

Capitolo 13

PROBABILITÀ CONDIZIONATE COERENTI

La nozione di *probabilità condizionata coerente* (Definizione 13.1.2) generalizza quella di *probabilità coerente*. Si ispira a uno schema di scommesse su eventi condizionati – una scommessa su $E|H$ è valida solo se l'ipotesi H è vera – adeguando la norma di coerenza (n° 13.1.1). Il *Teorema delle probabilità composte* (n° 13.1.4) – *condizione necessaria* per la coerenza – è il "ponte" che collega le probabilità di eventi definiti da una *stessa proposizione* in uno stato d'informazione *iniziale* e in uno *finale*, ed è equivalente alla definizione provvisoria di probabilità di $E|H$ se $P(H) > 0$ (Definizione 12.4.2). Nel § 13.2 si estende all'ambiente condizionato il *Teorema del prolungamento* (n° 13.2.1) e lo si utilizza per determinare le estensioni coerenti da $\mathcal{E}|H$ a $\mathcal{E}|H \wedge K$ (\mathcal{E} insieme di eventi assoluti): la scelta è *unica* se $P(H|K) > 0$ (*Teorema* 13.2.2), le scelte sono *arbitrarie* se $P(H|K) = 0$ (*Teorema* 13.2.3). Si studia poi il problema delle valutazioni *sequenziali* coerenti in ambiente finito (n° 13.2.5) – valutazione dei costituenti di una partizione prodotto usando il Teorema delle probabilità composte generalizzato (n° 13.1.5) –. Nel § 13.3 si prova che la proprietà di *chiusura per passaggio al limite* sussiste anche per le probabilità condizionate (*Teorema* 13.3.2) – non quella di *convessità* (13.3.1 *Contro-esempio*) –, si estende il metodo asintotico all'ambiente condizionato (n° 13.3.3) e si approfondisce la nozione di scelta a caso di un numero naturale (13.3.3 *Complemento*). Nel § 13.4 si studia la possibilità di ottenere valutazioni di probabilità assolute attraverso valutazioni condizionate (per *disintegrazione*). Nel § 13.5 si introduce la nozione di *conglomerabilità* – $P(E)$ compresa tra gli estremi delle $P(E|H)$, $H \in \mathcal{P}$ – e si studiano i suoi collegamenti con la disintegrabilità. Il § 13.6 chiude il capitolo mostrando che è possibile prolungare la valutazione in modo coerente anche in *difetto* (non in *errore!*) di deduzione (*deduzione tardiva*).

13.1 Nozione di probabilità condizionata coerente

L'ampia analisi del problema dell'evoluzione della valutazione fatta nel § 12.4 in termini intuitivi ha permesso di far emergere con sufficiente chiarezza i collegamenti che si ritiene

ragionevole richiedere come essenziali tra valutazioni in uno stato d'informazione iniziale (probabilità iniziali) e valutazioni in uno stato d'informazione incrementato (probabilità finali). Non può allora essere soddisfacente una nozione di probabilità condizionata che non tenga conto di queste esigenze. Risponde a questo requisito la nozione di probabilità condizionata coerente che andremo ora a introdurre estendendo in modo del tutto naturale la nozione di probabilità coerente (assoluta).

In accordo con l'impostazione sin qui seguita, si deve allora pensare di dover dare una valutazione di probabilità su un di insieme *eventi condizionati* \mathcal{D} non vuoto e per il resto qualsiasi (non soggetto cioè ad alcun vincolo, né di cardinalità né di struttura). È ovvio poi pensare di inquadrare il problema traendo ispirazione da uno schema di scommesse, che deve essere per forza più generale del precedente, in grado di tenere conto del fatto che gli eventi condizionati perdono di significato se si apprende che le loro ipotesi sono false.

13.1.1 Schema delle scommesse per valutazioni condizionate.

Cominciamo col dire cosa significa scommettere su un evento condizionato $E|H$. In analogia col caso degli eventi assoluti, si tratta ora di pensare che sia stata fissata una quota unitaria di scommessa p , pagando la quale si ha diritto di ricevere 1 (0) se *essendo vero* H , è vero E (\bar{E}) e di vedersi restituire la quota se H è (risulterà) falso. In altri termini, si paga p per ricevere 1 se è vero $E \wedge H$, 0 se è vero $\bar{E} \wedge H$ e p se è vero \bar{H} – il discorso è fatto con riferimento allo stato d'informazione iniziale α –. Il guadagno è allora (8.4.2 Esempio 1):

$$G = 1|E \wedge H| + 0|\bar{E} \wedge H| + p|\bar{H}| - p = \\ |E||H| + p(1 - |H|) - p = (|E| - p)|H|.$$

In parallelo con quanto fatto nello schema delle scommesse per valutazioni assolute, diamo ora le seguenti definizioni:

DEFINIZIONE 1. Scommessa su $E|H$.

Siano E, H eventi, $H \neq \emptyset$, p la quota di scommessa unitaria di $E|H$. Una

scommessa su $E|H$ di puntata pS è quella che dà luogo al guadagno $G = S(|E| - p)|H|$.

DEFINIZIONE 2. Combinazione di scommesse su $E_1|H_1, \dots, E_n|H_n$.

Siano $E_1, \dots, E_n, H_1, \dots, H_n$ eventi, H_1, \dots, H_n non impossibili, p_1, \dots, p_n le quote di scommessa unitarie di $E_1|H_1, \dots, E_n|H_n$. Una combinazione di scommesse (o scommessa) su questi ultimi n eventi, di puntate rispettive p_1S_1, \dots, p_nS_n , è quella che dà luogo al guadagno $G = \sum_{i=1}^n S_i(|E_i| - p_i)|H_i|$.

In analogia col caso assoluto (9.2.2 Definizione 1) qui le quote di scommessa unitarie vengono interpretate come probabilità dei corrispondenti eventi condizionati. Si dà cioè la seguente definizione.

DEFINIZIONE 3. Probabilità condizionata nello schema della scommesse.

La *probabilità (condizionata)* di un evento $E|H$ – per un soggetto – è la quota $P(E|H)$ che egli giudica equo scambiare con $|E \wedge H| + p|H|$.

Circa la *norma di coerenza*, bisogna ora tenere presente che il guadagno G può essere nullo perché nessuna delle scommesse sui singoli eventi della combinazione è valida – sono false tutte le ipotesi H_i – e le puntate vengono perciò tutte restituite. Il fatto che il guadagno sia nullo non rappresenta in generale una controindicazione. Lo è però in questo caso specifico, perché allora la condizione $\min G \leq 0 \leq \max G$ è soddisfatta comunque si scelgano le quote unitarie, anche al di fuori dell'intervallo $[0, 1]$. Nell'attuale contesto tale condizione non può essere presa come norma di coerenza condizionata senza pregiudizio per l'interpretazione delle quote unitarie come probabilità. Il rimedio consiste nel richiedere – come proposto da S. Holzer [6] – che le condizioni sul minimo e massimo vengano poste per il guadagno che si realizza quando *almeno una* delle scommesse è valida. Posto $H_0 = H_1 \vee \dots \vee H_n =$ *almeno una scommessa è valida*, al posto di G definito sulla partizione $\mathcal{P}_G(\{E_1, \dots, E_n, H_1, \dots, H_n\})$, si tratta cioè di considerare il guadagno condizionato $G|H_0$ di dominio $\mathcal{P}_G(\{E_1, \dots, E_n, H_1, \dots, H_n\})|H_0$ ⁴⁴. Si dà allora la seguente definizione.

44 Per una formulazione equivalente della norma di coerenza condizionata si veda E. Regazzini [9].

13.1.2 Definizione. Probabilità condizionata in via assiomatica

Sia \mathcal{D} un insieme non vuoto di eventi condizionati e $P(\cdot|\cdot)$ un'applicazione di \mathcal{D} in \mathbb{R} . Diremo che $P(\cdot|\cdot)$ è una **distribuzione di probabilità condizionata coerente** (anche **probabilità condizionata o solo probabilità**) se e solo se qualunque siano $E_1|H_1, \dots, E_n|H_n \in \mathcal{D}$ e $S_1, \dots, S_n \in \mathbb{R}$, posto

$$G = \sum_{i=1}^n S_i (|E_i| - P(E_i|H_i)) |H_i| \quad \text{e} \quad H_0 = \bigvee_{i=1}^n H_i,$$

riesce:

$$\min G|H_0 \leq 0 \leq \max G|H_0 \quad \text{ovvero} \quad \min G|H_0 \max G|H_0 \leq 0.$$

OSSERVAZIONE.

La definizione di probabilità condizionata coerente generalizza quella di probabilità coerente (assoluta). Si ritrova infatti la *Definizione 9.2.4* ponendo nella definizione attuale $\mathcal{D} = \mathcal{E}|\Omega$, con \mathcal{E} insieme di eventi assoluti, e osservando che $\mathcal{E}|\Omega = \mathcal{E}$ (12.3.3 a) e che per ogni combinazione di scommesse riesce $H_0 = \Omega$ e quindi $G|H_0 = G$.

Circa la struttura del guadagno $G|H_0$, osserviamo che G è un numero aleatorio trasformato degli indicatori $|E_1|, \dots, |E_n|, |H_1|, \dots, |H_n|$ e come tale è definito sulla partizione generata $\mathbb{P}_G(\{E_1, \dots, E_n, H_1, \dots, H_n\})$, cioè sugli *eventi elementari* del tipo $E'_1 \wedge \dots \wedge E'_n \wedge H'_1 \wedge \dots \wedge H'_n$, e che l'evento $H_0 = H_1 \vee \dots \vee H_n$ dipende logicamente da H_1, \dots, H_n ed è somma dei costituenti della partizione $\mathbb{P}_G(\{H_1, \dots, H_n\})$ del tipo $H'_1 \wedge \dots \wedge H'_n$ con almeno un H_i affermato. Di conseguenza riesce

$$(E'_1 \wedge \dots \wedge E'_n \wedge \bar{H}_1 \wedge \dots \wedge \bar{H}_n)|H_0 = (E'_1 \wedge \dots \wedge E'_n \wedge \bar{H}_0)|H_0 = \phi|H_0$$

e gli eventi elementari di $\mathbb{P}_G(\{E_1, \dots, E_n, H_1, \dots, H_n\})|H_0$ – quelli su cui è definito il guadagno $G|H_0$ – sono perciò del tipo $(E'_1 \wedge \dots \wedge E'_n \wedge H'_1 \wedge \dots \wedge H'_n)|H_0$ con almeno uno degli H_i affermato (quindi tali che almeno una scommessa sia valida).

COMPLEMENTO.

Ai fini della verifica della coerenza valgono gli accorgimenti previsti nel caso assoluto nei punti (i) e (ii) di 9.2.4 *Complemento*, sostituendo ivi G con $G|H_0$ e

interpretando gli insiemi di eventi come insiemi di eventi condizionati. Si ha cioè:

- i) la condizione di coerenza è verificata per una data combinazione di scommesse sugli eventi $E_1|H_1, \dots, E_n|H_n$ se per qualche evento elementare di $\mathbb{P}_G(\{E_1, \dots, E_n, H_1, \dots, H_n\})|H_0$ riesce $G|H_0 = 0$;
- (ii) se $P(\cdot|\cdot)$ è coerente su un insieme \mathcal{D} , allora è coerente anche su ogni suo sottoinsieme.

Sussiste inoltre (adattata) anche la 9.2.4 *Proposizione* (e la si prova in modo del tutto analogo a quello visto nel caso assoluto): la condizione $\min G|H_0 \leq 0$ ($\max G|H_0 \geq 0$) per ogni combinazione di scommesse è equivalente alla $\min G|H_0 \leq 0 \leq \max G|H_0$ della definizione.

Non vale invece, purtroppo, la proprietà (iii) di 9.2.4 *Complemento*. Al contrario di quanto accade per le probabilità assolute, la verifica della coerenza di $P(\cdot|\cdot)$ su un insieme finito $\mathcal{D} = \{E_1|H_1, \dots, E_n|H_n\}$ non può essere limitata all'esame delle scommesse in cui intervengono *tutti* gli eventi dell'insieme, perché le scommesse sugli eventi di sottoinsiemi di \mathcal{D} non possono essere qui interpretate come particolari scommesse su $E_1|H_1, \dots, E_n|H_n$. Invero, se si completa una combinazione di scommesse sugli eventi di un sottoinsieme di \mathcal{D} aggiungendo scommesse di vincita (e puntata) *nulla* in corrispondenza a eventi di \mathcal{D} mancanti nella combinazione, il guadagno G resta invariato, come già accadeva nel caso assoluto. Quella che *può cambiare*, però, è l'ipotesi condizionante relativa a tale guadagno – e con essa il guadagno condizionato – che nella scommessa di partenza è la somma logica H_0 di *parte* delle ipotesi H_1, \dots, H_n , mentre in quella completata è $H_1 \vee \dots \vee H_n$, somma logica di *tutte* le ipotesi. Sicché, se si sostituisce una scommessa su parte degli eventi di \mathcal{D} con la corrispondente scommessa completata (aggiungendo scommesse di vincita nulla), si finisce con l'esaminare $G|H_1 \vee \dots \vee H_n$ al posto di $G|H_0$, con $H_0 \Rightarrow H_1 \vee \dots \vee H_n$ (3.3.2j). Se $H_0 \neq H_1 \vee \dots \vee H_n$ – come può essere – si ha allora che l'immagine di $G|H_0$ è inclusa nell'immagine di $G|H_1 \vee \dots \vee H_n$ (12.3.5 *Nota 1*), e può allora capitare che sia $\min G|H_1 \vee \dots \vee H_n \leq 0$ ma non $\min G|H_0 \leq 0$. Esaminando $G|H_1 \vee \dots \vee H_n$ al posto di $G|H_0$ si concluderebbe allora che la $P(\cdot|\cdot)$ soddisfa le condizioni di coerenza in corrispondenza del sottoinsieme di eventi assegnati, mentre non le soddisfa.

Si badi che ciò non significa che per verificare se un'applicazione $P(\cdot|\cdot)$ è coerente su un insieme di eventi \mathcal{D} non si possa ragionare scegliendo *arbitrariamente* $E_1|H_1, \dots, E_n|H_n$ in \mathcal{D} e in loro corrispondenza *arbitrariamente* S_1, \dots, S_n in \mathbb{R} (anche nulli). Quello che non è corretto fare *in generale* è verificare se $P(\cdot|\cdot)$ soddisfa la condizione di coerenza in corrispondenza a un *assegnato* insieme di eventi (finito), considerando combinazioni di scommesse

come casi particolari di combinazione più ampie, ottenute scegliendo vincite *arbitrarie* per gli eventi assegnati e *nulle* per gli eventi aggiunti. Lo si può fare *in casi particolari*. Ad esempio in ambiente assoluto, perché allora per ogni combinazione di scommesse il guadagno da esaminare è sempre un guadagno assoluto. Lo si può fare anche se \mathcal{D} è un insieme di eventi condizionati a una medesima ipotesi H , perché allora è $H_0 = H$ per ogni combinazione di scommesse e i guadagni da esaminare sono tutti condizionati ad H .

13.1.3 Valutazioni coerenti su eventi condizionati a una sola ipotesi.

Nel caso particolare in cui tutti gli eventi di \mathcal{D} siano condizionati a una stessa ipotesi H , $\mathcal{D} = \mathcal{E}|H$ con \mathcal{E} insieme di eventi assoluti, le proprietà della probabilità $P(\cdot|H)$ sono quelle delle probabilità assolute. Si tratta in effetti di sostituire nei ragionamenti gli eventi del tipo $E = p|_{\alpha}$ con gli eventi del tipo $E|H = p|_{\alpha \wedge h}$ e i guadagni assoluti con i guadagni condizionati all'ipotesi H (come già segnalato al termine del precedente 13.1.2 *Complemento*). Concettualmente non cambia nulla. Tutto ciò che abbiamo detto per le probabilità assolute, vale anche per le probabilità condizionate a una sola ipotesi, e non è certamente il caso di rinenunciarle. Cambia però la prospettiva del discorso, perché rispetto al caso assoluto, in cui il quadro delle possibilità è statico (si ragiona in uno stato d'informazione *fissato*), ora si pensa di essere stati in uno stato d'informazione α e di trovarsi nello stato d'informazione incrementato $\alpha \wedge h$, e ciò induce a stabilire qualche confronto tra gli eventi del quadro iniziale e quelli del quadro finale.

Andiamo a riformulare in termini condizionati, a titolo d'esempio, alcune delle condizioni necessarie del *Teorema 9.3.3*.

Traducendo la 9.3.3c, per l'*additività* condizionata ad H si ha:

$$\text{se } E_1|H \wedge E_2|H = \phi|H, \text{ allora } P(E_1|H \vee E_2|H) = P(E_1|H) + P(E_2|H).$$

Per metterla in collegamento col quadro delle possibilità iniziale, osserviamo che supporre che $E_1|H \wedge E_2|H = \phi|H$ è equivalente a dire che $E_1 \wedge E_2 \wedge H = \phi$ (n° 12.3.4). Allora, l'*additività* condizionata ad H si può anche esprimere mediante la:

$$\text{se } E_1 \wedge E_2 \wedge H = \phi, \text{ allora } P(E_1 \vee E_2|H) = P(E_1|H) + P(E_2|H).$$

Analogamente, la proprietà di *monotonia* si esprime mediante la (traduzione della 9.3.3h)

$$\text{se } E_1|H \Rightarrow E_2|H, \text{ allora } P(E_1|H) \leq P(E_2|H)$$

oppure mediante la

$$\text{se } E_1 \wedge H \Rightarrow E_2 \wedge H, \text{ allora } P(E_1|H) \leq P(E_2|H),$$

che tiene conto che $E_1|H \Rightarrow E_2|H$ se e solo se $E_1 \wedge H \Rightarrow E_2 \wedge H$.

La proprietà di *normalizzazione* si può scrivere $P(\Omega|H) = 1$, ma anche $P(H|H) = 1$, che è la forma preferita, visto che se $H \Rightarrow E$ è $E|H = E \wedge H|H = H|H$, e H è allora il minimo evento che soddisfa questa proprietà.

Nelle considerazioni fatte fino a questo momento abbiamo avuto modo di verificare che la nozione di probabilità condizionata coerente estende quella di probabilità assoluta (13.1.2 *Osservazione*) e che le probabilità coerenti su insiemi di eventi condizionati a una medesima ipotesi sono del tutto simili (per quel che riguarda le loro proprietà) alla nozione di probabilità assoluta coerente. Si tratta di due esigenze che non potevano essere disattese. La prima infatti è richiesta dal fatto che gli eventi assoluti sono particolari eventi condizionati (12.3.3a). La seconda dal fatto che sono formalmente simili le descrizioni dell'incertezza. Anche se portano a quadri delle possibilità differenti, vengono fatte però con eventi che differiscono solo per lo stato d'informazione, ma non per le proposizioni che li definiscono.

Resta ancora da verificare che la nozione risponde anche alla terza esigenza: quella che collega valutazioni fatte in stati d'informazione diversi, uno più specifico dell'altro. Essa è espressa dalla relazione che abbiamo assunto come definizione provvisoria (*Definizione* 12.4.2), che qui conviene scrivere nella forma moltiplicativa $P(E \wedge H) = P(H) P(E|H)$, per renderla valida anche per $P(H) = 0$ (da $P(H) = 0$ e $E \wedge H \Rightarrow H$ segue infatti $P(E \wedge H) = 0$). Occorre anzi che la *Definizione* 13.1.2 risponda a questa esigenza in senso più generale, per tenere conto che nella realtà a un primo incremento d'informazione possono seguirne altri. Bisogna cioè sapersi muovere con coerenza partendo da uno stato d'informazione K già incrementato – o meglio da $\alpha \wedge K$ – per arrivare ad uno ancora più specifico $K \wedge H$. Il prossimo teorema risponde appieno a questa esigenza.

Vi è poi ancora l'esigenza – fondamentale dal punto di vista metodologico e più volte nominata – che sia possibile valutare in modo coerente gli eventi di ogni insieme e prolungare la valutazione su ogni soprainsieme. Questo accade ed è conseguenza del teorema del prolungamento che abbiamo dimostrato per le probabilità assolute (*Teorema 9.5.1*) e che, come abbiamo ivi anticipato, sussiste anche per le probabilità condizionate. L'enunciato è lo stesso – salvo che per gli insiemi che sono ora insiemi di eventi condizionati – e simile la dimostrazione, che vedremo comunque con sufficiente dettaglio nel prossimo paragrafo.

13.1.4 Teorema delle probabilità composte.

Siano E, H, K eventi (assoluti), $K \wedge H \neq \emptyset$. Allora, condizione necessaria affinché l'applicazione $P(\cdot|\cdot)$ sia una probabilità condizionata coerente è che sussista la:

$$P(E \wedge H|K) = P(H|K) P(E|K \wedge H). \quad (35)$$

DIMOSTRAZIONE.

Sia $\mathcal{D} = \{E \wedge H|K, H|K, E|K \wedge H\}$. Posto $p' = P(E \wedge H|K)$, $p'' = P(H|K)$, $p = P(E|K \wedge H)$, consideriamo le scommesse di puntate rispettive $p'S', p''S'', pS$ sugli eventi di \mathcal{D} . Si ottiene allora il guadagno:

$$G = S'(|E \wedge H| - p')|K| + S''(|H| - p'')|K| + S(|E| - p)|K \wedge H| = \\ ((S + S')|E \wedge H| + (S'' - pS)|H| - (p'S' + p''S''))|K|.$$

Riesce $H_0 = K \vee (K \wedge H) = K$. Scegliendo allora $S' = -S$, $S'' = pS$, si fa in modo che il guadagno diventi certo condizionatamente a K . Si ha cioè:

$$G|K = -(p'S' + p''S'') = S(p' - p''p).$$

Per rispettare la norma di coerenza si deve allora porre $p' = p''p$, e segue allora la (35). ■

COROLLARIO.

