Capitolo 4

VALUTAZIONE IN AMBIENTE FINITO

In questo capitolo vengono introdotte nozioni di base del calcolo delle probabilità tradizionale in *ambiente finito* (descritto con partizioni finite), inquadrando il discorso nell'ottica soggettiva per indagare sulle proprietà che è ragionevole dare a un grado di fiducia. Chi conosce dette nozioni può limitare la lettura a una rapida scorsa, trascurando esempi e dimostrazioni. Nel § 4.1 vengono presentati i *modelli simmetrici* in cui la probabilità degli eventi è il rapporto tra numero dei casi (costituenti) favorevoli e quello dei casi possibili. Si attribuisce così probabilità uguali ai casi possibili e la somma delle probabilità' dei casi favorevoli agli altri eventi. Ciò suggerisce di interpretare la probabilità come *massa unitaria* (§ 4.2) e di generalizzare il procedimento prevedendo che si possa distribuirla sui costituenti in modo qualsiasi. La probabilità, come le masse, è allora *non negativa, additiva* e qui a somma 1 (*normalizzata*) sui costituenti (*Definizione* 4.2.1). Si vedrà che queste proprietà caratterizzano le *probabilità coerenti* in ambiente finito (10.5.1 *Proposizione*). Nel § 4.3 viene introdotta la nozione di *probabilità qualitativa* e nel § 4.4 quella di *funzione peso*. Si tratta di valutazioni basate sul confronto delle attendibilità degli eventi a coppie, accontentandosi di indicare un ordine nel primo caso, esprimendolo anche quantitativamente mediante *quozienti di probabilità* (n° 4.4.1) nel secondo.

4.1 Modelli simmetrici

Il calcolo delle probabilità ha avuto inizio con lo studio di problemi che riguardano i giochi d'azzardo. Nei problemi di giochi lo schema è sostanzialmente quello che abbiamo visto nell'esempio del lancio del dado (2.1.1 *Esempio* 1)[20].

---

20 A problemi di giochi avevano rivolto la loro attenzione già gli algebristi italiani nel '500. Quello che può essere considerato il primo trattato di calcolo delle probabilità, l'opera "Ars conjectandi" di G. Bernoulli (1713), è però di epoca abbastanza posteriore.

Per descrivere la situazione si introduce un insieme $\{\omega_1, \ldots, \omega_n\}$ di *casi possibili* – una partizione finita in eventi elementari (3.4.1 *Nota e Terminologia*) – che vengono giudicati ugualmente probabili per ragioni di simmetria. Come segnalato in 3.5.1 *Commento,* il quadro delle possibilità è costituito dai casi possibili e dai loro sottoinsiemi. Gli elementi di tali sottoinsiemi sono detti **casi favorevoli** al loro corrispondente evento. In questi schemi, si assegna come *probabilità* di un evento il *rapporto* tra il numero dei suoi *casi favorevoli* e quello dei *casi possibili.* Ad ogni caso possibile ω_i viene allora attribuita probabilità $1/n$ – un caso è favorevole sugli n possibili – e agli insiemi di casi possibili (eventi) probabilità proporzionale alla loro cardinalità – con costante di proporzionalità $1/n$ –. In particolare, all'evento certo viene data così probabilità 1 e a quello impossibile probabilità 0 (perché nessun caso è a lui favorevole).

Andiamo ad illustrare il procedimento con alcuni esempi.

4.1.1 Esempi.

I primi cinque esempi si riferiscono a una ben determinata estrazione del lotto, che consiste nell'estrazione di 5 numeri, uno per volta, da ciascuna di 10 urne (**ruote**) contenenti i primi 90 numeri naturali (lotto italiano). Il risultato dell'estrazione su una ruota è quindi una *cinquina ordinata* di numeri naturali, senza ripetizioni. L'estrazione completa può allora essere riassunta in una matrice di 5 colonne e 10 righe (una per ruota). Ricordiamo inoltre che nella terminologia del lotto il termine *cinquina* (senza specificazioni) è usato come sinonimo di *insieme,* per indicare cioè 5 numeri prescindendo dal loro ordine di estrazione. Lo stesso dicasi per i termini *ambo, terno, quaterna,* che significano nell'ordine *insiemi* di 2, 3, 4 numeri.

1 *Calcolare la probabilità che il primo estratto sulla ruota di Venezia sia di una cifra.*

I casi possibili sono 90: *il primo estratto è i, $i = 1, \ldots, 90$.* Per ragioni di simmetria li giudichiamo equiprobabili. La probabilità del nostro evento *il primo estratto è di una cifra* è allora $9/90 = 1/10$.

2 *Calcolare la probabilità che il 27 venga estratto sulla ruota di Venezia.*

Possiamo considerare ora i 6 seguenti casi possibili: *il 27 viene estratto al primo posto , ... , il 27 viene estratto al quinto posto, il 27 non viene estratto.* Non possiamo però invocare ragioni di simmetria per dire che essi sono equiprobabili, perché è previsto che vengano estratti 5 numeri da un'urna che ne contiene 90, ed è allora evidente che il 27 ha più possibilità di non essere estratto che di esserlo. L'ultimo caso possibile dovrebbe perciò essere di probabilità superiore agli altri cinque messi assieme, e quindi a ciascuno di essi singolarmente. Se si vuole condurre il discorso in uno schema di simmetria, occorre allora cercare una descrizione più adeguata.
Risponde allo scopo quella che si può ottenere pensando di protrarre l'estrazione delle palline oltre la quinta fino al loro esaurimento. Questa descrizione non corrisponde a quanto avviene realmente in una estrazione eseguita ai fini del gioco del lotto. Ma non si vedono neanche motivi che impediscano a chi estrae, se ne ha voglia, di proseguire nell'estrazione (nel gioco della tombola, ad esempio, si estrae quanto occorre, andando ben oltre la quinta estrazione). Idealmente, quindi, è lecito pensare di eseguire estrazioni fino ad ottenere lo svuotamento dell'urna. Si possono allora considerare i casi possibili: *il 27 viene estratto al posto i*, $i = 1, \dots, 90$, per i quali è ragionevole riconoscere le condizioni di simmetria. I primi 5 sono favorevoli all'evento E = *il 27 viene estratto*, la cui probabilità è perciò $P(E) = 5/90$.
Una impostazione alternativa, che evita l'artifizio di dover immaginare un'estrazione completa, è la seguente: i numeri estratti su una ruota sono 5 su 90. In accordo con la premessa, i casi possibili si possono allora identificare con le *cinquine* (non ordinate, sottoinsiemi di cinque elementi) che si possono formare con i primi 90 numeri naturali e si possono definire con le proposizioni: *viene estratta la cinquina* $n_1, \dots, n_5$ *(il sotto-insieme* $\{n_1, \dots, n_5\}$*)*, $1 \le n_1 < \dots < n_5 \le 90$. I casi possibili sono $\binom{90}{5}$ e $\binom{89}{4}$ quelli favorevoli ad E, pari al numero delle cinquine che contengono il 27. Si ha pertanto:

$$P(E) = \binom{89}{4} \Big/ \binom{90}{5} = \frac{(89)_4}{4!} \; \frac{5!}{90\,(89)_4} = \frac{5}{90},$$

che è lo stesso risultato di prima.

3 *Calcolare la probabilità che sulla ruota di Venezia esca il terno 1, 19, 31.*

Il problema è analogo al precedente. Seguendo la seconda impostazione ivi proposta, ora abbiamo $\binom{87}{2}$ casi favorevoli: le cinquine che contengono 1, 19, 31 e due qualunque degli 87 numeri rimanenti. La probabilità dell'evento E = *esce il terno 1, 19, 31* è allora:

$$P(E) = \binom{87}{2}\Big/\binom{90}{5} = \frac{(87)_2}{2!}\,\frac{5!}{(90)_3\cdot(89)_2} = \frac{3\cdot4\cdot5}{90\cdot89\cdot88} = 0{,}00008512.$$

Si perviene naturalmente allo stesso risultato per qualsiasi altro terno.

4 *Calcolare la probabilità che il terno i, j, k esca su almeno una delle 10 ruote,* $1 \le i < j < k \le 90$.

Ragionando in termini di cinquine i casi possibili sono $\binom{90}{5}^{10}$, definiti dalle proposizioni $\bigvee_{h=1}^{10}$ *sulla ruota h esce la cinquina* $n_{1,h}, \dots, n_{5,h}$, $1 \le n_{1,h} < \dots < n_{5,h} \le 90$, $h = 1, \dots, 10$. L'insieme dei casi favorevoli è complementare di quello degli sfavorevoli (*Teorema* 3.4.3), la cui cardinalità è più facile da determinare. Cominciamo con l'osservare che è $\binom{90}{5} - \binom{87}{2}$ il numero delle cinquine che non contengono il terno i, j, k su una delle 10 ruote, qualunque essa sia. È allora $\left[\binom{90}{5} - \binom{87}{2}\right]^{10}$ il numero dei casi sfavorevoli all'uscita del terno su almeno una delle ruote, e quindi $\binom{90}{5}^{10} - \left[\binom{90}{5} - \binom{87}{2}\right]^{10}$ quello dei casi favorevoli alla sua uscita. Posto E = *esce il terno i, j, k su almeno una delle 10 ruote* si ha:

$$P(E) = 1 - \left[\left(\binom{90}{5} - \binom{87}{2}\right)\Big/\binom{90}{5}\right]^{10} = 1 - \left(1 - \frac{(5)_3}{(90)_3}\right)^{10} = 0{,}00085088.$$

5 *Calcolare la probabilità che sulla ruota di Venezia esca una cinquina in ordine crescente.*

La precedente descrizione in cinquine *non ordinate* è inadeguata per dare risposta all'attuale quesito. Bisogna ricorrere a una descrizione mediante cinquine *ordinate*, usando le proposizioni: *escono nell'ordine i cinque numeri* $n_1, \dots, n_5$, *distinti a due a due*, $1 \le n_1, \dots, n_5 \le 90$. I casi possibili sono allora $(90)_5$. Quelli favorevoli all'evento E = *esce una cinquina in ordine crescente* sono $\binom{90}{5}$, tanti quanti sono i sottoinsiemi di cinque numeri scelti tra i primi 90 naturali (ogni cinquina ordinata è una permutazione di cinque numeri; di esse una sola li elenca in modo ordinato). Allora si ha $P(E) = \binom{90}{5}/(90)_5 = 1/5! = 0{,}008333$.

6 *Con riferimento al gioco del poker, calcoliamo la probabilità dei seguenti eventi:*

E = *la mano servita a Tizio è una doppia coppia alle Donne*

D = *la mano servita a Tizio è una doppia coppia*
S = *la mano servita a Tizio è una scala (non reale)*

Le carte del poker sono 32, distinte in 8 ranghi, A, R, D, F, 10, 9, 8, 7 – elencati qui in ordine di valore decrescente –, e 4 semi (♥, ♦, ♣, ♠). All'inizio ogni giocatore riceve cinque carte. I punti del poker, tra i quali la *doppia coppia* e la *scala* nominati sopra, non dipendono dall'ordine in cui i giocatori ricevono le carte. La situazione è perciò analoga a quella del gioco del lotto, ove la realizzazione di ambo, terno, quaterna e cinquina non dipende dall'ordine in cui vengono estratti i numeri. Ai fini del calcolo della probabilità dei punti che possono essere serviti nel gioco del poker a un giocatore si può descrivere la situazione mediante sottoinsiemi di cinque carte. Essendo 32 le carte disponibili, i casi possibili sono allora $\binom{32}{5}$.

Il termine *coppia* indica due carte della stesso rango: due assi, due re, ... , due 7. La mano è una *doppia coppia* se è costituita da due coppie di ranghi diversi e da una carta di un terzo rango. *Doppia coppia alle Donne* significa che la coppia di rango più alto è costituita da due Donne.
Ciò precisato, contiamo i casi (le mani) favorevoli all'evento E scegliendo prima i ranghi e dopo i semi. Allora si ha: $\binom{4}{2}\binom{5}{1}\binom{4}{2}\binom{6}{1}\binom{4}{1}$. A parole: sono $\binom{4}{2}$ i modi di scegliere le due Donne della coppia assegnata; $\binom{5}{1}$ sono i modi di scegliere il rango della seconda coppia (di rango inferiore a quello delle Donne) e $\binom{4}{2}$ i modi di scegliere in questo rango i due semi della coppia; $\binom{6}{1}$ sono i modi di scegliere tra i 6 ranghi rimasti quello della quinta carta e $\binom{4}{1}$ modi di scegliere in esso il seme. Facendo il rapporto tra il numero dei casi favorevoli così determinato e quello dei casi possibili otteniamo: $P(E) = 0{,}0215$.

