Prima lezione di probabilità soggettiva

Donato Michele Cifarelli

a cura[1] *di*

Pietro Muliere *e* Piercesare Secchi

1. Introduzione

Nelle pagine che seguono ci apprestiamo a fornire gli elementi per giungere a fissare il nesso sostanziale tra significato della probabilità e sua definizione; questa, oltre che ad essere in linea col significato che generalmente si attribuisce a frasi del tipo *la probabilità dell'evento A è p*, risulterà, al tempo stesso, adeguatamente precisa da permettere di poggiarvi una teoria matematica della probabilità. Il tentativo di conciliare queste due esigenze è stato momento di grande rilevanza negli studi sulla probabilità e la diversità delle proposte avanzate per risolvere il problema ha dato luogo ad un dibattito sui fondamenti che ancor oggi è tutt'altro che esaurito. In effetti, le proposte per la sistemazione teorica dei principi della probabilità vanno da quelle basate su considerazioni empiriche - in genere collegate ad un significato oggettivo della probabilità - che si sono rilevate inadeguate ad essere integrate in un teoria matematica coerente, a presentazioni di tipo assiomatico, volutamente prive di nessi interpretativi, tendenti ad ancorare gli aspetti matematici della probabilità ad impostazioni già affermate in altri settori.

Per quanto riguarda la trattazione che stiamo per intraprendere, ci atterremo al punto di vista *soggettivistico* di Bruno de Finetti. Infatti, il lavoro ha lo scopo di introdurre la nozione di probabilità chiarendone il significato sostanziale attraverso il concetto di valutazione coerente.

Dopo aver analizzato i vari problemi inerenti l'esistenza di una distribuzione di probabilità e la possibilità della sua definizione sui più consueti campi di eventi, considereremo i legami con la probabilità introdotta seguendo l'impostazione assiomatica di Kolmogorov.

[1]Questi appunti didattici, di introduzione alla probabilità soggettiva secondo la lezione di de Finetti, sono cresciuti nel tempo sul tavolo del Prof. Cifarelli. Anno dopo anno, molti dei suoi studenti, e noi tra questi, hanno letto, utilizzato o citato una delle loro molteplici versioni. A noi è toccato il compito di rendere disponibile questo materiale agli studenti futuri, raccogliendolo, organizzandolo e sottoponendolo, nella nuova veste, al severo scrutinio del suo autore. (P.Muliere e P.Secchi)

2. La condizione di coerenza

Per giungere a dare una definizione di probabilità, si può procedere cercando di spiegare il significato che è possibile attribuire alla frase "la probabilità dell'evento A è p".

Una prima accezione è quella che origina dalla possibilità di pensare A come ad un evento che può realizzarsi in k possibili casi incompatibili, su un totale di N tutti *egualmente plausibili*, e di porre $p = \frac{k}{N}$, cioè interpretare p come il rapporto tra il numero dei *casi favorevoli* all'evento A $[k]$ ed il *numero dei casi possibili* $[N]$.

Altro possibile significato della valutazione di p discende dalla facoltà di pensare all'evento A come immerso in una famiglia di N altri eventi $A_1, A_2, \cdots, A_N$ in qualche senso *omogenei*, con N abbastanza grande. Nel caso in cui k eventi A_i si sono verificati, allora p può pensarsi dato dalla frequenza relativa k/N.

Gli esempi seguenti sono utili ad illustrare, nell'ordine, i due schemi.

Esempio 2.1. I numeri 1,2,3 vengono scritti "a caso" una volta: ovvero i sei modi in cui i tre numeri possono essere scritti sono egualmente plausibili. La probabilità che nell'ordinamento i numeri 1 e 2 appaiano contigui e nell'ordine dato è $p = \frac{2}{6}$ poiché tra i sei (N) ordinamenti differenti possibili quelli in cui si verifica l'evento sono due (k).

Esempio 2.2. Se in un lotto di $N = 5000$ unità di prodotto di un macchinario si rappresenta con $A_i, i = 1, ..., 5000$, l'evento che si manifesta qualora la i-esima unità sia "difettosa" e se nel lotto si osservano $k = 100$ unità difettose, la probabilità che una futura unità prodotta dal macchinario sia difettosa (evento $A_i = A_{5001}$) è

$$p = \frac{100}{5000} = 0.02.$$

Anche a volersi limitare a considerare situazioni del tipo indicato nei due schemi precedenti - che rispetto al problema della valutazione della probabilità presentano effettivi riferimenti oggettivi - non devono sfuggire certi aspetti soggettivi. In effetti, nel primo schema è soggettivo il giudizio sulla *eguale plausibilità* dei casi possibili mentre nel secondo è soggettivo il giudizio di *omogeneità* degli eventi considerati e, soprattutto, la scelta di far pesare esclusivamente la frequenza osservata sulla valutazione di probabilità. Lo sforzo di far discendere una definizione di probabilità da uno dei due precedenti schemi sembra, in ogni caso, inane. Il primo, infatti, presuppone di aver assegnato un significato alla locuzione *egualmente plausibile*. È possibile farlo senza cadere nel circolo vizioso creato interpretando la locuzione come *aventi la stessa probabilità*? Il secondo, che interpreta la probabilità p come "su un gran numero di prove l'evento A si verificherà nel p-percento dei casi", con le parole di de Finetti

> perché valesse come definizione, dovrebbe costituire una vera e propria profezia, nel qual caso potremmo dire senz'altro frequenza senza introdurre il termine probabilità che sarebbe superfluo. Ma invece sappiamo che con quella frase non si esclude nessuna possibilità, neppure che la frequenza scenda allo zero o raggiunga l'unità: essa non ha quindi alcun attributo per poter logicamente dar luogo a una definizione.

Ma, anche a prescindere da tali osservazioni, è sempre possibile ricorrere ai predetti schemi per dare una interpretazione alla probabilità di un evento intesa questa come una misura che indichi la maggior o minore attendibilità dello stesso evento? Per rispondere, si consideri la frase espressa da un tifoso: "la probabilità che il Milan vinca il campionato di calcio di serie A nell'anno 2011-2012 è 0.80". Si tratta di una espressione che ognuno comprende, ma sembra chiaro che il tifoso non pensa ad un'urna contenente, ad esempio, 100 esiti di campionati, tra i quali 80 per cento sono favorevoli al Milan, dalla quale viene estratto l'esito del campionato dell'anno in questione, e neanche pensa che quel campionato possa ripetersi, ad esempio, 100 volte osservando il Milan vincente 80 volte su 100. In altre parole, la valutazione espressa dal tifoso non può essere inquadrata in uno degli schemi precedenti. In più, tale valutazione esige il riferimento al soggetto (quel particolare tifoso) che tale valutazione di probabilità ha espresso (altri tifosi potrebbero valutare diversamente la probabilità in discussione). Per dare un significato effettivo alla frase del tifoso, si tratta allora di escogitare un metodo di misura della probabilità corrispondente alla sua opinione. Entro certi limiti l'atteggiamento del tifoso che ha provveduto alla valutazione di probabilità può interpretarsi nel senso che, se costui fosse obbligato ad accettare scommesse riguardanti l'esito del campionato, egli sarebbe disposto a pagare 80 centesimi di Euro (o altre unità monetarie) per ricevere 1 Euro qualora il Milan vincesse il campionato e niente in caso contrario. Questo significato della probabilità come *quota di scommessa* rappresenta il punto di partenza della definizione soggettiva di probabilità.

Definizione 2.3. Si dice scommessa di quota $p \in \mathbf{R}$ sull'evento A la scommessa secondo la quale, qualunque sia il numero reale $c \neq 0$, versando una somma cp si riceve un importo c se A si verifica e nulla in caso contrario.

Osservazione 2.4. Le parole *quota* e *importo* assumono nella definizione precedente un significato più generale di quello ordinario che presume che c e p siano numeri reali positivi.

In generale perciò, dire che la probabilità di un evento A, per un determinato soggetto, è eguale a p, significa che costui, se fosse obbligato ad impegnarsi in una scommessa sull'evento A in base ad una determinata quota, sceglierebbe quest'ultima eguale a p. Può un soggetto scegliere arbitrariamente e senza alcun vincolo la quota p? Con riferimento all'esempio del tifoso, è chiaro che se egli valutasse, ad esempio, la quota pari a $p > 1$ incorrerebbe in una perdita certa quando l'importo c della scomessa fosse positivo, poiché a fronte del pagamento della somma $cp > c$ ricaverebbe la somma c se il Milan vincesse il campionato e nulla in caso contrario. È compito della teoria della probabilità stabilire regole minime a cui la scelta della quota dovrebbe essere vincolata al fine di evitare conseguenze non desiderabili del tipo di quella precedentemente illustrata. Dalle condizioni di scommessa precedentemente richiamate non è arduo enunciare una regola (*condizione di coerenza*) atta a discriminare tra valutazioni (di quote) ammissibili e non ammissibili.

Sia $\mathcal{E}$ una classe qualunque di eventi e $P : \mathcal{E} \to \mathcal{R}$ una funzione a valori reali che ad ogni evento $E \in \mathcal{E}$ associa la relativa quota di scommessa, $P(E)$. Convenendo di chiamare *banco* il soggetto che effettua le valutazioni $P(E)$, per $E \in \mathcal{E}$, il suo *guadagno aleatorio* relativo alla scommessa sull'evento E d'importo $c \neq 0$, fissato positivo o negativo da un ipotetico *scommettitore*, è dato da

$$cP(E) - cI_E,$$

ove I_E è la funzione indicatrice che vale 1 quando E si verifica e 0 altrimenti. Analogamente

$$\sum_{i=1}^{n} c_i(P(E_i) - I_{E_i})$$

rappresenta il guadagno del banco corrispondente ad una combinazione (finita) di scommesse sugli eventi $E_1, E_2, \cdots, E_n$ d'importi non nulli $c_1, c_2, \cdots, c_n$ rispettivamente.

Sembra ragionevole che il banco scarti quelle funzioni P per le quali esista almeno una combinazione di scommesse con guadagno uniformemente negativo, cioè con guadagno che, indipendentemente dal verificarsi o meno degli eventi considerati, risulti sempre negativo. A causa dell'arbitrarietà del segno degli importi, la cui scelta è affidata allo scommettitore, il banco non avrà neppure convenienza a fissare P in modo che i guadagni di certe combinazioni siano uniformemente positivi. In tal caso infatti, basterebbe che lo scommettitore cambiasse i segni degli importi per ricadere nella prima situazione. L'arbitrarietà del segno degli importi fa sì che $cP(E) - cI_E$ non rappresenti il guadagno del banco inteso in senso proprio, ma, a seconda del segno di c, può rappresentare anche quello dello scommettitore; perciò, richiedendo che esso non risulti uniformemente negativo per ogni c diverso da zero si vuole che non risulti tale né per il banco né per lo scommettitore o, se si vuole, la valutazione di P deve restare inalterata qualora si scambino i ruoli tra banco e scomettitore.

È da queste considerazioni di carattere qualitativo che scaturisce il cosiddetto *principio di coerenza* che diamo nella seguente.

Definizione 2.5. La funzione reale $P : \mathcal{E} \to \mathcal{R}$ si dice coerente su $\mathcal{E}$ se, per ogni sottoclasse finita $\{E_1, ..., E_n\}$ di eventi in $\mathcal{E}$ e per ogni $c_1, ..., c_n$ reali e non nulli, si ha

$$\min \sum_{i=1}^{n} c_i(P(E_i) - I_{E_i}) \leq 0 \leq \max \sum_{i=1}^{n} c_i(P(E_i) - I_{E_i}) \tag{2.1}$$

dove min[max] indica il valore minimo[massimo] del guadagno al variare dei valori logici di $E_1, E_2, \cdots, E_n$.

L'aggettivo *coerente* sta ad indicare che solo con funzioni P soddisfacenti la condizione (2.1), il banco può essere certo che, comunque si scelgano gli eventi $E_1, E_2, \ldots, E_n$ della famiglia $\mathcal{E}$ e gli importi $c_1, c_2, \ldots, c_n$, il suo guadagno non sarà negativo in corrispondenza ad ogni realizzazione logicamente possibile degli eventi prescelti. Per ogni altra P, invece, esisterà almeno una scelta di eventi della classe $\mathcal{E}$ e di importi per cui il guadagno risulterà negativo qualunque sia la realizzazione degli eventi prescelti.