Se P è una probabilità coerente su $\{H, E \wedge H, E|H\}$ ($H \neq \emptyset$), allora sussiste la

(ponendo $K = \Omega$ nella (35)):

$$P(E \wedge H) = P(H) P(E|H) \quad (36)$$

PROPOSIZIONE.

Se P è una probabilità coerente su $\{E|H, E|K\}$ ($H \neq \emptyset$) ed $E \Rightarrow H \Rightarrow K$, allora riesce $P(E|K) \leq P(E|H)$.

DIMOSTRAZIONE.

Poiché $E = E \wedge H$ e $H = H \wedge K$, si ha:

$$P(E|K) = P(E \wedge H|K) = P(H|K) P(E|H) \leq P(E|H). \quad \blacksquare$$

COMMENTO.

Il grado di fiducia su un evento tende a rafforzarsi al crescere dello stato d'informazione. Nell'ipotesi però – si badi bene – che E implichi i due stati d'informazione e che lo stato d'informazione più specifico implichi quello più generico.

13.1.5 Generalizzazione del teorema delle probabilità composte.

Siano E_1, \dots, E_n eventi, H un'ipotesi, $H \wedge E_1 \wedge \dots \wedge E_{n-1} \neq \emptyset$, P probabilità coerente su $\{E_1|H, E_2|H \wedge E_1, \dots, E_n|H \wedge E_1 \wedge \dots \wedge E_{n-1}\}$. Allora si ha:

$$P(E_1 \wedge \dots \wedge E_n|H) = P(E_1|H) P(E_2|H \wedge E_1) \dots P(E_n|H \wedge E_1 \wedge \dots \wedge E_{n-1}). \quad (37)$$

DIMOSTRAZIONE.

La (37) ha senso, perché l'ipotesi $H \wedge E_1 \wedge \dots \wedge E_{n-1} \neq \emptyset$ implica che tutti i suoi eventi condizionanti siano diversi da \emptyset . Ciò premesso, la prova si raggiunge per induzione.

Ponendo nella (35) $E = E_2$, $H = E_1$ e $K = H$ si ottiene la (37) per $n = 2$ (base dell'induzione).

Passo induttivo.

Sia $P(E_1 \wedge \dots \wedge E_{n-1}|H) = P(E_1|H) \dots P(E_{n-1}|H \wedge E_1 \wedge \dots \wedge E_{n-2})$. Ponendo nella (35) $E = E_n$ e $H = E_1 \wedge \dots \wedge E_{n-1}$ e $K = H$, si ottiene

$$P(E_1 \wedge \dots \wedge E_n|H) = P(E_1 \wedge \dots \wedge E_{n-1}|H) P(E_n|H \wedge E_1 \wedge \dots \wedge E_{n-1}),$$

da cui per l'ipotesi induttiva segue la (37). ■

13.2 Prolungamenti coerenti

Abbiamo ricordato nel paragrafo precedente che il teorema del prolungamento sussiste anche per le probabilità condizionate e che lo si dimostra seguendo la linea vista nel caso assoluto, salvo dover ragionare in termini condizionati anziché assoluti. Il che comporta naturalmente qualche adattamento che sarà comunque segnalato adeguatamente.

13.2.1 Teorema. Prolungamento delle probabilità condizionate coerenti.

Siano \mathcal{D}_0 e \mathcal{D} insiemi di eventi condizionati, $\mathcal{D}_0 \subset \mathcal{D}$, $P_0(\cdot|\cdot)$ una probabilità coerente su \mathcal{D}_0 . Esiste allora una probabilità coerente $P(\cdot|\cdot)$ su \mathcal{D} , prolungamento di $P_0(\cdot|\cdot)$.

DIMOSTRAZIONE.

Come detto in premessa, la dimostrazione è simile a quella che si consegue per induzione transfinita nel caso assoluto (Teorema 9.5.1). Il *passo induttivo* è praticamente uguale. Basta ivi interpretare nelle coppie del tipo (\mathcal{F}, P) il simbolo \mathcal{F} come insieme di eventi condizionati e P come probabilità condizionata.

Nel *passo base*, in cui si tratta di provare che è possibile prolungare $P_0(\cdot|\cdot)$ su $\mathcal{D}_0 \cup \{E|H\}$, $E|H \notin \mathcal{D}_0$, si deve essenzialmente tenere conto che le condizioni di coerenza vanno riferite qui a guadagni condizionati anziché assoluti. In particolare, le definizioni di valutazione p «troppo grande» e «troppo piccola» per $E|H$ vanno date con riferimento alle combinazioni di scommesse su eventi di $\mathcal{D}_0 \cup \{E|H\}$ includenti la scommessa di vincita 1 su $E|H$, dicendo che la valutazione p è *troppo grande* se $G_p|K_0 < 0$ e *troppo piccola* se $G_p|K_0 > 0$, essendo G_p il guadagno della generica combinazione di scommesse e K_0 la somma logica delle ipotesi degli eventi presenti nella combinazione. Al posto dell'espressione (19) del guadagno assoluto abbiamo qui

$$G_p = (|E|-p)|H| + \sum_{i=1}^n S_i(|E_i|-P_0(E_i))|H_i|. \quad (38)$$

con $E_1|H_1, \dots, E_n|H_n$ arbitrari in \mathcal{D}_0 e S_1, \dots, S_n arbitrari in \mathbb{R} . Ponendo $S_1 = \dots = S_n = 0$ si trova $G_p = (|E|-p)|H|$, cioè il guadagno della scommessa su $E|H$ (di vincita 1), che ai fini dell'esame della coerenza della valutazione p non può però essere considerata come caso particolare della combinazione di

scommesse (38) (13.1.2 *Complemento*). Va considerata cioè come scommessa singola. Riesce allora $K_0 = H$ e si trova perciò (12.3.5 *Teorema, Corollario*):

$$G_p|H = |E|\backslash H - p|H|\backslash H = |E|\backslash H| - p|H|\backslash H| = |E|\backslash H| - p,$$

espressione analoga alla $G_p = |E| - p$ del caso assoluto. Analogamente a quel caso segue allora che i numeri negativi sono troppo piccoli e quelli maggiori di 1 troppo grandi.

Per provare i punti (i) e (ii) – che «se p' è troppo piccolo (grande), allora anche $p < p'$ ($p > p'$) è troppo piccolo (grande)» e che «gli insiemi dei numeri troppo piccoli e di quelli troppo grandi sono entrambi semirette aperte» – si deve considerare la differenza dei guadagni condizionati $(G_p - G_{p'})|K_0$, ove ora $K_0 = H \vee (H_1 \vee \dots \vee H_n) = H \vee H_0$. Si ha (12.3.5 *Corollario*)

$$(G_p - G_{p'})|H \vee H_0 = (p' - p)|H|\backslash |H \vee H_0| = (p' - p)|H|\backslash |H \vee H_0|,$$

da cui, essendo $(G_p - G_{p'})|H \vee H_0 = G_p|H \vee H_0 - G_{p'}|H \vee H_0$ (12.3.5 *Teorema*), si ottiene:

$$G_p|H \vee H_0 - G_{p'}|H \vee H_0 = (p' - p)|H|\backslash |H \vee H_0|.$$

Poiché $|H|\backslash |H \vee H_0|$ è 1 o 0, da qui si ricava che se $p < p'$ riesce $G_p|H \vee H_0 \geq G_{p'}|H \vee H_0$ e la disuguaglianza *debole* inversa se $p > p'$. La differenza col caso assoluto è che ivi le disuguaglianze valgono in senso *stretto*. Ciò non cambia però nella sostanza la linea dimostrativa.

Per quel che riguarda il punto (i) si ha: se $p < p'$ e p' è troppo piccolo risulta $G_p|(H \vee H_0) \geq G_{p'}|(H \vee H_0) > 0$ e allora anche p è troppo piccolo; se $p > p'$ e p' è troppo grande $G_p|(H \vee H_0) \leq G_{p'}|(H \vee H_0) < 0$ e allora anche p è troppo grande.

Per provare il punto (ii) nel caso dell'insieme dei troppo piccoli – per quello dei troppo grandi il discorso è simmetrico –, si suppone $\min G_{p'}|(H \vee H_0) > 0$ (p' troppo piccolo), si sceglie ε tale che $0 < \varepsilon < \min G_{p'}|(H \vee H_0)$ e si pone $p = p' + \varepsilon$. Si trova allora $G_p|(H \vee H_0) = G_{p'}|(H \vee H_0) - \varepsilon|H|\backslash |H \vee H_0| > 0$ e quindi $p > p'$ è troppo piccolo. L'insieme dei troppo piccoli è perciò una semiretta aperta come asserito.

Infine, per provare il punto (iii) – i numeri troppo piccoli sono minori dei troppo grandi –, come nel caso assoluto si prende un p troppo piccolo e un p' troppo grande, e si considera una scommessa su $E|H, E_1|H_1, \dots, E_n|H_n \in \mathcal{D}_0$ e $S_1, \dots, S_n \in \mathbb{R}$ con valutazione p – il cui guadagno G è dato dal secondo membro della (38) – e una scommessa su $E|H, A_1|K_1, \dots, A_r|K_r \in \mathcal{D}_0$ e $S'_1, \dots, S'_r \in \mathbb{R}$ con valutazione p' , il cui guadagno G' è dato da:

$$G' = (|E| - p')|H| + \sum_{i=1}^r S'_i (|A_i| - P_0(A_i))|K_i|$$

e si trova così il numero aleatorio

$$G - G' = (p' - p) |H| + \sum_{i=1}^n S_i (|E_i| - P_0(E_i)) |H_i| - \sum_{i=1}^r S'_i (|A_i| - P_0(A_i)) |K_i|,$$

che è descritto dall'applicazione a secondo membro di dominio

$$\mathbb{P}_G = \mathbb{P}_G(\{E_1, \dots, E_n, A_1, \dots, A_r, H, H_1, \dots, H_n, K_1, \dots, K_r\}),$$

in cui $G_0 = \sum_{i=1}^n S_i (|E_i| - P_0(E_i)) |H_i| + \sum_{i=1}^r (-S'_i) (|A_i| - P_0(A_i)) |K_i|$ è il guadagno di una scommessa su eventi di \mathcal{D}_0 .

Posto $H_0 = H_1 \vee \dots \vee H_n$ e $K_0 = K_1 \vee \dots \vee K_r$, per ipotesi riesce $G |H \vee H_0 > 0$ e $G' |H \vee K_0 < 0$. Allora si ha $G = 0$ ($G' = 0$) se le ipotesi H, H_1, \dots, H_n (H, K_1, \dots, K_r) sono tutte false e $G > 0$ ($-G' > 0$) altrimenti. Perciò, per ogni evento elementare della partizione \mathbb{P}_G in cui almeno una delle ipotesi $H, H_1, \dots, H_n, K_1, \dots, K_r$ è vera riesce $G - G' > 0$ – in quanto G o (vel) $-G'$ è positivo –. Pertanto si ha:

$$0 < (G - G') |H \vee H_0 \vee K_0 = (p' - p) |H |H \vee H_0 \vee K_0| + G_0 |H \vee H_0 \vee K_0|,$$

da cui si ottiene $(p - p') |H |H \vee H_0 \vee K_0| < G_0 |H \vee H_0 \vee K_0|$ e quindi

$$\begin{aligned} \min [(p - p') |H |H \vee H_0 \vee K_0|] &= \min (p - p', 0) < \\ \min G_0 |H \vee H_0 \vee K_0| &\leq \min G_0 |H_0 \vee K_0| \leq 0, \end{aligned}$$

ove la penultima disuguaglianza vale per 12.3.5 Nota 1 e l'ultima per la coerenza di P_0 su \mathcal{D}_0 . Deve allora essere $\min (p - p', 0) < 0$, cioè $p - p' < 0$ e quindi $p < p'$, come volevasi. ■

In virtù del teorema appena dimostrato, le procedure per la dimostrazione della coerenza basate sul teorema del prolungamento, descritte nel *Complemento 9.5.2* nel caso delle probabilità assolute, si possono applicare senza modifiche anche in ambiente condizionato. Abbiamo così ad esempio che se una probabilità è coerente su un insieme e il *dover* rispettare qualche condizione necessaria *obbliga* a prolungarla in un solo modo su un soprainsieme, allora il prolungamento è *necessariamente* una probabilità coerente (l'unica coerente con quella iniziale). Tenendo conto di ciò, siamo ad esempio in grado di provare che la condizione (35) del *Teorema 13.1.4* – e perciò anche la (36) del suo corollario –, oltre che *necessaria* affinché $P(E | K \wedge H)$ sia coerente con una valutazione data su $\{E \wedge H | K, H | K\}$, è pure

sufficiente se $P(H|K) > 0$. Ciò è conseguenza del fatto che in tal caso la condizione *necessaria* (35) si può esprimere in modo equivalente in forma di rapporto con la $P(E|K \wedge H) = P(E \wedge H|K)/P(H|K)$ e la probabilità dell'evento $E|K \wedge H$ viene così univocamente determinata dalle probabilità assegnate agli eventi $E \wedge H|K$ e $H|K$.

Da un punto di vista applicativo interessa estendere il discorso facendo riferimento a un ambiente più generale. Mostrando cioè che sussiste il seguente teorema, per la cui dimostrazione basta tener conto che nelle sue ipotesi vale il discorso precedente per ogni evento $E|K \in \mathcal{E}|K$ che compare nel suo enunciato.

13.2.2 Teorema.

Siano $\mathcal{E}|K$ un insieme di eventi condizionati a K , $H \in \mathcal{E}$, $H \wedge K \neq \emptyset$, $P_0(\cdot \wedge H|K)$ una probabilità coerente sull'insieme $\mathcal{E} \wedge H|K$, $P_0(H|K) > 0$. Allora l'applicazione $P(\cdot|A)$ definita dalla

$$P(\cdot|A) = \begin{cases} P_0(\cdot|K) & \text{se } A = K \\ \frac{P_0(\cdot \wedge H|K)}{P_0(H|K)} & \text{se } A = K \wedge H \end{cases}$$

è l'unica probabilità (condizionata) coerente su $\mathcal{D} = \mathcal{E} \wedge H|K \cup \mathcal{E}|K \wedge H$, prolungamento (estensione) di $P_0(\cdot|K)$.

COROLLARIO ($K = \Omega$).

Siano \mathcal{E} un insieme di eventi (assoluti), $H \in \mathcal{E}$ e $H \neq \emptyset$, $P_0(\cdot)$ una probabilità coerente su $\mathcal{E} \wedge H$, $P_0(H) > 0$. Allora l'applicazione $P(\cdot|H)$ definita da

$$P(E|H) = \frac{P_0(E \wedge H)}{P_0(H)}, \quad \text{per ogni } E|H \in \mathcal{E}|H \quad (39)$$

è l'unica estensione (condizionata) coerente di $P_0(\cdot)$ su $\mathcal{E}|H$.

Una immediata conseguenza del corollario è la seguente proposizione.

PROPOSIZIONE.

Siano P una probabilità su $\mathcal{A}_L(\mathcal{P})$, $\mathcal{P}^+ = \{H \in \mathcal{A}_L(\mathcal{P}) : P(H) > 0\}$. Allora l'applicazione $P(\cdot|\cdot)$ definita dalla (39) è l'unico prolungamento coerente di P su $\mathcal{A}_L(\mathcal{P})|\mathcal{P}^+$.

TERMINOLOGIA.

L'applicazione $P(\cdot|A)$ del teorema è stata definita a "tratti», su parti disgiunte ed esaustive del suo dominio \mathcal{D} , ponendola alternativamente uguale a due assegnate applicazioni, la prima di dominio $\mathcal{E} \wedge H|K$ e la seconda di dominio $\mathcal{E}|K \wedge H$. Capiterà spesso nel seguito di dover operare in modo analogo. Di dovere cioè considerare applicazioni definite a tratti su domini a due a due disgiunti e anche non. In quest'ultimo caso occorrerà naturalmente che le applicazioni siano **compatibili**, cioè che le loro restrizioni sull'intersezione dei loro domini coincidano.

Per agevolare l'esposizione conviene allora introdurre una terminologia adeguata. Se $\{P_i : i \in I\}$ è un insieme di applicazioni di rispettivi domini \mathcal{D}_i definite in modo compatibile, diremo **estensione congiunta delle applicazioni P_i** l'applicazione P di dominio $\mathcal{D} = \cup_{i \in I} \mathcal{D}_i$ che si ottiene ponendo $P = P_i$ su \mathcal{D}_i per ogni $i \in I$. Quella indicata nel teorema con $P(\cdot|A)$, ad esempio, è l'estensione congiunta delle due applicazioni $P_0(\cdot|K)$ di dominio $\mathcal{E}|K$ e $P_0(\cdot \wedge H|K)|P_0(H|K)$ di dominio $\mathcal{E} \wedge H|K$.

Nel precedente *Teorema* 13.2.2, l'ipotesi $P(H|K) > 0$ (precisamente $P_0(H|K) > 0$) – assieme alla conoscenza della probabilità su $\mathcal{E} \wedge H|K$ – obbliga la valutazione su $\mathcal{E}|K \wedge H$. Se $P(H|K) = 0$, invece, riesce anche $P(E \wedge H|K) = 0$ per ogni E – poiché $E \wedge H|K \Rightarrow H|K$ – e quindi sono di probabilità nulla tutti gli eventi di $\mathcal{E} \wedge H|K$. Il problema del prolungamento su $\mathcal{E}|K \wedge H$ non è allora trattabile seguendo l'impostazione del teorema precedente, perché diventa qui inutilizzabile il teorema delle probabilità composte nella forma di rapporto. I rapporti tra $P(E \wedge H|K)$ e $P(H|K)$ sono tutti indeterminati e questa indeterminazione di natura aritmetica si traduce in una libertà di scelta della probabilità su $\mathcal{E}|K \wedge H$ (nello stato d'informazione finale $K \wedge H$), rispetto a quella assegnata su $\mathcal{E}|K$ (nello stato d'informazione iniziale K). Ciò è quanto afferma il prossimo teorema.

13.2.3 Teorema.

Siano \mathcal{E} un insieme di eventi assoluti, K una ipotesi, $H \in \mathcal{E}$, $H \wedge K \neq \emptyset$, $K \not\Rightarrow H$, $P(\cdot|K)$ una probabilità coerente su $\mathcal{E}|K$, $P(H|K) = 0$, $P'(\cdot|K \wedge H)$ una probabilità coerente su $\mathcal{E}|K \wedge H$. Allora l'estensione congiunta di $P(\cdot|K)$ e $P'(\cdot|K \wedge H)$ è una probabilità coerente su $\mathcal{D} = \mathcal{E}|K \cup \mathcal{E}|K \wedge H$.

DIMOSTRAZIONE.

Il guadagno della generica scommessa su \mathcal{D} si scrive:

$$G = \sum_{i=1}^n S_i(|E_i| - P(E_i|K))|K| + \sum_{h=1}^s S'_h(|A_h| - P'(A_h|K \wedge H))|K \wedge H|.$$

Posto $G = G_1 + G_2$, con G_1 guadagno espresso dalla prima sommatoria e G_2 quello espresso dalla seconda, bisogna analizzare $G|H_0$, ove H_0 può essere $K \wedge H$ oppure $K = K \vee (K \wedge H)$.

Si presenta il caso $H_0 = K \wedge H$ quando mancano scommesse su eventi di $\mathcal{E}|K$. Allora $G = G_2$ è il guadagno di una scommessa su eventi di $\mathcal{E}|K \wedge H$ (13.1.2 *Complemento*). Poiché $P'(\cdot|K \wedge H)$ è coerente per ipotesi su $\mathcal{E}|K \wedge H$, si ottiene: $\min G|H_0 = \min G_2|K \wedge H \leq 0$. La norma di coerenza è perciò verificata, intanto, per tutte le scommesse di questo tipo.

Si presenta invece il caso $H_0 = K$ quando esistono scommesse su $\mathcal{E}|K$. Per provare che la norma di coerenza è soddisfatta anche in questo caso – e con ciò la tesi – mostreremo che la probabilità $P(\cdot|K)$ si prolunga in modo unico, e perciò coerente, su $\mathcal{E}|K \cup (\mathcal{E} \wedge H)|K$ e che il guadagno G si può interpretare anche come guadagno di una scommessa su eventi di questo insieme.

Circa il prolungamento della probabilità su $\mathcal{E}|K \cup (\mathcal{E} \wedge H)|K = \{\mathcal{E} \cup \mathcal{E} \wedge H\}|K$, abbiamo già notato in premessa che nell'ipotesi $P(H|K) = 0$, segue obbligatoriamente $P(E \wedge H|K) = 0$ per ogni evento E , e quindi anche per ogni $E \in \mathcal{E}$. La probabilità $P(\cdot|K)$ si prolunga allora come detto in modo unico da $\mathcal{E}|K$ a $\mathcal{E} \wedge H|K$. Indicheremo ancora con $P(\cdot|K)$ questo prolungamento.

