completa. Ad esempio nella ricerca scientifica non ha senso chiedersi la *probabilità assoluta* di una teoria, ma ha soltanto senso valutarne la credibilità rispetto ad una o più concorrenti. Quindi dalla (5.14) è possibile calcolarsi[2]

  $$P(H_1 \,|\, (H_1 \cup H_2) \cap E) = \frac{P(E \,|\, H_1) \cdot P(H_1)}{P(E \,|\, H_1) \cdot P(H_1) + P(E \,|\, H_2) \cdot P(H_2)}\,. \tag{5.16}$$

Riassumendo, il teorema di Bayes può essere sintetizzato dalla seguente relazione:

$$\text{probabilità finale} \;\propto\; \text{verosimiglianza} \times \text{probabilità iniziale},$$

ovvero

$$\text{probabilità a posteriori} \;\propto\; \text{verosimiglianza} \times \text{probabilità a priori}.$$

Esso permette di affrontare i cosiddetti problemi di *probabilità inversa*, primo fra tutti i problemi della probabilità delle cause. Infatti si può riformulare a parole dicendo che (vedi figura 5.4):

se ci sono più cause che possono produrre lo stesso effetto, la probabilità che l'effetto osservato sperimentalmente derivi da una delle cause è proporzionale alla probabilità di tale causa per la probabilità che essa produca l'effetto:

$$P(\text{Causa}_i \,|\, E) \propto P(E \,|\, \text{Causa}_i) \cdot P_\circ(\text{Causa}_i)\,. \tag{5.17}$$

[2] Si noti che $P(H_1 \cup H_2 \,|\, (H_1 \cup H_2) \cap E) = 1$ se $H_1 \cup H_2 \neq \emptyset$.

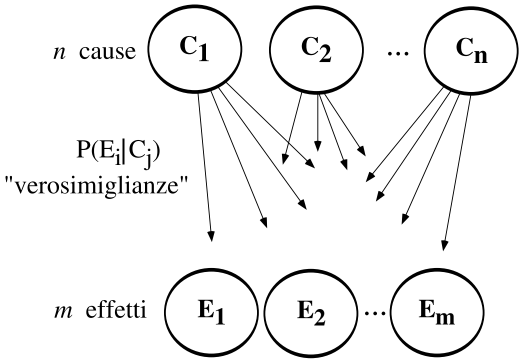

Figura 5.4: Relazioni cause-effetti viste in termini di condizionanti e eventi condizionati

5.4 Visione combinatoria del teorema di Bayes

Includiamo un esempio in cui si mostra come arrivare alla formula di Bayes da considerazioni combinatorie, ad uso di coloro che si trovano più a loro agio con esempi di questo tipo. Si tenga conto che tale visione è comunque fortemente limitativa in quanto non è di uso generale e, in particolare, non è utilizzabile nella valutazione della probabilità delle cause e quindi in tutti i casi di inferenza legata alle misure.

Supponiamo che il mercato delle automobili sia costituito da tre sole fabbriche A, B e C, le quali detengono rispettivamente le seguenti quote di mercato: 70%, 25% e 5%. Supponiamo inoltre di sapere che dopo dieci anni siano ancora funzionanti: il 6% delle auto A, il 22% delle B e il 75% delle C. Un amico ci dice di avere un'auto di oltre 10 anni. Qual'è la probabilità che sia una A?
Un modo di risolvere il problema è di considerare il numero di auto prodotte e quello di quelle circolanti dopo dieci anni, come schematizzato nella seguente tabella:

Marca	Quota	P($\geq$ 10 anni)	N_0	$N_{\geq 10}$
A	70 %	6 %	$0.7 \times N$	$0.7 \times 0.06 \times N$
B	25 %	22 %	$0.25 \times N$	$0.25 \times 0.22 \times N$
C	5 %	75 %	$0.05 \times N$	$0.05 \times 0.75 \times N$

Troviamo quindi che la probabilità cercata è:

$$P = \frac{0.06 \times 0.7}{0.06 \times 0.7 + 0.22 \times 0.25 + 0.75 \times 0.05} = 31\% \,.$$

Interpretando le varie frazioni come probabilità (nel senso di usare la valutazione di probabilità dai casi favorevoli e quelli possibili), possiamo riscrivere questo risultato come

$$P(A\,|\,F) = \frac{P(F\,|\,A) \cdot P(A)}{P(F\,|\,A) \cdot P(A) + P(F\,|\,B) \cdot P(B) + P(F\,|\,C) \cdot P(C)}\,,$$

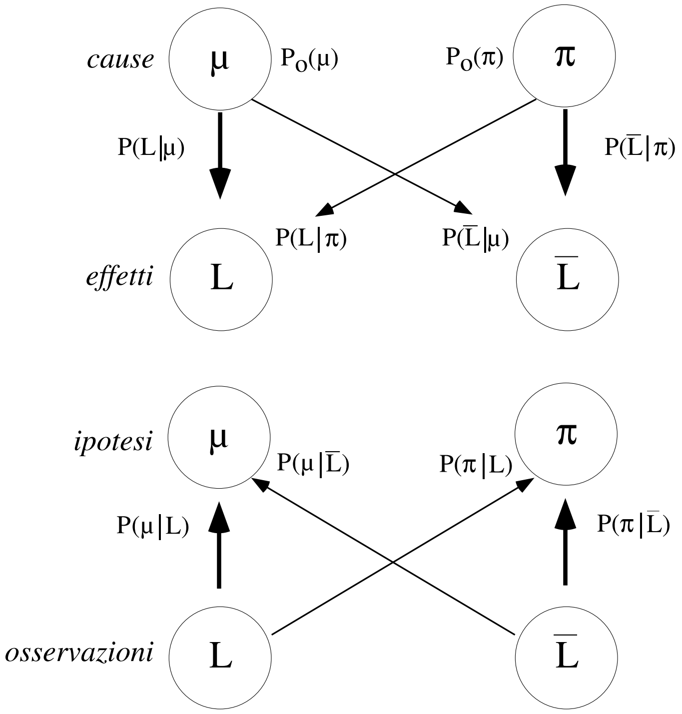

Figura 5.5: Schematizzazione in termini induttivi e deduttivi del problema dell'identificazione di particelle con strumento non ideale. Lo spessore delle freccie indica le direzioni favorite dei processi di deduzione e di induzione in questo esempio.

dove F sta per "funziona dopo 10 anni". Si riconosce in questa formula particolare la (5.5).

5.5 Esempi tipici di applicazione

Sebbene il teorema di Bayes sia strutturalmente molto semplice, se ne può capire fino in fondo la portata soltanto mediante opportuni esempi.

5.5.1 Classificazione di eventi e rapporto segnale rumore

Supponiamo che un rivelatore abbia un'efficienza di identificazione delle particelle μ del 95 % e una probabilità di confondere una particella π per un μ del 2 %. Se una particella viene identificata come μ si accende una lampadina (L). Conoscendo che le particelle del fascio contengono il 10 % di μ e il 90 % di π:

- quanto vale la probabilità che, se si accende la lampadina, sia passato un μ?

- quanto vale la probabilità che, se non si accende la lampadina, sia passato un π?

- Quanto vale il rapporto segnale rumore (S/N)? (Definiamo S/N come probabilità di μ diviso probabilità di π se si è accesa la lampadina, considerando la particella μ come "segnale", ovvero il fenomeno di interesse.)

- Come cambiano i risultati se si pongono sul fascio due contatori aventi le stesse caratteristiche e funzionanti indipendentemente?

Schematizziamo, con il formalismo delle probabilità condizionate, le informazioni che abbiamo a disposizione e le domande a cui vogliamo rispondere (vedi figura 5.5).

$P(L \,|\, \mu) = 0.95$ è la probabilità che si accenda la lampadina se l'ipotesi "μ" è vera, ossia se passa una vera particella μ. Essa viene stimata attraverso la frequenza relativa con la quale si accende la lampadina quando il rivelatore viene esposto ad un fascio "di soli μ" (ammettiamo che sia possibile). L'uso del valore di frequenza relativa implica che

  - le proprietà dei μ non cambiano con il tempo (e questo è pacifico);
  - il rivelatore e l'elettronica si comporteranno durante l'esperimento esattamente come si erano comportati durante il test (e questo può essere questionabile, ma assumiamo che sia vero).

$P(\overline{L} \,|\, \mu) = 0.05$ è la probabilità dell'evento complementare;

$P(L \,|\, \pi) = 0.02$ è la probabilità che la lampadina si accenda "per errore", nel senso che assumiamo che il rivelatore sia stato progettato per identificare i μ e quindi uno strumento perfetto darebbe[3] $P(L \,|\, \mu) = 1$ e $P(L \,|\, \pi) = 0$.

$P(\overline{L} \,|\, \pi) = 0.98$, complementare a $P(L \,|\, \pi)$. Lo strumento ideale darebbe $P(\overline{L} \,|\, \pi) = 1$. Anche per questo dato valgono le note a proposito di $P(L \,|\, \mu)$.

$P_\circ(\mu) = 0.10$ è la probabilità iniziale che una particella che arrivi al rivelatore sia un μ in assenza dell'informazione che la lampadina si sia accesa o no. $P_\circ(\mu)$ sta per $P(\mu \,|\, H_\circ)$, dove $H_\circ$ include tutta l'informazione sulla composizione del fascio.

$P_\circ(\pi) = 0.90$**:** probabilità iniziale dell'evento π. La somma $P_\circ(\mu) + P_\circ(\pi)$ deve dare 1, in quanto gli eventi μ e π formano una classe completa di ipotesi rispetto a questo problema e sono relativi allo stesso stato di informazione $H_\circ$.

$P(\mu \,|\, L)$**:** è la probabilità che la particella abbia attraversato il rivelatore sia un μ se la lampadina si è accesa. La complementare ad essa sarà $P(\pi \,|\, L)$.

[3] Si noti che $P(L \,|\, \mu) = 1$ non implica assolutamente $P(L \,|\, \pi) = 0$ in quanto esse sono probabilità di L relative a diversi stati di informazione.

Chiaramente ci sono tutte le condizioni di applicabilità del teorema di Bayes. Ne segue:

$$\begin{aligned}
P(\mu \,|\, L) &= \frac{P(L \,|\, \mu)\cdot P_\circ(\mu)}{P(L \,|\, \mu)\cdot P_\circ(\mu) + P(L \,|\, \pi)\cdot P_\circ(\pi)} \\
&= \frac{0.95\times 0.1}{0.95\times 0.1 + 0.02\times 0.9} = 0.84 \\
P(\pi \,|\, L) &= 1 - P(\mu \,|\, L) = 0.16 \\
P(\mu \,|\, \overline{L}) &= \frac{P(\overline{L} \,|\, \mu)\cdot P_\circ(\mu)}{P(\overline{L} \,|\, \mu)\cdot P_\circ(\mu) + P(\overline{L} \,|\, \pi)\cdot P_\circ(\pi)} \\
&= \frac{0.05\times 0.1}{0.05\times 0.1 + 0.98\times 0.9} = 0.0056 \\
P(\pi \,|\, \overline{L}) &= 1 - P(\mu \,|\, \overline{L}) = 0.994\,.
\end{aligned}$$

Siamo più sicuri che quando non si accende la luce sia passato un π di quanto non lo siamo quando identifichiamo un μ nel caso contrario. Per capirne il motivo studiamo il rapporto segnale rumore nei due casi:

$$\begin{aligned}
S/N(\mu/\pi) = \frac{P(\mu \,|\, L)}{P(\pi \,|\, L)} &= \frac{P(L \,|\, \mu)}{P(L \,|\, \pi)}\cdot\frac{P_\circ(\mu)}{P_\circ(\pi)} \\
&= 47.5\times\frac{1}{9} = 5.3 \\
S/N(\pi/\mu) = \frac{P(\pi \,|\, \overline{L})}{P(\mu \,|\, \overline{L})} &= \frac{P(\overline{L} \,|\, \pi)}{P(\overline{L} \,|\, \mu)}\cdot\frac{P_\circ(\pi)}{P_\circ(\mu)} \\
&= 19.6\times 9 = 176\,.
\end{aligned}$$

La migliore prestazione nella separazione dei π dai μ non è quindi dovuta alle caratteristiche del rivelatore, bensì alla composizione del fascio.

Questo ci insegna che quando le condizioni sono di *alto rumore*

$$P_\circ(S) \ll P_\circ(N) \tag{5.18}$$

l'esperimento deve essere *molto selettivo*:

$$P(E \,|\, S) \gg P(E \,|\, N)\,, \tag{5.19}$$

dove S, N e E indicano genericamente segnale, rumore (inglese "noise") e esito.

5.5.2 Uso iterativo del teorema di Bayes

Per rispondere all'ultima domanda del problema precedente si può pensare a due possibili approcci:

1. calcolare le probabilità condizionate $P(L_1 \cap L_2 \,|\, \mu)$ e $P(L_1 \cap L_2 \,|\, \pi)$ e usare il teorema di Bayes per trovare $P(\mu \,|\, L_1 \cap L_2)$;

2. utilizzare la probabilità finale condizionata dall'evento μ al posto di $P_\circ(\mu)$ e applicare il teorema di Bayes rispetto al secondo condizionamento di L_2.

Siccome le due soluzioni sono entrambe ragionevoli e non c'è nessun motivo per preferire una via rispetto all'altra, ci attendiamo che, se il metodo di aggiornamento bayesiano è ragionevole, dovremmo arrivare agli stessi risultati. In effetti questo è quanto succede.

