Capitolo 10

PROBABILITÀ CONCENTRATE E DIFFUSE

Con l'eccezione del § 10.6 – definizione alternativa di probabilità coerente – questo capitolo può essere visto come una rilettura con approfondimenti dei *Capitoli* 4 e 5, nell'ottica delle probabilità coerenti. Il discorso è fatto perciò con riferimento a probabilità definite su una partizione $\mathbb{P}$. Le probabilità *concentrate* – costituenti di probabilità positiva di somma 1 (10.1.2*b*) – si ottengono distribuendo *tutta* la probabilità a pezzetti sui costituenti di $\mathbb{P}$ e in modo arbitrario, sono σ-additive su $\mathcal{A}_L(\mathbb{P})$ (*Teorema* 10.2.1) e sono allora probabilità anche in senso classico. Se $\mathbb{P}$ è infinita esistono probabilità *diffuse* – tutti i costituenti sono di probabilità nulla (10.1.2*b*) – (*Teorema* 10.3.1). Contrariamente al caso classico ciò vale anche se $\mathbb{P}$ è numerabile (n° 10.5.2). Le probabilità coerenti su $\mathcal{A}_L(\mathbb{P})$ sono tutte misture di probabilità concentrate e diffuse (*Teorema* 10.4.2). Nel n° 10.3.2 viene introdotta la nozione di *distribuzione asintotica* – probabilità diffusa su un sottoinsieme numerabile di $\mathbb{P}$, ottenuta come limite di distribuzioni uniformi su segmenti finiti invadenti il sottoinsieme – usata in particolare per interpretare la *scelta a caso* nel numerabile (10.5.2 *Esempio*). Nel § 10.6, già citato, viene introdotta la nozione di *penalità* di $(E_1, \dots, E_n)$ su $(x_1, \dots, x_n)$ – quadrato della distanza del punto aleatorio $(|E_1|, \dots, |E_n|)$ dal punto certo $(x_1, \dots, x_n)$ (10.6.1 *Definizione*) – e quella corrispondente di *probabilità* – valutazione ammissibile se non si può diminuire la penalità di $(E_1, \dots, E_n)$ in qualche caso senza aumentarla in altri (*Definizione* 10.6.3) –. Si prova che gli insiemi delle probabilità coerenti secondo questo modello e quello delle scommesse sono uguali (*Teorema* 10.6.4) e che le valutazioni sui singoli eventi coincidono se si interpreta la penalità come perdita monetaria (*Teorema* 10.6.6).

10.1 Probabilità su partizioni

In 5.3.1 *Terminologia*, interpretando la probabilità come massa unitaria abbiamo introdotto le nozioni intuitive di probabilità concentrata e di probabilità diffusa. Daremo qui avanti una definizione precisa delle due nozioni in termini di probabilità coerenti. Occorre però premettere la seguente proposizione

10.1.1 Proposizione.

Gli eventi elementari di probabilità positiva di ogni partizione dell'evento certo sono al più un numerabile e la loro somma non può superare 1.

DIMOSTRAZIONE.

La proposizione sussiste anche nell'interpretazione della probabilità come massa unitaria. L'enunciato della sua prima parte coincide con quello della *Proposizione* 5.1.1. Per rendere valida la sua dimostrazione anche nell'attuale contesto bisogna dimostrare che se P è una probabilità coerente, allora al più n eventi elementari possono avere probabilità superiore a $1/(n+1)$ – condizione ovvia nell'interpretazione della probabilità come massa *unitaria* – o equivalentemente che la somma delle probabilità di un numero finito di eventi elementari non supera 1.
Sia allora $\mathbb{P}$ una partizione dell'evento certo, P *una* probabilità coerente su $\mathbb{P}$ e $\omega_1, \dots, \omega_n \in \mathbb{P}$. Da $\bigvee_{i=1}^{n} \omega_i \Rightarrow \Omega$ segue $P(\bigvee_{i=1}^{n} \omega_i) = \sum_{i=1}^{n} P(\omega_i) \le P(\Omega) = 1$ (9.3.3f,h,a) e con ciò la prima parte dell'enunciato è provata.

Se non esistono eventi elementari di probabilità positiva la seconda parte è banale ed è conseguenza della disuguaglianza appena dimostrata se l'insieme degli eventi elementari di probabilità positiva è finito. Altrimenti, posto che tali eventi siano $\omega_1, \dots, \omega_n, \dots$, si ha $\sum_{i=1}^{n} P(\omega_i) \le 1$ per ogni n e da qui segue l'asserto. ■

Ciò premesso, diamo la seguente definizione.

10.1.2 Definizione. Probabilità concentrata e probabilità diffusa.

Siano $\mathbb{P}$ *una partizione di* Ω, P *una probabilità su* $\mathbb{P}$, γ *la somma delle probabilità degli eventi elementari di probabilità positiva* ($0 \le \gamma \le 1$, *Proposizione* 10.1.1). *Diremo che:*

a) γ *è la* ***parte concentrata*** *e* $1-\gamma$ *la* ***parte diffusa*** *di* P *su* $\mathbb{P}$;

b) P *è una probabilità* ***concentrata*** *su* $\mathbb{P}$ *se* $\gamma = 1$, ***diffusa*** *se* $\gamma = 0$ *(se tutti gli eventi elementari di* $\mathbb{P}$ *hanno probabilità nulla),* ***parzialmente concentrata*** *o, equivalentemente,* ***parzialmente diffusa*** *se* $0 < \gamma < 1$.

Inoltre, se P *è una probabilità su* $\mathbb{P} \cup \mathscr{E}$, $\mathscr{E} \subset \mathscr{A}_{\mathrm{L}}(\mathbb{P})$, *allora diremo:*

c) ***parte concentrata*** *di* P *su* $E \in \mathscr{E}$ *(relativamente a* $\mathbb{P}$*) la somma delle* *probabilità dei componenti di E di probabilità positiva e* ***parte diffusa*** *il complemento a $P(E)$ di tale somma;*

d) *P è* ***concentrata*** *(****diffusa****,* ***parzialmente concentrata*** *o* ***parzialmente diffusa****) se è tale la sua restrizione su $\mathbb{P}$.*

In vista di studiare le probabilità coerenti su partizioni e loro prolungamenti coerenti, è utile premettere il seguente lemma.

10.1.3 Lemma.

Sia $\mathbb{P}$ una partizione di Ω, P una probabilità su $\mathbb{P}$, $\mathscr{E}^+ = \{\omega \in \mathbb{P} : P(\omega) > 0\}$. La valutazione di probabilità $P(E)$ di un qualunque evento E è soggetta alle seguenti limitazioni (condizioni necessarie per la coerenza):

$$\sum_{\omega \in \mathscr{E}^+} P(\omega)\, \delta(\omega \Rightarrow E) \le P(E) \le 1 - \sum_{\omega \in \mathscr{E}^+} P(\omega)\, \delta(\omega \Rightarrow \bar{E}). \qquad (20)$$

DIMOSTRAZIONE.

Se $\mathscr{E}^+$ è vuoto la disuguaglianza diventa $0 \le P(E) \le 1$, cioè la 9.3.3*b*, da cui segue l'asserto.

Se $\mathscr{E}^+$ non è vuoto, si possono numerare gli $\omega \in \mathscr{E}^+$ (*Proposizione* 10.1.1) e porre $\mathscr{E}^+ = \{\omega_i : i \in I\}$, con $I = \{1, \dots, n\}$ o $I = \mathbb{N}$. Per ogni $s \in I$ poniamo $E_s = \bigvee_{i \le s} \{\omega_i : \omega_i \Rightarrow E\}$. Riesce ovviamente $E_s \Rightarrow E$, da cui si ricava (9.3.3*f*, *h*) $P(E_s) = \sum_{i=1}^{s} P(\omega_i)\, \delta(\omega_i \Rightarrow E) \le P(E)$. Si ottiene allora la limitazione inferiore $\sum_{\omega \in \mathscr{E}^+} P(\omega)\, \delta(\omega \Rightarrow E)$ dell'enunciato del teorema ponendo $s = n$ se I è finito e facendo tendere s a $+\infty$ se $I = \mathbb{N}$ (la serie $\sum_{i=1}^{\infty} P(\omega_i)\, \delta(\omega_i \Rightarrow E)$ converge, perché serie a termini non negativi con le ridotte $P(E_s)$ limitate superiormente da 1).
Per quanto sin qui dimostrato, si ha poi $\sum_{\omega \in \mathscr{E}^+} P(\omega)\, \delta(\omega \Rightarrow \bar{E}) \le P(\bar{E})$. Segue allora la limitazione superiore, perché deve essere $P(\bar{E}) = 1 - P(E)$ in forza della 9.3.3*d*. ∎

NOTA.

Come detto esplicitamente nell'enunciato del *Lemma*, le limitazioni previste dalla (20) per $P(E)$ sono condizioni necessarie per la coerenza, e determinano perciò l'intervallo entro cui *devono* cadere le valutazioni coerenti per l'evento E. Non è detto però che *tutti* i valori di quell'intervallo siano valutazioni coerenti, ovvero che le condizioni (20) siano anche sufficienti.

Ad esempio, se P è *diffusa*, le limitazioni della (20) diventano 0 e 1, rispettivamente. Ciò visto, se E è tale che uno dei due insiemi $E \wedge \mathbb{P}$ o $\bar{E} \wedge \mathbb{P}$ sia *finito*, poiché P deve soddisfare le 9.3.3*d,f* la valutazione per E diventa obbligata: $P(E) = 0$ se è finito $E \wedge \mathbb{P}$ e $P(E) = 1$ se è finito $\bar{E} \wedge \mathbb{P}$. Pertanto, in questi due casi le condizioni (20) non sono sufficienti per la coerenza. Esse lo sono, invece, se $E \wedge \mathbb{P}$ e $\bar{E} \wedge \mathbb{P}$ sono entrambi infiniti. Vedremo infatti in 10.4.3 *Teorema* che in questa ipotesi è coerente scegliere per $P(E)$ un valore arbitrario in $[0, 1]$.

Un caso importante in cui la (20) è sufficiente per la coerenza si ha quando E è *logicamente dipendente* da $\mathbb{P}$ e P è una probabilità *concentrata*. In tal caso le due limitazioni coincidono e la probabilità è univocamente determinata. Il risultato è il contenuto della prossima proposizione.

10.1.4 Proposizione.

Siano P una probabilità concentrata sulla partizione $\mathbb{P}$ ed $\mathscr{E}^+$ l'insieme dei suoi eventi elementari di probabilità positiva. Allora esiste unico il prolungamento di P su $\mathscr{A}_L(\mathbb{P})$ ed è dato da $P(\cdot) = \sum_{\omega \in \mathscr{E}^+} P(\omega)\, \delta(\omega \Rightarrow \cdot)$.

DIMOSTRAZIONE.

Sia $E \in \mathscr{A}_L(\mathbb{P})$. In queste ipotesi ogni $\omega \in \mathscr{E}^+$ implica E oppure $\bar{E}$. Allora si ha $\sum_{\omega \in \mathscr{E}^+} P(\omega)\, \delta(\omega \Rightarrow E) + \sum_{\omega \in \mathscr{E}^+} P(\omega)\, \delta(\omega \Rightarrow \bar{E}) = 1$, da cui si ricava $1 - \sum_{\omega \in \mathscr{E}^+} P(\omega)\, \delta(\omega \Rightarrow \bar{E}) = \sum_{\omega \in \mathscr{E}^+} P(\omega)\, \delta(\omega \Rightarrow E)$. Le due limitazioni perciò coincidono e bisogna allora scegliere $P(E) = \sum_{\omega \in \mathscr{E}^+} P(\omega)\, \delta(\omega \Rightarrow E)$, che è una valutazione coerente per E (*Complemento* 9.5.2).

10.2 Probabilità concentrate

Nella *Proposizione* 10.1.4 del precedente paragrafo abbiamo visto che se P è una probabilità concentrata sulla partizione $\mathbb{P}$, essa ammette un solo prolungamento coerente su $\mathscr{A}_L(\mathbb{P})$: quello che si ottiene attribuendo a ogni suo evento E la somma delle probabilità dei relativi componenti di probabilità positiva. Quello che la proposizione citata non dice – né lo fa il

Lemma 10.1.3 da cui essa deriva – è se è coerente *scegliere arbitrariamente* un sottoinsieme non vuoto di $\mathbb{P}$ e *distribuire a piacere* l'intera probabilità a "pezzetti" sui suoi eventi elementari, come abbiamo fatto nel dare le definizioni intuitive di probabilità in ambiente finito (*Definizione* 4.2.1) e in quello numerabile (*Definizione* 5.2.1). Risponde positivamente a questo quesito il prossimo teorema.

10.2.1 Teorema. Caratterizzazione delle probabilità concentrate

Siano $\mathbb{P}$ una partizione di Ω, $(p_i)_{i\in I}$ una sequenza – se $I=\{1,\ldots,n\}$ – o una successione – se $I=\mathbb{N}$ – di numeri non negativi di somma 1, $\{\omega_i \in \mathbb{P}^\phi : i\in I\}$ *un sottoinsieme non vuoto di eventi elementari. L'applicazione $P(\cdot)=\sum_{i\in I} p_i\,\delta(\omega_i \Rightarrow \cdot)$, di dominio $\mathcal{A}_L(\mathbb{P})$ è una probabilità σ-additiva, la sua restrizione $P_{|\mathbb{P}}$ è una probabilità concentrata su $\mathbb{P}$ e P è l'unico prolungamento coerente di $P_{|\mathbb{P}}$ su $\mathcal{A}_L(\mathbb{P})$.*

DIMOSTRAZIONE.

L'applicazione P è definita sull'algebra (e σ-algebra) $\mathcal{A}_L(\mathbb{P})$. Proveremo che essa è una probabilità mostrando che soddisfa le condizioni *a), b), c)* del *Teorema* 9.4.2.
La *a)* è immediata: si ha $P(\Omega)=\sum_{i\in I} p_i\,\delta(\omega_i \Rightarrow \Omega)=\sum_{i\in I} p_i = 1$.
La *b)* è ovvia.
Resta da provare la *c)*. Tenendo conto che le somme delle serie convergenti a termini non negativi soddisfano le stesse proprietà delle somme ordinarie, le prove dell'additività e della σ-additività di P si possono ottenere seguendo una medesima linea dimostrativa. Basta indicare per questo con J l'insieme dei numeri naturali o un loro segmento iniziale e considerare un insieme di eventi a due a due incompatibili $\{E_h : h\in J\}\subset \mathcal{A}_L(\mathbb{P})$. Posto allora $E=\bigvee_{h\in J} E_h$ e tenuto conto che nelle nostre ipotesi riesce $\delta(\omega\Rightarrow E)=\sum_{h\in J}\delta(\omega\Rightarrow E_h)$ (9.4.1 *Proposizione, b*), si ha:

$$P\left(\bigvee_{h\in J} E_h\right)=P(E)=\sum_{i\in I} p_i\,\delta(\omega_i\Rightarrow E)=\sum_{i\in I} p_i \sum_{h\in J}\delta(\omega_i\Rightarrow E_h)=$$
$$\sum_{h\in J}\sum_{i\in I} p_i\,\delta(\omega_i\Rightarrow E_h)=\sum_{h\in J} P(E_h).$$

È così provato che P è una probabilità σ-additiva di dominio $\mathcal{A}_L(\mathbb{P})$. È allora una probabilità anche la sua restrizione $P_{|\mathbb{P}}$ (9.2.4 *Complemento*). Resta allora a questo punto da dimostrare che $P_{|\mathbb{P}}$ è concentrata e che P è il suo unico prolungamento coerente su $\mathcal{A}_L(\mathbb{P})$.
Che $P_{|\mathbb{P}}$ sia concentrata lo si vede osservando che per ogni $\omega \in \mathbb{P}$, riesce $\omega_i \Rightarrow \omega$ se e solo se $\omega = \omega_i$. Allora $P_{|\mathbb{P}}(\omega) = P(\omega) = \sum_{i \in I} p_i \, \delta(\omega_i \Rightarrow \omega) = 0$ se $\omega \in \mathbb{P} - \{\omega_i : i \in I\}$, $P_{|\mathbb{P}}(\omega_i) = p_i$ se $i \in I$. Segue $\sum_{i \in I} P_{|\mathbb{P}}(\omega_i) = 1$ e quindi la $P_{|\mathbb{P}}$ è concentrata come si voleva.
Che P sia l'unico prolungamento coerente di $P_{|\mathbb{P}}$ su $\mathcal{A}_L(\mathbb{P})$ è conseguenza della *Proposizione* 10.1.4. ■

Ha interesse per il seguito mettere in evidenza la seguente proposizione.