La possibilità di interpretare G come guadagno di una scommessa su $\{\mathcal{E} \cup \mathcal{E} \wedge H\}|K$, è dovuto poi al fatto che è costante e nulla la $P(\cdot|K)$ su $\mathcal{E} \wedge H|K$. Posto $S' = -\sum_{h=1}^s S'_h P'(A_h|K \wedge H)$, questo permette, infatti, di modificare la scrittura di G_2 in modo da presentarlo come guadagno di una scommessa su $\mathcal{E}|K \cup (\mathcal{E} \wedge H)|K$ (non serve modificare la scrittura di G_1 , perché è già il guadagno di una scommessa su $\mathcal{E}|K$). Precisamente si ha:

$$G_2 = \sum_{h=1}^s S'_h(|A_h| - P'(A_h|K \wedge H))|K \wedge H| =$$

$$\sum_{h=1}^s S'_h (|A_h \wedge H| - |H| P'(A_h|K \wedge H)) |K| = \\ \sum_{h=1}^s S'_h (|A_h \wedge H| - P(A_h \wedge H|K)) |K| + S'(|H| - P(H|K)) |K|.$$

Scritto in questa forma, G_2 risulta essere il guadagno di una scommessa sugli eventi $A_1 \wedge H|K, \dots, A_s \wedge H|K \in \mathcal{E} \wedge H|K$ e sull'evento $H|K \in \mathcal{E}|K$.

In conclusione, come preannunziato, $G = G_1 + G_2$ può essere letto come guadagno di una scommessa su eventi di $\mathcal{E}|K \cup \mathcal{E} \wedge H|K$. Poiché $P(\cdot|K)$ è coerente su questo insieme, segue: $\min G|H_0 = \min G|K \leq 0$ e ciò completa la prova. ■

CONVENZIONI.

Per agevolare l'esposizione conviene estendere anche in ambiente condizionato le convenzioni fatte in 9.5.1 *Terminologia* in relazione ai prolungamenti in ambiente assoluto. Così, se P_0 è una probabilità o applicazione su \mathcal{D}_0 , talvolta diremo che « P è un prolungamento di P_0 su \mathcal{D} » sottintendendo che il dominio di P è $\mathcal{D}_0 \cup \mathcal{D}$. Inoltre, quando i domini dei prolungamenti sono chiaramente espressi o facilmente sottintesi, useremo spesso lo stesso simbolo per indicare la probabilità (applicazione) iniziale e i suoi prolungamenti.

13.2.4 Esempi.

Gli esempi di questo numero fanno riferimento a modelli d'estrazione da un'urna o ad essi assimilabili (ad esempio il lanci di uno o più dadi, lanci di una moneta) e vengono risolti usando il *Teorema 13.1.5* (delle probabilità composte generalizzato). Fatta eccezione per il secondo quesito dell'*Esempio 5* e per quello dell'*Esempio 6*, le risposte vengono date nell'ipotesi che gli oggetti presenti nell'urna all'atto di ogni estrazione siano estratti a caso – *ipotesi di simmetria in ogni singola estrazione* –.

- 1 Calcolare la probabilità che in una estrazione del lotto su una ruota esca la cinquina ordinata (n_1, \dots, n_5) .

In 4.1.1 *Esempio 5*, abbiamo risolto un problema collegato con l'estrazione del lotto su una ruota ponendoci in ipotesi di simmetria sull'esito dell'*estrazione completa* (equiprobabilità della cinque ordinate). Il teorema delle probabilità composte ci permette però di pervenire al medesimo risultato seguendo fedelmente nella descrizione la modalità di estrazione, ponendoci ora, come detto in premessa, in ipotesi di simmetria sugli esiti delle *singole estrazioni*. Posto $(X_i = n_i) = n_i$ è il numero estratto all'*i-esimo posto*, $\mathbb{P}_i = \{X_i = 1, \dots, X_i = 90\}$, $i = 1, \dots, 5$, si ha che \mathbb{P}_i descrive

l'esito dell'estrazione dell' i -esimo numero della cinquina. L'evento che a noi interessa è allora un evento elementare – il generico – della partizione prodotto $\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_5$, ovvero $(X_1 = n_1) \wedge \dots \wedge (X_5 = n_5)$, con $1 \leq n_i \leq 90$ e gli n_i a due a due diversi. Allora si ha: $P((X_1 = n_1) \wedge \dots \wedge (X_5 = n_5)) = P(X_1 = n_1) P(X_2 = n_2 | X_1 = n_1) \dots P(X_5 = n_5 | X_1 = n_1 \wedge \dots \wedge X_4 = n_4)$. Nelle nostre ipotesi da qui si ricava $P(X_1 = n_1) = 1/90$ e $P(X_{h+1} = n_{h+1} | X_1 = n_1 \wedge \dots \wedge X_h = n_h) = 1/(90-h)$, $h = 1, \dots, 4$, perché sono $90-h$ i numeri nell'urna quando viene estratto l' $(h+1)$ -esimo numero e n_{h+1} è ancora nell'urna – è l'ipotesi condizionante che informa che esso non è stato ancora estratto –. Si ottiene allora $P(X_1 = n_1 \wedge \dots \wedge X_5 = n_5) = 1/(90)_5$ qualunque sia la cinquina (n_1, \dots, n_5) . Gli eventi elementari $(X_1 = n_1) \wedge \dots \wedge (X_5 = n_5)$ sono dunque equiprobabili di somma 1 – perché sono in numero di $(90)_5$, tanti quante le cinquine ordinate che si possono formare con i primi 90 numeri naturali –, e si perviene perciò a una distribuzione di probabilità coerente (quella assunta per ipotesi in 4.1.1 *Esempio 5*).

Come si è fatto cenno all'inizio, nel gioco del lotto spesso non interessa l'ordine di estrazione. Basta allora riferirsi ad una descrizione che usa cinquine non ordinate al posto di quelle ordinate, meno fine perciò di quella introdotta. Se vogliamo però tenere valido l'assunto di estrazione a caso dei singoli numeri della cinquina, bisogna che la valutazione relativa alla partizione meno fine sia coerente con quella qui calcolata. Osserviamo a tal fine, che ogni cinquina ordinata "genera" mediante permutazione $5!$ cinquine equivalenti (con gli stessi numeri) che nel loro insieme danno luogo ad una cinquina non ordinata. Poiché la probabilità che esca la cinquina (n_1, \dots, n_5) è $1/(90)_5$, per la condizione di additività deve essere $5!/(90)_5 = \binom{90}{5}$ la probabilità che esca la cinquina $\{n_1, \dots, n_5\}$. Questi argomenti giustificano nell'ottica attuale l'ipotesi di simmetria fatta in relazione alla descrizione mediante cinquine non ordinate in 4.1.1 *Esempi 2, 3*. Si tratta in effetti di un trasferimento coerente alle cinquine non ordinate della condizione di simmetria supposta qui per i numeri presenti nell'urna all'atto di ogni estrazione.

Osserviamo ancora che, con la terminologia introdotta poc'anzi, possiamo dire che nel gioco del lotto l'estrazione della cinquina è fatta a caso, sia che si pensi in termini di cinquine ordinate che in quelli di cinquine non ordinate. A proposito di queste ultime, le precedenti considerazioni mettono anche in chiaro che – ai fini della valutazione – quando non interessa l'ordine è indifferente supporre che i numeri vengano estratti a caso uno per volta o pensare che l'estrazione avvenga – sempre a caso – a gruppi di 5. Questa conclusione ha carattere generale, come si vedrà in dettaglio nei modelli d'estrazione con due alternative (qui le alternative

sono 90, a priori in ogni estrazione) in 14.2.2 *Complemento* e con più alternative nel n° 14.3.2.

Il procedimento seguito qui sopra per calcolare la probabilità di $(X_1 = n_1) \wedge \dots \wedge (X_5 = n_5)$ mette in evidenza che il giudizio di simmetria sul risultato globale dell'estrazione si giustifica *se e solo se* si ritiene che *in ogni singola estrazione i numeri presenti nell'urna abbiano la stessa probabilità di essere estratti*. Il *se* lo provano i calcoli appena eseguiti. Il *solo se* lo provi il lettore calcolando le probabilità $P(X_1 = n)$, $P(X_2 = m | X_1 = n)$, eccetera, partendo dall'ipotesi che gli eventi $(X_1 = n_1) \wedge \dots \wedge (X_5 = n_5)$ siano equiprobabili – si può sfruttare il fatto che l'equiprobabilità si trasmette agli eventi elementari condizionati, atteso che le ipotesi $(X_1 = n_1)$, $(X_2 = n_1) \wedge (X_1 = n_2)$, ..., $(X_1 = n_1) \wedge \dots \wedge (X_4 = n_4)$ sono logicamente dipendenti dalla partizione che descrive l'estrazione in cinque ordinate (12.4.2 *Nota*) –.

È inoltre evidente che questo procedimento – basato su una puntuale analisi delle modalità di estrazione – è più appagante di quelli precedenti, e sarà d'ora in poi usato sistematicamente nella descrizione dei problemi collegati con modelli d'estrazione o ad essi riconducibili (già nei prossimi esempi, come detto in premessa).

- 2 *Con riferimento ad una estrazione del lotto sulle 10 ruote, si definisca il generico evento elementare che descrive il risultato e si calcoli la sua probabilità.*

Il risultato di un'estrazione completa è una composizione dei risultati relativi alle 10 ruote. Indicata con \mathbb{P}_h la partizione che descrive il risultato dell' h -esima ruota, $h = 1, \dots, 10$, in analogia con quanto visto nel precedente *Esempio 1*, gli eventi elementari di \mathbb{P}_h sono del tipo $\bigwedge_{i=1}^5 (X_{h,i} = n_{h,i})$, $n_{h,1}, \dots, n_{h,5}$ tra 1 e 90 a due a due distinti. L'esito dell'estrazione completa è allora descritto dalla partizione prodotto $\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_{10}$, il cui generico evento elementare è dato da $\bigwedge_{h=1}^{10} \bigwedge_{i=1}^5 (X_{h,i} = n_{h,i})$.

Per calcolare la sua probabilità useremo ancora il teorema delle probabilità composte, la cui applicazione comporta il calcolo del prodotto di 50 fattori. I primi 5 riguardano l'estrazione della prima ruota e – in analogia con l'esempio precedente – il loro prodotto è $1/(90)_5$. Fissiamo ora l'attenzione sui fattori relativi all'estrazione della k -esima ruota, $k > 1$. Posto $H_k = \bigwedge_{h=1}^{k-1} \bigwedge_{i=1}^5 (X_{h,i} = n_{h,i})$, il primo di essi è $P(X_{k,1} = n_{k,1} | H_k) = 1/90$, perché sono 90 le palline nell'urna k al momento della sua prima estrazione. Per gli altri si ha $P(X_{k,i} = n_{k,i} | H_k \wedge (X_{k,1} = n_{k,1} \wedge \dots \wedge X_{k,i-1} = n_{k,i-1})) = 1/(90-i+1)$, perché poi il calcolo si ripete come nel caso dell'*Esempio 1*;

il contributo di questi cinque fattori è quindi ancora $1/(90)_5$, e ciò accade per ogni ruota. Sicché riesce:

$$P(\bigwedge_{h=1}^{10} \bigwedge_{i=1}^5 (X_{h,i} = n_{h,i})) = \left(\frac{1}{(90)_5} \right)^{10}$$

- 3 *Descrivere il risultato del lancio di n dadi e calcolare la probabilità del generico evento elementare.*

Posto $(X_i = h)$ = il punto realizzato dal dado i -esimo è h , il generico evento elementare è $\bigwedge_{i=1}^n (X_i = h_i)$, $1 \leq h_i \leq 6$. Se assimiliamo i dadi a urne contenenti 6 palline (facce) numerate da 1 a 6 e supponiamo che il risultato realizzato da ciascun dado non venga influenzato dalla presenza degli altri dadi, si trova:

$$\begin{aligned} P(\bigwedge_{i=1}^n (X_i = h_i)) &= P(\bigwedge_{i=1}^{n-1} (X_i = h_i)) P(X_n = h_n | \bigwedge_{i=1}^{n-1} (X_i = h_i)) = \\ &= P(\bigwedge_{i=1}^{n-1} (X_i = h_i)) \frac{1}{6} = \dots = \left(\frac{1}{6} \right)^n. \end{aligned}$$

Resta così giustificata l'ipotesi di simmetria per i risultati del lancio di n dadi fatta in alcuni degli *Esempi* del n° 4.1.1 per $n=2$.

Si ottiene lo stesso risultato se, anziché lanciare una volta n dadi, si lancia un dado n volte, purché si supponga che in ogni lancio le facce si presentino a caso, indipendentemente perciò dai risultati realizzati nei lanci precedenti.

- 4 *Si eseguono tre estrazioni con rimessa da un'urna contenente 8 palline bianche e 2 nere. Posto « E_h = la pallina estratta nella h -esima estrazione è bianca», si determini la distribuzione di probabilità sulla partizione generata $\mathbb{P}_G(\{E_1, E_2, E_3\})$. Rispondere alla stessa domanda nell'ipotesi che le estrazioni avvengano senza rimessa.*

Sia al solito $E'_1 \wedge E'_2 \wedge E'_3$ il generico costituente di $\mathbb{P}_G(\{E_1, E_2, E_3\})$. Allora si ha $P(E'_1 \wedge E'_2 \wedge E'_3) = P(E'_1) P(E'_2 | E'_1) P(E'_3 | E'_1 \wedge E'_2)$. Questo vale formalmente in entrambe le modalità d'estrazione – solo formalmente però, perché gli eventi sono riferiti a stati d'informazione differenti e sono perciò diversi –. È facile verificare che si trovano i risultati riportati nella tabella più avanti, in cui gli $\omega_{h,i,k}$ sono gli eventi $E'_1 \wedge E'_2 \wedge E'_3$ scritti con altra notazione: i pedici indicano nell'ordine se E_1, E_2, E_3 sono affermati (pedice 1) o negati (pedice 0). Ad esempio, nel modello con rimessa riesce

$$P(E_1 \wedge \bar{E}_2 \wedge E_3) = P(\omega_{1,0,1}) = \frac{8}{10} \frac{2}{10} \frac{8}{10} = 0,128$$

e in quello senza rimessa

$$P(\bar{E}_1 \wedge \bar{E}_2 \wedge E_3) = P(\omega_{0,0,1}) = \frac{2}{10} \frac{1}{9} \frac{8}{8} = 0,022.$$

Si legge nella tabella che le probabilità degli eventi elementari con lo stesso numero di affermazioni sono uguali. Si tratta di un caso particolare di una importante proprietà – la scambiabilità (*Cap. 16, Vol. II*) – dei modelli d'estrazione con due alternative che sono studiati nel *Cap 14*.

	$\omega_{1,1,1}$	$\omega_{1,1,0}$	$\omega_{1,0,1}$	$\omega_{0,1,1}$	$\omega_{1,0,0}$	$\omega_{0,1,0}$	$\omega_{0,0,1}$	$\omega_{0,0,0}$
con rimessa	0,512	0,128	0,128	0,128	0,032	0,032	0,032	0,008
senza rimessa	0,467	0,156	0,156	0,156	0,022	0,022	0,022	0,000

È interessante osservare ancora in proposito che in entrambe le modalità d'estrazione riesce $P(E_1) = P(E_2) = P(E_3) = 8/10$ (anche questo rientra nella proprietà di scambiabilità). Cosa largamente intuibile nel modello con rimessa, perché prima di ogni estrazione viene ristabilita la composizione iniziale dell'urna. Lasciamo comunque al lettore il compito di una verifica completa. Ci limitiamo qui a mostrare a titolo d'esempio che l'affermazione è vera per $P(E_3)$ nel modello senza rimessa, in cui si ha:

$$\begin{aligned} P(E_3) &= P(\omega_{1,1,1}) + P(\omega_{1,0,1}) + P(\omega_{0,1,1}) + P(\omega_{0,0,1}) = \\ &= \frac{8}{10} \frac{7}{9} \frac{6}{8} + \frac{8}{10} \frac{2}{9} \frac{7}{8} + \frac{2}{10} \frac{8}{9} \frac{7}{8} + \frac{2}{10} \frac{1}{9} \frac{8}{8} = \frac{576}{720} = \frac{8}{10}. \end{aligned}$$

Le ipotesi di questo esempio sono quelle di 4.1.1 *Esempi 10, 11*. Cambiano però le domande. Ivi è richiesto di calcolare la probabilità di avere 0, 1, 2, 3 successi (estrazioni di pallina bianca) e il problema è stato risolto rendendo idealmente distinguibili le palline – numerandole da 1 a 10 – e ponendosi in ipotesi di simmetria sull'esito della terna di estrazioni – sulle terne ordinate dei primi 10 numeri, con ripetizione o senza in dipendenza dalla modalità d'estrazione –. Posto come allora ($S=n$) = *il numero di successi è n*, è facile verificare che si ritrovano qui gli stessi risultati. Nella modalità con rimessa si ha ad esempio:

$$\begin{aligned} P(S_n=2) &= P(\omega_{1,1,0}) + P(\omega_{1,0,1}) + P(\omega_{0,1,1}) = \\ &= \frac{8}{10} \frac{8}{10} \frac{2}{10} + \frac{8}{10} \frac{2}{10} \frac{8}{10} + \frac{2}{10} \frac{8}{10} \frac{8}{10} = \frac{576}{720} = \frac{384}{1000}. \end{aligned}$$

L'attuale ipotesi che in ogni estrazione le palline nell'urna abbiano la stessa probabilità di essere estratte è dunque equivalente a quella di simmetria ricordata poc'anzi.

- 5 Tre palline indistinguibili vengono collocate una alla volta in un cassetto scelto tra quattro disponibili. Descrivere il risultato della collocazione e calcolare le relative probabilità nelle due seguenti ipotesi sulla modalità di scelta del cassetto nei singoli passi:
- (i) il cassetto viene scelto a caso:
 - (ii) il cassetto viene scelto con probabilità proporzionale a $1+n$, essendo $n =$ numero di palline presenti nel cassetto (inizialmente a caso).

Indichiamo con $C_{h,i,j,k}$ = il primo cassetto contiene h palline, il secondo i , il terzo j ed il quarto k . Le collocazioni possibili sono allora quelle che si ottengono scegliendo gli indici verificanti le condizioni $0 \leq h, i, j, k \leq 3$, $h+i+j+k = 3$. Perciò 4 collocazioni prevedono che le palline vengano messe in un unico cassetto ($C_{3,0,0,0}$, $C_{0,3,0,0}$, $C_{0,0,3,0}$, $C_{0,0,0,3}$), 12 che due palline vengano messe in un cassetto e una in un altro (ad es. $C_{2,0,1,0}$) – perché si può scegliere in 4 modi il cassetto che contiene due palline e poi in 3 quello che ne contiene una – e infine ancora 4 collocazioni prevedono che le palline vengano messe una per cassetto ($C_{1,1,1,0}$, $C_{1,1,0,1}$, $C_{1,0,1,1}$, $C_{0,1,1,1}$). Per descrivere l'operazione di collocazione in dettaglio, immaginiamo le palline *distinguibili*, etichettandole ad esempio a, b, c . Possiamo allora introdurre gli eventi A_h (B_h, C_h) = la pallina a (b, c) è collocata nel cassetto h , $h = 1, 2, 3, 4$. In questa descrizione il generico evento elementare che descrive una collocazione con palline distinguibili è $A_h \wedge B_i \wedge C_j$, $1 \leq h, i, j \leq 4$. Si ha:

$$P(A_h \wedge B_i \wedge C_j) = P(A_h) P(B_i | A_h) P(C_j | A_h \wedge B_i).$$

Fin qui il discorso vale in entrambe le modalità d'estrazione. Andiamo ora a distinguere i due casi.

Modalità (i).

Poiché in ogni passo il cassetto viene scelto a caso indipendentemente dal suo contenuto, si ha $P(A_h \wedge B_i \wedge C_j) = (1/4)^3 = 1/4^3$, perché ora i cassettei a disposizione all'atto di ogni singola collocazione sono sempre 4 e quindi $P(A_h) = P(B_i | A_h) = P(C_j | A_h \wedge B_i) = 1/4$. La distribuzione di probabilità sulla partizione di eventi elementari $A_h \wedge B_i \wedge C_j$ è perciò uniforme, in accordo col fatto che è 4^3 il numero di tali eventi elementari – tante quante sono le terne ordinate con ripetizione dei numeri 1, 2, 3, 4 –. Per calcolare le probabilità degli eventi $C_{h,i,j,k}$ è sufficiente contare i loro casi favorevoli e fare il rapporto con i 4^3 casi possibili.

Riesce $C_{3,0,0,0} = A_1 \wedge B_1 \wedge C_1$, e quindi $P(C_{3,0,0,0}) = 1/4^3$, e lo stesso vale per le altre tre collocazioni che concentrano le tre palline in un cassetto.

Si ha poi $C_{2,1,0,0} = (A_1 \wedge B_1 \wedge C_2) \vee (A_1 \wedge B_2 \wedge C_1) \vee (A_2 \wedge B_1 \wedge C_1)$ e quindi

$P(C_{2,1,0,0}) = 3/4^3$, e lo stesso vale per le altre 11 collocazioni che prevedono due palline in un cassetto e una in un altro.

Infine, sono 6 i casi favorevoli di $C_{1,1,1,0}$ – le permutazioni di a, b, c –, quindi riesce $P(C_{1,1,1,0}) = 6/4^3$, e lo stesso vale per le altre tre collocazioni che prevedono una pallina per cassetto (uno vuoto).

Riesce naturalmente $4P(C_{3,0,0,0}) + 12P(C_{2,1,0,0}) + 6P(C_{1,1,1,0}) = 1$, com'è facile verificare.

Modalità (ii).

