Indipendenza stocastica senza paradossi

Lucio Crisma – Università di Trieste

Sommario. La nozione di indipendenza stocastica ha un ruolo centrale in ogni teoria delle probabilità. Quella di n eventi sta alla base di ogni sua estensione. In termini intuitivi, n eventi sono stocasticamente indipendenti se per la valutazione della probabilità di ciascuno di essi non importa conoscere il valore logico degli altri. La condizione si traduce in uguaglianze tra probabilità assolute e condizionate, per soddisfare le quali è necessario che la probabilità si fattorizzi su ogni prodotto logico di parte degli n eventi. Nella teoria classica (e anche da de Finetti) questa fattorizzazione è assunta come definizione di indipendenza stocastica. Essa ne coglie il significato intuitivo se gli eventi sono di probabilità strettamente compresa tra 0 e 1; altrimenti è fonte di situazioni paradossali – di indipendenza stocastica in presenza di dipendenza logica – che pongono la definizione ben lontana dal cogliere il significato intuitivo della nozione. Lo nota lo stesso de Finetti, aggiungendo però che "sembra difficile migliorarla senza uscire dalle possibilità realistiche". Gli argomenti svolti in [2] – è riportati qui in sintesi a partire dal n° 3 – mostrano però che trattando il problema in termini di probabilità condizionate coerenti nel modo più generale è possibile dare appieno consistenza teorica al significato intuitivo della nozione. Lo studio mette in evidenza che l'indipendenza logica degli eventi che vengono confrontati (in ogni caso a gruppi finiti) è prerequisito necessario (com'era da attendersi) e anche sufficiente (cosa meno scontata ancorché auspicabile) per garantire l'esistenza di probabilità in grado di realizzare l'indipendenza stocastica. La fattorizzazione – di fondamentale importanza nelle applicazioni e scelta come definizione nella teoria classica – è in questo contesto condizione necessaria per l'indipendenza stocastica, ma non sufficiente. Differentemente dal caso classico, in questa impostazione non tutte le probabilità che si fattorizzano sono compatibili con l'indipendenza stocastica. Ed è proprio grazie a ciò che non si corre il rischio di andare incontro a situazioni paradossali.

1 Indipendenza stocastica di due eventi.

La nozione di «indipendenza stocastica di un evento E da una informazione H (evento non impossibile)» si spiega usualmente in modo intuitivo dicendo che la probabilità di E non cambia se si apprende che «H è vero» o, in una versione più aderente al significato intuitivo, se si apprende «il valore logico di H». Tradotto in termini formali si ha dunque che «l'evento E è stocasticamente indipendente dall'evento H» se «$P(E|H) = P(E)$» nella prima versione (debole), se «$P(E|H) = P(E|\bar{H}) = P(E)$» nella seconda (forte). In entrambe le versioni l'indipendenza stocastica di E da H implica che la probabilità si fattorizzi sulla partizione generata dai due eventi $\{E \wedge H, \bar{E} \wedge H, E \wedge \bar{H}, \bar{E} \wedge \bar{H}\}$. Infatti, la condizione $P(E|H) = P(E)$, comune alla due versioni, implica $P(E \wedge H) = P(H)\,P(E|H) = P(E)\,P(H)$ ed è noto (e facile da provare) che se la probabilità di un costituente della partizione generata da due eventi si fattorizza, allora si fattorizzano anche le probabilità degli altri tre.

Due eventi E_1, E_2 sono poi considerati *stocasticamente indipendenti* (in termini intuitivi) se e solo se *ciascuno dei due è stocasticamente indipendente dall'altro*. Per quanto appena visto, la fattorizzazione della probabilità sulla partizione generata da E_1, E_2 (basta su $E_1 \wedge E_2$) è una condizione necessaria per l'indipendenza stocastica dei due eventi. Ricordando che in generale riesce $P(E|H) = P(E \wedge H)\,/\,P(H)$ se $P(H) > 0$ (definizione classica di probabilità condizionata), si trova poi che la condizione è anche sufficiente in senso debole se E_1, E_2 hanno probabilità positiva, in senso forte se non hanno probabilità estrema (se $E_1, \bar{E}_1, E_2, \bar{E}_2$ sono tutti di probabilità positiva). Si ha dunque che se E_1, E_2 non sono di probabilità estrema, la fattorizzazione della probabilità di $E_1 \wedge E_2$ interpreta appieno la nozione intuitiva di indipendenza stocastica di due eventi nelle sua forma forte (e anche in quella debole ovviamente). Altrimenti – se uno almeno degli eventi ha probabilità estrema –, la probabilità si fattorizza sulla loro partizione generata, ma le condizioni intuitive di indipendenza stocastica sono deducibili solo in parte. Infatti, se $P(E_2) = 1$ riesce $P(E_1 \wedge \bar{E}_2) \leq P(\bar{E}_2) = 0 = P(E_1)\,P(\bar{E}_2)$, quindi anche $P(E_1 \wedge E_2) = P(E_1)\,P(E_2)$; si trova allora $P(E_1|E_2) = P(E_2 \wedge E_1)\,/\,P(E_2) = P(E_1)$, mentre $P(E_1|\bar{E}_2)$ è indeterminata $(0/0)$. Se $P(E_2) = 0$ si trova invece che è indeterminata $P(E_1|E_2)$ e che riesce $P(E_1|\bar{E}_2) = P(E_1)$. A conclusioni analoghe si perviene poi, per simmetria, nei riguardi delle probabilità $P(E_2|E_1)$ e $P(E_2|\bar{E}_1)$ quando è estrema la probabilità di E_1.

In conclusione: la fattorizzazione della probabilità di $E_1 \wedge E_2$ (e di $E_1 \wedge \bar{E}_2$, $\bar{E}_1 \wedge E_2$, $\bar{E}_1 \wedge \bar{E}_2$) è una condizione necessaria per l'indipendenza stocastica di E_1 e E_2 (intesa in senso intuitivo) e "quasi sufficiente". Nella teoria classica essa è assunta di fatto anche come "sufficiente", perché si dà la seguente definizione.

Gli eventi E_1, E_2 *sono stocasticamente indipendenti se e solo se* $P(E_1 \wedge E_2) = P(E_1)\,P(E_2)$,

Abbiamo visto che se uno (almeno) di due eventi ha probabilità estrema la probabilità si fattorizza. Dando questa definizione, gli eventi sono allora stocasticamente indipendenti. Con riferimento al significato intuitivo della nozione, ci si comporta come se i casi $0/0$ fossero *determinati* e *di valore* $P(E_1)$ $(P(E_2))$ se lo 0 a denominatore è $P(E_2)$ o $P(\bar{E}_2)$ $(P(E_1)$ o $P(\bar{E}_1))$. La scelta non è però priva di conseguenze. Nei casi di probabilità estrema essa può condurre infatti a situazioni paradossali, alcune delle quali sono illustrate nel prossimo numero.

2 Indipendenza stocastica per fattorizzazione. Esempi paradossali.

Esempio 1.

Ogni evento di probabilità nulla è stocasticamente indipendente da se stesso in senso classico, ma non in quello intuitivo.

Se l'evento E è di probabilità nulla riesce $P(E \wedge E) = P(E)P(E)$, mentre $E|E$ è certo e quindi $P(E|E) = 1 \neq 0 = P(E)$.

Esempio 2.

Ogni evento di probabilità estrema e la sua negazione sono stocasticamente indipendenti in senso classico, ma non in quello intuitivo della versione forte.

Se E è di probabilità estrema riesce $P(E \wedge \bar{E}) = P(E)P(\bar{E})$, laddove E e $\bar{E}$ non possono essere stocasticamente indipendenti nel senso intuitivo della versione forte. Basta osservare per questo che $\bar{E}$ non può essere stocasticamente indipendente da E, perché $\bar{E}|E$ è impossibile, $\bar{E}\,|\,\bar{E}$ certo e allora $P(\bar{E}|E) = 0 \neq 1 = P(\bar{E}|\bar{E})$.

Esempio 3.

Un'urna sferica dotata di oblò con apertura regolabile contiene una pallina bianca (più grande) e una nera. Un ago è libero di ruotare attorno a un asse ortogonale (come quello di una bussola). Si eseguano due estrazioni con rimessa di una pallina per volta. La prima con apertura dell'oblò di diametro maggiore di quello delle palline; la seconda con la medesima apertura se «nella prima estrazione esce pallina bianca e, fatto ruotare l'ago, questo si ferma in direzione ovest» (ipotesi H), con apertura di diametro intermedio a quello delle palline altrimenti (ipotesi $\bar{H}$).