La condizione di coerenza (2.1) descritta nella Definizione 2.5 può essere espressa, equivalentemente, mediante il riferimento ai guadagni corrispondenti agli eventi elementari relativi alla data famiglia di eventi $\{E_1, E_2, \ldots, E_n\}$. Rappresentiamo in $S = \{0, 1\}^n$ l'insieme di tutti i modi secondo i quali gli elementi della famiglia $\{E_1, E_2, \cdots, E_n\}$ potrebbero manifestarsi. Per esempio: data la famiglia $\{E_1, E_2\}$ rappresentiamo con $(1, 1)$ la situazione in cui si manifesta sia E_1 che E_2, con $(0, 1)$ quella in cui E_1 non si manifesta mentre si manifesta E_2, con $(1, 0)$ la situazione in cui si manifesta E_1 ma non E_2 e, infine, con $(0, 0)$ quella in cui né E_1 né E_2 si manifestano. Osserviamo che non necessariamente tutti gli elementi di S corrispondono a manifestazioni logicamente possibili per gli elementi della famiglia $\{E_i\}_{i=1}^n$; infatti, nell'esempio precedente, qualora il manifestarsi di E_1 implichi che anche E_2 si manifesti, il punto $(1, 0)$ corrisponde ad una situazione impossibile. Indichiamo con $\Omega := \{\omega_1, \omega_2, \cdots, \omega_s\}$ il sottoinsieme di elementi di S corrispondenti ai modi logicamente possibili per il manifestarsi degli elementi della famiglia $\{E_i\}_{i=1}^n$; gli elementi ω_i sono gli *eventi elementari* relativi alla famiglia. Notiamo che gli elementi della famiglia $\{E_i\}_{i=1}^n$ sono identificati da sottoinsiemi di Ω. Precisamente, il generico elemento E_i della famiglia $\{E_i\}_{i=1}^n$ è identificato dal sottoinsieme di eventi elementari il cui manifestarsi implica il manifestarsi di E_i; usando lo stesso simbolo E_i anche per questo sottinsieme di Ω scriveremo $E_i = \bigcup_{\omega_j \subseteq E_i} \{\omega_j\}$. Si consideri ora il guadagno $\sum_{i=1}^n c_i(P(E_i) - I_{E_i})$; poiché verificandosi l'evento elementare ω_k, lo stesso è dato dal *ricavo* certo $\sum_{i=1}^{n} c_i P(E_i)$ meno gli *esborsi* c_i corrispondenti alle scommesse sugli eventi implicati da ω_k, si avrà

$$\sum_{i=1}^{n} c_i (P(E_i) - I_{E_i}) = \sum_{k=1}^{s} I_{\omega_k} (\sum_{i=1}^{n} c_i P(E_i) - \sum_{i:E_i \supseteq \omega_k} c_i).$$

Posto allora, per $k = 1, 2, ..., s$,

$$g_k = \sum_{i=1}^{n} c_i P(E_i) - \sum_{i:E_i \supseteq \omega_k} c_i,$$

la condizione di coerenza (2.1) della Definizione 2.5 potrà esprimersi anche mediante

$$\min\{g_1, g_2, \cdots, g_s\} \leq 0 \leq \max\{g_1, g_2, \cdots, g_s\}. \tag{2.2}$$

Osservazione 2.6. Dalla mera definizione di coerenza discende che se P è coerente su $\mathcal{E}$ la sua restrizione a $\mathcal{E}^* \subset \mathcal{E}$ è ivi coerente.

Esempio 2.7. Si consideri la classe finita di eventi $\mathcal{E} = \{E_1, E_2, \cdots, E_n\}$ costituente una partizione dell'evento certo e sia $P : \mathcal{E} \to \mathcal{R}$ tale che $P(E_i) = \frac{1}{n}$ per $i = 1, 2, ..., n$. Sono tali quote coerenti?

Poiché gli eventi elementari relativi alla partizione $\{E_1, E_2, \cdots, E_n\}$ coincidono con gli eventi stessi, il generico guadagno g_k è dato da

$$g_k = \sum_{i=1}^{n} c_i \frac{1}{n} - \sum_{i:E_i \supseteq \omega_k} c_i = \frac{1}{n} \sum_{i=1}^{n} c_i - c_k.$$

Qualunque sia la scelta degli importi $c_1, ..., c_n$, si ha dunque che $\sum_{k=1}^{n} g_k = 0$; ne discende che vi saranno guadagni g_k non negativi e guadagni non positivi e ciò implica che

$$\min\{g_1, g_2, \cdots, g_n\} \leq 0 \leq \max\{g_1, g_2, \cdots, g_n\}.$$

Alla stessa conclusione di coerenza si perverrebbe, come il lettore potrà verificare, qualora le quote $P(E_i)$, anziché pari a $\frac{1}{n}$, fossero rispettivamente uguali alle quantità $p_1, ..., p_n$ tutte non negative e tali che $\sum_{i=1}^{n} p_i = 1$.

Esempio 2.8. Sia $\mathcal{E} = \{E_1, E_2, E_3\}$; assumiamo che gli eventi E_1 ed E_2 siano incompatibili e che l'evento E_3 si manifesti se e solo se si manifesta E_1 oppure E_2. Sono le quote $P(E_1) = 0.30,\ P(E_2) = 0.60,\ P(E_3) = 0.80$ coerenti sulla famiglia data?

Come si può immediatamente controllare, gli eventi elementari relativi alla famiglia $\mathcal{E}$ sono dati da ω_1 = (1,0,1), ω_2= (0,1,1), ω_3 = (0,0,0). Fissati gli importi c_1, c_2, c_3, in corrispondenza dei tre eventi elementari si hanno i seguenti guadagni:

$$\begin{aligned}
g_1 &= 0.30c_1 + 0.60c_2 + 0.80c_3 - c_1 - c_3 = -0.70c_1 + 0.60c_2 - 0.20c_3; \\
g_2 &= 0.30c_1 + 0.60c_2 + 0.80c_3 - c_2 - c_3 = -0.80c_1 + 0.40c_2 - 0.20c_3; \\
g_3 &= 0.30c_1 + 0.60c_2 + 0.80c_3.
\end{aligned}$$

È immediato scorgere che esistono importi c_1, c_2, c_3 (ad esempio $c_1 = c_3 = 10, c_2 = -20$) che rendono tutti i guadagni g_1, g_2, g_3 negativi contro la condizione di coerenza.

Esempio 2.9. Sia $\mathcal{E}$ =$\{E_1, E_2, \cdots, E_m, \cdots\}$ una partizione numerabilmente infinita dell'evento certo e siano $0 = P(E_1) = P(E_2) = \cdots = P(E_m) = \cdots$ le quote fissate. Sono esse coerenti?

Si consideri una qualunque sottoclasse finita $\mathcal{E}^* = \{E_1^*, ..., E_n^*\}$ di eventi di $\mathcal{E}$ con rispettivi importi $c_1^*, \cdots, c_n^*$. La classe degli eventi elementari relativa ad $\mathcal{E}^*$ è costituita dalle n-uple $\omega_1 = (1, 0, ..., 0), ..., \omega_n = (0, ..., 0, 1)$ con l'aggiunta della n-upla $\omega_{n+1} = (0, ..., 0)$ corrispondente alla situazione in cui nessun elemento di $\mathcal{E}^*$ si manifesta; osserviamo che questa situazione è possibile poiché $\mathcal{E}$ è una partizione numerabilmente infinita dell'evento certo. In corrispondenza degli eventi elementari relativi a $\mathcal{E}^*$ abbiamo i guadagni:

$$g_k = \sum_{i=1}^{n} c_i^* P(E_i^*) - \sum_{i: E_i^* \supset \omega_k} c_i^* = -c_k^*$$

per $k = 1, ..., n$ mentre

$$g_{n+1} = \sum_{i=1}^{n} c_i^* P(E_i^*) - \sum_{i: E_i^* \supset \omega_k} c_i^* = 0.$$

Poiché evidentemente

$$\min\{-c_1^*, \cdots, -c_n^*, 0\} \leq 0 \leq \max\{-c_1^*, \cdots, -c_n^*, 0\}$$

ne discende che la condizione (2.2) è soddisfatta per ogni scelta di $c_1^*, \cdots, c_n^*$. Per l'arbitrarietà della sottoclasse considerata segue che $P(E_i) = 0$ per ogni $E_i \in \mathcal{E}$ è coerente. Lasciamo al lettore il compito di verificare invece che le quote $P(E_i) = p > 0$ per ogni $E_i \in \mathcal{E}$ non sono coerenti.

A conclusione di questa sezione giova far osservare che uno stesso soggetto non può valutare in modo differente, e coerentemente, le quote relative agli stessi eventi. Infatti, la valutazione P è stata definita come funzione! La seguente precisazione ribadisce la legittimità della Definizione 2.5.

Proposizione 2.10. *Sia $\mathcal{E}$ una famiglia di eventi e siano P e P^*, due valutazioni. Allora la condizione di coerenza implica $P(E) = P^*(E)$ per ogni $E \in \mathcal{E}$.*

Dimostrazione. Si considerino due combinazioni di scommesse simultaneamente attuate da un soggetto sugli eventi$\{E_1, E_2, \cdots, E_n\}$ della famiglia $\mathcal{E}$ con quote rispettivamente $P(E_i)$ e $P(E_i^*)$ e con importi $c_1, c_2, \cdots, c_n$ e $c_1^*, c_2^*, \cdots, c_n^*$. Il guadagno del banco è

$$G = \sum_{i=1}^{n} c_i(P(E_i) - I_{E_i}) + \sum_{i=1}^{n} c_i^*(P^*(E_i) - I_{E_i})$$

Se ora si fissa $c_i^* = -c_i$, allora G diventa

$$G = \sum_{i=1}^{n} c_i(P(E_i) - P^*(E_i))$$

e se $P(E_i) \neq P^*(E_i)$ per almeno un indice i, allora si potranno scegliere i corrispondenti importi c_i in modo che $G > 0$ contro la coerenza. L'arbitrarietà della famiglia finita considerata permette di pervenire alla tesi. ■

3. Definizione di Probabilità

Nelle pagine che precedono abbiamo convenuto di misurare il grado di attendibilità di un evento da parte di un individuo mediante una quota di scommessa attuata sotto certe condizioni. La condizione di coerenza espressa nella Definizione 2.5 appare difficilmente eliminabile in vista delle conseguenze negative che dovrebbero essere sopportate da coloro che la trascurassero. D'altro canto, imporre, in generale, ulteriori condizioni restrittive alla valutazione delle quote di scommessa non sembra giustificato, mentre restrizioni più o meno forti potranno essere adottate, di volta in volta, a seconda dei problemi affrontati e nelle circostanze che ne suggeriscono la soluzione. È però opportuno osservare che il metodo di misura delle quote attraverso il meccanismo della scommessa ha dei limiti intrinseci dovuti al diverso grado di *avversione al rischio* provato da un soggetto in connessione alla maggiore o minore entità degli importi posti in gioco. Per questa ragione è consigliabile riferirsi a condizioni di scommesse coinvolgenti importi non grandi in modo da annullare gli effetti dovuti al suddetto fenomeno. D'altra parte, la scommessa su uno o più eventi svolge il ruolo di uno strumento di misura che, pur traendo origine dagli schemi dei giochi d'azzardo, si applica alla valutazione di quote connesse a eventi qualunque relativi a fenomeni di carattere fisico, economico, o di qualsiasi altro tipo. La scommessa, infatti, costituisce un criterio operativo per definire la quota secondo un punto di vista comune a campi di indagine completamente diversi. Si pensi, ad esempio, alla "durezza" dei materiali solidi per la quale esiste una definizione operativa che coinvolge un ipotetico esperimento; per misurare la durezza, infatti, si fa riferimento alla resistenza che il materiale opporrebbe quando si tentasse di intaccarlo con mezzi meccanici. Ma, naturalmente, non è necessario, per riferirsi al concetto di durezza, immaginare di avere eseguito detto esperimento. Così, per la valutazione delle quote, non si richiede che la scommessa sia effettivamente attuata; si chiede invece che, nell'ipotesi che si scommettesse, le quote fissate non fossero tali da originare, con certezza, una perdita o un guadagno da parte di uno degli opponenti.

Tutto ciò premesso, fissiamo la fondamentale

Definizione 3.1. Data una classe $\mathcal{E}$ di eventi, ogni funzione $P : \mathcal{E} \to \mathcal{R}$ soddisfacente la condizione di coerenza (2.1) è detta probabilità o distribuzione di probabilità su $\mathcal{E}$; il valore $P(E)$ assunto dalla funzione P in corrispondenza di un evento $E \in \mathcal{E}$ è detto probabilità di E.

Naturalmente, la definizione appena proposta solleva il problema della esistenza di una probabilità per una qualunque classe di eventi considerata. Di questo, come dell'altro problema riguardante la possibilità di riformulare la condizione di coerenza in modo da prestarsi ad una più agevole trattazione matematica, si discuterà nelle prossime sezioni.

Osservazione 3.2. La condizione di coerenza (2.1) coinvolge solo sottoclassi finite di elementi di $\mathcal{E}$. Pertanto possiamo affermare che $P : \mathcal{E} \to \mathbf{R}$ è una probabilità su $\mathcal{E}$ se e solo se è una probabilità su ogni sottoclasse finita di $\mathcal{E}$.

Conviene mostrare subito come dalla Definizione 3.1 di probabilità discendano alcune sue fondamentali proprietà.

Proposizione 3.3. *Sia P una probabilità definita su una classe di eventi $\mathcal{E}$ contenente l'evento certo Ω e l'evento impossibile $\emptyset$.*

(a) *La probabilità di un qualunque evento E è compresa nell'intervallo $[0, 1]$, ovvero $P(E) \in [0, 1]$ per ogni $E \in \mathcal{E}$.*
(b) *La probabilità dell'evento certo Ω è eguale ad uno ($P(\Omega) = 1$) mentre la probabilità dell'evento impossibile $\emptyset$ è uguale a zero ($P(\emptyset) = 0$).*
(c) *Se gli eventi $E_1, E_2, \ldots, E_n$ e $\bigcup_{i=1}^n E_i$ appartengono ad $\mathcal{E}$ e gli eventi E_i sono a due a due incompatibili allora*

$$P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i).$$

Dimostrazione. Osservato che l'espressione del guadagno da una scommessa di importo c sul singolo evento $E \in \mathcal{E}$ è data da $c(P(E) - I_E)$ con possibili determinazioni $c(P(E) - 1)$, se E si verifica, e $cP(E)$ se E non si verifica, la (2.1) impone che sia

$$\min\{c(P(E) - 1), cP(E)\} \leq 0 \leq \max\{c(P(E) - 1), cP(E)\}$$

per ogni $c \neq 0$, ovvero che i due guadagni non abbiano lo stesso segno:

$$c(P(E) - 1) \times cP(E) = c^2 P(E)[P(E) - 1] \leq 0.$$

Da questa si trae $0 \leq P(E) \leq 1$ che dimostra la (a).