Proposizione.

Siano $\mathbb{P}$ e $\mathbb{P}'$ partizioni, $\mathbb{P}'$ finita, P una probabilità concentrata su $\mathbb{P}$, $\mathcal{E}^+ \subset \mathbb{P}$ l'insieme degli eventi elementari di probabilità positiva, P' un prolungamento coerente di P su $\mathbb{P} \cup (\mathbb{P} \wedge \mathbb{P}')$. Allora la restrizione di P' su $\mathbb{P} \wedge \mathbb{P}'$ è una probabilità concentrata.

Dimostrazione.

Poiché P' è un prolungamento di P, per ogni $\omega \in \mathbb{P}$ si ha $P'(\omega) = P(\omega)$. Ciò premesso sia $\mathbb{P}' = \{e_1, \dots, e_n\}$. Per ogni $\omega \in \mathbb{P}$ riesce allora $\omega = \bigvee_{i=1}^{n} (\omega \wedge e_i)$ e quindi $P(\omega) = P'(\omega) = \sum_{i=1}^{n} P'(\omega \wedge e_i)$, la seconda uguaglianza valendo perché P' è coerente e perciò additiva.
Ora, se $\omega \in \mathbb{P} - \mathcal{E}^+$ si ha $P(\omega) = 0$ e quindi $P'(\omega \wedge e_1) = \dots = P'(\omega \wedge e_n) = 0$. Pertanto, gli eventi elementari di $\mathbb{P} \wedge \mathbb{P}'$ di probabilità P' positiva appartengono all'insieme numerabile $\mathcal{E}^+ \wedge \mathbb{P}'$. Sommando le probabilità degli eventi di questo insieme – la somma è una serie a termini non negativi – si ottiene $\sum_{\omega \in \mathcal{E}^+} \sum_{i=1}^{n} P'(\omega \wedge e_i) = \sum_{\omega \in \mathcal{E}^+} P'(\omega) = 1$, da cui segue che P' è concentrata su $\mathbb{P} \wedge \mathbb{P}'$ come volevasi. ■

Nota.

Ripercorrendo la dimostrazione si vede che affinché il risultato sussista non occorre che $\mathbb{P}'$ sia finito. Basta in effetti che sia finito l'insieme $\omega \wedge \mathbb{P}'$ per ogni $\omega \in \mathcal{E}^+$. Al di fuori di questa ipotesi – un po' più generale – si ha che se $\mathbb{P}'$ è infinita, allora tutte le probabilità P concertate su $\mathbb{P}$ ammettono prolungamenti P' *non concentrati* su $\mathbb{P} \wedge \mathbb{P}'$. Si veda in proposito 10.4.1 *Nota*.

10.3 Probabilità diffuse

Il precedente *Teorema* 10.2.1 mostra che *esistono infinite* probabilità concentrate su *ogni partizione* non nulla. La probabilità diffusa su una partizione, invece, o *non esiste,* quando la partizione è finita (ovvio), altrimenti è *unica* – l'applicazione identicamente nulla –. Sussiste infatti il seguente teorema.

10.3.1 Teorema.

Sia $\mathbb{P}$ *una partizione di* Ω *di cardinalità infinita. Allora l'applicazione identicamente nulla è una probabilità diffusa su* $\mathbb{P}$.

DIMOSTRAZIONE.
Sia $\{\omega_1, \dots, \omega_n, \dots\}$ un insieme numerabile di eventi elementari di $\mathbb{P}$. Indichiamo con P_n la probabilità concentrata di dominio $\mathbb{P}$ che si ottiene equidistribuendo tutta la probabilità sugli n eventi elementari $\omega_1, \dots, \omega_n$. Per ogni $\omega \in \mathbb{P}$ allora si ha $P_n(\omega) = 1/n$ se $\omega = \omega_i$, $i = 1, \dots, n$ e $P_n(\omega) = 0$ altrimenti. Per *Teorema* 10.2.1 la probabilità P_n si prolunga in modo unico su $\mathcal{A}_L(\mathbb{P})$ ponendo $P_n(E) = (1/n)\sum_{i=1}^{n} \delta(\omega_i \Rightarrow E)$ per ogni suo evento E. Al variare di n in $\mathbb{N}$ si ottiene così una successione di probabilità di comune dominio $\mathcal{A}_L(\mathbb{P})$. Indicato con $\mathcal{D} \subset \mathcal{A}_L(\mathbb{P})$ il suo insieme di convergenza, per ogni $E \in \mathcal{D}$ poniamo $P(E) = lim_{n \to +\infty} P_n(E)$. L'applicazione P è allora una probabilità coerente (9.3.2 *Corollario*) e diffusa, come andiamo a verificare. Infatti, scelto $\omega \in \mathbb{P}^{\phi}$, se $\omega \notin \{\omega_1, \dots, \omega_n, \dots\}$, riesce $\delta(\omega_i \Rightarrow \omega) = 0$ per ogni $i \in \mathbb{N}$, perciò $P_n(\omega) = 0$ per ogni n, e quindi $P(\omega) = 0$. Altrimenti, posto che sia $\omega = \omega_h$, per $n \geq h$ si ha $P_n(\omega_h) = 1/n$, da cui passando al limite si ricava ancora $P(\omega) = P(\omega_h) = 0$. ■

10.3.2 Complemento. *Distribuzioni asintotiche.*

Ai fini della dimostrazione del *Teorema* si sarebbe potuto definire l'applicazione P solo su $\mathbb{P}$ anziché su un soprainsieme $\mathcal{D} \subset \mathcal{A}_L(\mathbb{P})$. Si è preferito fare il discorso nella forma estesa perché in questo modo si ottiene contestualmente una distribuzione di probabilità diffusa coerente non solo sugli eventi elementari, ma anche su loro collettività – eventi logicamente dipendenti da $\mathbb{P}$ –, di cui terremo conto per le prossime considerazioni in questo numero e anche nel seguito. La procedura, prescrivendo che si operi privilegiando un numerabile di eventi elementari, appare sicuramente più adatta per introdurre probabilità diffuse su eventi logicamente dipendenti da partizioni numerabili piuttosto che da quelle di cardinalità superiore, in relazione alle quali essa appare invero alquanto sbilanciata. In effetti, esse sono state in origine introdotte con riferimento a partizioni numerabili, dando a loro il nome di *probabilità asintotiche* – per ricordare che vengono determinate con operazioni di passaggio al limite su successioni di probabilità (*metodo asintotico*) –. Vedremo tuttavia che in qualche misura risultano utili anche per trarre conclusioni di carattere generale. Anche se il significato di probabilità asintotica è sufficientemente chiaro in questo momento – subito dopo la dimostrazione del precedente teorema –, per agevolare i futuri riferimenti conviene ricordarlo dando una precisa definizione.

Definizione. Probabilità asintotica.

Sia $\mathbb{P}$ una partizione infinita, $(\omega_1, \dots, \omega_n, \dots)$ una successione numerabile di eventi elementari di $\mathbb{P}$ distinti a due a due e P_n la probabilità uniforme su $\{\omega_1, \dots, \omega_n\}$. Si dà allora il nome di ***probabilità asintotica*** *di* ***base*** *$(\omega_1, \dots, \omega_n, \dots)$ all'applicazione $P = \lim_{n \to +\infty} P_n$ di dominio l'insieme $\mathfrak{D}$ degli eventi di $\mathcal{A}_L(\mathbb{P})$ per i quali il limite esiste.*

Ciò premesso, andiamo ora ad approfondire lo studio delle probabilità asintotiche provando che sussiste la seguente proposizione.

Proposizione.

a) *Le probabilità asintotiche dipendono dall'ordinamento degli eventi della relativa base, che per questo motivo è stata definita come successione di eventi elementari (a due a due distinti!) e non come insieme.*

b) *Il dominio $\mathfrak{D}$ di ogni probabilità asintotica con base in $\mathbb{P}$ contiene gli eventi finiti e cofiniti*[36] *di $\mathcal{A}_L(\mathbb{P})$ – in particolare Ω (cofinito) e ϕ (finito) –.*

---

36 La nozione di cofinito è riferita usualmente ai sottoinsiemi di un universo U: *un sottoinsieme di U è* ***cofinito*** *se il suo complementare è finito.* Quando gli eventi sono definiti come sottoinsiemi di un universo, scelto a rappresentare l'evento certo, la nozione si riferisce ovviamente anche a essi come caso particolare. La si può però estendere in modo del tutto naturale agli eventi logicamente dipendenti da una partizione $\mathbb{P}$, anche se definiti come classi di proposizioni equivalenti, riferendo i termini finito e cofinito al numero di eventi elementari di $\mathbb{P}$ che li compongono.

c) $\mathfrak{D}$ *è <u>chiuso</u> rispetto alla <u>negazione</u> e alla <u>somma logica finita</u> di eventi <u>a due a due incompatibili</u>.*

d) $\mathfrak{D}$ *<u>non è chiuso</u> rispetto alla <u>somma logica finita</u> di eventi <u>compatibili</u>. Esso <u>non è</u> perciò un'<u>algebra</u> ed è incluso <u>propriamente</u> in* $\mathcal{A}_L(\mathbb{P})$.

e) *L'intersezione dei domini delle probabilità asintotiche con base in una medesima partizione* $\mathbb{P}$ *è l'insieme degli eventi finiti e cofiniti di* $\mathcal{A}_L(\mathbb{P})$.

DIMOSTRAZIONE.

Sia $\{\omega_1, \dots, \omega_n, \dots\} \subset \mathbb{P}$ un insieme di eventi elementari distinti a due a due.

Prova della a).
Sia $E = \bigvee_{i=1}^{\infty} \omega_{2i}$. Usando la base $(\omega_1, \dots, \omega_n, \dots)$ – ordinamento naturale degli indici –, posto $n = 2h + r$, $r = 0, 1$, si ottiene $P_n(E) = h/(2h + r)$. Facendo tendere n – e quindi h – all'infinito si ricava $P(E) = 1/2$.
Usando la base che si ottiene alternando due ω_n di indice dispari con uno di indice pari la base diventa $\omega_1, \omega_3, \omega_2, \omega_5, \omega_7, \omega_4, \omega_9, \omega_{11}, \omega_6, \dots$. Posto $n = 3h + r$, $r = 0, 1, 2$, si trova $P_n'(E) = h/(3h + r)$. Facendo tendere n all'infinito si ricava $P'(E) = 1/3 \neq 1/2 = P(E)$.

Prova della b).
Siano E finito (cofinito) e s il massimo indice per cui $\omega_s \Rightarrow E$ ($\omega_s \Rightarrow \bar{E}$). Allora, per ogni $n \geq s$ riesce $P_n(E) \leq s/n$ $\left(P_n(E) \geq 1 - s/n\right)$. Facendo divergere n si trova che la successione $(P_n(E))$ converge a $P(E) = 0$ $(P(E) = 1)$.

Prova della c).
Sussistono le $P_n(\bar{E}) = 1 - P_n(E)$ e $P_n(E_1 \vee \dots \vee E_n) = P_n(E_1) + \dots + P_n(E_n) = 1$. Passando al limite si ha l'asserto.

Prova della d).
Per le argomentazioni che seguono in questo punto serve osservare che gli insiemi di eventi elementari del tipo $\{\omega_i : 2^n \leq i < 2^{n+1}\}$ hanno cardinalità 2^n. Ciò premesso, sia $E = \bigvee_{i=1}^{\infty} \omega_{2i}$, $A = \bigvee\{\omega_i : 2^{2h} \leq i < 2^{2h+1}, h = 0, 1, \dots\}$ e $B = (\bar{E} \wedge A) \vee (E \wedge \bar{A})$. Abbiamo visto provando la *a)* che $P(E) = 1/2$. Mostriamo ora che anche $P(B) = 1/2$, e quindi che $E, B \in \mathfrak{D}$. Può essere utile a tal fine dettagliare A e $\bar{A}$ per i primi valori di h, per i quali si trova:

$$A = \omega_1 \vee (\omega_4 \vee \dots \vee \omega_7) \vee (\omega_{16} \vee \dots \vee \omega_{31}) \vee (\omega_{64} \vee \dots \vee \omega_{127}) \vee \dots ,$$
$$\bar{A} = (\omega_2 \vee \omega_3) \vee (\omega_8 \vee \dots \vee \omega_{15}) \vee (\omega_{32} \vee \dots \vee \omega_{63}) \vee (\omega_{128} \vee \dots \vee \omega_{255}) \vee \dots$$

Si vede facilmente, aiutandosi anche con lo sviluppo qui sopra, che il numero

dei componenti di A di indice dispari inclusi in un segmento iniziale $\{\omega_1, \dots, \omega_n\}$ (quelli di $\bar{E} \wedge A$) è uguale o supera di 1 quello dei componenti di indice pari (quelli di $E \wedge A$) – il primo componente è ω_1 (di indice dispari) e i componenti iniziali di ogni intervallo successivo di componenti di A hanno indice pari e quelli finali indice dispari –. Pertanto, qualunque sia n, la differenza tra i componenti di $\bar{E} \wedge A$ e $E \wedge A$ inclusi in un segmento iniziale $\{\omega_1, \dots, \omega_n\}$ è 0 o 1. Simmetricamente, per i componenti di $\bar{A}$ avviene il viceversa: in ogni segmento $\{\omega_1, \dots, \omega_n\}$ quelli di indice pari sono in numero uguale o di 1 superiore a quelli di indice dispari. È perciò 0 o 1 la differenza tra i componenti di $E \wedge \bar{A}$ e $\bar{E} \wedge \bar{A}$. Di conseguenza, per ogni n la differenza tra i componenti di $B = (\bar{E} \wedge A) \vee (E \wedge \bar{A})$ e $\bar{B} = (E \wedge A) \vee (\bar{E} \wedge \bar{A})$ inclusi in un qualunque segmento $\{\omega_1, \dots, \omega_n\}$ è o 0 o 1 o 2, mentre la loro somma è ovviamente n. Il numero dei componenti di B inclusi in tale segmento è pertanto compreso tra $n/2$ e $n/2+2$. Segue allora $lim_{n \to +\infty} P_n(B) = 1/2$, come volevasi.
Proviamo ora che $E \vee B \notin \mathcal{D}$, mostrando così contemporaneamente che $\mathcal{D}$ non è chiuso rispetto la somma logica di eventi compatibili e che è incluso propriamente in $\mathcal{A}_L(\mathbb{P})$. In virtù della *c)* è equivalente a tal fine provare che $\bar{E} \wedge \bar{B} = \bar{E} \wedge \bar{A} \notin \mathcal{D}$. Posto $F = \bar{E} \wedge \bar{A}$ – a parole: F = *somma degli* ω_i *di indice dispari relativi agli intervalli* $2^{2h-1} \leq i < 2^{2h}$, $h \geq 1$ –, si tratta di far vedere che la successione $(P_n(F))$ non converge. Consideriamo a tal fine le due sottosuccessioni $(P_{2n}(F))$, $(P_{2n+1}(F))$, $n \geq 0$. I componenti di F di indice $i \leq 2^{2n}$ sono in numero di $1+2^2+\dots+(2^2)^{n-1} = (2^{2n}-1)/3$ e sono tali anche quelli di indice $i \leq 2^{2n+1}$, perché gli ω_i del prossimo intervallo – quelli per cui $2^{2n} \leq i < 2^{2n+1}$ – sono componenti di A. Allora si ottiene $P_{2n}(E) = (2^{2n}-1)/(3 \times 2^{2n})$ e $P_{2n+1}(E) = (2^{2n}-1)/(3 \times 2^{2n+1})$, da cui si ricava $lim_{n \to +\infty} P_{2n}(E) = 1/3$ e $lim_{n \to +\infty} P_{2n+1}(E) = 1/6$. Abbiamo trovato così due sottosuccessioni convergenti verso limiti diversi e la successione $(P_n(F))$ non converge.

Prova della e).
Per la *b)* l'insieme degli eventi finiti e cofiniti di $\mathcal{A}_L(\mathbb{P})$ è incluso in ogni dominio $\mathcal{D}$ e quindi anche nella loro intersezione.
Basta perciò provare che se $C \in \mathcal{A}_L(\mathbb{P})$ è infinito assieme alla sua negazione, esiste allora una base di $\mathbb{P}$ tale che in sua corrispondenza la successione $(P_{2n}(C))$ non converge. Costruiamo per questo una base sostituendo in quella del precedente punto *d)* gli eventi elementari ω_i che implicano $F = \bar{E} \wedge \bar{A}$ con componenti di C e quelli che implicano $\bar{F}$ con componenti di $\bar{C}$, conservando gli indici – mantenendo così l'ordine –. L'attuale successione $(P_{2n}(C))$ coincide allora con la precedente $(P_{2n}(F))$ – con diverso significato del simbolo P_{2n} – e come quella quindi non ammette limite.