Supponiamo che sia nulla la probabilità che l'ago si fermi in una direzione assegnata (un solo caso favorevole su una infinità continua) e che le palline nell'urna (nel nostro caso due) di diametro minore di quello dell'oblò abbiano probabilità positiva di essere estratte. Posto E_h = *esce pallina bianca al colpo h,* $h = 1, 2,$ allora si ha: $P(H) = 0$, $0 < P(E_1) < 1, P(E_2) = P(H)P(E_2|H) + P(\bar{H})P(E_2|\bar{H}) = 0 \cdot P(E_2|H) + 1 \cdot 0 = 0$. Segue $P(E_1 \wedge E_2) = P(E_1)P(E_2)$ e i due eventi sono perciò stocasticamente indipendenti in senso classico. Non in senso intuitivo, però, perché $E_1|E_2$ è certo e riesce perciò $P(E_1|E_2) = 1 > P(E_1)$.

ESEMPIO 4.

Siano X, Y numeri aleatori reali logicamente indipendenti ($X=x \wedge Y=y$ *possibile per ogni* $x, y \in \mathbb{R}$*) e stocasticamente indipendenti in senso classico, con funzioni di ripartizione marginali continue. Sia inoltre Z il numero aleatorio coincidente con X (con Y) se la sua determinazione è irrazionale (razionale).*
Allora, sono stocasticamente indipendenti in senso classico anche i numeri aleatori Z, Y *e quindi in particolare la coppie di eventi* $Z=z$ *e* $Y=y$ *per ogni* $z, y \in \mathbb{R}$ $(P(Z=z \wedge Y=y) = 0 = P(Z=z)P(Y=y))$. *Non sono però stocasticamente indipendenti in senso intuitivo gli eventi* $Z=r$ *e* $Y=r$ *se r è razionale.*

Siano $F_X(x) = P(X \leq x)$ e $F_Y(y) = P(Y \leq y)$, $x, y \in \mathbb{R}$, le funzioni di ripartizione di X e Y. Allora, per definizione classica di indipendenza stocastica, $F(x, y) = P(X \leq x \wedge Y \leq y) = F_X(x)F_Y(y)$ è la loro funzione di ripartizione congiunta. Le funzioni $F_X(\cdot)$ e $F_Y(\cdot)$, continue, definiscono probabilità numerabilmente additive (σ-additive) e nulle sulle rispettive partizioni $\{X=x : x \in \mathbb{R}\}$, $\{Y=y : y \in \mathbb{R}\}$. È allora nulla la probabilità delle somme (unioni) numerabili di loro eventi, in particolare quella degli eventi *X è razionale* e *Y è razionale*. Pertanto si ha $F_Z(z) = P(Z \leq z) = P((Z \leq z) \wedge (Z \text{ è } razionale)) + P((Z \leq z) \wedge (Z \text{ è } irrazionale)) = P((Y \leq z) \wedge (Y \text{ è } razionale)) + P((X \leq z) \wedge (X \text{ è } irrazionale)) = P((X \leq z) \wedge (X \text{ è } razionale)) + P((X \leq z) \wedge (X \text{ è } irrazionale)) = P(X \leq z) = F_X(z)$. Con passaggi analoghi si trova $P(Z \leq z \wedge Y \leq y) = P(X \leq z \wedge Y \leq y) = F_X(x)F_Y(y) = F_Z(z)F_Y(y)$. I numeri aleatori Z, Y sono allora stocasticamente indipendenti in senso classico. Sono perciò tali pure gli eventi $Z=r$ e $Y=r$ per r razionale, che in tal caso sono uguali e non sono perciò stocasticamente indipendenti in senso intuitivo (*Esempio* 1).

In tutti quattro gli esempi precedenti le situazioni paradossali riguardano coppie di eventi non logicamente indipendenti – sono uguali nel primo esempio, uno negazione dell'altro nel secondo, E_2 implica E_1 nel terzo, $(Z=r)=(Y=r)$ per ogni r razionale nel quarto – uno (almeno) di probabilità estrema. Condizione quest'ultima che implica la fattorizzazione della probabilità del loro prodotto (intersezione) e quindi la loro indipendenza stocastica in senso classico. Ciò in contrasto col significato intuitivo della nozione, perché i legami di natura logica impongono scelte di probabilità condizionate differenti da quelle richieste dalle condizioni intuitive di indipendenza stocastica. Nel caso dei numeri aleatori del quarto esempio, i paradossi si manifestano perché Z e Y sono uguali se razionali, e in tal caso dipendono perciò funzionalmente (è $Z=Y$). Si tratta però di dipendenza funzionale (quindi logica) *parziale* e "trascurabile", nel senso che Z e Y sono *logicamente indipendenti* con probabilità 1 (quando sono irrazionali). Il prossimo esempio è dovuto a Bruno de Finetti e riguarda un caso di due numeri aleatori stocasticamente indipendenti che sono funzione l'uno dell'altro *con certezza* (in condizioni perciò di massima dipendenza logica). L'esempio va però inquadrato nella teoria delle probabilità coerenti (in quella classica perde di significato).

È noto in proposito che de Finetti introduce la nozione di probabilità coerente interpretandola come quota di scommessa unitaria. La probabilità di E, $P(E)$, è la quota che dà diritto di ricevere 1 se E si verifica, 0 se non si verifica; $SP(E)$ è poi la quota che dà diritto di ricevere l'importo S (pagare S se negativo) se E si verifica, 0 altrimenti. Ogni assegnazione di quote di scommessa unitarie agli eventi di un insieme $\mathcal{E}$ determina un'applicazione di $\mathcal{E}$ in $\mathbb{R}$. Essa è una probabilità se in sua corrispondenza nessuna scommessa su un numero finito di eventi di $\mathcal{E}$ produce guadagno (aleatorio) positivo (*norma di coerenza*). Si trova allora che le probabilità coerenti devono essere *normalizzate* (la probabilità dell'evento certo è 1), *non negative* e *additive.* L'additività completa (numerabile), invece, non è richiesta, ma ammissibile: le condizioni di coerenza e additività completa non sono cioè contraddittorie.

Circa indipendenza stocastica dei numeri aleatori, iniziando da quella di una coppia di numeri X, Y, de Finetti in [4] così si esprime (Cap. VI, 9.3):

> … «qualunque cosa si apprenda riguardo a *X*, l'opinione riguardo *Y* rimane inalterata» o più «tecnicamente» «ogni evento riguardante *Y* è stocasticamente indipendente da ogni evento riguardante *X*»
> ……………………………………………………………………………………
> E qui si presenta la solita questione: *quali eventi* riteniamo di comprendere in questa definizione? Si potrebbe dire «tutti» (… ma sappiamo che è un'astrazione pressoché inimmaginabile); si potrebbe (come i fautori della concezione «forte») dire «tutti quelli lebesguiani o almeno boreliani» (… ma ciò contrasta con le obiezioni contro l'additività completa …); si potrebbe limitarsi agli intervalli (e ciò che ne segue …). La questione non è però, si badi, di discutere e decidere quale risposta dia la definizione *vera*; la soluzione migliore sarebbe forse di considerare tutte tre le definizioni (o magari più) distinguendo «indipendenza completa», «forte», «debole»; ci limiteremo invece a far uso di quella debole perché e la sola che non richieda conoscenze eccessivamente non realistiche …
> Basta ammettere che siano indipendenti gli eventi del tipo $X \leq x$ con quelli del tipo $Y \leq y$ (per x e y qualunque) per concludere che è allora $F(x, y) = F_X(x)\, F_Y(y)$; …

ovvero che *X, Y* sono stocasticamente indipendenti.

La definizione di de Finetti coincide dunque *formalmente* con la definizione classica. Solo formalmente, però, perché la nozione di de Finetti va interpretata nell'ambito della teoria delle probabilità coerenti, in cui le classi delle funzioni di ripartizione sono più ampie di quelle della teoria classica. Sono inoltre generalmente più numerosi, per le funzioni di ripartizione in comune (quelle classiche), anche i rispettivi prolungamenti (probabilità) ammissibili. Nella teoria classica devono essere completamente (numerabilmente) additivi; basta siano semplicemente additivi in quella delle probabilità coerenti. Ai fini della lettura del prossimo esempio interessa infine segnalare che sono funzioni di ripartizione coerenti (ma non classiche) per numeri aleatori *numerabili* e coppie di numeri aleatori *numerabili* con determinazioni *dense* in $\mathbb{R}$ e $\mathbb{R}^2$, rispettivamente, nel primo caso *tutte* (e sole) le funzioni $F(\cdot)$ comprese tra 0 e 1, monotone non decrescenti (v. ad es. [1], *Teorema* 11.1.2), nel secondo *tutte* (e sole) le funzioni $F(\cdot,\cdot)$ comprese tra 0 e 1, monotone non decrescenti in entrambe le variabili e tali che sia $0 \leq F(x, y) - F(x, \eta) - F(\xi, y) + F(\xi, \eta) \leq 1$ per ogni $\xi < x$ e $\eta < y$, (v. ad es. [1], *Teorema* 11.2.2).