Nel caso dell'evento certo il guadagno è certo ed è dato da $c(P(\Omega) - 1)$. Segue che deve essere $c(P(\Omega) - 1) = 0$ per ogni $c \neq 0$ e dunque $P(\Omega) = 1$. Quando l'evento è impossibile, il guadagno della scommessa è altrettanto certo ed è dato da $cP(\emptyset)$ e dunque, dovendo essere $cP(\emptyset) = 0$ per ogni $c \neq 0$, segue che $P(\emptyset) = 0$. Ciò dimostra la (b).

Per dimostrare la (c), consideriamo la famiglia $\{E_1, E_2, \cdots, E_n, \bigcup_{i=1}^n E_i\}$. Una scommessa d'importo costante c su ognuno degli eventi $E_1, E_2, \cdots, E_n$ e di importo $-c$ su $\bigcup_{i=1}^n E_i$ produce il guadagno certo

$$c\left(\sum_{i=1}^n (P(E_i) - I_{E_i})\right) - c\left(P(\bigcup_{i=1}^n E_i) - I_{\bigcup E_i}\right) = c\left(\sum_{i=1}^n P(E_i) - P(\bigcup_{i=1}^n E_i)\right)$$

essendo $\sum_{i=1}^n I_{E_i} = I_{\bigcup E_i}$ per l'incompatibiltà degli eventi E_i. La condizione di coerenza impone allora $c\,(\sum_{i=1}^n P(E_i) - P(\bigcup_{i=1}^n E_i)) = 0$ per ogni $c \neq 0$ da cui la tesi di (c). ■

Il lettore avrà certamente colto come le proprietà della probabilità, espresse dalla Proposizione 3.3, siano semplici conseguenze della condizione di coerenza e costituiscano, pertanto, solo condizioni necessarie per la sua sussistenza. Che esse possano essere, con qualche precisazione riguardante la classe degli eventi $\mathcal{E}$, anche condizioni sufficienti è fatto che verrà appurato nella Sezione 6.

Esempio 3.4. Si riprenda l'Esempio 2.8. Le valutazioni ivi espresse non rappresentano le probabilità degli eventi considerati in quanto esse non rispettano la (c) della Proposizione 3.3.

Esempio 3.5 (Problema di Ellesberg). Un'urna contiene 90 palline di colore Rosso, Blu o Giallo. Si sa che il numero delle palline Rosse è 30 mentre delle restanti 60 non è noto quante siano quelle Blu e quante quelle Gialle. Indicando con R, B e G gli eventi corrispondenti al verificarsi di una pallina Rossa, Blu o Gialla nell'estrazione di una pallina dall'urna, le valutazioni $P(R) = 1/3$, $P(B \bigcup G) = 2/3$, $P(G) = 1/3$, $P(R \bigcup B) > 2/3$ sono coerenti?

Per la condizione necessaria di coerenza deve essere $P(B \bigcup G) = P(B) + P(G) = 2/3$; ne discende $P(B) = 2/3 - P(G) = 1/3$. D'altro canto, per la stessa condizione, deve essere anche $P(R) + P(B) = P(R \bigcup B) > 2/3$ da cui discende $P(B) > 1/3$. Ne consegue l'incoerenza delle valutazioni.

Esempio 3.6. Sia $\mathcal{E} = \{E_1, E_2, \cdots, E_n\}$ una famiglia di eventi e $\Omega = \{\omega_1, \omega_2, \cdots, \omega_s\}$ la classe degli eventi elementari relativi a $\mathcal{E}$. Posto che si sia assegnata la probabilità sulla classe degli eventi elementari mediante una funzione P, come nell'Esempio 2.7, $P(\{\omega_k\}) = p_k \geq 0$ per $k = 1, 2, ..., s$, con $\sum_{k=1}^s P(\{\omega_k\}) = 1$, la valutazione $P_A(E_i) := \sum_{\omega_k \subset E_i} P(\{\omega_k\})$, per $i = 1, 2, .., n$ è una probabilità sulla classe di eventi $\mathcal{E}$?

Il guadagno sulla combinazione di scommesse sugli eventi $\{E_1, E_2, \cdots, E_n\}$ di importi rispettivi $c_1, ..., c_n$ è dato da:

$$\sum_{k=1}^s I_{\omega_k}(\sum_{i=1}^n c_i P_A(E_i) - \sum_{\{i: E_i \supset \omega_k\}} c_i) = \sum_{k=1}^s I_{\omega_k} g_k.$$

Si osservi ora che

$$\begin{aligned} M &= \sum_{k=1}^{s} g_k P(\{\omega_k\}) \\ &= \sum_{i=1}^{n} c_i P_A(E_i) - \sum_{k=1}^{s} P(\{\omega_k\}) \sum_{\{i: E_i \supset \omega_k\}} c_i \\ &= \sum_{i=1}^{n} c_i P_A(E_i) - \sum_{i=1}^{n} c_i \sum_{\{k: \omega_k \subset E_i\}} P(\{\omega_k\}) \\ &= 0, \end{aligned}$$

l'ultima eguaglianza essendo dovuta alla definizione di $P_A(E_i)$. Dalla precedente e dal fatto che M è combinazione lineare convessa di $g_1, g_2, \cdots g_s$ segue che :

$$\min\{g_1, g_2, \cdots g_s\} \leq 0 \leq \max\{g_1, g_2, \cdots g_s\}$$

e con ciò la conclusione che P_A è una distribuzione di probabilità sulla classe $\mathcal{E}$.

Nella prossima sezione ci occuperemo del problema, già annunciato, dell'esistenza di una distribuzione di probabilità. A tale, ma anche ad altri scopi, è importante il risultato espresso dalla seguente:

Proposizione 3.7. *Sia $(P_N)_{N=1}^{\infty}$ una successione di distribuzioni di probabilità su una classe di eventi $\mathcal{E}$ e sia $S = \{E \in \mathcal{E} : lim_{N\to\infty} P_N(E) =: P(E)$ esistente$\}$ una classe non vuota. Allora la funzione P, definita su $S \subseteq \mathcal{E}$, è una distribuzione di probabilità sulla classe S.*

Dimostrazione. Considerati $E_1, E_2, \cdots, E_n$ in S e $c_1, c_2, \cdots, c_n$ differenti da zero, segue che, per $N = 1, 2, ...,$

$$\min\{\sum_{i=1}^{n} c_i(P_N(E_i) - I_{E_i})\} \leq 0 \leq \max\{\sum_{i=1}^{n} c_i(P_N(E_i) - I_{E_i})\}$$

poiché P_N è una probabilità. Passando al limite per $N \to \infty$ si ottiene la condizione di coerenza espressa per P su $E_1, ..., E_n$ e la tesi segue dall'arbitrarietà di $E_1, E_2, \cdots, E_n$ in S. ■

Il contenuto della proposizione precedente, espresso a parole, è che il limite di una successione di distribuzioni di probabilità è ancora una distribuzione di probabilità. Inoltre, vale la pena di osservare che la tesi della proposizione continua a sussistere anche se le P_N, al variare di N, fossero definite su classi differenti di eventi purché tutte contenute in $\mathcal{E}$.

Esempio 3.8. Consideriamo la classe $\mathcal{E}$ dei sottoinsiemi di $\mathcal{N} = \{1, 2, 3, ...\}$. Ricordando gli esempi 2.7 e 3.6, è facile verificare come, per $N = 1, 2, ...,$

sia una probabilità su $\mathcal{E}$ la funzione P_N definita, per $E \in \mathcal{E}$, dalla

$$P_N(E) = \frac{|E \bigcap \{1, ..., N\}|}{N}$$

ove $|G|$ indica la cardinalità del generico sottoinsieme G di $\mathcal{N}$. È inoltre agevole appurare che

$$lim_{N \to \infty} P_N(E) = P(E) = 0$$

se E è un sottoinsieme finito di $\mathcal{N}$, mentre $lim_{N \to \infty} P_N(E) = 1$ se E è il complementare di un sottoinsieme finito, ovvero un sottoinsieme cofinito di $\mathcal{N}$. Dunque P definita sulla classe $\mathcal{S}$ dei sottoinsiemi finiti e cofiniti di $\mathcal{N}$ da $P(E) = 0$ oppure 1, a seconda che E sia finito o cofinito, è una distribuzione di probabilità.

4. Esistenza di una distribuzione di probabilità

Abbiamo già sollevato il problema della necessità di appurare l'esistenza di una distribuzione di probabilità, cioè se, data una classe arbitraria di eventi, sia possibile definire su di essa una funzione a valori reali soddisfacente la condizione di coerenza. Vedremo che la risposta positiva riposa sulla soluzione, pure positiva, del cosiddetto *problema di estensione o prolungamento* di una distribuzione di probabilità. Si tratta di una circostanza del massimo interesse teorico a motivo della naturale necessità di mantenere le valutazioni già effettuate su una classe di eventi $\mathcal{E}$ quando questa subisca un allargamento per l'esigenza di considerare nuovi eventi che in un primo momento non erano stati considerati. Per chiarire, si supponga che un individuo, nell'affrontare una certa questione, abbia ritenuto necessario considerare la classe degli eventi $\{E_1, E_2, \ldots, E_{n_1}\}$ a cui, coerentemente, ha assegnato la distribuzione di probabilità P e che, in una successiva fase della ricerca, lo stesso individuo si veda costretto a maggior approfondimenti, per prendere in esame ulteriori eventi $\{E_1^{'}, E_2^{'}, \ldots, E_{n_2}^{'}\}$. Sarà possibile assegnare una distribuzione di probabilità sulla nuova classe di eventi $\{E_1, E_2, \ldots, E_{n_1} E_1^{'}, E_2^{'}, \ldots, E_{n_2}^{'}\}$ in modo da rispettare le valutazioni di probabilità già effettuate per gli eventi$\{E_1, E_2, \ldots, E_{n_1}\}$? Ad esempio, se si è valutata pari a p la probabilità di $E_1 \bigcup E_2$ è possibile assegnare una probabilità ad E_1 in modo da salvaguardare la precedente probabilità p?

Definizione 4.1. Siano $\mathcal{E}$ e $\mathcal{E}_1$ due classi di eventi con $\mathcal{E} \subset \mathcal{E}_1$. La funzione P_1 definita su $\mathcal{E}_1$ si dice estensione o prolungamento della funzione P, definita su $\mathcal{E}$, se $P_1(E) = P(E)$ per ogni $E \in \mathcal{E}$.

Obiettivo di questa sezione è dimostrare che ogni distribuzione di probabilità su $\mathcal{E}$ ammette un prolungamento su $\mathcal{E}_1$ che è ancora una distribuzione di probabilità. A tal fine si procederà per gradi enunciando due proposizioni la prima delle quali di fondamentale importanza perché, da sola, risolve già il problema del prolungamento in un caso speciale.

Proposizione 4.2. *Sia $\mathcal{E}$ una classe di eventi su cui è definita una distribuzione di probabilità P e sia $\mathcal{F} = \{E_1, E_2, \cdots, E_m\}$ una famiglia finita di eventi che non abbia elementi in comune con $\mathcal{E}$. Allora esiste una distribuzione di probabilità P_1 su $\mathcal{E}_1 = \mathcal{E} \bigcup \mathcal{F}$ tale che $P_1(E) = P(E)$ per ogni $E \in \mathcal{E}$.*

Dimostrazione. Basterà dimostrare che l'estensione può eseguirsi relativamente ad un elemento di $\mathcal{F}$, sia esso E_1; la tesi seguirà riapplicando, sequenzialmente, il procedimento.

Si indichino con p la probabilità di E_1 e con V_+ e V_- gli insiemi dei valori di p per cui esistono, rispettivamente, una combinazione di scommesse con guadagni uniformemente positivi e una combinazione di scommesse con guadagni uniformemente negativi. Ovvero:

$$V_+ = \{p \in \mathbf{R} : \text{esistono } E'_1, ..., E'_{n_1} \in \mathcal{E} \text{ e } c'_1, ..., c'_{n_1} \text{ non nulli}$$
$$\text{tali che } \min\{p - I_{E_1} + \sum_{i=1}^{n_1} c_i{}'(P(E'_i) - I_{E'_i})\} > 0\}$$

e, analogamente,

$$V_- = \{p \in \mathbf{R} : \text{esistono } E''_1, ..., E''_{n_2} \in \mathcal{E} \text{ e } c''_1, ..., c''_{n_2} \text{ non nulli}$$
$$\text{tali che } \max\{p - I_{E_1} + \sum_{i=1}^{n_2} c_i{}''(P(E''_i) - I_{E''_i})\} < 0\}.$$

Gli insiemi V_+ e V_- sono certamente non vuoti in quanto ogni $p > 1$ appartiene a V_+ e ogni $p < 0$ appartiene a V_-. Inoltre, se $p \in V_+$ e $q > p$ allora $q \in V_+$ e, se $p \in V_-$ e $q < p$, allora $q \in V_-$; dunque V_+ e V_- sono intervalli.