I casi favorevoli all'evento D si contano in modo analogo al precedente. Essi sono: $\binom{8}{2}\binom{4}{2}\binom{4}{2}\binom{6}{1}\binom{4}{1}$. A parole: sono $\binom{8}{2}$ modi di scegliere due ranghi per le due coppie; sono poi $\binom{4}{2}$ i modi di scegliere in uno dei due ranghi i due semi per una delle coppie e il secondo $\binom{4}{2}$ ha lo stesso significato per l'altra coppia; il $\binom{6}{1}\binom{4}{1}$ si spiega in modo analogo a quello dell'evento E. Allora si trova $P(D) = 0{,}1201$.

Il termine *scala* indica che le cinque carte di una mano sono di ranghi consecutivi, essendo l'asso oltre che di rango massimo, anche di rango minimo (precede quello del 7). I cinque ranghi di una scala minima sono A, 7, 8, 9, 10 e quelli di una massima 10, F, D, R, A. Sicché, con riferimento al rango le scale sono di cinque tipi: quelle che cominciano con Asso, con 7, ... , con 10. Se le carte in scala sono dello stesso seme, la scala è detta *reale*. Le altre sono scale semplici (spesso nel gergo del poker scale e basta).

Per contare il numero di tutte le possibili scale teniamo conto del fatto che sono 5 i modi di cominciare una scala e che ogni scala è costituita da carte di 5 ranghi. Allora, sono 5 i modi di scegliere il rango iniziale della scala e restano così determinati i suoi quattro ranghi successivi. Per ciascuno di essi sono $\binom{4}{1}$ i modi di scegliere il seme. Pertanto, è $5\binom{4}{1}^5$ il numero delle mani in scala. Tra queste mani ci sono anche le scale reali, per contare le quali, come per le altre, bisogna scegliere prima il rango d'inizio – in 5 modi –, ma scegliere poi un solo seme per tutti i cinque ranghi. Il numero delle mani in scala reale è allora $5\binom{4}{1}$. Segue allora che i casi favorevoli all'evento S sono $5\binom{4}{1}^5 - 5\binom{4}{1}$, numero che rapportato a quello dei casi possibili fornisce: $P(S) = 0{,}0253$.

7 *Con riferimento al gioco del bridge, calcoliamo la probabilità dei seguenti eventi:*

T = *Tizio riceve una mano tricolore*
B = *Tizio riceve una mano bilanciata*

Le carte del bridge sono 52. A quelle del poker si aggiungono cinque carte per ogni seme (20 in tutto) di ranghi 6, 5, 4, 3, 2, in ordine decrescente e inferiore al 7. Le carte vengono distribuite a quattro giocatori fino all'esaurimento del mazzo. Anche nel bridge il gioco non dipende dall'ordine in cui i giocatori ricevono le carte. Con riferimento a quelle che riceve uno di essi, l'incertezza può essere allora descritta mediante i sottoinsiemi di 13 carte. I casi (le mani) possibili sono allora $\binom{52}{13}$. Ha molta importanza nel gioco del bridge quella che nel gergo si chiama *distribuzione della mano* e che si esprime con una quaterna di numeri, ciascuno dei quali indica il numero di carte di un seme, senza precisare quale. Ad esempio, la quaterna 6.5.1.1 indica una mano che ha 6 carte di un seme, 5 di un secondo seme e una per ciascuno degli altri due. Le mani che hanno distribuzioni 4.3.3.3, 4.4.3.2, 5.3.3.2, sono dette *bilanciate*. Le altre distribuzioni sono classificate in base al numero di semi con più di tre carte e sono dette:

tricolore quelle con *tre semi* con più di 3 carte (le distribuzioni 4.4.4.1 e 5.4.4.0),
bicolore quelle con *esattamente due semi* con più di 3 carte, esclusa la bilanciata 4.4.3.2.
monocolore quelle con *esattamente un seme* con più di 3 carte, esclusa la bilanciata 5.3.3.2.

Come indicato esplicitamente sopra, le mani bilanciate sono di tre tipi e quelle tricolore di due. I tipi di mani bicolore e monocolore sono molto più numerosi, tanto che per distinguerli si usa aggiungere il numero delle carte dei semi che superano 3. Ad esempio, *bicolore 6.5* – è una mano distribuita 6.5.1.1 o 6.5.2.0 –, *monocolore di 7 carte* – è una mano distribuita 7.2.2.2 o 7.3.2.1 o 7.3.3.0 –, eccetera.

Per il conteggio delle mani che hanno una assegnata distribuzione conviene scegliere prima il seme e poi il rango.

Ciò premesso, i casi favorevoli all'evento T sono le mani che hanno distribuzione 4.4.4.1 o (aut) 5.4.4.0. Cominciamo a contare le mani con distribuzione 4.4.4.1. Esse sono $\binom{4}{3}\binom{13}{4}^3\binom{1}{1}\binom{13}{1}$. A parole: i modi di scegliere i 3 semi di 4 carte sono $\binom{4}{3}$, per ciascuno dei quali sono $\binom{13}{4}$ i modi di scegliere i 4 ranghi; sono perciò $\binom{13}{4}^3$ i modi di scegliere i 4 ranghi per tutti tre i semi scelti; $\binom{1}{1}$ è l'unico modo rimasto per la scelta del seme di una carta e $\binom{13}{1}$ sono i modi di scegliere in esso il rango della carta medesima. Analogamente le mani distribuite 5.4.4.0 sono $\binom{4}{1}\binom{13}{5}\binom{3}{2}\binom{13}{4}^2$. Il numero dei casi favorevoli all'evento T è la somma di questi due numeri. Facendo il rapporto con $\binom{52}{13}$ – numero dei casi possibili – si ricava $P(T) = 0{,}0424$.

Le distribuzioni favorevoli a B sono le 4.3.3.3, 4.4.3.2, 5.3.3.2. I rispettivi casi favorevoli sono $\binom{4}{1}\binom{13}{4}\binom{3}{3}\binom{13}{3}^3$, $\binom{4}{2}\binom{13}{4}^2\binom{2}{1}\binom{13}{3}\binom{1}{1}\binom{13}{2}$, $\binom{4}{1}\binom{13}{5}\binom{3}{2}\binom{13}{3}^2\binom{1}{1}\binom{13}{2}$ e i casi favorevoli a B sono la loro somma. Facendo il rapporto con il numero dei casi possibili si trova $P(B) = 0.4760$.

8 *Con riferimento al lancio di due dadi, si calcoli la probabilità che il punto somma sia s, $s = 2, \dots, 12$.*

I casi possibili per il punto somma sono 11: *la somma è s*, $s = 2, \dots, 12$. Non si ravvisano però elementi che possano giustificare un giudizio di simmetria che porti a considerare gli undici casi equiprobabili. Al contrario, vi sono elementi palesi di asimmetria che suggeriscono una valutazione non bilanciata. Ad esempio, il *punto somma* 2 si ottiene in un solo modo: quando entrambi i dadi realizzano punto 1; il *punto somma* 4, invece, si ottiene addizionando 1 e 3 o 2 e 2, e quindi in tre modi: se entrambi i dadi realizzano punto 2, oppure se uno di essi realizza punto 1 e l'altro punto 3 o viceversa. In effetti, i dadi hanno una loro individualità – fisicamente sono distinti – anche quando non siamo in grado di riconoscerli. Pertanto, almeno idealmente possiamo pensare a un 1° e a un 2° dado, in modo che il risultato del lancio possa essere visto come una *coppia ordinata* (i, j), con i punto realizzato dal 1° dado e j dal 2°. I casi possibili sono allora:

il risultato del lancio è (i,j), $1 \le i,j \le 6$. In questa descrizione si possono ravvisare quelle condizioni di simmetria che andavamo cercando, tali da giustificare una valutazione di equiprobabilità degli attuali 36 casi possibili. Venendo ora al punto somma, ci possiamo aiutare con la matrice qui a fianco, ai cui margini (sinistro e superiore) si leggono i punti i, j realizzati dal 1° e dal 2° dado, mentre i suoi elementi sono il punto somma $i+j$, il quale risulta essere lo stesso lungo la diagonale secondaria e le linee ad essa parallele. Posto ω_s = *il punto somma è* s, dalla matrice si ricava:

$i \backslash j$	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$P(\omega_2) = P(\omega_{12}) = \frac{1}{36}, \quad P(\omega_3) = P(\omega_{11}) = \frac{2}{36}, \quad P(\omega_4) = P(\omega_{10}) = \frac{3}{36},$$

$$P(\omega_5) = P(\omega_9) = \frac{4}{36}, \quad P(\omega_6) = P(\omega_8) = \frac{5}{36}, \quad P(\omega_7) = \frac{6}{36}.$$

I prossimi esempi si riferiscono a estrazioni da un'urna che contiene palline fisicamente uguali, distinte solo per colore (quelle dello stesso colore non sono distinguibili).

9 *Da un'urna contenente 8 palline bianche e 2 nere, ne viene estratta una. Qual è la probabilità dell'evento E = «la pallina estratta è bianca»?*

I risultati possibili sono due: *la pallina estratta è bianca, la pallina estratta è nera*. È evidente che mancano le condizioni di simmetria, perché le possibilità di estrarre bianca superano quelle di estrarre nera. Per ricondurci a uno schema di simmetria, possiamo immaginare che le palline siano numerate: da 1 a 8 le bianche e 9, 10 le nere. I casi possibili sono allora 10: *esce pallina numerata* i, $i = 1, \ldots, 10$, che ora possiamo giudicare equiprobabili. In questa descrizione l'evento E è definito dalla proposizione $\bigvee_{h=1}^{8}$ *esce la pallina numero* h. I casi favorevoli ad E sono 8 e quindi $P(E) = 8/10$.

10 *Dall'urna del precedente Esempio 9, si eseguano in sequenza 3 estrazioni di una pallina per volta, seguite da reimbussolamento (rimessa). Detto successo l'estrazione di pallina bianca, calcolare le probabilità di 0, 1, 2, 3 successi.*

Per il computo dei casi possibili, osserviamo preliminarmente che al momento di ciascuna estrazione le palline nell'urna sono 10. Distinte le palline mediante numerazione come nell'esempio precedente, in ogni estrazione la pallina può essere scelta in 10 modi. Il risultato di un'estrazione completa è allora una terna di numeri da 1 a 10 con possibilità di ripetizioni (disposizione con ripetizione di 10 elementi di classe 3)[21]. I casi possibili sono perciò 10^3: *esce la terna* (h, i, k), $1 \le h, i, k \le 10$. Indicato con S il numero dei successi, dobbiamo calcolare la probabilità degli eventi definiti dalle proposizioni $S=n$, $n=0, 1, 2, 3$. Per calcolare i casi favorevoli a $S=n$, osserviamo preliminarmente che per formare una terna bisogna occupare 3 posti coi risultati della prima, seconda e terza estrazione, nell'ordine. Una terna favorevole a $S=n$ ha n posti occupati da palline bianche e i rimanenti $3-n$ da palline nere. I posti per le palline bianche si possono scegliere in $\binom{3}{n}$ modi e, per ogni scelta di posti, si possono collocare su di essi palline bianche in 8^n modi, e sui rimanenti palline nere in 2^{3-n} modi (disposizioni con ripetizione di 8 elementi di classe n e di 2 elementi di classe $3-n$, rispettivamente). Il numero dei casi favorevoli a $S=n$ è allora $\binom{3}{n} 8^n 2^{3-n}$. Pertanto si ha:

$$P(S=n) = \binom{3}{n} 8^n 2^{3-n} \big/ 10^3 = \binom{3}{n}\left(8/10\right)^n \left(2/10\right)^{3-n}.$$

Risulta naturalmente:

$$\sum_{n=0}^{3} P(S=n) = \sum_{n=0}^{3} \binom{3}{n}\left(8/10\right)^n \left(2/10\right)^{3-n} = \left(\frac{8}{10}+\frac{2}{10}\right)^3 = 1.$$