ESEMPIO 5 ([4], Cap. VI, 9.4).

Siano R, S numeri aleatori logicamente indipendenti con determinazioni i numeri razionali e $X = R + S\sqrt{2}$, $Y = R + S\sqrt{3}$. *I due numeri aleatori X, Y sono dipendenti funzionalmente (in corrispondenza biiettiva). È tuttavia corretto supporre X e Y stocasticamente indipendenti secondo la definizione di de Finetti.*

La coppia di equazioni $x = u + v\sqrt{2}$ e $y = u + v\sqrt{3}$ introduce una corrispondenza *biiettiva continua* tra i piani $\sigma(u, v)$ e $\sigma(x, y)$. A insiemi aperti di un piano corrispondono allora insiemi aperti dell'altro e di conseguenza a insiemi densi insiemi densi. In particolare, all'insieme delle determinazioni di (R, S) – delle coppie razionali – numerabile e denso in $\sigma(u, v)$, corrisponde l'insieme delle determinazioni della coppia (X, Y), $\{(r + s\sqrt{2}, r + s\sqrt{3}) : r, s \text{ razionali}\}$, numerabile e denso in $\sigma(x,y)$. Sono conseguentemente numerabili e dense in $\mathbb{R}$ le determinazioni sia di X sia di Y e gli eventi $(x_1 < X \leq x_2) \wedge (y_1 < Y \leq y_2)$ sono possibili per ogni $x_1 < x_2$ e $y_1 < y_2$. Si prova allora – usando le condizioni che caratterizzano le funzioni di ripartizione coerenti riportate sopra (nella premessa all'esempio) – che in queste ipotesi è coerente scegliere le funzioni di ripartizione marginali separatamente in modo arbitrario e prolungarle per fattorizzazione su $\{X \leq x \wedge Y \leq y : (x,y) \in \mathbb{R}^2\}$, facendo così in modo che i due numeri aleatori siano stocasticamente indipendenti secondo la definizione di de Finetti.

Come enunciato, i due numeri sono però funzionalmente dipendenti. Infatti, se x è una determinazione di X, la coppia di razionali che la determina è unica, perché se $x = r + s\sqrt{2} = r' + s'\sqrt{2}$, non può essere $s \neq s'$, poiché allora si otterrebbe che $\sqrt{2} = (r - r')/(s - s')$ è razionale (assurdo!), mentre se $s = s'$ si ha banalmente anche $r = r'$. Allora, fissata una determinazione x di X, essa determina una coppia di razionali (r, s) e questa determina $y = r + s\sqrt{3}$.

Il numero aleatorio X è dunque funzione di Y (e viceversa) ed è tuttavia coerente supporre X e Y stocasticamente indipendenti. Quanto sia paradossale questa situazione è messo in evidenza scegliendo le marginali continue, com'è coerente fare perché X e Y sono densi in $\mathbb{R}$. In tal caso riesce $P(X{=}x) = P(Y{=}y) = 0$ per ogni x, y. Poiché X è funzione di Y si ha poi $(X = \xi) = (Y = \eta)$ per ogni determinazione (ξ, η) della coppia aleatoria (X, Y). Segue $X = \xi | Y = \eta$ certo, quindi $P(X = \xi | Y = \eta) = 1 > 0 = P(X = \xi)$ e analogamente $P(Y = \eta | X = \xi) > P(Y = \eta)$.

De Finetti porta questo esempio dopo aver definito X e Y stocasticamente indipendenti quando la loro funzione di ripartizione congiunta è prodotto delle rispettive funzioni di ripartizione marginali e dopo aver osservato:

> Si noti però quanto poco tale definizione risponda al concetto intuitivo di indipendenza stocastica ...

Fa poi seguire l'esempio e la seguente conclusione:

> Si può avere dipendenza logica (addirittura biunivoca) e insieme indipendenza stocastica (ripartizionale); bisogna tener presente quanto c'è di insoddisfacente in questa definizione dal punto di vista logico, anche se sembra difficile migliorarla senza uscire dall'ambito delle possibilità realistiche

Ma è proprio uscendo dall'ambito delle possibilità realistiche che si riesce invece a dare una definizione di indipendenza stocastica che interpreta appieno la nozione intuitiva e che consente anche di spiegare l'esempio in modo non paradossale. L'argomento è sviluppato in breve nei prossimi numeri nell'ambito della teoria delle probabilità condizionate coerenti. Ricordiamo qui in proposito che al pari delle probabilità assolute anche le condizionate possono essere interpretate come quote di scommessa unitaria. Con l'avvertenza però che una scommessa su un evento $E|H$ è valida solo se si verifica l'ipotesi H; altrimenti la puntata viene rimborsata. In dettaglio, $P(E|H)$ ($SP(E|H)$) è la quota che dà diritto di ricevere 1 (S) se si verifica $E \wedge H$, 0 se si verifica $\bar{E} \wedge H$, $P(E|H)$ ($SP(E|H)$) se si verifica $\bar{H}$. Questa definizione è sufficiente per dimostrare il teorema delle probabilità composte e stabilire così il collegamento tra probabilità coerenti condizionate a ipotesi diverse, una più specifica (implicante) l'altra. In primo luogo, il collegamento tra probabilità assolute e condizionate a una ipotesi. Il teorema è fondamentale (anche nella teoria classica) ed è in larga misura sufficiente nelle applicazioni per trattare problemi di probabilità in cui intervengono eventi condizionati. Forse per questo de Finetti non dà un definizione di probabilità condizionata in termini globali, con riferimento cioè a insiemi di eventi condizionati qualunque, come ha fatto nel caso assoluto. Una nozione siffatta è però essenziale per sviluppare in modo rigoroso una teoria delle probabilità condizionate. Risponde allo scopo la seguente definizione, dovuta a Silvano Holzer [5], che estende nel modo più naturale quella di de Finetti.

Con riferimento a un insieme di eventi condizionati $\mathcal{D}$, consideriamo le applicazioni $P(\cdot|\cdot)$ di $\mathcal{D}$ in $\mathbb{R}$ e, per ogni $E|H \in \mathcal{D}$, attribuiamo a $P(E|H)$ il significato di quota di scommessa unitaria. Come nel caso assoluto consideriamo solo scommesse su un numero finito di eventi. Diciamo poi *valida* una scommessa – i cui guadagni vengano calcolati usando una medesima $P(\cdot|\cdot)$ – se non si conclude con la restituzione di tutte le puntate (se risulta vera perciò almeno una delle ipotesi degli eventi oggetto di scommessa). Ciò posto, le probabilità (condizionate) coerenti su $\mathcal{D}$ sono le applicazioni $P(\cdot|\cdot)$, in cui corrispondenza nessuna scommessa valida produce guadagno positivo (*norma di coerenza condizionata*).

3 Indipendenza stocastica di *n* eventi.

Nei numeri precedenti abbiamo discusso della nozione di indipendenza stocastica nel caso più semplice di *due* eventi e della sua estensione a *due* numeri aleatori (a insiemi di eventi che parlano delle loro determinazioni). La nozione interessa però in generale con riferimento a più *eventi* e a più *insiemi di eventi,* di numerosità *qualsiasi.* Tutte le definizioni che vengono date in proposito in letteratura sono fondate sul concetto di «indipendenza stocastica di un *numero finito* di eventi». È allora questo il concetto che va studiato per primo. Nella teoria classica gli eventi di un insieme finito sono stocasticamente indipendenti se la probabilità del prodotto di qualsiasi parte di essi si fattorizza. Sappiamo che già nel caso di due eventi questa definizione non è in grado di interpretare appieno il significato *intuitivo* del concetto, che per più eventi è quello di indipendenza stocastica di *ciascuno* dai *rimanenti.* Frase che non ha senso, questa, se non si spiega prima cosa significa «indipendenza stocastica di un evento da altri (in numero finito)», come si è fatto nel n° 1 per «E stocasticamente indipendente da H». Si è detto allora che ciò accade quando la probabilità di E non cambia se si apprende il valore logico di H e lo si è tradotto in termini formali nelle uguaglianze $P(E|H) = P(E|\bar{H}) = P(E)$. Analogamente, «l'evento E è stocasticamente indipendente dagli eventi $E_1, \dots, E_n$» – in termini intuitivi – quando la probabilità di E non cambia se si apprendono i valori logici di $E_1, \dots, E_n$. Indicato con E' il generico evento E o la sua negazione $\bar{E}$, l'insieme dei 2^n prodotti $E'_1 \wedge \dots \wedge E'_n$ è la partizione generata da $E_1, \ldots, E_n$, che indicheremo con $\mathbb{P}_G(E_1,\ldots,E_n)$ quando non interessa escludere gli eventuali prodotti (costituenti) impossibili, e con $\mathbb{P}_G^\phi(E_1,\ldots,E_n)$ altrimenti. Ciò premesso, conoscere il valore logico di $E_1, \ldots, E_n$ significa allora sapere quale costituente di $\mathbb{P}_G^\phi(E_1,\ldots,E_n)$ è vero. La traduzione formale della nozione di indipendenza stocastica di un evento da n eventi è data perciò dalla seguente definizione.