La prova è così completa. ■

10.4 Rappresentazione delle probabilità su $\mathscr{A}_{\mathrm{L}}(\mathbb{P})$

Nei due numeri precedenti abbiamo studiato i *casi limite* delle probabilità definite su di una partizione $\mathbb{P}$: quello delle probabilità concentrate e quello delle diffuse. Abbiamo visto che esistono *infinite* probabilità concentrate su $\mathbb{P}$, ciascuna delle quali ammette *unico* prolungamento coerente su $\mathscr{A}_{\mathrm{L}}(\mathbb{P})$. Abbiamo altresì provato che se $\mathbb{P}$ è infinita, la probabilità diffusa su $\mathbb{P}$ è *unica* – è l'applicazione nulla –. In compenso essa ammette *infiniti* prolungamenti coerenti su $\mathscr{A}_{\mathrm{L}}(\mathbb{P})$. Infatti, come visto nel corso della dimostrazione di 10.3.2 *Proposizione, a,* la probabilità diffusa su $\mathbb{P}$ ammette più probabilità asintotiche – ivi ne vengono messe in evidenza due –, perciò più prolungamenti coerenti su $\mathscr{A}_{\mathrm{L}}(\mathbb{P})$, e allora ne ammette infiniti in virtù di 9.5.1 *Complemento.*

Ampliando ora lo studio delle probabilità di dominio $\mathbb{P}$, cominciamo con l'osservare che *ogni mistura* di una probabilità concentrata e di una diffusa dà luogo a una probabilità coerente su $\mathbb{P}$ (*Teorema* 9.3.1), con parte concentrata pari a γ se questo è il coefficiente della mistura che compete alla probabilità concentrata. Se $0 < \gamma < 1$ – quindi $\mathbb{P}$ infinita –, la probabilità che si ottiene è *parzialmente concentrata (diffusa).* In termini più precisi, queste considerazioni si riassumono nella seguente proposizione.

10.4.1 Proposizione. Caratterizzazione delle probabilità su partizioni infinite.

Sia $\mathbb{P}$ *una partizione di cardinalità infinita. Un'applicazione* P *di* $\mathbb{P}$ *in* $[0,1]$ *è una probabilità coerente se e solo se attribuisce valori positivi ad al più un numerabile di suoi eventi elementari e la somma di tali valori non supera* 1.

DIMOSTRAZIONE.

Se P è l'applicazione nulla, essa è una probabilità (diffusa) e non c'è altro da dimostrare (*Teorema* 10.3.1). Altrimenti, se γ è la somma dei valori positivi di P, $(1/\gamma)P$ è una probabilità concentrata su $\mathbb{P}$. Una mistura di questa con la probabilità diffusa (identicamente nulla) con coefficienti γ e $1-\gamma$, rispettivamente, dà luogo a una probabilità coerente su $\mathbb{P}$, coincidente con l'applicazione assegnata. Il viceversa è già stato provato nella *Proposizione* 10.1.1. ■

COMPLEMENTO.

Le probabilità concentrate su $\mathbb{P}$ *sono le sole* σ*-additive su* $\mathcal{A}_L(\mathbb{P})$.
Ovvio se $\mathbb{P}$ ha cardinalità finita. Facile corollario dell'attuale *Proposizione* se $\mathbb{P}$ è *numerabile*. Supposto infatti che P sia una probabilità *non* concentrata su $\mathbb{P} = \{\omega_1, \dots, \omega_n, \dots\}$, allora si ha: $P(\bigvee_{n=1}^{\infty} \omega_n) = P(\Omega) = 1 > \sum_{n=1}^{\infty} P(\omega_n)$ e perciò P non è σ-additiva. È forse intuitivo a questo punto, ma tutt'altro che semplice da provare, che il risultato valga anche se la cardinalità di $\mathbb{P}$ è superiore al numerabile. Per il caso continuo se ne è già fatto cenno in 5.3.3 *Commento.*

NOTA.

Si è affermato in 10.2.1 *Nota* che una probabilità P concentrata su una partizione $\mathbb{P}$ ammette prolungamenti *non concentrati* sulla partizione prodotto $\mathbb{P} \wedge \mathbb{P}'$ – $\mathbb{P}'$ partizione – se per un $\omega \in \mathbb{P}$ tale che $P(\omega) > 0$, l'insieme $\omega \wedge \mathbb{P}'$ è *infinito.* Ciò e conseguenza del fatto che in quest'ultima ipotesi ogni probabilità P su $\mathbb{P}$ – concentrata o no – può essere prolungata in modo coerente sulla partizione $\mathbb{P} \cup (\omega \wedge \mathbb{P}')$ distribuendo la probabilità $P(\omega)$ *a piacere* su $\omega \wedge \mathbb{P}'$, purché lo si faccia in modo che il prolungamento P' (coincidente con P su $\mathbb{P} - \{\omega\}$ sia tale che $P'(\bigvee(\omega \wedge \mathbb{P}')) = P(\omega)$. Fatta salva questa condizione, che si distribuisca la probabilità *a pezzetti* – su un insieme finito o numerabile di eventi di $\omega \wedge \mathbb{P}'$, attribuendo probabilità nulla a tutti gli altri –, che la si distribuisca *diffusa* su $\omega \wedge \mathbb{P}'$ – ad esempio col metodo asintotico, equiripartendo prima $P(\omega)$ sui segmenti iniziali di una base $\{\omega \wedge e_1, \dots, \omega \wedge e_n, \dots\} \subset \omega \wedge \mathbb{P}'$ e passando poi al limite per $n \to +\infty$ – o che lo si faccia in *modo misto* – misturando due delle assegnazioni precedenti –, si ottiene sempre un'applicazione P' su $(\mathbb{P} - \{\omega\}) \cup (\omega \wedge \mathbb{P}')$ che attribuisce probabilità positiva a un numerabile al più di eventi elementari, la cui somma non supera quella degli eventi elementari di $\mathbb{P}$ di probabilità positiva, che a sua volta per ipotesi non supera 1. Per la caratterizzazione precedente l'applicazione P' è allora una probabilità coerente. ■

Passando ora ai prolungamenti su $\mathcal{A}_L(\mathbb{P})$ delle probabilità di dominio $\mathbb{P}$, osserviamo che ogni mistura del prolungamento di una probabilità concentrata – unico – e di un prolungamento della probabilità diffusa dà luogo a una probabilità coerente su $\mathcal{A}_L(\mathbb{P})$ (*Teorema* 9.3.1). Il prossimo teorema dimostra che le probabilità su $\mathcal{A}_L(\mathbb{P})$ si possono scomporre sempre come mistura di una *componente (probabilità) concentrata* e di una *componente diffusa.*

10.4.2 Teorema. Componenti concentrata e diffusa delle probabilità su $\mathscr{A}_L(\mathbb{P})$.

Sia $\mathbb{P}$ una partizione di Ω. Allora, ogni probabilità P su $\mathscr{A}_L(\mathbb{P})$ si può ottenere come mistura di una probabilità concentrata e di una diffusa. Salvo che per i casi limite – P concentrata o diffusa – la rappresentazione è unica.

DIMOSTRAZIONE.

Che la mistura di una probabilità concentrata e di una diffusa su $\mathscr{A}_L(\mathbb{P})$ sia una probabilità coerente è stato visto poc'anzi in premessa. Qui si tratta perciò di provare il viceversa. Sia allora P una probabilità su $\mathscr{A}_L(\mathbb{P})$ *parzialmente* concentrata (escludiamo cioè per il momento i casi limite). Indicati con ω_i, $i \in I$, gli eventi di $\mathbb{P}$ di probabilità positiva e con $\gamma = \sum_{i\in I} P(\omega_i)$ la parte concentrata di P su $\mathbb{P}$, $0 < \gamma < 1$, per ogni $E \in \mathbb{P}$ poniamo:

$$\mu_c(E) = \textstyle\sum_{i\in I} P(\omega_i)\, \delta(\omega_i \Rightarrow E), \quad \mu_d(E) = P(E) - \mu_c(E).$$

Le applicazioni μ_c e μ_d sono univocamente determinate dalla probabilità P, riesce $0 \le \mu_c(E) \le P(E)$ (*Lemma* 10.1.3) e quindi anche $0 \le \mu_d(E) \le P(E)$. Inoltre, $\mu_c(\omega) = P(\omega)$ per ogni $\omega \in \mathbb{P}$. Consideriamo ora l'applicazione $P_c = (1/\gamma)\,\mu_c$, univocamente determinata da P, perché tale è la μ_c. Per quanto osservato poc'anzi si ha $P_c(\omega) \ge 0$ per ogni $\omega \in \mathbb{P}$ e $\sum_{i\in I} P_c(\omega_i) = (1/\gamma) \sum_{i\in I} \mu_c(\omega_i) = 1$. L'applicazione P_c è dunque una probabilità concentrata su $\mathbb{P}$ e, per ogni $E \in \mathscr{A}_L(\mathbb{P})$, il suo prolungamento è unico e dato da (*Teorema* 10.2.1):

$$P_c(E) = \textstyle\sum_{i\in I} P_c(\omega_i)\, \delta(\omega_i \Rightarrow E) = (1/\gamma) \sum_{i\in I} P(\omega_i)\, \delta(\omega_i \Rightarrow E).$$

Riassumendo, abbiamo intanto che P_c è una probabilità concentrata su $\mathscr{A}_L(\mathbb{P})$ e univocamente determinata da P.
Ciò visto, poniamo ora $P_d = (1-\gamma)^{-1}\,\mu_d$. Anche l'applicazione P_d è univocamente determinata da P, perché tale è la μ_d. Inoltre, essa è una probabilità diffusa su $\mathscr{A}_L(\mathbb{P})$, come andiamo qui di seguito a dimostrare. Cominciamo con l'osservare che si può scrivere: $P_d = (1-\gamma)^{-1}(P - \mu_c) = (1-\gamma)^{-1}(P - \gamma P_c)$. Ciò premesso, la prova che P_d è una probabilità si consegue usando il *Teorema* 9.4.2. Si ha:

a) $P_d(\Omega) = (1-\gamma)^{-1}(P(\Omega) - \gamma P_c(\Omega)) = (1-\gamma)^{-1}(1-\gamma) = 1$;

b) per ogni $E \in \mathscr{A}_L(\mathbb{P})$ riesce $P_d(E) \ge 0$, perché, come visto sopra, $\mu_d(E) \ge 0$;

c) se $E_1, E_2 \in \mathscr{A}_L(\mathbb{P})$ sono incompatibili, allora per l'additività di P e P_c si ottiene:

$$P_d(E_1 \vee E_2) = (1-\gamma)^{-1}(P(E_1 \vee E_2) - \gamma P_c(E_1 \vee E_2)) =$$
$$(1-\gamma)^{-1}(P(E_1) - \gamma P_c(E_1)) + (1-\gamma)^{-1}(P(E_2) - \gamma P_c(E_2)) = P_d(E_1) + P_d(E_2).$$

Pertanto, P_d è coerente su $\mathscr{A}_L(\mathbb{P})$. Proviamo che essa è diffusa. Per ogni $\omega \in \mathbb{P}$ si ha $P_d(\omega) = (1-\gamma)^{-1}(P(\omega) - \mu_c(\omega)) = (1-\gamma)^{-1}(P(\omega) - P(\omega)) = 0$.

In conclusione, poiché anche γ è univocamente determinata dalla probabilità P, se questa è parzialmente concentrata la si può ottenere in modo unico come mistura di una probabilità concentrata e una diffusa su $\mathbb{P}$. Riesce precisamente: $P = \mu_c + \mu_d = \gamma P_c + (1-\gamma) P_d$, con γ parte concentrata di P su $\mathbb{P}$.
Come si vede, si può ottenere questa rappresentazione anche nei casi limite $\gamma = 0$ e $\gamma = 1$. In questi casi però la rappresentazione non è unica. Se $\gamma = 0$ si ottiene $P = P_d$ e si può scegliere P_c in modo arbitrario. Simmetricamente, se $\gamma = 1$ è $P = P_c$ e ora è P_d a essere arbitraria (purché $\mathbb{P}$ sia infinita). ■

10.4.3 Parte concentrata di una probabilità e limitazioni.

Il teorema del prolungamento (n° 9.5.1) assicura che ogni valutazione di probabilità coerente su un insieme si può sempre prolungare in modo coerente su ogni suo soprainsieme. In particolare, abbiamo visto nel corso della dimostrazione del teorema citato – e messo in evidenza in 9.5.1 *Corollario* – che se è data una probabilità su un insieme $\mathscr{E}$ e si aggiunge un unico evento E, allora l'insieme delle valutazioni coerenti per E è un intervallo, di cui non è detto però, nella dimostrazione, come si fa a determinare in modo costruttivo i suoi estremi né si conoscono regole di carattere generale che insegnino a farlo. Si riesce a farlo in ipotesi particolari. Ad esempio se, come vedremo in questo numero, il problema riguarda una partizione $\mathbb{P}$ è un evento E generici, nell'ipotesi che siano note entità e distribuzione della parte concentrata di una probabilità assegnata su $\mathbb{P}$ e le condizioni di dipendenza logica dell'evento dalla partizione medesima. Della parte di probabilità diffusa è nota l'entità – per complemento a 1 della parte concentrata –, ma nulla della sua distribuzione. È il caso di osservare a questo proposito che assegnazioni concrete di valutazioni diffuse non possono essere che parziali – come ad esempio nel caso delle distribuzioni asintotiche (10.3.2 *Definizione*) –. Il teorema del prolungamento assicura infatti l'esistenza di probabilità coerenti su ogni insieme, e quindi anche su *tutto* $\mathscr{A}_L(\mathbb{P})$ (9.5.1 *Proposizione*), ma non si conoscono modi costruttivi per ottenerle.
Nelle ipotesi dette sopra sulla probabilità assegnata su $\mathbb{P}$, abbiamo già visto nel *Lemma* 10.1.3 che per ogni evento devono essere soddisfatte le disuguaglianze (20), le quali in corrispondenza di ogni evento indicano l'intervallo entro cui è *necessario* scegliere la sua probabilità per essere coerenti. Abbiamo

altresì visto che in generale il rispetto di tali limitazioni non è *sufficiente* a garantire la coerenza. Condizioni di natura logica possono infatti imporre una restrizione dell'intervallo indicato nella (20). Andiamo ora ad approfondire questo aspetto perfezionando l'enunciato del lemma citato con l'aggiunta di tutte le possibili ipotesi sul legame logico di E da $\mathbb{P}$, ed indicando in corrispondenza di ogni ipotesi l'intervallo entro cui la scelta della probabilità di E oltre che necessaria è anche sufficiente. Sussiste in proposito il seguente teorema.

Teorema.

Siano $\mathbb{P}$ una partizione di Ω, E un evento, P una probabilità su $\mathbb{P}$, $\mathcal{E}^+$ l'insieme degli eventi elementari di $\mathbb{P}$ di probabilità positiva. Allora, condizione necessaria affinché una valutazione di probabilità $P(E)$ sia coerente è che soddisfi le seguenti limitazioni:

$$\sum_{\omega\in\mathcal{E}^+} P(\omega)\,\delta(\omega\Rightarrow E) \leq P(E) \leq 1-\sum_{\omega\in\mathcal{E}^+} P(\omega)\,\delta(\omega\Rightarrow \bar{E}). \qquad (21)$$

Tali limitazioni sono anche sufficienti se P è concentrata su $\mathbb{P}$ oppure se E e $\bar{E}$ sono entrambi infiniti nella partizione $\mathbb{P}\wedge\{E,\bar{E}\}$ (hanno infiniti componenti in quella rappresentazione).
Altrimenti l'intervallo di coerenza si restringe di una quantità pari alla probabilità diffusa su $\mathbb{P}$ e diminuisce (aumenta) di tale quantità il suo estremo superiore (inferiore) se E ($\bar{E}$) è finito in $\mathbb{P}\wedge\{E,\bar{E}\}$.

Dimostrazione.