11 *Poniamo gli stessi quesiti dell'Esempio 10, con riferimento però a una sequenza di 3 estrazioni senza rimessa.*

Pensando le palline numerate da 1 a 10, in analogia con quanto accade nel gioco del lotto abbiamo qui un'estrazione di 3 numeri naturali scelti tra i primi 10 (al posto di 5 scelti tra i primi 90). In termini di terne non ordinate i casi possibili sono $\binom{10}{3}$: *esce la terna (l'insieme)* $\{n_1, n_2, n_3\}$, $1 \le n_1 < n_2 < n_3 \le 10$.
Anche i casi favorevoli agli eventi $S=n$ devono allora essere contati in termini di terne non ordinate, ovvero di sottoinsiemi di 3 elementi. L'evento $S=0$ non ha casi favorevoli – è l'evento impossibile –, perché

21 Anche nel caso del lancio dei due dadi dell'*Esempio* 6, i risultati sono disposizioni con ripetizione: di 6 elementi di classe 2.

non è possibile scegliere 3 numeri tra 9 e 10. Perciò $P(S=0)=0$. I sottoinsiemi favorevoli a $S=n$, $n=1, 2, 3$, sono costituiti da n numeri scelti tra i primi 8 e $3-n$ tra gli ultimi due. Il loro numero è quindi $\binom{8}{n}\binom{2}{3-n}$. Allora si ha:

$$P(S=n) = \binom{8}{n}\binom{2}{3-n}\Big/\binom{10}{3} = \binom{8}{n}\binom{2}{3-n}\Big/120.$$

Risulta poi:

$$\sum_{n=1}^{3} P(S=n) = \frac{1}{120}\sum_{n=1}^{3}\binom{8}{n}\binom{2}{3-n} = \frac{1}{120}(8+56+56) = 1.$$

12 *Un'urna contiene 1 pallina bianca, 2 nere, 3 rosse e 4 verdi. Si eseguano quattro estrazioni senza rimessa. Calcolare la probabilità dei seguenti eventi:*

E_1 = *le palline sono tutte bianche*
E_2 = *le palline sono dello stesso colore*
E_3 = *le palline sono tutte di colore diverso*
E_4 = *una pallina è bianca, due sono verdi e una di altro colore*

Anche in questo esempio immaginiamo le palline numerate da 1 a 10. specificamente: 1 la pallina bianca, 2, 3 le nere, 4, 5, 6 le rosse, 7, 8, 9, 10 le verdi. Come nell'*Esempio 11*, in termini di quaterne non ordinate (sottoinsiemi di 4 elementi) i casi possibili sono $\binom{10}{4}=210$.
L'evento E_1 è impossibile: una sola pallina può essere bianca. Il numero dei casi favorevoli è 0 e quindi $P(E_1)=0$.
Le palline sono dello stesso colore solo se sono verdi. C'è un solo caso favorevole: *esce il sottoinsieme* $\{7, 8, 9, 10\}$. Perciò $P(E_2)=1/210$.
Dire che *le palline sono tutte di colore diverso* equivale ad affermare che *una pallina è bianca, una nera, una rossa e una verde.* I sottoinsiemi favorevoli a questo evento sono $1\cdot2\cdot3\cdot4$ – la pallina bianca può essere scelta in un solo modo, quella nera in due, quella rossa in tre e la verde in quattro – e perciò $P(E_3)=24/210$.
Per l'evento E_4 si ha: un posto della terna è occupato da una pallina bianca in 1 modo, due posti da palline verdi in $\binom{4}{2}$ modi e il posto rimanente da una pallina nera o rossa in $\binom{5}{1}$ modi. I casi favorevoli sono allora $1\cdot6\cdot5$ e perciò $P(E_4)=30/210=1/7$.

13 *Due palline dello stesso colore vengono collocate una alla volta in un cassetto scelto fra tre disponibili. Elencare le collocazioni possibili e calcolare la loro probabilità nelle due seguenti ipotesi:*

(i) *i cassetti possono contenere più palline;*
(ii) *i cassetti possono contenere al più una pallina.*

Indichiamo con (h, i, j) il risultato della collocazione che prevede h palline nel 1° cassetto, i nel 2° e j nel 3°.

Nell'ipotesi (i) sono possibili i casi:

$A_{h,i,j}$ = *il risultato della collocazione è* (h, i, j), $0 \leq h, i, j \leq 2$, $h+i+j=2$.

Ad operazione ultimata avremo dunque uno dei sei seguenti risultati:

$$(2, 0, 0), (0, 2, 0), (0, 0, 2), (1, 1, 0), (1, 0, 1), (0, 1, 1).$$

Per vedere se è giustificato esprimere un giudizio di simmetria su questi risultati, non dobbiamo scordare che le palline sono fisicamente distinte anche se per noi indistinguibili (come negli esempi precedenti). Immaginiamo perciò di etichettarle a, b. Allora, abbiamo a disposizione 3 scelte per collocare la pallina a e – a seguire – ancora 3 scelte per collocare la pallina b (è come estrarre un cassetto e reimbussolarlo). Pertanto, i casi possibili, che ora possono essere giudicati equiprobabili, sono in numero di $3^2 = 9$, contro i 6 elencati in precedenza. Indicando le palline con le loro etichette, possiamo rappresentare graficamente i nove risultati come segue:

ab		
	ab	
		ab
a	b	
a		b
	a	b
b	a	
b		a
	b	a

Pertanto, un solo caso è favorevole a ciascuna delle prime tre collocazioni e due a ciascuna delle altre tre. Si ricava allora:

$$P(A_{2,0,0}) = P(A_{0,2,0}) = P(A_{0,0,2}) = \frac{1}{9}, \; P(A_{1,1,0}) = P(A_{1,0,1}) = P(A_{0,1,1}) = \frac{2}{9}.$$

Nell'ipotesi (ii) sono possibili i casi:

$B_{h,i,j}$ = *il risultato della collocazione è* (h, i, j), $0 \leq h, i, j \leq 1$, $h+i+j=2$

e di conseguenza, rispetto a prima, sono possibili solo i tre risultati (1, 1, 0), (1, 0, 1), (0, 1, 1). Il giudizio di equiprobabilità appare qui giustificato dalla simmetria di questi risultati (due cassetti occupati e uno vuoto). Tuttavia è sempre prudente rifarsi a uno schema che interpreta fedelmente le operazioni che vengono eseguite. Nel nostro caso, conviene rimettersi nell'ottica precedente: dare cioè individualità alle palline e poi collocarle rispettando l'attuale vincolo. Abbiamo allora 3 scelte per la pallina a e – a seguire – 2 per la pallina b (nulla cambia nel risultato se si colloca prima la b e poi la a). Le collocazioni da giudicare simmetriche sono così 6 – le ultime rappresentate sopra graficamente – e a coppie danno luogo alle tre collocazioni indistinguibili, che risultano pertanto equiprobabili. Riesce allora:

$$P(B_{1,1,0}) = P(B_{1,0,1}) = P(B_{0,1,1}) = \frac{2}{6} = \frac{1}{3}.$$

4.2 Probabilità come massa

Le condizioni poste per trattare gli esempi del paragrafo appena concluso prevedono che nei modelli simmetrici la descrizione dell'incertezza sia fatta mediante una partizione finita $\{\omega_1, \ldots, \omega_n\}$ di casi possibili e che agli eventi – ai sottoinsiemi di casi possibili – venga assegnata probabilità uguale al rapporto tra i casi a loro favorevoli e i casi possibili. Si può pervenire allo stesso risultato interpretando la probabilità come una *massa unitaria* da equiripartire tra i casi possibili. In tal caso infatti ogni caso possibile riceve massa pari a $1/n$ e, poiché la massa è additiva, quella che compete agli altri eventi è la somma delle masse dei casi a loro favorevoli; cioè, a conti fatti, al rapporto tra i casi a favorevoli e i casi possibili, come prima. Questa interpretazione della probabilità come massa suggerisce un'ovvia estensione delle valutazioni in ambiente finito, prevedendo che la massa possa essere distribuita sui casi possibili (le alternative) anche in modo ineguale. Distribuzioni siffatte le abbiamo del resto già considerate parlando dei risultati 1, X, 2 di una partita di calcio in 2.1.1 *Esempio* 3 ed anche in alcuni esempi del n° 4.1.1 precedente. Come nel caso dell'*Esempio* 8, ove con riferimento al punto realizzato lanciando due dadi abbiamo considerato le alternative ω_i = *il punto somma realizzato è* i, $i = 2, \ldots, 12$, anche se in quell'esempio per attribuire la probabilità agli ω_i abbiamo ricondotto il problema a uno schema di simmetria introducendo una descrizione più dettagliata, nella quale gli ω_i sono somme delle alternative $\omega_{i,j}$ = *il risultato del lancio è* (i,j), $1 \le i, j \le 6$.

A fronte di una descrizione dell'incertezza mediante un insieme finito di alternative, considereremo dunque ammissibili le valutazioni che si ottengono secondo quanto previsto dalla seguente definizione.

4.2.1 Definizione. Probabilità in ambiente finito.

Sia $\mathbb{P} = \{\omega_1, \dots, \omega_n\}$ *una partizione finita in eventi elementari* (3.4.1 *Nota e terminologia*). *Diremo* ***distribuzione di probabilità*** *(o semplicemente probabilità) P ogni valutazione (applicazione) che si ottiene ripartendo a piacere la massa (probabilità) 1 tra gli eventi elementari* $\omega_1, \dots, \omega_n$ *e prolungandola poi per additività sugli eventi logicamente dipendenti dalla partizione (sull'algebra* $\mathcal{A}_L(\mathbb{P})$*), ovvero sulle somme logiche degli eventi elementari di* $\mathbb{P}$ *(Teorema* 3.4.3*). Sicché si ha:*

$$P(\omega_i) \geq 0,\ i = 1, \dots, n,\ \sum_{i=1}^{n} P(\omega_i) = 1 \quad e \quad P(E) = \sum_{\omega \in \mathcal{E}} P(\omega),$$

per ogni $E = \vee\mathcal{E}$, $\mathcal{E} \subset \{\omega_1, \dots, \omega_n\}$.

NOTA.

Osserviamo che la condizione $P(\omega_i) \geq 0$ considera lecito porre $P(\omega_i) = 0$, pur essendo ω_i possibile. Si ammette con ciò che *eventi possibili* possano avere probabilità *nulla* – come l'evento impossibile – e perciò anche probabilità 1 – come l'evento certo (quando le loro negazioni hanno probabilità 0) –. L'opportunità di consentire di attribuire probabilità nulla (oppure 1) a *eventi possibili* sarà giustificata nel prossimo capitolo (*Proposizione* 5.1.1). Qui segnaliamo invece che questa definizione caratterizza le probabilità coerenti in ambiente finito. Si veda in proposito 10.5.1 *Proposizione*.

4.2.2 Teorema. Proprietà della probabilità.

Siano $\mathbb{P} = \{\omega_1, \dots, \omega_n\}$ *una partizione di* Ω *in eventi elementari,* $E, E_1, \dots, E_n$ *eventi logicamente dipendenti da* $\mathbb{P}$ *e P una probabilità. Allora si ha:*

a) $P(\Omega) = 1,\ P(\phi) = 0$ — *condizione di normalizzazione*

b) $0 \leq P(E) \leq 1$ — *condizione di non negatività*

c) Se $E_1 \wedge E_2 = \phi$, *allora* $P(E_1 \vee E_2) = P(E_1) + P(E_2)$ — *proprietà additiva* *d)* $P(\bar{E}) = 1 - P(E)$

e) $P(E_1 \vee \ldots \vee E_n) = \sum_{i=1}^{n} P(E_h) - \sum_{h<i}^{1,n} P(E_h \wedge E_i) + \sum_{h<i<j}^{1,n} P(E_h \wedge E_i \wedge E_j) - \ldots$
$+ (-1)^{n+1} P(E_1 \wedge \ldots \wedge E_n)$ *formula d'inclusione-esclusione*

f) *Se* $E_i \wedge E_j = \phi$, $1 \le i < j \le n$, *allora* $P(E_1 \vee \ldots \vee E_n) = P(E_1) + \ldots + P(E_n)$

g) $P(E_1 \vee \ldots \vee E_n) \le P(E_1) + \ldots + P(E_n)$ *proprietà subadditiva*

h) *Se* $E_1 \Rightarrow E_2$, *allora* $P(E_1) \le P(E_2)$ *proprietà di monotonia*

Dimostrazione.