Ora viene scelto a caso solo il primo cassetto. Negli altri due passi la distribuzione cambia con la funzione $1+n$. Ragionando in termini di pesi sin dal primo passo, abbiamo che la somma dei pesi è 4 all'inizio, 5 nella seconda estrazione (in cui tre cassette hanno peso 1 e uno peso 2) e 6 nella terza estrazione (in cui tre cassette hanno peso 1 e uno peso 3 oppure due cassette hanno peso 1 e due peso 2). Ciò premesso, in ogni caso si ha:

$$P(A_h \wedge B_i \wedge C_j) = P(A_h)P(B_i|A_h)P(C_j|A_h \wedge B_i) = \frac{1}{4} \frac{r}{5} \frac{s}{6},$$

ove $r=2$ e $s=3$ se $h=i=j$; $r=2$ e $s=1$ se $h=i \neq j$; $r=1$ e $s=2$ se $h \neq i=j$ o $i \neq h=j$; $r=s=1$ se h, i, j sono a due a due diversi. Per la proprietà commutativa del prodotto sono perciò uguali le probabilità degli eventi elementari $A_h \wedge B_i \wedge C_j$ con lo stesso numero di indici uguali. Scegliendone uno per tipo si trova (h, i, j diversi a due a due):

$$P(A_h \wedge B_h \wedge C_h) = (1/4) \times (2/5) \times (3/6) = 1/20,$$

$$P(A_h \wedge B_h \wedge C_j) = (1/4) \times (1/5) \times (2/6) = 1/60,$$

$$P(A_h \wedge B_i \wedge C_j) = (1/4) \times (1/5) \times (1/6) = 1/120.$$

Allora si ha: $P(C_{3,0,0,0}) = P(A_1 \wedge B_1 \wedge C_1) = 1/20$, e lo stesso per le altre tre collocazioni che concentrano le tre palline in un cassetto; $P(C_{2,1,0,0}) = 3P(A_1 \wedge B_1 \wedge C_2) = 3 \times 1/60 = 1/20$ e lo stesso per le altre undici collocazioni che mettono due palline in un cassetto e una in un altro; $P(C_{1,1,1,0}) = 6P(A_1 \wedge B_1 \wedge C_1) = 6 \times 1/120 = 1/20$ e lo stesso per le altre tre collocazioni che mettono una pallina per cassetto.

In conclusione, se le palline sono indistinguibili e ciascuna di esse esercita una forza attrattiva di uguale intensità sulle altre nella misura sopra precisata – aumenta cioè la probabilità di scelta del cassetto in cui viene collocata incrementando di 1 la corrispondente funzione peso –, allora le 20 collocazioni possibili sono equiprobabili. Tutto avviene dunque come se sui risultati possibili $C_{h,i,j,k}$ della collocazione si fosse espresso un giudizio di simmetria.

I due problemi proposti in questo esempio saranno trattati in generale nel § 14.4 assieme a un terzo modello di collocazione, in cui è posta la condizione che i cassetti non possano contenere più di una pallina. I tre modelli sono usati in fisica per lo studio con metodi statistici del comportamento di particelle subatomiche.

- 6 *Calcolare la distribuzione di probabilità del tempo d'attesa del primo successo in una partita a testa e croce che dura indefinitamente, se in ogni lancio è p la probabilità che esca testa (successo), qualunque siano i risultati dei lanci precedenti.*

Posto $E_i = \text{esce testa al lancio } i\text{-esimo}$ e $\omega_n = \text{testa esce per la prima volta al lancio } n\text{-esimo}$, si ha $\omega_1 = E_1$ e $\omega_n = \bar{E}_1 \wedge \dots \wedge \bar{E}_{n-1} \wedge E_n$. Riesce perciò $P(\omega_1) = p$ e $P(\omega_n) = P(\bar{E}_1)P(\bar{E}_2|\bar{E}_1) \dots P(\bar{E}_{n-1}|\bar{E}_1 \wedge \dots \wedge \bar{E}_{n-2})P(E_n|\bar{E}_1 \wedge \dots \wedge \bar{E}_{n-1}) = p q^{n-1}$, con $q = 1 - p$, come anticipato nella Nota di 5.2.1 *Esempio 3*. Posto $\omega_\infty = \bar{E}_1 \wedge \dots \wedge \bar{E}_{n-1} \wedge \dots = \text{testa non esce mai}$, si deve porre $P(\omega_\infty) = 0$, perché $\sum_{n=1}^{\infty} p q^{n-1} = 1$.

13.2.5 Valutazioni sequenziali coerenti.

La procedura messa in atto per valutare le situazioni d'incertezza relative agli esempi del precedente n° 13.2.4, resa possibile dall'applicazione del teorema delle probabilità composte, è un esempio eccellente di approccio dinamico ai problemi dell'incertezza. Tutti i casi ivi considerati prevedono infatti che venga effettuata una sequenza di esperimenti – estrazioni di un numero per volta nel gioco del lotto (*Esempi 1, 2*), lancio di n dadi o un dado per n volte (*Esempio 3*), estrazioni di una pallina per volta (*Esempio 4*), collocazioni di un oggetto per volta in cassetti (*Esempio 5*) – e che i risultati vengano descritti prima singolarmente, ciascuno mediante una sua partizione, e poi in modo congiunto mediante la corrispondente partizione prodotto.

In parallelo, in tutti gli esempi è previsto che si proceda in modo sequenziale anche per la valutazione. In ogni passo la valutazione è limitata al risultato (aleatorio) di un esperimento. All'inizio (primo passo) la valutazione è data sulla prima partizione in termini assoluti e nei passi successivi sulle altre in termini

condizionati all'esito dei passi precedenti. Si è fatto ciò al fine di sfruttare al meglio le informazioni sulla modalità d'estrazione della sequenza – senza o con rimessa – e su quella di scelta nelle singole estrazioni – «a caso» o «pesata» (la seconda limitatamente al punto (ii) dell'*Esempio 5*) –. Usando il teorema delle probabilità composte generalizzato (n° 13.1.5) si perviene infine alla distribuzione di probabilità sulla partizione che descrive il risultato congiunto, prodotto delle partizioni precedenti.

La procedura sequenziale vista negli esempi si presta però a essere usata in condizioni del tutto generali. Indicata con \mathbb{P}_i la partizione che descrive il risultato (l'estrazione) i -esimo, la procedura prevede allora che al primo passo venga valutata la distribuzione di probabilità su \mathbb{P}_1 , al secondo le distribuzioni condizionate su $\mathbb{P}_2 | \omega$ per ogni $\omega \in \mathbb{P}_1$ (la probabilità su $\mathbb{P}_2 | \mathbb{P}_1^\phi$), ..., all' n -esimo la probabilità su $\mathbb{P}_n | (\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_{n-1})^\phi$, e che si valuti infine la probabilità su $\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_n$ usando il teorema delle probabilità composte. La correttezza (coerenza) della procedura è stata giustificata negli esempi a posteriori – verificando che è 1 la somma delle probabilità degli eventi elementari delle rispettive partizioni prodotto –. La coerenza della stessa in generale è assicurata invece dal prossimo teorema, che riguarda la valutazione sequenziale con riferimento alla descrizione mediante partizioni di cardinalità finita e per il resto qualsiasi – negli esempi citati e nei modelli d'estrazione in generale (Cap. 14) le partizioni hanno cardinalità uguale –.

TEOREMA.

Siano $\mathbb{P}_1, \dots, \mathbb{P}_n$ partizioni finite, $P_1(\cdot) = P_1(\cdot | \Omega), \dots, P_n(\cdot | \cdot)$ probabilità (condizionate) coerenti su

$$\mathbb{P}_1, \mathbb{P}_2 | \mathbb{P}_1^\phi, \dots, \mathbb{P}_n | (\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_{n-1})^\phi,$$

rispettivamente. Allora si ha:

a) l'estensione congiunta $P(\cdot | \cdot)$ di $P_1(\cdot | \Omega), \dots, P_n(\cdot | \cdot)$ (definita su $\mathbb{P}_1 \cup (\mathbb{P}_2 | \mathbb{P}_1^\phi) \cup \dots \cup (\mathbb{P}_n | (\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_{n-1})^\phi)$) è una proba-

bilità coerente;

b) l'estensione $P(\cdot|\cdot)$ di cui in a) si prolunga in modo unico su $\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_n$ ponendo

$$P(\omega_1 \wedge \dots \wedge \omega_n) = P_1(\omega_1) P_2(\omega_2 | \omega_1) \dots \\ \dots P_n(\omega_n | \omega_1 \wedge \dots \wedge \omega_{n-1}) \quad (40)$$

per ogni $\omega_1 \in \mathbb{P}_1, \dots, \omega_n \in \mathbb{P}_n$.

Commento. Resta così determinata la probabilità anche su $\mathcal{A}_L(\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_n)$ per additività (10.5.1 Proposizione), e su $\{E|H: E, H \in \mathcal{A}_L(\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_n), P(H) > 0\}$ ponendo $P(E|H) = P(E \wedge H) / P(H)$ (Teorema 13.2.2).

DIMOSTRAZIONE.

In seguito i simboli a_i, b_i indicano eventi elementari di \mathbb{P}_1 .

Proviamo intanto che le probabilità $P_j(\cdot|\Omega), \dots, P_n(\cdot|\cdot)$ sono compatibili, cioè che se un evento appartiene sia a $\mathbb{P}_h | \mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_{h-1}$ che a $\mathbb{P}_k | \mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_{k-1}$, allora ad esso viene assegnata in entrambi i passi h e k la stessa probabilità. Ciò accade perché, come mostreremo, se due di tali eventi sono uguali, allora essi sono o l'evento impossibile o quello certo e le probabilità a loro assegnate, coerenti per ipotesi, sono perciò o entrambe 0 o entrambe 1. Sia infatti $\omega | a_1 \wedge \dots \wedge a_{h-1} = e | b_1 \wedge \dots \wedge b_{k-1}$, $\omega \in \mathbb{P}_h, e \in \mathbb{P}_k, 1 \leq h < k \leq n$. Allora si ha che $a_1 \wedge \dots \wedge a_{h-1}$ e $b_1 \wedge \dots \wedge b_{k-1}$ sono eventi *non impossibili* (in quanto ipotesi) e *uguali* (12.3.3 e). Segue allora che $(a_1 \wedge \dots \wedge a_{h-1}) \wedge (b_1 \wedge \dots \wedge b_{k-1}) = (a_1 \wedge \dots \wedge a_{h-1}) \neq \phi$ e che i costituenti a_i, b_i di \mathbb{P}_1 sono *eventi elementari*, per ogni $i = 1, \dots, h-1$; anzi lo stesso evento elementare, perché altrimenti riesce $(a_1 \wedge \dots \wedge a_{h-1}) \wedge (b_1 \wedge \dots \wedge b_{k-1}) = \phi$ contro la precedente conclusione. Pertanto si ha

$$a_1 \wedge \dots \wedge a_{h-1} = b_1 \wedge \dots \wedge b_{h-1} \wedge b_h \wedge \dots \wedge b_{k-1} = a_1 \wedge \dots \wedge a_{h-1} \wedge b_h \wedge \dots \wedge b_{k-1},$$

da cui segue $a_1 \wedge \dots \wedge a_{h-1} \Rightarrow b_h \wedge \dots \wedge b_{k-1}$ e quindi anche $a_1 \wedge \dots \wedge a_{h-1} \Rightarrow b_h$. Allora riesce $b_h | a_1 \wedge \dots \wedge a_{h-1} = \Omega | a_1 \wedge \dots \wedge a_{h-1}$, perciò $b_i | a_1 \wedge \dots \wedge a_{h-1} = \phi | a_1 \wedge \dots \wedge a_{h-1}$ per ogni altro $b_i \in \mathbb{P}_h$, e quindi $\omega | a_1 \wedge \dots \wedge a_{h-1} = e | b_1 \wedge \dots \wedge b_{k-1}$ è certo se $\omega = b_h$, impossibile altrimenti.

Prova della a) (per induzione).

Poniamo $\mathcal{D}_1 = \mathbb{P}_1, \mathcal{D}_m = \mathbb{P}_1 \cup \mathbb{P}_2 | (\mathbb{P}_1)^\phi \cup \dots \cup \mathbb{P}_m | (\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_{m-1})^\phi, m = 2, \dots, n$, e indichiamo con $P(\cdot|\cdot)$ l'estensione congiunta di $P_j(\cdot|\Omega), \dots, P_m(\cdot|\cdot)$

– come fatto nell'enunciato –, Allora, per ipotesi si ha intanto che $P_l(\cdot)$ è coerente su $\mathbb{P}_l = \mathcal{D}_l$ (base dell'induzione).

Andiamo poi a provare che se $P(\cdot|\cdot)$ è coerente su \mathcal{D}_m , allora lo è anche su \mathcal{D}_{m+1} , $m = 1, \dots, n-1$ (passo induttivo). Cominciamo con l'osservare a tal fine che una scommessa su eventi di $\mathbb{P}_h | (\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_{h-1})^\phi$, $h \leq m$, riguarda eventi di \mathbb{P}_h e di $(\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_{h-1})^\phi$, e il suo guadagno è perciò definito su $\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_h$, e quindi anche su $\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_m$. Su questo insieme è allora definito anche il guadagno di una scommessa su \mathcal{D}_m .

Sia ora G il guadagno di una scommessa su eventi (condizionati) di \mathcal{D}_{m+1} e H_0 la somma delle loro ipotesi condizionanti. Il guadagno $G|H_0$ verifica la condizione di coerenza per ipotesi se la scommessa riguarda o solo eventi di \mathcal{D}_m o solo eventi di $\mathbb{P}_{m+1} | (\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_m)^\phi$. Altrimenti possiamo scrivere $G = G_1 + G_2$, con G_1 scommessa su eventi di \mathcal{D}_m e G_2 scommessa su eventi di $\mathbb{P}_{m+1} | (\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_m)^\phi$. Si tratta di dimostrare che G verifica la condizione di coerenza anche in questo caso, e basta provare per questo che riesce $\min G|H_0 \leq 0$. Indichiamo a tal fine con K_0 la somma delle ipotesi degli eventi di \mathcal{D}_m oggetto di scommessa. Si ha allora $\min G_1|K_0 \leq 0$ (per l'ipotesi induttiva) e scelto un $\omega = a_1 \wedge \dots \wedge a_m \in \mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_m$ in cui $G_1|K_0$ assume valore minimo riesce $\min G_1|K_0 = \min G_1|\omega = G_1|\omega$ e $\omega \Rightarrow K_0 \Rightarrow H_0$. Segue perciò:

$$\min G|H_0 \leq \min G|K_0 \leq \min G|\omega = \min (G_1 + G_2)|\omega = G_1|\omega + \min G_2|\omega.$$

Ora, il guadagno G_2 riguarda eventi di $\mathbb{P}_{m+1} | (\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_m)^\phi$ ed è quindi del tipo $G_2 = \sum_i S_i (|e_i| - P_{m+1}(e_i|h_i))/h_i$ – somma estesa a un numero finito di termini –, con $e_i \in \mathbb{P}_{m+1}$ e $h_i \in (\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_m)^\phi$. Pertanto il generico termine della sommatoria calcolato in ω è 0 se $h_i \neq \omega$ e $S_i (|e_i| - P_{m+1}(e_i|\omega))/\omega$ altrimenti. Sicché, $G_2(\omega) = 0$ se tutti gli h_i sono diversi da ω , e in tal caso riesce $\min G|H_0 = G_1|\omega \leq 0$. Altrimenti $G_2(\omega)$ è una somma di addendi del tipo $S_i (|e_i| - P_{m+1}(e_i|\omega))/\omega$ ed è quindi il guadagno di una scommessa su eventi di $\mathbb{P}_{m+1} | (\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_m)^\phi$, tutti condizionati alla medesima ipotesi ω , e conseguentemente riesce $\min G_2|\omega \leq 0$, perché $P_{m+1}(\cdot|\omega)$ è coerente su $\mathbb{P}_{m+1} | (\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_m)^\phi$ per ipotesi. In conclusione si ha $\min G|H_0 \leq 0$ in ogni caso e quindi l'applicazione $P(\cdot|\cdot)$ è una probabilità coerente su \mathcal{D}_{m+1} . Ciò prova il passo induttivo e con esso la a).

Prova della b).

L'applicazione definita dalla (40) è la scrittura con i simboli attuali della condizione necessaria (37) del teorema delle probabilità composte generalizzato (n° 13.1.5), ed è perciò l'unico prolungamento coerente su $\mathbb{P}_1 \wedge \dots \wedge \mathbb{P}_n$ dell'estensione congiunta di $P_1(\cdot|\Omega), \dots, P_n(\cdot|\cdot)$ – di $P(\cdot|\cdot)$ –, che è coerente su \mathcal{D}_n come appena dimostrato. ■

13.3 Chiusura delle probabilità condizionate per passaggi al limite

Abbiamo visto che l'insieme delle probabilità assolute coerenti, definite su un *medesimo* insieme di eventi, è convesso e chiuso per passaggi al limite (*Teoremi* 9.3.1, 9.3.2). Quest'ultima proprietà si conserva anche per le probabilità condizionate (*Teorema* 13.3.2 qui avanti) ma non quella di convessità, a meno di ipotesi particolari sull'insieme degli eventi (del dominio) su cui è definita la probabilità. Ad esempio, se gli eventi sono condizionati a una medesima ipotesi o se il dominio soddisfa le condizioni previste dal prossimo *Lemma* 13.3.1.

Se $\mathcal{D} = \mathcal{E}|H$, \mathcal{E} insieme di eventi assoluti e $H \neq \emptyset$, la situazione è concettualmente uguale a quella che si ha in ambiente assoluto e la prova della convessità dell'insieme delle probabilità $P(\cdot|H)$ di dominio \mathcal{D} ha bisogno solo di qualche aggiustamento formale rispetto a quella fatta nel *Teorema* 9.3.1, dovuto al fatto che il guadagno G si scompone qui nella differenza di due numeri aleatori anziché in quella della vincita aleatoria e della puntata certa. Ci si riconduce però a questa situazione con riferimento al guadagno $G|H$, che è quello che si deve analizzare ai fini della coerenza. In dettaglio, si ha cioè $G = V - c_P |H|$, ove $V = \sum_{i=1}^n S_i |E_i| |H|$ e $c_P = S_i P(E_i|H)$, e quindi $G|H = V|H - c_P |H| |H| = \sum_{i=1}^n S_i |E_i| |H| - c_P$, in cui solo c_P dipende dalla valutazione $P(\cdot|H)$. Segue allora che al variare di $P(\cdot|H)$ i guadagni condizionati si comportano come quelli assoluti in 9.2.4 *Lemma* – qui raggiungono valore massimo (minimo) in corrispondenza a uno stesso evento elementare di $\mathbb{P}_G(\{E_1, \dots, E_n\})|H$ – e per provare la convessità dell'insieme delle $P(\cdot|H)$ di dominio \mathcal{D} si può ripetere la dimostrazione del *Teorema* 9.3.1.

Come anticipato poc'anzi, un altro caso di convessità dell'insieme delle valutazioni condizionate è quello previsto dal prossimo lemma, che viene messo qui in evidenza perché sarà utilizzato in seguito.

13.3.1 Lemma.

Siano P_1 e P_2 applicazioni di dominio \mathcal{E} , insieme di eventi assoluti, $P(\cdot|\cdot)$ un'applicazione di dominio \mathcal{D} , insieme di eventi condizionati, \mathcal{E} e \mathcal{D} disgiunti. Se le estensioni congiunte di P_1 , $P(\cdot|\cdot)$ e di P_2 , $P(\cdot|\cdot)$ su $\mathcal{E} \cup \mathcal{D}$ sono probabilità coerenti, allora è coerente sul medesimo insieme ogni loro mistura.

DIMOSTRAZIONE.

Osserviamo preliminarmente che se P è una probabilità su \mathcal{E} , il guadagno di

una combinazione di scommesse su eventi di $\mathcal{E} \cup \mathcal{D}$ valutato con l'estensione congiunta di $P, P(\cdot|\cdot)$, si può scrivere $G = V - c + G'$, con $V - c$ guadagno delle scommesse su eventi (assoluti) di \mathcal{E} – differenza tra vincita *Valeatoria* e costo *c certo* – e G' guadagno *aleatorio* delle scommesse su eventi (condizionati) di \mathcal{D} . Sicché, posto $X = V + G'$, si ha $G = X - c$. Pertanto G e X differiscono per un numero certo e hanno perciò in comune i "punti" di minimo e di massimo – eventi elementari della partizione generata da tutti gli eventi che intervengono nella scommessa, o come eventi o come ipotesi – i quali dipendono dalla probabilità $P(\cdot|\cdot)$, ma non dalla probabilità P . Se si tiene fissa $P(\cdot|\cdot)$, come previsto per le tre estensioni congiunte dell'enunciato, la situazione è allora analoga a quella sfruttata per dimostrare la convessità dell'insieme delle probabilità assolute di un comune dominio (*Teorema 9.3.1*), che può essere perciò qui riproposta con gli opportuni adattamenti.