Le limitazioni (21) sono le (20) del *Lemma* 10.1.3, che afferma che esse sono condizione necessaria per la coerenza. Qui resta da dimostrare perciò la parte dell'enunciato che riguarda la sufficienza. Poiché sappiamo che le scelte coerenti per la probabilità di E costituiscono un intervallo chiuso (9.5.1 *Corollario*), è sufficiente per questo determinare la minima e massima valutazione coerente per E, nelle singole ipotesi previste dal teorema. Cominciamo col considerare i prolungamenti coerenti di P su $\mathbb{P}\wedge\{E,\bar{E}\}$. Per ciascuno di essi *è necessario* scegliere un'applicazione P' non negativa tale che $P'(\omega\wedge E)+P'(\omega\wedge\bar{E})=P(\omega)$, per ogni $\omega\in\mathbb{P}$. Più specificamente si deve porre

$$P'(\omega\wedge E)=\begin{cases} P(\omega) & \text{se } \omega\Rightarrow E \\ 0 & \text{se } \omega\Rightarrow\bar{E} \\ p_\omega\leq P(\omega) & \text{se } \omega \text{ compatibile con } E \text{ e } \bar{E}\end{cases}$$

e a seguire $P'(\omega\wedge\bar{E})=P(\omega)-P'(\omega\wedge E)$. Riesce così in particolare $P'(\omega\wedge E)=$

$P'(\omega\wedge \bar{E})=0$ se $P(\omega)=0$ – come deve essere – e P' può assumere perciò valori positivi solo sui costituenti $\omega\wedge E$ e $\omega\wedge \bar{E}$ con componenti $\omega\in \mathcal{E}^+$. La somma dei valori positivi di P' – se esistono – è allora

$$\textstyle\sum_{\omega\in \mathcal{E}^+}\ [P'(\omega\wedge E)+P'(\omega\wedge \bar{E})] = \sum_{\omega\in \mathcal{E}^+}\ P(\omega)\leq 1, \qquad (22)$$

la disuguaglianza valendo perché P è una probabilità su $\mathbb{P}$. L'applicazione P' è perciò una probabilità coerente su $\mathbb{P}\wedge\{E,\bar{E}\}$ per ogni scelta dei p_ω (*Proposizione* 10.4.1). È altresì evidente che ogni prolungamento di P su $\mathbb{P}\wedge\{E,\bar{E}\}$ soddisfa le condizioni poste come necessarie. Con la procedura descritta sopra si ottengono perciò tutti e soli i prolungamenti coerenti di P su $\mathbb{P}\wedge\{E,\bar{E}\}$. Questo significa che se delle limitazioni per $P(E)$ sono sufficienti per qualche prolungamento P', lo sono anche per la sua restrizione P.

Ciò premesso, andiamo ora a determinare la restrizione che si deve imporre alle limitazioni (21) affinché – oltre che necessarie – esse diventino anche sufficienti per una scelta della probabilità di E coerente con la valutazione P, studiando il problema per casi con l'ausilio dei prolungamenti P'.

Sia P concentrata.
L'uguaglianza delle due sommatorie nella (22) dice che se P è concentrata su $\mathbb{P}$, allora P' è concentrata su $\mathbb{P}\wedge\{E,\bar{E}\}$. In questa ipotesi, essendo E logicamente dipendente da $\mathbb{P}\wedge\{E,\bar{E}\}$, posto

$$\mathcal{E}=\{\omega\in \mathcal{E}^+ : \omega\wedge E \text{ e } \omega\wedge\bar{E} \text{ possibili}\},$$

riesce (*Proposizione* 10.1.4):

$$P'(E)=\textstyle\sum_{\omega\in \mathcal{E}^+}\ P'(\omega\wedge E)\ \delta(\omega\wedge E\Rightarrow E) =$$
$$\textstyle\sum_{\omega\in \mathcal{E}^+}\ P(\omega)\ \delta(\omega\Rightarrow E)+\sum_{\omega\in \mathcal{E}}\ p_\omega\delta(\omega\wedge E\Rightarrow E).$$

La minima (massima) scelta coerente per la probabilità di E si ottiene allora ponendo $p_\omega=0$ ($p_\omega=P(\omega)$) per ogni $\omega\in\mathcal{E}$, da cui si ricavano facilmente le due limitazioni previste nella (21). Ciò prova l'asserto nell'ipotesi che P sia concentrata.

Sia P diffusa.
In questo caso le limitazioni della (21) diventano 0, 1. Si ha poi che, dovendo soddisfare la condizione $P'(\omega\wedge E)+P'(\omega\wedge\bar{E})=P(\omega)=0$ per ogni $\omega\in\mathbb{P}$, l'unico prolungamento P' coerente di P su $\mathbb{P}\wedge\{E,\bar{E}\}$ è l'applicazione identicamente nulla, e perciò anche P' è una probabilità diffusa.
Come già osservato in 10.1.3 *Nota*, da qui segue subito che la scelta di $P(E)$ è obbligata se E o $\bar{E}$ sono finiti in $\mathbb{P}\wedge\{E,\bar{E}\}$ – $P(E)=0$ nel primo caso e $P(E)=1$ nel secondo –. L'intervallo $[0,1]$ degenera in uno degli estremi (si riduce di 1,

della quantità di probabilità diffusa), e ciò prova quanto enunciato nelle ipotesi E e $\bar{E}$ finiti e nel caso particolare che P sia diffusa interamente.
Se E o $\bar{E}$ sono entrambi infiniti in $\mathbb{P} \wedge \{E, \bar{E}\}$ – $E \wedge \mathbb{P}$, $\bar{E} \wedge \mathbb{P}$ infiniti –, possiamo considerare un prolungamento asintotico di P' – che indichiamo ancora con P' – di base $(E \wedge \omega_1, \dots, E \wedge \omega_n, \dots)$, $\omega_1, \dots, \omega_n, \dots \in \mathbb{P}$ (10.3.2 *Definizione*). Allora si ha $E' = \bigvee_{n=1}^{\infty} (E \wedge \omega_n) \Rightarrow E$ e $P'(E') = 1$ – perché $P_n(E') = 1$ per ogni n –. Segue $P'(E) \geq P'(E') = 1$. Simmetricamente, sia P'' la probabilità asintotica di base $(\bar{E} \wedge e_1, \dots, \bar{E} \wedge e_n, \dots)$, $e_1, \dots, e_n, \dots \in \mathbb{P}$. Procedendo come sopra si ricava allora $P''(\bar{E}) = 1$, ovvero $P''(E) = 0$. Scelto ora $p \in [0,1]$ e posto $P = pP' + (1-p)P''$, per il *Teorema* 9.3.1 si ha che P è una probabilità coerente su E e che riesce $P(E) = p$. È dunque coerente scegliere per la probabilità di E ogni numero di $[0,1]$, e questo completa la prova quando la probabilità P è diffusa.

Sia P parzialmente concentrata (diffusa).
Indicata con $\gamma = \sum_{\omega \in \mathcal{E}^+} P(\omega)$ la parte concentrata di P, supponiamo ora che sia $0 < \gamma < 1$. Si può allora scrivere $P = \gamma P_c + (1-\gamma) P_d$, con P_c componente concentrata e P_d componente diffusa di P su $\mathbb{P}$ (*Teorema* 10.4.2). Tenuto conto che gli insiemi degli eventi elementari di probabilità positiva relativi a P e P_c coincidono (ovvio) e che P_c è una probabilità concentrata, per quanto visto sopra è coerente scegliere $P_c(E)$ nei limiti

$$\sum_{\omega \in \mathcal{E}^+} P_c(\omega)\, \delta(\omega \Rightarrow E) \leq P_c(E) \leq 1 - \sum_{\omega \in \mathcal{E}^+} P_c(\omega)\, \delta(\omega \Rightarrow \bar{E}),$$

da cui, tenuto conto che $\gamma P_c(\cdot) = \mu_c(\cdot) = \sum_{\omega \in \mathcal{E}^+} P(\omega)\, \delta(\omega \Rightarrow \cdot)$ (*Teorema* 10.4.2), si ricava

$$\sum_{\omega \in \mathcal{E}^+} P(\omega)\, \delta(\omega \Rightarrow E) \leq \gamma P_c(E) \leq \gamma - \sum_{\omega \in \mathcal{E}^+} P(\omega)\, \delta(\omega \Rightarrow \bar{E}). \quad (23)$$

Per quel che riguarda P_d, invece, essa si prolunga in modo unico se E è finito o cofinito in $\mathbb{P} \wedge \{E, \bar{E}\}$ e riesce $P_d(E) = 0$ e $P_d(E) = 1$, rispettivamente. Altrimenti, se E e $\bar{E}$ sono ivi infiniti, si può scegliere per $P_d(E)$ un valore arbitrario in $[0,1]$. Tenuto conto di questi risultati si ottengono le conclusioni articolate nei seguenti sottocasi.

Siano $E \wedge \mathbb{P}$ e $\bar{E} \wedge \mathbb{P}$ infiniti.
Dalla $0 \leq P_d(E) \leq 1$ si ricava $0 \leq (1-\gamma) P_d(E) \leq 1 - \gamma$ e sommando membro a membro nella (23) si ottiene la (21) e con ciò l'asserto nell'attuale ipotesi.

Sia $E \wedge \mathbb{P}$ finito.
Si deve porre $P_d(E) = 0$ e quindi $0 \leq (1-\gamma) P_d(E) \leq 0$. Sommando nella (23) si ottiene

$$\sum_{\omega \in \mathcal{E}^+} P(\omega)\, \delta(\omega \Rightarrow E) \leq P(E) \leq 1 - \sum_{\omega \in \mathcal{E}^+} P(\omega)\, \delta(\omega \Rightarrow \bar{E}) - (1-\gamma),$$

cioè la tesi.

Sia $\bar{E} \wedge \mathbb{P}$ *finito.*
Si deve porre $P_d(E) = 1$ e quindi $1 - \gamma \leq (1 - \gamma) P_d(E) \leq 1 - \gamma$. Sommando nella (23) ora si ottiene

$$\sum_{\omega \in \mathcal{E}^+} P(\omega)\, \delta(\omega \Rightarrow E) + (1 - \gamma) \leq P(E) \leq 1 - \sum_{\omega \in \mathcal{E}^+} P(\omega)\, \delta(\omega \Rightarrow \bar{E}),$$

e anche qui la tesi.

La prova è così completa. ■

NOTA. *Limitazioni e relazioni logiche.*

Se la probabilità P è concentrata su $\mathbb{P}$ – come sempre accade ad esempio se $\mathbb{P}$ è finita – ed E è *logicamente dipendente* da $\mathbb{P}$, allora la sua probabilità è univocamente determinata. Lo abbiamo osservato già in 10.1.3 *Nota* e ribadito nel *Teorema* 10.2.1 che caratterizza le probabilità concentrate. Il teorema attuale lo conferma, aggiungendo però che le limitazioni (21) – le (20) del *Lemma* 10.1.3 – sono anche sufficienti, ovvero *non migliorabili,* nel senso che l'intervallo *chiuso* da esse determinato è l'insieme di tutte e sole le valutazioni di probabilità di E coerenti con la distribuzione data su $\mathbb{P}$. Nel corso della sua dimostrazione si vede anche che l'ampiezza di tale intervallo è determinata esclusivamente dalla probabilità che si trova concentrata sugli eventi elementari $\omega \in \mathbb{P}$ di tipo 3 per E (n° 3.4.2) – e che può essere ripartita a piacere tra E e $\bar{E}$ –. Le limitazioni diventano allora 0 e 1 se E è *logicamente indipendente* da $\mathbb{P}$, e non pongono perciò alcun vincolo alla scelta di $P(E)$. Se E è *logicamente semidipendente,* generalmente le limitazioni si restringono. Basta (e occorre) per questo che *almeno un* evento elementare di probabilità positiva *non sia* di tipo 3 per E. Si possono però presentare anche le due situazioni limite: *tutti* gli eventi elementari di probabilità positiva sono di tipo 3 per E, e allora le limitazioni rimangono 0 e 1 (come nel caso di logica indipendenza); *nessun* evento elementare di probabilità positiva è di tipo 3 per E, e allora le due limitazioni coincidono e la probabilità di E è univocamente determinata (come nel caso di logica dipendenza).
Se P è diffusa su $\mathbb{P}$, rivisitando la dimostrazione fatta in corrispondenza di questa ipotesi si vede che per poter scegliere in modo coerente la probabilità di E non importa conoscere il legame di dipendenza logica di E da $\mathbb{P}$. Occorre invece conoscere la cardinalità di E e $\bar{E}$ in $\{E, \bar{E}\} \wedge \mathbb{P}$ (che è comunque, anche questo, un legame di natura logica). La scelta di $P(E)$ non è soggetta a vincoli se E e $\bar{E}$ sono entrambi infiniti in $\{E, \bar{E}\} \wedge \mathbb{P}$ ed è obbligata altrimenti – $P(E) = 0$ o $P(E) = 1$ a seconda che E sia finito o cofinito –.

10.5 Applicazioni

Il teorema del prolungamento assicura che ogni probabilità coerente su un insieme si può sempre estendere in modo coerente su ogni suo soprainsieme. Interessa ora mettere l'accento sulla scelta dei prolungamenti, che, come visto per un evento nel precedente 10.4.3 *Teorema* (e nella *Nota* che lo segue), talvolta è libera – in $[0,1]$ ovviamente –, altre volte è soggetta a limitazioni più strette, talvolta fino al punto di essere obbligata. Questo aspetto dei problemi della valutazione è una delle caratteristiche dell'impostazione soggettiva, o meglio, dell'approccio dinamico ai problemi dell'incertezza proposto da de Finetti per rendere la teoria aderente alle procedure e alle esigenze applicative. Lo studio fatto nei paragrafi precedenti di questo capitolo ci mette ora in grado di approfondire e comprendere meglio questa peculiarità della teoria. Di precisare ad esempio che visto da questa angolazione il calcolo delle probabilità appare come una *ricerca di vincoli entro cui muoversi per esprimere con coerenza il proprio grado di fiducia.* La mancanza di più possibilità di scelta è solo un caso limite anche se molto importante. Esso indica – quando presente – che nel quadro descrittivo fissato le probabilità dei suoi eventi sono tutte determinate. A questo livello di descrizione e valutazione il calcolo delle probabilità diventa analogo a quello classico. Come in quella impostazione si tratta infatti di calcolare probabilità ben determinate – e non intervalli di coerenza –. Per chi sposa la teoria soggettiva ciò è però conseguenza di scelte soggettive precedenti e di accettazione della norma di coerenza. Se egli trovasse che le probabilità *calcolate* non lo convincono – non gli sembrano cioè rispondenti alle sue opinioni e sensazioni –, è nelle scelte iniziali che dovrebbe ricercare le ragioni della sua insoddisfazione. Se dopo una seconda e più accurata analisi le giudicasse ancora rispondenti al suo grado di fiducia, non gli resterebbe altro che accettare – per coerenza – le conclusioni.

Come s'intuisce già da queste poche righe, la ricerca di vincoli entro cui muoversi in condizioni di coerenza è un problema di larghissimo respiro e in generale di non semplice soluzione, anche in casi non troppo complessi, soprattutto se si pretende di trovare risposte esaustive. Lo testimonia la dimostrazione di 10.4.3 *Teorema,* che appare piuttosto complicata se messa a confronto con la sostanziale semplicità del problema proposto: prolungamento su un evento di una probabilità coerente su una partizione.

I risultati trovati nei paragrafi precedenti pongono comunque le basi per trattare in modo completo, in termini di probabilità coerenti, il problema della valutazione in *ambiente finito.* Essi sono pure sufficienti per mettere a fuoco abbastanza bene il medesimo problema anche in *ambiente numerabile.* Consentono invece di trarre solo conclusioni parziali per gli ambienti più che numerabili. Sono tuttavia sufficienti a dare consistenza ai ragionamenti intuitivi svolti in ambiente continuo nel *Cap.* 5. Questo spiega la diversa ampiezza che abbiamo destinato all'approfondimento dei tre casi nei prossimi tre numeri di questo paragrafo.

10.5.1 Probabilità in ambiente finito.

Il *Teorema* 10.2.1 caratterizza le probabilità concentrate su partizioni e sui corrispondenti insiemi di eventi logicamente dipendenti. In particolare rimangono così caratterizzate *tutte* le probabilità coerenti su partizioni finite, perché ogni probabilità P coerente su una partizione finita $\mathbb{P} = \{\omega_1, \dots, \omega_n\}$ è una probabilità concentrata. Per l'additività delle probabilità deve essere infatti $P(\Omega) = P(\vee_{i=1}^{n} \omega_i) = \sum_{i=1}^{n} P(\omega_i) = 1$. Come corollario del *Teorema* 10.2.1 abbiamo perciò la seguente caratterizzazione.