Le proprietà *a)* e *b)* sono conseguenza immediata della definizione.

Prova della c).
Posto $E_1 = \bigvee \mathcal{E}_1$, $E_2 = \bigvee \mathcal{E}_2$, riesce $E_1 \vee E_2 = \bigvee(\mathcal{E}_1 \cup \mathcal{E}_2)$ (*Proposizione* 3.5.1) e per definizione si ha

$$P(E_1 \vee E_2) = \sum_{\omega \in \mathcal{E}_1 \cup \mathcal{E}_2} P(\omega) =$$
$$\sum_{\omega \in \mathcal{E}_1} P(\omega) + \sum_{\omega \in \mathcal{E}_2} P(\omega) = P(E_1) + P(E_2),$$

valendo la seconda uguaglianza perché da $E_1 \wedge E_2 = \phi$ segue che $\mathcal{E}_1$ e $\mathcal{E}_2$ sono disgiunti (*Proposizione* 3.5.1).

Le proprietà successive si provano usando le prime tre appena dimostrate.

Prova della d).
Riesce $E \wedge \bar{E} = \phi$ e $E \vee \bar{E} = \Omega$ (3.2.5*c*). Per le *c)* e *a)* si ottiene $P(E \vee \bar{E}) = P(E) + P(\bar{E}) = P(\Omega) = 1$, da cui segue la *d)*.

Prova della e).
Si ottiene per induzione (base $n = 2$).
Per $n = 2$ si deve mostrare che $P(E_1 \vee E_2) = P(E_1) + P(E_2) - P(E_1 \wedge E_2)$. Riesce intanto $E_1 \vee E_2 = E_1 \vee [(E_1 \wedge E_2) \vee (\bar{E}_1 \wedge E_2)] = E_1 \vee (\bar{E}_1 \wedge E_2)$ (3.2.5*e,c,i* e 3.3.2*i,f*) con E_1 e $\bar{E}_1 \wedge E_2$ incompatibili. In forza della *c)* riesce allora:

$$P(E_1 \vee E_2) = P(E_1) + P(\bar{E}_1 \wedge E_2) = P(E_1) + P(\bar{E}_1 \wedge E_2) + P(E_1 \wedge E_2) - P(E_1 \wedge E_2).$$

Ancora per la *c)* si ha $P(\bar{E}_1 \wedge E_2) + P(E_1 \wedge E_2) = P[\bar{E}_1 \wedge E_2) \vee (E_1 \wedge E_2)] = P(E_2)$ e quindi la tesi[22].

---

22 La $P(E_1 \vee E_2) = P(E_1) + P(E_1) - P(E_1 \wedge E_2)$ si interpreta in termini di massa come segue. Sommando le masse dei componenti di E_1 e di E_2, quelle dei

Per semplicità di scrittura mostriamo ora come si sfrutta il risultato base per ottenere la *e)* nel caso $n = 3$. Il passo induttivo, solo formalmente più complicato, è lasciato al lettore per esercizio. Riesce:

$$P(E_1 \vee E_2 \vee E_3) = P(E_1 \vee E_2) + P(E_3) - P[(E_1 \vee E_2) \wedge E_3] =$$
$$P(E_1 \vee E_2) + P(E_3) - [P(E_1 \wedge E_3) + P(E_2 \wedge E_3) - P(E_1 \wedge E_2 \wedge E_3)] =$$
$$P(E_1) + P(E_2) - P(E_1 \wedge E_2) + P(E_3) - P(E_1 \wedge E_3) - P(E_2 \wedge E_3) + P(E_1 \wedge E_2 \wedge E_3),$$

e quindi la conclusione.

Prova della f).
Riesce (3.3.2*i*) $E_{i_1} \wedge \ldots \wedge E_{i_s} \Rightarrow E_{i_1} \wedge E_{i_s} = \phi$ per ogni $2 \leq s \leq n$ e $1 \leq i_1 < \ldots < i_s \leq n$, ove l'uguaglianza vale per ipotesi. Pertanto $\phi \Rightarrow E_{i_1} \wedge \ldots \wedge E_{i_s} \Rightarrow \phi$ (3.3.2*g*), quindi $E_{i_1} \wedge \ldots \wedge E_{i_s} = \phi$ (3.3.2*c*) e dalle *e)* e *a)* segue allora la *f)*.

Prova della g).
Anche questa si prova per induzione a partire dalla *e)*.
Base. Conseguenza di *e), b),* che per $n = 2$ porgono:

$$P(E_1 \vee E_2) = P(E_1) + P(E_2) - P(E_1 \wedge E_2) \leq P(E_1) + P(E_2).$$

Passo induttivo. Sia $P(E_1 \vee \ldots \vee E_{n-1}) \leq P(E_1) + \ldots + P(E_{n-1})$, $n > 2$. Usando il risultato base si ricava

$$P(E_1 \vee \ldots \vee E_n) \leq P(E_1 \vee \ldots \vee E_{n-1}) + P(E_n) \leq P(E_1) + \ldots + P(E_{n-1}) + P(E_n).$$

Prova della h).
Poiché $E_1 \wedge E_2 = E_1$, si ha $E_2 = (E_1 \wedge E_2) \vee (\bar{E}_1 \vee E_2) = E_1 \vee (\bar{E}_1 \vee E_2)$. Segue:

$$P(E_2) = P(E_1) + P(\bar{E}_1 \vee E_2) \geq P(E_1).$$

La dimostrazione del teorema è così completa. ■

Esempi.

Tutti quattro gli esempi che seguono vengono risolti ponendosi in ipotesi di simmetria. Gli ultimi tre sono molto noti in letteratura.

---

componenti comuni a E_1 e E_2 vengono contate due volte. Per ottenere la somma corretta delle masse dei componenti di $E_1 \vee E_2$, esse devono perciò essere tolte una volta. Lasciamo al lettore il compito di interpretare in modo analogo la formula d'inclusione-esclusione per $n = 3$.

1 *Tizio gioca il terno con ambo 1,2,3 sulla ruota di Venezia. Trovare la probabilità che vinca (vince se esce il terno o uno dei tre ambi).*

Introduciamo i seguenti eventi:

E = *Tizio vince*
T = *esce il terno*
$A_{h,i}$ = *esce l'ambo h,i e non il terzo numero,* $1 \leq h < i \leq 3$

Si ha $E = T \vee A_{1,2} \vee A_{1,3} \vee A_{2,3}$ con gli eventi della somma logica a due a due incompatibili. Riesce $P(T) = \binom{87}{2} / \binom{90}{5} = (5)_3 / (90)_3$. Per motivi di simmetria è inoltre evidente che i tre eventi $A_{h,i}$ sono equiprobabili. Riesce $P(A_{h,i}) = \binom{87}{3} / \binom{90}{5} = 85\,(5)_2 / (90)_3$. Per la 4.2.2*f* allora si trova:

$$P(E) = P(T) + P(A_{1,2}) + P(A_{1,3}) + P(A_{2,3}) =$$
$$(5)_3 / (90)_3 + 3 \cdot 85 \cdot (5)_2 / (90)_3 = 0{,}00732.$$

2 *Dire se è più probabile fare almeno un 6 lanciando 4 volte un dado o almeno un doppio 6 lanciando 24 volte due dadi.*

L'inizio della moderna teoria delle probabilità si fa risalire a una famosa corrispondenza tra i matematici Blaise Pascal (1623, 1662) e Pierre de Fermat (1601, 1665), originata da alcuni problemi posti a Pascal dal Cavaliere di Mèré, un noto giocatore dell'epoca. Uno di questi problemi è quello proposto nell'esempio. Il Cavaliere di Mèré pensava che le due probabilità dovessero essere uguali. Forse perché la probabilità di fare doppio 6 lanciando due dadi è sei volte inferiore a quella di fare 6 lanciando una volta un dado (1/36 contro 1/6). Interpretando le probabilità in termini frequentisti si dovrebbe vedere allora un doppio 6 circa ogni 36 lanci di una coppia di dadi e un 6 circa ogni 6 lanci di un dado. Per dare luogo a situazioni paragonabili in termini di probabilità in partite di più lanci, i numeri di questi dovrebbero allora essere inversamente proporzionali alle probabilità di realizzare il punto desiderato in un singolo lancio. Mettendosi in questo ordine di idee, ai fini della valutazione delle probabilità degli eventi E_6 = *lanciando 4 volte un dado si realizza almeno un 6* e $E_{6,6}$ = *lanciando 24 volte due dadi si realizza almeno un doppio 6,* 4 lanci di un dado dovrebbero "equivalere" a 24 lanci di una coppia di dadi. Questo ragionamento (era quello del Cavaliere di Mèré?) è quasi deterministico, e non sorprende perciò che la risposta possa essere negativa. Vedremo in effetti che le probabilità di E_6 è maggiore – anche se di poco – a quella di $E_{6,6}$. Per rispondere a entrambi i quesiti conviene calcolare la probabilità della negazione dei rispettivi eventi.

Calcolo della probabilità di E_6.
Sono 6 i risultati possibili per ciascun dado e quindi 6^4 i risultati possibili – eventi elementari – del lancio congiunto. Per ogni dado sono poi 5 i punti possibili diversi da 6 e quindi 5^4 i casi favorevoli all'evento $\bar{E}_6$ = *in nessuno dei quattro lanci si realizza 6.* Per la 4.2.2*d* allora si ha: $P(E_6) = 1 - P(\bar{E}_6) = 1 - (5/6)^4 = 0{,}5178$.

Calcolo della probabilità di $E_{6,6}$. Sono 36 i punti possibili quando si lancia una coppia di dadi – i «punti coppia» di 4.1.1*Esempio* 8 –. Sono perciò 36^{24} i casi possibili e 35^{24} quelli favorevoli all'evento $\bar{E}_{6,6}$. Segue: $P(E_{6,6}) = 1 - P(\bar{E}_{6,6}) = 1 - (35/36)^{24} = 0{,}4914$.

3 Problema dei compleanni.

Calcolare la probabilità che in un gruppo di n persone almeno due abbiano il compleanno nello stesso giorno dell'anno (evento E).

Semplifichiamo il problema prescindendo dagli anni bisestili. Supponiamo cioè che in ogni anno le date di nascita possibili siano 365 – escludendo dal gruppo i nati il 29 febbraio –. Allora riesce $P(E) = 1$ se $n > 365$. Altrimenti, per rispondere alla domanda possiamo immaginare un'urna contenente i 365 giorni dell'anno, dalla quale andiamo ad eseguire n estrazioni con rimessa per attribuire la data di nascita a ciascuna delle n persone del gruppo. Sono allora 365^n le attribuzioni (i casi) possibili. Circa il numero dei casi favorevoli, anche in questo problema è più facile il calcolo di quelli sfavorevoli, ovvero dei casi favorevoli a $\bar{E}$ = *le date di nascita delle n persone sono tutte diverse.* Si ottiene allora che questi sono $(365)_n$, perché sono 365 i modi di scegliere la data di nascita di una prima persona, 364 i modi di scegliere quella di una seconda persona diversa dalla prima, eccetera. Pertanto si ha $P(E) = 1 - (365)_n / 365^n$.
È agevole eseguire i calcoli per piccoli valori di n. Per $n = 2, 3, 5, 10$, si trova nell'ordine 0,0027, 0,0082, 0,0271, 0,1169. Al crescere di n i calcoli finiscono col diventare troppo onerosi per un calcolo manuale. Con l'ausilio di un computer il calcolo diventa invece banale per ogni $n \leq 365$. Per $n = 22, 23, 24$,40 si trova 0,4757, 0,5073, 0,5383, 0,8912. Come si vede, già in un gruppo di 23 persone si ha probabilità superiore a $1/2$ di trovare almeno una coppia di persone che hanno la stessa data di nascita. Abbiamo scritto «già» perché sembra che siano i più quelli che, dovendo rispondere intuitivamente, pensano che per raggiungere quella probabilità il gruppo dovrebbe essere più numeroso.