Ciò premesso, siano coerenti le estensioni congiunte di $P_1, P(\cdot|\cdot)$ e $P_2, P(\cdot|\cdot)$ e consideriamo una loro mistura con coefficienti $\lambda, \bar{\lambda} = 1 - \lambda$ che, tenuto conto che $\lambda P(\cdot|\cdot) + \bar{\lambda} P(\cdot|\cdot) = P(\cdot|\cdot)$ e ponendo $P = \lambda P_1 + \bar{\lambda} P_2$, dà luogo all'estensione congiunta $P, P(\cdot|\cdot)$. Indichiamo con G il guadagno di una scommessa \heartsuit su eventi di $\mathcal{E} \cup \mathcal{D}$ valutata con tale estensione. Se essa riguarda solo eventi di \mathcal{D} il corrispondente guadagno verifica le condizioni di coerenza perché è valutato con la $P(\cdot|\cdot)$, coerente per ipotesi. Altrimenti riesce $H_0 = \Omega$, e bisogna allora far vedere che valgono le condizioni di coerenza per il guadagno G incondizionatamente. Indicati con E_1, \dots, E_n gli eventi di \mathcal{E} presenti nella scommessa \heartsuit e con S_1, \dots, S_n i corrispondenti importi, riesce

$$c = \sum_{i=1}^n S_i P(E_i) = \lambda \sum_{i=1}^n S_i P_1(E_i) + \bar{\lambda} \sum_{i=1}^n S_i P_2(E_i) = \lambda c_1 + \bar{\lambda} c_2,$$

con ovvio significato di c_1 e c_2 . Segue allora

$$G = X - c = (\lambda + \bar{\lambda})X - (\lambda c_1 + \bar{\lambda} c_2) = \lambda(X - c_1) + \bar{\lambda}(X - c_2)$$

ove $G_1 = X - c_1$ ($G_2 = X - c_2$) è il guadagno di una scommessa sugli stessi eventi e le stesse vincite di \heartsuit , valutata però con l'estensione congiunta di $P_1, P(\cdot|\cdot)$ ($P_2, P(\cdot|\cdot)$). Se ω' è un "punto" di minimo per X (e quindi per G, G_1, G_2), allora si ricava

$$\min G = G(\omega') = \lambda G_1(\omega') + \bar{\lambda} G_2(\omega') = \lambda \min G_1 + \bar{\lambda} \min G_2 \leq 0,$$

la disuguaglianza valendo per la coerenza delle estensioni congiunte $P_1, P(\cdot|\cdot)$ e $P_2, P(\cdot|\cdot)$ su $\mathcal{E} \cup \mathcal{D}$. Per la 9.2.4 *Proposizione* segue la coerenza dell'estensione congiunta $\lambda P_1 + \bar{\lambda} P_2, P(\cdot|\cdot)$, come volevasi. ■

CONTROESEMPIO.

Abbiamo detto in premessa che l'insieme delle probabilità condizionate definite

su uno stesso dominio non è in generale convesso. Lo prova il seguente controesempio. Consideriamo la partizione in eventi elementari $\{A, B, C\}$ e su di essa le probabilità coerenti P_1 e P_2 così definite:

$$P_1(A) = 0,3, \quad P_1(B) = 0,5, \quad P_1(C) = 0,2,$$

$$P_2(A) = 0,1, \quad P_2(B) = 0,3, \quad P_2(C) = 0,6,$$

È allora coerente su $\{A, B, C\}$ la combinazione lineare convessa $P = 0,4P_1 + 0,6P_2$, per la quale si ha: $P(A) = 0,18, P(B) = 0,38, P(C) = 0,44$.

Consideriamo ora l'insieme degli eventi $\{A, B, A|A \vee B\}$ e manteniamo nei tre casi le valutazioni date per A e B . In forza del teorema delle probabilità composte, si deve allora porre $P_1(A|A \vee B) = 3/8, P_2(A|A \vee B) = 1/4, P(A|A \vee B) = 9/28$. Quelle ora calcolate sono le uniche valutazioni coerenti per $A|A \vee B$ in tutti i tre casi, perché imposte da condizioni necessarie per la coerenza – additività e teorema delle probabilità composte in forma di rapporto –. Come miscela di $P_1(A|A \vee B)$ e $P_2(A|A \vee B)$ si ottiene invece:

$$0,4P_1(A|A \vee B) + 0,6P_2(A|A \vee B) = 0,3 \neq P(A|A \vee B).$$

La miscela $0,4P_1(\cdot|\cdot) + 0,6P_2(\cdot|\cdot)$ delle probabilità coerenti $P_1(\cdot|\cdot)$ e $P_2(\cdot|\cdot)$ su $\{A|\Omega, B|\Omega, A|A \vee B\}$ non è dunque una probabilità coerente. ■

13.3.2 Teorema. Chiusura delle probabilità condizionate per passaggio al limite.

Sia $(P_x(\cdot|\cdot))$ una famiglia di probabilità condizionate coerenti su \mathcal{D} ed esista il $\lim_{x \rightarrow x_0} P_x(E|H)$ per ogni $E|H \in \mathcal{D}$. Allora anche $P(\cdot|\cdot) = \lim_{x \rightarrow x_0} P_x(\cdot|\cdot)$ è una probabilità condizionata coerente su \mathcal{D} .

DIMOSTRAZIONE.

Sia $G_x = \sum_{i=1}^n S_i(|E_i| - P_x(E_i|H_i))/|H_i|$ il guadagno della generica scommessa su $E_1|H_1, \dots, E_n|H_n \in \mathcal{D}$ valutata con $P_x(\cdot|\cdot)$. Come nel caso assoluto si ricava facilmente $G = \sum_{i=1}^n S_i(|E_i| - P(E_i|H_i))/|H_i| = \lim_{x \rightarrow x_0} G_x$, e quindi anche $G|H_0 = \lim_{x \rightarrow x_0} G_x|H_0$ ($H_0 = H_1 \vee \dots \vee H_n$).

Ciò premesso, indichiamo con $g(\cdot|H_0)$ e $g_x(\cdot|H_0)$ le applicazioni di dominio $\mathbb{P}_G|H_0 = \mathbb{P}_G(\{E_1, \dots, E_n, H_1, \dots, H_n\})|H_0$ che definiscono i guadagni G e G_x e supponiamo per assurdo che sia $\min G|H_0 = m > 0$. Riesce allora $g(\omega|H_0) > m$ per ogni evento elementare $\omega|H_0 \in \mathbb{P}_G|H_0$. Esiste allora un intorno I_{x_0} tale che, per ogni $x \in I_{x_0}$ e ogni $\omega|H_0 \in \mathbb{P}_G|H_0$, si ha $g_x(\omega|H_0) \geq m/2 > 0$, contro l'ipotesi, perché le probabilità $P_x(\cdot|H_0)$ sono coerenti. ■

NOTA.

Non è possibile ripetere per questo teorema la dimostrazione fatta nel caso delle probabilità assolute, perché i guadagni condizionati ad H_0 che si devono qui esaminare sono somma di due numeri aleatori, il secondo dei quali, $\sum_{i=1}^n P_x(E_i|H_i)/|H_i|$, dipende dalla valutazione. Gli eventi elementari in cui vengono raggiunti i valori minimo e massimo variano allora con la valutazione, e ciò rende inapplicabile la tecnica usata nel caso assoluto.

13.3.3 Metodo asintotico in ambiente condizionato.

Cominciamo col dare una generalizzazione del metodo asintotico nel caso numerabile. Sia allora $\mathbb{P} = \{\omega_1, \dots, \omega_i, \dots\}$ una partizione di cardinalità numerabile in eventi elementari e P_n probabilità concentrata uniformemente sui primi suoi n eventi elementari. Sicché, posto $A_n = \omega_1 \vee \dots \vee \omega_n$, per ogni $E \in \mathcal{A}_L(\mathbb{P})$ si ha $P_n(E) = P_n(E \wedge A_n) = \frac{1}{n} \sum_{i=1}^n \delta(\omega_i \Rightarrow E) = \frac{1}{n} \text{card}(E \wedge A_n)$, ove $\text{card}(E \wedge A_n)$ indica il numero dei componenti – eventi elementari di \mathbb{P} – dell'evento $E \wedge A_n$. Sia poi $P_n(\cdot | \cdot)$ un prolungamento di P_n su $\mathcal{A}_L(\mathbb{P}) | \mathcal{A}_L^\phi(\mathbb{P})$. Preso $H \in \mathcal{A}_L^\phi(\mathbb{P})$, da un certo n in poi riesce $P_n(H) > 0$. Perciò, per ogni $H \in \mathcal{A}_L^\phi(\mathbb{P})$ la $P_n(\cdot | H)$ risulta definitivamente data dal rapporto $P_n(\cdot \wedge H) / P_n(H)$. Se n è abbastanza grande – dipendente da H – riesce cioè:

$$P_n(E|H) = \frac{P_n(E \wedge H)}{P_n(H)} = \frac{P_n(E \wedge H \wedge A_n)}{P_n(H \wedge A_n)} = \frac{\text{card}(E \wedge H \wedge A_n)}{\text{card}(H \wedge A_n)}.$$

Per ogni $E, H \neq \phi$ logicamente dipendenti da \mathbb{P} ha allora senso studiare il limite al divergere di n della successione $(P_n(E|H))$. Posto

$$\mathcal{C} = \{E|H \in \mathcal{A}_L(\mathbb{P}) | \mathcal{A}_L^\phi(\mathbb{P}) : \text{la successione } (P_n(E|H)) \text{ converge}\}$$

e $P(E|H) = \lim_{n \rightarrow \infty} P_n(E|H)$ per ogni $E|H \in \mathcal{C}$, l'applicazione $P(\cdot | \cdot)$ di dominio \mathcal{C} è una probabilità coerente in virtù del Teorema 13.3.2.

Se la partizione \mathbb{P} ha cardinalità superiore al numerabile, analogamente al caso assoluto la procedura potrà essere applicata con riferimento a sottoinsiemi numerabili di eventi elementari $\{\omega_1, \dots, \omega_i, \dots\} \subset \mathbb{P}$. In tal caso riesce però $P_n(H) = 0$ per ogni n se nessun ω_i implica H , e bisogna perciò limitare la scelta delle ipotesi H agli eventi di $\mathcal{A}_L^\phi(\mathbb{P})$ che hanno almeno una componente in $\{\omega_1, \dots, \omega_i, \dots\}$. Per ciascuna di tali ipotesi si ha in particolare $P(E|H) = 0$ per tutti gli E privi di componenti in $\{\omega_1, \dots, \omega_i, \dots\}$, perché allora anche $E \wedge H$ è privo di componenti in $\{\omega_1, \dots, \omega_i, \dots\}$ e riesce $P_n(E|H) = P_n(E \wedge H) / P_n(H) = 0$ da un certo n in poi – per $n \geq m$ con m tale che $\omega_m \Rightarrow H$ –.

COMPLEMENTO. *Perfezionamento della nozione di scelta a caso di un naturale.*

Il risultato trovato sopra permette di precisare ampliandolo il significato di scelta a caso di un numero naturale (10.5.2 *Esempio*). Posto $\mathbb{P} = \{N=i : i \in \mathbb{N}\}$, cominciamo con l'osservare per questo che riesce $P_n(\cdot | \Omega) = P_n(\cdot)$, da cui segue che in questo modo si introduce la distribuzione asintotica sull'insieme di Dynkin (di eventi assoluti) $\mathcal{D} \subset \mathcal{C}$. Inoltre, per ogni $H \neq \emptyset$ finito, se $n \geq \max \{i : \omega_i \Rightarrow H\}$ riesce $H \wedge A_n = H$, e per ogni $E \in \mathcal{A}_L(\mathbb{P})$ si ricava allora $P_n(E|H) = \text{card}(E \wedge H) / \text{card}(H) = P(E|H)$. In particolare, se $E = \omega_i \Rightarrow H$ riesce $P(\omega_i|H) = 1 / \text{card}(H)$. Sicché, il metodo asintotico ampliato agli eventi condizionati di $\mathcal{A}_L(\mathbb{P}) | \mathcal{A}_L^\emptyset(\mathbb{P})$ introduce la distribuzione asintotica di base \mathbb{P} (10.3.2 *Definizione*) – sul corrispondente insieme di Dynkin – e la distribuzione uniforme (l'equiprobabilità) sulle partizioni $\mathbb{P}|H \subset \mathcal{C}$ per ogni $H \neq \emptyset$ logicamente dipendente da \mathbb{P} finito. È allora coerente intendere che **scegliere a caso nel numerabile** significhi *introdurre la distribuzione asintotica su \mathcal{D} e l'equiprobabilità su $\mathbb{P}|H$ per ogni $H \in \mathcal{A}_L^\emptyset(\mathbb{P})$ finito.*

13.4 Valutazione per disintegrazione

La nozione di probabilità condizionata è stata introdotta ponendo l'accento sull'esigenza di aggiornare con coerenza una valutazione preesistente, a seguito di incremento d'informazione. Vista la questione in questa ottica, si può trarre l'impressione che le probabilità condizionate abbiano interesse solo quando si viene in possesso di effettiva nuova informazione. Anche se è vero che questo è il primo motivo di interesse per il concetto – e l'approccio più naturale per trovare suggerimenti per la sua definizione – è pure vero che si può parlare di incremento d'informazione anche in via ipotetica, come fatto del resto quasi sempre già in precedenza. Lo si può fare ad esempio per decidere se, a certi fini, vale la pena di procurarsi una certa informazione (tenere presente che di solito le informazioni costano!). Ma lo si può fare – e ciò è sicuramente più importante – anche per ottenere attraverso valutazioni condizionate, una valutazione di probabilità assoluta. Anzi, l'uso delle probabilità condizionate a questi fini è molto frequente. Lo si spiega riflet-

tendo sul fatto che le valutazioni di probabilità condizionate spesso risultano più agevoli, perché vengono effettuate in stati di informazione più specifici di quello iniziale, che conferiscono perciò contorni più precisi al quadro delle possibilità. Abbiamo del resto già utilizzato valutazioni condizionate per ottenere probabilità assolute nel n° 13.2.5 precedente trattando della valutazione sequenziale in ambiente finito. A ben guardare, già la formulazione del teorema delle probabilità composte è una regola che insegna a ricavare la probabilità di un evento assoluto – un prodotto logico – per mezzo delle probabilità di eventi condizionati. Le prossime considerazioni riguardano un altro modo di valutare – anche questo molto importante – che si pone su questa medesima linea. Esso insegna a costruire probabilità assolute mediante un'adeguata "manipolazione" di valutazioni, sia assolute sia condizionate.

13.4.1 Approccio al problema.

Prendiamo le mosse considerando un evento E e una partizione $\{H_1, \dots, H_n\}$ in eventi elementari. Poiché $E = (E \wedge H_1) \vee \dots \vee (E \wedge H_n)$, se operiamo con probabilità coerenti *deve* essere:

$$P(E) = P(E \wedge H_1) + \dots + P(E \wedge H_n) = P(H_1)P(E|H_1) + \dots + P(H_n)P(E|H_n). \quad (41)$$

Circa la probabilità, abbiamo detto che essa deve essere coerente. Questa affermazione è vaga, perché non precisa su quale dominio è definita la probabilità. È implicitamente chiaro, però, che essa deve essere comunque definita per tutti gli eventi che vengono presi in considerazione nelle due uguaglianze precedenti. Pertanto, il suo dominio deve essere costituito almeno dall'insieme di eventi

$$\{E, H_1, \dots, H_n, E \wedge H_1, \dots, E \wedge H_n, E|H_1, \dots, E|H_n\}.$$

Si pone allora in modo naturale il problema di stabilire quali sono le probabilità coerenti su questo insieme. Studieremo una generalizzazione di questo problema, legata al fatto che abitualmente interessa considerare insiemi di eventi al posto di un unico evento e partizioni di Ω qualunque – e non solo finite –.

Mettendoci in questo ordine di idee, consideriamo un insieme \mathcal{E} di eventi assoluti e una partizione \mathbb{P} di ipotesi (in eventi elementari assoluti). Interessa allora studiare le probabilità coerenti sull'insieme di eventi $\mathcal{E} \cup \mathbb{P} \cup (\mathcal{E} \wedge \mathbb{P}) \cup$

$\mathcal{E}|\mathbb{P}$ – in termini condizionati, perché sono tali gli eventi dell'ultimo insieme –. Mostriamo in proposito che se \mathbb{P} è finita, assegnate a piacere una probabilità coerente $P(\cdot)$ su \mathbb{P} e probabilità condizionate coerenti $P(\cdot|H)$ su $\mathcal{E}|H$ per ogni $H \in \mathbb{P}$, resta univocamente determinata la probabilità (coerente) su $\mathcal{E} \cup (\mathcal{E} \wedge \mathbb{P})$, è data da $P(E \wedge H) = P(H)P(E|H)$ su $\mathcal{E} \wedge \mathbb{P}$ e dalla (41) su \mathcal{E} , ed è σ -additiva se sono tali le $P(\cdot|H)$. Dimosteremo anzi un risultato più generale, che vale cioè per partizioni qualsiasi, purché la probabilità $P(\cdot)$ assegnata a \mathbb{P} sia concentrata (Teorema 13.4.3). Premessa indispensabile per questo è il seguente teorema.

13.4.2 Teorema.

Siano \mathcal{E} un insieme di eventi (assoluti), \mathcal{H} un insieme di ipotesi a due a due incompatibili, $P(\cdot)$ una probabilità su $\mathcal{F} \subset \mathcal{A}_L(\mathcal{H})$ non vuoto e $P(\cdot|H)$ una probabilità condizionata su $\mathcal{E}|H$ per ogni $H \in \mathcal{H}$. È allora coerente su $\mathcal{F} \cup \mathcal{E}|\mathcal{H}$ l'estensione congiunta $P(\cdot|\cdot)$ della $P(\cdot)$ e delle $P(\cdot|H)$, $H \in \mathcal{H}$.

DIMOSTRAZIONE.

Cominciamo col provare che è coerente la restrizione della $P(\cdot|\cdot)$ su $\mathcal{E}|\mathcal{H}$. Considerata una generica combinazione di scommesse su eventi di $\mathcal{E}|\mathcal{H}$, indichiamo con $\{E_1, \dots, E_n\} \subset \mathcal{E}$ l'insieme dei componenti (eventi assoluti) usati nel ruolo di *eventi* negli eventi (condizionati) che fanno parte della combinazione, e con $\{H_1, \dots, H_s\} \subset \mathcal{H}$ quello delle relative *ipotesi*. Più eventi di $\{E_1, \dots, E_n\}$ possono essere ovviamente condizionati a una medesima ipotesi – e vale anche il viceversa (un E_i può essere condizionato a più ipotesi), ma qui interessa tenere conto del primo aspetto –. Indichiamo allora con I_j l'insieme degli indici degli eventi di $\{E_1, \dots, E_n\}$ che sono condizionati all'ipotesi H_j . Il guadagno della scommessa data allora si scrive:

$$G' = \sum_{j=1}^s \sum_{i \in I_j} S_i^{(j)} (|E_i| - P(E_i|H_j)) |H_j| = \sum_{j=1}^s G_j,$$

ove $G_j = \sum_{i \in I_j} S_i^{(j)} (|E_i| - P(E_i|H_j)) |H_j|$ è la parte del guadagno G' riguardante le scommesse sugli eventi condizionati alla medesima ipotesi H_j . Ora, poiché le ipotesi H_1, \dots, H_s sono incompatibili a due a due, riesce $G_j|H_l = 0$ se $j=2, \dots, n$ – perché allora $|H_j||H_l| = 0$ – e $G_l|H_l = \sum_{i \in I_l} S_i^{(l)} (|E_i||H_l| - P(E_i|H_l))$ – perché $|H_l||H_l| = 1$ –. Pertanto, posto al solito $H_0 = H_1 \vee \dots \vee H_s$, si ottiene $G'|H_l = \sum_{j=1}^s G_j|H_l = G_l|H_l$ e quindi

$$\min G'|H_0 \leq \min G'|H_l = \min G_l|H_l \leq 0,$$

l'ultima disuguaglianza valendo perché $P(\cdot|H_j)$ è coerente per ipotesi su $\mathcal{E}|H_j$. Segue allora $\min G'|H_0 \leq 0$, e quindi che $P(\cdot|\cdot)$ è coerente su $\mathcal{E}|\mathcal{H}$.

Andiamo ora a provare che $P(\cdot|\cdot)$ è coerente anche su $\mathcal{F} \cup \mathcal{E}|\mathcal{H}$. A tale scopo è sufficiente studiare il guadagno G di una scommessa in cui compaiono sia eventi di \mathcal{F} che eventi di $\mathcal{E}|\mathcal{H}$ – la restrizione della $P(\cdot|\cdot)$ a \mathcal{F} è coerente per ipotesi e quella a $\mathcal{E}|\mathcal{H}$ per quanto appena dimostrato –. Fermo restando il significato di G' – la parte di scommessa su eventi di $\mathcal{E}|\mathcal{H}$ –, riesce $G = G'' + G'$, con $G'' = \sum_{h=1}^r S_h(|A_h| - P(A_h))$, essendo A_1, \dots, A_r gli eventi di \mathcal{F} oggetto di scommessa. Inoltre per tale scommessa riesce $H_0 = \Omega$ e quindi si tratta di esaminare il minimo (o massimo) di G incondizionatamente. Per la coerenza di P su \mathcal{F} , si ha intanto $\min G'' \leq 0$. Indicato con K un evento elementare della partizione generata da $\{A_1, \dots, A_r\}$ in cui G'' assume valore minimo, si ha: $G|K = G''|K + G'|K = \min G'' + G'|K$, da cui segue $\min G \leq \min G|K = \min G'' + \min G'|K$. Essendo $\min G'' \leq 0$, per dimostrare la tesi basta allora provare che $\min G'|K \leq 0$. Osserviamo per questo che K , come ogni altro evento elementare della partizione generata da $\{A_1, \dots, A_r\}$, è un evento logicamente dipendente dalla partizione generata da \mathcal{H} – i cui eventi elementari sono gli $H \in \mathcal{H}$ e $\neg \bigvee \mathcal{H}$ –, perché $\{A_1, \dots, A_r\} \subset \mathcal{F} \subset \mathcal{A}_1(\mathcal{H})$. Pertanto, per ogni $H \in \mathcal{H}$ si ha o $H \Rightarrow \bar{K}$ cioè $H \wedge K = \emptyset$ e quindi $|H||K = 0$, oppure $H \Rightarrow K$. Ciò premesso si possono presentare due casi. O tutte le ipotesi che compaiono nella scommessa di guadagno G' implicano \bar{K} , e allora $G'|K = 0$, quindi $\min G'|K \leq 0$, e da qui segue la tesi. O esiste un'ipotesi H_j che implica K , e allora $\min G'|K \leq \min G'|H_j$ e $G'|H_j = G_j|H_j$, ove G_j è il guadagno di una scommessa su eventi di $\mathcal{E}|H_j$, come convenuto sopra. Poiché per ipotesi $P(\cdot|H_j)$ è coerente su $\mathcal{E}|H_j$, finalmente si trova $\min G'|K \leq \min G_j|H_j \leq 0$, cioè $\min G'|K \leq 0$ anche in questo caso, e ciò completa la prova. ■

COMMENTO.