Proposizione. Caratterizzazione delle probabilità in ambiente finito.

Sia $\mathbb{P} = \{\omega_1, \dots, \omega_n\}$ *una partizione di* Ω. *Le probabilità* P *coerenti su* $\mathbb{P}$ *si ottengono, tutte e sole, scegliendo n numeri non negativi* $p_1, \dots, p_n$ *di somma* 1 *e ponendo* $P(\omega_i) = p_i$, $i = 1, \dots, n$. *La* P *si prolunga in modo unico per additività su ogni evento* $E \in \mathcal{A}_{\mathrm{L}}(\mathbb{P})$ *ponendo* $P(E) = \sum_{i=1}^{n} p_i\, \delta(\omega_i \Rightarrow E)$.

Nota.

Questa proposizione mostra che in ambiente finito – inteso come insieme degli eventi logicamente dipendenti da una partizione finita – la nozione di probabilità coerente coincide formalmente (per le proprietà) con quella classica (*Definizione* 4.2.1). Gli esempi presentati nel *Cap.* 4 sono perciò esempi di valutazione coerente dell'incertezza.
La coincidenza formale tra le due teorie – che devono essere sempre confrontate con riferimento a descrizioni dell'incertezza mediante *algebre di eventi logicamente dipendenti da una partizione* – si ferma però all'ambiente finito.
Se la partizione è numerabile, l'ambiente è l'intera σ-algebra degli eventi logicamente dipendenti da essa e si ha perciò coincidenza dal punto di vista descrittivo. Non da quello della valutazione, però, perché in virtù del *Teorema* 10.2.1 abbiamo che le probabilità della teoria classica – tutte concentrate come recita la *Definizione* 5.2.1 – sono probabilità coerenti, ma non sussiste il viceversa. Il *Teorema* 10.3.1 e la *Proposizione* 10.4.1 mostrano infatti che esistono probabilità diffuse e parzialmente diffuse su partizioni di cardinalità qualsiasi, e quindi anche numerabile.

Se la descrizione viene fatta mediante una partizione di cardinalità superiore al numerabile, la coincidenza viene allora a mancare già a livello descrittivo, perché nella teoria classica la probabilità è una misura σ-additiva su una σ-algebra inclusa propriamente in quella degli eventi logicamente dipendenti dalla partizione (5.3.3 *Commento*).

ESEMPI.

Il primo esempio tratta di un problema di verifica della coerenza di una valutazione e successivamente della determinazione degli intervalli delle valutazioni coerenti per assegnati eventi, compatibili con la data valutazione.
Alla verifica della coerenza di una valutazione è dedicato anche l'*Esempio* 3.
L'*Esempio* 2 approfondisce invece il problema della determinazione di intervalli di coerenza. In esso vengono determinati gli intervalli relativi a cinque eventi, compatibili con una probabilità assegnata su una partizione. Si fa vedere poi che scelte *singolarmente coerenti* per due dei dati eventi, possono dare luogo a una scelta *congiunta incoerente*. La cosa non dovrebbe sorprendere. Abbiamo visto infatti, nei paragrafi precedenti, che ogni scelta coerente pone vincoli per scelte coerenti successive. È allora naturale attendersi che la scelta di un valore preciso – ancorché coerente – per la probabilità di uno degli eventi coinvolti nell'analisi possa provocare una contrazione degli intervalli di coerenza trovati per gli altri. Ciò e messo concretamente in evidenza nella parte conclusiva dell'esempio.
L'*Esempio* 4 può essere letto come un caso di valutazione fatta con l'aiuto di opinioni espresse da esperti e anche come introduzione all'interpretazione della probabilità come penalità. A quest'ultimo argomento è dedicato il prossimo § 10.6 che conclude il presente capitolo.

1 *Gli eventi E_1, E_2, E_3 sono incompatibili ed esaustivi. Verificare che la seguente valutazione di probabilità è ammissibile:*

$$P(E_1) = 0{,}5, \quad P(E_2) = 0{,}6, \quad P((E_1 \wedge \bar{E}_2 \wedge E_3) \vee (\bar{E}_1 \wedge E_2 \wedge \bar{E}_3)) = 0{,}2.$$

Trovare poi le limitazioni di probabilità per gli eventi:

a) *E, sapendo che $E_1 \wedge E_2 \Rightarrow E \Rightarrow E_2 \vee (E_1 \wedge E_3)$,*
b) E_3.

Poiché E_1, E_2, E_3 sono incompatibili ed esaustivi si ha $E_1 \wedge E_2 \wedge E_3 = \phi$, e $\neg(E_1 \vee E_2 \vee E_3) = \bar{E}_1 \wedge \bar{E}_2 \wedge \bar{E}_3 = \phi$. La partizione generata da E_1, E_2, E_3 ha perciò 6 costituenti possibili: quelli che contengono almeno una affermazione e almeno una negazione. Poniamo:

$$P(E_1 \wedge E_2 \wedge \bar{E}_3) = p_1, \quad P(E_1 \wedge \bar{E}_2 \wedge E_3) = p_2, \quad P(\bar{E}_1 \wedge E_2 \wedge E_3) = p_3,$$
$$P(E_1 \wedge \bar{E}_2 \wedge \bar{E}_3) = p_4, \quad P(\bar{E}_1 \wedge E_2 \wedge \bar{E}_3) = p_5, \quad P(\bar{E}_1 \wedge \bar{E}_2 \wedge E_3) = p_6.$$

Affinché le $p_1, \dots, p_6$ siano probabilità compatibili con la valutazione data sopra devono essere soddisfatte le seguenti condizioni:

$$p_1 + p_2 + p_4 = 0{,}5, \quad p_1 + p_3 + p_5 = 0{,}6, \quad p_2 + p_5 = 0{,}2,$$
$$p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1, \quad p_i \geq 0, \quad i = 1, \dots, 6.$$

Abbiamo un sistema di quattro equazioni in sei incognite. Scelte p_1 e p_2 come parametri, la soluzione è data da:

$$p_3 = 0{,}4 - p_1 + p_2, \quad p_4 = 0{,}5 - p_1 - p_2, \quad p_5 = 0{,}2 - p_2, \quad p_6 = p_1 - 0{,}1. \quad (24)$$

Imponendo le condizioni di non negatività alle p_1, p_2 e ai secondi membri della soluzione ora riportata si ottiene che la regione delle valutazioni ammissibili per p_1 e p_2 è rappresentata dal pentagono tratteggiato in *Fig. 14* di vertici (0,1 ,0), (0,4 ,0), (0,45 ,0,05), (0,3 ,0,2), (0,1 ,0,2). La regione ammissibile non è dunque vuota e perciò la valutazione fatta sugli eventi $E_1, E_2, (E_2 \wedge \bar{E}_2 \wedge E_3) \vee (\bar{E}_1 \wedge E_2 \wedge \bar{E}_3)$ è coerente.

Passando alla seconda parte, cominciamo con l'osservare che l'evento E del punto *a)* è logicamente semidipendente dalla partizione generata da $\{E_1, E_2, E_3\}$. L'evento elementare $E_1 \wedge E_2 \wedge \bar{E}_3$ è di tipo 1, mentre $\bar{E}_1 \wedge \bar{E}_2 \wedge E_3, E_1 \wedge \bar{E}_2 \wedge \bar{E}_3$ sono di tipo 2 – segue da $\neg[E_2 \vee (E_1 \wedge E_3)] = [(\bar{E}_1 \wedge \bar{E}_2) \vee (\bar{E}_2 \wedge \bar{E}_3)] \Rightarrow \bar{E}$ –. Pertanto, per ogni scelta ammissibile delle $p_1, \dots, p_6$, si ha $p_1 \leq P(E) \leq 1 - (p_6 + p_4) = 1 - (0{,}4 - p_2) = 0{,}6 + p_2$. Tenuto conto che i punti di ottimo delle funzioni lineari soggette al vincolo che il punto (p_1, p_2) appartenga al pentagono della *Fig. 14* vengono raggiunti in un vertice, si ha che la massima limitazione inferiore si raggiunge in (0,45 , 0,05) e quella superiore in (0,3 , 0,2). Si ottiene allora: $0{,}45 \leq P(E) \leq 0{,}8$.

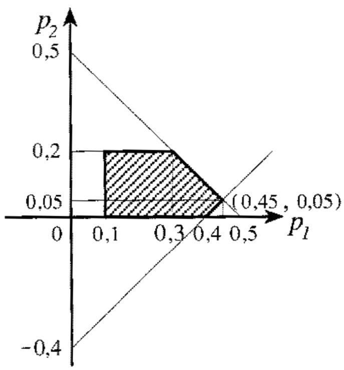

Figura 14

L'evento E_3 del punto *b)* dipende logicamente dalla partizione $\mathbb{P}_G(\{E_1, E_2, E_3\})$, ma la sua valutazione non è stata data. Essa dipende dalla probabilità che viene assegnata su questa partizione mediante le scelte ammissibili di $p_1, \dots, p_6$, quelle ottenute cioè scegliendo la coppia (p_1, p_2) nel pentagono della

Fig. 14 e calcolando le altre p_i in loro funzione mediante le (24). In corrispondenza ad ogni scelta di tale coppia si ottiene $P(E_3)=p_2+p_3+p_6=2p_2+0{,}3$. Per ricavare le limitazioni inferiore e superiore si procede come prima per l'evento E, salvo che ora si devono trovare minimo e massimo di una medesima funzione. Inoltre, poiché la funzione dipende solo da p_2 basta ricordare che il minimo valore consentito per p_2 è 0 e il massimo 0,2. Si ricava allora $0{,}3 \leq P(E_3) \leq 0{,}7$.

2 *Una macchina produce in sequenza 5 oggetti a ogni fornitura di materia prima. Gli oggetti possono essere buoni o difettosi. L'esperienza sulla produzione di numerose sequenze porta a valutare – sulla base delle frequenze osservate – la probabilità degli eventi* ω_n = *«il numero di pezzi buoni nella produzione della prossima sequenza è n»,* $n = 0, 1, \ldots, 5$, *al modo seguente:* $P(\omega_0) = 0.002$, $P(\omega_1) = 0.028$, $P(\omega_2) = 0.077$, $P(\omega_3) = 0.198$, $P(\omega_4) = 0.335$, $P(\omega_5) = 0.360$. *Determinare le limitazioni di probabilità – l'intervallo di coerenza – per i seguenti eventi:*

E_1 = *il primo pezzo prodotto è difettoso,*
E_2 = *il numero dei pezzi buoni è dispari,*
E_3 = *nella produzione dei 5 pezzi si alternano pezzi buoni e difettosi o viceversa.*
E_4 = *gli ultimi 2 pezzi prodotti sono buoni,*
E_5 = *il numero di pezzi buoni è uguale a quello dell'ultima sequenza prodotta.*

Si ha:
$\omega_0 \Rightarrow E_1$, $\omega_5 \Rightarrow \bar{E}_1$, da cui: $0.002 \leq P(E_1) \leq 0.640$,
$E_2 = \omega_1 \vee \omega_3 \vee \omega_3$ è logicamente dipendente da $\{\omega_0, \ldots, \omega_5\}$ e quindi di probabilità determinata: $P(E_2) = P(\omega_1) + P(\omega_3) + P(\omega_3) = 0{,}586$,
$\omega_n \Rightarrow \bar{E}_3$, $n = 0, 1, 4, 5$, da cui $0 \leq P(E_3) \leq 0.275$,
$\omega_5 \Rightarrow E_4$, $\omega_n \Rightarrow \bar{E}_4$, $n = 0, 1$ da cui $0.360 \leq P(E_4) \leq 0.970$,
E_5 è logicamente indipendente da $\{\omega_0, \ldots, \omega_5\}$ e quindi $0 \leq P(E_5) \leq 1$.

Ci chiediamo ora se è coerente attribuire probabilità 0.6 a E_1 e 0.5 a E_4. Entrambe le scelte sono tra quelle ammissibili per i due eventi. Abbiamo però spiegato intuitivamente – in premessa di questi esempi – che la scelta di una valutazione per un evento può obbligare a restringere le scelte per altri. Andiamo verificare che ciò accade nel nostro caso specifico. Costruiamo per questo la partizione $\{\omega_0, \ldots, \omega_5\} \wedge \{E_1, \bar{E}_1\}$, da cui E_1 è logicamente dipendente. Tenuto conto delle implicazioni $\omega_0 \Rightarrow E_1$ e $\omega_5 \Rightarrow \bar{E}_1$, essa è data da:

$$\{\omega_0, \omega_1 \wedge E_1, \ldots, \omega_4 \wedge E_1, \omega_1 \wedge \bar{E}_1, \ldots, \omega_4 \wedge \bar{E}_1, \omega_5\}. \qquad (25)$$

Posto $p_n = P(\omega_n \wedge E_1)$, $n = 1, \ldots, 4$, per coerenza è necessario e sufficiente porre:

$$P(E_1) = 0.002 + p_1 + p_2 + p_3 + p_4, \quad 0 \le p_n \le P(\omega_n).$$

Andiamo a determinare le limitazioni per l'evento E_4 nella nuova situazione. È subito visto che gli eventi elementari che implicano E_4 sono ora ω_5 (come prima) e $\omega_4 \wedge E_1$ (nuovo) e quelli che implicano $\bar{E}_4$ ω_0, $\omega_1 \wedge E_1$, $\omega_1 \wedge \bar{E}_1$ (gli ultimi due al posto di ω_1) e $\omega_2 \wedge \bar{E}_1$ (nuovo). Riesce allora $0.360 + p_4 \le P(E_4) \le 0.970 - (0{,}077 - p_2)$. Nella nuova situazione l'intervallo di coerenza relativo a E_4 non cambia se $p_4 = 0$ e $p_2 = 0{,}077$. Ma allora $P(E_1) = 0.079 + p_1 + p_3$ e l'intervallo di coerenza relativo a E_1 si restringe – e notevolmente –. Riesce $0.079 \le P(E_1) \le 0{,}305$. Circa il quesito iniziale, osserviamo che scelto $P(E_1) = 0{,}6$, si tratta di vedere se sotto questa condizione la minima limitazione per $P(E_4)$ – che si ottiene attribuendo a p_4 il più piccolo valore possibile – è minore di 0.5. Poiché riesce $p_4 = P(E_1) - (0.002 + p_1 + p_2 + p_3) = 0.6 - (0.002 + p_1 + p_2 + p_3)$, tale minimo si ottiene attribuendo a p_1, p_2, p_3 i massimi valori consentiti. Si ricava allora $p_4 = 0.295$. Nelle condizioni di scelta più favorevoli deve quindi essere $P(E_4) \ge 0.655$, e non si può perciò scegliere $P(E_4) = 0.5$. In altri termini, la valutazione $P(E_1) = 0.6$, $P(E_4) = 0.5$ *non è coerente* con quella data agli eventi elementari della partizione $\{\omega_0, \ldots, \omega_5\}$.
Osserviamo ancora che la scelta di una valutazione per E_1 – nei limiti consentiti dalla coerenza – non modifica la probabilità di E_2, che è ben determinata dalla distribuzione assegnata su $\{\omega_0, \ldots, \omega_5\}$, e che resta perciò invariata in ogni prolungamento. Le limitazioni per E_5 rimangono 0, 1, perché questo evento è logicamente indipendente anche dalla partizione (25). Resta uguale a 0 la limitazione inferiore per E_3, mentre quella superiore diventa $0.275 - (0{,}077 - p_2)$, perché alle implicazioni precedenti si aggiunge la $\omega_2 \wedge \bar{E}_1 \Rightarrow \bar{E}_3$.

3 *Con riferimento al bilancio economico nazionale del prossimo anno si considerino i due eventi:*

E_1 = *il reddito nazionale aumenterà almeno del 2%,*

E_2 = *la bilancia dei pagamenti registrerà un saldo passivo.*

Un economista esprime la sua valutazione fornendo i seguenti rapporti di probabilità

$$\frac{P(E_1)}{P(\bar{E}_1)} = 3, \quad \frac{P(E_1)}{P(E_1 \wedge E_2)} = 2{,}5, \quad \frac{P(E_2)}{P(E_1 \wedge E_2)} = 1{,}6$$

e, in termini assoluti, le seguenti probabilità:

$$P(E_1) = 0{,}75, \quad P(E_2) = 0{,}50, \quad P(E_1 \wedge E_2) = 0{,}25, \quad P(E_1 \vee E_2) = 0{,}90.$$

Verificare che quell'economista non è coerente e calcolare la probabilità coerente con la valutazione fatta mediante i tre rapporti.