3.1 Definizione. E stocasticamente indipendente da $E_1, \ldots, E_n$.

Siano $E, E_1, \ldots, E_n$ eventi. Diremo che E è ***stocasticamente indipendente*** *da $E_1, \ldots, E_n$ se e solo $P(E|\omega) = P(E)$ per ogni $\omega \in \mathbb{P}_G^\phi(E_1,\ldots,E_n)$. Ovvero, se e solo se la probabilità su $\{E\} \cup \{E|\omega : \omega \in \mathbb{P}_G^\phi(E_1,\ldots,E_n)\}$ è costante.*

Gli eventi certo e impossibile sono stocasticamente indipendenti da ogni insieme di eventi. Essi conservano infatti il loro valore (rimangono certo e impossibile) a seguito di ogni incremento d'informazione e conservano perciò anche le rispettive probabilità. Riesce cioè $P(\Omega|H) = P(\Omega) = 1$ e $P(\phi|H) = P(\phi) = 0$ per ogni $H \neq \phi$.

Se E è possibile, si possono presentare casi diversi, dipendenti dal suo grado di dipendenza logica da $E_1, \ldots, E_n$, Osserviamo prima in proposito che se si apprende che $\omega \in \mathbb{P}_G^\phi(E_1,\ldots,E_n)$ è vero – e si conoscono quindi i valori logici di $E_1, \ldots, E_n$ – il valore logico di E è *deducibile* se ω implica E o $\bar{E}$ (E è vero nel primo caso, falso nel secondo), *non deducibile* altrimenti. Pertanto, E è *logicamente dipendente* da $E_1,\ldots,E_n$ (di valore deducibile da quelli di $E_1, \ldots, E_n$) se il suo valore è deducibile per ogni ω, *logicamente indipendente* se non è deducibile per alcun ω, *logicamente semidipendente* se è deducibile per qualche ω (non per tutti). Inoltre, la semidipendenza può essere *unilaterale* (esistono ω che implicano o solo E o solo $\bar{E}$) o *bilaterale* (esistono sia ω che implicano E sia ω che implicano $\bar{E}$).

Circa la possibilità di considerare un evento E *possibile* stocasticamente indipendente da $E_1, \ldots, E_n$ si hanno allora i seguenti casi.

(i) *E è logicamente dipendente o semidipendente bilateralmente da $E_1,\ldots,E_n$.* Esiste cioè ω_1 che implica E e ω_2 che implica $\bar{E}$. Allora E non può essere stocasticamente indipendente da $E_1, \ldots, E_n$, perché $E|\omega_1$ è certo e $E|\omega_2$ impossibile, quindi $P(E|\omega_1) \neq P(E|\omega_2)$ e $P(E|\cdot)$ non è perciò costante su $\mathbb{P}_G^\phi(E_1,\ldots,E_n)$.

È il caso degli *Esempi* 1 e 2 del n° 2, in cui E è logicamente dipendente da se stesso e rispettivamente dalla sua negazione.

(ii) *E è logicamente semidipendente unilateralmente da $E_1, \dots, E_n$*. Allora E può essere stocasticamente indipendente da $E_1, \dots, E_n$, ma solo se la sua probabilità è estrema: 1 se qualche ω implica E, 0 se qualche ω implica $\bar{E}$ ([2], 15.2.4*c*).
È il caso dell'*Esempio* 3 del n° 2, in cui E_1 e semidipendente unilateralmente da E_2, perché E_2 implica E_1 e $\bar{E}_2$ è compatibile sia con E_1 sia con $\bar{E}_1$. L'evento E_1 diventerebbe stocasticamente indipendente da E_2 se e solo se fosse $P(E_1) = 1$ (nell'esempio è $P(E_1) < 1$).

(iii) *E è logicamente indipendente da $E_1, \dots, E_n$*. Allora si può attribuire all'evento E probabilità arbitraria e supporlo stocasticamente indipendente da $E_1, \dots, E_n$: tutte le costanti tra 0 e 1 su $\{E\} \cup \mathbb{P}_G^{\phi}(E_1, \dots, E_n)$ – estremi inclusi – sono cioè probabilità coerenti) ([2], 15.2.4*d*).

Torniamo a riflettere a questo punto sul significato intuitivo di indipendenza stocastica in ottica soggettiva (la probabilità è una valutazione di un soggetto). Chi *suppone* «E stocasticamente indipendente da $E_1, \dots, E_n$» *giudica* «irrilevanti ai fini della valutazione della probabilità di E le informazioni sui valori logici di $E_1, \dots, E_n$». Egli effettua così di fatto una valutazione di probabilità, perché decide che la probabilità su $\{E|\omega : \omega \in \mathbb{P}_G^{\phi}(E_1, \dots, E_n)\}$ deve essere costante e uguale a quella di E. Quest'ultima va valutata – come recita il giudizio (non importa se prima o dopo averlo espresso) – senza tener conto di eventuali informazioni sui valori di $E_1, \dots, E_n$. Ad esempio, nel caso di una sequenza di estrazioni *con rimessa* da un'urna di composizione nota contenente palline bianche e nere, il giudizio di indipendenza stocastica del risultato di una estrazione da quelli di altre è dovuto al fatto che, in virtù della modalità d'estrazione, si sa che la composizione dell'urna è la stessa (e nota) in tutte le estrazioni. Al fine del giudizio non interessa conoscere l'esatta composizione. Cosa invece indispensabile per valutare la probabilità degli eventi *esce pallina bianca* (o *nera*) nelle singole estrazioni. Analogamente, in una partita a testa e croce l'ipotesi di indipendenza stocastica del risultato di un lancio da quelli di altri lanci esprime un giudizio sulla modalità dei lanci (si giudica che vengano effettuati tutti nelle stesse condizioni), mentre la probabilità di *testa* (*croce*) dipende esclusivamente dal grado di regolarità della moneta.

In conclusione: chi giudica E stocasticamente indipendente da $E_1, \ldots, E_n$ decide di attribuire agli eventi $E|\omega$, $\omega \in \mathbb{P}_G^{\phi}(E_1, \ldots, E_n)$, probabilità uguale a quella di E *qualunque essa sia* (purché coerente). Il caso (iii) è l'unico in grado di dare alla nozione questo significato. Si ha allora che *l'indipendenza logica di E da* $E_1, \ldots, E_n$ è *condizione necessaria* (abbiamo visto anche sufficiente) per dare consistenza alla *nozione intuitiva* di *indipendenza stocastica di E da* $E_1, \ldots, E_n$.

Poiché la nozione di «indipendenza stocastica di un evento da altri» sta alla base di tutte le altre nozioni di indipendenza stocastica – sono tutte espresse da *insiemi* di *condizioni base* di quel tipo –, si ha allora che *l'indipendenza logica* degli eventi delle condizioni base dai rispettivi eventi rimanenti è una *condizione necessaria* per l'introduzione di nozioni di indipendenza stocastica conformi al loro significato intuitivo. Si potrebbe allora pensare di mettere nelle definizioni l'indipendenza logica come prerequisito all'indipendenza stocastica. È però preferibile non inserire nelle definizioni alcuna condizione di natura logica, perché le condizioni base richieste nelle medesime sono di natura probabilistica (sono uguaglianze di probabilità). Ed è bene quindi caratterizzare prima la classe delle probabilità che soddisfano le condizioni base – usando per questo esclusivamente proprietà della probabilità – e individuare poi in essa le probabilità che soddisfano condizioni aggiuntive di indipendenza logica (necessariamente di tipo (iii)) che le rendono adatte a interpretare il significato intuitivo di indipendenza stocastica.