1. Seguendo il primo metodo si ha:

$$\begin{aligned} P(L_1 \cap L_2 \,|\, \mu) &= P(L_1 \,|\, \mu) \cdot P(L_2 \,|\, \mu) = 0.95 \times 0.95 = 90.25\,\% \\ P(L_1 \cap L_2 \,|\, \pi) &= P(L_1 \,|\, \pi) \cdot P(L_2 \,|\, \pi) = 0.02 \times 0.02 = 0.04\,\% \\ P(\mu \,|\, L_1 \cap L_2) &= 99.6\,\%\,. \end{aligned}$$

   (Si ricordi l'ipotesi di indipendenza fra le risposte dei due rivelatori. I conti vengono lasciati per esercizio.)

2. Nel secondo caso abbiamo invece:

$$\begin{aligned} P(\mu \,|\, L_1 \cap L_2) &= \frac{P(L_2 \,|\, \mu) \cdot P(\mu \,|\, L_1)}{P(L_2 \,|\, \mu) \cdot P(\mu \,|\, L_1) + P(L_2 \,|\, \pi) \cdot P(\pi \,|\, L_1)} \\ &= \frac{0.95 \times 0.84}{0.95 \times 0.84 + 0.02 \times 0.16} = 99.6\,\% \end{aligned}$$

Si ottiene quindi lo stesso valore di probabilità finale. Questo è un grosso pregio di questo metodo. Infatti come è naturale pensare, le conclusioni scientifiche possono dipendere dalle ipotesi iniziali e dalle informazioni sperimentali, ma non devono dipendere dall'uso che si fa delle informazioni stesse.

L'uso iterativo del teorema di Bayes può essere riassunto dicendo che

> *la probabilità iniziale di una inferenza è pari alla probabilità finale dell'inferenza precedente.*

5.6 ↻ Statistica bayesiana: imparare dall'esperienza

Nell'esempio precedente si era partiti da una stima di probabilità fondata da una esperienza sperimentale più o meno "solida" di tipo "oggettivo" (cioè sulla quale è difficile non convenire[4]). Essa veniva riaggiornata alla luce di verosimiglianze degli eventi osservati.

Consideriamo un altro esempio nel quale il punto di partenza è molto più vago e assolutamente soggettivo (cioè non è, in genere, nemmeno in linea di principio "intersoggettivo"): una persona ascolta un comune amico, collega

[4] ❄In realtà non è difficile convincersi che anche in quei casi, di veramente oggettivo, nel senso che le conclusioni seguano necessariamente dalle osservazioni, c'è ben poco. Riprendiamo il caso di $P(\mu \,|\, L) \propto P(L \,|\, \mu) \cdot P_\circ(\mu)$. Si può immaginare $P_\circ(\mu)$ stimata dalle frequenze relative mediante un rivelatore "perfetto", il quale era stato testato precedentemente con un fascio "puro" di particelle μ, etc. etc. L'inesistenza nella pratica di tali stati di preparazione sperimentale inficia ogni tentativo di passare in modo univoco dalle osservazioni alle conclusioni. Il motivo per il quale ciononostante tali conclusioni vengono considerate "oggettive" è che "è molto improbabile incontrare una persona ragionevole e con cognizione di causa dei problemi di sperimentazione che non si dichiari d'accordo che quella procedura sia corretta". Siamo quindi di fronte a un caso di "intersoggettività assoluta".

o familiare, rispondere al telefono e tenta di indovinare l'identità dell'interlocutore. A secondo della conoscenza che egli ha dell'amico, dell'ora della telefonata, della lingua usata, del tono della voce e del soggetto della conversazione (indichiamo con I l'insieme di tutte queste informazioni), attribuirà una certa probabilità all'evento "l'interlocutore è la persona "X_i". Egli comincia a prendere in considerazione alcune persone e a scartarne altre, finché la massa di informazioni diventa tale da poter essere in grado di identificare con la quasi certezza la persona (anche se non se ne conosce l'identità, come ad esempio "quel signore che gioca a tennis con papà").

Questa esperienza è capitata a tutti e non è difficile riconoscere che il processo induttivo utilizzato si basi sulla valutazione della probabilità relativa delle singole persone secondo lo schema di Bayes:

$$P(X_i\,|\,I) \propto P(I\,|\,X_i)\cdot P_\circ(X_i)\,,$$

dove con $P_\circ(X_i)$ è stata indicata la probabilità iniziale che una persona possa telefonare e con $P(X_i\,|\,I)$ la probabilità finale condizionata dallo stato di informazione I.

Nello schema bayesiano la probabilità finale si riaggiorna ogni qualvolta intervengono nuove osservazioni. $P(I\,|\,X_i)$ invece indica quanto è verosimile che le informazioni siano prodotte dall'interazione con la persona X_i. Per esplicitare il fatto che l'informazione I dipende dal tempo e che sia la probabilità iniziale che la verosimiglianza dipendono dalla conoscenza iniziale $I_\circ$, la relazione dovrebbe essere scritta come

$$P(X_i\,|\,I(t)\cap I_\circ) \propto P(I(t)\,|\,X_i\cap I_\circ)\cdot P(X_i|I_\circ)\,.$$

5.7 ⥁ Il caso del sospetto baro

Poiché daremo una notevole importanza all'approccio bayesiano nella parte dedicata all'inferenza statistica facciamo ora un simpatico esempio relativo ai giochi. Anche se alcuni aspetti possono suonare ameni, questo esempio va preso molto sul serio e studiato molto attentamente, poiché presenta molti paralleli con le problematiche che entrano nelle inferenze su questioni scientifiche.

5.7.1 I "fatti"

Supponiamo che un signore incontri un bel giorno al bar sotto casa un vecchio amico (e non una persona qualsiasi!) e che questo lo inviti a giocarsi l'aperitivo estraendo da un mazzo di carte da gioco una carta ciascuno. Colui che estrae la carta dal valore più alto vince. In caso di parità si estrae di nuovo.

Supponiamo che l'amico vinca. Qual'è la probabilità che egli sia un baro?

Nelle settimane successive ogni tanto i due si rincontrano, si giocano l'aperitivo nello stesso modo e il vecchio amico vince ogni volta. Quanto vale, ad ogni nuova vittoria, la probabilità che l'amico sia diventato un baro?

5.7.2 Riaggiornamento della probabilità

Le ipotesi a cui siamo interessati sono quelle della partizione dell'evento certo $\Omega = \{B, O\}$, ovvero dalle ipotesi che l'amico sia Baro o Onesto. Indichiamo con V_n l'evento "vince n volte consecutive". Supponiamo, per semplificare i conti, che se l'amico è baro vinca sempre ($P(V_n \,|\, B) = 1$), mentre se è onesto vinca secondo le leggi della probabilità, ossia $P(V_n \,|\, O) = (1/2)^n$. Applicando la formula di Bayes otteniamo la probabilità dell'ipotesi Baro subordinata alle n vincite:

$$\begin{aligned} P(B \,|\, V_n) &= \frac{P(V_n \,|\, B) \cdot P_\circ(B)}{P(V_n \,|\, B) \cdot P_\circ(B) + P(V_n \,|\, O) \cdot P_\circ(O)} \\ &= \frac{1 \cdot P_\circ(B)}{1 \cdot P_\circ(B) + \left(\frac{1}{2}\right)^n \cdot P_\circ(O)} . \end{aligned} \tag{5.20}$$

Rimane da assegnare la probabilità iniziale $P_\circ(B)$. Chiaramente un estraneo che ci invitasse a giocare d'azzardo si renderebbe immediatamente sospetti e la probabilità iniziale sarebbe ritenuta prossima a 1. Volendo essere generosi, in quanto si tratta pur sempre di un vecchio amico, fissiamo un valore basso, pari al 5 %: $P_\circ(B) = 0.05$, $P_\circ(O) = 0.95$.

La seguente tabella riporta i valori della probabilità in funzione del numero di vittorie consecutive:

n	$P(B \,\vert\, V_n)$ (%)	$P(O \,\vert\, V_n)$ (%)
0	5.0	95.0
1	9.5	90.5
2	17.4	82.6
3	29.4	70.6
4	45.7	54.3
5	62.7	37.3
6	77.1	22.9
...	...	...

Come naturale, all'aumentare del numero di vittorie consecutive, cresce il sospetto che il vecchio amico stia imbrogliando.

5.7.3 Confronto fra inferenza diretta e inferenza iterativa

Per seguire meglio come nell'approccio bayesiano la probabilità si riaggiorni dopo ogni dato sperimentale, indichiamo con $P(B \,|\, V_{n-1})$ la probabilità assegnata dopo la vincita precedente, e applichiamo ricorsivamente la formula di Bayes. Essendo $P(V \,|\, B) = 1$ e $P(V \,|\, O) = 1/2$ le probabilità di vincere in un singolo tentativo, otteniamo:

$$\begin{aligned} P(B \,|\, V_n) &= \frac{P(V \,|\, B) \cdot P(B \,|\, V_{n-1})}{P(V \,|\, B) \cdot P(B \,|\, V_{n-1}) + P(V \,|\, O) \cdot P(O \,|\, V_{n-1})} \\ &= \frac{1 \cdot P(B \,|\, V_{n-1})}{1 \cdot P(B \,|\, V_{n-1}) + \frac{1}{2} \cdot P(O \,|V_{n-1})} . \end{aligned} \tag{5.21}$$

Il risultato interessante è che si ottengono esattamente gli stessi valori di probabilità ottenuti direttamente, come già mostrato nell'esempio precedente del rivelatore di particelle e come sarà mostrato più formalmente nel paragrafo 5.9

5.7.4 Dipendenza dalla probabilità iniziale

È importante notare la dipendenza della probabilità finale da quelle iniziale. Riportiamo nella seguente tabella le conclusioni a cui sia arriva, in funzione di n per i diversi valori di $P_\circ(B)$:

$P_\circ(B)$	$P(B \mid V_n)$ (%)			
	$n = 5$	$n = 10$	$n = 15$	$n = 20$
$1\,\%$	24	91	99.7	99.99
$5\,\%$	63	98	99.94	99.998
$50\,\%$	97	99.90	99.997	99.9999

5.7.5 Pregiudizio, indizi e conclusioni

La tabella precedente mostra che, come è naturale, all'inizio la probabilità è dominata da $P_\circ$, poi a mano a mano che gli indizi (leggi "osservazioni sperimentali") a favore dell'ipotesi "Baro" aumentano i pregiudizi iniziali diventano sempre meno influenti. Anche la dipendenza delle conclusioni dal pregiudizio e dagli indizi sperimentali segue la prassi di quanto avviene nella ricerca:

- se una teoria risulta già molto credibile dalla comunità scientifica - sia per l'autorevolezza di chi l'ha proposta, sia per ragioni estetiche, sia perché riesce a descrivere in modo economico le osservazioni precedenti - sono sufficienti poche "verifiche" affinché essa sia accettata;

- se una teoria è ben consolidata da secoli di sperimentazione, non bastano poche osservazioni "poco probabili" all'interno della teoria stessa per convincere la comunità scientifica ad abbandonarla. In tal caso si tende a credere di più che sia sbagliato l'esperimento o che si tratti soltanto di una fluttuazione statistica.

In particolare se $P_\circ(B)$ è nulla, la probabilità finale rimarrà sempre nulla. Anche questo è abbastanza naturale: se la persona è di provata fiducia, anche se si verificano degli eventi ritenuti "a priori" poco probabili bisogna semplicemente constatare la loro occorrenza.

5.7.6 Probabilità e decisione

Lo schema di riaggiornamento bayesiano della probabilità modifica il grado di fiducia di una certa ipotesi, ma non può mai giungere a certezze (a meno di non partire già da certezze, ma questa è materia di dogmatici e non di persone di scienza). Alla fine, avendo tratto conclusioni probabilistiche su un certo evento bisogna prendere delle decisioni che sono responsabilità di chi agisce. Ad esempio, nel caso del sospetto baro, si può

- accettare di giocare la $(n+1)$-ma partita con probabilità $P(B \mid V_n)$ di perdere con sicurezza;

- si può rifiutare il prossimo gioco, con probabilità $P(O \mid V_n)$ di offendere l'amico innocente.

Nei casi scientifici la scelta può essere fra mancare il riconoscimento di una grossa scoperta (non pubblicando i risultati) e farsi deridere dalla comunità scientifica dicendo sciocchezze.

5.8 ❄Recupero e superamento del metodo di falsificazione

Abbiamo visto nell'esempio precedente che se $P_\circ(H_i)$ è uguale a zero essa rimarrà sempre tale. Vediamo quali sono, più in generale, le condizioni affiché la probabilità finale sia 0 o 1. Ricordiamo ancora una volta lo schema

probabilità finale $\propto$ **verosimiglianza** $\times$ **probabilità iniziale**.

Ne segue che:

- affinché la probabilità finale si annulli è sufficiente che sia nulla la verosimiglianza relativa ad una qualsiasi osservazione;
- affinché si raggiunga l'assoluta certezza (probabilità finale uguale a 1) è necessario invece
  - che sia diversa da zero la verosimiglianza;
  - ma che sia uguale a 1 la probabilità iniziale[5]

Quindi, se delle osservazioni sono assolutamente incompatibili con una teoria, questa è *falsificata*:

$$P(\text{osservazioni} \mid \text{teoria}) = 0 \quad \Longrightarrow \quad P(\text{teoria} \mid \text{osservazioni}) = 0\,,$$

Al contrario, affinché una teoria sia dichiarata certa è necessario, oltre che spieghi i dati sperimentali, che essa sia ritenuta certa fin dall'inizio. Questa è una posizione dogmatica inaccettabile. Quindi, una teoria può essere falsificata ma non può essere mai ritenuta come certa ("la teoria finale") fintanto che altre teorie sono concepibili e finché esiste il genere umano.