Una probabilità coerente deve verificare la condizione $P(E_1 \vee E_2) = P(E_1) + P(E_2) - P(E_1 \wedge E_2)$ (9.3.3*e*). Nel nostro caso il primo membro vale 0,90, mentre per il secondo si trova valore 1. La valutazione non è perciò coerente.
Passando alla correzione del calcolo, cominciamo con l'osservare che da $P(E_1)/P(\bar{E}_1) = 3$ si ricava $P(E_1) = 0{,}75$, in accordo con l'economista.
Per ottenere il resto scriviamo $P(E_1) = P(E_1 \wedge E_2) + P(E_1 \wedge \bar{E}_2)$, da cui dividendo membro a membro per $P(E_1 \wedge E_2)$ – sicuramente diverso da 0, data la valutazione dei rapporti – ricaviamo $2{,}5 = 1 + P(E_1 \wedge \bar{E}_2)/P(E_1 \wedge E_2)$ e quindi $P(E_1 \wedge \bar{E}_2)/P(E_1 \wedge E_2) = 1{,}5$. Analogamente, scambiando il ruolo di E_1 e E_2 si trova $P(\bar{E}_1 \wedge E_2)/P(E_1 \wedge E_2) = 0{,}6$. Per trovare il rapporto tra le probabilità di $\bar{E}_1 \wedge \bar{E}_2$ e $E_1 \wedge E_2$ scriviamo

$$P(E_1) = 3\,P(\bar{E}_1) = 3\,P(\bar{E}_1 \wedge E_2) + 3\,P(\bar{E}_1 \wedge \bar{E}_2),$$

da cui, dividendo per $P(E_1 \wedge E_2)$, si ricava $P(\bar{E}_1 \wedge \bar{E}_2)/P(E_1 \wedge E_2) = 7/30$.
Ora con riferimento alla partizione generata da $\{E_1, E_2\}$ deve essere $P(E_1 \wedge E_2) + P(E_1 \wedge \bar{E}_2) + P(\bar{E}_1 \wedge E_2) + P(\bar{E}_1 \wedge \bar{E}_2) = 1$, da cui si ottiene $P(E_1 \wedge E_2)\,(1 + 1{,}5 + 0{,}6 + 7/30) = 1$ e quindi $P(E_1 \wedge E_2) = 0{,}30$. Le probabilità degli altri eventi elementari allora sono $P(E_1 \wedge \bar{E}_2) = 0{,}45$, $P(\bar{E}_1 \wedge E_2) = 0{,}18$, $P(\bar{E}_1 \wedge \bar{E}_2) = 0{,}07$. La risposta corretta pertanto è:

$$P(E_1) = 0{,}75,\ P(E_2) = 0{,}48,\ P(E_1 \wedge E_2) = 0{,}30,\ P(E_1 \vee E_2) = 0{,}93.$$

4 *Per l'incontro di calcio tra le squadre A e B, un allibratore offre per i seguenti eventi*

E_1 = *A vince l'incontro*
E_2 = *l'incontro si conclude con meno di due marcature*
E_3 = *B marca reti e non vince*
E_4 = *una delle due squadre vince marcando almeno due reti,*

le quote unitarie $q(E_1) = 1{,}75$, $q(E_2) = 2{,}30$, $q(E_3) = 3{,}85$, $q(E_4) = 1{,}65$.

Il signor Bianchi – per niente appassionato di calcio – è invitato dagli amici – tutti esperti, perché tifosi accaniti – a partecipare a un gioco in cui il perdente dovrà offrire una cena. Si tratta di un gioco congegnato come segue: ciascun giocatore attribuisce valori v_i agli eventi E_i e viene penalizzato di $p = \sum_{i=1}^{4} (|E_i| - v_i)^2$. Perde chi realizza penalità più alta.

Il signor Bianchi osserva che per realizzare penalità nulla bisogna avere la fortuna di indovinare i valori di tutti gli eventi e porre $v_i = 1$ in corrispondenza degli E_i veri, $v_i = 0$ in corrispondenza di quelli falsi. Giocando in questo modo, però, se si commettono errori – anche uno solo – si viene penalizzati almeno di 1, e si ottiene così la stessa penalità ($p = 1$) o peggio ($p > 1$) di chi ponesse $v_i = 1/2$ per ogni i. Sulla base di queste considerazioni il signor Bianchi si convince: (i) che è troppo rischioso cercare di indovinare il valore degli eventi – dare ai v_i valori 0 o 1 –, (ii) che è ragionevole ritenere che un esperto abbia probabilità elevata di realizzare penalità inferiore a 1, sfruttando per questo in modo adeguato il suo stato d'informazione – visto che può realizzare penalità 1 ignorandolo del tutto –. A lume di naso ritiene che sfruttare lo stato d'informazione significhi porre v_i uguale al proprio grado di fiducia sul verificarsi di E_i ($v_i = P(E_i)$). Lo conforta in questo l'interpretazione nei casi limite, in cui o si sia quasi sicuri di azzeccare il valore di un evento E o al contrario si sia pressoché ugualmente incerti sui due valori possibili. Nel primo caso conviene cercare di indovinare avendo fondata speranza di realizzare penalità 0 – ponendo $v = 1$ ($v = 0$) se si propende per il valore vero (falso) di E –; nel secondo la prudenza consiglia di preferire la penalità certa e abbastanza contenuta $1/4$ – che si ottiene ponendo $v = 1/2$ – piuttosto che correre il rischio di subire penalità elevata[37].
A causa della sua scarsa conoscenza della situazione in esame, per essere competitivo il signor Bianchi decide di basare la sua valutazione affidandosi a quella dell'allibratore – che è senza dubbio quella di un esperto – il quale si è espresso dando sugli eventi E_i le offerte unitarie. Egli sa che le offerte dagli allibratori non sono quelle corrispondenti al loro grado di fiducia. Offrono di meno di quanto dovrebbero, perché vogliono giocare

---

37 Il signor Bianchi ha ragione. Il procedimento seguito per valutare le penalità nel gioco con amici si inquadra nello schema delle penalizzazioni che viene trattato nel § 10.6. La probabilità è ivi interpretata come valore da attribuire agli eventi per rendere minima, secondo le proprie opinioni e informazioni, l'«aspettativa» della penalizzazione. Seguendo questa interpretazione si perviene alla *Definizione* 10.6.3, che risulta equivalente alla *Definizione* 9.2.4 suggerita dallo schema delle scommesse (*Teorema* 10.6.4).

in condizioni di vantaggio (da banco). Il signor Bianchi sa – per aver letto il § 9.1 – che se le offerte fossero eque nel giudizio dell'allibratore, i loro reciproci sarebbero le sue probabilità. Essendo però ritoccate al ribasso, i loro reciproci superano le probabilità medesime. Il signor Bianchi sa in proposito – sempre per aver letto il § 9.1 – che il vantaggio dell'allibratore si misura con la percentuale di allibramento, che viene calcolata sulla base delle offerte fatte per le alternative di una partizione. Nel nostro caso gli eventi $E_1, \dots, E_4$ non sono alternative di una partizione. È verosimile però che le offerte che li riguardano siano state determinate sulla base della valutazione delle offerte degli eventi elementari della partizione da loro generata. Il signor Bianchi ipotizza a questo punto che l'allibratore abbia espresso la sua valutazione di probabilità su $\mathbb{P}_G(\{E_1, \dots, E_4\})$, abbia poi maggiorato la valutazione moltiplicando le probabilità dei corrispondenti eventi elementari per la percentuale di allibramento desiderata, ottenendo così una funzione peso su $\mathbb{P}_G(\{E_1, \dots, E_4\})$ e per additività sui relativi eventi logicamente dipendenti. Abbia determinato infine le offerte per gli eventi $E_1, \dots, E_4$ come reciproci dei rispettivi pesi. Sulla base di queste ipotesi e della conoscenza delle offerte fatte per tali eventi, il signor Bianchi cerca di risalire almeno approssimativamente alle loro probabilità (secondo l'allibratore).
Bisogna cominciare per questo, col determinare la partizione generata da $\{E_1, \dots, E_4\}$ ed è comodo farlo per gradi (6.2.3 *Esempio*). Per la partizione generata da $\{E_1, E_2\}$ abbiamo i seguenti quattro eventi elementari: $E_1 \wedge E_2 =$ *A vince 1-0*, $\bar{E}_1 \wedge E_2 =$ *l'incontro termina 0-0 o B vince 1-0*, $E_1 \wedge \bar{E}_2 =$ *A vince marcando almeno due reti*, $\bar{E}_1 \wedge \bar{E}_2 =$ *l'incontro termina in parità con marcature oppure B vince marcando almeno due reti*. Passando alla partizione $\mathbb{P}_G(\{E_1, E_2, E_3\}) = \mathbb{P}_G(\{E_1, E_2\}) \wedge \{E_3, \bar{E}_3\}$, osserviamo che $E_2 \wedge E_3 = \phi$ e che perciò $E_2 \wedge \bar{E}_3 = E_2$, $\bar{E}_2 \wedge E_3 = E_3$. Questo significa che i due costituenti $E_1' \wedge E_2 \wedge E_3$ sono impossibili. Si ottengono allora i seguenti 6 costituenti possibili:

$\omega_1 = E_1 \wedge E_2 \wedge \bar{E}_3 =$ *A vince 1-0,*
$\omega_2 = \bar{E}_1 \wedge E_2 \wedge \bar{E}_3 =$ *l'incontro termina 0-0 o B vince 1-0,*
$\omega_3 = E_1 \wedge \bar{E}_2 \wedge E_3 =$ *A vince subendo reti,*
$\omega_4 = E_1 \wedge \bar{E}_2 \wedge \bar{E}_3 =$ *A vince h-0,* $h \geq 2$,
$\omega_5 = \bar{E}_1 \wedge \bar{E}_2 \wedge E_3 =$ *l'incontro termina pari con marcature,*
$\omega_6 = \bar{E}_1 \wedge \bar{E}_2 \wedge \bar{E}_3 =$ *B vince marcando almeno due reti.*

Il completamento dello sviluppo col prodotto per la partizione $\{E_4, \bar{E}_4\}$ non serve, perché l'evento E_4 è logicamente dipendente dalla partizione $\{\omega_1, \dots, \omega_6\}$ (*Teorema* 6.3.3). È subito visto infatti che $\omega_1, \omega_2, \omega_5$

implicano $\bar{E}_4$, mentre $\omega_3, \omega_4, \omega_6$ implicano E_4. Tutti gli E_i sono dunque logicamente dipendenti da questa partizione. Si ha: $E_1 = \omega_1 \vee \omega_3 \vee \omega_4$, $E_2 = \omega_1 \vee \omega_2$, $E_3 = \omega_3 \vee \omega_5$, $E_4 = \omega_3 \vee \omega_4 \vee \omega_6$. Indicati con $\pi_i = 1/q(\omega_i)$ i reciproci delle offerte unitarie dell'allibratore sugli eventi elementari ω_i, nelle ipotesi del signor Bianchi riesce

$$\begin{cases} \pi_1 + \pi_3 + \pi_4 = 0{,}5714 \\ \pi_1 + \pi_2 = 0{,}4348 \\ \pi_3 + \pi_5 = 0{,}2597 \\ \pi_3 + \pi_4 + \pi_6 = 0{,}6061 \end{cases}$$

ove i secondi membri delle uguaglianze sono nell'ordine i reciproci delle quote offerte dall'allibratore per gli eventi E_1, E_2, E_3, E_4. Sommando membro a membro la seconda e quarta uguaglianza e aggiungendo π_5 si trova che l'offerta unitaria complessiva sui sei eventi elementari – la percentuale di allibramento – è $1{,}0409 + \pi_5$, con π_5 che non può superare 0,2597. Questa limitazione non è migliorabile. Lo si prova risolvendo il sistema in funzione di π_5 e π_6 – si trova $\pi_1 = -0{,}0337 + \pi_6$, $\pi_2 = 0{,}4685 + \pi_6$, $\pi_3 = 0{,}2597 - \pi_5$, $\pi_4 = 0{,}3464 + \pi_5 - \pi_6$ – e imponendo ai pesi di essere positivi. La percentuale di allibramento è perciò compresa tra 4,09% e 30,06%. Per arrivare a una stima del suo valore, il signor Bianchi deve dare a questo punto una *sua* valutazione sulla ripartizione dell'offerta 0,2597 tra π_3 e π_5. Vista la sua scarsa conoscenza della situazione, il signor Bianchi non si sente di privilegiare nella ripartizione uno dei due pesi – degli eventi elementari ω_3 e ω_5 a cui sono destinati – e decide perciò di ripartire equamente l'offerta 0,2597. Arrotondando, egli arriva così a stimare la percentuale di allibramento pari al 17%. Normalizzando i pesi π_i – dividendoli per 1,17 – e arrotondando alla seconda decimale egli perviene allora alla valutazione della seguenti probabilità per gli E_i (4.4.2 *Proposizione*): $P(E_1) = 0.49$, $P(E_2) = 0.37$, $P(E_3) = 0.22$, $P(E_4) = 0.52$.
Auguriamo a questo punto all'ingegnoso signor Bianchi buona fortuna! In fondo, basta arrivare penultimi.

10.5.2 Probabilità in ambiente numerabile.

La valutazione di probabilità mediante distribuzioni *concentrate*, *obbligatoria* se l'ambiente è finito, ha rilevante interesse nelle applicazioni anche in ambiente numerabile. La situazione che ivi si configura è analoga a quella del caso finito. In entrambi gli ambienti la massa unitaria viene distribuita a pezzetti sugli eventi elementari di $\mathbb{P}$, assegnando probabilità positiva a tutti o parte di essi. La probabilità degli eventi logicamente dipendenti da $\mathbb{P}$ si ottiene poi sommando le probabilità dei loro componenti, in senso ordinario o della serie. La somma per serie sarà utilizzata nella (21) di 10.4.3 *Teorema* – se occorre e sempre nell'ipotesi che la probabilità sia concentrata – anche per ottenere le limitazioni di probabilità per eventi logicamente semidipendenti da una partizione numerabile.

Vi è però un aspetto che in generale rende diversa la situazione del caso numerabile da quella del caso finito. Se $\mathbb{P}$ è numerabile, usando le probabilità concentrate si esprimono sempre giudizi molto sbilanciati. La cosa è stata già segnalata in 5.2.1 *Nota,* dopo aver dato la definizione provvisoria di probabilità nel numerabile in accordo con la nozione classica (5.2.1 *Definizione*). Si è ivi sottolineato che le valutazioni concentrate in ambiente numerabile sono le sole significative quando l'ambiente è una idealizzazione di un ambiente finito, come spesso accade nelle applicazioni. Lo possono essere però anche quando la descrizione riguarda un ambiente autenticamente numerabile. Come ad esempio nel problema del primo successo in una partita a testa e croce, che abbiamo trattato con riferimento a una moneta non perfetta in 5.2.3 *Esempio* 1 e che tratteremo in generale in 13.2.4 *Esempio* 6.

Finché si considerano probabilità concentrate, la teoria delle probabilità coerenti coincide con quella assiomatica di Kolmogorov, come già ricordato in 10.5.1 *Nota*. Analogamente agli esempi del *Cap.* 4 nel caso finito, anche gli esempi presentati nel *Cap.* 5 sono esempi di valutazione coerente dell'incertezza.

Se si ha bisogno di esprimere gradi di fiducia più equilibrati, si dovrà ricorrere a valutazioni mediante probabilità diffuse o parzialmente diffuse, esistenti nella teoria delle probabilità coerenti anche in corrispondenza di partizioni di cardinalità numerabile. Si può ad esempio dare un *giudizio di equiattendibilità* (qualitativo) degli eventi elementari, che per essere equiattendibili *devono* essere equiprobabili, e quindi di probabilità nulla. Di ciò si è già detto brevemente anche in 5.3.5 *Complemento,* ove si è anche sottolineata l'anomalia della teoria classica, che consente di dare senso alla nozione di equiattendibilità – al problema della *scelta a caso* se la situazione di incertezza è descritta come scelta di un oggetto da un insieme (5.3.5 *Nota*) – in ambiente *finito* e in quello *continuo,* ma non in quello *numerabile*, di cardinalità intermedia! È un'anomalia che, potendo scegliere probabilità diffuse anche su partizioni numerabili, nella teoria delle probabilità coerenti si può eliminare. La questione è illustrata per la scelta a caso di un numero naturale nell'esempio qui di seguito e per la scelta di un numero razionale in 11.3.6 *Esempio* 2.