Ciò premesso, andiamo ad approfondire l'argomento per la nozione di indipendenza stocastica di n eventi, che in accordo con quanto detto in premessa di questo numero, è data dalla seguente definizione.

3.2 Definizione. $E_1, \ldots, E_n$ stocasticamente indipendenti.

Siano $E_1, \ldots, E_n$ *eventi. Diremo che* $E_1, \ldots, E_n$ *sono* ***stocasticamente indipendenti*** *se e solo se ciascuno di essi è stocasticamente indipendente dai rimanenti. Se riesce cioè* $P(E_h|\omega) = P(E_h)$ *per ogni* $\omega \in \mathbb{P}_G^{\phi}(\{E_1, \ldots, E_n\} - \{E_h\})$, $h = 1, \ldots, n$.

Il seguente teorema dà consistenza alla definizione e al giudizio di indipendenza stocastica degli eventi $E_1, \ldots, E_n$.

3.3 Teorema ([2], 15.3.5 *Complemento*).

Se gli eventi $E_1, \dots, E_n$ sono logicamente indipendenti – i costituenti di $\mathbb{P}_G(E_1, \dots, E_n)$ sono tutti possibili e ogni evento è perciò logicamente indipendente dai rimanenti –, è coerente scegliere a piacere le loro probabilità (separatamene tra 0 e 1, estremi inclusi) e supporli stocasticamente indipendenti (è coerente cioè giudicare $E_1, \dots, E_n$ stocasticamente indipendenti).

La *Definizione* 3.2 è consistente anche in casi di dipendenza logica, come quelli dei prossimi due esempi.

Esempio 1.

Sia $E_1 \wedge \dots \wedge E_n$ l'unico costituente impossibile della partizione generata dagli eventi $E_1, \dots, E_n$.

Poiché gli eventi $\bar{E}_h \wedge \bigwedge_{i \neq h}^{1,n} E_i'$ sono possibili per ogni h, sono allora tali anche gli eventi $\bigwedge_{i \neq h}^{1,n} E_i'$. Ogni E_h risulta perciò logicamente semidipendente unilateralmente dagli altri $n-1$ eventi, perché $\bigwedge_{i \neq h}^{1,n} E_i$ implica $\bar{E}_h$ e ogni altro $\bigwedge_{i \neq h}^{1,n} E_i'$ è compatibile sia con E_h sia con $\bar{E}_h$. Le condizioni di semidipendenza unilaterale lasciano aperta la possibilità di considerare $E_1, \dots, E_n$ stocasticamente indipendenti. Poiché $E_h | \bigwedge_{i \neq h}^{1,n} E_i$ è impossibile, riesce $P(E_h | \bigwedge_{i \neq h}^{1,n} E_i) = 0$ e si deve allora porre a tal fine $P(E_h | \bigwedge_{i \neq h}^{1,n} E_i') = P(E_h) = 0$, $h = 1, \dots, n$.

Occorre vedere a questo punto se questa valutazione è coerente. Osserviamo intanto che le condizioni $P(E_1) = \dots P(E_n) = 0$ implicano $P(\bar{E}_1) = \dots = P(\bar{E}_n) = 1$, quindi $P(\bar{E}_1 \wedge \dots \wedge \bar{E}_n) = 1$, e che è coerente dare probabilità 1 a $\bar{E}_1 \wedge \dots \wedge \bar{E}_n$. Si deve dare in tal caso probabilità 0 agli altri costituenti di $\mathbb{P}_G(E_1, \dots, E_n)$. Restano così determinate le probabilità di $E_1, \dots, E_n$ – nulle come richiesto – ma anche quelle degli $n 2^n$ prodotti $\bigwedge_{i \neq h}^{1,n} E_i'$, per i quali riesce $P(\bigwedge_{i \neq h}^{1,n} \bar{E}_i) = 1$ e $P(\bigwedge_{i \neq h}^{1,n} E_i') = 0$ se $E_i' = E_i$ per almeno un i. Usando il teorema delle probabilità composte si ricava allora $P(E_h | \bigwedge_{i \neq h}^{1,n} \bar{E}_i) = P(E_h)\ (= 0)$ – come richiesto – mentre la probabilità di $E_h | \bigwedge_{i \neq h}^{1,n} E_i'$ non è ricavabile se qualche E_i' è affermato. Si prova però che è coerente porre $P(E_h | \bigwedge_{i \neq h}^{1,n} E_i') = 0 = (P(E_h))$ – ultime condizioni richieste – e con questa scelta (e solo con questa) gli eventi $E_1, \dots, E_n$ sono stocasticamente indipendenti ([2], 15.3.6 *Esempio* 2).

ESEMPIO 2.

Siano $E_1, \dots, E_n$ eventi possibili, E_1 e E_2 incompatibili, $E_1, E_3, \dots, E_n$ e $E_2, E_3, \dots, E_n$ logicamente indipendenti.

Sono $E_1 \wedge E_2 \wedge E_3 \wedge \dots \wedge E_n$ impossibile, $E_1 \wedge E_3 \wedge \dots \wedge E_n$ e $E_2 \wedge E_3 \wedge \dots \wedge E_n$ possibili, $E_1 \wedge E_2 \wedge \bigwedge_{i \neq h}^{3,n} E_i$ impossibili, $h = 3, \dots, n$. Ciò premesso, i due eventi $E_1 | E_2 \wedge E_3 \wedge \dots \wedge E_n$ e $E_2 | E_1 \wedge E_3 \wedge \dots \wedge E_n$ sono impossibili e affinché $E_1, \dots, E_n$ siano stocasticamente indipendenti bisogna allora porre $P(E_1) = P(E_1 | E_2 \wedge \bigwedge_{i \neq h}^{3,n} E_i') = P(E_2) = P(E_2 | E_1 \wedge \bigwedge_{i \neq h}^{3,n} E_i') = 0$. Si prova che scegliendo a piacere le probabilità degli eventi $E_3, \dots, E_n$ e imponendo per ciascuno di essi le condizioni di indipendenza stocastica – ovvero $P(E_h | \omega) = P(E_h)$, $\omega \in \mathbb{P}_G^{\phi}(\{E_1, \dots, E_n\} - \{E_h\})$, $h = 3, \dots, n$ – si ottiene una probabilità coerente ([2], 15.3.6 *Esempio* 3). In corrispondenza a ciascuna di tali probabilità gli eventi $E_1, \dots, E_n$ sono allora stocasticamente indipendenti.

Interessa ancora segnalare che sussistono la seguenti proposizioni.

3.4 Proposizione ([2], *Teorema* 15.3.2,*d*).

Se gli eventi $E_1, \dots, E_n$ sono stocasticamente indipendenti, allora sono tali anche gli eventi di ogni loro sottoinsieme.

3.5 Proposizione ([2], *Teorema* 15.3.3).

Se gli eventi $E_1, \dots, E_n$ sono stocasticamente indipendenti la probabilità si fattorizza su $\mathbb{P}_G(E_1, \dots, E_n)$. Per ogni $E_1' \wedge \dots \wedge E_n' \in \mathbb{P}_G(E_1, \dots, E_n)$ riesce cioè $P(E_1' \wedge \dots \wedge E_n') = P(E_1') \dots P(E_n')$.

La fattorizzazione della probabilità su $\mathbb{P}_G(E_1, \dots, E_n)$ è dunque condizione necessaria per l'indipendenza stocastica di $E_1, \dots, E_n$. In virtù della precedente *Proposizione* 3.4, essa è anche equivalente alla fattorizzazione della probabilità sull'insieme $\{E_{i_1} \wedge \dots \wedge E_{i_s} : \{E_{i_1}, \dots, E_{i_s}\} \subset \{1, \dots, n\}\}$ che, come già ricordato, è la definizione della nozione classica di indipendenza stocastica di n eventi. Sicché, *la nozione classica di indipendenza stocastica è condizione necessaria per l'indipendenza stocastica di $E_1, \dots, E_n$* secondo la *Definizione* 3.2. In altri termini, eventi stocasticamente indipendenti secondo la *Definizione* 3.2 sono stocasticamente indipendenti anche in senso classico. È facile verificare che vale pure il viceversa se gli eventi non sono di probabilità estrema; ma non altrimenti, come mostrano gli esempi del n° 2.