Dimostreremo che:

(a) $\sup V_- \leq \inf V_+$;
(b) ogni punto dell'intevallo $[\sup V_-, \inf V_+] \subseteq [0, 1]$ è un valore ammissibile per la probabilità p di E_1.

Per dimostrare (a) procediamo per assurdo supponendo che $\inf V_+ < \sup V_-$. Esisterebbe allora $p^* \in V_- \cap V_+$ tale che, per opportuni $E'_1, ..., E'_{n_1}, E''_1, ..., E''_{n_2} \in \mathcal{E}$ e $c'_1, ..., c'_{n_1}, c''_1, ..., c''_{n_2}$ non nulli accade che

$$p^* - I_{E_1} + \sum_{i=1}^{n_1} c'_i(P(E'_i) - I_{E'_i}) > 0$$

e

$$p^* - I_{E_1} + \sum_{i=1}^{n_2} c''_i(P(E''_i) - I_{E''_i}) < 0$$

qualunque siano i valori logici degli eventi $E'_1, ..., E'_{n_1}, E''_1, ..., E''_{n_2}$; di conseguenza

$$\min(\sum_{i=1}^{n_1} c'_i(P(E'_i) - I_{E'_i}) - \sum_{i=1}^{n_2} c''_i(P(E''_i) - I_{E''_i})) > 0$$

contro l'ipotesi che P sia una probabilità su $\mathcal{E}$.

Osserviamo che $\sup V_+ \geq 0$ poiché $p \in V_+$ se $p < 0$; analogamente $\inf V_- \leq 1$. Inoltre i valori di p dell'intervallo aperto $(\sup V_-, \inf V_+)$ non appartengono né a V_- né a V_+; essi sono quindi valori ammissibili per la probabilità di E_1. Per dimostare (b) basterà allora verificare che $p' = \sup V_- \notin V_-$ e che $p'' = \inf V_+ \notin V_+$.

Procediamo, ancora per assurdo, assumendo che $p'' \in V_+$; esistono allora $E'_1, ..., E'_{n_1} \in \mathcal{E}$ e $c'_1, ..., c'_{n_1}$ non nulli per cui

$$\min\{p'' - I_{E'_1} + \sum_{i=1}^{n_1} c'_i(P(E'_i) - I_{E'_i})\} > 0.$$

Dunque esiste $\epsilon > 0$ sufficientemente piccolo e tale che

$$\min\{p'' - \epsilon - I_{E'_1} + \sum_{i=1}^{n_1} c'_i(P(E'_i) - I_{E'_i})\} > 0$$

ossia $p'' - \epsilon \in V_+$ che è falsa essendo $p'' = \inf V_+$. Similmente si dimostra che $p' \notin V_-$. ■

Proposizione 4.3. *Siano $\mathcal{E}$ e $\mathcal{E}_1$ due classi di eventi con $\mathcal{E} \subseteq \mathcal{E}_1$, e sia P una probabilità su $\mathcal{E}$. Allora le condizioni seguenti sono equivalenti:*

(a) *Esiste una probabilità su $\mathcal{E}_1$ che coincide con P sulla classe $\mathcal{E}$.*
(b) *Per ogni sottoinsieme finito $\mathcal{F}$ di $\mathcal{E}_1$, esiste una probabilità su $\mathcal{F}$ che coincide con P su $\mathcal{F} \cap \mathcal{E}$.*

Dimostrazione. Il fatto che la proposizione (a) implichi la (b) è di verifica immediata. Basterà quindi dimostrare che la proposizione (b) implica la (a).

Si rappresenti ogni probabilità su $\mathcal{E}_1$ come un punto dello spazio prodotto $\mathcal{S} = [0,1]^{\mathcal{E}}$ munito della topologia prodotto. Fissato un sottoinsieme finito $\mathcal{F}$ di $\mathcal{E}_1$, si indichi con $\mathcal{P}_{\mathcal{F}}$ il sottoinsieme di $\mathcal{S}$ costituito dai prolungamenti da $\mathcal{E}$ a $\mathcal{E}_1$ di P che sono probabilità su $\mathcal{F}$. Dalla condizione (b) discende che, qualunque sia $\mathcal{F}$, $\mathcal{P}_{\mathcal{F}}$ è un non vuoto .

Verifichiamo che $\mathcal{P}_{\mathcal{F}}$ è chiuso nella topologia prodotto di $\mathcal{S}$. Infatti, scelto f^* nella chiusura $\overline{\mathcal{P}_{\mathcal{F}}}$ di $\mathcal{P}_{\mathcal{F}}$ è palese che $f^* \in \mathcal{S}$. Dimostriamo che f^* coincide con P su $\mathcal{E}$ e che f^* è una probabilità su $\mathcal{F}$.

Per verificare la prima affermazione si osservi che, fissato $E \in \mathcal{E}$, l'insieme

$$\mathcal{I}_{n,E} = \{g \in \mathcal{S} : |g(E) - f^*(E)| < \frac{1}{n}\}$$

è un intorno di f^* per ogni $n = 1, 2, ...$ Poiché f^* appartiene alla chiusura di $\mathcal{P}_{\mathcal{F}}$, esiste una successione $\{g_n\}_1^\infty$ in $\mathcal{P}_{\mathcal{F}}$ per la quale $g_n \in$

$I_{n,E}$ per $n = 1, 2, \ldots$. Quindi, per $n = 1, 2, \ldots$, $|g_n(E) - f^*(E)| < \frac{1}{n}$ ovvero, essendo $g_n(E) = P(E)$,

$$|P(E) - f^*(E)| < \frac{1}{n}.$$

Ciò dimostra che $P(E) = f^*(E)$ per ogni $E \in \mathcal{E}$.

Si deve ora provare che f^* è una probabilità su $\mathcal{F}$. Per fissare le idee sia $\mathcal{F} = \{E_1, E_2, \cdots, E_m\}$. Allora, per $n = 1, 2, \ldots$, l'insieme

$$\mathcal{I}_n = \{g \in \mathcal{S} : |g(E_i) - f^*(E_i)| < \frac{1}{n}, \;\; i = 1, 2, \ldots, m\}$$

è un intorno di $f^* \in \overline{\mathcal{P}_\mathcal{F}}$ e, di conseguenza, esiste una successione $\{g_n\}_1^\infty$ di elementi di $\mathcal{P}_\mathcal{F}$ tale che $|g_n(E_i) - f^*(E_i)| < \frac{1}{n}$ per ogni $i = 1, 2, \ldots m$ e $n \geq 1$. La restrizione di f^* a $\mathcal{F}$ è quindi il limite di probabilità su $\mathcal{F}$ e, per la Proposizione 3.7, è essa stessa una probabilità su $\mathcal{F}$.

In definitiva $\mathcal{P}_\mathcal{F}$ è chiuso nella topologia prodotto di $\mathcal{F}$.

Si considerino ora $n \geq 1$ sottoclassi finite e non vuote di $\mathcal{E}_1$: siano esse $\mathcal{F}_1, \mathcal{F}_2, \cdots, \mathcal{F}_n$. Si ponga $\mathcal{F} = \bigcup_{k=1}^n \mathcal{F}_k$; grazie alla (b), l'insieme $\mathcal{P}_\mathcal{F}$ è non vuoto. Inoltre se $f \in \mathcal{P}_\mathcal{F}$, allora $f \in \mathcal{P}_{\mathcal{F}_k}$ per ogni $k = 1, 2, \ldots, n$ e perciò $f \in \cap_{k=1}^n \mathcal{P}_{\mathcal{F}_k}$. Ciò prova che la famiglia di chiusi $\{\mathcal{P}_\mathcal{F} : \mathcal{F}$ è una famiglia finita e non vuota di elementi di $\mathcal{E}_1\}$ gode della proprietà dell'intersezione finita. Allora, essendo $\mathcal{S}$ compatto, in forza ad un fondamentale teorema di Tychonoff, l'intersezione di tutti gli elementi della famiglia è non vuota. Perciò, se g è un elemento di questa intersezione, g è una probabilità su ogni sottoclasse finita di $\mathcal{E}_1$. Inoltre $g = P$ su $\mathcal{E}$. Ovvero g è una probabilità su $\mathcal{E}_1$ che coincide con P su $\mathcal{E}$. ■

Siamo ora in grado di dimostrare il seguente fondamentale teorema.

Teorema 4.4 (Teorema di estensione)**.** *Siano $\mathcal{E}$ e $\mathcal{E}_1$ due classi di eventi con $\mathcal{E} \subseteq \mathcal{E}_1$ e sia P una probabilità su $\mathcal{E}$. Allora esiste una probabilità P_1 su $\mathcal{E}_1$ che è una estensione di P.*

Dimostrazione. La proposizione 4.2 mostra che sussiste la condizione (b) della proposizione 4.3. ■

Il contenuto del Teorema 4.4, espresso a parole, afferma che se, entro il campo in cui sono assegnate, le valutazioni sono coerenti, nessuna difficoltà insorge che impedisca l'esistenza di valutazioni coerenti in un campo qualunque, coincidenti con quelle già esistenti dovunque queste ultime erano state effettuate.

Ritorniamo ora al problema dell'esistenza di una probabilità, la cui analisi avevamo momentaneamente rimandato. Supponiamo assegnata una classe di eventi $\mathcal{E}_1$ col proposito di definire su di essa una distribuzione di probabilità. Per fare ciò, si consideri un suo elemento E e, avvalendoci della Proposizione 3.7, fissiamo la sua probabilità $P(E)$ uguale a 0 oppure 1, a seconda che E sia impossibile o certo, oppure

$P(E)$ in $[0,1]$ se E è diverso dall'evento impossibile e dall'evento certo. Su $\mathcal{E}_1$ si può allora definire una distribuzione di probabilità per prolungamento da $\mathcal{E} = \{E\}$ a $\mathcal{E}_1$ di P in virtù del Teorema 4.4. Ciò è sufficiente per la dimostrazione del seguente:

Teorema 4.5. *Data una classe di eventi $\mathcal{E}$, esiste almeno una distribuzione di probabilità su $\mathcal{E}$.*

Il Teorema 4.5 si applica , in particolare, alla situazione in cui $\mathcal{E}$ è la classe di tutti i sottoinsiemi dell'insieme degli eventi elementari Ω.

Gli esempi seguenti, illustrano, con riferimento a classi finite, alcune situazioni in cui il prolungamento è variamente articolato (unico, non unico).

Esempio 4.6. Data la famiglia finita di eventi $\mathcal{E} = \{E_1, E_2, \cdots, E_n\}$ e la probabilità P definita su di essa, si prolunghi P alla famiglia $\mathcal{E}_1 = \mathcal{E} \bigcup \{\omega_1, \omega_2, \cdots, \omega_s\}$, essendo $\omega_1, ..., \omega_s$ gli eventi elementari relativi alla famiglia $\mathcal{E}$.

Indichiamo con P_1 un prolungamento di P a $\mathcal{E}_1$. Poiché, per $i = 1, ..., n$, $E_i = \bigcup_{\omega_j \subseteq E_i} \{\omega_j\}$, le probabilità $P_1(\{\omega_j\})$, per la (c) della Proposizione 3.3, devono soddisfare le equazioni seguenti

$$\begin{aligned}
\sum_{\{j:\omega_j \subseteq E_1\}} P_1(\{\omega_j\}) &= P_1(E_1) &= P(E_1) \\
\sum_{\{j:\omega_j \subseteq E_2\}} P_1(\{\omega_j\}) &= P_1(E_2) &= P(E_2) \\
&\cdots \\
\sum_{\{j:\omega_j \subseteq E_n\}}^{(n)} P_1(\{\omega_j\}) &= P_1(E_n) &= P(E_n)
\end{aligned} \tag{4.1}$$

con il vincolo

$$\sum_{j=1}^{s} P_1(\{\omega_j\}) = 1. \tag{4.2}$$

Poiché, per ipotesi, P è una probabilità, in base al Teorema di estensione il precedente sistema di equazioni ammetterà almeno una soluzione per le probabilità $P_1(\omega_j)$; osserviamo che il sistema è composto da $n+1$ equazioni in s incognite con $n \leq s \leq 2^n$.

Esempio 4.7. Si consideri il lancio di un dado regolare e gli eventi

$$\begin{aligned}
E_1 &:= \text{appare una faccia con punteggio} \leq 4; \\
E_2 &:= \text{appare una faccia con punteggio} \leq 5; \\
E_3 &:= \text{appare una faccia con punteggio 3,4 o 5.}
\end{aligned}$$

Siano $P(E_1) = 2/3$, $P(E_2) = 5/6$, $P(E_3) = 1/2$ (valutazione coerente).