Si noti come, al contrario del metodo usuale di falsificazione, l'induzione probabilistica basata sull'aggiornamento bayesiano, permette di classificare in ordine di credibilità le teorie non falsificate. Questo è in accordo con il modo di procedere "de facto" della ricerca scientifica[6].

[5] Questo si vede bene dalla (5.13). Ad essere precisi, si può ottenere $P(H_i|E) = 1$ anche quando $P(E|\overline{H}_i) = 0$, ovvero quando le osservazioni sono incompatibili con tutte le ipotesi alternative ad H_i prese in considerazione (e non tutte quelle possibili, comprese quelle formulabili nel futuro). Detto in altre parole, in questo caso l'esperimento falsifica tutte le ipotesi complementari formulate fino a quel momento. Quindi una teoria può in effetti assumere un valore "certezza provvisoria" se è la sola, fra quelle formulate, in grado di spiegare le osservazioni sperimentali.

[6] ❄ Se ad ogni passo dello sviluppo delle idee scientifiche ci si fosse dovuti attenere solo alle ipotesi certe …non ci sarebbe stato progresso scientifico, ma una situazione di stallo. La comunità scientifica per fortuna si muove mediamente nella direzione che sembra più credibile, con i successi e gli inevitabili insuccessi che le decisione in stato di incertezza comportano.

5.9 ↻ Osservazioni indipendenti e prodotto delle verosimiglianze

Negli esempi precedenti abbiamo incontrato dei casi di aggiornamento della probabilità ottenuto mediante l'uso iterativo della formula di Bayes. Formalizziamo il risultato, considerando di n osservazioni indipendenti. Trascurando l'inessenziale fattore di normalizzazione otteniamo:

$$\begin{aligned}
P(H_k \,|\, E_2 \cap E_1) &\propto P(E_2 \,|\, H_k) \cdot P(H_k \,|\, E_1) \\
&\propto P(E_2 \,|\, H_k) \cdot P(E_1 \,|\, H_k) \cdot P_\circ(H_k) \\
P(H_k \,|\, E_3 \cap E_2 \cap E_1) &\propto P(E_3 \,|\, H_k) \cdot P(H_k \,|\, E_2 \cap E_1) \\
&\propto P(E_3 \,|\, H_k) \cdot P(E_2 \,|\, H_k) \cdot P(E_1 \,|\, H_k) \cdot P_\circ(H_k) \\
\dots\dots\dots\dots & \quad \dots\dots\dots\dots\dots\dots\dots\dots \\
P(H_k \,|\, E_1 \cap \cdots \cap E_n) &\propto \Pi_{i=1}^n P(E_i \,|\, H_k) \cdot P_\circ(H_k) \qquad (5.22) \\
&\propto P(E_1 \cap \cdots \cap E_n \,|\, H_k) \cdot P_\circ(H_k)\,. \qquad (5.23)
\end{aligned}$$

Quindi la probabilità finale dipende dalla probabilità iniziale per la verosimiglianza delle osservazioni sperimentali, indipendentemente dal loro ordine.

5.10 ❄Fattore di Bayes e incremento logaritmico delle quote di scommessa

Se si parte da una situazione iniziale di incertezza su due ipotesi H_1 e H_2, l'osservazione di una successione di eventi $\{E_1, E_2, \cdots, E_n\}$ altera le probabilità relative ma non può mai arrivare a conclusione certe fintanto le verosimiglianze $P(E_i \,|\, H_j)$ sono diverse da zero. Può accadere invece che gli eventi osservati favoriscano fortemente una delle due ipotesi e quindi che, ad esempio, dopo un certo numero di eventi, $P(H_1 \,|\, H_1 \cup H_2)$ sia molto prossima a 0. Sebbene il valore di zero non sia mai raggiunto è chiaro che 0.1 è ben diverso da 0.001 o da 10^{-9}. Quando le probabilità sono molto piccole quello che interessa veramente è l'ordine di grandezza e quindi sembrerebbe naturale considerarne il logaritmo. Ma questo non funziona per le probabilità molto prossime a 1. Il rapporto delle probabilità è invece un numero compreso fra 0 e infinito che meglio si presta ad essere trattato con i logaritmi, in quanto $\log P(H_1)/P(H_2)$ va da meno infinito a più infinito e vale 0 quando $P(H_1) = P(H_2)$.

Applicando questa idea alla (5.14) si ottiene

$$\log \frac{P(H_i \,|\, E)}{P(H_j \,|\, E)} = \log \frac{P(H_i)}{P(H_j)} + \log \frac{P(E \,|\, H_i)}{P(E \,|\, H_j)}\,, \qquad (5.24)$$

Ossia il logaritmo dei rapporti di credibilità di due eventi è modificato da incrementi pari al logaritmo del fattore di Bayes. Se sono stati osservati n eventi indipendenti si ha

$$\begin{aligned}
\log \frac{P(H_i \,|\, E_1 \cap \cdots \cap E_n)}{P(H_j \,|\, E_1 \cap \cdots \cap E_n)} &= \log \frac{P(H_i)}{P(H_j)} + \log \frac{P(E_1 \,|\, H_i)}{P(E_1 \,|\, H_j)} + \cdots \\
&\quad + \cdots + \log \frac{P(E_n \,|\, H_i)}{P(E_n \,|\, H_j)}\,.
\end{aligned} \qquad (5.25)$$

5.11 ↻ Indifferenza iniziale e massima verosimiglianza

Nel caso che si sia in condizione di indifferenza rispetto ad una classe di ipotesi alternative (ovvero $P(H_i)/P(H_j) = 1\ \forall i, j$) le conclusioni dipendono "soltanto" dalle verosimiglianze. L'ipotesi più probabile diventa allora quella rispetto alla quale è massima la verosimiglianza di osservare E (questo può essere visto molto bene dall'incremento logaritmo dell'evidenza della (5.25) ove il primo termine al secondo membro si annulla).

Questo fatto può indurre a pensare che si possa utilizzare questo criterio al fine di eliminare dal processo di induzione la dipendenza delle probabilità iniziali. È invece ovvio che questa è una cattiva interpretazione del risultato ottenuto, nel senso che la semplificazione delle probabilità iniziali non rende il metodo "oggettivo", ma piuttosto richiede un giudizio di equiprobabilità da parte di chi valuta la probabilità. Inoltre anche le verosimiglianze, essendo probabilità condizionate, dipendono da chi le valuta.

Ritorneremo sul criterio di massima verosimiglianza al momento di trattare i risultati di misure, ma dovrebbe essere chiaro finora che, riprendendo approssimativamente una frase di de Finetti, "se è vero che l'induzione basata sull'inferenza bayesiana è fondata sulla sabbia (le probabilità iniziali) quella basata su un criterio di massimizzazione della verosimiglianza è basata sul niente" (si veda ad esempio l'ultimo problema di questo capitolo).

5.12 ❄Problema della verificabilità ed estensione del concetto di evento

C'è un altro aspetto importante che compare nel problema del sospetto baro e nell'applicazione della probabilità alle ipotesi scientifiche.

Abbiamo detto nel capitolo 2 che un evento deve essere non ambiguo e verificabile (almeno in linea di principio). Nel caso del sospetto baro la verifica significa che lui ad un certo punto confessi di aver imbrogliato, oppure che sia colto in flagrante. Si può invece immaginare che per motivi vari il gioco sia sospeso per cause di forza maggiore e quindi la persona rimanga sempre nel dubbio. Oppure può capitare che egli improvvisamente smetta di essere "così fortunato" e il suo amico pensi che si è trattato soltanto di una coincidenza casuale ... o che semplicemente abbia smesso di imbrogliare. In questi casi l'evento rimarrà per sempre nello stato di incertezza. Si può allora parlare ancora di "evento"? Infatti la definizione di evento data a suo tempo (capitolo 1) prevedeva la verificabilità (almeno in principio) del contenuto di verità dell'affermazione. Quello è il caso delle previsioni meteorologiche, finanziarie, sportive, e così via. Sicuramente non va bene invece per la misura di grandezze fisiche o per stabilire il livello di credibilità di una teoria. Ad esempio, non si può minimamente pensare di arrivare un giorno ad "aprire" un elettrone e di leggere su una etichetta il valore della sua massa e carica elettrica! Questa difficoltà può essere aggirata credendo nell'idea del progresso della scienza, per cui si può ipotizzare che nel futuro si potranno escogitare metodi più raffinati per determinare una grandezza fisica fondamentale, tali da dare "valori più esatti" rispetto a quelli attuali. Ma la scappatoia non può funzionare con le grandezze contingenti che sono misurabili soltanto in un certo istante e in un certo contesto. Chiaramente il problema diventa metafisico, ma non va ignorato. La sola possibilità costruttiva è di considerarlo nel senso ipotetico-normativo delle scommesse già introdotto a proposito delle scommesse coerenti.

Se si agisce sempre in modo coerente (questa volta l'aggettivo ha il significato ordinario) per tutte le valutazioni di probabilità, sia che gli eventi siano verificabili domani, fra 10 anni o mai, si può dare alle affermazioni probabilistiche lo stesso significato di grado di fiducia, senza che la non verificabilità immediata inducano a bluffare per proprio tornaconto. Ma a questo punto ricerca scientifica, questioni metafisiche e morali si fondono ... La scommessa coerente, nonostante la sua vaghezza, sembra essere l'unica possibilità per uscire dall'impasse, credo[7].

5.13 Ricapitolando

- Il teorema di Bayes gioca un ruolo cruciale in tutti i problemi di probabilità inversa, primo fra tutti il problema delle probabilità delle cause, viste come eventi condizionanti degli effetti osservati.
- Nella sua formulazione più semplice esso afferma che

  $$P(H_i \,|\, E) \quad \propto \quad P(H_i \,|\, E) \cdot P(H_i)\,,$$

  cioè

  > la *probabilità finale di* H_i (subordinata all'ipotesi che E sia vero) è proporzionale alla *verosimiglianza di* E (subordinata all'ipotesi che H sia vera) per la *probabilità iniziale di* H_i.

[7] Un punto di vista alternativo è quello di rinunciare a dare un valore di realtà alle ipotesi fisiche (e quindi anche ai "valori veri" di grandezze fisiche, per quanto detto nel capitolo 1). Questo punto di vista è espresso molto bene da de Finetti:

> *"E illusorio attribuire a una teoria o a una legge un significato apodittico ma tuttavia esiste chiaramente un significato pragmatico in quanto essa induce ad attendere che certi fatti si svolgano nel modo che noi riteniamo conforme all'idea che di tale teoria o legge ci siamo fatti. La formulazione di una teoria, di una legge, è un anello - in certa misura infido perché metafisico ma tuttavia spesso necessario come tentativo di sintesi semplificativa di cose complesse - del processo mentale per cui passiamo dall'osservazione di fatti passati alla previsione di fatti futuri. In definitiva è solo dei fatti, dei singoli fatti, che ha senso parlare. È ai fatti, che (se sono futuri, e se comunque ne ignoriamo l'esito) possiamo attribuire una probabilità".*

Detto in parole povere, si tratta di convenire se la predittività delle leggi fisiche sia da intendersi rispetto ad osservazioni future o rispetto a intangibili "valori veri". Schematizzando nei processi di misura ai quali siamo interessati, questo vorrebbe dire che non ha senso parlare della probabilità che il valore vero di una grandezza fisica sia in un certo intervallo, subordinatamente alle osservazioni sperimentali. Sarebbe rilevante parlare soltanto della probabilità che, in un esperimento futuro, l'indicazione dello strumento cada in un certo intervallo, subordinatamente alle osservazioni precedenti e alla conoscenza delle condizioni di tale esperimento. Anche se questo punto di visto sembra più realista di quello dell'estensione del concetto di evento illustrato nel testo, è innegabile il vantaggio pratico di considerare i valori di grandezze fisiche alla stessa stregua degli eventi reali. Per questo motivo la teoria dell'incertezza di misura che sarà sviluppata nella terza parte del testo si baserà su tale approccio e sull'ipotesi di esistenza di "valori veri" di grandezze fisiche.

- Quando gli eventi H_i costituiscono una classe completa di ipotesi il teorema assume la sua forma più conosciuta di

$$P(H_i\,|\,E) \;=\; \frac{P(E\,|\,H_i)\cdot P(H_i)}{\sum_j P(E\,|\,H_j)\cdot P(H_j)} \quad \text{se} \quad \left\{ \begin{array}{l} H_i \cap H_j = \emptyset \;\; (i \neq j) \\ \bigcup_{i=1}^n H_i = \Omega \end{array} \right.$$

- Il teorema di Bayes può essere usato iterativamente considerando la probabilità iniziale per la prossima inferenza uguale alla probabilità finale dell'inferenza precedente. È notevole il fatto che i risultati dipendano soltanto dalle osservazioni sperimentali (oltre che dalle probabilità a priori), ma non dall'ordine con il quale sono considerati.

- La statistica bayesiana consiste nell'uso sistematico del teorema di Bayes per aggiornare i gradi di fiducia delle diverse ipotesi.

- La dipendenza dalle probabilità iniziali è fortemente ridotta quando si hanno a disposizione molte informazioni sperimentali. In tale caso le conclusioni sono determinate dai dati sperimentali.

- Lo schema di aggiornamento del grado di fiducia mediante il meccanismo bayesiano (pregiudizio + indizi $\rightarrow$ conclusioni) è molto simile a quello utilizzato comunemente nella ricerca scientifica.

- Il metodo di falsificazione è recuperato nel caso in cui una ipotesi ha verosimiglianza nulla di produrre una certa osservazione. Le varie teorie che non sono state falsificate sono invece classificate secondo la loro credibilità (soggettiva o intersoggettiva).