4 Problema delle concordanze.

Da un'urna contenente n palline numerate da 1 a n si eseguono estrazioni senza rimessa di una pallina per volta fino a vuotare l'urna. Calcolare la probabilità che per almeno una pallina il numero che la contraddistingue coincida con quello dell'ordine d'estrazione.

Posto E_i = *all'i-esima estrazione esce la pallina numero i*, si deve calcolare la probabilità dell'evento $E = E_1 \vee \ldots \vee E_n$. Per rispondere al quesito, osserviamo che un'adeguata descrizione dell'attuale situazione d'incertezza può essere fatta mediante la partizione di generico evento elementare *l'esito dell'estrazione è l'n-pla* $(h_1, \ldots, h_n)$, ove $(h_1, \ldots, h_n)$ è una permutazione dei primi n numeri naturali. I casi possibili sono perciò $n!$. Vediamo ora quello che si può dire dei casi favorevoli (componenti) degli eventi $E_{i_1} \wedge \ldots \wedge E_{i_s}$, $1 \leq i_1 < \ldots < i_s \leq n$. Un componente di $E_{i_1} \wedge \ldots \wedge E_{i_s}$ si ottiene ponendo le palline $i_1, \ldots, i_s$ nei posti i_1-esimo, ..., i_s-esimo e le altre come si vuole nei rimanenti $n-s$ posti. E ciò può essere fatto in $(n-s)!$ modi. Sono perciò $(n-s)!$ i componenti di $E_{i_1} \wedge \ldots \wedge E_{i_s}$ e il loro numero dipende – come si vede – da *quanti* sono gli indici $i_1, \ldots, i_s$, ma non da *quali*. Ciò premesso, usando la formula d'inclusione-esclusione (4.2.2*e*), si ottiene:

$$P(E) = \binom{n}{1}\frac{(n-1)!}{n!} - \binom{n}{2}\frac{(n-2)!}{n!} + \ldots + (-1)^{n+1}\binom{n}{n}\frac{0!}{n!} =$$
$$= 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots + (-1)^{n+1}\frac{1}{n!}.$$

La probabilità $P(E)$ è dunque la somma ridotta n-esima della serie di Taylor della funzione $1-e^x$ calcolata in -1. Poiché si tratta di una serie a termini alterni e decrescenti in valore assoluto, il primo dei termini trascurati fornisce una limitazione dell'errore che si commette - nei riguardi della somma della serie - arrestando la somma a quel punto. Risulta allora $P(E) \simeq 1 - 1/e$ e l'errore che si commette non supera in valore assoluto $1/(n+1)!$. Ad esempio: per $n=5$ l'errore è inferiore a 0.0014, per $n=7$ è inferiore a 0.000025 (2.5×10^{-5}), per $n=10$ inferiore a 2.5×10^{-7}. Pertanto, da un certo n in poi $P(E)$ è praticamente costante e riesce $P(E) = 0.6321$. Nella tabella a fianco sono riportati i valori di $P(E)$ per $n = 1, \ldots, 8$, arrotondati alla quinta cifra decimale.

	$P(E)$
1	1.00000
2	0.50000
3	0.66667
4	0.62500
5	0.63333
6	0.63134
7	0.63214
8	0.63212

Segnaliamo che non è agevole risolvere questo problema calcolando la probabilità di $\bar{E}$ e poi per complemento a 1 quella di E. Almeno a prima vista non si scorgono infatti vie semplici per contare i casi favorevoli a $\bar{E}$, al contrario di quanto accade, come abbiamo visto, per quelli favorevoli agli eventi $E_{i_1} \wedge \ldots \wedge E_{i_s}$. Il ricorso alla formula di inclusione-esclusione pare perciò, qui, praticamente obbligato.

4.3 Approccio alla valutazione qualitativa

È abbastanza naturale immaginare che una valutazione di probabilità possa essere preceduta da una fase di valutazione di natura *qualitativa*. Immaginare cioè che chi è chiamato a formulare una valutazione di probabilità su un insieme di eventi, in una prima analisi delle sue aspettative sulla loro possibilità di essere veri sia portato a graduarli in termini qualitativi. A dire cioè quali eventi ritiene più *attendibile* che siano veri e quali meno. Lo farà anzi tanto più volentieri, quanto più complesso è il problema che deve (o vuole) esaminare. Solo in un secondo momento, a seguito di una più o meno attenta e minuziosa indagine, potrà arrivare a *quantificare* le sue propensioni esprimendo con un numero il suo grado di fiducia (la probabilità) sugli eventi in esame. Ciò in effetti è già stato fatto in occasione della valutazione dei modelli simmetrici (*Esempi* 4.1.1), ove l'*ipotesi di simmetria* altro non è che un giudizio di *equiattendibilità* (giudizio qualitativo) che si traduce in *equiprobabilità* (giudizio quantitativo) se, com'è naturale fare, si conviene che eventi **ugualmente attendibili** debbano avere **uguale probabilità**.

Un giudizio *quantitativo* del grado di fiducia sugli eventi di un insieme $\mathscr{A}$ – una probabilità – si esprime assegnando a ogni evento un numero compreso tra 0 e 1; introducendo cioè un'*applicazione* di $\mathscr{A}$ in $[0,1]$. Un giudizio di natura *qualitativa* prevede invece che gli eventi di $\mathscr{A}$ vengano messi a confronto *a due a due* al fine di stabilire se giudicarli *ugualmente attendibili* o uno *più attendibile* dell'altro. Poiché l'obiettivo dei due tipi di giudizio è il medesimo – graduare l'incertezza –, è del tutto naturale richiedere che le rispettive proprietà non siano tra di loro contraddittorie. Per evitare che ciò possa accadere bisogna rendere confrontabili i due metodi di valutazione. Osserviamo per questo che assegnata una probabilità su un insieme di eventi $\mathscr{A}$ – mediante un'applicazione –, si può introdurre in $\mathscr{A}$ la relazione binaria *non meno probabile*, dicendo che E_1 è *non meno probabile* di E_2 se e solo se riesce $P(E_1) \geq P(E_2)$. Si può poi cercare di esprimere le proprietà della probabilità P per mezzo della relazione *non meno probabile* – quindi tramite la relazione $\geq$ dei numeri reali – e chiedere poi che le valutazioni qualitative siano definite da relazioni che soddisfino proprietà analoghe. Si perviene così alla prossima definizione (B. de Finetti [2]). La rispondenza delle proprietà in essa elencate con quelle della probabilità P – della corrispondente relazione *non meno probabile* – è messa in luce nel *Commento* che segue la definizione.

4.3.1 Definizione. Valutazione qualitativa di probabilità.

Sia $\geq$**att** *(non meno attendibile di) una relazione binaria tra gli eventi di un qualche insieme. Diremo che* $\geq$att *è una* ***valutazione qualitativa di probabilità*** *se soddisfa le seguenti proprietà:*

a) *dati due eventi E_1, E_2 riesce o $E_1 \geq$att E_2 o $E_2 \geq$att E_1 (due eventi sono sempre confrontabili);*

b) *la relazione* $\geq$att *è riflessiva e transitiva;*

c) *se $E_1 \geq$att E_2 e $E_2 \geq$att E_1, si dice che E_1 e E_2 sono ugualmente attendibili e si scrive $E_1 =$att E_2;*
se $E_1 \geq$att E_2 e non $E_2 \geq$att E_1, si dice che E_1 è più attendibile di E_2 e si scrive $E_1 >$att E_2;

d) *se E è incompatibile sia con E_1 sia con E_2 $(E \wedge E_1 = E \wedge E_2 = \phi)$, allora:*

$$E_1 \geq_{\text{att}} E_2 \quad \textit{se e solo se} \quad E_1 \vee E \geq_{\text{att}} E_2 \vee E;$$

e) *l'evento certo (impossibile) è più (meno) attendibile di ogni altro evento, ovvero se E è possibile riesce $\Omega >$att $E >$att ϕ.*

NOTA.

Nella definizione è lasciato indeterminato l'insieme di definizione della relazione $\geq$att. Va sottinteso che quando tale insieme è noto, le singole proprietà vanno prese in considerazione solo se tutti gli eventi che in esse compaiono appartengono a tale insieme.

COMPLEMENTO.

Usando le proprietà *c)* e *d)* si possono dedurre proprietà analoghe alla *d)* anche per le due relazioni >att e =att. Sia infatti $E \wedge E_1 = E \wedge E_2 = \phi$.

Se $E_1 >$att E_2, allora per la *c)* $E_1 \geq$att E_2 e non $E_2 \geq$att E_1, da cui per la *d)* segue $E_1 \vee E \geq$att $E_2 \vee E$. Ora, se fosse $E_2 \vee E \geq$att $E_1 \vee E$, sempre per la *d)* si avrebbe che $E_2 \geq$att E_1 contro l'ipotesi (assurdo!). Si ha perciò *non* $E_2 \vee E \geq$att $E_1 \vee E$, e quindi $E_1 \vee E >$att $E_2 \vee E$, cioè la tesi. Il viceversa si prova in modo analogo.

La dimostrazione dell'estensione della *d)* alla relazione $=_{\text{att}}$ è analoga. ■

4.3.2 Commento.

Nella *Definizione* 4.3.1 le proprietà *a)* e *b)* dicono che la relazione $\geq_{\text{att}}$ è – nel suo insieme di definizione – un *preordine debole totale*, come lo è la relazione $\geq$ negli insiemi numerici.

La *c)* definisce le relazioni *ugualmente attendibile* e *più attendibile* ricalcando strettamente le analoghe proprietà della relazione $\geq$.

La *d)* interpreta in termini qualitativi la proprietà additiva delle probabilità quantitative (*Teorema* 4.2.2). Nelle ipotesi della definizione citata sugli eventi E, E_1, E_2 si ha infatti: $P(E \vee E_1) = P(E) + P(E_1)$ e $P(E \vee E_2) = P(E) + P(E_2)$, da cui segue $P(E \vee E_1) \geq P(E \vee E_2)$ se e solo se $P(E_1) \geq P(E_2)$ e per analogia si scrive la *d)*.

La *e)* infine risponde a un'esigenza di natura logica che abbiamo già avuto modo di segnalare in *Nota* 7 a piè di pagina: quella di considerare un evento E_2 *più attendibile* di E_1 se E_1 *implica* E_2 e *non viceversa;* quando cioè è *logicamente* provato che E_2 ha *più possibilità* di essere vero di E_1 (E_2 può essere vero senza che lo sia E_1, ma non viceversa). Ciò è detto esplicitamente dalla *e)* per i casi particolari in cui si confrontano Ω e ϕ – i due eventi di valore noto – tra di loro o con gli eventi possibili. È deducibile da essa e dalla *d)* nel caso generale, come mostra la prossima proposizione.

PROPOSIZIONE.

Siano E_1, E_2 *eventi,* $E_1 \Rightarrow E_2$, $E_1 \neq E_2$. *Allora* $E_2 >_{\text{att}} E_1$.

DIMOSTRAZIONE.