Gli eventi elementari relativi alla famiglia $\{E_1, E_2, E_3\}$ sono dati da:

$$\begin{aligned}
\omega_1 &= (1,1,1) = \text{appare il 3 o il 4;}\\
\omega_2 &= (1,1,0) = \text{appare il numero 1 o il 2;}\\
\omega_3 &= (0,1,1) = \text{appare il 5;}\\
\omega_4 &= (0,0,0) = \text{appare il 6.}
\end{aligned}$$

Per $j = 1, ..., 4$ si indichino con q_j le probablità degli eventi ω_j; il sistema di equazioni (4.1) e (4.2) diventa

$$\begin{aligned}
2/3 &= q_1 + q_2\\
5/6 &= q_1 + q_2 + q_3\\
1/2 &= q_1 + q_3\\
1 &= q_1 + q_2 + q_3 + q_4.
\end{aligned}$$

La soluzione è unica e data da $q_1 = 1/3, q_2 = 1/3, q_3 = 1/6, q_4 = 1/6$ e ciò mostra che il prolungamento di P da $\{E_1, E_2, E_3\}$ a $\{E_1, E_2, E_3, \omega_1, \omega_2, \omega_3, \omega_4\}$ dato da $P_1(E_1) = 2/3, P_1(E_2) = 5/6, P_1(E_3) = 1/2, P_1(\{\omega_1\}) = P_1(\{\omega_2\}) = 1/3, P_1(\{\omega_3\}) = P_1(\{\omega_4\}) = 1/6$ è unico.

Esempio 4.8. Con riferimento al lancio di un dado regolare siano:

$$\begin{aligned}
E_1 &= \text{appare una faccia con punteggio pari;}\\
E_2 &= \text{appare una faccia con punteggio inferiore a 4.}
\end{aligned}$$

Gli eventi elementari sono dati da :

$$\begin{aligned}
\omega_1 &= \text{appare il 2;}\\
\omega_2 &= \text{appare il 4 o il 6;}\\
\omega_3 &= \text{appare il numero 1 o il 3;}\\
\omega_4 &= \text{appare il 5}
\end{aligned}$$

e, supponendo $P(E_1) = 1/3,\ P(E_2) = 1/2$, le loro probabilità $q_1, ..., q_4$ devono soddisfare il sistema :

$$\begin{aligned}
1/3 &= q_1 + q_2;\\
1/2 &= q_1 + q_3;\\
1 &= q_1 + q_2 + q_3 + q_4.
\end{aligned}$$

Poiché la soluzione è data da $q_1 = \lambda - 1/6,\ q_2 = 1/2 - \lambda,\ q_3 = 2/3 - \lambda,\ q_4 = \lambda$ con $1/6 \leq \lambda \leq 1/2$, il prolungamento di P da $\{E_1, E_2\}$ a $\{E_1, E_2, \omega_1, \omega_2, \omega_3, \omega_4\}$ è dato da P_1 con $P_1(E_1) = 1/3,\ P_1(E_2) = 1/2,\ P_1(\omega_1) = \lambda - 1/6,\ P_1(\omega_2) = 1/2 - \lambda,\ P_1(\omega_3) = 2/3 - \lambda,\ P_1(\omega_4) = \lambda$ con $1/6 \leq \lambda \leq 1/2$ e non è pertanto unico.

Esempio 4.9. Siano E_1 ed E_2 e $P(E_1),\ P(E_2)$ come nell'esempio precedente e si consideri il nuovo evento $E_3 =$ "appare un punteggio dispari". È unico il prolungamento di P alla classe $\{E_1, E_2, E_3\}$?

Poiché il nuovo evento E_3 dipende logicamente da $\{E_1, E_2\}$ ed è $E_3 = \{\omega_3, \omega_4\}$, la sua probabilità verrà assegnata tramite $P_1(E_3) = q_3 + q_4 = 2/3$ ed il prolungamento è dunque unico.

Si noti che la circostanza che E_3 dipenda logicamente da $\{E_1, E_2\}$ non implica, in generale, che il prolungamento sia unico. Se si considera come nuovo evento $E_4 =$ "appare un punteggio diverso da 2" si avrebbe $E_4 = \{\omega_2, \omega_3, \omega_4\}$ e quindi $P_1(E_4) = q_2 + q_3 + q_4 = 7/6 - \lambda$, con $1/6 \le \lambda \le 1/2$.

Esempio 4.10. Si supponga di avere assegnato $P(E_1 \bigcup E_2) = p$, con $0 \le p \le 1$. La probabilità P_1 su $\{E_1, E_1 \cup E_2\}$ definita da $P_1(E_1 \cup E_2) = p$, $P_1(E_1) = p/2$ è una distribuzione di probabilità che prolunga P?

Evidentemente P_1 è un prolungamento di P; dobbiamo solo verificare che sia una probabilità. Consideriamo gli eventi elementari relativi alla famiglia $\{E_1, E_1 \bigcup E_2\}$. Essi sono dati da $\omega_1 = (1,1)$, $\omega_2 = (0,1)$, $\omega_3 = (0,0)$. Qualora siano assegnate $P_1(E_1 \bigcup E_2) = p$ e la probabilità $P_1(E_1)$ dell'evento E_1, le probabilità q_1, q_2, q_3, degli eventi elementari $\omega_1, \omega_2, \omega_3$ devono soddisfare il sistema di equazioni:

$$\begin{aligned} p &= q_1 + q_2; \\ P_1(E_1) &= q_1; \\ 1 &= q_1 + q_2 + q_3 \end{aligned}$$

che ha soluzione unica fornita da: $q_1 = P_1(E_1)$, $q_2 = p - P_1(E_1)$, $q_3 = 1 - p$. Dunque qualunque valore dell'intervallo chiuso $[0, p]$ è ammissibile per $q_2 = P_1(E_1)$, in particolare il valore $p/2$. Si noti, ed è circostanza generale, che essendo $E_1 \subseteq (E_1 \bigcup E_2)$ si è concluso con

$$P_1(E_1) \le P_1(E_1 \cup E_2) = P(E_1 \cup E_2) = p.$$

Per terminare questa sezione è utile ottenere una relazione tra le probabilità di eventi quando questi risultano collegati mediante i loro indicatori. Quanto diremo costituisce un elemento importante della nozione di previsione di un numero aleatorio.

Sia $\{E_1, E_2, \cdots, E_n, A\}$ una classe di eventi e si sappia che I_A, l'indicatore dell'evento A, è una combinazione lineare di indicatori di tutti, parte o combinazioni di eventi E_i. Ad esempio, nel caso $A = E_1 \bigcup E_2$, sarà $I_A = I_{E_1} + I_{E_2} - I_{E_1 \bigcap E_2}$. Sia dunque, in generale, $I_A = \sum_{i=1}^n c_i I_{E_i}$ con c_i numeri reali. Allora $P(A) = \sum_{i=1}^n c_i P(E_i)$.

Infatti, si immagini una scommessa d'importo unitario su A e di importi $-c_1, -c_2, \cdots, -c_n$ sugli eventi E_i; senza perdere in generalità assumiamo che gli importi siano non nulli. Il guadagno risulta

$$P(A) - I_A - \sum_{i=1}^n c_i (P(E_i) - I_{E_i}) = P(A) - \sum_i c_i P(E_i)$$

e la condizione di coerenza rimarrà soddisfatta se e solo se

$$P(A) - \sum_{i=1}^{n} c_i P(E_i) = 0.$$

Vedremo nella sezione che segue una importante applicazione di questa anticipazione.

5. Altre proprietà della probabilità

Abbiamo già avuto modo, mediante la Proposizione 3.3, di considerare alcune delle proprietà di cui deve godere la probabilità; esse furono dedotte dal principio di coerenza. In questa sezione stabiliremo alcune altre conseguenze dello stesso principio le quali sono utili per valutare la probabilità di eventi inizialmente non inclusi nella classe a cui fu assegnata una certa distribuzione di probabilità P. Concettualmente, si tratta di regole per il prolungamento di P a certi eventi che sovente capita di prendere in esame. Ma le suddette conseguenze possono anche essere riguardate come proprietà di cui necessariamente deve godere una distribuzione di probabilità P sul campo in cui è definita. Per poter tener conto di queste due interpretazioni, non si preciserà se gli eventi coinvolti appartengono al dominio di P oppure non vi appartengono e se si tratta di descrivere proprietà di P oppure di fissare come P debba essere prolungata ad eventi non inclusi nel suo dominio. Ad esempio, la proprietà (c) della Proposizione 3.3 è stata data rispettando la prima delle suddette interpretazioni, ma potrà anche leggersi nel senso che, assegnata la probabilità P sugli eventi della classe $\{E_1, E_2, \cdots, E_n\}$, il suo prolungamento P_1 alla classe $\{E_1, ..., E_n, \bigcup_1^n E_i\}$, è dato da $P_1(E_i) = P(E_i)$, per $i = 1, 2, ..., n$, mentre $P_1(\bigcup_1^n E_i) = \sum_{i=1}^n P_1(E_i)$.

D'ora innanzi, si denoterà sempre col simbolo P la probabilità degli eventi in esame.

Proposizione 5.1.

(i) *Se* $E_1 \subseteq E_2$ *allora* $P(E_1) \leq P(E_2)$.

(ii) *Se* E^c *è l'evento complementare di* E, *allora* $P(E^c) = 1 - P(E)$.

(iii) *Se* $E_1, E_2, ..., E_n, ...$ *è una partizione numerabilmente infinita dell'evento* E, *allora* $P(E) \geq \sum_{i=1}^n P(E_i)$ *per ogni* $n = 1, 2, ...$ *e, di conseguenza,* $P(E) \geq \sum_{i=1}^\infty P(E_i)$.

(iv) $P(E_1 \bigcup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$.

(v) $P(\bigcup_{i=1}^n E_i) = \sum_{j=1}^n (-1)^{j+1} \sum_{1 \leq i_1 ... \leq i_j \leq i_n} P(E_{i_1} \cap E_{i_2} \cap ... \cap E_{i_j})$.

(vi) $P(\bigcup_{i=1}^n E_i) \leq \sum_{i=1}^n P(E_i)$.

(vii) *Dati gli eventi* $E_1, E_2, ...E_n$ *sia* $A_m^{(n)}$ *l'evento che è verificato se e solo se esattamente* $m \leq n$ *eventi* E_i *si verificano; allora*

$$P(A_m^{(n)}) = \sum_{j=m}^{n} \binom{j}{m} (-1)^{j+m} \sum_{1 \leq i_1 < \cdots < i_j \leq n} P(E_{i_1} \cap E_{i_2} \cap ... \cap E_{i_j}).$$

(viii) *Dati gli eventi* $E_1, E_2, ..., E_n$ *sia* $B_m^{(n)}$ *l'evento che è se e solo se almeno* $m \leq n$ *eventi* E_i *si verificano; allora*

$$P(B_m^{(n)}) = \sum_{i=m}^{n} \binom{j-1}{m-1} (-1)^{j+m} \sum_{1 \leq i_1 < \cdots < i_j \leq n} P(E_{i_1} \bigcap E_{i_2} \bigcap ... \bigcap E_{i_j}).$$

Dimostrazione. Per dimostare la (i) si osservi che $E_2 = E_1 \cup (E_2 \cap {E_1}^c)$ con E_1 e $(E_2 \cap {E_1}^c)$ incompatibili; perciò per la (c) della Proposizione 3.3 si avrà

$$P(E_2) = P(E_1) + P(E_2 \cap {E_1}^c) \geq P(E_1)$$

per la (a) della stessa Proposizione.

La (ii) segue dalla constatazione che E e E^c sono incompatibili e $E \bigcup E^c$ è l'evento certo Ω; la (b) e (c) della Proposizione 3.3 danno allora

$$P(E) + P(E^c) = P(\Omega) = 1.$$

Per dimostare la (iii) osserviamo che, per ipotesi, $E_i \cap E_j = \emptyset$ per ogni $i \neq j$ e inoltre $E = \bigcup_{i=1}^{\infty} E_i = (\bigcup_{i=1}^{n} E_i) \bigcup (\bigcup_{i=n+1}^{\infty} E_i)$. Dunque

$$P(E) = P(\bigcup_{i}^{n} E_i) + P(\bigcup_{n+1}^{\infty} E_i) \geq P(\bigcup_{i}^{n} E_i) = \sum_{i=1}^{n} P(E_i)$$

per ogni $n \geq 1$. Il risultato segue dal passaggio al limite per $n \to \infty$.

La (iv) segue dalle scomposizioni

$$E_1 \cup E_2 = E_1 \cup (E_2 \cap {E_1}^c)$$

$$E_2 = (E_1 \cap E_2) \cup (E_2 \cap {E_1}^c)$$

in eventi incompatibili. Infatti, per la Proposizione 3.3,

$$P(E_1 \cup E_2) = P(E_1) + P(E_2 \cap {E_1}^c)$$

mentre

$$P(E_2) = P(E_1 \cap E_2) + P(E_2 \cap {E_1}^c);$$

sottraendo membro a membro la seconda dalla prima si ottiene la tesi.

La (v) è una generalizzazione della (iv) che si può dimostrare per induzione su n. Infatti, riducendosi alla (iv), essa è vera per $n = 2$.

Supponiamo che sia vera per un intero $n > 2$ e consideriamo:

$$
\begin{aligned}
&P(\bigcup_{i=1}^{n+1} E_i) \\
&\quad = P\left((\bigcup_{i=1}^{n} E_i) \bigcup E_{n+1}\right) \\
&\quad = P(\bigcup_{1}^{n} E_i) + P(E_{n+1}) - P((\bigcup_{1}^{n} E_i) \bigcap E_{n+1}) \\
&\quad = \sum_{j=1}^{n} (-1)^{j+1} \sum_{i_0 < i_1 < \ldots < i_j \le n} P(E_{i_1} \bigcap \ldots \bigcap E_{i_j}) + P(E_{n+1}) - P(\bigcup_{1}^{n} (E_i \cap E_{n+1})) \\
&\quad = \sum_{j=1}^{n+1} (-1)^{j+1} \sum_{1 \le i_0 < i_1 < \ldots < i_j \le n+1} P(E_{i_1} \bigcap \ldots \bigcap E_{i_j})
\end{aligned}
$$

che è la (v) con $n+1$ al posto di n.