ESEMPIO. *Scelta a caso di un numero naturale.*

Il problema della scelta a caso di un numero aleatorio naturale N ha avuto origine storicamente con lo studio della "frequenza" dei numeri primi nell'insieme $\mathbb{N}$ dei numeri naturali. Cominciamo con l'osservare in proposito che nei primi 10 numeri naturali i primi sono 1, 2, 3, 5, 7 e la loro *frequenza relativa* è perciò $1/2 = 0{,}5$. Nei primi 30 numeri naturali, ai cinque primi precedenti si aggiungono 11, 13, 17, 19, 23, 29, e la frequenza relativa è allora $11/30 = 0{,}366$. Allargando l'analisi ai segmenti iniziali di cento, cinquecento, mille numeri naturali si ottengono nell'ordine le frequenze relative 0,26, 0,192, 0,169. Si nota così che la frequenza relativa decresce al crescere di n, anche se in modo non monotono, dal momento che ogni volta che si incontra un numero primo la frequenza, in quel passo, aumenta. La sensazione che se ne ricava, comunque, è che al crescere di n i numeri primi diventino sempre più radi. Quale sarà il loro comportamento al limite? Ovvero, come si comporta la frequenza relativa al divergere di n? Esiste una frequenza limite? Per rispondere al quesito è stata proposta in letteratura una procedura di applicazione più generale, adatta cioè per studiare il comportamento della frequenza relativa in $\mathbb{N}$ dei numeri di un insieme $I \subset \mathbb{N}$ *generico*. Andremo ora ad esporla interpretandola – com'è naturale nel nostro contesto – in termini di probabilità coerenti.

Cominciamo con l'osservare che la frequenza relativa dei numeri di un sottoinsieme di numeri naturali I nei primi n numeri naturali può essere interpretata come probabilità di scegliere a caso nel segmento iniziale $\{1, \dots, n\}$ un numero di $I \cap \{1, \dots, n\}$. Ragionando nell'ambiente infinito descritto dalla partizione $\mathbb{P} = \{N = i : i \in I\}$, tale frequenza può essere anche interpretata come la probabilità $P_n(N \in I)$, essendo P_n la probabilità concentrata ed equidistribuita sui primi n eventi $(N = 1), \dots, (N = n)$ di $\mathbb{P}$. Facendo tendere n a $+\infty$, per gli $E = (N \in I)$ per cui esiste il limite si ottiene la probabilità asintotica di base $(N = 1, \dots, N = n, \dots)$ (10.3.2 *Definizione*), che "estende" il giudizio di equiattendibilità – espressa con l'equiprobabilità degli elementi dei segmenti iniziali (di cardinalità finita) – a tutti i numeri naturali. Essa è conseguenza diretta dell'approccio dinamico allo studio del comportamento delle frequenze relative, e appare pertanto quella che meglio si presta a interpretare – ai fini di questi problemi – il significato di «scelta a caso di un numero naturale». Lo conferma anche, ad esempio, il calcolo della probabilità degli eventi $E_h = N$ *è multiplo di* h, $h = 2, 3, \dots$, per i quali risulta $P(E_h) = 1/h$. Per ogni n si trovano infatti q e r tali che $n = qh + r$, $r < h$, da cui si ricava $P_n(E_h) = q/(qh + r)$. Poiché q tende a $+\infty$ se n tende a $+\infty$, allora si trova:

$$\lim_{n \to +\infty} P_n(E_h) = \lim_{q \to +\infty} \frac{q}{qh + r} = \frac{1}{h}.$$

Con riferimento alla frequenza dei numeri primi nei naturali, da cui ha preso le mosse il nostro discorso, ci limitiamo a segnalare che essa esiste ed è nulla, che riesce cioè $lim_{n \to +\infty} P_n$(*N è un numero primo*) = 0.
I risultati relativi agli eventi *N è multiplo di h* fanno apparire appropriato il ricorso alla distribuzione asintotica per realizzare la nozione di scelta a caso di un numero naturale. Si tratta tuttavia di una realizzazione parziale, perché le probabilità asintotiche sono definite su un sottoinsieme proprio di $\mathcal{A}_L(\mathbb{P})$ (10.3.2 *Proposizione, d*), che qui prende il nome di insieme di Dynkin. La valutazione dipende inoltre dalla decisione di mantenere ai numeri della base il loro ordinamento naturale. Sappiamo infatti in proposito che modificando l'ordinamento si perviene di solito a risultati differenti (10.3.2 *Proposizione, a*). Ad esempio, ordinando i numeri naturali facendo seguire ogni dispari da due pari, si ottiene $P(E_2) = 2/3$ (conseguentemente è $1/3$ la probabilità che il numero scelto sia dispari). Analogamente, se uno ha "simpatia" per i numeri primi e decide di manifestarla alternando nella base un numero primo con uno non primo, allora ricava che la probabilità che venga scelto un numero primo è $1/2$.
È interessante infine ricordare che i domini delle probabilità asintotiche – e quindi anche l'insieme di Dynkin – non sono algebre (10.3.2 *Proposizione, d*). Ciò significa che la **distribuzione asintotica** – in letteratura questo termine è riservato alla probabilità di questo esempio – *non è una probabilità additiva in senso classico,* se si pretende che sia definita su un'algebra. È però una probabilità coerente, e quindi restrizione di qualche probabilità additiva; se non altro dei suoi prolungamenti coerenti su $\mathcal{A}_L(\mathbb{P})$ (*Teorema* 9.5.1).

COMPLEMENTO. *Valutazione mediante quozienti di probabilità nel numerabile.*

Un modo abbastanza naturale di assegnare la probabilità P su una partizione finita $\{\omega_1, \dots, \omega_n\}$ è quello di esprimere la valutazione in modo indiretto dando i quozienti delle probabilità dei suoi eventi elementari. Abbiamo discusso l'argomento nel § 4.4, mostrando che una tale valutazione si sintetizza con l'introduzione di un'applicazione non negativa π proporzionale alla probabilità, da cui si ottiene la P per normalizzazione, ponendo cioè $P = k\pi$ e calcolando k imponendo che sia $k(\pi(\omega_1) + \dots + \pi(\omega_n)) = 1$ (4.4.2 *Proposizione*). Nella teoria classica questa procedura può essere utilizzata anche nel caso numerabile, perché allora la probabilità è σ-additiva e la somma delle probabilità degli eventi elementari è 1 come nel caso finito (*Proposizione* 5.2.2). In tale teoria è dunque equivalente assegnare la probabilità su partizioni finite o numerabili *direttamente* – distribuendola a pezzetti sugli eventi elementari – o valutandola *indirettamente* – dando i quozienti di probabilità della coppie di eventi elementari –. Se la partizione è finita ciò è vero anche nella teoria delle probabilità

coerenti. Come già segnalato in 5.2.2*Nota,* e come andiamo ora a giustificare, non lo è invece se la partizione è numerabile.
Sia dunque $\mathbb{P} = \{\omega_1, \ldots, \omega_n, \ldots\}$ una partizione numerabile. Le valutazioni coerenti su $\mathbb{P}$ sono tutte e sole quelle che assegnano la probabilità P in modo che riesca $0 \leq \sum_{n=1}^{\infty} P(\omega_n) \leq 1$ (*Proposizione* 10.4.1). Si ha allora intanto che la procedura precedente non può essere applicata se si valutano tutti gli eventi elementari di probabilità nulla. Altrimenti, come fatto sopra, si introdurrà un'applicazione π non negativa, non identicamente nulla e tale che riesca $P = k\pi$. Si ha allora $k\sum_{n=1}^{\infty} \pi(\omega_n) = \sum_{n=1}^{\infty} P(\omega_n) = \gamma$, ove γ è la parte concentrata di P. Questa parte della probabilità non è compresa nella valutazione espressa in termini di quozienti con l'applicazione π. Se non si aggiunge altro, allora, quello che si è in grado di dire in questa situazione è che sono di probabilità positiva tutti e soli gli eventi elementari per cui riesce $\pi(\omega_n) > 0$ e che in loro corrispondenza si ha $P(\omega_h) / P(\omega_n) = \pi(\omega_h) / \pi(\omega_n)$ per ogni h naturale. Per poter dire qual è la probabilità degli ω_n occorre esprimere una valutazione per γ. In tal caso riesce $P(\omega_n) = \gamma\pi(\omega_n) / \sum_{i=1}^{\infty} \pi(\omega_i)$, e si ritrova poi il risultato classico solo se $\gamma = 1$. Altrimenti si conosce la parte concentrata della distribuzione e rimane ovviamente da valutare la parte diffusa $1 - \gamma$.

10.5.3 Coerenza delle probabilità definite da densità.

È sufficiente collegare quanto si è detto in 5.3.3*Proposizioni* 1, 2 per le probabilità definite – ivi in senso intuitivo – mediante integrali di funzioni di densità e il *Teorema* 9.4.2. Entrambe le proposizioni citate del n° 5.3.3 segnalano che le applicazioni definite sull'algebra dei sottoinsiemi misurabili di una regione di $\mathbb{R}^n$ misurabile – secondo Peano-Jordan e secondo Lebesgue, rispettivamente – soddisfano le proprietà *a), b), c)* del *Teorema* 9.4.2 e quindi siamo in grado di dire ora che sono probabilità coerenti. Le applicazioni introdotte usando la misura di Lebesgue sono anche probabilità σ-additive.

10.6 Probabilità come penalizzazione

Ci sono due aspetti che possono mettere a disagio chi è chiamato ad esprimere il suo grado di fiducia sugli eventi di un dato insieme, dovendo poi sottostare alle regole previste nello schema delle scommesse. Da un lato la *scelta della scommessa* che è fatta dal competitore, al quale è concessa la facoltà di decidere su quali eventi scommettere e quanto puntare per ciascuno. Dall'altro l'*obbligo di accettare scommesse* anche quando si ha solo la prospettiva di perdere o pareggiare. Della cosa ci siamo già occupati mostrando che entrambe le sensazioni di disagio possono essere attenuate: usando la norma di coerenza vincolata la prima (*Complemento* 9.2.3) e tenendo conto per la seconda che nei casi in cui si può solo perdere (o vincere) e pareggiare, la probabilità di perdere (vincere) è nulla (*Complemento* 9.5.4).

Anche per questo motivo – ma soprattutto per fornire uno strumento alternativo di misura e di indagine del grado di fiducia – de Finetti ha proposto un secondo schema per la valutazione della probabilità: lo **schema delle penalizzazioni** che, come vedremo, risulta equivalente a quello della scommesse (*Teoremi* 10.6.4 e 10.6.6). In questo schema si attenua la prima situazione di disagio – è lasciata al competitore la facoltà di scegliere gli eventi su cui calcolare la penalizzazione, ma non di influire sul suo valore – e svanisce la seconda, perché non è possibile migliorare la penalizzazione in qualche caso senza peggiorarla in qualche altro (come dire: non è possibile «guadagnare» senza andare incontro alla possibilità di «perdere»).

Come si vede, le considerazioni fatte qui sopra si riferiscono all'utilizzo dello schema delle scommesse e di quello delle penalizzazioni come *strumenti di misura* del grado di fiducia. Esse non hanno perciò alcuna conseguenza di natura teorica. La teoria si sviluppa, infatti, sulla base dell'*assioma* 9.2.4 o di quello equivalente 10.6.3, i quali, proprio in quanto assiomi, non hanno bisogno di essere interpretati a tutti i costi.

10.6.1 Le penalità.

Nello schema delle penalizzazioni è previsto che chi è chiamato ad esprimere il suo grado di fiducia sugli eventi di un insieme $\mathscr{E}$ sia obbligato a farlo sapendo di dover subire una *penalità* calcolata in corrispondenza a un insieme di eventi $\{E_1, \dots, E_n\} \subset \mathscr{E}$ scelto da un competitore, essendo la penalità medesima definita come segue.

Definizione. Penalità di un punto di $\mathbb{R}^n$ su $(E_1, \dots, E_n)$.

Siano $x_1, \dots, x_n$ numeri reali. Dicesi ***penalità*** *del punto <u>certo</u> $(x_1, \dots, x_n)$ su $(E_1, \dots, E_n)$ il quadrato della sua distanza dal punto <u>aleatorio</u> $(|E_1|, \dots, |E_n|)$: $L_x = \sum_{i=1}^{n} (|E_i| - x_i)^2$.*

Il punto aleatorio (n-pla aleatoria) $(|E_1|, \dots, |E_n|)$ – e quindi anche la penalità – è definito sulla partizione generata $\mathcal{P}_G(\{E_1, \dots, E_n\})$ dall'applicazione che fa corrispondere all'evento elementare $E_1' \wedge \dots \wedge E_n'$ il punto $(v_1, \dots, v_n)$, con $v_i = 0$ se $E_i' = \bar{E}_i$ e $v_i = 1$ se $E_i' = E_i$. Valutare la penalità L_x significa

allora calcolare le distanze del punto $(x_1, \ldots, x_n)$ da tutti i punti possibili $(v_1, \ldots, v_n)$. Messo questo in chiaro, risulta allora evidente che il punto $(x_1, \ldots, x_n)$ è sicuramente preferibile al punto $(y_1, \ldots, y_n)$, ai fini della penalità su $(E_1, \ldots, E_n)$, se il primo è più vicino del secondo a *tutti* i punti possibili di $(|E_1|, \ldots, |E_n|)$. Se riesce cioè $L_x < L_y$. Non esistono invece ragioni *oggettive* per giudicare uno dei due punti preferibile all'altro, se $(x_1, \ldots, x_n)$ è più vicino a qualche punto possibile di $(|E_1|, \ldots, |E_n|)$ e $(y_1, \ldots, y_n)$ a qualche altro. In questo caso un ordine di preferibilità può essere espresso solo in base a valutazioni *soggettive*, dipendenti dal grado di fiducia che si dà al verificarsi dei punti possibili.
Ciò premesso, andiamo ora ad approfondire l'analisi fatta dal signor Bianchi nella parte iniziale di 10.5.1 *Esempio* 4. Osserviamo intanto che la penalità di x su Ω è $(|\Omega| - x)^2 = (1-x)^2$, e che la si può annullare ponendo $x = 1 = P(\Omega)$ – scelta non migliorabile –. Analogamente la penalità di x su ϕ è x^2 e la si annulla ponendo $x = 0 = P(\phi)$. Sicché, nei casi di *certezza* sul valore dell'evento, la scelta più conveniente per il punto x è quella di porlo uguale alla probabilità dell'evento, valutata secondo lo schema delle scommesse. Se E è possibile, i punti possibili in $\mathbb{R}$ del punto aleatorio $|E|$ sono 0, 1. La penalità di x su $|E|$ è x^2 se E è falso e $(1-x)^2$ se E è vero. Allora, se x è esterno all'intervallo $[0, 1]$, entrambi i valori della penalità possono essere migliorati, avvicinando ad esempio x all'estremo più prossimo. Si può migliorare invece solo uno dei due valori, e a scapito dell'altro, se $0 \leq x \leq 1$. Le scelte dei valori esterni a $[0, 1]$ non sono perciò convenienti. Per quelli appartenenti a $[0, 1]$ non sussistono elementi oggettivi per stabilire un ordine di preferibilità. Vanno perciò considerati tutti ammissibili. Come accennato in generale, una scelta effettiva sarà fatta sulla base di opinioni e valutazioni soggettive. Possiamo concordare in proposito col signor Bianchi che conviene rischiare scegliendo valori prossimi a 0 (1) se siamo quasi sicuri che E sia falso (vero), rischiare sempre di meno – spostandoci verso il punto medio dell'intervallo – man mano che l'incertezza sul valore logico di E aumenta, fino a scegliere al limite $x = 1/2$ quando si è ugualmente incerti su tale valore (che è poi la probabilità che si dà in questa ipotesi ad E nello schema delle scommesse).
Le considerazioni sin qui svolte inducono a pensare che la scelta più conveniente sia quella di porre sempre $x = P(E)$. Ciò è avvalorato dalla prossima argomentazione, dal *Teorema* 10.6.4, ma soprattutto dal *Teorema* 10.6.6. Per il momento comunque abbiamo visto che conviene richiedere che i punti (numeri) x ammissibili ai fini della penalità su E soddisfino le proprietà di normalizzazione e di non negatività – come la probabilità nello schema della scommesse–. Una ulteriore analisi ci permette di aggiungere che conviene richiedere che essi soddisfino anche la proprietà additiva. Supponiamo infatti

che gli eventi E_1 e E_2 siano incompatibili e andiamo a esaminare la penalità di un generico punto (x_1, x_2, x) su $(E_1, E_2, E_1 \vee E_2)$. I punti possibili del punto aleatorio $(|E_1|, |E_2|, |E_1 \vee E_2|)$ sono $(0,0,0)$, $(1,0,1)$, $(0,1,1)$, ai quali corrisponde in $\mathbb{R}^3$ il triangolo appartiene al piano di equazione $x = x_1 + x_2$, com'è facile verificare. Per le proprietà degli insiemi convessi ogni punto (di $\mathbb{R}^3$) non appartenente al triangolo può essere "avvicinato" a tutti tre i suoi vertici, diminuendo così la penalità qualunque cosa accada. Ciò non succede invece per i punti appartenenti al triangolo, nessuno dei quali può essere infatti "avvicinato" a qualche suo vertice senza essere "allontanato" da qualche altro[38]. I punti di tale triangolo sono perciò – tutti e soli – quelli ammissibili ai fini della penalizzazione su $(|E_1|, |E_2|, |E_1 \vee E_2|)$ – cioè quelli che non si possono escludere perché riconosciuti con criteri oggettivi meno convenienti di qualche altro punto –. Le loro coordinate verificano la condizione $x = x_1 + x_2$, come preannunziato.