Poiché $E_1 \Rightarrow E_2$ riesce $E_1 \wedge E_2 = E_1$. Segue allora $E_2 = (E_1 \wedge E_2) \vee (\bar{E}_1 \wedge E_2) = E_1 \vee (\bar{E}_1 \wedge E_2)$, con $\bar{E}_1 \wedge E_2 \neq \phi$, perché $E_1 \neq E_2$ (3.3.2*c*). Allora per la *e)* riesce $\bar{E}_1 \wedge E_2 >_{\text{att}} \phi$. Inoltre, poiché E_1 è incompatibile con $\bar{E}_1 \wedge E_2$ e con ϕ, usando la *d)* si ricava $E_2 = E_1 \vee (\bar{E}_1 \wedge E_2) >_{\text{att}} E_1$. ■

In accordo con il discorso iniziale di questo numero, supponiamo ora che sia data una valutazione qualitativa $\geq_{\text{att}}$ su un insieme $\mathscr{A}$ e che si desideri ottenere a partire da essa una probabilità P che concordi con l'ordinamento degli eventi in termini di attendibilità. Se Ω e ϕ appartengono a $\mathscr{A}$, si deve intanto porre $P(\Omega) = 1$ e $P(\phi) = 0$ (*Teorema* 4.2.2*a*). Osserviamo poi che, in quanto preordine totale, la relazione $\geq_{\text{att}}$ *ripartisce* l'insieme $\mathscr{A}$ in classi di eventi ugualmente attendibili e le *ordina* in ragione dell'attendibilità riposta sul verificarsi dei loro eventi. Agli eventi di ogni classe dovrà allora essere attribuita la *medesima probabilità*, che possiamo perciò identificare anche come probabilità della rispettiva classe. Le probabilità delle classi devono essere poi scelte in modo che il loro ordinamento – in termini di probabilità – *concordi* con quello determinato dalla relazione $\geq_{\text{att}}$. Cosa che però non è sempre possibile fare in **senso forte**, richiedendo cioè che se $E_1 >_{\text{att}} E_2$, sia $P(E_1) > P(E_2)$. A causa della proprietà *e)*, infatti, in generale le valutazioni qualitative sono più "raffinate" di quelle quantitative. Ad esempio, mentre $\Omega >_{\text{att}} E$ per ogni E possibile, accanto a $P(\Omega) > P(E)$ è consentito anche porre $P(\Omega) = P(E)$ (4.2.1 *Nota*). Più in generale, se $E_1 \neq E_2$ e $E_1 \Rightarrow E_2$, riesce $E_2 >_{\text{att}} E_1$ (4.3.2 *Proposizione*), mentre sarà $P(E_1) = P(E_2)$ se $\bar{E}_1 \wedge E_2$ è giudicato di probabilità nulla, come si può fare se $\bar{E}_1 \wedge E_2 \neq \Omega$ (4.2.1 *Nota*).

Vedremo più avanti che se la descrizione è fatta con una partizione infinita, gli eventi elementari di probabilità positiva sono al più un numerabile (*Proposizione* 5.1.1). Segue allora che se la partizione ha cardinalità superiore al numerabile, esistono infiniti eventi elementari di probabilità nulla. Indicati con ω_1 e ω_2 due di essi, si ha $\omega_1 \vee \omega_2 >_{\text{att}} \omega_1$, mentre $P(\omega_1 \vee \omega_2) = P(\omega_1) + P(\omega_2) = 0 = P(\omega_1)$. Questo mostra che in ambiente infinito più che numerabile non è *mai* possibile introdurre probabilità quantitative che mantengano in *senso forte* l'ordinamento associato a una valutazione qualitativa. In tal caso ci si deve allora per forza accontentare di realizzare la condizione in **senso debole**. Di richiedere cioè che sia $P(E_1) \geq P(E_2)$ se $E_1 >_{\text{att}} E_2$. Il problema è ripreso e approfondito in ambiente finito nel prossimo numero.

4.4 Quozienti di probabilità e funzioni peso

In ambiente finito, quando la descrizione dell'incertezza è fatta mediante partizioni, il modo più semplice di esprimere una valutazione di probabilità in modo completo – su tutti gli eventi logicamente dipendenti dalle partizioni – è quello di distribuire la probabilità (massa) 1 a "pezzetti" sugli eventi elementari. In virtù della *Definizione* 4.2.1, la valutazione rimane allora determinata per additività per tutti gli eventi logicamente dipendenti (i sottoinsiemi) delle partizioni. Chi sceglie di valutare in questo modo, ma desidera esprimersi prima in termini qualitativi, è sufficiente che lo faccia limitandosi a confrontare in termini di attendibilità gli eventi elementari. Limitandosi cioè ad introdurre una valutazione qualitativa sulla partizione scelta per descrivere l'incertezza. La questione è trattata in dettaglio nel prossimo numero.

4.4.1 Quozienti di probabilità.

Sia $\mathbb{P} = \{\omega_1, \dots, \omega_n\}$ una partizione e $\geq_{att}$ una valutazione qualitativa definita su di essa. Tenendo conto della discussione fatta in *Commento* 4.3.2, di 4.3.2 *Proposizione* e della *Definizione* 4.2.1, per determinare su $\mathbb{P}$ una probabilità quantitativa che conservi l'ordine di attendibilità di quella qualitativa, si dovrà porre $P(\omega_i) \geq 0$ per ogni i, $P(\omega_h) \geq P(\omega_i)$ se $\omega_h >_{att} \omega_i$, $P(\omega_h) = P(\omega_i)$ se $\omega_h =_{att} \omega_i$ e $\sum_{i=1}^{n} P(\omega_i) = 1$. Poiché gli addendi di questa sommatoria sono non negativi, uno di essi deve essere positivo. Almeno quello o quelli di attendibilità massima, mentre gli altri, in accordo con la *Definizione* 4.2.1 e con l'ordine di attendibilità, possono essere anche tutti di probabilità nulla (4.2.1 *Nota*). Per concordare con l'ordinamento determinato dalla valutazione qualitativa non sono richieste altre condizioni. Salvo che nel caso particolare dell'equiattendibilità (giudizio di simmetria) – in cui si deve porre $P(\omega_i) = 1/n$ per ogni i –, esse non bastano per determinare la probabilità P su $\{\omega_1, \dots, \omega_n\}$, cioè le probabilità $P(\omega_i)$. Queste andranno scelte dopo un supplemento di analisi, rispettando le condizioni di uguaglianza e disuguaglianza viste sopra. Cosa che si può fare esprimendosi, anziché sulle probabilità direttamente, sui loro rapporti, come andiamo ad illustrare qui di seguito.

Il problema che si pone quando un soggetto decide di confrontare due eventi per stabilire quale sia a suo giudizio il rapporto delle loro probabilità, è analogo a quello che si incontra quando si hanno da misurare delle grandezze. In generale, misurare significa mettere a confronto una grandezza con un'altra ad essa omogenea – scelta come unità di misura – in vista di stabilire se sono uguali o *se* e di *quanto* la prima supera o è inferiore alla seconda. Di dire cioè qual è il loro *rapporto*. Questo metodo del confronto a coppie per dare una misura a una grandezza, può essere utilmente adoperato anche per analizzare le proprie opinioni e sensazioni al fine di far emergere il grado di fiducia. Così, per giungere a una valutazione di P su $\{\omega_1, \dots, \omega_n\}$, si possono mettere a confronto le coppie di eventi ω_i, ω_h giudicati di probabilità positiva – gli eventuali di probabilità nulla sono già valutati –, al fine di stabilire qual è nel nostro giudizio il rapporto tra le probabilità $P(\omega_i)$ e $P(\omega_h)$, ovvero il loro

quoziente di probabilità $P(\omega_i)/P(\omega_h)$[23].
Osserviamo prima che le disuguaglianze e uguaglianze viste sopra prevedono in linea di principio che si facciano $(n^2-n)/2$ confronti – escludiamo di confrontare gli eventi ω_i con se stessi e mettiamo a confronto le coppie ω_i e ω_h in un solo verso –. In pratica, però, il numero di confronti può essere ridotto a $n-1$ e – allo stesso tempo – la procedura di valutazione può essere notevolmente semplificata e resa più precisa se mettiamo le classi di alternative equiattendibili in ordine di attendibilità non-crescente ed elenchiamo le alternative di ogni classe consecutivamente in un ordine qualsiasi. Otteniamo così una permutazione degli ω_i che – riassegnando eventualmente gli indici – possiamo sempre elencare nell'ordine naturale. Si ha allora $\omega_1 \geq_{\text{att}} \omega_2 \geq_{\text{att}} \ldots \geq_{\text{att}} \omega_n$, con ω_1 alternativa di *attendibilità massima,* e quindi di *probabilità massima* e *positiva*. Posto $q_i = P(\omega_i)/P(\omega_1)$, riesce $q_1 = 1$. Per conservare l'ordinamento di attendibilità deve essere $q_i \geq q_{i+1}$ – la sequenza (q_i) deve essere monotona non-crescente – e $q_i = q_{i+1}$ se $\omega_i =_{\text{att}} \omega_{i+1}$ – assegnando così un quoziente comune alle alternative di ogni classe di equiattendibilità –. Una volta valutati questi quozienti, la P resta univocamente determinata (e restano così determinati anche gli altri quozienti). Sussiste infatti in proposito la seguente proposizione.

PROPOSIZIONE.

Siano $\mathbb{P} = \{\omega_1, \ldots, \omega_n\}$ *una partizione di* Ω *in alternative* (3.4.1*Nota e terminologia*) *e* $\geq_{\text{att}}$ *una valutazione qualitativa su* $\mathbb{P}$ *tale che* $\omega_1 \geq_{\text{att}} \omega_2 \geq_{\text{att}} \ldots \geq_{\text{att}} \omega_n$. *Scelta allora una sequenza* $1, q_2, \ldots, q_n$ *non negativa, monotona non crescente, tale che* $q_i = q_{i+1}$ *se* $\omega_i =_{\text{att}} \omega_{i+1}$ *e posto* $q_i = P(\omega_i)/P(\omega_1)$, $i = 1, \ldots, n$, *resta determinata una probabilità* P *su* $\mathbb{P}$ *per la quale riesce:*

$$P(\omega_i) = q_i / \textstyle\sum_{h=1}^{n} q_h . \qquad (2)$$

Le condizioni $q_i = q_{i+1}$ *se* $\omega_i =_{\text{att}} \omega_{i+1}$ *e quelle di non negatività e monotonia della sequenza* (q_i) *sono necessarie e sufficienti affinché la probabilità* P *sia concorde in senso debole con la valutazione qualitativa* $\geq_{\text{att}}$.

---

23 Si tratta di una procedura largamente praticata nel mondo delle scommesse, anche se le valutazioni che vengono ivi espresse si discostano del grado di fiducia, perché gli allibratori sono disposti a giocare solo ponendosi in condizioni di vantaggio (a loro giudizio). L'argomento viene approfondito nel § 9.1. Si riveda anche in proposito il cenno fatto al termine dell'*Esempio* 3 del n° 2.1.1, quello che riguarda la partita di calcio.

DIMOSTRAZIONE.

Si ha intanto $P(\omega_i) = P(\omega_1)\, q_i$ per ogni $i = 1, \ldots, n$. Inoltre, poiché deve essere $\sum_{h=1}^{n} P(\omega_h) = 1$ (condizione di *normalizzazione*), si ricava $\sum_{h=1}^{n} q_h P(\omega_1) = P(\omega_1) \sum_{h=1}^{n} q_h = 1$, da cui si ottiene la (2). Ciò prova che P è una probabilità secondo le condizioni richieste dalla *Definizione* 4.2.1.
La necessità per P delle condizioni $q_i = q_{i+1}$ se $\omega_i =_{\text{att}} \omega_{i+1}$ e quelle di non negatività e monotonia delle q_i è ovvia.
La sufficienza segue immediatamente dalla (2). Se $\omega_h =_{\text{att}} \omega_i$, riesce infatti $q_h = q_i$ per costruzione e quindi $P(\omega_h) = P(\omega_i)$, come deve essere, atteso che abbiamo convenuto che alternative ugualmente attendibili debbano essere di probabilità uguale. Se invece $\omega_h >_{\text{att}} \omega_i$, sempre per costruzione riesce $q_h \geq q_i$, da cui segue $P(\omega_h) \geq P(\omega_i)$, in accordo – in senso debole – con l'ordinamento di attendibilità. ■

OSSERVAZIONE.

Per valutare i rapporti q_i abbiamo scelto di elencare le alternative in ordine di attendibilità non-crescente e di confrontarle tutte con ω_1, attribuendo così a questa alternativa il ruolo "privilegiato" di *unità di misura*. Abbiamo scelto cioè di dare misura 1 a ω_1, riservandoci di dare alle altre alternative, in subordine e per confronto, misure (rapporti) $q_i \leq 1$. Ciò è stato fatto per due motivi. Il primo è che in questo modo basta porre $q_1 = 1$ e scegliere una sequenza di rapporti $q_1, \ldots, q_n$ monotona non-crescente e rispettare la condizione $q_i = q_{i+1}$ se $\omega_i =_{\text{att}} \omega_{i+1}$ per trovarsi concordi con l'ordinamento di attendibilità. Il secondo è dovuto invece al fatto che ω_1 – alternativa di attendibilità massima – è sicuramente di probabilità positiva, condizione indispensabile per dare significato ai rapporti $P(\omega_i)/P(\omega_1)$. Condizione, questa, che sussiste però anche se si affida il ruolo di unità di misura a una qualunque delle alternative di probabilità positiva. Nelle ipotesi precedenti, se ω_r è una di esse, scegliendola come unità di misura si genera una sequenza di quozienti $q_i = P(\omega_i)/P(\omega_r)$, che per concordare con l'ordinamento di attendibilità dovrà essere monotona non-crescente e tale che $q_i = q_{i+1}$ se $\omega_i =_{\text{att}} \omega_{i+1}$ – come prima –, ma con $q_r = 1$ e quindi $q_i \geq 1$ se $i < h$, $q_i \leq 1$ se $i > h$. Normalizzando si trovano per le $P(\omega_i)$ ancora le (2).