La (vi) è vera per $n = 2$; per induzione si perviene alla tesi.

Per dimostrare la (vii) osserviamo che

$$I_{A_0^{(n)}} = 1 - I_{\cup_1^n E_i}$$

mentre, per $m \ge 1$,

$$I_{A_m^{(n)}} = \sum_{j=m}^{n} \binom{j}{m} (-1)^{j-m} \sum_{1 \le i_0 < i_1 < \ldots < i_j \le n} I_{E_{i_1}} \cdots I_{E_{i_j}}.$$

Per quanto riportato nella parte finale della precedente sezione, si deduce allora che:

$$P(A_0^{(n)}) = 1 - P(\bigcup_{1}^{n} E_i)$$

e

$$P(A_m^{(n)}) = \sum_{j=m}^{n} \binom{j}{m} (-1)^{j-m} \sigma_j \tag{5.1}$$

avendo indicato con

$$\sigma_j = \sum_{1 \le i_0 < i_1 < \ldots < i_j \le n} P(E_{i_1} \cap E_{i_2} \ldots \cap E_{i_j}). \tag{5.2}$$

Infine, per dimostrare la (viii), si può osservare che $B_m^{(n)} = \bigcup_{k=m}^n A_k^{(n)}$ e che gli eventi $A_k^{(n)}$ sono a due a due incompatibili. Pertanto,

$$\begin{aligned}
P(B_m^{(n)}) &= \sum_{k=m}^{n} P(A_k^{(n)}) \\
&= \sum_{k=m}^{n}\sum_{j=k}^{n} \binom{j}{k}(-1)^{j-k}\sigma_j \\
&= \sum_{j=m}^{n}\sigma_j \sum_{k=m}^{j} \binom{j}{k}(-1)^{j-k} \\
&= \sum_{j=m}^{n}\sigma_j \{\sum_{k=0}^{j} \binom{j}{k}(-1)^{j-k} + \sum_{k=0}^{m-1} \binom{j}{k}(-1)^{j-k+1}\} \\
&= \sum_{j=m}^{n}\sigma_j \binom{j-1}{m-1}(-1)^{j+m}.
\end{aligned}$$

■

Prima di riportare due esempi di applicazione delle formule appena stabilite conviene fare qualche osservazione a proposito delle (iii), (v), (vii) e (viii).

La (iii), nel caso particolare in cui E è l'evento certo Ω, fornisce $P(\Omega) = 1 \geq \sum_{i=1}^{\infty} P(E_i)$. Si noti la possibilità del caso in cui la disuguaglianza valga strettamente: un esempio estremo di questa circostanza si è visto nell'Esempio 2.9.

Per quanto riguarda la (v) si può osservare che, nel caso in cui

$$P(E_{i_1} \bigcap ... \bigcap E_{i_j}) = P(E_1 \bigcap ... \bigcap E_j)$$

per ogni $j = 1, ..., n$ e $1 \leq i_1 < ... < i_j \leq n$, essa diviene:

(ix) $P(\bigcup_{i=1}^n E_i) = \sum_{j=1}^n (-1)^{j+1}\binom{n}{j} P(E_1 \bigcap ... \bigcap E_j)$

mentre la (vii) e la (viii) diventano, rispettivamente,

(x) $P(A_m^{(n)}) = \sum_{j=1}^n \binom{j}{m}(-1)^{j+m}\binom{n}{j} P(E_1 \bigcap ... \bigcap E_j)$, per $1 \leq m \leq n$;

(xi) $P(B_m^{(n)}) = \sum_{j=m}^n \binom{j-1}{m-1}(-1)^{j+m}\binom{n}{j} P(E_1 \bigcap ... \bigcap E_j)$, per $1 \leq m \leq n$.

Infine sono di pratica utilità alcune disuguaglianze relative a $P(A_m^{(n)})$ e $P(B_{(m)}^n)$ che coinvolgono le quantità σ_j introdotte nella (5.2):

(xii) $\sigma_m - (m+1)\sigma_{m+1} \leq P(A_m^{(n)}) \leq \sigma_m$,

(xiii) $\sigma_m - m\sigma_{m+1} \leq P(B_m^{(n)}) \leq \sigma_m$.

Per dimostrare la (xii) si consideri la (5.1) che fornisce $P(A_m^{(n)})$ come funzione lineare delle quantità $\sigma_m, ..., \sigma_{m+1}$. Da essa, per ricorrenza, si ottiene:

$$\sigma_j = \sum_{k=j}^{n} \binom{k}{j} P(A_k^{(n)}).$$

Si ha poi

$$P(A_m^{(n)}) - \sigma_m = \sum_{j=m+1}^{n} \binom{j}{m} (-1)^{j-m} \sigma_j < 0$$

e

$$P(A_m^{(n)}) - \sigma_m + (m+1)\sigma_{m+1} = \sum_{j=m+2}^{n} \binom{j}{m} (-1)^{j-m} \sigma_j > 0$$

essendo, per $t = m+1, m+2$,

$$\begin{aligned}
&\sum_{j=t}^{n} \binom{j}{m} (-1)^{j-m} \sigma_j \\
&\quad = \sum_{j=t}^{n} \binom{j}{m} (-1)^{j-m} \sum_{k=j}^{n} \binom{k}{j} P(A_k^{(n)}) \\
&\quad = \sum_{k=t}^{n} P(A_k^{(n)}) \sum_{j=t}^{n} \binom{j}{m} \binom{k}{j} (-1)^{j-m} \\
&\quad = \sum_{k=t}^{n} \binom{k}{m} P(A_k^{(n)}) \sum_{r=t-m}^{k-m} \binom{k-m}{r} (-1)^r \\
&\quad = \begin{cases} -\sum_{k=m+1}^{n} \binom{k}{m} P(A_k^{(n)}) \le 0 & \text{quando } t = m+1, \\ \sum_{k=m+2}^{n} (k-m-1) \binom{k}{m} P(A_k^{(n)}) \ge 0 & \text{quando } t = m+2. \end{cases}
\end{aligned}$$

Allo stesso modo si ottiene la disuguaglianza (xiii).

Esempio 5.2. I primi $n \ge 1$ interi vengono scritti a caso, in un certo ordine. Si può assumere che ognuno dei possibili ordinamenti, che sono in numero di $n!$, abbia la stessa probabilità. Si dirà che vi è una coincidenza se l'intero r occupa la r-esima posizione nell'ordinamento scritto. Qual è la probabilità che l'ordinamento scritto presenti almeno una coincidenza?

Detto E_i, per $i = 1, 2, ..., n$, l'evento che si verifica se una coincidenza accade nella posizione $i-$esima, allora

$$P(E_i) = \frac{(n-1)!}{n!} = \frac{1}{n}$$

mentre

$$P(E_i \cap E_j) = \frac{(n-2)!}{n!}$$

se $1 \le i < j < n$. In generale

$$P(E_{i_1} \cap E_{i_2} \cap .... \cap E_{i_j}) = \frac{(n-j)!}{n!}$$

se $1 \leq i_1 < i_2 < .... < i_j \leq n$. La (ix) fornisce allora

$$P(\bigcup_{i=1}^{n} E_i) = \sum_{j=1}^{n} (-1)^{j+1} \binom{n}{j} \frac{(n-j)!}{n!} = \sum_{j=1}^{n} (-1)^{j+1} \frac{1}{j!},$$

mentre la probabilità di nessuna coincidenza è

$$P(A_0^{(n)}) = 1 - P(\bigcup_{i=1}^{n} E_i).$$

Si osservi come la probabilitâ di almeno una coincidenza venga assai debolmente influenzata da n. Se poi $n \to \infty$, si ha

$$\lim_{n\to\infty} P(\bigcup_{1}^{n} E_i) = \sum_{j=1}^{\infty} (-1)^{j+1} \frac{1}{j!} = 1 - \frac{1}{e} \simeq 0,6321$$

L'esempio precedente si presta a numerose variazioni. Per concludere questa sezione ci limiteremo a considerare la seguente.

Esempio 5.3. Si supponga di avere a disposizione n urne numerate da 1 a n, ed n palline pure numerate da 1 ad n. Le palline vengono inserite a caso nelle urne in modo che ognuna di esse accolga una sola pallina. Se una pallina viene deposta nell'urna che reca la stessa numerazione si avrà una coincidenza. Qual è la probabilità di avere m coincidenze nell'ipotesi che tutte le assegnazioni abbiano la stessa probabilità?

Come nell'esempio precedente, detto E_i l'evento che si verifica se una coincidenza avviene nell'urna i-esima, allora

$$P(A_m^{(n)}) = \sum_{j=m}^{n} \binom{j}{m} (-1)^{j+m} \binom{n}{j} \frac{(n-j)!}{n!} \simeq \frac{1}{m!e}$$

essendo n grande rispetto a m. Lo schema può applicarsi per la soluzione di problemi analoghi.

6. La caratterizzazione di una probabilità definita su un'algebra

La definizione di probabilità adottata nella Sezione 3 si basa sulla condizione di coerenza contenuta nella Definizione 3.1 la quale, se da un punto di vista interpretativo appare - per le ragioni addotte- pienamente soddisfacente, è in generale di difficile verifica nei casi in cui la famiglia degli eventi che si considera presenti complessità appena superiori a quelle degli esempi illustrati nelle sezioni precedenti. Le condizioni (a),(b),(c) della Proposizione 3.3 sarebbero certamente più adatte per questo ufficio, ma, generalmente, esse non possono riguardarsi come equivalenti alla condizione di coerenza. Esse, infatti, costituiscono solo condizioni necessarie e non sono, in generale, sufficienti ad assicurare la coerenza per via della loro formulazione che coinvolge certe operazioni sugli eventi e presuppone la chiusura della famiglia degli stessi rispetto alle predette operazioni. È circostanza fortunata però che, dotando la famiglia degli eventi di conveniente struttura, le condizioni (a), (b) e (c) , opportunamente riformulate, risultino anche sufficienti.

La struttura cui alludiamo è quella di algebra di eventi; punto di partenza per la sua definizione è la considerazione di un insieme di elementi Ω e di una classe $\mathcal{A}$ di suoi sottoinsiemi.

Definizione 6.1. Una classe $\mathcal{A}$ non vuota di sottoinsiemi di Ω è un'algebra se gode delle seguenti proprietà:

(i) $\Omega \in \mathcal{A}$
(ii) Se $E \in \mathcal{A}$, allora l'insieme E^c complementare di E è un elemento di $\mathcal{A}$.
(iii) Se $E_i \in \mathcal{A}$ per $i = 1, ..., n$, allora

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

Osserviamo che dalle (i), (ii) e (iii) discendono anche le seguenti proprietà:

(iv) L'insieme vuoto $\emptyset \in \mathcal{A}$.
(v) Se $E_i \in \mathcal{A}$ per $i = 1, ..., n$, allora

$$\bigcap_{i=1}^{n} E_i \in \mathcal{A}.$$

La (iv) discende dalla (i) e dal fatto che l'insieme vuoto $\emptyset$, o *evento impossibile*, è il complementare dell'insieme Ω, o *evento certo*. Per ottenere la (v) si usino le leggi di de Morgan e le proprietà (ii) e (iii).

Teorema 6.2. *Sia $\mathcal{A}$ un'algebra di eventi. Allora $P : \mathcal{A} \to \mathbf{R}$ è una distribuzione di probabilità su $\mathcal{A}$ se e solo se:*

(a) $P(E) \geq 0$ *per ogni* $E \in \mathcal{A}$;
(b) $P(\Omega) = 1$;
(c) *Per ogni* $n \geq 1$ *e per ogni sottofamiglia finita* $E_1, ..., E_n$ *di eventi di* $\mathcal{A}$ *a due a due incompatibili, si ha*

$$P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i).$$

Dimostrazione. La Proposizione 3.3 contiene la deduzione di (a), (b) e (c) dal principio di coerenza. Dimostriamo allora che (a), (b) e (c) sono sufficienti per la coerenza.