- Come nel caso in cui si ha una certa quantità di osservazioni sperimentali e pregiudizi non troppo rigidi le conclusioni sono essenzialmente determinate dalle verosimiglianze, lo stesso si verifica quando c'è un giudizio di indifferenza a priori rispetto alle varie ipotesi. In questi casi le ipotesi più probabili sono quelle che massimizzano la verosimiglianza. Questo giustifica il criterio di massima verosimiglianza, se almeno una di queste condizioni è soddisfatta.

- L'uso dei concetti e dei metodi probabilistici nell'inferenza di grandezze fisiche e nel test di ipotesi scientifiche richiede di rilasciare la condizione di verificabilità degli eventi (la possibilità alternativa, forse filosoficamente più fondata ma più pesante da perseguire, sarebbe quella di interessarsi soltanto a previsioni e ad affermazioni probabilistiche di osservabili). Nel seguito assumeremo, essenzialmente per ragioni di comodo, di ipotizzare l'esistenza di "valori veri" di grandezze fisiche.

5.14 Problemi

1. Ad una persona vengono date due monete, apparentemente identiche, ma una di esse è sbilanciata in modo tale da presentare testa nel 60 % dei casi (come valutato da misure precedenti). Egli prende una delle due monete, la lancia tre volte, ottenendo sempre testa. Quanto vale la probabilità che la moneta scelta sia quella truccata?

2. Il 60% degli studenti di un liceo sono maschi e il 40% femmine. L' 80% delle ragazze fuma, mentre l'80% dei ragazzi non fuma. Quanto vale la probabilità che uno studente scelto a caso risulti maschio e fumatore? Incontrando lungo il corridoio uno studente, qual'è la probabilità che sia un fumatore? Sapendo invece che uno studente fuma, quanto vale la probabilità che sia un ragazzo? (Risolvere il problema senza considerare un ipotetico numero totale di studenti)

3. Si osserva un certo evento il quale, nell'ambito di una certa teoria assunta vera, ha probabilità dello 0.1 %. Quanto vale la probabilità che tale teoria sia vera, alla luce di tale osservazione?

4. Secondo due teorie, le quali godono dello stesso grado di credibilità, la probabilità che si verifichi un certo evento è pari a $e^{-\lambda}$ in cui il parametro λ vale 1 per la teoria Th_1 e 2 per la teoria Th_2. Come cambia la credibilità delle due teorie nell'ipotesi che si verifichi tale evento?

5. Una scatola può contenere con pari probabilità un anello d'oro o uno d'argento, indistinguibili al tatto. Si aggiunge un anello d'oro nella scatola (anch'esso indistinguibile), si agita e si estrae a caso un anello. Esso è d'oro. Quanto vale la probabilità che nella scatola ci sia rimasto ancora un anello d'oro?

6. Una ditta cerca dell'acqua nel sottosuolo mediante misure di conducibilità del terreno. L'esperienza passata ha mostrato che, nei terreni in cui non c'è acqua, la probabilità che la conducibilità superi un certo valore critico è dell'1 %, mentre è praticamente 1 nel caso di acqua. Le misure indicano un valore di conducibilità ben al di sopra di quello critico. Inoltre i contadini della zona ritengono molto probabile che in quel terreno si possa trovare l'acqua. Quantificando il "molto probabile" in 90 % valutare la probabilità che scavando ci sia effettivamente dell'acqua.

7. Una persona ritiene che la probabilità che una squadra vinca un incontro di calcio sia dell' 80 %, mentre la probabilità che essa passi in vantaggio entro il termine del primo tempo sia del 60 %. Egli crede inoltre che se la squadra termini il primo tempo in vantaggio ha una probabilità di vincere del 90 %. Di quanto aumenta la probabilità che la squadra avesse terminato il primo tempo in vantaggio, qualora egli venisse a sapere che la squadra ha vinto l'incontro?

8. Un paziente presenta un sintomo che potrebbe essere prodotto da quattro possibili malattie M_1, M_2, M_3 e M_4, fra loro incompatibili. L'incidenza di quelle malattie sulla popolazione è rispettivamente dello 0.5, 0.1, 0.01 e 0.1 % e le frequenze con le quali esse presentano quel sintomo sono il 5, il 50, il 95 e il 100 %. Da quale malattia è più probabilmente affetto il paziente?

9. Sui dati del problema precedente. Ammettiamo che sia possibile determinare in modo certo l'esatta patologia ma attraverso analisi specifiche, lunghe e molto costose. Esiste invece la possibilità di effettuare un test rapido ed economico che dà un risultato negativo se è presente M_4, mentre dà un risultato positivo per le altre malattia, ma con probabilità rispettivamente del 5, 30 e 90 %.

   Il test dà risultato positivo. In quale ordine dovranno essere prescritte le analisi costose per identificare la malattia? (Assumiamo che la pericolosità delle malattie sia simile e che nessuna di esse presenti carattere di urgenza.)

10. Un medico di famiglia riscontra in un paziente dei sintomi che possono essere causati, con probabilità 90 %, dalla malattia M_1 e, con probabilità 50 % dalla malattia M_2. Le patologie M_1 e M_2 sono abbastanza rare ed inoltre il medico sa che ci sono altre malattie che potrebbero provocare i sintomi riscontrati, ma non è sufficientemente informato su di esse. Sa bene invece che le due patologie non si presentano insieme e che M_2 ha una incidenza sulla popolazione 10 volte M_1. In base a queste informazioni quanto vale la probabilità che il paziente sia afflitto da M_1.

11. In un paese di 60 milioni di abitanti si vuole effettuare un test di AIDS sull'intera popolazione. Stime preliminari indicano che il probabile numero di persone infette dal virus HIV è di 50 mila. Supponiamo di effettuare il test mediante delle analisi di laboratorio che abbiano le seguenti caratteristiche: la probabilità che una persona infetta risulti positiva al test è del 99.9%, mentre la probabilità che una persona non infetta risulti negativa è del 99.8%. Si discutano gli effetti di tale test sull'intera popolazione. Ad esempio: quanto vale la probabilità che una persona dichiarata positiva dal test sia effettivamente infetta? Quanto vale la probabilità che una persona dichiarata negativa

sia effettivamente sana? E' possibile, con questi dati effettuare veramente un test su tutta la popolazione? Perché è preferibile analizzare i così detti "gruppi a rischio"? Quali sono i vantaggi e gli svantaggi di effettuare due test indipendenti, definendo positivi coloro che risultano positivi ad entrambi?

12. Sui dati del problema precedente: un dottore è seriamente in dubbio che un paziente possa essere infetto dal virus HIV e lo invita a sottoporsi a un test dal quale risulta positivo. Quanto vale il grado di fiducia del medico che il paziente sia infetto?

13. "Agli impiegati piace il lifting" è il titolo di un trafiletto pubblicato sul Corriere della Sera del 12/11/94: "Plastiche facciali, lifting, seni o semplici nasi nuovi. la chirurgia estetica conquista sempre nuovi clienti. …secondo un sondaggio della scuola di medicina estetica, infatti, solo '8 % degli operati appartiene al mondo dello spettacolo. Al primo posto si collocano a sorpresa, gli impiegati (18 %), seguiti da artigiani e commercianti (16 %), professionisti e casalinghe (15 %), dirigenti (12 %). Perfino gli operai (11 %) superano in numero gli artisti". Perché il giornalista si dice sorpreso dei risultati del sondaggio? (Provare, ad esempio, a rispondere alla domanda, espressa in modo canonico: "quanto vale la probabilità che una persona sia un artista, nell'ipotesi che la persona si sia sottoposta ad un intervento di chirurgia estetica?")

14. Un investigatore di provincia, autodidatta in statistica inferenziale, dovendo affrontare un caso difficile si documenta sulla probabilità che persone appartenenti a diverse categorie possano aver commesso un certo crimine. Questo gli dovrebbe servire a rivolgere le indagine cominciando da certi ambienti anziché altri. Egli divide le persone in: a) artigiani; b) frequentatori di palestre; c) divorziati; e) frequentatori di night club; f) donatori di sangue; g) insegnanti; h) titolari di agenzie assicurative; i) mariti infedeli. Egli si documenta anche sulla percentuale di questi gruppi di persone fra la popolazione. Può usare la formula di Bayes?

15. Un telespettatore, avendo risposto correttamente ad una serie di quiz, ha diritto ad un premio consistente in due anelli, da estrarre nel seguente modo: tre custodie identiche contengono una due anelli d'oro, un'altra uno d'oro e uno d'argento e la terza due d'argento; il telespettatore sceglie una custodia e la valletta estrae a caso un anello e lo mostra alle telecamere; successivamente il fortunato telespettatore ha il diritto di scegliere se far estrarre il secondo anello dalla stessa custodia o da un'altra.
Supponendo che il primo anello sia d'oro, dove conviene estrarre il secondo?

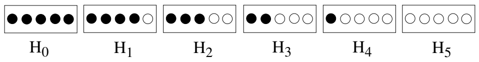

Figura 5.6: Problema delle sei scatole.

16. Per un esperimento di fisica si sta progettando un rivelatore per identificare una nuova particella che è attesa essere prodotta insieme ad un fondo 10^7 volte superiore. Assumendo che il rivelatore riesca ad identificare con la quasi certezza la particella di interesse, trovare l'errore massimo che esso può commettere nella reiezione del fondo al fine di avere un rapporto segnale/rumore di 10.

17. Riprendiamo il problema delle due scatole di cui una contiene 8 palline bianche e 2 nere e l'altra 2 bianche e 8 nere. Si estrae una pallina da una scatola scelta a caso e, senza guardarla, la si ripone nell'altra scatola. Successivamente si estrae una pallina da quest'ultima scatola. Calcolare la probabilità che la pallina sia bianca facendo i conti dettagliati. Quanto vale invece la probabilità che la seconda pallina sia bianca se <u>si è visto</u> che la prima pallina è bianca?

18. Consideriamo sei scatole le quali contengono 5 palline di cui una tutte nere (N), la seconda 4 nere e una bianca (B), e così via, fino alla sesta che contiene tutte palline bianche (vedi figura 5.6). Indichiamole con H_0, H_1, …, H_5, associando l'indice di ordine al numero di palline bianche. Si sceglie a caso una delle scatole. Poi si sceglie a caso una pallina, la si guarda e la si reintroduce nella scatola. Quindi si estrae di nuovo una pallina, e così via. Valutare la probabilità che la scatola scelta sia ciascuna delle H_i dopo ciascuna estrazione, sapendo che l'ordine di osservazione delle palline è N, B, N. Quali sarebbero le conclusioni dopo la terza estrazione se l'ordine fosse stato N, N, B? (Rispondere a questa domanda sia in modo intuitivo che facendo i conti). Annotazioni da inserire in vari punti delle dispensa

19. Seguito del problema precedente. Quanto vale la probabilità che al quarto tentativo si estrai una pallina bianca?

20. In un test a scelta multipla consistente in 10 quesiti, ciascuno con 3 possibili soluzioni, i candidati

devono rispondere esattamente a tutte le domande. Da studi preliminari è noto che le frazioni di persone in grado di rispondere correttamente ad almeno 10, 9, 8, 7, 6 e 5 domande sono rispettivamente 5, 15, 35, 75, 95 e 100 %. E' anche noto che i candidati rispondono a caso quando non sanno risolvere il quesito, nella speranza di aumentare il punteggio. Quanto vale la probabilità che se un candidato ha risposto correttamente a tutte le domande abbia tirato a indovinare almeno una risposta? E almeno due risposte?

21. Riprendiamo l'esempio del sospetto baro discusso nel paragrafo 5.7. Rendiamo il problema un po' più realistico supponendo che un baro non vinca sempre. Fissiamo ad esempio $P(V|B) = 0.70$ e supponiamo che la sequenza di Vincite e Non_vincite del sospetto baro sia: N, V, V, V, N, V, N, V, V, V. Valutare la probabilità che egli sia effettivamente un baro dopo ogni risultato partendo da una probabilità iniziale di $P_\circ(B) = 0.50$.

22. Dai dati del problema precedente: valutare direttamente la probabilità che la persona sospetta bari sapendo soltanto che ha vinto 7 volte e ha perso 3 volte.

23. Un termometro digitale avente una risoluzione di 1 grado è stato appena tolto dalla sua confezione con cui è arrivato dal fornitore[8] Dall'esperienza precedente si sa che quel tipo di termometri ha il 60 % di probabilità di indicare la temperatura giusta mentre ha una probabilità del 20 % di sbagliare di $+1\,°C$ e il 20 % di sbagliare di $-1\,°C$. La sensazione fisiologica dello sperimentatore è tale che costui creda, dalla sua esperienza passata, che la temperatura ambiente sia molto probabilmente 20 o 21 $°C$ (40 % di probabilità per ciascuno dei valori). Egli crede quattro volte di meno ai valori contigui a questi e si sente praticamente sicuro di escludere valori inferiori a 19 o superiori a 22 $°C$. Il termometro fornisce una lettura di 21 $°C$. Come vengono modificati i gradi di fiducia sulle diverse temperature?

24. Sul problema precedente: discutere come cambiano le conclusioni se il termometro segna 18 $°C$, fermo restante il giudizio a priori dello sperimentatore. Che conclusioni ne trarrebbe invece una persona meno esperta che ha idee più vaghe sul valore della temperatura ambiente?

25. Sempre sul problema del termometro: si immagini che lo sperimentatore "esperto" sappia che il 5 % quel tipo di strumenti arriva difettoso dal costruttore. In tale caso la lettura può segnare qualsiasi valore fra 0 e 99 °C con la stessa probabilità. Se a parità di condizioni fisiologiche descritte precedentemente il termometro segna 18°C, quanto vale la probabilità che il termometro sia rotto? Quanto sono credibili i vari valori di temperatura sotto queste condizioni?