Queste proprietà di fattorizzazione sussistono per tutte le probabilità che soddisfano le condizioni base di indipendenza stocastica della *Definizione* 3.2, che sono le $P(E_h | \omega) = P(E_h)$, $\omega \in \mathbb{P}_G^\phi(\{E_1, \dots, E_n\} - \{E_h\})$, $h = 1, \dots, n$. Non importa se gli E_h sono *tutti* logicamente indipendenti dai rimanenti (caso del *Teorema* 3.3) o sono *in parte* logicamente semidipendenti unilateralmente (come nell'*Esempio* 2) o sono *solo* di quest'ultimo genere (come nell'*Esempio* 1).

4 Nozione generale di indipendenza stocastica.

Nella sua accezione più generale, la nozione di indipendenza stocastica riguarda *collezioni* di *insiemi di eventi*. Identificando gli eventi con i sottoinsiemi di una partizione (nozione classica), si deve intendere dato un insieme di partizioni e per ogni partizione un corrispondente insieme di eventi logicamente dipendenti (un insieme di suoi sottoinsiemi). Definendo gli eventi mediante proposizioni della logica (impostazione logica della descrizione dell'incertezza) gli insiemi di eventi possono essere introdotti (definiti) invece direttamente mediante un insieme di proposizioni. Si può tuttavia ricondurre la descrizione alla precedente, perché ogni insieme di eventi genera una partizione in grado di descriverli come somma di costituenti ([1], *Proposizione* 6.3.1).

Ciò premesso, gli insiemi di eventi $\mathcal{E}_1, \dots, \mathcal{E}_n$ sono considerati stocasticamente indipendenti – in senso intuitivo – se sono tali gli eventi $E_1, \dots, E_n$ per ogni scelta di $E_1 \in \mathcal{E}_1, \dots, E_n \in \mathcal{E}_n$. Se gli insiemi di eventi sono in numero infinito è usuale poi considerarli stocasticamente indipendenti se ciò accade per ogni loro parte finita. Nell'impostazione classica la nozione viene formalizzata con riferimento alla totalità degli eventi logicamente dipendenti dalle corrispondenti partizioni se queste sono finite o numerabili; con riferimento a una loro conveniente σ-algebra altrimenti. La definizione che segue ricalca l'enunciato di quella classica, dalla quale però si distingue perché le collezioni di insiemi di eventi sono qualsiasi (come vuole il significato intuitivo) e la nozione di indipendenza stocastica di n eventi è quella introdotta con la *Definizione* 3.2.

4.1 Definizione. Collezioni di insiemi di eventi stocasticamente indipendenti.

Sia $\mathscr{C}$ una collezione di insiemi di eventi. Diremo allora che gli insiemi di $\mathscr{C}$ sono ***stocasticamente indipendenti*** *se e solo se per ogni* $\{\mathscr{E}_1, \dots ,\mathscr{E}_n\} \subset \mathscr{C}$ *e* $E_1 \in \mathscr{E}_1, \dots , E_n \in \mathscr{E}_n$, *gli eventi* $E_1, \dots ,E_n$ *sono stocasticamente indipendenti.*

NOTA.

Se gli eventi $E_1 , \dots , E_n$ sono stocasticamente indipendenti secondo la *Definizione* 3.2, sono tali anche gli eventi di ogni loro sottoinsieme ([1], *Proposizione* 15.3.2*d*). La presente definizione estende allora la 3.2, com'è subito visto ponendo in essa $\mathscr{C} = \{\{E_1\} , \dots , \{E_n\}\}$.

Come accade ogni qualvolta si dà una definizione, si pone anche per questa il problema della sua consistenza e della sua capacità di interpretare il significato del concetto che si vuole introdurre. Qui quello intuitivo di indipendenza stocastica di una collezione di insiemi di eventi. Si tenga conto per questo che *giudicare* gli insiemi di eventi di un insieme $\mathscr{C}$ *stocasticamente indipendenti* significa *ritenere irrilevante* per la valutazione della probabilità su ciascuno di tali insiemi *avere informazioni* su valori logici di eventi appartenenti a una parte finita degli altri. Si tratta allora di vedere se la definizione è in grado di cogliere questo significato. Deve essere coerente, a tal fine, *assegnare* le probabilità marginali sugli insiemi di $\mathscr{C}$ *a piacere,* ciascuna in modo indipendente dalle altre, e *imporre* che siano soddisfatte le condizioni base di indipendenza stocastica previste nella definizione. Il prossimo *Teorema* 4.2 fornisce una condizione necessaria e sufficiente affinché sia soddisfatta la prima esigenza. Il *Teorema* 4.3 successivo mostra che la medesima condizione è sufficiente anche per la seconda. Per una lettura corretta di entrambi i teoremi bisogna aver presente: (1) che le algebre di un insieme $\mathscr{A}$ sono *logicamente indipendenti* se sono tali gli eventi degli insiemi ottenuti scegliendo un evento in ogni algebra; (2) che le algebre di un insieme $\mathscr{A}$ sono *finito-logicamente indipendenti* se sono tali le algebre di ogni suo sottoinsieme finito (il concetto è equivalente a quello del punto (1) se $\mathscr{A}$ è finito); (3) che *l'algebra generata* da un *insieme finito* di eventi è la totalità delle *somme* di costituenti della partizione da essi generata; (4) che *l'algebra generata* da un *insieme qualsiasi* di eventi è l'unione delle algebre generate dai suoi insiemi finiti; (5) che *assegnazione di probabilità marginali su* $\mathfrak{E}$ (collezione di insiemi di eventi) significa scelta di una probabilità coerente per ogni insieme di $\mathfrak{E}$, ciascuna effettuata in modo indipendente dalle altre.

Circa l'esistenza di probabilità in generale, ricordiamo che essa è assicurata per ogni insieme di eventi ed è conseguenza del teorema del prolungamento, dimostrato da de Finetti per le probabilità coerenti assolute [3] ed esteso alle probabilità condizionate da Holzer [5]. Il teorema afferma che se una probabilità è coerente su un insieme di eventi, può allora essere prolungata in modo coerente su ogni altro insieme. L'esistenza di probabilità coerenti su ogni insieme segue allora dall'esistenza di probabilità coerenti sui singleton degli insiemi.

4.2 Teorema ([2],15.5.5 *Teorema*).

Siano $\mathfrak{E}$ *una collezione di insiemi di eventi e* $\mathfrak{A}_G(\mathfrak{E}) = \{ \mathcal{A}_G(\mathcal{E}) : \mathcal{E} \in \mathfrak{E} \}$ *l'insieme delle algebre generate dagli insiemi di* $\mathfrak{E}$. *Allora, le assegnazioni di probabilità marginali su* $\mathfrak{E}$ *sono* <u>*tutte*</u> *coerenti (determinano una probabilità coerente su* $\cup\mathfrak{E}$*) se e solo se le algebre di* $\mathfrak{A}_G(\mathfrak{E})$ *sono finito-logicamente indipendenti).*

Ogni probabilità su $\cup\mathfrak{E}$ (unione degli insiemi di $\mathfrak{E}$) *determina* le probabilità marginali. Se le algebre generate dagli insiemi di $\mathfrak{E}$ sono finito-logicamente indipendenti vale anche il viceversa. In questa ipotesi infatti, gli *eventi possibili* di $\cup\mathfrak{E}$ appartengono a *uno solo* degli insiemi di $\mathfrak{E}$ e ogni assegnazione di probabilità marginali *determina* perciò un'applicazione di $\cup\mathfrak{E}$ in $(0\,,1)$. Il teorema afferma che essa è una probabilità coerente.

Sicché: se le algebre generate dagli insiemi di $\mathfrak{E}$ sono finito-logicamente indipendenti, le probabilità su $\cup\mathfrak{E}$ si ottengono tutte e sole assegnando a piacere le probabilità sugli insiemi di $\mathfrak{E}$. Il prossimo teorema ci dice che nella medesima ipotesi, ciascuna di quelle probabilità può essere prolungata in modo da soddisfare *l'ipotesi di indipendenza stocastica* per gli insiemi di $\mathfrak{E}$: ovvero, indicata con P una di esse, per ogni $\{\mathcal{E}_1,\dots,\mathcal{E}_n\}\subset\mathfrak{E}$, $E\in\mathcal{E}$, $E_1\in\mathcal{E}_1,\dots, E_n\in\mathcal{E}_n$, $\omega\in \mathbb{P}^{\phi}_{G}(E_1,\dots,E_n)$ é coerente porre $P(E|\,\omega)=P(E)$.

4.3 Teorema ([2], segue da 15.5.8 *Corollario*).

Se $\mathscr{G}$ è una collezione di insiemi di eventi, P una probabilità su $\cup\mathscr{G}$ e le algebre generate dagli insiemi di $\mathscr{G}$ sono finito-logicamente indipendenti, allora P è compatibile (coerente) con l'ipotesi di indipendenza stocastica degli insiemi di $\mathscr{G}$ (è coerente prolungare P imponendo le relative condizioni di indipendenza stocastica).