COMPLEMENTO.

In alternativa alle procedure precedenti, in cui si tiene fisso uno dei termini del confronto, si può pensare ad altre che non soddisfano questa condizione. Ad esempio, in certe circostanze può essere che le alternative appaiano tanto più facili da confrontare quanto più prossime siano giudicate le loro attendibilità.

Elencate le alternative in ordine di attendibilità non-crescente – come prima – la procedura di valutazione più fedele a questo modo di sentire è allora quella di mettere a confronto le coppie di alternative consecutive. Le eventuali alternative giudicate di probabilità nulla si troveranno allora elencate in coda. Per le altre, posto che siano s, si ha $0 < P(\omega_{i+1}) / P(\omega_i) = r_i \leq 1$, $r_i = 1$ se $\omega_i =_{\text{att}} \omega_{i+1}$, $i = 1, \ldots, s-1$, da cui si ottengono $s-1$ equazioni lineari linearmente indipendenti del tipo $P(\omega_{i+1}) = r_i P(\omega_i)$, che assieme alla condizione di normalizzazione consentono di completare il calcolo con la determinazione delle probabilità delle s alternative di probabilità positiva.
Anche in questo caso, come in quello dei procedimenti con termine di confronto fisso, le condizioni poste per le r_i – allora per le q_i – sono necessarie – ovvio – e sufficienti affinché la corrispondente valutazione di probabilità sia concorde con l'ordinamento di attendibilità.

La sufficienza si giustifica cominciando con l'osservare che, come abbiamo visto, le alternative di probabilità positiva precedono quelle di probabilità nulla e che per queste ultime il rispetto dell'ordine di attendibilità è ovvio. Per provare che ciò accade anche per quelle di probabilità positiva, confrontiamo ω_h con ω_j per $1 \leq h < j \leq s$. Si ha:

$$P(\omega_j) = r_{j-1} P(\omega_{j-1}) = r_{j-1} r_{j-2} P(\omega_{j-2}) = r_{j-1} \ldots r_h P(\omega_h), \tag{3}$$

da cui si ricava $P(\omega_h) \leq P(\omega_j)$, perché i rapporti r_i sono tutti non superiori a 1, e in particolare è $P(\omega_h) = P(\omega_j)$ se $\omega_h =_{\text{att}} \omega_j$, perché allora $r_h = \ldots = r_{j-1} = 1$ per costruzione. ■

NOTA.

Se l'ambiente è finito, è data una valutazione qualitativa e siamo disposti a giudicare tutte le alternative di probabilità positiva, esistono allora probabilità che concordano in *senso forte* con l'ordine di attendibilità. Ciò anzi accade anche se si attribuisce probabilità nulla alle alternative di attendibilità minima (e solo a loro). Nelle notazioni e ipotesi precedenti, in entrambi i tipi di procedimento ivi considerati – quello iniziale e quello del precedente complemento –, ciò si realizza mantenendo le rispettive condizioni di uguaglianza $q_i = q_{i+1}$ e $r_i = 1$, se $\omega_i =_{\text{att}} \omega_{i+1}$, e rimpiazzando altrimenti le disuguaglianze *deboli* $q_i \geq q_{i+1}$ e $r_i \leq 1$ con le disuguaglianze *forti* $q_i > q_{i+1}$ e $r_i < 1$.
Il problema di trovare valutazioni quantitative che concordano in senso forte con l'ordinamento di attendibilità di una valutazione qualitativa in presenza di alternative di probabilità nulla e di attendibilità diversa dalla minima, sarà ripreso nel § 17.2, Vol. II e risolto positivamente in ambiente finito usando *probabilità condizionate*.

4.4.2 Funzioni peso.

Riprendiamo in esame il primo dei due procedimenti di valutazione *indiretta* della probabilità in ambiente finito visti nel numero appena concluso. In corrispondenza alla partizione $\{\omega_1, \dots, \omega_n\}$ si è scelta un'alternativa ω_r di probabilità positiva come termine fisso per il confronto dei rapporti $q_i = P(\omega_i)/P(\omega_r)$, $i = 1, \dots, n$, che in forma moltiplicativa si scrivono $P(\omega_i) = q_i P(\omega_r)$. Si legge allora che i quozienti di probabilità $q_1, \dots, q_n$ sono numeri non negativi, uno almeno positivo – almeno $q_r = 1$ –, proporzionali all'n-pla delle probabilità $P(\omega_1), \dots, P(\omega_n)$. Usando la (2) del numero precedente si vede allora che essi sono in grado di misurare attraverso i rapporti q_i/q_j i quozienti di probabilità di ogni coppia (ω_i, ω_j), per cui riesce $P(\omega_j) > 0$. Risultato che si può ovviamente ottenere anche usando al posto dell'n-pla $(q_1, \dots, q_n)$ una qualunque altra n-pla di numeri non negativi – basterebbe dello stesso segno o nulli – ad essa proporzionale. In relazione a ciò si dà allora la seguente definizione.

Definizione. Funzione peso.

Sia $\mathbb{P} = \{\omega_1, \dots, \omega_n\}$ una partizione in eventi elementari. Chiameremo ***funzione peso*** *ogni applicazione π di dominio $\{\omega_1, \dots, \omega_n\}$, non negativa, non identicamente nulla e proporzionale secondo un fattore positivo alla probabilità P di pari dominio. I numeri $\pi(\omega_1), \dots, \pi(\omega_n)$ si dicono* ***pesi*** *di $\omega_1, \dots, \omega_n$, rispettivamente.*

Sussiste allora la seguente proposizione.

Proposizione.

Sia $\{\omega_1, \dots, \omega_n\}$ una partizione e π una sua funzione peso. Posto $\pi(\omega_i) = \pi_i$, allora si ha $P(\omega_i)/P(\omega_j) = \pi_i/\pi_j$, $1 \le i, j \le n$ e $\pi_j > 0$, e riesce:

$$P(\omega_h) = \pi_h \Big/ \textstyle\sum_{i=1}^{n} \pi_i, \qquad h = 1, \dots, n. \tag{4}$$

Le funzioni peso si ripartiscono in classi di funzioni proporzionali. Le funzioni di ogni classe determinano per ***normalizzazione*** *– mediante l'operazione* (4) *– la medesima distribuzione di probabilità, che a sua volta è una funzione peso, l'unica nella classe con pesi di somma* 1. *La corrispondenza tra le distribuzioni di probabilità su* $\{\omega_1, \dots, \omega_n\}$ *e le classi di funzioni peso ad esse proporzionali è perciò biunivoca.*

DIMOSTRAZIONE.

Per definizione di funzione peso riesce $P(\omega_i) = k\,\pi_i$ per ogni i e per definizione di probabilità $\sum_{i=1}^{n} P(\omega_i) = 1$. Allora si ha $\sum_{i=1}^{n} P(\omega_i) = k \sum_{i=1}^{n} \pi_i = 1$ e quindi $k = 1/\sum_{i=1}^{n} \pi_i$. Sostituendo questo valore di k in $P(\omega_h) = k\,\pi_h$, che vale per ogni h, si ottiene la (4). Per la seconda parte della tesi osserviamo che usando nella (4) i pesi $c\,\pi_1, \dots, c\,\pi_n$ al posto dei $\pi_1, \dots, \pi_n$, $c > 0$, la costante c si semplifica e si ottiene perciò lo stesso risultato. Il resto è ovvio. ■

COMPLEMENTO. *Proprietà della funzione peso.*

La funzione peso è sostanzialmente una probabilità *non normalizzata,* nel senso che il suo prolungamento per additività sugli eventi somma di alternative soddisfa le proprietà della probabilità ad essa associata, con la sola eccezione della condizione di normalizzazione, che richiede alla somma delle probabilità delle alternative di essere 1, mentre alla somma dei pesi è richiesto soltanto di essere positiva. Il risultato è facile conseguenza della proporzionalità tra la probabilità e le funzioni peso corrispondenti, ma merita tuttavia di essere visto con qualche dettaglio.

Cominciamo per questo col prolungare la funzione peso π sugli eventi somma di alternative. Con notazione analoga a quella usata per la probabilità nella *Definizione* 4.2.1, poniamo:

$$E = \vee \mathcal{E},\ \mathcal{E} \subset \{\omega_i, \dots, \omega_n\},\ \pi(E) = \textstyle\sum_{\omega \in \mathcal{E}} \pi(\omega).$$

Posto ancora $s_\pi = \sum_{i=1}^{n} \pi(\omega_i)$, per la (4) abbiamo $\pi(\omega) = s_\pi P(\omega)$ per ogni $\omega \in \{\omega_1, \dots, \omega_n\}$. Segue allora:

$$\pi(E) = \textstyle\sum_{\omega \in \mathcal{E}} \pi(\omega) = \sum_{\omega \in \mathcal{E}} s_\pi P(\omega) = s_\pi P(E).$$

Da qui si vede intanto che $P(E) = \pi(E)/s_\pi$, cioè che la (4) si estende agli eventi somma di alternative.

È facile poi provare che π soddisfa proprietà analoghe a quelle della probabilità. Basta per questo moltiplicare per s_π membro a membro le proprietà del *Teorema* 4.2.2 e ricordare che $s_\pi P(E) = \pi(E)$. Le prime quattro allora diventano:

a') $\pi(\Omega) = s_\pi$, $\pi(\phi) = 0$ (segue da $s_\pi P(\Omega) = s_\pi$ e $s_\pi P(\phi) = 0$),

b') $0 \leq \pi(E) \leq s_\pi$ (segue da $0 \leq s_\pi P(E) \leq s_\pi$),

c') *Se* $E_1 \wedge E_2 = \phi$, *allora* $\pi(E_1 \vee E_2) = \pi(E_1) + \pi(E_2)$
(segue da $s_\pi P(E_1 \vee E_2) = s_\pi P(E_1) + s_\pi P(E_2)$),

d') $\pi(\bar{E}) = s_\pi - \pi(E)$ (segue da $s_\pi P(\bar{E}) = s_\pi - s_\pi P(E)$.

L'elenco delle altre quattro proprietà del *Teorema* 4.2.2 può essere ormai facilmente completato dal lettore.

4.4.3 Nota. *Aspetti operativi.*

Come messo in evidenza dalla formula (4) di 4.4.2 *Proposizione*, le funzioni peso esprimono in modo *indiretto* una valutazione di probabilità su una partizione finita $\{\omega_1, \ldots, \omega_n\}$, la medesima se le funzioni peso sono proporzionali. Data perciò una funzione peso – una n-pla *arbitraria* di numeri *non negativi*, di cui *uno almeno positivo* (4.4.2 *Definizione*) –, la determinazione della corrispondente distribuzione di probabilità può essere fatta utilizzando la funzione peso originale o altra ad essa proporzionale, che ai fini pratici potrà essere scelta in modo di semplificare lo sviluppo dei calcoli previsti nella formula (4). L'uso della formula (4) porta alla determinazione preliminare della distribuzione di probabilità sulle alternative della partizione $\{\omega_1, \ldots, \omega_n\}$ e, a seguire, al calcolo delle probabilità degli eventi logicamente dipendenti dalla partizione che interessano – di somme logiche di alternative –.

Un secondo accorgimento da tenere presente per lo sviluppo dei calcoli è suggerito dalle proprietà della funzione peso viste in 4.4.2 *Complemento* – analoghe a quelle della probabilità – che consentono di interpretare le funzioni peso come misure di probabilità *non* normalizzate. Dal punto di vista operativo ciò significa che il calcolo delle probabilità degli eventi logicamente dipendenti dalla partizione può essere effettuato senza bisogno di determinare preliminarmente la distribuzione di probabilità associata a una data funzione peso. Basta infatti operare con i pesi come fossero probabilità e normalizzare i risultati a posteriori servendosi dell'uguaglianza $P(E) = \pi(E)/s_\pi$. Questa modalità di calcolo sarà conveniente se gli eventi da valutare non sono troppo numerosi, e potrà comunque essere adottata se non si ha interesse di determinare la distribuzione su $\{\omega_1, \ldots, \omega_n\}$.