26. Risolvere il problema precedente nel caso che le valutazioni a priori consentano, benché con probabilità molto piccola il valore di 18°C: $P(T = 18), \ldots P(T = 23)$ =1, 10, 39, 39, 10, 1 %.

27. Sempre sullo stesso termometro e in base alle probabilità a priori del problema precedente. Quanto vale la probabilità che il termometro sia rotto se indica 20, 17 o 5°C?

28. A proposito dei problemi precedenti: "ma non sarebbe meglio cambiare termometro?"; "bisogna fare tutti questi conti per un problema così semplice?".

29. Il gene responsabile del carattere di poter arrotolare la lingua ("R") è dominante rispetto a quello dominante di non poterla arrotolare ("r"). Chiamiamo *roller* un individuo che abbia tale capacità. Supponiamo che un bambino, orfano di madre, sia roller. Il padre, anch'egli roller, si è successivamente risposato con una donna proveniente da una famiglia nella quale non c'era nessun roller. Dal nuovo matrimonio sono nati tre figli di cui soltanto uno mostra la capacità di arrotolare la lingua. Sapendo che nella regione in cui vive questa famiglia i geni alleli responsabili del carattere sono diffusi con circa la stessa frequenza, calcolare la probabilità che la mamma del bambino fosse roller.

30. Continuazione del problema precedente: come cambia la probabilità dalla conoscenza che il nonno materno è roller?

31. Un sedicente sensitivo, davanti ad una commissione di "scienziati scettici", riesce a fare delle previsioni che la commissione ritiene abbiano una probabilità di 10^{-8} di essere indovinate per caso. Assumiamo che la commissione (scettica!) valuti in 10^{-8} la probabilità di avere di fronte un sensitivo genuino. Come varia l'opinione della commissione di fronte a tale constatazione sperimentale?

    Successivamente un prestigiatore professionista fa notare alla commissione che se il sedicente sensitivo fosse un abile imbroglione avrebbe avuto una

[8]Questo tipo di strumenti, che sembrano una vera assurdità per misure di routine, vogliono essere, "mutatis mutandis", una metafora di quanto avviene comunemente nella ricerca avanzata.

probabilità del 99 % di ingannare la commissione senza essere scoperto. Come si tiene conto di questa nuova informazione?

Il sensitivo viene di nuovo invitato ad esibirsi di fronte alla commissione costituita anche dal prestigiatore, ma egli si rifiuta. Quali conclusioni deve trarne la commissione?

32. Analizzare i problemi 11, 13 e 30 (prima domanda) assumendo che le conclusioni siano quelle che massimizzano la verosimiglianza.

Capitolo 1

1. a) vero; b) falso; c) e d) incerte, con d) molto meno probabile di c).
2. c) vera; a) e b) incerte, con b) molto più probabile che a).
3.
   - Tutti tranne E_3 e E_6;
   - i due terni hanno la stessa probabilità;
   - E_8.
4. 50 % (non servono conti: spostare palline alla cieca non cambia lo stato di conoscenza; il problema è equivalente all'estrazione da una scatola contenente 10 palline bianche e 10 nere).
5. 50 % (come il caso precedente).

Capitolo 2

1. $P(E_1) = P(E_2) = 1/4$, $P(E_3) = 1/2$ e non $P(E_1) = P(E_2) = P(E_3) = 1/3$.
2. $1/36$.
3. $2/36 = 1/18$.
4. $(1/2)^{10} = 1/1024 \approx 0.1\,\%$.
5. La probabilità di aprire le valigie al primo tentativo è la stessa in quanto bisogna indovinare tutte le cifre: $(1/10)^6$, ovvero un milionesimo. Le cose cambiano effettuando più tentativi. Nella valigia con una sola serratura si ha la certezza soltanto con un milione di tentativi, mentre per l'altra ne sono sufficienti 2000.
6. Stessa probabilità: $(1/3)^{12} = 1/531441 \approx 1.88 \times 10^{-6}$.
7. $(1/106)^{20} = 3.1 \times 10^{-41}$. Per confronto si pensi che il nostro universo esiste da circa $5 \times 10^{17}\,s$!
8. Anche se il numero di triplette è lo stesso, esse non sono tutte equiprobabili: $P(11) = 27/216$, $P(12) = 25/216$.
9. $p = (\pi a^2/4)/a^2 = \pi/4$.
10. I dati non mostrano chiari segni di aumento o diminuzione della frequenza. In mancanza di ulteriori informazioni si può ragionevolmente assegnare una probabilità costante all'evento (vedremo nel seguito dei criteri per valutare più quantitativamente questa ipotesi). Ne segue che $p \approx 20370/21896 \approx 93.0\,\%$.
11. I dati mostrano una crescita abbastanza lineare della frequenza con il passare degli anni (riportare ad esempio i dati su un foglio a quadretti). Si può quindi stimare per l'anno successivo una probabilità di circa $1.8\,\%$ (tracciare una retta ad occhio fra i punti).
12. Il fatto che la frequenza indichi il 100 % di pioggia non influenza il grado di fiducia sul bel tempo il giorno dopo, basato su fatti molto più contingenti. (Ben altra sarebbe la situazione se negli anni precedenti avesse piovuto spesso contro ogni previsione meteorologica. . . !)
13. Con questo stato di informazione: $P(1) > P(X) > P(2)$ (per ulteriori raffinamenti bisogna sapere chi gioca ed essere esperti di calcio).
14. $P(1) > P(X) > P(2)$.
15. L'informazione è irrilevante.
16. Se la Tua risposta è "$63.5\,\%$" c'è un'altissima probabilità che non Tu abbia capito niente del concetto di probabilità. Questa non è semplicemente "la frequenza", ma dipende dallo stato di informazione riguardante l'evento in questione ("hai studiato?", "ti senti sicuro?", etc.). Eventualmente il $63.5\,\%$ è la valutazione della probabilità che può dare l'esaminatore se è in mancanza di ulteriori informazioni sullo studente (ad esempio la probabilità che il 15° studente dei 30 prenotati sia promosso, a patto di non sapere nemmeno il nome dello studente).
17. Chiaramente non si può dire 100 % e 0. Immetteresti sul mercato il primo farmaco, o scoraggeresti la sperimentazione del secondo, con queste informazioni?

18. (a) Lo spazio degli eventi (indicato da qui in poi con Ω) è $\Omega = \{(G,g),(G,g),(G,g),(G,g)\}$, con tutti gli elementi equiprobabili. Si ha quindi con certezza il genotipo (G,g), da cui il carattere giallo.

    (b) Combinando (G,g) con (G,g) si ottiene
    $\Omega = \{(G,G),(G,g),(g,G),(g,g)\}$,
    da cui $P(\text{“giallo”}) = 3/4$.

    (c) Combinando (G,g) con (g,g) si ottiene
    $\Omega = \{(G,g),(G,g),(g,g),(g,g)\}.$,
    da cui $P(\text{“giallo”}) = 1/2$.

    (d) $P(\text{“verde”}) = 1/2$.

    (e) 3^a generazione: combiniamo $\{(G,G),(G,g),(g,G)\}$ con sé stesso. Per comodità riscriviamo lo spazio delle possibilità $\{(G,G), 2\cdot(G,g)\}$, in quanto (G,g) e (g,G) sono indistiguibili. Il nuovo spazio degli incroci è costituito da:
    $\{(G,G)\cdot(G,G), 4\cdot(G,G)\cdot(G,g), 4\cdot(G,g)(G,g)\}$. Si hanno quindi in totale 36 coppie di geni di cui 16 del tipo (G,G), 16 del tipo (G,g) e soltanto 4 del tipo (g,g):
    $P(\text{“Giallo omozigota”}) = 4/9$,
    $P(\text{“Giallo eterozigota”}) = 4/9$,
    $P(\text{“Verde”}) = 1/9$.
    4^a generazione: I gialli della 3^a generazione selezionati erano con pari probabilità omo- e etero-zigoti, quindi combinando $\{(GG),(Gg)\}$ con sé stessi otteniamo $\{(GG)\cdot(GG), 2\cdot(GG)\cdot(Gg),(Gg)\cdot(Gg)\}$. Delle 16 possibilità, 9 danno luogo a (GG), 6 a (Gg) e una sola a (g,g):
    $P(\text{“Giallo omozigota”}) = 9/16$,
    $P(\text{“Verde”}) = 1/16$.
    5^a generazione: Dei gialli della 4^a generazione selezionati i $9/15$ erano (G,G) e i $6/15$ erano (Gg). Si può quindi pensare di combinare 9 (G,G) con 6 (Gg). Delle 900 possibilità ce ne sono 576 (GG) e solo 36 (gg):
    $P(\text{“Giallo omozigota”}) = 16/25$,
    $P(\text{“Verde”}) = 1/25$.
    Da questi esempi si deduce una regola generale valida quando, incrociando due individui provenienti da linee pure, una con allele dominante e l'altra con allele recessivo, si lasciano incrociare nel seguito soltanto gli individui che presentano il carattere dominante:
    $P(\text{Carattere recessivo}) = 1/n^2$;
    $P(\text{Omozigote dominante}) = (n-1)^2/n^2$.
    Per avere piante che diano almeno al $99.9\,\%$ semi gialli bisogna arrivare a 32 generazioni.

19. (a) 0;

    (b) 1/2;

    (c) Si può seguire la percentuale delle caratteristiche al succedersi delle generazioni con la tabella 16.1, dove sono indicate le frequenza dei geni e le probabilità, ottenute normalizzando al numero totale di coppie di geni in ciascuna generazione:

    La frazione di roani va come $1/2^n$ in funzione del numero di generazione. Questo sistema di incroci, nel caso di non dominanza, tende a selezionare individui omozigoti.

20. $P(F) = 44.9\,\%$; $P(F\,|\,S) = 38.5\,\%$; $P(S\,|\,M) = 53.3\,\%$.

21. Sì , le probabilità sono diverse. Per la rossa, la bionda e la mora esse valgono rispettivamente 1/4, 1/3 e 1/2 (si pensi allo spazio delle possibilità del sesso dei figli, ordinati in età, e a come tale spazio viene ristretto dalle informazioni relative a ciascuna signora).

22. 12 %; 44 %; 27 %.
    In questi casi conviene fare una tabellina del tipo

gen.	Rosso	Roano	Bianco
0		Rr (1)	
1	RR (1/4)	RR $2 \cdot Rr$ rr $2 \cdot Rr$ (1/2)	rr (1/4)
2	$4 \cdot RR$ $6 \cdot RR$ (3/8)	$2 \cdot RR$ $4 \cdot Rr$ $2 \cdot rr$ $4 \cdot Rr$ (1/4)	$4 \cdot rr$ $6 \cdot rr$ (3/8)
3	$24 \cdot RR$ $28 \cdot RR$ (7/16)	$4 \cdot RR$ $8 \cdot Rr$ $4 \cdot rr$ $8 \cdot Rr$ (1/8)	$24 \cdot rr$ $28 \cdot rr$ (7/16)
	...	...	...

Tabella 16.1: Evoluzione dello spazio delle possibilità dei geni R e r (vedi testo).

	fum	$\overline{fum}$	$fum + \overline{fum}$
M	120	480	600
F	320	80	400
$M + F$	440	560	1000

23. Tirando a caso conviene puntare sul pari (52 % di probabilità di vincere).
    In realtà, l'avversario ha il 60 % di probabilità di tirare dispari. Tirando pari e scommettendo sul dispari (o viceversa) si raggiunge una probabilità di vincita del 60 %.

24. Chiaramente la scelta è indifferente.

25. L'offerta è ora conveniente (P=2/3). Si pensi all'offerta equivalente di poter scegliere fra la propria scatola e le altre due quando tutte le scatole sono ancora chiuse.

26. 2000.

27. A è favorito.
    Dalle quote di scommessa si calcolano $P(A) = 4/9$ e $P(B) = 2/9$), con coerenti in quanto $P(A) + P(B) < 1$ (l'allibratore fa il suo lavoro e non beneficenza).
    Rispettivamente 4/3 e 4 volte la puntata.

28. Gli si propone di scommettere 8000 contro 2000 (lire) su Testa. Nel secondo caso gli si propone di scommettere 7000 contro 3000 su Croce.

29. Probabilità della singola vincita: $p_\circ = 3.33 \times 10^{-8}$. $M = p_\circ \times 6\,000\,000\,000 + p_\circ \times 3\,000\,000\,000 + \cdots \quad + p_\circ \times (102 \times 250\,000\,000) + p_\circ \times (252 \times 50\,000\,000)$.

$$\frac{M}{\text{Puntata}} = \frac{\text{Monte premi}}{\text{(costo biglietto)}\cdot\text{(numero biglietti)}} = 0.363$$

30. P(Controllo in un mese) ≈ 0.25;
    Speranza matematica: $0.25 \times 60000 = 15000\, lire$, molto inferiore al costo dell'abbonamento. Il ragazzo potrebbe preferire di rischiare.

31. Il criterio è chiaramente di suddividere la posta in parti proporzionali alle rispettive probabilità di vittoria all'istante dell'interruzione. Il numero massimo di partite che sarebbero occorse per il termine del gioco è 4. Il numero delle possibili sequenze di vincita/perdita dei due giocatori nelle quattro partite è pari a 16 e, per l'ipotesi che i giocatori sono di pari abilità, sono tutte equiprobabili. Analizzando le sequenze si trova: $P(A) = 11/16$ e $P(B) = 5/16$.