Unendo risultati dei due teoremi abbiamo:

(i) *la Definizione* 4.1 è *consistente:* esistono cioè probabilità che soddisfano le condizioni di indipendenza stocastica in essa richieste;

(ii) *la finito-logica indipendenza delle algebre generate dagli insiemi di un insieme di eventi è condizione necessaria (e anche sufficiente) per dare consistenza formale alla nozione intuitiva di indipendenza stocastica (intesa come giudizio di indipendenza stocastica di insiemi di eventi):* in quella ipotesi (e solo allora), infatti, si possono assegnare separatamente le probabilità marginali sugli insiemi e prolungare la valutazione imponendo le condizioni per la loro indipendenza stocastica.

Sussistono inoltre le due seguenti proposizioni:

(iii) *è coerente giudicare stocasticamente indipendenti le algebre di eventi di un insieme $\mathscr{A}$ se e solo se sono finito-logicamente indipendenti* (basta porre nei due teoremi $\mathscr{G} = \mathscr{A}$ e tener conto del punto (ii));

(iv) *se* $\{\mathscr{A}(\mathscr{E}) : \mathscr{E} \in \mathscr{G}\}$ *è un insieme di algebre di eventi finito-logicamente indipendenti che "includono" gli insiemi di* $\mathscr{G}$ ($\mathscr{E} \subset \mathscr{A}(\mathscr{E})$ *per ogni* $\mathscr{E} \in \mathscr{G}$), *è coerente giudicare gli insiemi di* $\mathscr{G}$ *stocasticamente indipendenti* (basta ricordare che l'algebra generata da un insieme è la minima algebra che lo contiene, da cui segue $\mathscr{A}_G(\mathscr{E}) \subset \mathscr{A}(\mathscr{E})$, per ogni $\mathscr{E} \in \mathscr{G}$, e quindi che le algebre di $\{\mathscr{A}_G(\mathscr{E}) : \mathscr{E} \in \mathscr{G}\}$ sono finito-logicamente indipendenti)

4.4 Nota. *Estendibilità del giudizio di indipendenza stocastica.*

Sia $\{\mathscr{A}(\mathscr{E}) : \mathscr{E} \in \mathscr{G}\}$ un insieme di algebre di eventi che "includono" gli insiemi di $\mathscr{G}$. Se le algebre $\mathscr{A}(\mathscr{E})$ sono finito-logicamente indipendenti, è coerente supporre che gli insiemi di $\mathfrak{E}$ siano stocasticamente indipendenti (precedente punto (iv)). Per realizzare concretamente la loro indipendenza stocastica bisogna assegnare (valutare) le probabilità marginali sui singoli insiemi, dare così una probabilità P_0 su $\cup\mathfrak{E}$ e porre infine $P_0(E|E_1\wedge\ldots\wedge E_n)=P_0(E)$ per ogni $E\in\mathcal{E}$, $E_1\in\mathcal{E}_1,\ldots,E_n\in\mathcal{E}_n$ e ogni $\{\mathcal{E}_1,\ldots,\mathcal{E}_n)\}\subset\mathfrak{E}$. Sappiamo che nelle ipotesi attuali è coerente supporre stocasticamente indipendenti anche le algebre di $\{\mathcal{A}(\mathcal{E}):\mathcal{E}\in\mathfrak{E}\}$ (precedente punto (iii)). In casi come questo, dopo aver introdotto l'indipendenza stocastica per gli insiemi di $\mathfrak{E}$, si può avere interesse di estenderla anche alle algebre corrispondenti senza modificare quella degli insiemi. Occorre a tal fine prolungare le probabilità date agli insiemi sulle rispettive algebre, ottenendo così una probabilità P su $\cup\mathfrak{A}$, prolungamento di P_0 su $\cup\mathfrak{E}$, e porre $P(E|E_1\wedge\ldots\wedge E_n)=P(E)$ per ogni $E\in\mathcal{A}(\mathcal{E})$, $E_1\in\mathcal{A}(\mathcal{E}_1),\ldots,E_n\in\mathcal{A}(\mathcal{E}_n)$ e ogni $\{\mathcal{A}(\mathcal{E}_1),\ldots,\mathcal{A}(\mathcal{E}_n)\}\subset\mathfrak{A}$. Queste condizioni diventano ovviamente quelle poste per gli insiemi di $\mathfrak{E}$ quando $E, E_1,\ldots,E_n$ sono eventi di tali insiemi (perché in tal caso la P diventa la restrizione P_0).

5 Numeri aleatori stocasticamente indipendenti.

A ogni numero aleatorio X è associata un'algebra $\mathcal{A}_L(X)$ di eventi *logicamente dipendenti* da X che parlano delle sue determinazioni: quelli logicamente dipendenti dalla partizione $\{X=x : x\in\mathbb{R}\}$, che sono del tipo $(X\in\mathfrak{I})=\bigvee_{x\in\mathfrak{I}}(X=x)$ con $\mathfrak{I}\subset\mathbb{R}$. Ciò premesso, ricordiamo che nella teoria classica e in quella di de Finetti la nozione di indipendenza stocastica per numeri aleatori viene data con riferimento a *convenienti* sottoinsiemi di eventi logicamente dipendenti dai numeri aleatori: gli eventi $X\in\mathfrak{I}$ misurabili (probabilizzabili) secondo Lebesgue e rispettivamente secondo Peano-Jordan (prolungamenti unici di una probabilità assegnata su $\{X\leq x : x\in\mathbb{R}\}$, per σ-additività i primi e additività i secondi). Per quanto visto nel precedente n° 4, nella presente impostazione la nozione è invece significativa con riferimento a sottoinsiemi *qualsiasi* di tali eventi, anche alla loro *totalità*. Ha senso allora dare la seguente definizione, caso particolare della *Definizione* 4.1.

5.1 Definizione. Numeri aleatori stocasticamente indipendenti.

Sia $\mathfrak{X}$ una collezione di insiemi di numeri aleatori. Diremo allora che i *numeri aleatori di $\mathscr{X}$ sono* ***stocasticamente indipendenti*** *se e solo se sono tali le algebre di* $\{\mathscr{A}_{\mathrm{L}}(X) : X \in \mathscr{X}\}$.

Osservato che l'algebra generata da un'algebra è l'algebra stessa e tenuto conto dei punto (ii) del n° 4, si ha che questa definizione è consistente e interpreta il significato intuitivo di indipendenza stocastica dei numeri aleatori di $\mathscr{X}$ se e solo se le algebre di $\{\mathscr{A}_{\mathrm{L}}(X) : X \in \mathscr{X}\}$ sono finito-logicamente indipendenti: allora e solo allora è possibile scegliere separatamente le distribuzioni di probabilità degli $X \in \mathscr{X}$ – sulle corrispondenti algebre $\mathscr{A}_{\mathrm{L}}(X)$ – e supporre che gli X siano stocasticamente indipendenti.

La necessità della condizione è intuitiva, Sembrerebbe altrimenti davvero strano: che legami di natura logica tra i numeri aleatori – tra loro eventi logicamente dipendenti – non inducessero alcun vincolo sulla scelta delle relative distribuzioni di probabilità. Meno scontata appare invece la sua sufficienza, che riguarda poi non solo la coerenza di ogni assegnazione di probabilità marginale, ma anche il prolungamento che si ottiene in sua corrispondenza imponendo le condizioni di indipendenza stocastica. Non sorprende invece, una volta acquisito questo risultato, che sia sufficiente la *finito-logica* indipendenza dei numeri aleatori anche quando l'insieme $\mathscr{X}$ è infinito (che non si debba cioè richiedere in tal caso la *logica* indipendenza), perché sia la norma di coerenza sia le condizioni di indipendenza stocastica mettono a confronto risultati che si ottengono operando con gli eventi dei sottoinsiemi finiti di $\{\mathscr{A}_{\mathrm{L}}(X) : X \in \mathscr{X}\}$; esattamente com'è richiesto per la verifica della finito-logica indipendenza.

Realizzazione dell'indipendenza stocastica.