Ogni probabilità coerente su $\mathcal{F} \cup \mathcal{E}|\mathcal{H}$ determina per restrizione una probabilità coerente su \mathcal{F} e una su ogni insieme $\mathcal{E}|H$, $H \in \mathcal{H}$. Il teorema ora dimostrato mostra che grazie alla particolare struttura del dominio – \mathcal{H} insieme di eventi incompatibili a due a due e $\mathcal{F} \subset \mathcal{A}_1(\mathcal{H})$ – è vero anche il viceversa: scegliendo a piacere una probabilità su \mathcal{F} e una su ogni $\mathcal{E}|H$, $H \in \mathcal{H}$, la loro estensione congiunta su $\mathcal{F} \cup \mathcal{E}|\mathcal{H}$ è una probabilità coerente.

COROLLARIO.

Siano \mathcal{E} un insieme di eventi e \mathbb{P} una partizione di ipotesi. Allora, assegnate a piacere una probabilità $P(\cdot)$ su \mathbb{P} e una $P(\cdot|H)$ per ogni $H \in \mathbb{P}$, la loro esten-

sione congiunta è coerente e si prolunga in modo unico su $\mathcal{E} \wedge \mathbb{P}$ ponendo $P(E \wedge H) = P(H)P(E|H)$ per ogni $E \in \mathcal{E}$ e $H \in \mathbb{P}$.

DIMOSTRAZIONE.

Posto $\mathcal{H} = \mathcal{F} = \mathbb{P}$ nel Teorema 13.4.2, si ha intanto che l'estensione congiunta della $P(\cdot)$ e delle $P(\cdot|H)$, $H \in \mathbb{P}$, – estensione di dominio $\mathbb{P} \cup \mathcal{E}|\mathbb{P}$ – è una probabilità coerente. Dovendo rispettare il teorema delle probabilità composte si deve porre $P(E \wedge H) = P(H)P(E|H)$ per ogni $E \wedge H \in \mathcal{E} \wedge \mathbb{P}$ e si ottiene così il suo unico prolungamento coerente su $\mathcal{E} \wedge \mathbb{P}$. ■

Usando 13.4.2 Corollario, siamo ora in grado di dimostrare il risultato preannunciato nel n° 13.4.1. Che sussiste cioè il seguente teorema.

13.4.3 Teorema. Disintegrazione su partizioni.

Siano \mathcal{E} un insieme arbitrario di eventi non vuoto e \mathbb{P} una partizione di ipotesi. Allora, assegnate a piacere una probabilità concentrata $P(\cdot)$ su \mathbb{P} e una $P(\cdot|H)$ su $\mathcal{E}|H$ per ogni $H \in \mathbb{P}$ di probabilità positiva, posto $\mathcal{H}^+ = \{H : P(H) > 0\}$, l'applicazione P di dominio \mathcal{E} definita dalla

$$P(\cdot) = \sum_{H \in \mathcal{H}^+} P(H)P(\cdot|H) \quad (42)$$

è l'unico prolungamento coerente su \mathcal{E} dell'estensione congiunta delle probabilità $P(\cdot)$ e $P(\cdot|H)$, ed è σ -additiva se sono tali le $P(\cdot|H)$, per ogni $H \in \mathcal{H}^+$. Viceversa, ogni probabilità di dominio $\mathcal{E} \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$, la cui restrizione su \mathbb{P} è concentrata, è esprimibile mediante la (42).

DIMOSTRAZIONE.

Posto $H_0 = \neg \bigvee \mathcal{H}^+$, la probabilità $P(\cdot)$, concentrata su \mathbb{P} , determina (per prolungamento) una probabilità concentrata sulla partizione $\{H_0\} \cup \mathcal{H}^+$, che assegna probabilità 0 ad H_0 e coincide con P su \mathcal{H}^+ . Scelta a piacere una probabilità $P(\cdot|H_0)$ coerente su $\mathcal{E}|H_0$, in virtù di 13.4.2 Corollario, resta allora determinata una probabilità P su $(\mathcal{E} \wedge (\{H_0\} \cup \mathcal{H}^+)) \cup (\{H_0\} \cup \mathcal{H}^+) \cup$

$\mathcal{E} \setminus (\{H_0\} \cup \mathcal{H}^+)$, tale che $P(E \wedge H) = P(H)P(E|H)$ per ogni $E \in \mathcal{E}$ e $H \in (\{H_0\} \cup \mathcal{H}^+)$ – in particolare $P(E \wedge H_0) = P(H_0)P(E|H_0) = P(H_0) = 0$ –. Rimane perciò determinata anche una probabilità *concentrata* sulla partizione $(\{H_0\} \cup \mathcal{H}^+) \wedge \{E, \bar{E}\}$, per ogni E (10.2.1 *Proposizione*), e quindi σ -additiva sull'algebra degli eventi logicamente dipendenti dalla partizione medesima. Poiché per ogni E riesce $E = (E \wedge H_0) \vee \bigvee (E \wedge \mathcal{H}^+)$, si trova che se $E \in \mathcal{E}$, l'unico modo per essere coerenti è porre $P(E) = \sum_{H \in \mathcal{H}^+} P(E \wedge H) = \sum_{H \in \mathcal{H}^+} P(H)P(E|H)$, come vuole la (42). Resta così provata la prima parte della tesi. Il viceversa è ovvia conseguenza della detta unicità, in virtù della quale se una probabilità coerente su $\mathcal{E} \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$ ha restrizione su \mathbb{P} concentrata, allora la sua restrizione su \mathcal{E} è necessariamente il prolungamento secondo la (42) della sua restrizione su $\mathbb{P} \cup \mathcal{E}|\mathbb{P}$.

Rimane da provare che se le $P(\cdot|H)$ sono σ -additive sui rispettivi insiemi $\mathcal{E}|H$, allora anche la P a primo membro della (42) è σ -additiva su \mathcal{E} . Siano per questo E_1, \dots, E_h, \dots eventi di \mathcal{E} a due a due incompatibili e sia $E = \bigvee_{h=1}^{\infty} E_h \in \mathcal{E}$. Per quanto appena provato, le probabilità degli E_h e della loro somma logica E sono univocamente determinate dalla (42). Inoltre, poiché la valutazione è coerente, la loro serie è convergente, perché deve essere $\sum_{h=1}^{\infty} P(E_h) \leq 1$ (*Proposizione* 10.4.1). Ciò premesso, usando la (42) si trova:

$$\begin{aligned} P(E) &= \sum_{H \in \mathcal{H}^+} P(H)P(E|H) = \sum_{H \in \mathcal{H}^+} P(H)P(\bigvee_{i=1}^{\infty} E_i|H) = \\ &= \sum_{H \in \mathcal{H}^+} P(H) \sum_{i=1}^{\infty} P(E_i|H) = \sum_{i=1}^{\infty} \sum_{H \in \mathcal{H}^+} P(H)P(E_i|H) = \\ &= \sum_{i=1}^{\infty} [P(E_i \wedge H_0) + \sum_{H \in \mathcal{H}^+} P(E_i \wedge H)] = \sum_{i=1}^{\infty} P(E_i), \end{aligned}$$

l'ultima uguaglianza valendo perché P è σ -additiva perché concentrata su $(\{H_0\} \cup \mathcal{H}^+) \wedge \{E_i, \bar{E}_i\}$ (10.2.1 *Proposizione*). Con ciò la prova è completa. ■

13.4.4 Probabilità \mathbb{P} -disintegrabili.

Il risultato del precedente teorema dice in sostanza che la valutazione di probabilità su \mathcal{E} può essere ottenuta eseguendo "frammenti" di valutazione – in modo concentrato su \mathbb{P} e a piacere su $\mathcal{E}|H$ per ogni $H \in \mathbb{P}$ – e "assemblandoli" poi con la (42). La portata operativa di questa procedura è palese: essa permette di pervenire all'espressione del grado di fiducia sugli eventi di \mathcal{E} in via indiretta, passando attraverso fasi di valutazione effettuate in stati d'informazione più specifici di quello relativo alla valutazione finale, generalmente perciò più facili da esprimere direttamente. La procedura suggerisce inoltre la seguente definizione.

DEFINIZIONE. Probabilità \mathbb{P} -disintegrabile.

Siano \mathcal{E} un insieme di eventi, \mathbb{P} una partizione di ipotesi e $P(\cdot|\cdot)$ una probabilità su $\mathcal{E} \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$. Diremo che $P(\cdot|\cdot)$ è \mathbb{P} -**disintegrabile** se la sua restrizione su \mathcal{E} è l'unico prolungamento coerente della sua restrizione su $\mathbb{P} \cup \mathcal{E}|\mathbb{P}$.

TERMINOLOGIA.

Al posto di \mathbb{P} -disintegrabile, diremo in seguito anche *disintegrabile su \mathbb{P}* o semplicemente *disintegrabile*, purché il dominio $\mathcal{E} \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$ della $P(\cdot|\cdot)$ sia espresso esplicitamente o anche chiaramente sottinteso.

Tenuto conto che comunque si assegnino le probabilità su \mathbb{P} e sugli insiemi $\mathcal{E}|H$, per ogni $H \in \mathbb{P}$, la loro estensione congiunta su $\mathbb{P} \cup \mathcal{E}|\mathbb{P}$ è coerente (Teorema 13.4.2), con questa terminologia il Teorema 13.4.3 dice allora che una condizione sufficiente affinché una probabilità su $\mathcal{E} \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$ sia \mathbb{P} -disintegrabile è che la sua restrizione su \mathbb{P} sia concentrata, e in tal caso la probabilità su \mathcal{E} è data dalla (42). Se \mathbb{P} ha cardinalità finita, la condizione è anche necessaria, perché le distribuzioni di probabilità su partizioni finite sono *tutte* concentrate. Ma ciò è "quasi vero" in generale, cioè a meno che gli eventi $E \in \mathcal{E}$ non siano tutti finiti o cofiniti nelle rispettive partizioni $\{E, \bar{E}\} \wedge \mathbb{P}$ (Nota 36 a piè di pagina). Lo provano le due seguenti proposizioni.

PROPOSIZIONE 1.

Siano \mathcal{E} un insieme di eventi non vuoto, \mathbb{P} una partizione di ipotesi, $P(\cdot)$ una probabilità su \mathbb{P} e $P(\cdot|H)$ una probabilità su $\mathcal{E}|H$ per ogni $H \in \mathbb{P}$ di probabilità positiva ($\mathcal{E}|H$ è vuoto se P è diffusa). Se $E \in \mathcal{E}$ è finito in $\{E, \bar{E}\} \wedge \mathbb{P}$, è allora unico il prolungamento dell'estensione congiunta di $P(\cdot)$, $P(\cdot|\cdot)$ su $\{E, \bar{E}\} \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$. La probabilità $P(E)$ è nulla se P è diffusa su \mathbb{P} ed è data invece dalla (42) altrimenti. Gli stessi risultati valgono naturalmente per \bar{E} se E è cofinito in $\{E, \bar{E}\} \wedge \mathbb{P}$ e $\bar{E} \in \mathcal{E}$.

L'unicità del prolungamento sull'evento $E = \bigvee (E \wedge \mathbb{P})$, se $E \wedge \mathbb{P}$ è finito, è conseguenza dell'unicità del prolungamento della $P(\cdot|\cdot)$ su $\mathcal{E} \wedge \mathbb{P}$ – che sussiste qualunque sia la probabilità su $\mathbb{P} \cup \mathcal{E}|\mathbb{P}$ – e dell'additività del medesimo. Il resto della tesi è ovvio. Si tratta di sommare un numero finito di addendi: di zeri se P è diffusa, di addendi del tipo $P(H)P(E|H)$ altrimenti. ■

PROPOSIZIONE 2.

Siano \mathcal{E} un insieme di eventi non vuoto, \mathbb{P} una partizione di ipotesi, $P(\cdot)$ una probabilità non concentrata su \mathbb{P} e $P(\cdot|H)$ una probabilità su $\mathcal{E}|H$ per ogni $H \in \mathbb{P}$. Se E e \bar{E} sono entrambi infiniti in $\{E, \bar{E}\} \wedge \mathbb{P}$ per qualche $E \in \mathcal{E}$, esistono allora infiniti prolungamenti su $\mathcal{E} \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$ dell'estensione congiunta di $P(\cdot)$, $P(\cdot|\cdot)$.

Proviamo intanto l'asserto nell'ipotesi che la probabilità P di dominio $\mathbb{P} \cup \mathcal{E}|\mathbb{P}$ – l'estensione congiunta di $P(\cdot)$, $P(\cdot|\cdot)$ – abbia restrizione $P(\cdot)$ diffusa su \mathbb{P} . Proveremo che in tal caso è coerente prolungare P su $\{E\} \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$ attribuendo ad E probabilità arbitraria tra 0 e 1. Avremo con ciò dimostrato che sono infiniti i prolungamenti della P su tale insieme e quindi anche su ogni suo soprainsieme. In particolare quindi anche su $\mathcal{E} \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$, come asserito.

Ora, poiché il prolungamento della P su $(\mathcal{E} \wedge \mathbb{P}) \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$ – che indicheremo in seguito ancora con P – è obbligato dal teorema delle probabilità composte, una valutazione per E è coerente con quella su $\mathbb{P} \cup \mathcal{E}|\mathbb{P}$ se e solo se lo è anche con quella su $(\mathcal{E} \wedge \mathbb{P}) \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$. Possiamo perciò provare l'asserto mostrando che è coerente prolungare la P su $\{E\} \cup (\mathcal{E} \wedge \mathbb{P}) \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$, ponendo $P(E) = p$, $0 \leq p \leq 1$. Per mostrare che l'applicazione che così si ottiene è coerente, basta ovviamente esaminare i guadagni delle scommesse che prevedono che si punti anche su E . Osserviamo intanto in proposito che è coerente la restrizione di tale applicazione su $\{E\} \cup \mathbb{P}$, perché la distribuzione su \mathbb{P} è identicamente nulla per ipotesi e allora la scelta coerente per E è arbitraria (10.4.3 Teorema). È perciò coerente anche la restrizione su $\{E\} \cup (\mathcal{E} \wedge \mathbb{P}) \cup \mathbb{P}$ (prolungamento della restrizione precedente) perché deve essere $P(E \wedge H) \leq P(H) = 0$ per ogni $H \in \mathbb{P}$. Resta con ciò da provare che soddisfano le condizioni di coerenza i guadagni che si ottengono puntando su E e su eventi di $(\mathcal{E} \wedge \mathbb{P}) \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$, di cui almeno uno di $\mathcal{E}|\mathbb{P}$, per i quali $H_0 = \Omega$. Detto G uno di tali guadagni, si tratta perciò di provare che $\min G \leq 0$. Posto G' il guadagno della parte di scommessa che riguarda eventi di $(\mathcal{E} \wedge \mathbb{P}) \cup \mathbb{P}$, riesce

$$G = S(|E| - p) + G' + \sum_{i=1}^n S_i(|E_i| - P(E_i|H_i))|H_i|,$$

$E_i|H_i, \dots, E_n|H_n \in \mathcal{E}|\mathbb{P}$. Tenendo conto che $P(E_i \wedge H_i) = P(H_i) = 0$ per ogni $E_i \in \mathcal{E}$ e $H_i \in \mathbb{P}$, il guadagno G si può scrivere:

$$G = S(|E| - p) + G' + \sum_{i=1}^n (-S_i P(E_i|H_i))|H_i| + \sum_{i=1}^n S_i |E_i \wedge H_i|.$$

Posto in questa forma, G appare allora come il guadagno di una scommessa su eventi di $\{E\} \cup (\mathcal{E} \wedge \mathbb{P}) \cup \mathbb{P}$, ove abbiamo dimostrato che la restrizione dell'applicazione definita su $\{E\} \cup (\mathcal{E} \wedge \mathbb{P}) \cup \mathbb{P} \cup \mathcal{E}|\mathbb{P}$ è coerente. Si ha allora

$\min G \leq 0$, come volevasi. Come anticipato sopra, abbiamo con ciò provato che se la restrizione $P(\cdot)$ di P su \mathbb{P} è una probabilità diffusa, i suoi prolungamenti da $\mathbb{P} \cup \mathcal{E} | \mathbb{P}$ a $\mathcal{E} \cup \mathbb{P} \cup \mathcal{E} | \mathbb{P}$ sono infiniti.

Resta da provare che ciò accade anche se la $P(\cdot)$ è *parzialmente diffusa (concentrata)* su \mathbb{P} . Indicata a tal fine con γ la parte concentrata della $P(\cdot)$ su \mathbb{P} , con P_c e P_d le corrispondenti componenti (probabilità) concentrata e diffusa, dall'ipotesi segue $0 < \gamma < 1$. La componente P_c è perciò ben determinata (Teorema 10.4.2) e riesce $P(H) = \gamma P_c(H) + (1 - \gamma) P_d(H)$, per ogni $H \in \mathbb{P}$. Le estensioni congiunte di $P_c, P(\cdot | \cdot)$ e $P_d, P(\cdot | \cdot)$, entrambe di dominio $\mathbb{P} \cup \mathcal{E} | \mathbb{P}$, sono coerenti e la prima si prolunga in modo *unico* su $\mathcal{E} \cup \mathbb{P} \cup \mathcal{E} | \mathbb{P}$ (Teorema 13.4.3), mentre la seconda, come appena visto, ammette *infiniti* prolungamenti sul medesimo insieme. Indicata con P' la restrizione su $\mathcal{E} \cup \mathbb{P}$ del primo prolungamento e con P'' quella di uno dei secondi, sostituendo \mathcal{E} con $\mathcal{E} \cup \mathbb{P}$ e \mathcal{D} con $\mathcal{E} | \mathbb{P}$ nel Lemma 13.3.1 si ricava che la mistura con pesi $\gamma, 1 - \gamma$ delle estensioni congiunte $P', P(\cdot | \cdot)$ e $P'', P(\cdot | \cdot)$ è una valutazione di probabilità coerente su $\mathcal{E} \cup \mathbb{P} \cup \mathcal{E} | \mathbb{P}$. È coerente cioè l'estensione congiunta di

$$\gamma P' + (1 - \gamma) P'', \quad \gamma P_c + (1 - \gamma) P_d, \quad \gamma P(\cdot | \cdot) + (1 - \gamma) P(\cdot | \cdot),$$

ovvero quella di $\gamma P' + (1 - \gamma) P'', P(\cdot), P(\cdot | \cdot)$, ove si legge che $\gamma P' + (1 - \gamma) P''$ è un prolungamento coerente su \mathcal{E} dell'estensione congiunta di $P(\cdot), P(\cdot | \cdot)$. Ciò accade comunque si scelga il prolungamento P'' fra gli infiniti possibili, e ciò completa la prova della proposizione. ■

Osserviamo infine che il Teorema 13.4.3 non dice nulla sull'unicità della rappresentazione della probabilità su \mathcal{E} mediante la (42). E ciò per il semplice motivo che questo risultato non sussiste, perché mentre $P(\cdot)$ è l'unico prolungamento coerente su \mathcal{E} della restrizione della $P(\cdot | \cdot)$ su $\mathbb{P} \cup \mathcal{E} | \mathbb{P}$, non sussiste il viceversa. Vediamolo con un controesempio.

CONTROESEMPIO.

Sia $\mathcal{E} = \{E\}$, $\mathbb{P} = \{H, \bar{H}\}$, E e H possibili. Posto $P(E) = a$, a arbitrario tra 0 e 1, $P(\cdot)$ è coerente su $\{E\}$ ed esistono perciò suoi prolungamenti su $\{H, \bar{H}\} \cup \{E | \bar{H}\} \cup \{E | H\}$ che, per quanto appena visto, determinano $P(E)$ mediante la (42). Si tratta di far vedere che esiste più di una rappresentazione di questo tipo. Poniamo per questo $P(H) = p$, $P(E | H) = b$, $P(E | \bar{H}) = c$. Poiché $P(E)$ è mistura di $P(E | H)$ e $P(E | \bar{H})$, si deve scegliere $0 \leq b \leq a \leq c \leq 1$. Determinando a con la (42) si trova $a = pb + (1 - p)c$, da cui si ricava $p = (c - a) / (c - b)$. Sono quindi ∞^2 le scelte di $b = P(E | H)$ e $c = P(E | \bar{H})$ nei limiti stabiliti che – calcolato il corrispondente p – conducono allo stesso valore di $a = P(E)$.