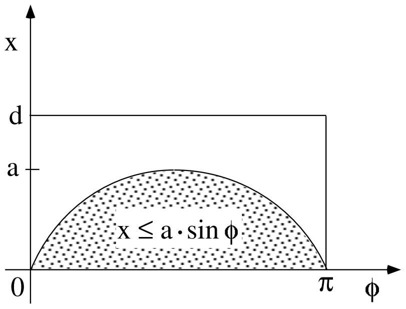

Figura 16.6: Spazio delle possibilità per l'esperimento dell'ago di Buffon.

32. Chiamando x la distanza dal centro dell'ago con la linea più vicina e ϕ l'angolo formato fra ago e linea, i possibili valori di x e di ϕ sono quelli del quadrato di dimensioni πd. La condizione di toccare è verificata se il centro dell'ago dista meno di $a \sin\phi$ dalla linea (vedi figura 16.6). Quindi $P = 2\,a/(\pi d)$.
33. La risposta dipende dall'ipotesi di come sono state preparate le scatole. Se sono state prese due palline a caso da una grande scatola che conteneva entrambi i colori in egual numero la probabilità vale 1/4. Se invece erano state preparate tre scatole di cui due con colori uguali e una con colori diversi la probabilità vale 1/3.
34. Ovviamente la risposta dipende dalla conoscenza di chi ha preparato le buste: fra compagni di classe $100'000$ potrebbero rappresentare l'importo più alto, mentre per una trasmissione televisiva potrebbe essere il contrario. Il paradosso nasce dal fatto che se si considerano i due importi *equiprobabili* il valore atteso di guadagno in caso di scambio sarebbe (chiamando v il valore dell'assegno) $G = -v + 1/2 \times (10\,v) + 1/2 \times (v/10) = 4.05v$. Ma siccome questo ragionamento può essere fatto prima ancora di aprire la busta, sembrerebbe che convenga cambiare sempre!
35. Se si è in condizione di incertezza vuol dire che la previsione di guadagno è nulla. Considerando un importo unitario sulla prima busta si ha: $-1 + p \cdot f + (1-p)/f = 0$. Ne segue che $p = 1/(1+f)$. Nel nostro caso: 9 %.
36. Si tiene le 10000 lire!
37. La previsione di guadagno in funzione di ϵ è

$$\mathrm{I\!P}(G(\epsilon)) = -p\,[1-(p+\epsilon)]^2 - (1-p)\,[p+e]^2 = \mathrm{I\!P}(G(0)) - \epsilon^2\,.$$

Conviene non bluffare ($\epsilon = 0$).

38. In questo caso la previsione di guadagno è

$$\mathrm{I\!P}(G(\epsilon)) = -p\,[1-(p+\epsilon)] - (1-p)\,[p+e] = \mathrm{I\!P}(G(0)) - \epsilon(1-2p)\,.$$

Se $p = 0.5$ la penalizzazione è indifferente da quanto si dichiara. Se è > 0.5 o < 0.5 conviene dichiarare rispettivamente di più o di meno (possibilmente 1 o 0). Quindi non è una penalizzazione atta ad addestrare le persone.

39. La probabilità ottenuta dal rapporto di aree è una estensione della "definizione" classica per infiniti eventi (i punti del piano) ciascuno di probabilità nulla. Questo argomento sarà ripreso più rigorosamente nella seconda parte del testo.

Capitolo 3

1. 300.
2. 2160 volte.
3. 3.76 % (=$14^2 \times 15^2/19^4 = 4900/19^4$).
4. $\binom{10}{6} = 210$: la schedina costa $168'000$ lire.

5. In tutti i problemi i casi possibili (assunti equiprobabili) sono dati da r disposizioni (con possibili ripetizioni) di n oggetti con rispettivamente: $n = r = 2$; $r = 2, n = 6$; $n = 2$, $r = 10$; $n = 3, r = 12$; $n = 106, r = 20$. Nei problemi 1, 4, 6 e 7 si considera un solo stato possibile. Nel 3 ce ne sono due.
6. 6˙760˙000 (in realtà alcune lettere non vengono usate nelle targhe perché si confonderebbero con i numeri).
7. Il risultato precedente va moltiplicato per il numero di combinazioni di 6 elementi presi 2 a 2 (o, con identico risultato, 4 a 4), pari a 15. Quindi si possono avere 101˙400˙000 targhe diverse.
8. Le possibili sequenze di calzini sono $30!$. Di esse $30 \times 9 \times (28!)$ hanno due calzini dello stesso colore nelle prime due posizioni:
$$P(n = 2) = \frac{30 \times 9 \times (28!)}{30!} = \frac{30 \times 9}{30 \times 29} = 0.31\,.$$
La condizione dei primi due diversi e il terzo uguale a uno dei primi due si verifica in $30 \times 20 \times 18 \times (27!)$ sequenze, mentre quella dei primi tre calzini disuguali e del quarto uguale a uno dei primi tre si verifica in $30 \times 20 \times 10 \times (27!)$ sequenze.
Se si richiede invece che una sequenza comprenda fra i primi 4 calzini almeno uno dello stesso colore del primo estratto il loro numero è: $30 \times 9 \times (28!) + 30 \times 20 \times 9 \times (27!)$ $+ 30 \times 20 \times 18 \times 9 \times (26!)$, da cui segue una probabilità del 68 %.
9. 1/18, 1/400.5, 1/11˙748, 1/511˙038 e 1/43˙949˙268. Il rapporto fra speranza matematica e puntata è sempre più sfavorevole: 0.605, …, 0.022: chissà perché!
10. È sufficiente esplicitare la formula binomiale e confrontare le espressioni.
11. Il numero di colonne possibili è $3^{13} = 1594323$, ma la probabilità <u>non</u> è $6.3\,10^{-7}$, in quanto esse non sono equiprobabili.

Capitolo 4

1. $\Omega = \{A, B, C\}$:
   (a) $E_1 = A \cap \overline{B} \cap \overline{C}$;
   (b) $E_2 = A \cap B \cap C$;
   (c) $E_3 = \overline{A} \cap \overline{B} \cap \overline{C} = \overline{A \cup B \cup C} = \emptyset$;
   (d) $E_4 = A \cup B \cup C = \Omega$;
   (e) $E_5 = \overline{A \cap B \cap C} = \overline{A} \cup \overline{B} \cup \overline{C}$;
   (f) $E_6 = A \cap B \cap (C \cup \overline{C}) = (A \cap B \cap C) \cup (A \cap B \cap \overline{C})$, ovvero $A \cap B$;
   (g) $E_7 = B$;
   (h) $E_8 = A \cap \overline{B}$;
   (i) $E_9 = (A \cap B \cap C \cup)(\overline{A} \cap \overline{B} \cap \overline{C}) = A \cap B \cap C \cup (\overline{A \cup B \cup C}) = A \cap B \cap C \cup \emptyset = A \cap B \cap C$;
   (j) $E_{10} = A \cup (B \cup C) = \Omega$;
2. I dati del problema sono: $P(A) = 1/3$, $P(B) = 1/2$ e $P(B|A) = 1/2$.
   (a) $P(\overline{A}) = 2/3$;
   $P(\overline{B}) = 1/2$;
   $P(A \cap B) = P(A) \cdot P(B|A) = 1/6$;
   $P(A \cup B) = 2/3$;
   $P(\overline{B}|A) = 1/2$;
   $P(\overline{A} \cup B) = P(\overline{A}) + P(B) - P(\overline{A} \cap B) = 5/6$;
   $P(A \cup \overline{B}) = P(A) + P(\overline{B}) - P(A \cap \overline{B}) = 2/3$;
   $P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B}) = 1/3$;
   $P(\overline{A} \cup \overline{B}) = P(\overline{A}) + P(\overline{B}) - P(\overline{A} \cap \overline{B}) = 5/6$;
   $P(\overline{A} \cap B) = P(B) \cdot P(\overline{A}|B) = P(B) \cdot (1 - P(A|B))$
   $= P(B) \cdot (1 - P(A \cap B)/P(B)) = P(B) - P(A \cap B) = 1/3$;
   $P(A \cap \overline{B}) = P(A) \cdot P(\overline{B}|A) = P(A) \cdot (1 - P(B|A)) = 1/6$.
   La seguente tabella aiuta a visualizzare le varie probabilità:

	A	$\overline{A}$	$A\cup\overline{A}$
B	1/6	1/3	1/2
$\overline{B}$	1/6	1/3	1/2
$B\cup\overline{B}$	1/3	2/3	1

(b) Provarci…

(c) $P^2(B) = 1/4$; $P(A)\cdot P(\overline{B}) = 1/6$.

3. No. Perde 1000 lire qualsiasi cosa accada.
4. No. Perde 1000 lire qualsiasi cosa accada.
5. Sì. $P(E_1 \cup E_2)$ è compresa fra 0 e 1, non inferiore a ciascuna delle $P(E_i)$ e non superiore alla loro somma.
6. Sì. $P(A\cap B) \geq 0.3$.
7. 19/40 (ci sono 3 figure per seme).
8. No. La probabilità di $A\cap B$ non può essere superiore a quella di A o di B. Quanto varrebbe $P(A\,|\,B)$ in questo caso?
9. No: $P(\text{vantaggio} \cap \text{vittoria}) > P(\text{vantaggio})$ viola le regole della probabilità.
10. $P(0) = 4/7\times 3/6\times 2/5 = 24/210 = 0.114$;
    $P(1) = P(RNN) + P(NRN) + P(NNR) = 3/7\times 4/6\times 3/5 + 4/7\times 3/6\times 3/5 + 4/7\times 3/6\times 3/5 = 108/210 = 0.514$;
    $P(2) = P(RRN) + P(RNR) + P(NRR) = 3/7\times 2/6\times 4/5 + 3/7\times 4/6\times 2/5 + 4/7\times 3/6\times 2/5 = 72/210 = 0.343$;
    $P(3) = P(RRR) = 3/7\times 2/6\times 1/5 = 6/210 = 0.029$;
    Naturalmente deve valere $P(0) + P(1) + P(2) + P(3) = 1$.
11. $P(0) = P(NNN) = (4/7)^3$;
    $P(1) = P(RNN) + P(NRN) + P(NNR) = 3\times(4/7)^2\times 3/7$;
    $P(2) = P(RRN) + P(RNR) + P(NRR) = 3\times 4/7\times(3/7)^2$;
    $P(3) = P(RRR) = (3/7)^3$;
    Cosa ricordano le quattro espressioni trovate?
12. $P(B_1) = 1/3$; $P(R_2) = 1/4$; $P(A) = 1/12$; $P(E_1) = P(F_1) = 2/3$; $P(E_2) = P(F_2) = P(H) = 1/2$.
    $P(E_1\cap F_1) \neq P(E_1)\cdot P(F_1)$: dipendenti;
    $P(E_2\cap F_2) = P(E_2)\cdot P(F_2)$: indipendenti;
    $P(E_2\cap H) = P(E_2)\cdot P(H)$: indipendenti;
    $P(E_2\cap F_2\cap H) \neq P(E_2)\cdot P(F_2)\cdot P(H)$: dipendenti.
13. $P(\text{sopravvive a } n \text{ prove}) = (5/6)^n$;
14. Ogni giocatore ha la stessa probabilità di morire:
    $P(G_1) = 1/6$;
    $P(G_2)(= P(G_2\cap\overline{G_1}) = P(G_2|\overline{G_1})\cdot P(\overline{G_1})) = 5/6\times 1/5 = 1/6; \ldots$;
    $P(G_6) = 5/6\times 4/5\times 3/4\times 2/3\times 1/2\times 1 = 1/6$.
    Il modo più semplice per capirlo è di pensare alla posizione dove si trova il proiettile dopo aver fatto ruotare il tamburo. Se si trova nella prima muore il primo, se nella seconda muore il secondo, e così via.
15. $P(6)$ su 4 lanci vale $51.8\,\%$, mentre $P(12)$ su 24 doppi lanci vale $49.1\,\%$ (conviene ragionare sugli eventi complementari).
16. 25 lanci.
17. Ovviamente $1/6$. Volendo fare i conti:
    $$\begin{aligned} P(E_{25}\,|\,\overline{E_1\cap\cdots\cap E_{24}}) &= \frac{P(E_{25}\cap\overline{E_1\cap\cdots\cap E_{24}})}{P\overline{E_1\cap\cdots\cap E_{24}})} \\ &= \frac{\frac{1}{6}\left(\frac{5}{6}\right)^{24}}{\left(\frac{5}{6}\right)^{24}} = \frac{1}{6}\,. \end{aligned}$$
18. È conseguenza del fatto che E_i implichi $E_1\cup E_2\cup E_3$ e della mutua incompatibilità degli eventi. Questa formula rappresenta una estensione alla valutazione combinatoria della probabilità nei casi di non equiprobabilità degli eventi.