Si consideri una combinazione di scommesse sugli eventi $E_1, E_2, ...E_n \in \mathcal{A}$ d'importi non nulli $c_1, c_2, ..., c_n$. Il guadagno corrispondente è dato da

$$\sum_{k=1}^{s} I_{\omega_k}(\sum_{i=1}^{n} c_i P(E_i) - \sum_{i:\omega_k \subseteq E_i} c_i) = \sum_{k=1}^{s} I_{\omega_k} g_k$$

dove $\omega_1, \omega_2, ..., \omega_s$ indicano gli eventi elementari relativi alla famiglia $\{E_1, E_2, ..., E_n\}$ e, per $k = 1, ..., s$, il guadagno relativo all'evento elementare ω_k è g_k. Dalla definizione di evento elementare e da quella di algebra segue che $\{\omega_k\} \in \mathcal{A}$ per ogni $k = 1, 2, ..., s$. Pertanto $P(\{\omega_k\})$ è assegnata ed in forza di (a), (b) e (c) segue che $P(\{\omega_k\}) \geq 0$, per $k = 1, 2, ..., s$ e inoltre $1 = P(\Omega) = \sum_{k=1}^{s} P(\{\omega_k\})$. In conseguenza di ciò, $\sum_{k=1}^{s} g_k P(\{\omega_k\})$ è una combinazione lineare convessa dei guadagni $g_1, g_2, ..., g_s$ il cui valore è nullo. Infatti,

$$\begin{aligned}
\sum_{k=1}^{s} g_k P(\{\omega_k\}) &= \sum_{k=1}^{s} P(\{\omega_k\})(\sum_{i=1}^{n} c_i P(E_i) - \sum_{i:\omega_k \subseteq E_i} c_i) \\
&= \sum_{i=1}^{n} c_i P(E_i) - \sum_{i=1}^{n} c_i \sum_{\{k:\omega_k \subseteq E_i\}} P(\{\omega_k\}) \\
&= 0.
\end{aligned}$$

L'ultima uguaglianza deriva dalla (c) applicata agli eventi elementari che implicano E_i. Da questo fatto discende che

$$\min\{g_1, g_2, ..., g_s\} \leq 0 \leq \max\{g_1, g_2, ..., g_s\}$$

e dunque, per l'arbitrarietà della famiglia $\{E_1, E_2, ..., E_n\}$, la coerenza di P su $\mathcal{A}$. ■

Da un punto di vista strettamente matematico il teorema precedente afferma che una funzione di insieme, definta su un'algebra, è una distribuzione di probabilità se, e soltanto se, essa soddisfa (a), (b) e (c). Dunque, su un piano sostanziale, la concezione soggettivistica della probabilità, basata sulla coerenza delle valutazioni, lascia piena libertà nella scelta della probabilità P definita su un'algebra purché essa non violi la coerenza, ovvero le condizioni (a), (b) e (c) del teorema. Si potrebbe perciò prescindere dalla condizione di coerenza, che sostanzia queste condizioni, ed incamminarsi sul binario di una impostazione assiomatica. Se non abbiamo optato per questo cammino è perché, in primo luogo, ci è apparsa primaria la necessità di giustificare le proprietà di cui deve godere la probabilità mediante un criterio, quello di coerenza, difficilmente confutabile e di portata affatto generale. In secondo luogo, non abbiamo voluto imporre alla probabilità la condizione di avere un'algebra come dominio.

Concludiamo questa sezione con alcuni esempi di assegnazione di distribuzioni di probabilità su algebre.

Esempio 6.3. Si consideri una partizione finita $\{E_1, E_2, ..., E_n\}$ dell'evento certo Ω e sia $\mathcal{A}$ l'algebra da essa generata, cioè la classe di tutte le unioni - compresa quella vuota - di elementi della partizione. Per definire una distribuzione di probabilità su $\mathcal{A}$ si può procedere nel modo seguente. Si fissa una probabilità su $\{E_1, E_2, ...E_n\}$ che, per le (a),

(b) e (c) del Teorema 6.2, deve soddisfare $P(E_i) \geq 0$, per $i = 1, ..., n$, e $\sum_{i=1}^{n} P(E_i) = 1$.

Ora, se $E \in \mathcal{A}$ allora, per costruzione, esisterà un sottoinsieme $\Gamma(E)$ degli indici $\{1, 2, ..., n\}$ tale che $E = \bigcup_{i \in \Gamma(E)} E_i$. Per la (b) del Teorema 6.2, la probabilità di E deve assegnarsi mediante

$$P(E) = \sum_{i \in \Gamma(E)} P(E_i).$$

Se la famiglia di partenza non fosse una partizione dell'evento certo, allora si può procedere determinando la famiglia degli eventi elementari $\Omega = \{\omega_1, \omega_2, ..., \omega_s\}$ relativa alla famiglia $\{E_1, ..., E_n\}$ a cui si assegneranno le probabilità $P(\{\omega_k\}) \geq 0$ per $k = 1, ..., s$, con la condizione $\sum_{k=1}^{s} P(\{w_k\}) = 1$. Quindi si proseguirà assegnando ad ogni elemento E dell'algebra generata dalla famiglia Ω degli eventi elementari la probabilitâ $P(E) = \sum_{\omega_k \subseteq E} P(\{\omega_k\})$.

Osservazione 6.4. Sia $\{E_k, k \in \Gamma\}$ una partizione dell'evento certo in cui l'insieme degli indici Γ è infinito anche non numerabile. Allora la classe degli elementi della partizione a cui è possibile assegnare probabilità posititva è necessariamente finita o, al più, numerabilmente infinita. Infatti, per $n \geq 1$, si consideri la classe ϕ_n degli elementi della partizione con probabilitâ contenuta nell'intervallo $(\frac{1}{n+1}, \frac{1}{n}]$, ovvero

$$\phi_n = \{E_k, k \in \Gamma : \frac{1}{n+1} < P(E_k) \leq \frac{1}{n}\}.$$

La classe $\bigcup_{n=1}^{\infty} \phi_n$ coincide con quella degli elementi della partizione che hanno probabilità positiva. Ma osserviamo che, qualunque sia $n \geq 1$, la classe ϕ_n contiene, al più, n elementi. Infatti, ϕ_1 non può contenere più di un elemento della partizione; se contenesse due o più elementi della partizione, la somma delle loro probabilità sarebbe maggiore di 1. Analogamente, ϕ_2 non può contenere più di due elementi della partizione, ϕ_3 più di tre etc... Perciò, $\bigcup_{n=1}^{\infty} \phi_n$, unione di una infinità numerabile di insiemi finiti, avrà potenza non superiore a quella del numerabile.

Esempio 6.5. Si consideri una partizione dell'evento certo $\{E_k, k \in \Gamma\}$ in cui l'insieme Γ degli indici è infinito e può avere anche la potenza del continuo. L'algebra $\mathcal{A}$ generata dalla partizione è costituita dalla classe di tutte le unioni finite - compresa quella vuota - di elementi della partizione e delle loro negazioni. Ci proponiamo di indicare un procedimento per assegnare una distribuzione di probabilità su $\mathcal{A}$.

Si tratta, in primo luogo, di assegnare una probabilità agli elementi della partizione di partenza. A tal fine, sia Γ^* un sottoinsieme finito (eventualmente vuoto) o numerabilmente infinito dell'insieme degli indici Γ e si assegni $P(E_i) \geq 0$ per $i \in \Gamma^*$, con la condizione che sia $\sum_{i \in \Gamma^*} P(E_i) \leq 1$. Inoltre, per ogni $i \in \Gamma - \Gamma^*$, si ponga $P(E_i) = 0$. Se $E \in \mathcal{A}$, e quindi $E = \bigcup E_k$, il rispetto di (a)-(c) del Teorema 6.2

impone $P(E) = \sum P(E_k)$. Naturalmente, come abbiamo sottolineato, Γ^* può essere vuoto ed in tal caso $P(E) = 0$ per ogni evento E definito come unione finita di elementi della partizione di partenza.

Per esemplificare, se la partizione di partenza è data dall'insieme degli interi $\{1, 2, 3, ...\}$, allora distribuzioni possibili sull'algebra generata possono essere ottenute a partire da

$$P(\{i\}) = p(1-p)^{i-1}, \text{ per } i = 1, 2, ...$$

essendo $0 < p < 1$. Oppure $P(\{i\}) = 0$ per $i = 1, 2, ...$; in quest'ultimo caso si avrebbe $P(E) = 0$ se E è finito; in particolare, per ogni $N = 1, 2, ...$, si avrebbe $P(\{1, 2, 3, ..., N\}) = 0$ e, di conseguenza, e $P(\{N, N+1, ...\}) = 1$.

Esempio 6.6. Si consideri la classe degli intervalli $(a, b]$ e (c, ∞) ottenuti facendo variare a, b, c in modo che sia $-\infty \leq a \leq b < \infty$ e $-\infty \leq c < \infty$. La più piccola algebra $\mathcal{A}$ che contiene questa classe di sottoinsiemi di $\mathbf{R}$ è costituita dall'insieme delle unioni finite degli intervalli della classe scelti in modo tale da risultare a due a due disgiunti. Per assegnare una distribuzione di probabilità su tale algebra si può procedere definendo una funzione $\bar{F}$ su $\bar{\mathbf{R}} = [-\infty, +\infty]$, il completamento di $\mathbf{R}$ con l'aggiunta dei due punti $-\infty$ e $+\infty$, con le seguenti proprietà:

(i) $\bar{F}(y) \geq \bar{F}(x)$ se $y > x$;
(ii) $\bar{F}(-\infty) = 0, \bar{F}(\infty) = 1$.

Si noti che

$$\lim_{x \to +\infty} \bar{F}(x) \leq \bar{F}(\infty) = 1$$

e

$$\lim_{x \to -\infty} \bar{F}(x) \geq \bar{F}(-\infty) = 0.$$

Per ogni $-\infty \leq a \leq b < \infty$, si definisca

$$P((a, b]) = \bar{F}(b) - \bar{F}(a)$$

e

$$P((a, \infty)) = 1 - \bar{F}(a).$$

Se $E \in \mathcal{A}$, esistono $(a_k, b_k], k = 1, ..., n$, intervalli disgiunti, e aperti anche a destra nel caso $b_k = \infty$, tali che $E = \bigcup_{k=1}^n (a_k, b_k]$; si assegni ad E la probabilità

$$P(E) = \sum_{k=1}^n P((a_k, b_k]) = \sum_{k=1}^n \{\bar{F}(b_k) - \bar{F}(a_k)\}$$

È bene osservare che, malgrado la rappresentazione di E mediante intervalli non sia unica (ad esempio: $E = (1, 3] \cup (5, 9] = (1, 2] \cup (2, 3] \cup$

$(5,7]\cup(7,9]$) la precedente assegnazione della probabilità ad E è unica. Si considerino infatti due rappresentazioni per E, ovvero si assuma che

$$E=\bigcup_{k=1}^{n}(a_k,b_k]=\bigcup_{j=1}^{m}(a'_j,b'_j].$$

Allora, per ogni $k=1,2,...,n$, si potrà scrivere

$$(a_k,b_k]=\bigcup_{j=1}^{m}(a_k,b_k]\bigcap(a'_j,b'_j]$$

e pertanto, essendo l'unione disgiunta,

$$P((a_k,b_k])=\sum_{j=1}^{m}P((a_k,b_k]\bigcap(a'_j,b'_j])$$

da cui

$$\sum_{k=1}^{n}P((a_k,b_k])=\sum_{k=1}^{n}\sum_{j=1}^{m}P((a_k,b_k]\bigcap(a'_j,b'_j]).$$

Allo stesso modo, per $j=1,...,m$, si potrà scrivere

$$(a'_j,b'_j]=\bigcup_{k=1}^{n}(a_k,b_k]\bigcap(a'_j,b'_j]$$

e quindi anche

$$\sum_{j=1}^{m}P((a'_j,b'_j])=\sum_{j=1}^{m}\sum_{k=1}^{n}P((a_k,b_k]\bigcap(a'_j,b'_j]),$$

da cui l'asserita uguaglianza.

È agevole mostrare ora che P, come sopra definita, soddisfa le condizioni (a),(b) e (c) del Teorema 6.2. La (a) e la (b) sono di verifica immediata; per la (c), si considerino gli elementi dell'algebra $E_1,E_2,...,E_n$ a due a due incompatibili. Naturalmente $E_0=\bigcup_{i=1}^{n}E_i$ appartiene all'algebra e pertanto si rappresenterà come unione finita di intervalli disgiunti,

$$E_0=\bigcup_{h=1}^{m}(a_{0_h},b_{0_h}].$$

Inoltre, per $i=1,...,n$,

$$E_i=\bigcup_{j=1}^{m_i}(a_{i_j},b_{i_j}]$$

ove gli intervalli $(a_{i_j},b_{i_j}]$ sono a due a due disgiunti. Ma allora, per ogni $h=1,...,m$, si ha

$$(a_{0_h},b_{0_h}]=\bigcup_{(h)}(a_{i_j},b_{i_j}],$$

essendo $(h) = \{i_j : (a_{i_j}, b_{ij}] \subseteq (a_{0_h}, b_{0_h}]\}$, e quindi

$$P((a_{0_h}, b_{0_h}]) = \sum_{(h)} P((a_{i_j}, b_{i_j}]).$$

Di conseguenza

$$P(E_0) = \sum_{h=1}^{m} P((a_{0_h}, b_{0_h}]) = \sum_{h=1}^{m} \sum_{(h)} P((a_{i_j}, b_{i_j}])$$

da cui, riordinando le somme,

$$P(E_0) = \sum_{i=1}^{n} \sum_{j=1}^{m_i} P((a_{i_j}, b_{i_j}]) = \sum_{i=1}^{n} P(E_i).$$

Nei problemi concreti sarà necessario porre molta cura nella scelta della funzione $\bar{F}$, detta *funzione di ripartizione* su $\bar{R}$, perché, una volta assegnata P con il procedimento appena descritto, da

$$\bar{F}(x) = P((-\infty, x])$$

per ogni $x \in R$, segue che la funzione di ripartizione fornisce i valori della distribuzione di probabilità sulle semirette $(-\infty, x]$.