ESEMPLI.

- 1 Un'urna contiene b palline bianche e r rosse. Si eseguono due estrazioni di una pallina per volta a caso. Posto $E_i = \text{«esce bianca all}'i\text{-esima estrazione»}$, calcolare la probabilità di E_2 nelle due seguenti ipotesi sulla modalità d'estrazione: (i) con rimessa, (ii) senza rimessa.

Usando la (42) con le due ipotesi $H_1 = E_1, H_2 = \bar{E}_1$ e $\mathcal{G} = \{E_2\}$, si ottiene $P(E_2) = P(E_1)P(E_2|E_1) + P(\bar{E}_1)P(E_2|\bar{E}_1)$. Nelle due ipotesi sulle modalità d'estrazione si ricava:

$$(i) \quad P(E_2) = \frac{b}{b+r} \frac{b}{b+r} + \frac{r}{b+r} \frac{b}{b+r} = \frac{b}{b+r} \frac{b+r}{b+r} = \frac{b}{b+r} = P(E_1)$$

$$(ii) \quad P(E_2) = \frac{b}{b+r} \frac{b-1}{b+r-1} + \frac{r}{b+r} \frac{b}{b+r-1} = \frac{b}{b+r} \frac{b-1+r}{b+r-1} = \frac{b}{b+r} = P(E_1)$$

Il lettore provi che in entrambe le modalità d'estrazione riesce anche $P(E_3) = P(E_1)$, in accordo con quanto già visto nel caso particolare $b=8$ e $r=2$ in 13.2.4 Esempio 4, usando qui la formula di disintegrazione sulla partizione $\{E_1 \wedge E_2, E_1 \wedge \bar{E}_2, \bar{E}_1 \wedge E_2, \bar{E}_1 \wedge \bar{E}_2\}$.

Commento.

Riteniamo utile cogliere l'occasione per segnalare un'insidia a cui si va incontro se viene a mancare una corretta interpretazione dello stato d'informazione in cui si effettua la valutazione, che è quello che si ha all'inizio, *prima di eseguire le singole estrazioni*. In quel momento, se la modalità d'estrazione è quella con rimessa, è allora nota la composizione dell'urna – numero totale di palline e ripartizione in bianche e rosse –, sia all'atto della prima che in quello della seconda estrazione. Altrimenti – nella modalità senza rimessa –, la composizione è nota solo all'atto della prima estrazione. All'atto della seconda la si conosce, invece, in modo incompleto: è noto il numero totale delle palline nell'urna ($b+r-1$), ma la composizione è aleatoria: $b-1$ bianche, r rosse o b bianche, $r-1$ rosse. Ciò può essere motivo di confusione se, suggestionati dall'ordine cronologico delle estrazioni, si sostituisce erroneamente all'evento E_2 l'evento $E_2|E_1$ ($E_2|\bar{E}_1$) come se fosse noto che al primo colpo è stata estratta pallina bianca (rossa). Allora sì che le cose cambierebbero, perché al posto di $P(E_2)$ si verrebbe a calcolare $P(E_2|E_1) = (b-1)/(a+b-1)$ ($P(E_2|\bar{E}_1) = b/(a+b-1)$).

- 2 Da un'urna contenente 2 palline rosse, e palline bianche in numero pari al punto realizzato lanciando un dado, se ne estrae una. Calcolare la proba-

bilità che sia bianca.

Posto $E =$ la pallina estratta è bianca e $H_i =$ il punto realizzato dal dado è i , $i = 1, \dots, 6$, usando la (42) – disintegrando su $\{H_1, \dots, H_6\}$ – si ha:

$$P(E) = \sum_{i=1}^6 P(H_i)P(E|H_i) = \frac{1}{6} \left(\frac{1}{3} + \frac{2}{4} + \dots + \frac{6}{8} \right) = 0,59405.$$

- 3 Da un'urna contenente 2 palline rosse, e palline bianche in numero pari al tempo d'attesa del primo successo (esce testa) in una partita a testa e croce, se ne estrae una. Calcolare la probabilità che sia bianca.

Poniamo $E =$ la pallina estratta è bianca e $H_i =$ il primo successo si verifica all' n -esimo lancio, $H_\infty =$ testa non esce mai. Nella consueta ipotesi di simmetria riesce $P(H_i) = 1/2^i$, $P(H_\infty) = 0$ (5.2.3 Esempio 1) e la corrispondente distribuzione sulla partizione $\{H_1, \dots, H_i, \dots, H_\infty\}$ è perciò concentrata. Usando la (42) si trova:

$$P(E) = \sum_{i=1}^{\infty} P(H_i)P(E|H_i) = \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^i \frac{i}{i+2}. \quad (43)$$

Arrestando la somma all' n -esimo termine si commette un errore minore di $1/2^n$, come segue dalle:

$$\sum_{i=n+1}^{\infty} \left(\frac{1}{2}\right)^i \frac{i}{i+2} < \sum_{i=n+1}^{\infty} \left(\frac{1}{2}\right)^i = \left(\frac{1}{2}\right)^n.$$

Per $n = 10$ si trova $P(E) = 0,454$ a meno di 1×10^{-3} . Come si vede, dovendo puntare sul colore della pallina estratta conviene farlo sul rosso, anche se solo nell'ipotesi H_1 il rosso prevale sul bianco ed è $P(\bar{E}|H_1) = 2/3$. Nell'ipotesi H_2 le probabilità di rosso e bianco sono uguali e nelle ipotesi successive il bianco prevale sul rosso in modo sempre più netto. Non tanto però da far recuperare alla probabilità di E – misturando le relative probabilità condizionate come previsto dalla (43) – quanto si perde rispetto all'equo se il primo successo si verifica al primo lancio ($1/2 - 1/3 = 1/6$). Ciò è dovuto al fatto che le probabilità delle ipotesi favorevoli all'uscita di pallina bianca decrescono molto rapidamente (in modo esponenziale), e il divario a favore delle bianche descritto da tali ipotesi diventa presto irrilevante in termini di valutazione. Se ne ha conferma considerando il problema analogo in cui l'urna contiene solo una pallina rossa. In tal caso conviene certamente puntare sul bianco, perché nell'ipotesi peggiore – la H_1 – si ha la stessa probabilità di estrarre pallina bianca o rossa. Si tratta comunque di una condizione favorevole abbastanza contenuta, perché a fronte della valutazione $P(E|H_i) = P(E|\bar{H}_i) = 0,5$ a si

trova $P(E) = 0,613$.

- 4 *Uno spedizioniere desidera assicurare un lotto di merci che deve essere portato a destinazione via mare. Per semplicità, supponiamo che la merce vada perduta tutta o niente; che sia cioè così deperibile da risultare inseribile anche a seguito di un lieve danneggiamento. Come calcolare il premio di assicurazione?*

Nelle ipotesi poste, come s'intuisce, per il calcolo del premio è sufficiente conoscere il valore della merce – che sarà coperto in toto o in parte, secondo contratto – e la probabilità dell'evento $E = \text{durante il trasporto la merce si deteriora}$. È ovvio immaginare che le condizioni del mare – più in generale quelle meteorologiche – abbiano influenza sulla valutazione della probabilità di sinistro (perdita della merce) ed è naturale ritenere che essa aumenti col peggiorare delle condizioni atmosferiche. È altresì evidente che la conoscenza dello stato delle condizioni atmosferiche renderebbe più facile la valutazione di tale probabilità. Per tradurre in termini operativi queste considerazioni, bisogna dire però qualche cosa di meno vago sul significato da dare a «condizioni atmosferiche». Occorre anzi darne una descrizione ben definita, anche se necessariamente in una forma approssimata. Se è lo stato del mare che interessa – come qui supporremo – esso si può descrivere adoperando la usuale scala delle «forze» (ad es. mare forza 1, mare forza 2, forza 3, ecc.). O più semplicemente – a titolo d'esempio ciò è sufficiente – esaurendo la descrizione con le tre condizioni (eventi a due a due incompatibili): $C = \text{il mare è calmo}$, $M = \text{il mare è mosso}$, $B = \text{il mare è burrascoso}$, $C \vee M \vee B = \Omega$. L'assicuratore – è lui che deve valutare $P(E)$ – per arrivare ad esprimere il suo grado di fiducia, se non ha sufficiente dimestichezza con questi problemi, può cercare aiuto rivolgendosi ad esperti. Può consultare, ad esempio, armatori, imprenditori ed ingegneri navali – anche statistiche, se ve ne sono – per pervenire alle valutazioni di $E|C$, $E|M$ e $E|B$; esperti di meteorologia, invece, per giungere alla valutazione delle probabilità degli eventi C , M , B che descrivono le condizioni del mare durante il viaggio (supposto qui di breve durata per rendere accettabile l'ipotesi che lo stato del mare rimanga invariato durante il viaggio). "Assemblando" come prescritto dalla (42) queste valutazioni egli sarà in grado a questo punto di costruire la sua $P(E)$.

Per una esemplificazione numerica, supponiamo che il nostro assicuratore sia giunto alle seguenti valutazioni:

$$\begin{array}{lll} P(E|C) = 0,001 & P(E|M) = 0,008 & P(E|B) = 0,031 \\ P(C) = 0,17 & P(M) = 0,80 & P(B) = 0,03 \end{array}$$

Da qui si ricava:

$$P(E) = 0,17 \times 0,001 + 0,80 \times 0,008 + 0,03 \times 0,031 = 0,0075.$$

13.5 Proprietà conglomerativa

Prendiamo in esame la formula (42) del *Teorema* 13.4.3. Tenuto conto che $\sum_{H \in \mathfrak{H}^+} P(H) = 1$, si ricava

$$P(E) = \sum_{H \in \mathfrak{H}^+} P(H)P(E|H) \geq \inf_{H \in \mathfrak{H}^+} P(E|H) \sum_{H \in \mathfrak{H}^+} P(H) = \\ \inf_{H \in \mathfrak{H}^+} P(E|H) \geq \inf_{H \in \mathfrak{H}} P(E|H)$$

e simmetricamente $P(E) \leq \sup_{H \in \mathfrak{H}} P(E|H)$. Nelle nostre ipotesi – *probabilità concentrata* su \mathbb{P} – sussiste dunque la seguente proprietà: *la probabilità di ogni evento assoluto E – dell'insieme \mathcal{E} , che con riferimento al teorema citato è arbitrario – si colloca tra le valutazioni date ai corrispondenti eventi $E|H$ condizionati alle ipotesi $H \in \mathbb{P}$ di probabilità positiva, e quindi anche tra le valutazioni degli $E|H$ condizionati alla totalità delle ipotesi H di \mathbb{P}* . In letteratura questa proprietà è detta *conglomerativa*. Si dà precisamente la seguente definizione.

13.5.1 Definizione. Probabilità \mathbb{P} -conglomerabile.

Sia \mathcal{E} un insieme di eventi, \mathbb{P} una partizione di ipotesi e P una probabilità su $\mathcal{E} \cup \mathcal{E}|\mathbb{P}$. Diremo che P è **\mathbb{P} -conglomerabile in \mathcal{E}** , se per ogni $E \in \mathcal{E}$ riesce:

$$\inf_{H \in \mathbb{P}} P(E|H) \leq P(E) \leq \sup_{H \in \mathbb{P}} P(E|H). \quad (44)$$

COMMENTO.

Il termine «conglomerare» sintetizza il significato della condizione (44). Esso è riferito alla probabilità $P(E)$ dell'evento E , per dire che questa è "conglobata" nell'insieme delle valutazioni condizionate, nel senso che $P(E)$ appartiene al minimo intervallo contenente tutte le valutazioni condizionate $P(E|H)$, $H \in \mathbb{P}$. Rispetto alla proprietà di disintegrabilità, che nella forma espressa dalla (42) del *Teorema* 13.4.3 fornisce l'unico valore coerente della probabilità di E , la proprietà conglomerativa si limita a indicare un intervallo entro cui si deve trovare tale probabilità. Essa è perciò *meno informativa* di quella di disinte-

grabilità. Il che è naturale, perché la proprietà conglomerativa si riferisce a una *valutazione di probabilità più vaga* – riguarda gli eventi di $\mathcal{E}|\mathbb{P}$ – di quella che serve per utilizzare la formula di disintegrabilità (42), che ha bisogno che sia esplicitata anche la distribuzione di probabilità su \mathbb{P} . Interessa osservare ancora che l'intervallo di indeterminazione degenera in un punto se (e solo se) le $P(E|H)$ sono tutte uguali. È allora determinata, e uguale alla valutazione congiunta degli eventi $E|H$, anche la probabilità di E ⁴⁵. È utile per il seguito mettere in evidenza questo caso particolare mediante la seguente proposizione.

PROPOSIZIONE.

Siano \mathcal{E} un insieme di eventi, \mathbb{P} una partizione di ipotesi, P una probabilità \mathbb{P} -conglomerabile in \mathcal{E} . Allora, per ogni $E \in \mathcal{E}$ per cui $P(E|\cdot)$ è costante su \mathbb{P} si ha $P(E) = P(E|H)$, H arbitrario in \mathbb{P} .

TERMINOLOGIA.

Come nel caso della \mathbb{P} -disintegrabilità (13.4.4 *Terminologia*), quando la partizione \mathbb{P} è espressa o chiaramente sottintesa, al posto di \mathbb{P} -conglomerabile in \mathcal{E} diremo più semplicemente *conglomerabile (in \mathcal{E})*.

È naturale a questo punto chiedersi se la proprietà valga in condizioni più generali, con riferimento cioè a probabilità non concentrate. Osserviamo intanto in proposito che una condizione sufficiente per la conglomerabilità è che la probabilità sia disintegrabile per gli eventi di \mathcal{E} – cosa che si poteva attendere dopo aver letto il precedente commento –, cioè che la probabilità sia concentrata come detto in premessa – e in questo caso la proprietà sussiste per gli eventi di ogni insieme – oppure che \mathcal{E} sia costituito da eventi finiti o cofiniti (13.4.4 *Proposizione 1*). Sussiste infatti il seguente teorema.

45 Non è detto che si possa scegliere sempre – per ogni $H \in \mathbb{P}$ – le probabilità $P(E|H)$ tutte uguali senza incorrere in incoerenze. Basti pensare per questo a un evento E logicamente semidipendente da \mathbb{P} in modo bilaterale. Si deve allora porre $P(E|H) = 1$ in corrispondenza alle ipotesi H che implicano E , e $P(E|H) = 0$ per quelle che implicano \bar{E} . È coerente invece scegliere un valore comune per le $P(E|H)$ – e per giunta arbitrario in assenza di altre condizioni – se E è logicamente indipendente da \mathbb{P} (conseguenza della (21) di 10.4.3 *Teorema*). È coerente farlo – scegliere le $P(E|H)$ uguali – anche se E è logicamente semidipendente in modo unilaterale, ma con scelta obbligata ed estrema: 0 o 1 nei due casi.

13.5.2 Teorema.

Sia \mathcal{E} un insieme di eventi, \mathbb{P} una partizione di ipotesi, P una probabilità disintegrabile su $\mathcal{E} \cup \mathbb{P} \cup \mathcal{E} | \mathbb{P}$. Allora la restrizione di P su \mathbb{P} è conglomerabile in \mathcal{E} .

DIMOSTRAZIONE.

Come abbiamo ricordato in premessa di questo paragrafo, P è disintegrabile su $\mathcal{E} \cup \mathbb{P} \cup \mathcal{E} | \mathbb{P}$ in due casi: se la restrizione di P su \mathbb{P} è concentrata – e allora l'asserto è già stato dimostrato in premessa – oppure se \mathcal{E} è costituito da eventi E finiti o cofiniti in $\mathbb{P} \wedge \{E, \bar{E}\}$ (e \mathbb{P} è perciò infinita). Dimostriamo che la probabilità è conglomerativa anche in questo caso.

Se E è finito, esiste allora un'ipotesi H tale che $E \wedge H = \emptyset$ e quindi $P(E|H) = 0$. Segue $P(E) \geq \inf_{H \in \mathbb{P}} P(E|H)$. Poiché vale la (42) del Teorema 13.4.3 si ha poi

$$P(E) = \sum_{H \in \mathbb{P}} P(H)P(E|H) \leq \sup_{H \in \mathcal{H}^+} P(E|H) \sum_{H \in \mathcal{H}^+} P(H) \leq \sup_{H \in \mathbb{P}} P(E|H).$$

e quindi la tesi quando E è finito.

Se E è cofinito, l'evento \bar{E} si trova nelle condizioni dell'evento E del caso appena esaminato. Si ha pertanto $P(\bar{E}|H) = 0$, quindi $P(E|H) = 1$ e

$$0 \leq P(\bar{E}) \leq \sup_{H \in \mathbb{P}} P(\bar{E}|H) = \sup_{H \in \mathbb{P}} (1 - P(E|H)) = 1 - \inf_{H \in \mathbb{P}} P(E|H),$$

da cui segue: $\inf_{H \in \mathbb{P}} P(E|H) \leq P(E) \leq 1 = \sup_{H \in \mathbb{P}} P(E|H)$ e con ciò la prova è completa. ■

Abbiamo osservato in 13.5.1 *Commento* che a livello di definizione la proprietà conglomerativa è meno restrittiva di quella di disintegrabilità, e il teorema appena dimostrato ne è una conferma. Si pongono a questo punto due questioni. La prima è quella di vedere se la nozione di conglomerabilità è effettivamente meno restrittiva di quella di disintegrabilità – se la conglomerabilità *non implica* la disintegrabilità – e in caso affermativo se esistono probabilità *non conglomerabili*. Rispondono affermativamente a entrambi i quesiti i due esempi che costituiscono il prossimo complemento.

13.5.3 Complemento.

ESEMPIO 1. *La conglomerabilità non implica la disintegrabilità.*

Con riferimento alla scelta a caso di un numero naturale nel senso ampliato di

13.3.3 *Complemento*, poniamo $\omega_i = (N = i)$, e $\mathbb{P} = \{\omega_1, \dots, \omega_n, \dots\}$. Consideriamo poi la partizione $\mathcal{H} = \{H_1, \dots, H_i, \dots\}$, meno fine di \mathbb{P} , che si ottiene sommando coppie di ω_i , uno dispari e il pari successivo – ponendo cioè $H_i = \omega_{2i-1} \vee \omega_{2i}$ –. In accordo col complemento citato, scegliere a caso nel numerabile significa introdurre in \mathcal{D} – insieme di Dynkin – la distribuzione asintotica e l'equiprobabilità su $\mathbb{P}|H$ per ogni $H \in \mathcal{A}_L^\phi(\mathbb{P})$ finito. Questa probabilità non è \mathcal{H} -disintegrabile, perché è diffusa su \mathcal{H} ed esistono eventi $E \in \mathcal{D}$ tali che E e \bar{E} sono infiniti in $\mathcal{H} \wedge \{E, \bar{E}\}$ (è tale ad esempio l'evento N è pari). Tuttavia essa è \mathcal{H} -conglomerabile in \mathcal{D} . Lo proveremo ragionando per casi sulla base della dipendenza logica di E da \mathcal{H} . Poniamo per questo $I_E = \inf_{H \in \mathcal{H}} P(E|H)$ e $S_E = \sup_{H \in \mathcal{H}} P(E|H)$. Osserviamo poi che se H_i è di tipo 1 per E ($H_i \Rightarrow E$), allora $P(E|H_i) = P(H_i|H_i) = 1$; se è di tipo 2 ($H_i \Rightarrow \bar{E}$), allora $P(E|H_i) = P(\phi|H_i) = 0$; infine se è di tipo 3, $E \wedge H_i$ e $\bar{E} \wedge H_i$ sono possibili, costituiti da un solo evento elementare di \mathbb{P} (uno dei due componenti di H_i) e riesce perciò $P(E|H_i) = P(\bar{E}|H_i) = 1/2$. Pertanto, per l'intervallo $[I_E, S_E]$ si possono presentare sei casi, di cui tre degeneri, che andiamo ad esaminare qui di seguito, in relazione al comportamento delle ipotesi.

- (i) *Le H_i sono tutte di tipo 1 ($E = \Omega$)*. Allora $I_E = P(\Omega) = S_E = 1$ e segue perciò l'asserto.
- (ii) *Le H_i sono tutte di tipo 2 ($E = \phi$)*. Allora $I_E = P(\phi) = S_E = 0$ e segue ancora l'asserto.
- (iii) *Le H_i sono tutte di tipo 3 (E logicamente indipendente da \mathcal{H})*. Allora $I_E = S_E = 1/2$. Posto poi $E_n = \omega_1 \vee \dots \vee \omega_n$, riesce $\text{card}(E \wedge H_i) = 1$, per ogni ipotesi H_i , e quindi $\text{card}(E \wedge E_{2n}) = n$. Si trova perciò

$$P(E) = \lim_{n \rightarrow \infty} \text{card}(E \wedge E_n) / n = \lim_{n \rightarrow \infty} \text{card}(E \wedge E_{2n}) / 2n = 1/2,$$

ove il primo limite esiste per definizione di $P(E)$ (10.5.2 *Esempio*) e il secondo – uguale al primo – perché limite di una sottosuccessione della precedente. Segue $I_E = P(E) = S_E = 1/2$ e con ciò l'asserto.

- (iv) *Esistono H_i di tipo 1 e 2 (E logicamente dipendente o semidipendente bilateralmente da \mathcal{H})*. Allora $I_E = 0$ e $S_E = 1$ e l'asserto è banale.
- (v) *Esistono ipotesi di tipo 1 e 3, ma non di tipo 2 (E logicamente semidipendente inferiormente da \mathcal{H})*. Allora $I_E = 1/2$ e $S_E = 1$. Argomentando come in (iii) si trova poi $\text{card}(E \wedge E_{2n}) \geq n$, quindi $\text{card}(E \wedge E_{2n}) / 2n \geq 1/2$ e, passando al limite per $n \rightarrow \infty$, $P(E) \geq 1/2$, da cui segue l'asserto.