19. $P(A) = 1/18 = 5.56\,\%$; $P(B) = 17/18^2 = 5.25\,\%$; $P(C) = 17^2/18^3 = 4.96\,\%$. Le quote di B e di C devono essere rispettivamente $17/18 \times 10\,000 = 9\,444$ e $(17/18)^2 \times 10\,000 = 8\,920$ lire.
20. Le probabilità vanno da 2.3×10^{-9} a 2.0×10^{-5}, da confrontarsi con $(1/3)^{13} = 6.3 \times 10^{-7}$. La conoscenza del gioco altera i rapporti di probabilità - pari a 1 in caso di assoluta ignoranza - fino a un fattore $10\,000$. La colonna con tutti 1 ha probabilità 3.4×10^{-6}, l'altra 1.8×10^{-7}
21. Il rapporto delle puntate deve essere 9.7 a 1 (paga di più quello che gioca su tutti 1).
22. Chiamiamo P_2, P_3 e P_4 le probabilità che ci siano due calzini dello stesso colore prendendone rispettivamente 2, 3 e 4. Chiamiamo $P(n=2)$, $P(n=3)$ e $P(n=4)$ la probabilità che l'evento "due calzini uguali" si verifichi <u>esattamente</u> con l'estrazione del secondo, terzo e quarto calzino.

$$\begin{aligned}
P(n=2) &= P(\text{“colore qualsiasi”}) \cdot P(\text{“colore precedente”}) = 1 \times 9/29 \\
P(n=3) &= P(\text{“colore qualsiasi”}) \cdot P(\text{“}\overline{\text{colore precedente}}\text{”}) \\
&\qquad \cdot P(\text{”uno dei colori precedenti”}) = 1 \times 20/29 \times 18/28 \\
P(n=4) &= P(\text{“colore qualsiasi”}) \cdot P(\text{“}\overline{\text{colore precedente}}\text{”}) \\
&\quad \cdot P(\text{“}\overline{\text{uno dei colori precedenti}}\text{”}) \cdot P(\text{“colore qualsiasi”}) \\
&= 1 \times 20/29 \times 10/28 \times 1
\end{aligned}$$

$P_2 = 31\,\%$, $P_3 = 75\,\%$ e $P_4 = 100\,\%$. Il risultato di P_4 è facilmente intuibile in quanto su quattro calzini di tre colori diversi ce ne devono essere necessariamente almeno due dello stesso colore.

Probabilità che su 4 estrazioni, non capiti un calzino dello stesso colore del primo: $20/29 \times 19/28 \times 18/27 = 0.312$. Quindi la probabilità complementare è del $68.8\,\%$. In questo caso si raggiunge la certezza soltanto alla 22^a estrazione (quando tutti gli altri calzini sono estratti).

23. Le probabilità che un giocatore e due giocatore ricevano tutti "lisci" è:
P_1("zero pezzi") $= 28/40 \times 27/39 \times \cdots \times 19/31 = 1.5\,\%$,
P_2("zero pezzi") $= 28/40 \times 27/39 \times \cdots \times 9/21 = 2.2 \times 10^{-5}$,
P_3("zero pezzi") $= 0$.
24. Si chiami O la scatola che ha 8 palline bianche e 2 nere, D l'altra scatola, B_1 e N_1 l'evento "pallina bianca" e "pallina nera" nella prima estrazione e B_2 l'evento "pallina bianca" nella seconda estrazione:

$$\begin{aligned}
P(B) &= P(B_2 \cap (B_1 \cap O)) + P(B_2 \cap (N_1 \cap O)) \\
&= +P(B_2 \cap (B_1 \cap D)) + P(B_2 \cap (N_1 \cap D)) \\
&= P(O) \cdot P(B_1|O) \cdot P(B_2|B_1 \cap O) + P(O) \cdot P(N_1|O) \cdot P(B_2|N_1 \cap O) \\
&\quad +P(D) \cdot P(B_1|D) \cdot P(B_2|B_1 \cap D) + P(D) \cdot P(N_1|D) \cdot P(B_2|N_1 \cap D) \\
&= \frac{1}{2} \times \frac{8}{10} \times \frac{3}{11} + \frac{1}{2} \times \frac{2}{10} \times \frac{2}{11} + \frac{1}{2} \times \frac{2}{10} \times \frac{9}{11} + \frac{1}{2} \times \frac{8}{10} \times \frac{8}{11} = \frac{1}{2}\,.
\end{aligned}$$

25. $P(\text{“nero”}) = 3/4$; $P(\text{“rosso”}) = 1/4$. Essendo gli eventi indipendenti $P(\text{“5 neri”}) = (3/4)^5 = 23.7\,\%$.
26. Innanzitutto nessuna persona che abbia un'idea delle condizioni ambientali per lo sviluppo di forme evolute di vita animale scommetterebbe alla pari su tale possibilità. In secondo luogo, non si può applicare la regola di moltiplicazione delle probabilità in quanto gli eventi sono tutt'altro che indipendenti. Per finire, ogni volta che il conto di probabilità dà un numero che non corrisponde al proprio grado di fiducia (e che viene quindi percepito come paradosso) si sta mettendo in forte dubbio che il conto sia sbagliato.
27. $65\,\%$.
28. 31/288 = $10.8\,\%$.
29. 27.5

30. La probabilità che il premio sia nell'altra scatola è pari a

$$\begin{aligned} P(\text{premio altra scatola}) &= P(\text{premio altra scatola} \,|\, \text{bluff}) \cdot P(\text{bluff}) \\ &\quad + P(\text{premio altra scatola} \,|\, \overline{\text{bluff}}) \cdot P(\overline{\text{bluff}}) \ : \end{aligned}$$

conviene sempre cambiare. Se $P(\text{bluff}) = 1/2$ la probabilità vale $7/12 = 58\,\%$.

31. Chiamando E = "25 al secondo estratto" e H = "25 al primo estratto", si ha: $P(E\,|\,H) = 0$, $P(E\,|\,\overline{H}) = 1/89$, $P(H) = 1/90$ e $P(\overline{H}) = 89/90$. Ne segue $P(E) = 1/90$, come si poteva pensare intuitivamente.

32. Chiamando E = "16 al terzo estratto" e H = "16 al primo estratto" (sottoindendendo $H_\circ$ = "47 al secondo estratto"): si ha: $P(E\,|\,H) = 0$, $P(E\,|\,\overline{H}) = 1/88$, $P(H) = 1/89$ e $P(\overline{H}) = 88/89$. Ne segue $P(E) = 1/89$.

    Se si sapesse anche che il primo estratto era 58 $P(E) = 1/88$ (anche se il risultato è intuitivo, si raccomanda, per esercizio, formalizzare la risposta anche in questo caso mediante l'uso della formula di disintegrazione).

33. Nel primo caso (scatola trovata casualmente vuota) si ha, intendendo con *"terza"* "il premio è nella terza scatola" e *"scelta"* "il premio è nella scatola scelta dal concorrente":

$$\begin{aligned} P(terza) &= P(terza\,|\,scelta) \cdot P(scelta) + P(terza\,|\,\overline{scelta}) \cdot P(\overline{scelta}) \\ &= 0 \times \frac{1}{3} + \frac{1}{2} \times \frac{2}{3} = \frac{1}{3}\,, \end{aligned}$$

da cui

$$P(terza\,|\,scelta \cup terza) = \frac{P(terza \cup (scelta \cup terza))}{P(scelta \cup\ terza)} = \frac{1/3}{2/3} = \frac{1}{2}\,.$$

Nel secondo caso (intenzionalmente viene aperta una scatola vuota):

$$\begin{aligned} P(terza|\,|\,H_{c.v.}) &= P(terza\,|\,scelta, H_{c.v.}) \cdot P(scelta\,|\,H_{c.v.}) \\ &\quad + P(terza\,|\,\overline{scelta}, H_{c.v.}) \cdot P(\overline{scelta}, |\,H_{c.v.}) \\ &= 0 \times \frac{1}{3} + 1 \times \frac{2}{3} = \frac{2}{3}\,, \end{aligned}$$

dove $H_{c.v.}$ sta per l'ipotesi "il conduttore dice la verità affermando di sapere dove sta il premio" (chiaramente $P(\overline{scelta}\,|\,H_{c.v.}) = P(\overline{scelta})$).

Capitolo 5

1. 63.3 %.
2. $P(M) = 0.6$; $P(F) = 0.4$); $P(\text{fum}\,|\,M) = 0.2$; $P(\text{fum}|F) = 0.8$.
   $P(\text{fum}) = P(\text{fum} \cap M) + P(\text{fum} \cap F)$.
   $P(\text{fum} \cap M) = P(M) \cdot P(\text{fum}\,|\,M) = 0.12$;
   $P(\text{fum} \cap F) = 0.32$;
   ne segue $P(\text{fum}\,|\,M \cup F) = 0.44$.
   Per la seconda domanda si utilizza la formula di Bayes:

$$P(M\,|\,\text{fum}) = \frac{P(\text{fum}\,|\,M) \cdot P(M)}{P(\text{fum}\,|\,M) \cdot P(M) + P(\text{fum}\,|\,F) \cdot P(F)} = 0.27\,.$$

   È istruttivo riscrivere la formula, utilizzando le probabilità composte, come

$$P(M\,|\,\text{fum}) = \frac{P(\text{fum} \cap M)}{P(\text{fum} \cap M) + P(\text{fum} \cap F)} = \frac{P(\text{fum} \cap M)}{P(\text{fum})}\,,$$

   che corrisponde al modo di ragionare più immediato (I passaggi della formula seguono a ritroso la dimostrazione del teorema di Bayes).

3. La domanda non ha senso. Dipende dall'esistenza e dalla credibilità di altre teorie che dovrebbero spiegare l'evento.
4. $P(\text{Th}_1)/P(\text{Th}_2) = 2.7$, $P(\text{Th}_1\,|\,\text{Th}_1 \cup \text{Th}_2) = 73\,\%$.

5. 2/3. Infatti indicando con H_1 "la scatola che contiene all'origine un anello d'oro", con H_2 "... un anello d'argento" e con O "si estrae un anello d'oro", si tratta di valutare $P(H_1 \,|\, O)$.

6. $$\frac{P(H_2O \,|\, \text{conduc.} \cap \text{contad.})}{P(\overline{H_2O} \,|\, \text{conduc.} \cap \text{contad.})} = \frac{1}{0.01}\,\frac{90}{10} = \frac{900}{1}\,,$$
   ovvero $P(H_2O \,|\, \text{conduc.} \cap \text{contad.}) = 99.9\,\%$. Si presti attenzione al fatto che questo risultato deriva dall'aver assunto l'indipendenza delle due informazioni, ma in realtà il contadino potrebbe aver utilizzato inconsciamente la conducibilità del terreno, la quale è legata alla sua umidità e quindi alla vegetazione locale.

7. Di un fattore 1.13 ($= P(\text{Vitt} \,|\, \text{Vant})/P(\text{Vitt})$), ossia $P(\text{Vant} \,|\, \text{Vitt}) = 67.5\,\%$.

8. $P(M_i|\text{Incidenza, sintomi}) =$ 13.5, 27.1, 5.1, 54.2

9. Si comincerà dalle analisi per accertare M_2. Infatti $P(M_i|\text{Incidenza, sintomi, test})$ vale, rispettivamente, 5.0, 60.7, 34.3, 0 %

10. Impossibile valutare $P(M_1 \,|\, S)$ in quanto M_1 e M_2 non sono esaustivi ($M_1 \cup M_2 \neq \Omega$). Al più si può dire che $P(M_1 \,|\, S)/P(M_2 \,|\, S) = 0.18$, ovvero $P(M_1 \,|\, S \cap (M_1 \cup M_2)) = 15.3\,\%$.

11. Con i simboli I, S, A_p e $\overline{A}_p$ per indicare *Infetto, Sano, Positivo* e *Negativo*, i dati del problema sono $P_\circ(I) = 8.3 \times 10^{-4}$, $P(A_p|I) = 0.999$, $P(\overline{A}_p|S) = 0.998$, ovvero anche $P(\overline{A}_p|I) = 0.001$ e $P(A_p|S) = 0.002$.
    Ne segue: $P(I|A_p) = 0.294$, $P(S|A_p) = 0.706$, $P(I|\overline{A}_p) = 8.3 \times 10^{-7}$ e $P(S|\overline{A}_p) = 0.99999917$. La maggior parte dei "positivi" sono sani.
    Come mai? Analizziamo il rapporto segnale/rumore:
    $$\frac{s}{n} = \frac{P(I|A_p)}{P(S|A_p)} = \frac{P_\circ(I)}{P_\circ(S)} \cdot \frac{P(A_p|I)}{P(A_p|S)}\,.$$
    Per ottenere $s/n \gg 1$ (invece di 0.41), essendo $P_\circ(I)/P_\circ(S) = 8 \times 10^{-4}$, dobbiamo avere $P(A_p|I)/P(A_p|S) \gg 10^3$. Ma $P(A_p|I)$ è già prossimo ad uno e quindi si può solo intervenire su $P(A_p|S)$ e farlo diventare $\ll 10^{-3}$.
    Chiamando A'_p l'esito positivo a <u>due</u> analisi indipendenti: $P(A'_p|I) = 0.998$ e $P(A'_p|S) = 4 \times 10^{-6}$. Ne segue $P(I|A'_p) = 0.995$ e $P(S|A'_p) = 0.005$: un po' meglio.
    Prendere un campione "a rischio" significa invece aumentare in s/n il fattore $P_\circ(I)/P_\circ(S)$.

12. Quantificando il "serio dubbio" in $P_\circ(I) = 1/2$ si ha che $P(I|A_p) = 99.8\,\%$.

13. Il giornalista è stato tratto in inganno dal fatto che $P(\text{chirurgia plastica} \,|\, \text{artista}) \gg P(\text{chirurgia plastica} \,|\, \text{impiegato})$, senza considerare che $P(\text{impiegato}) \gg P(\text{artista})$!

14. No. I vari gruppi non formano una classe completa di ipotesi. Inoltre, siccome non sono mutualmente esclusivi non è nemmeno possibile valutare probabilità relative.