10.6.2 Probabilità come penalità.

La discussione fatta nel numero precedente suggerisce le due seguenti definizioni, che corrispondono alle 9.2.2 *Definizione* 1, *Definizione* 2 dello schema delle scommesse.

Definizione 1. Probabilità nello schema delle penalizzazioni.

La probabilità di un evento E è il valore P(E) che si ritiene conveniente attribuire in corrispondenza ad E, sapendo di dover subire la penalità (aleatoria) $L = (|E| - P(E))^2$.

---

38 Ogni poliedro convesso $\mathcal{C}$ di $\mathbb{R}^n$ è *caratterizzato* dalla proprietà ora usata, che in generale si esprime come segue:

> *un poliedro $\mathcal{C}$ è convesso se e solo se i suoi punti non possono essere "spostati" senza aumentare la distanza da qualche suo vertice e diminuirla da qualche altro.*

Interessa inoltre, per i prossimi sviluppi di questo paragrafo, richiamare anche quest'altra caratterizzazione:

> *i punti del poliedro convesso $\mathcal{C}$ si ottengono tutti e soli come mistura dei suoi vertici.*

Definizione 2. Norma di coerenza.

Una valutazione di probabilità P definita su un insieme di eventi $\mathcal{E}$ dicesi coerente se e solo se qualunque siano $E_1, \dots, E_n \in \mathcal{E}$, non esiste un punto $(x_1, \dots, x_n) \in \mathbb{R}^n$, la cui penalità su $(E_1, \dots, E_n)$ sia minore di quella del punto $(P(E_1), \dots, P(E_n))$.

Come abbiamo fatto nello schema delle scommesse, possiamo anche qui esprimere le due definizioni del numero precedente in forma assiomatica dando la seguente definizione.

10.6.3 Definizione. Probabilità in via assiomatica nello schema delle penalizzazioni.

Sia $\mathcal{E}$ un insieme non vuoto di eventi e P un'applicazione di $\mathcal{E}$ in $\mathbb{R}$. Diremo che P è una probabilità coerente (secondo lo schema delle penalizzazioni), se e solo se, qualunque siano $E_1, \dots, E_n \in \mathcal{E}$, posto $L_P = \sum_{i=1}^{n} (|E_i| - P(E_i))^2$ e $L_x = \sum_{i=1}^{n} (|E_i| - x_i)^2$, non esistono $x_1, \dots, x_n \in \mathbb{R}$ tali che sia $L_x < L_P$.

Complemento.

Dalla definizione segue subito che se una probabilità è coerente su un insieme $\mathcal{E}$, essa lo è anche su ogni suo sottoinsieme $\mathcal{E}'$, perché ogni penalità L_x su eventi di $\mathcal{E}'$ è anche una penalità su eventi di $\mathcal{E}$.

Nota.

Al pari di quanto accade nello schema delle scommesse, anche qui la valutazione P è di natura *soggettiva,* come già del resto messo in evidenza in premessa discutendo di valutazione in casi particolari, mentre la norma di coerenza è di natura *oggettiva.* La verifica della coerenza di un'applicazione P di dominio $\mathcal{E}$ è fatta infatti eseguendo calcoli algebrici (calcolando distanze di punti in qualche spazio $\mathbb{R}^n$).

10.6.4 Teorema.

L'applicazione P definita sull'insieme di eventi $\mathcal{E}$ è coerente secondo lo schema delle penalizzazioni se e solo se lo è anche secondo lo schema delle scommesse.

DIMOSTRAZIONE.

Poiché in entrambi gli schemi la condizione di coerenza è definita con riferimento ai sottoinsiemi finiti di $\mathscr{E}$, è sufficiente provare che l'equivalenza enunciata sussiste per le probabilità definite su un suo sottoinsieme generico $\{E_1, \dots, E_n\} \subset \mathscr{E}$. Si tratta allora di provare che gli insiemi delle probabilità coerenti su $\{E_1, \dots, E_n\}$ secondo l'uno e l'altro schema sono uguali. È utile a questi fini considerare l'insieme $\{E_1, \dots, E_n\}$ ordinato, perché allora le probabilità coerenti su di esso – o meglio su $(E_1, \dots, E_n)$ – sono definite da n-ple di numeri reali e si prestano perciò ad essere rappresentate geometricamente come punti di $\mathbb{R}^n$.
Ciò premesso, cominciamo col determinare l'insieme delle probabilità coerenti secondo lo schema delle penalizzazioni. La linea dimostrativa è sostanzialmente quella che abbiamo visto nei casi particolari del n° 10.6.1. Abbiamo ivi osservato preliminarmente che il punto aleatorio $(/E_1/, \dots, /E_n/)$ è definito dall'applicazione di dominio $\mathbb{P}_G(\{E_1, \dots, E_n\})$ che fa corrispondere al generico evento elementare $E_1' \wedge \dots \wedge E_n'$ il punto $(v_1, \dots, v_n)$ di $\mathbb{R}^n$, con $v_i = 0$ se $E_i' = \bar{E}_i$ e $v_i = 1$ se $E_i' = E_i$. I punti possibili che così si ottengono sono vertici di un *poliedro convesso* $\mathscr{D}$ – sottoinsieme dell'ipercubo unitario (l'ipercubo medesimo se $E_1, \dots, E_n$ sono logicamente indipendenti) –. I punti $(x_1, \dots, x_n)$ che corrispondono a valutazioni coerenti su $(E_1, \dots, E_n)$ sono per definizione quelli che non si possono "avvicinare" a tutti i vertici del poliedro $\mathscr{D}$, e quindi tutti e soli i punti di $\mathscr{D}$ stesso (*Nota* 38 a piè di pagina), il quale è perciò l'insieme di tutte e sole le valutazioni coerenti su $(E_1, \dots, E_n)$ secondo lo schema delle penalizzazioni.
Passando ora alla determinazione dell'insieme delle probabilità coerenti secondo lo schema delle scommesse, osserviamo intanto che esse sono tutte e sole quelle che si ottengono come restrizione su $\{E_1, \dots, E_n\}$ delle probabilità coerenti su $\mathscr{A}_L(\mathbb{P}_G(\{E_1, \dots, E_n\})$. Per la 10.5.1 *Proposizione* che le caratterizza, queste sono tutte e sole le applicazioni

$$P(\cdot) = \sum P(E_1' \wedge \dots \wedge E_n')\, \delta(E_1' \wedge \dots \wedge E_n' \Rightarrow \cdot), \tag{26}$$

ove la sommatoria si intende estesa agli eventi elementari di $\mathbb{P}_G(\{E_1, \dots, E_n\})$ – anche nel seguito della dimostrazione –, $\delta(E_1' \wedge \dots \wedge E_n' \Rightarrow \cdot)$ è la probabilità concentrata su $E_1' \wedge \dots \wedge E_n'$, $P(E_1' \wedge \dots \wedge E_n')$ sono numeri non negativi di somma 1 e per il resto arbitrari – le probabilità che vengono assegnate agli eventi elementari –. Osserviamo ora che la probabilità $\delta(E_1' \wedge \dots \wedge E_n' \Rightarrow \cdot)$ determina su $(E_1, \dots, E_n)$, tramite la (26), la probabilità

$$(\delta(E_1' \wedge \dots \wedge E_n' \Rightarrow E_1), \dots, \delta(E_1' \wedge \dots \wedge E_n' \Rightarrow E_n)), \tag{27}$$

che nella rappresentazione geometrica è il vertice di $\mathcal{D}$ corrispondente all'evento elementare $E_1' \wedge \ldots \wedge E_n'$. Misturando i vettori (27) con i coefficienti $P(E_1' \wedge \ldots \wedge E_n')$ e tenendo conto della (26), per ogni scelta delle probabilità degli eventi elementari – di una probabilità su $\mathbb{P}_G(\{E_1, \ldots, E_n\}))$ – si trova il vettore di componenti $P(E_h) = \sum P(E_1' \wedge \ldots \wedge E_n')\, \delta(E_1' \wedge \ldots \wedge E_n' \Rightarrow E_h)$, $h = 1, \ldots, n$. In termini geometrici, le probabilità coerenti su $(E_1, \ldots, E_n)$ si ottengono dunque tutte e sole come mistura dei vertici di $\mathcal{D}$, e il loro insieme è perciò $\mathcal{D}$ stesso (*Nota* 38 a piè di pagina), come nello schema delle scommesse.

Con ciò la tesi è provata. ∎

COROLLARIO.

Sia $\mathcal{E}$ finito, $\mathcal{E} = \{E_1, \ldots, E_n\}$. Allora, affinché un'applicazione P sia una probabilità coerente secondo lo schema delle penalizzazioni è sufficiente esaminare la penalità del punto $(P(E_1), \ldots, P(E_n))$ su $(E_1, \ldots, E_n)$ – non occorre cioè estendere l'analisi ai sottoinsiemi di $\mathcal{E}$ –.

DIMOSTRAZIONE.

Se la penalità del punto $(P(E_1), \ldots, P(E_n))$ su $(E_1, \ldots, E_n)$ verifica la condizione di coerenza secondo lo schema delle penalizzazioni, allora esso appartiene al poliedro convesso $\mathcal{D}$ introdotto nella dimostrazione del teorema, in virtù del quale la valutazione $(P(E_1), \ldots, P(E_n))$ è coerente anche secondo lo schema della scommesse. Riesce allora coerente secondo lo schema delle scommesse – e quindi pure secondo lo schema delle penalizzazioni (per il teorema usato all'inverso) – anche la valutazione $(P(E_{i_1}), \ldots, P(E_{i_s}))$ per ogni $\{E_{i_1}, \ldots, E_{i_s}\} \subset \{E_1, \ldots, E_n\}$. ∎

In virtù del teorema precedente, che afferma che ogni probabilità coerente secondo la *Definizione* 9.2.4 è coerente anche secondo la *Definizione* 10.6.3, la teoria delle probabilità coerenti può essere sviluppata sulla base di uno o entrambi gli assiomi da esse introdotti. Volendo anzi – per precisione formale – si potrebbe scegliere una delle due definizioni come assioma e considerare l'altra come un suo teorema. Noi abbiamo sin qui sviluppato la teoria sulla base dell'assioma 9.2.4 e continueremo a farlo anche in seguito, perché, dei due, è l'assioma che si è rivelato più maneggevole per gli sviluppi teorici. Occorrerà però per questo perfezionare la *Definizione* 9.2.4 per renderla adatta alla trattazione del problema della valutazione degli eventi condizionati (*Definizione* 13.1.2)..

Circa l'uso dei due schemi come strumento di misura del grado di fiducia, tenuto conto delle considerazioni fatte in apertura di questo paragrafo, può essere che per motivi di ordine psicologico si preferisca lo schema delle penalizzazioni o che la scelta possa variare a seconda delle circostanze. È importante però sapere che se si seguono le regole comportamentali elencate nel prossimo numero 10.6.5 i due schemi risultano equivalenti anche sotto questo aspetto operativo. Il successivo *Teorema* 10.6.6 mostra infatti che seguendo tali regole si perviene – almeno in teoria – allo stesso risultato.

10.6.5 Convenzioni e notazioni.

A fini interpretativi, supporremo che il soggetto abbia una sua scala di preferibilità sugli importi aleatori definita dalla relazione binaria *X è preferito almeno quanto Y,* che indicheremo con $X \underset{p}{\geq} Y$ o $Y \underset{p}{\leq} X$, soddisfacente le seguenti proprietà:

a) *dati due numeri aleatori X, Y riesce o* $X \underset{p}{\geq} Y$ *o* $Y \underset{p}{\geq} X$;

b) *la relazione è riflessiva e transitiva;*

c) $X \underset{p}{\geq} Y$ *se e solo se* $X+Z \underset{p}{\geq} Y+Z$, *per ogni numero aleatorio Z;*

d) $X \underset{p}{\geq} Y$ *se e solo se* $aX \underset{p}{\geq} aY$ *se* $a>0$ *e* $aY \underset{p}{\geq} aX$ *se* $a<0$;

e) *se X e Y sono numeri certi,* $\underset{p}{\geq}$ *coincide con* $\geq$ *dei numeri reali.*

Le proprietà sono analoghe a quelle della relazione $\geq$ dei numeri reali, la quale però oltre che riflessiva e transitiva è anche antisimmetrica, cioè tale che «$x \geq y$ *e* $y \geq x$ *se e solo se* $x = y$».
Qui la condizione «$X \underset{p}{\geq} Y$ e $Y \underset{p}{\geq} X$» definisce invece la relazione «*X è preferito quanto Y*», che indicheremo $X \underset{p}{\sim} Y$. È facile provare che $\underset{p}{\sim}$ è una equivalenza e che anche per essa sussistono le proprietà *c)* e *d)* – potendo esprimere la *d)* più semplicemente mediante la «*se* $a \neq 0$, *allora* $X \underset{p}{\sim} Y$ *se e solo se* $aX \underset{p}{\sim} aY$» –. Inoltre, se X e Y sono numeri certi, $\underset{p}{\sim}$ coincide con $=$.
La condizione «$X \underset{p}{\geq} Y$ *e non* $Y \underset{p}{\geq} X$» definisce invece la preferibilità in senso forte, cioè la relazione *X è preferito a Y,* che indicheremo con $X \underset{p}{>} Y$ o con $Y \underset{p}{<} X$. Si può facilmente provare che essa è transitiva, che soddisfa le proprietà *c)* e *d)* e che diventa la relazione $>$ o $<$ del campo reale se X e Y sono entrambi numeri certi.

10.6.6 Teorema.

Se un soggetto interpreta nell'assioma 9.2.4 *la P come quota di scommessa e nell'assioma* 10.6.3 *la penalità L come perdita monetaria, allora le sue rispettive valutazioni sono uguali.*

DIMOSTRAZIONE.

Sia $\mathcal{E}$ un insieme di eventi, P_1 (P_2) una probabilità su $\mathcal{E}$ secondo lo schema delle scommesse (penalizzazioni). Si tratta allora di provare che se si seguono le regole di preferibilità del n° 10.6.5, interpretando P ed L come previsto nell'enunciato, si deve porre $P_1(E) = P_2(E)$ per ogni $E \in \mathcal{E}$. Per semplicità di scrittura poniamo a tal fine $P_1(E) = p_1$, $P_2(E) = p_2$.
Allora riesce $p_1 \underset{p}{\sim} |E|$, perché p_1 è la quota di scommessa unitaria su E. e $L_{p_2} \underset{p}{\leq} L_{p_1}$, perché L_{p_2} è la penalità preferita tra le penalità L_p. Pertanto si ha $(|E|-p_2)^2 \underset{p}{\leq} (|E|-p_1)^2$, da cui segue $|E|^2-2p_2|E|+p_2^2 \underset{p}{\leq} |E|^2-2p_1|E|+p_1^2$ e quindi $(p_1-p_2)\,|E| \underset{p}{\leq} (p_1-p_2)(p_1+p_2)/2$. Condizione questa che sussiste se $p_1=p_2$ (ovvio) e solo allora. Infatti se $p_1<p_2$ ($p_1>p_2$) si ricava $|E| \underset{p}{\geq} (p_1+p_2)/2>p_1$ ($E| \underset{p}{\leq} (p_1+p_2)/2<p_1$), da cui, sostituendo il simbolo $>$ ($<$) con $\underset{p}{>}$ ($\underset{p}{<}$) – perché i numeri a confronto sono reali – si ottiene $|E| \underset{p}{>} p_1$ ($|E| \underset{p}{<} p_1$) (assurdo!). ■