Circa la scelta di una funzione peso – della valutazione –, come ampiamente discusso nel n° 4.4.1, può accadere che la valutazione quantitativa venga espressa mediante quozienti di probabilità, a partire magari da una preliminare valutazione qualitativa. Nella parte introduttiva del n° 4.4.2 abbiamo osservato che se i quozienti di probabilità sono quelli ottenuti confrontando le alternative con una fissa di esse, essi risultano allora proporzionali alle rispettive probabilità. La loro sequenza costituisce perciò una funzione peso. Se i quozienti sono numeri razionali – come accade spesso in pratica –, in accordo con quanto osservato all'inizio di questa *Nota,* li potremo moltiplicare per un conveniente fattore comune – ad esempio per il loro minimo comune multiplo – ottenendo così una funzione peso a valori naturali, sicuramente più semplice di quella originale per il calcolo delle corrispondenti probabilità. Si veda in proposito 4.4.4 *Esempio* 4.

In 4.4.1 *Complemento* abbiamo segnalato che, oltre quella ora ricordata, si può pensare anche di seguire altre regole per la scelta dei quozienti di probabilità. Abbiamo considerato con dettaglio, in proposito, quella che prevede che le alternative vengano preliminarmente elencate in ordine di attendibilità non-decrescente, come nella regola precedente, ma valutate poi usando i quozienti di probabilità di coppie di alternative che in quell'ordinamento sono consecutive. Questa regola di confronto sequenziale può essere usata anche in assenza di un ordinamento qualitativo preliminare degli ω_i. È sufficiente per questo mettere da parte le eventuali alternative di probabilità nulla – che non hanno bisogno di essere ulteriormente valutate – e tenere conto che in assenza di un ordinamento qualitativo i quozienti $r_i = P(\omega_{i+1})/P(\omega_i)$ relativi alle alternative di probabilità positiva sono soggetti alla sola condizione di essere positivi – non si ha da rispettare la monotonia –.
Importa ancora segnalare che in questi casi – in cui non è prevista un'alternativa fissa come termine di confronto – la sequenza (r_i) non è una funzione peso. Per la (3) riesce infatti $P(\omega_j) = r_{j-1} \dots r_1 P(\omega_1)$. La funzione peso è allora data da $\pi_1 = q_1 = 1$, $\pi_i = 0$ se $P(\omega_i) = 0$, $\pi_j = r_1 \dots r_{j-1}$ se $P(\omega_j) > 0$ (è sottinteso che le alternative di probabilità positiva precedono nell'elenco quelle di probabilità nulla). Si vedano in proposito 4.4.4 *Esempio* 5, in cui gli ω_i sono elencati in ordine di attendibilità, e 4.4.4 *Esempio* 6, ove la regola di valutazione sequenziale è accolta solo in parte.

4.4.4 Esempi.

Come abbiamo visto, la funzione peso è uno strumento a disposizione del soggetto per indagare sul suo grado di fiducia. Gli esempi che seguono illustrano l'uso di questo strumento. Spesso si arriva alla determinazione di una funzione peso a seguito di confronti tra gli ω_i di una partizione (*Esempi* 4 , 5) o, più in generale, facendo intervenire anche somme di ω_i (*Esempio* 6).

1 In termini di funzione peso, il giudizio di simmetria sui casi possibili di un insieme $\{\omega_1, \ldots, \omega_n\}$ è espresso dalla classe delle funzioni costanti positive.

2 *Con riferimento a* 4.1.1 *Esempio* 12 *(urna con una pallina bianca, 2 nere, 3 rosse e 4 verdi) si estragga una pallina e in ipotesi di simmetria, si calcoli la probabilità che essa sia di un dato colore.*

Numerate le palline da 1 a 10 e indicato con ω_h l'evento *viene estratta la pallina numero h,* possiamo esprimere il giudizio di simmetria introducendo su $\{\omega_1, \ldots, \omega_{10}\}$ la funzione peso 1. Indicati con B, N, R, V gli eventi *esce bianca*, *esce nera*, *esce rossa*, *esce verde*, i loro pesi (4.4.2 *Complemento*) sono allora dati dal numero delle palline del rispettivo colore. Pertanto si ha: $\pi(B) = 1$, $\pi(N) = 2$, $\pi(R) = 3$, $\pi(V) = 4$. Poiché $s_\pi = 10$, da qui si ricava:

$$P(B) = 0{,}1, \quad P(N) = 0{,}2, \quad P(R) = 0{,}3, \quad P(V) = 0{,}4.$$

Gli eventi B, N, R, V sono esaustivi e a due a due incompatibili. Sono perciò una partizione in una descrizione che considera indistinguibili le palline di uno stesso colore. Pertanto, la quaterna $(\pi(B), \pi(N), \pi(R), \pi(V))$ è una funzione peso su $\{B, N, R, V\}$. In virtù della (4), la probabilità dei quattro eventi si calcola allora dividendo i loro pesi per $\pi(B) + \pi(N) + \pi(R) + \pi(V)$. Si ottengono i valori calcolati in precedenza, perché $\pi(B) + \pi(N) + \pi(R) + \pi(V) = \pi(\omega_1) + \ldots + \pi(\omega_{10}) = s_\pi$.
Questo risultato, visto qui nell'esempio, vale in generale. Ovvero, se $\{\omega_1, \ldots, \omega_n\}$ è una partizione su cui è data una funzione peso π e $E_1, \ldots, E_s$ sono eventi somma di alternative esaustivi e a due a due incompatibili, allora prolungando π sulle somme di alternative – in particolare su $E_1, \ldots, E_s$ – si ha che $\pi(E_1), \ldots, \pi(E_s)$ è una funzione peso su $\{E_1, \ldots, E_s\}$ e riesce $\pi(E_1) + \ldots + \pi(E_s) = \pi(\omega_1) + \ldots + \pi(\omega_n)$. La facile verifica è lasciata al lettore per esercizio.

3 *Generalizzazione dell'Esempio 2.*

Si estragga una pallina da un'urna contenente n_1 palline di tipo 1, n_2 di tipo 2, ..., n_s di tipo s, $n = n_1 + \ldots + n_s$. Calcolare la probabilità che la pallina estratta sia di tipo h nell'ipotesi che le n palline abbiano la stessa probabilità di essere estratte.

Posto E_h = *la pallina estratta è di tipo h,* riesce $\pi(E_h) = n_h$, da cui segue $P(E_h) = n_h / n$.

4 *Con riferimento a una corsa di cavalli un allibratore dà come favoriti i cavalli A, B, C, D. Indicati con gli stessi simboli gli eventi «A vince», … , «D vince», e con E l'evento «vince un outsider», egli esprime la sua valutazione sull'insieme delle alternative $\{A, B, C, D, E\}$ mediante i seguenti quozienti di probabilità:*

$$\frac{P(B)}{P(A)} = 0{,}7, \quad \frac{P(C)}{P(A)} = 0{,}5, \quad \frac{P(D)}{P(A)} = 1{,}2, \quad \frac{P(E)}{P(A)} = 0{,}1.$$

Calcolare la distribuzione di probabilità su $\{A, B, C, D, E\}$.

Poiché l'alternativa A è il termine di confronto fisso per le altre alternative, i quozienti $P(A)/P(A) = 1, 0{,}7, 0{,}5, 1{,}2, 0{,}1$ costituiscono una funzione peso su $\{A, B, C, D, E\}$, che ai fini del calcolo conviene sostituire con quella di pesi $10, 7, 5, 12, 1$, ad essa equivalente. Normalizzando (dividendo i pesi per 35) si ottiene:

$$P(A) = 0{,}286, \; P(B) = 0{,}200, \; P(C) = 0{,}143, \; P(D) = 0{,}343, \; P(E) = 0{,}029.$$

5 *Le squadre A, B, C, D partecipano a un torneo di calcio all'italiana – ogni squadra incontra tutte le altre –. Un allibratore giudica le squadre favorite per la vittoria nell'ordine indicato e decide di esprimere la sua valutazione sugli eventi (alternative) «A vince», «B vince», «C vince», «D vince» mediante quozienti di probabilità. A tal fine egli ritiene di poter dare una valutazione più accurata mettendo a confronto coppie di squadre di forza il più possibile simile, ovvero coppie di squadre consecutive nel suo ordine di preferenza. In accordo con questo atteggiamento egli fornisce allora i seguenti quozienti di probabilità:*

$$\frac{P(B)}{P(A)} = 0{,}8, \quad \frac{P(C)}{P(B)} = 0{,}9, \quad \frac{P(D)}{P(C)} = 0{,}5.$$

Calcolare la distribuzione di probabilità su $\{A, B, C, D\}$.

Completati con $P(A)/P(A) = 1$, i quozienti $1, 0{,}8, 0{,}9, 0{,}5$ non costituiscono una funzione peso, perché ora non è fisso il termine di confronto. Essi ci permettono però di risalire a una funzione peso osservando che riesce: $P(B) = 0{,}8\, P(A)$, $P(C) = 0{,}9 \times 0{,}8\, P(A) = 0{,}72\, P(A)$, $P(D) = 0{,}5 \times 0{,}72\, P(A) = 0{,}36\, P(A)$, da cui si ricava

$$\frac{P(B)}{P(A)} = 0{,}8, \quad \frac{P(C)}{P(A)} = 0{,}72, \quad \frac{P(D)}{P(A)} = 0{,}36$$

e quindi i pesi $\pi(A) = 100$, $\pi(B) = 80$, $\pi(C) = 72$, $\pi(D) = 36$. Normalizzando (somma dei pesi 288) si ottiene:

$$P(A) = 0{,}347, \quad P(B) = 0{,}278, \quad P(C) = 0{,}250, \quad P(D) = 0{,}125.$$

6 *Con riferimento all'Esempio 4, un appassionato scommettitore dà come favoriti i cavalli A, B, C, D in questo ordine (che in parte contrasta con quello dell'allibratore dell'Esempio 4). Indicati anche qui con gli stessi simboli A, B, C, D, E gli eventi «A vince», ..., «D vince», «vince un outsider», questo scommettitore esprime la sua valutazione sull'insieme delle alternative $\{A, B, C, D, E\}$ mediante i seguenti quozienti di probabilità:*

$$\frac{P(B)}{P(A)} = \frac{3}{5}, \quad \frac{P(C)}{P(B)} = \frac{2}{3}, \quad \frac{P(D)}{P(C)} = \frac{5}{8}, \quad \frac{P(E)}{P(\bar{E})} = \frac{1}{18}.$$

Calcolare la distribuzione di probabilità su $\{A, B, C, D, E\}$.

Anche in questo esempio i quozienti che vengono forniti non costituiscono una funzione peso, ma si può costruirne una operando in modo analogo al precedente, tenendo anche conto che $\bar{E}$ è costituito dalle alternative A, B, C, D (*Teorema* 3.4.3) e che riesce perciò $P(\bar{E}) = P(A) + P(B) + P(C) + P(D)$. Allora si ha: $P(B) = 3/5\,P(A)$, $P(C) = 2/3 \times 3/5\,P(A) = 2/5\,P(A)$, $P(D) = 5/8 \times 2/5\,P(A) = 1/4\,P(A)$, $P(E) = 1/18 \times (1 + 3/5 + 2/5 + 1/4) \times P(A) = 1/8\,P(A)$, da cui si ricavano i quozienti di probabilità

$$\frac{P(B)}{P(A)} = \frac{3}{5}, \quad \frac{P(C)}{P(A)} = \frac{2}{5}, \quad \frac{P(D)}{P(A)} = \frac{1}{4}, \quad \frac{P(E)}{P(A)} = \frac{1}{8},$$

e moltiplicando per 40 si trovano i pesi di A, B, C, D, E espressi nell'ordine da $40, 24, 16, 10, 5$. Normalizzando si ricava:

$$P(A) = 0{,}421, \; P(B) = 0{,}253, \; P(C) = 0{,}168, \; P(D) = 0{,}105, \; P(E) = 0{,}053.$$