15. Siano A, B e C i cofanetti che contengono rispettivamente i 2, 1 e 0 anelli d'oro. Indicando con O l'anello d'oro, con O_1 l'evento "il primo anello è d'oro" e con O_2 l'evento "il secondo anello è d'oro":
    $P_\circ(A) = P_\circ(B) = P_\circ(C) = 1/3$;
    $P(O_1|A) = 1$, $P(O_1|B) = 1/2$ e $P(O_1|C) = 0$;
    $P(A|O_1) = 2/3$, $P(B|O_1) = 1/3$ e $P(C|O_1) = 0$;
    $$\begin{aligned}
    P(O_2 \cap \text{“Stessa”} \,|\, O_1) &= P(A \,|\, O_1) \cdot P(O_2 \,|\, A) + P(B \,|\, O_1) \cdot P(O_2 \,|\, B) + P(C \,|\, O_1) \cdot P(O_2 \,|\, C) \\
    &= \frac{2}{3} \times 1 + \frac{1}{3} \times 0 + 0 \times 0 = \frac{2}{3} \\
    P(O_2 \cap \text{“Diversa”} \,|\, O_1) &= P(A \,|\, O_1) \cdot P(O_2 \,|\, B \cup C) + P(B \,|\, O_1) \cdot P(O_2 \,|\, A \cup C) \\
    &\quad + P(C \,|\, O_1) \cdot P(O_2 \,|\, A \cup B) \\
    &= \frac{2}{3} \times \frac{1}{4} + \frac{1}{3} \times \frac{1}{2} + 0 \times \frac{3}{4} = \frac{1}{3}\,.
    \end{aligned}$$

16. $P(\text{“falso segnale”} \,|\, \text{rumore}) < 10^{-8}$.

17. Si chiami O la scatola che ha 8 palline bianche e 2 nere, D l'altra scatola, B_1 e N_1 l'evento "pallina bianca" e "pallina nera" nella prima estrazione e B_2 l'evento "pallina bianca" nella seconda estrazione: Se invece la prima pallina estratta era bianca:

$$\begin{aligned}
P(B|B_1) &= P(O|B_1)\cdot P(B_2|B_1\cap O) + P(D|B_1)\cdot P(B_2|B_1\cap D)\,;\\
P(O|B_1) &= \frac{P(B_1|O)\cdot P(O)}{P(B_1|O)\cdot P(O)+P(B_1|D)\cdot P(D)}\\
P(D|B_1) &= \frac{P(B_1|D)\cdot P(D)}{P(B_1|O)\cdot P(O)+P(B_1|D)\cdot P(D)}\\
P(B|B_1) &= \frac{P(B_1|O)\cdot P(O)}{P(B_1|O)\cdot P(O)+P(B_1|D)\cdot P(D)}\cdot P(B_2|B_1\cap O)\\
&\quad +\frac{P(B_1|D)\cdot P(D)}{P(B_1|O)\cdot P(O)+P(B_1|D)\cdot P(D)}\cdot P(B_2|B_1\cap D)\\
&= \frac{21}{55}\,.
\end{aligned}$$

Come ci si aspetta intuitivamente, è più alta la probabilità di ottenere una pallina nera, in quanto l'operazione più probabile è stata quella di estrarre una pallina bianca da dove ce ne sono 8 e metterla dove ce ne sono solo 2.

18. Inizialmente $P(H_i) = 1/6 = 0.167$, con $i = 0, 1, \ldots 5$.
Succesivamente, indicando con

- E_k la k-ma estrazione, i cui esiti possono essere i Colori Bianco o Nero;
- $I_l = E_1 \cap E_2 \cap \cdots \cap E_l$ lo stato di informazione dopo la l-ma estrazione,

$$P(H_i|I_n) \quad\propto\quad P(E_n|H_i\cap I_{n-1})\cdot P(H_i|I_{n-1})\,,$$

con

$$\begin{aligned}
P(H_i|I_\circ) &= \frac{1}{6}\\
P(C=B|H_i\cap I_l) &= \frac{i}{5}\\
P(C=N|H_i\cap I_l) &= \frac{5-i}{5}\,.
\end{aligned}$$

Dopo $E_1 = N$: $P(H_i\,|\,I_1)$ = 1/3=0.333; 4/15=0.267; 3/15=0.200; 2/15=0.133; 1/15=0.067; 0.
Dopo $E_2 = B$: $P(H_i\,|\,I_2) = 0$; 2/10; 3/10; 3/10; 2/10; 0.
Dopo $E_3 = N$: $P(H_i\,|\,I_3) = 0$; 0.32; 0.36; 0.24; 0.08; 0.

19. $P(B\,|\,I_3) = \sum_i P(B|H_i)\cdot P(H_i\,|\,I_3) = 41.6\,\%$.

20. Applicando il teorema di Bayes:

$$\begin{aligned}
P(V=10-n\,|\,R=10) &\propto P(R=10|V=10-n)\cdot P_\circ(V=10-n)\\
&\propto \left(\frac{1}{3}\right)^n P_\circ(V=10-n)\,,
\end{aligned}$$

dove V sta per "risposte note veramente", R per "risposte giuste nel test" e n "risposte tirate a indovinare".
Probabilità iniziali: $P_\circ(10) = 0.05$, $P_\circ(9) = 0.10$, $P_\circ(8) = 0.20$, $P_\circ(7) = 0.40$, $P_\circ(6) = 0.20$, $P_\circ(5) = 0.05$. Segue:

$$\begin{aligned}
P(V\le 9\,|\,R=10) &= 1-\frac{1\times 0.05}{1\times 0.05+\frac{1}{3}\times 0.10+\frac{1}{9}\times 0.20+\cdots} = 0.59\\
P(V\le 8\,|\,R=10) &= 1-\frac{1\times 0.05+\frac{1}{3}\times 0.10}{1\times 0.05+\frac{1}{3}\times 0.10+\frac{1}{9}\times 0.20+\cdots} = 0.32
\end{aligned}$$

21. Dopo l'n-mo gioco abbiamo che $P(B\,|\,G_n)$ vale: 0.37, 0.46, 0.54, 0.62, 0.50, 0.58, 0.45, 0.54, 0.62, 0.69.

22.
$$\frac{P(B\,|\,7V,3P)}{P(\overline{B}\,|\,7V,3P)} = \frac{(0.7)^7 \times (0.3)^3}{(0.5)^7 \times (0.5)^3} \times \frac{50}{50} = 2.28\,,$$
da cui segue lo stesso risultato del 69 %.

23. Sono diverse da zero le probabilità a posteriori di 20, 21 e 22 $^\circ C$, rispettivamente 23.5, 70.6 e 5.9 %.

24. $P(T_v = 19) = 1$. L'inesperto direbbe invece che le temperature possibili sono 17, 18 e 19 $^\circ C$, con probababilità del 20, 60 e 20 %.

25. La probabilità che il termometro sia rotto è del 2.6 %. Il risultato può essere giustificato intuitivamente pensando che la probabilità iniziale è bassa e inoltre il valore letto è ancora ragionevole. Più esattamente, indicando con R, G, L e T rispettivamente "rotto", "giusto", "lettura" e "temperatura" (vera):
$$P(R\,|\,L=18) = \frac{P(L=18\,|\,R)\cdot P(R)}{P(L=18\,|\,R)\cdot P(R) + P(L=18\,|\,G)\cdot P(G)}\,.$$
Poiché l'evento "la temperatura assume un valore qualsiasi" è un evento certo (Ω), possiamo scrivere:
$$\begin{aligned} P(L=18\,|\,G) &= P(L=18 \cap \text{“}T\text{qualsiasi}'' \,|\, G) \\ &= P(L=18 \cap (\ldots \cup T=19 \cup T=20, \cup \ldots)\,|\,G) \\ &= \ldots + P(L=18 \cap (T=19)\,|\,G) + P(L=18 \cap (T=20)\,|\,G) + \ldots \\ &= P(L=18\,|\,(T=19), G)\cdot P(T=19)\,, \end{aligned}$$
in quanto tutti gli altri termini sono nulli, o perché nulla la probabilità a priori della temperatura o perché nulla la probabilità condizionata di osservare una lettura di 18 da temperature superiori a 19 °C.
In conclusioni:
$$P(R\,|\,L=18) = \frac{0.01 \times 0.05}{0.01 \times 0.05 + 0.2 \times 0.1 \times 0.95} = 2.6\,\%\,.$$
È interessante notare come il dubbio sul funzionamento del termometro cambi di nuovo i gradi di fiducia delle diverse temperature. Infatti nell'ipotesi che esso sia rotto la sua lettura non riaggiorna la probabilità iniziale. Riassumendo i risultati raggiunti nella seguente tabella abbiamo (con le probabilità in %):

condizione	T			
	19	20	21	22
iniziale	10	40	40	10
$L = 18 \cap G$	100	0	0	0
$L =$ qualsiasi $\cap\, R$	10	40	40	10
$L = 18 \cap (G \cup R)$	97.7	1.0	1.0	0.3

26. La probabilità che il termometro sia rotto scende ora al 2.0 %. La tabella precedente si trasforma in:

condizione	T					
	18	19	20	21	22	23
iniziale	1	10	39	39	10	1
$L = 18 \cap G$	13.0	87.0	0	0	0	0
$L =$ qualsiasi $\cap\, R$	1	10	39	39	10	1
$L = 18 \cap (G \cup R)$	12.8	85.5	0.8	0.8	0.2	0.02

27. 0.16 %, 1, 1. Se la lettura capita al centro dell'intervallo che ci si attendeva si tende a credere anche che il termometro funzioni bene. Diverso il caso quando la lettura cade al margine di tale intervallo. Se poi essa si discosta molto si è praticamente sicuri del malfunzionamento dello strumento. Nella schematizzazione rozza del problema la probabilità non cambia se la lettura è 17 o 5°C. In realtà la mente non opera mai così

schematicamente e magari si è ancora disposti a sbagliarsi ("forse ho caldo perché sono agitato"). Nel caso di lettura pari a 5 °C il termometro viene dichiarato "oggettivamente" rotto, in quanto tutti, freddolosi e calorosi, esperti o principianti, converranno che "non può essere".

28. Obiezioni del tutto lecite. Il fatto è che nella ricerca di frontiera bisogna basare le conclusioni sui rivelatori che si hanno e che già rappresentano quanto di più sofisticato si possa costruire. Vedremo che le cose si semplificano enormemente nei "casi tranquilli" di laboratorio. In ogni caso, questo modo di procedere è quello degli sperimentali esperti, i quali, in base alle loro aspettative e alla conoscenze degli strumenti, "riprovano", "ricalibrano", "cambiano strumento", etc.

29. Il padre è eterozigote (Rr) in quanto ha avuto almeno un figlio che non arrotola la lingua (rr) da una donna (rr).
La madre del bambino può essere (RR) (Rr) o (rr). Poiché in quella regione i due alleli R e r sono equiprobabili, le probabilità iniziali per la madre sono $P_\circ(RR) = P_\circ(rr) = \frac{1}{2}P_\circ(Rr) = 1/4$ (consideriamo indistinguibili (Rr) e (rR)). Applicando le leggi di Mendel otteniamo le probabilità che il bambino sia *roller* condizionate dal genotipo della madre, e da queste le probabilità finali dei genotipi della madre:

$$\begin{aligned}
P(\text{roller}\,|\,RR) &= 1;\\
P(\text{roller}\,|\,Rr) &= 3/4;\\
P(\text{roller}\,|\,rr) &= 1/2;\\
P(RR\,|\,\text{roller}) &= \frac{P(\text{roller}\,|\,RR)\cdot P_\circ(RR)}{P(\text{roller}\,|\,RR)\cdot P_\circ(RR)+\ldots}\\
&= \frac{1\times\frac{1}{4}}{1\times\frac{1}{4}+\frac{3}{4}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{4}} = \frac{1}{3}\\
P(Rr\,|\,\text{roller}) &= \frac{1}{2}\\
P(rr\,|\,\text{roller}) &= \frac{1}{6}
\end{aligned}$$

Ne segue una probabilità di $5/6$ che la mamma fosse stata roller.

30. I dati del problema indicano che il nonno può avere i geni (Rr) o (RR) con probabilità che stanno fra di loro in proporzione 2:1. La nonna può essere invece (RR) (Rr) o (rr), con proporzioni 1:2:1.
Delle 48 possibilità con cui la madre del bambino può ereditare i geni dei genitori ce ne sono 16 che danno luogo a (RR), 8 a (rr) e 24 a (Rr).
Le probabilità iniziali dei geni della madre del bambino diventano ora: $P_\circ(RR) = 1/3$, $P_\circ(Rr) = 1/2$ e $P_\circ(rr) = 1/6$.
Seguono: $P(RR|roller) = 0.42$, $P(Rr|roller) = 0.47$ e $P(rr|roller) = 0.11$.
Come ci si poteva attendere, sapere che la madre del bambino avesse un genitore roller aumenta le probabilità che anche lei lo fosse (dall'83 % all'89 %). Come cambierebbe ulteriormente la probabilità, se in quella regione la frequenza degli alleli di tipo R fosse il doppio di quelli del tipo r?

31. Dopo il primo esperimento la probabilità sale al 50 % (lo scetticismo della commissione comincia a vacillare). L'osservazione del prestigiatore riporta essenzialmente la probabilità al livello iniziale (1.01×10^{-8}, ad essere precisi…).
Il rifiuto a presentarsi manda definivamente a zero la probabilità, se la commissione assume, ragionevolmente, che P("non si presenta" | "imbroglione") $= 1$, mentre P("non si presenta" | "onesto") $= 0$.

32. Ne risulterebbe che: una persona dichiarata positiva sia praticamente infetta; una persona operata di chirurgia plastica sia un artista; chi riesce a ingannare una commissione di scienziati sia un sensitivo.

Capitolo 6

1. La variabile casuale può assumere con uguale probabilità i valori x_i = -1, -2, -3, 2, 6, 10: $f(x_i) = 1/6$.