- (vi) Esistono ipotesi di tipo 2 e 3, ma non di tipo 1 (E logicamente semi-dipendente superiormente da \mathcal{H}). Allora $I_E = 0$ e $S_E = 1/2$. Ora riesce $\text{card}(E \wedge E_{2n})/2n \leq 1/2$, perciò $0 = I_E \leq P(E) \leq S_E = 1/2$ e con ciò ancora l'asserto. La prova è così completa. ■

ESEMPIO 2. *Un caso di non conglomerabilità.*

Nelle ipotesi del precedente *Esempio 1* – distribuzione asintotica in senso ampliato sulla partizione numerabile $\mathbb{P} = \{\omega_1, \dots, \omega_n, \dots\}$ – consideriamo la partizione \mathcal{H} , meno fine di \mathbb{P} , che si ottiene sommando gli eventi elementari di \mathbb{P} a gruppi di tre, facendo in modo che per ciascun evento elementare di \mathcal{H} siano presenti un evento elementare di indice pari e due di indice dispari. Ad esempio, se eseguiamo l'operazione di raggruppamento percorrendo la successione (ω_n) nell'ordine naturale e poniamo $(i, j, k) = \omega_i \vee \omega_j \vee \omega_k$, otteniamo:

$$H_1 = (2, 1, 3), H_2 = (4, 5, 7), \dots, H_n = (2n, 4n-3, 4n-1), \dots$$

Posto $E = N$ è pari, riesce come sappiamo $P(E) = 1/2$, mentre è $P(E|H_n) = P(\omega_{2n}|H_n) = 1/3$ per ogni n . Per $E \in \mathcal{D}$ non vale perciò la proprietà conglomerativa (più specificamente, la 13.5.1 *Proposizione*).

13.5.4 Complemento. Disintegrabilità e conglomerabilità in ambiente condizionato a una sola ipotesi.

Come già segnalato nel n° 13.1.3, tutti i risultati che si conseguono ragionando con eventi assoluti, continuano a valere se lo stato d'informazione α viene sostituiti da una stato incrementato $\alpha \wedge H$. Ciò vale naturalmente anche per le nozioni di disintegrabilità e di conglomerabilità riferite a eventi condizionati a una medesima ipotesi e per i relativi risultati. Sarebbe perciò inutilmente gravoso ripetere tutto quanto è stato detto sull'argomento in termini condizionati, perché tutto è facilmente adattabile – nelle definizioni, negli enunciati dei teoremi e nelle dimostrazioni – sostituendo gli eventi assoluti con eventi condizionati a una medesima ipotesi. È sufficiente perciò dare qui a titolo d'esempio le versioni condizionate a un'ipotesi H delle definizioni di disintegrabilità e di conglomerabilità e quella del *Teorema 13.4.3*, che avremo occasione di richiamare più spesso.

DEFINIZIONE 1. Probabilità $(\mathbb{P}|H)$ -disintegrabile.

Siano \mathcal{E} un'insieme di eventi, H un'ipotesi, \mathbb{P} una partizione di ipotesi e P una

probabilità su $\mathcal{E}|H \cup \mathbb{P}|H \cup \mathcal{E}|(\mathbb{P} \wedge H)$. Diremo che P è $(\mathbb{P}|H)$ -disintegrabile se la sua restrizione su $\mathcal{E}|H$ è l'unico prolungamento coerente della sua restrizione su $\mathbb{P}|H \cup \mathcal{E}|(\mathbb{P} \wedge H)$.

TEOREMA.

Siano \mathcal{E} un insieme arbitrario di eventi non vuoto, \mathbb{P} una partizione di ipotesi e K un'ipotesi. Allora, assegnate a piacere una probabilità concentrata $P(\cdot|K)$ su $\mathbb{P}|K$ e una $P(\cdot|K \wedge H)$ per ogni $H|K \in \mathbb{P}|K$ di probabilità positiva, posto $\mathcal{H}^+ = \{H|K : P(H|K) > 0\}$, l'applicazione P di dominio $\mathcal{E}|K$ definita dalla

$$P(\cdot|K) = \sum_{H|K \in \mathcal{H}^+} P(H|K) P(\cdot|K \wedge H) \quad (45)$$

è l'unico prolungamento coerente su $\mathcal{E}|K$ dell'estensione congiunta delle probabilità $P(\cdot|K)$ e $P(\cdot|K \wedge H)$, ed è σ -additiva se sono tali le $P(\cdot|K \wedge H)$, per ogni $H|K \in \mathcal{H}^+$. Viceversa, ogni probabilità di dominio $\mathcal{E}|K \cup \mathbb{P}|K \cup \mathcal{E}|(\mathbb{P} \wedge K)$, la cui restrizione su $\mathbb{P}|K$ è concentrata, è esprimibile mediante la (45).

DEFINIZIONE 2. Probabilità $(\mathbb{P}|K)$ -conglomerabile.

Siano \mathcal{E} un insieme di eventi, \mathbb{P} una partizione di ipotesi, K logicamente indipendente da \mathbb{P} e P una probabilità su $\mathcal{E}|K \cup \mathcal{E}|(\mathbb{P} \wedge K)$. Diremo che P è $(\mathbb{P}|K)$ -conglomerabile in $\mathcal{E}|K$, se per ogni $E|K \in \mathcal{E}|K$ riesce:

$$\inf_{H \in \mathbb{P}} P(E|K \wedge H) \leq P(E|K) \leq \sup_{H \in \mathbb{P}} P(E|K \wedge H).$$

13.6 Deduzione tardiva

Lo stato di informazione ha un ruolo fondamentale nella nozione di evento. Lo abbiamo sottolineato sin dall'inizio in 2.2.2 *Commento*, e messo in tutta evidenza nel § 12.2 introducendo la nozione di stati ugualmente informativi e nel § 12.3 trattando della sua evoluzione. Abbiamo però trascurato sin qui di approfondire un aspetto che abbiamo segnalato già in *Nota 12* a piè di pagina, per il semplice fatto che solo ora siamo in grado di commentarlo in maniera adeguata.

Parlando delle fasi descrittive dei problemi in condizioni di incertezza, si è detto che la conoscenza dei valori delle proposizioni di un insieme \mathcal{L} dipende dallo stato di informazione α in cui uno si trova. Nella nota citata si è detto anche però che in realtà essa dipende da quanto uno è *in grado di derivare dall'ipotesi* α in forza della sua *personale* capacità deduttiva. E che

non sarebbe sicuramente realistico considerare tale capacità illimitata, in grado cioè di garantire *in ogni caso* che saranno individuate tutte le proposizioni di \mathcal{L} il cui valore è deducibile a partire dall'ipotesi α . Per rendere operativa la teoria bisogna allora ammettere che il soggetto possa passare dalla fase descrittiva a quella della valutazione, dopo aver acquisito solo parte delle conoscenze potenzialmente raggiungibili in uno stato d'informazione. Circa la valutazione, affinché la teoria conservi consistenza, importa che l'aggiunta di *deduzioni tardive* non impedisca di prolungare la valutazione *in coerenza* con le scelte precedenti. Come andremo ora a mostrare, la questione è un aspetto particolare del problema del prolungamento della valutazione a seguito di incremento di informazione, trattato nei paragrafi precedenti. Per semplicità – e perché lo si può sempre fare – supporremo qui nel seguito che \mathcal{L} sia un linguaggio generato (§ 2.3, n° 3.2.1).

In generale, come abbiamo visto, un incremento di informazione β porta nuova conoscenza e fa nascere tra gli eventi relazioni che prima non sussistevano. Formalmente una relazione nasce perché una proposizione p incerta sub α – per la quale non sussiste $|\alpha p$ – è dedotta vera sub $\alpha \wedge \beta$ – sussiste ora la $|\alpha \wedge \beta p$ –. Il caso della deduzione tardiva rientra in questo schema. Per intendere bene il discorso, è opportuno mettere in evidenza nei simboli – ma solo in questo paragrafo – la dipendenza dell'insieme degli eventi dal grado di deduzione raggiunto *nel tempo* dal soggetto. Scriveremo perciò $\mathcal{L}_{|\alpha,t}$ e $p_{|\alpha,t}$ per sottolineare, nell'ordine, che l'insieme di eventi e l'evento vanno riferiti allo *stato d'informazione α nell'istante t* . Ammetteremo inoltre che le *conoscenze acquisite non vengono mai più dimenticate*. Sicché, se è $\tau > t$ e non arriva nuova informazione nell'intervallo temporale $[t, \tau]$, allora il quadro delle possibilità $\mathcal{L}_{|\alpha,\tau}$ coincide – tolto il riferimento temporale – con $\mathcal{L}_{|\alpha,t}$. L'ipotesi non è da poco. Può essere però ritenuta ragionevole in molte circostanze pratiche (o perché il problema viene esaminato in un arco temporale sufficientemente breve o perché le conoscenze raggiunte vengono registrate, ...).

Ciò premesso, il problema della deduzione tardiva si inquadra nello schema generale come segue. Sia $\mathcal{L}_{|\alpha,t}$ l'insieme degli eventi individuati da un soggetto all'istante t e $p_{|\alpha,t}$ un evento possibile. Tale cioè che non sia a lui noto in quell'istante il valore di p , o perché non è deducibile nell'ipotesi α – e allora $p_{|\alpha,t}$ è possibile per tutti – o perché è deducibile ma non da lui dedotto. Se questo è il caso, può allora capitare che il nostro soggetto venga ad apprendere in un istante t' successivo a t che p è vera e *non altro* (per sua deduzione o perché informato da altro soggetto). Il suo insieme di eventi all'istante t' diventa allora $\mathcal{L}_{|\alpha \wedge p,t'}$. Il caso della deduzione tardiva rientra perciò nello schema generale dell'evoluzione della nozione di evento per incremento d'informazione, ponendo ivi l'incremento $\beta = p$. Andiamo ad illustrare la

questione con degli esempi.

ESEMPIO 1.

Supponiamo che all'istante t un soggetto non si sia accorto che $a \leftrightarrow b$ è deducibile vera sub α . Per lui, allora, gli eventi $a|_{\alpha,t}$ e $b|_{\alpha,t}$ sono diversi e l'evento $(a \leftrightarrow b)|_{\alpha,t}$ è possibile, perché essendo deducibile che « $a \leftrightarrow b$ è vera», non è possibile che abbia dedotto che « $a \leftrightarrow b$ è falsa» (il nostro soggetto difetta in deduzione, ma non commette errori!). Ha senso perciò considerare in ogni istante $\tau \geq t$ l'insieme di eventi $\mathcal{L}_{|\alpha \wedge (a \leftrightarrow b), \tau}$, ove all'epoca τ lo stato d'informazione $\alpha \wedge (a \leftrightarrow b)$ è *ipotetico* se il soggetto non ha ancora dedotto che « $a \leftrightarrow b$ è vera», altrimenti è della stessa natura di α (*reale* se α è reale). In entrambi i casi, osservato che $[a \wedge (a \leftrightarrow b)] \leftrightarrow [b \wedge (a \leftrightarrow b)]$ è una tautologia e posto per brevità $h = (a \leftrightarrow b)$, si ha (12.3.3 c):

$$a|_{\alpha \wedge h, \tau} = (a \wedge h)|_{\alpha \wedge h, \tau} = (b \wedge h)|_{\alpha \wedge h, \tau} = b|_{\alpha \wedge h, \tau}.$$

Sono perciò uguali i due eventi $a|_{\alpha \wedge h, \tau}$ e $b|_{\alpha \wedge h, \tau}$, e vanno considerati come *eventi condizionati* e coincidenti con $a|_{\alpha \wedge h, t}$ e $b|_{\alpha \wedge h, t}$ se il soggetto non ha ancora dedotto all'epoca τ che « $a \leftrightarrow b$ è vera» – il suo insieme di eventi assoluti è invece $\mathcal{L}_{|\alpha, \tau}$ e coincide con quello iniziale $\mathcal{L}_{|\alpha, t}$ –, come eventi assoluti altrimenti. Se all'istante $t' > t$ il soggetto deduce che « $a \leftrightarrow b$ è vera» e null'altro rispetto a quanto ha fatto in t , l'incremento d'informazione diventa reale e il quadro delle possibilità effettivo diventa $\mathcal{L}_{|\alpha \wedge h, t'}$, i cui eventi sono *quelli di $\mathcal{L}_{|\alpha, t}$ condizionati a $h|_{\alpha, t}$* . Inoltre, gli stati d'informazione α e $\alpha \wedge h$ sono ugualmente informativi (12.2.3 b) – mentre non lo erano prima – e perciò si ha $p|_{\alpha, t'} = p|_{\alpha \wedge h, t'} = p|_{\alpha \wedge h, t}$ per ogni $p \in \mathcal{L}$. In particolare riesce $a|_{\alpha, t'} = a|_{\alpha \wedge h, t} = b|_{\alpha \wedge h, t} = b|_{\alpha, t'}$, mentre era $a|_{\alpha, t} \neq b|_{\alpha, t'}$.

Esaminiamo ora la questione dal punto di vista probabilistico. Poiché $p|_{\alpha, t'} = p|_{\alpha \wedge (a \leftrightarrow b), t}$ per ogni $p \in \mathcal{L}$, si ha naturalmente $P(p|_{\alpha, t'}) = P(p|_{\alpha \wedge (a \leftrightarrow b), t})$.

Se $P(a \leftrightarrow b|_{\alpha, t}) = 0$, la valutazione su $\mathcal{L}_{|\alpha \wedge (a \leftrightarrow b), t'}$ non è condizionata da quella effettuata su $\mathcal{L}_{|\alpha, t}$ (Teorema 13.2.3) e si è liberi perciò di scegliere le probabilità di $\mathcal{L}_{|\alpha \wedge (a \leftrightarrow b), t'}$ ex novo. Si dovrà comunque porre $P(a|_{\alpha, t'}) = P(b|_{\alpha, t'})$.

Se $P(a \leftrightarrow b|_{\alpha, t}) > 0$, l'aggiornamento della valutazione dovrà essere effettuato invece rispettando il teorema delle probabilità composte. Per ogni $p \in \mathcal{L}$ allora si ha $P(p|_{\alpha, t'}) = P(p|_{\alpha \wedge (a \leftrightarrow b), t}) = P((p \wedge a \leftrightarrow b)|_{\alpha, t}) / P(a \leftrightarrow b|_{\alpha, t})$. Da qui, ponendo $p|_{\alpha, t} = E_t$, $a|_{\alpha, t} = A_t$, $b|_{\alpha, t} = B_t$, $a \leftrightarrow b|_{\alpha, t} = h|_{\alpha, t} = H_t$ e tenendo conto che $H_t = (A_t \leftrightarrow B_t) = (A_t \wedge B_t) \vee (\bar{A}_t \wedge \bar{B}_t)$ (3.2.2 Definizione, 3.2.5 Proposizione, b), allora si ha:

$$P(p|_{\alpha, t'}) = P(E_t | H_t) = \frac{P(E_t \wedge H_t)}{P(H_t)} = \frac{P(E_t \wedge A_t \wedge B_t) + P(E_t \wedge \bar{A}_t \wedge \bar{B}_t)}{P(A_t \wedge B_t) + P(\bar{A}_t \wedge \bar{B}_t)},$$

Quella ora determinata è dunque la probabilità che il nostro soggetto, per rispettare il teorema delle probabilità composte – ed essere coerente con la valutazione espressa in t –, deve attribuire "a priori" (all'epoca t e in ogni epoca successiva prima di t') all'evento definito da p *condizionatamente all'ipotesi* che la proposizione $a \leftrightarrow b$ sia vera. Essa è anche la probabilità che egli attribuirà "a posteriori" (all'epoca t' , apprendendo che $a \leftrightarrow b$ è vera) all'evento $p|_{\alpha, t'}$. In particolare per l'evento $a|_{\alpha, t'} = b|_{\alpha, t'}$ si trova:

$$P(a|_{\alpha, t'}) = P(A_t | H_t) = \frac{P(A_t \wedge B_t)}{P(A_t \wedge B_t) + P(\bar{A}_t \wedge \bar{B}_t)} = P(B_t | H_t) = P(b|_{\alpha, t'}).$$

ESEMPIO 2.

Supponiamo che venga dedotto tardivamente che l'evento $A \rightarrow B$ è vero sub α . Indichiamo anche qui con t' l'epoca della deduzione dell'asserzione $|\bar{\alpha} (A \rightarrow B)$, con t un'epoca anteriore, con E_t il generico evento all'epoca t e con $E_{t'}$ il suo corrispondente all'epoca t' , ovvero $E_{t'} = E_t | (A \rightarrow B)_t$. Anche in questo caso, se $P((A \rightarrow B)_t) = 0$ si è liberi di scegliere le probabilità degli eventi dell'ambiente $\mathcal{L}|_{\alpha \wedge (a \rightarrow b), t'}$, e questa sarà sempre la conclusione se l'evento che descrive la deduzione tardiva è giudicato di probabilità nulla all'epoca t . Nel nostro caso particolare si dovrà comunque valutare $P(A_{t'}) \leq P(B_{t'})$. Se $P((A \rightarrow B)_t) = P(\neg(A_t \wedge \bar{B}_t)) = P(\bar{A}_t \vee B_t) > 0$ (3.2.2 *Definizione*, 3.2.5 *Proposizione*, a), in virtù del teorema delle probabilità composte si ottiene:

$$P(E_{t'}) = P(E_t | \neg(A_t \wedge \bar{B}_t)) = \frac{P(E_t \wedge (\bar{A}_t \vee B_t))}{1 - P(A_t \wedge \bar{B}_t)}.$$

In particolare, per gli eventi $A_{t'}$ e $B_{t'}$ si ottiene:

$$P(A_{t'}) = \frac{P(A_t \wedge B_t)}{1 - P(A_t \wedge \bar{B}_t)}, \quad P(B_{t'}) = \frac{P(B_t)}{1 - P(A_t \wedge \bar{B}_t)}$$

e risulta perciò $P(A_{t'}) \leq P(B_{t'})$, come doveva essere.

ESEMPIO 3. *Deduzione tardiva dell'eshaustività e dell'incompatibilità.*

Sia \mathcal{E} un insieme di eventi. Dedurre tardivamente che \mathcal{E} è un insieme di eventi esaustivo significa considerare in un certo istante t l'evento $\bigvee \mathcal{E}_t$ possibile. Ci limitiamo ad osservare che se all'epoca $t' > t$ si deduce che gli eventi di \mathcal{E} sono esaustivi, si avrà al solito che $E_{t'} = E_t | \bigvee \mathcal{E}_t$ e in particolare $\bigvee \mathcal{E}_{t'} = \bigvee \mathcal{E}_t | \bigvee \mathcal{E}_t$ e quindi $P(\bigvee \mathcal{E}_{t'}) = 1$, sia quando $P(\bigvee \mathcal{E}_t) = 0$ sia quando $P(\bigvee \mathcal{E}_t) > 0$.

Dedurre tardivamente che gli eventi di \mathcal{E} sono incompatibili significa invece considerare in un certo istante t l'evento $\bigwedge \mathcal{E}_t$ possibile. Anche qui ci limitiamo ad osservare che se all'epoca $t' > t$ si deduce che gli eventi di \mathcal{E} sono incompatibili,

si avrà $E_{t'} = E_t | (\neg \wedge \mathcal{E}_t)$ e in particolare $\wedge \mathcal{E}_{t'} = \wedge \mathcal{E}_t | (\neg \wedge \mathcal{E}_t) = \phi_t | (\neg \wedge \mathcal{E}_t)$ e quindi $P(\wedge \mathcal{E}_{t'}) = 0$, sia quando $P(\wedge \mathcal{E}_t) = 0$ sia quando $P(\wedge \mathcal{E}_t) > 0$.

Le considerazioni svolte provano dunque che dopo una prima analisi sugli aspetti descrittivi dell'incertezza si può passare a una fase di valutazione senza preoccuparsi di dedurre tutto il deducibile, sapendo di avere la possibilità di aggiustare la valutazione con coerenza tutte le volte che ci si accorge di non aver dedotto qualche cosa di deducibile. Il fatto di *poter essere coerenti* in casi come questo è un aspetto di grande importanza dal punto di vista teorico. Lo abbiamo già anticipato nella parte finale del *Commento 9.5.3*, ove però abbiamo osservato che anche se i difetti di deduzione non impediscono di continuare nella valutazione in modo coerente, in taluni casi può essere preferibile per motivi psicologici riprendere in esame il problema e procedere all'aggiustamento della valutazione senza porsi l'obbligo di rispettare il teorema delle probabilità composte. Ma anche senza ignorarlo. Le valutazioni coerenti a seguito di deduzione tardiva, in quanto collegate a misure del grado di fiducia frutto di analisi precedenti possono rappresentare comunque una utile indicazione orientativa. Si badi però che l'utilizzo del teorema delle probabilità composte presuppone che l'evento oggetto di deduzione tardiva sia stato valutato all'inizio. In difetto di ciò, o si rinuncia a usare il teorema delle probabilità composte o ci si sforza di dare al momento una valutazione della probabilità di tale evento cercando di rivivere le condizioni psicologiche e di conoscenza iniziali. Cosa che peraltro generalmente non sarà facile da realizzare in modo sufficientemente fedele.