Esempio 6.7. Si consideri la classe degli intervalli di $\mathbf{R}^n$, $(a, b] = (a_1, b_1] \times ... \times (a_n, b_n] = \times_{k=1}^{n} (a_k, b_k]$ con $-\infty \leq a_k \leq b_k \leq \infty$ con l'intesa che gli intervalli sono aperti anche a destra nel caso in cui $b_k = \infty$. Come nell'esempio precedente, la classe delle unioni finite e disgiunte di tali intervalli è un'algebra. Per probabilizzarla si può procedere assegnando una funzione di ripartizione $\bar{F}$ su $\bar{\mathbf{R}}^n$, l'insieme di tutte le n-uple $(x_1, ..., x_n)$ con $x_i \in \bar{\mathbf{R}}$ per $i = 1, ..., n$, con le proprietà:

(i) fissati due punti $(x_1, x_2, ...x_n)$ e $(y_1, y_2, ...y_n)$ di $\bar{R}^n$, con $x_i \leq y_i$ per $i = 1, ..., n$, si abbia

$$\triangle_{x_1}^{y_1} ... \triangle_{x_n}^{y_n} \bar{F}(z_1, z_2, ...z_n) = \sum (-1)^{\sum_1^n \varepsilon_k} \bar{F}(z_1, z_2, ...z_n) \geq 0$$

con $z_k = \varepsilon_k x_k + (1 - \varepsilon_k) y_k$ ed ε_k che assumono i valori 0 oppure 1 indipendentemente l'uno dall'altro;

(ii) $\bar{F}(\infty, ..., \infty) = 1$ e $\bar{F}(x_1, x_2, ..., x_n) = 0$ se almeno una componente x_i risulta eguale ad $-\infty$.

Sugli intervalli $(a, b]$ si definisca la funzione

$$P((a, b]) = \triangle_{a_1}^{b_1} ... \triangle_{a_n}^{b_n} \bar{F}(z_1, z_2, ...z_n)$$

mentre all'elemento $E = \bigcup_1^n (a_k, b_k]$ dell'algebra si assegni $P(E) = \sum_{k=1}^{n} ((a_k, b_k])$. Così come abbiamo proceduto nel caso di intervalli di $\mathbf{R}$ nell'Esempio 6.3, si può dimostrare che l'assegnazione proposta soddisfa le condizioni (a), (b) e (c) del Teorema 6.2 e dunque essa costituisce una probabilità sull'algebra generata dagli intervalli di $\mathbf{R}^n$. Si noti, anche in questo caso, come

$$\bar{F}(x_1, x_2, ..., x_n) = P((-\infty, x_1] \times (-\infty, x_2] \times .... \times (-\infty, x_n]).$$

A chiusura della sezione osserviamo che gli Esempi 6.6 e 6.7 costituiscono schemi sufficientemente generali di assegnazione di distribuzioni di probabilità su algebre applicabili alla valutazione di probabilità di eventi relativi, rispettivamente, a numeri aleatori e vettori aleatori.

7. Probabilitâ e frequenza

Giunti a questo punto una domanda che sorge spontanea è la seguente: esistono collegamenti naturali tra la probabilità soggettivamente valutata e la frequenza? In questa sezione giustificheremo, in termini estremamente elementari, questo collegamento; una più profonda discussione sull'intimo legame tra probabilità e frequenza farebbe invece riferimento ad uno schema generale, quello della scambiabilità, introdotto a questo scopo dal de Finetti. Vale comunque la pena di notare immediatamente che questo nesso tra probabilità valutata soggettivamente e frequenza dipende da atteggiamenti che sono sempre soggettivi.

Sia $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ lo spazio dei risultati incompatibili di un certo esperimento, ovvero lo spazio degli eventi elementari relativi ad una classe finita di eventi, e sia $p(\omega_i) \geq 0$ la probabilità del i-esimo elemento di Ω. Evidentemente si ha che $\sum_{i=1}^{N} p(\omega_i) = 1$. Supponiamo ora di considerare $n \geq 1$ eventi o sottoinsiemi di Ω, $E_1, ..., E_n$, con probabilità $P(E_i), i = 1, 2, ..., n$, e sia π_j la probabilità che esattamente j eventi E_i si verifichino, con $j = 0, 1, ..., n$.

Definiamo

$$I_{ik} = \begin{cases} 1 & \text{se } \omega_k \in E_i, \\ 0 & \text{se } \omega_k \notin E_i; \end{cases}$$

allora

$$P(E_i) = \sum_{\omega_k \in E_i} p(\omega_k) = \sum_{k=1}^{N} I_{ik} p(\omega_k)$$

Indichiamo ora con $j(\omega_k)$ il numero degli eventi E_i implicati da ω_k, cioè che si manifestano se si manifesta ω_k,

$$j(\omega_k) = \sum_{i=1}^{n} I_{ik}.$$

Segue allora che "mediamente" il numero di eventi E_i che si verificano è dato da:

$$\sum_{j=0}^{n} j\pi_j = \sum_{k=1}^{N} j(\omega_k) p(\omega_k),$$

cioè:

$$\begin{aligned}\sum_{j=0}^{n} j\pi_j &= \sum_{k=1}^{N} j(\omega_k)p(\omega_k) = \sum_{k=1}^{N}\left(\sum_{i=1}^{n} I_{ik}\right)p(\omega_k) = \\ &= \sum_{i=1}^{n}\sum_{k=1}^{N} I_{ik}p(\omega_k) = \sum_{i=1}^{n} P(E_i)\end{aligned}$$

da cui, dividendo per n entrambi i membri, si ha

$$\frac{1}{n}\sum_{i=1}^{n} P(E_i) = \sum_{j=0}^{n}\frac{j}{n}\pi_j.$$

Se il soggetto valuta $\pi_k = 1$, si avrà

$$\frac{1}{n}\sum_{i=1}^{n} P(E_i) = \frac{k}{n}$$

Inoltre se $P(E_i) = p$ per $i = 1, ..., n$ si ha

$$P(E_i) = \frac{k}{n}$$

per $i = 1, ..., n$. Cioè: la probabilità del generico evento E_i coincide con la frequenza relativa del numero di eventi $E_1, ..., E_n$ che il soggetto ritiene si verifichi con certezza.

Se le probabilità $P(E_i)$ fossero diverse, la conoscenza della frequenza attesa fornirebbe una base per i necessari aggiustamenti che consentono di giungere alla valutazione di $P(E_i)$.

Il procedimento descritto comprende tre fasi:

(i) la scelta di una classe di eventi contenenti quello che si vuole considerare;
(ii) la previsione della frequenza;
(iii) il confronto tra le probabilità dei singoli eventi e la previsione precedente.

È facile individuare, nella prima e nella terza delle fasi del procedimento, gli elementi soggettivi della valutazione. La prima fase, consistente nella scelta della classe di eventi, contiene un notevole elemento di arbitrarietà. Il fatto di considerare eventi "simili" può facilitare il confronto tra frequenza e probabilità, potendo ritenersi le $P(E_i)$ circa uguali, ma rimane una cosa sempre incerta e soggetta ad un giudizio di natura soggettiva. Per quanto concerne la terza fase le difficoltà sono strettamente collegate alla prima, nel senso che una scelta opportuna della classe di eventi può condurre il soggetto interessato a ritenere di valore costante le $P(E_i)$: si tratta comunque di un giudizio di natura soggettiva. La seconda fase, quella della previsione della frequenza, è accompagnata dall'ipotesi che la frequenza resti all'incirca costante all'aumentare di n. Cruciale è la giustificazione di questa ipotesi nell'ambito dell'impostazione soggettiva; questa problematica viene trattata approfonditamente solo introducendo il concetto di scambiabilità.

Prima di terminare questo sezione facciamo tuttavia osservare che si può giustificare immediatamente, in termini soggettivi, la "definizione" classica di probabilità fondata sul rapporto del numero dei casi favorevoli sul numero dei casi possibili. In effetti se si ha una classe completa di n eventi incompatibili (ed esaustivi) e se essi sono giudicati ugualmente probabili, in virtù della (c) della Proposizione 3.3, ciascuno di essi avrà probabilità $p = \frac{1}{n}$ e l'unione di m tra essi avrà probabilità m/n. Questo criterio di valutazione è applicabile solo se l'individuo giudica ugualmente probabili i casi considerati e ciò è dovuto ancora ad un giudizio soggettivo, che le abituali considerazioni di simmetria (tipiche dei giochi d'azzardo) possono spiegare, ma che non possono assolutamente trasformare in qualcosa di oggettivo.

8. L'impostazione assiomatica di Kolmogorov.

Nei paragrafi precedenti abbiamo esposto sommariamente i primi elementi della teoria delle probabilità dal punto di vista soggettivo. Presentiamo ora brevemente l'impostazione assiomatica dovuta a Kolmogorov ed esposta nella fondamentale monografia del 1933. La convinzione di fondo di Kolmogorov si regge sul fatto che la teoria della probabilità può essere sviluppata assiomaticamente prescindendo dalle questioni inerenti il significato di probabilità. Il matematico sovietico mostrò che la teoria della probabilità si occupa di entità astratte che non necessitano, almeno a livello di sviluppo della teoria, di alcuna interpretazione. Sfruttando l'analogia tra probabilità e misura, tra speranza matematica e integrale nel senso di Lebesgue, l'assiomatica di Kolmogorov si discosta da quella che sarebbe naturale dedurre dalle proposizioni dei precedenti paragrafi. Il punto di partenza è la considerazione di un insieme di elementi Ω, detto insieme dei risultati dell'esperimento, e di una classe di sottoinsiemi (eventi) di Ω costituenti una σ-algebra e indicata con $\mathcal{F}$.

Definizione 8.1. Un'algebra $\mathcal{F}$ di sottoinsiemi di Ω è detta σ-algebra se, per una arbitraria successione di insiemi $\{E_i, i = 1, 2, ..., n, ...\} \in \mathcal{F}$ si ha
$$\bigcup_{i=1}^{\infty} E_i \in \mathcal{F}.$$

La definizione di probabilità P viene fissata da Kolmogorov mediante i seguenti assiomi:

(A.1) ad ogni elemento $E \in \mathcal{F}$ si associa un numero reale $P(E)$ compreso tra 0 e 1;

(A.2) per l'evento certo Ω si ha $P(\Omega) = 1$;

(A.3) per ogni successione $E_1, E_2, \ldots$ di elementi di $\mathcal{F}$ a due a due incompatibili si ha

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

La probabilità così definita è detta *completamente additiva*; ovviamente ogni probabilità di questo tipo è finitamente additiva, ovvero vale la (c) del Teorema 6.2, ed è quindi coerente nel senso di de Finetti, ma, in generale, non vale il contrario. È immediato infatti verificare che l'additività completa non può essere dedotta da quella finita con un passaggio al limite per n. Infatti, se $E_1, ..., E_n, ...$ è una successione di elementi di $\mathcal{F}$ a due a due incompatibili, l'additività finita di P porge

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

per ogni $n \geq 1$. Poiché il membro di sinistra della precedente uguaglianza è sempre minore o eguale ad 1, il membro di destra deve convergere quando $n \to \infty$, cioè

$$\lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} E_i\right) = \lim_{n\to\infty} \sum_{i=1}^{n} P(E_i) = \sum_{i=1}^{\infty} P(E_i)$$

e dunque l'additività completa richiederebbe

$$\sum_{i=1}^{\infty} P(E_i) = P\left(\bigcup_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} E_i\right)$$

ovvero

$$P\left(\lim_{n\to\infty} \bigcup_{i=1}^{n} E_i\right) = \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} E_i\right).$$

Ma questa operazione, come è ben noto, non sempre è legittima sicché la proprietà (A.3), che possiamo chiamare di continuità, è appunto avanzata a questo scopo.

Vale infine la pena di osservare che, se è vero che l'assunzione di completa additività permette alla teoria della probabilità di mutuare agili strumenti dalla teoria della misura, primo tra tutti l'integrale secondo Lebesgue, essa d'altra parte costituisce un fattore di disturbo nel senso che la probabilità, così intesa, non può, in generale, essere estesa ad ogni sottoinsieme di Ω.

9. Bibliografia

(1) de Finetti, B. (1930), Funzione caratteristica di un fenomeno aleatorio. *Atti Reale Accademia dei Lincei*, serie VI, Mem.4, pp. 86-133.

(2) de Finetti, B. (1931), Sul significato soggettivo della probabilità, *Fund. Math.*, 17, 298-329.

(3) DE FINETTI, B. (1937), La prévision, ses lois logiques, ses sources subjectives, *Ann. Inst. H. Poincaré*, 7, 1-68.
(4) DE FINETTI, B. (1949), Sull'impostazione assiomatica del calcolo delle probabilità, *Annali Triestini dell'Università di Trieste*, 19, 29-81, Trad. inglese in de Finetti (1972): *Probability, Induction and Statistics*, Wiley.
(5) DE FINETTI, B. (1970),*Teoria della probabilità*, Einaudi, Torino.
(6) DE FINETTI, B. (1974), Teoria della probabilità, in *Repertorio di matematiche (III)*, a cura di M.Villa, CEDAM, Padova.
(7) KOLMOGOROV, A. N. (1950), *Foundations of the theory of probability*, Chelsea Publ. co., New York.

DONATO MICHELE CIFARELLI
ISTITUTO DI METODI QUANTITATIVI
UNIVERSITÀ "L. BOCCONI"
20100 MILANO - ITALY

PIETRO MULIERE
ISTITUTO DI METODI QUANTITATIVI
UNIVERSITÀ "L. BOCCONI"
20100 MILANO - ITALY
e-mail: pietro.muliere@uni-bocconi.it

PIERCESARE SECCHI
DIPARTIMENTO DI MATEMATICA "F. BRIOSCHI"
POLITECNICO DI MILANO
PIAZZA LEONARDO DA VINCI 32
20123 MILANO - ITALY
email: secchi@mate.polimi.it