Per realizzare l'indipendenza stocastica dei numeri aleatori finito-logicamente indipendenti di un insieme $\mathscr{X}$ bisogna scegliere una probabilità su ogni algebra di $\{\mathscr{A}_{\mathrm{L}}(X) : X \in \mathscr{X}\}$ e prolungare poi la valutazione così ottenuta su ogni evento $E | E_1 \wedge \ldots \wedge E_n$, con $E \in \mathscr{A}_{\mathrm{L}}(X)$, $E_1 \in \mathscr{A}_{\mathrm{L}}(X_1)$, … , $E_n \in \mathscr{A}_{\mathrm{L}}(X_n)$, $E_1 \wedge \ldots \wedge E_n \neq \phi$, ponendo $P(E | E_1 \wedge \ldots \wedge E_n) = P(E)$. Resta così determinata per fattorizzazione anche la probabilità sull'insieme $\{E_1 \wedge \ldots \wedge E_n : E_1 \in \mathscr{A}_{\mathrm{L}}(X_1), \ldots, E_n \in \mathscr{A}_{\mathrm{L}}(X_n), \{X_1, \ldots, X_n\} \subset \mathscr{X}\}$ (segue da *Proposizione* 3.5). Dal punto di vista operativo, il calcolo delle probabilità degli eventi $E_1 \wedge \ldots \wedge E_n$ di questo insieme si esegue come nella teoria classica ponendo $P(E_1 \wedge \ldots \wedge E_n) = P(E_1) \ldots P(E_n)$; qui però lo si fa in quanto condizione necessaria per l'indipendenza stocastica e non per sua definizione. Si tenga inoltre presente che, al contrario del caso classico, prolungamento e fattorizzazione sussistono per ogni scelta di eventi (in numero finito) nelle corrispondenti algebre. Non è detto tuttavia che sia possibile effettuare tutti i calcoli, perché ciò è completamente fattibile solo quando si è in grado di dare concretamente le probabilità sulle algebre $\mathscr{A}_L(X)$. Tali probabilità esistono sicuramente in conseguenza del teorema del prolungamento (come già ricordato), ma sono realizzabili *in modo completo* solo se si concentra l'intera probabilità su un numero finito o numerabile di costituenti delle corrispondenti partizioni $\{X = x : x \in \mathbb{R}\}$. Come si deve fare *sempre* se *X è finito* e lo si fa *usualmente* nelle applicazioni se *X è numerabile* (*sempre* nella teoria classica, perché le probabilità sono allora σ-additive). Se *X* è *continuo* è usuale invece scegliere probabilità non concentrate (continue almeno in parte). La loro realizzabilità *completa* in tal caso è "un'astrazione pressoché inimmaginabile", per dirla con de Finetti (testo citato nel n° 2). In questi casi ci si deve accontentare di realizzare valutazioni parziali come nel caso classico. Viene naturale farlo seguendo lo stesso approccio, operando però con probabilità coerenti. Si assegnerà allora una probabilità su $\{X \leq x : x \in \mathbb{R}\}$ – una funzione di ripartizione $F(\cdot)$ (*probabilità coerente*) –, determinando in questo modo via prolungamento la probabilità degli eventi $X \in \mathcal{J}$, con $\mathcal{J}$ Peano-Jordan misurabili mediante la *F*. Si procederà all'occorrenza a eventuali ulteriori prolungamenti, quanto serve e quanto si saprà fare.

Quando i numeri aleatori sono continui (anche solo in parte) la *realizzabilità completa della loro* indipendenza stocastica è dunque una *astrazione pressoché inimmaginabile,* ma solo perché, come si è detto, è tale *l'assegnazione* di probabilità *complete* sulle algebre degli eventi logicamente dipendenti da numeri aleatori continui. È palese tuttavia la straordinaria importanza teorica e metodologica del teorema del prolungamento di de Finetti: chi valuta in modo coerente la probabilità su un insieme di eventi, grazie a tale teorema sa che esistono suoi prolungamenti coerenti su ogni soprainsieme; ove non sia in grado di realizzali, sa comunque che se opera con coerenza non corre il rischio di andare incontro a contraddizioni. Di valore simile è allora anche l'importanza teorica e metodologica dei *Teoremi* 4.2 e 4.3 nei riguardi della nozione di indipendenza stocastica: della sua coerenza e del suo grado di realizzabilità, che peraltro dipende esclusivamente dal grado di realizzabilità delle probabilità marginali.

Forme deboli di indipendenza stocastica

La *Definizione* 5.1 introduce la nozione di indipendenza stocastica dei numeri aleatori di un insieme $\mathscr{X}$ con riferimento alla *totalità* degli eventi logicamente dipendenti dai rispettivi numeri aleatori: quelli delle algebre $\mathscr{A}_L(X)$. In corrispondenza agli stessi numeri aleatori è però possibile introdurre forme di indipendenza stocastica più deboli, riguardanti solo *parte* degli eventi di tali algebre. Si può cioè studiare il problema dell'indipendenza stocastica di collezioni di insiemi di eventi del tipo $\{\mathfrak{I}(X) : \mathfrak{I}(X) \subset \mathscr{A}_L(X), X \in \mathscr{X}\}$. Gli insiemi $\mathfrak{I}(X)$ non possono essere stocasticamente indipendenti se le loro algebre generate $\mathscr{A}_G(X)$ non sono finito-logicamente indipendenti. Altrimenti, è coerente supporli stocasticamente indipendenti ed è altresì coerente fare altrettanto per le loro algebre generate; più in generale, per le algebre degli insiemi del tipo $\{\mathscr{A}(X) : \mathfrak{I}(X) \subset \mathscr{A}(X) \subset \mathscr{A}_L(X), X \in \mathscr{X}\}$, se esse sono finito-logicamente indipendenti (*Nota* 4.4). In quanto a terminologia, riserveremo la dizione «indipendenza stocastica di numeri aleatori di $\mathscr{X}$» al caso $\mathfrak{I}(X) = \mathscr{A}_L(X)$; se qualche $\mathfrak{I}(X)$ è incluso propriamente in $\mathscr{A}_L(X)$ seguiremo invece la terminologia introdotta con la *Definizione* 4.1: diremo semplicemente che si tratta di «indipendenza stocastica di insiemi di eventi», aggiungendo se occorre «logicamente dipendenti da numeri aleatori».

Poiché in corrispondenza ai numeri aleatori è usuale introdurre le distribuzioni di probabilità mediante funzioni di ripartizione, viene naturale esaminare la nozione di indipendenza stocastica nel caso degli insiemi $\mathfrak{I}(X) = \{X \leq x : x \in \mathbb{R}\}$. Le loro algebre generate sono la totalità delle *somme finite* di *eventi base* del tipo $X \leq x$, $\xi < X \leq x$, $X > x$. S'intende allora facilmente che le algebre di $\{\mathscr{A}_G(\{X \leq x : x \in \mathbb{R}\}) : X \in \mathscr{X}\}$ risultano finito-logicamente indipendenti se e solo se sono possibili i prodotti finiti di eventi base. In questa ipotesi è coerente supporre stocasticamente indipendenti sia gli insiemi $\{X \leq x : x \in \mathbb{R}\}$, $X \in \mathscr{X}$, sia le corrispondenti algebre generate.

Per realizzare queste indipendenze stocastiche bisogna dare le relative probabilità marginali e imporre poi le corrispondenti condizioni di indipendenza stocastica. Le probabilità marginali sono determinate dalle funzioni di ripartizione dei singoli numeri aleatori, non solo sugli insiemi $\{X \le x : x \in \mathbb{R}\}$, ma per prolungamento anche sulle loro algebre generate. Gli eventi di queste sono infatti esprimibili come somme finite di eventi base a due a due incompatibili e le loro probabilità sono perciò somme di probabilità di eventi base, per i quali riesce $P(X \le x) = F(x)$, $P(\xi < X \le x) = F(\xi) - F(x)$, $P(X > x) = 1 - F(x)$.

Riesaminiamo in questa ottica l'*Esempio* 5 del n° 2 (di de Finetti). Si verifica subito che le algebre $\mathcal{A}_G(\{X \le x : x \in \mathbb{R}\})$ e $\mathcal{A}_G(\{Y \le y : y \in \mathbb{R}\})$ sono logicamente indipendenti. È perciò coerente supporre stocasticamente indipendenti gli *insiemi* $\{X \le x : x \in \mathbb{R}\}$ e $\{Y \le y : y \in \mathbb{R}\}$. Per realizzare la loro indipendenza stocastica bisogna assegnare le funzione di ripartizione $F_X(\cdot)$, $F_Y(\cdot)$ e porre $P(X \le x \mid Y \le y) = F_X(x)$, $P(Y \le y \mid X \le x) = F_Y(y)$ per ogni $x, y \in \mathbb{R}$. Segue allora che *deve* essere $F(x, , y) = F_X(x) F_Y(y)$. I numeri aleatori X, Y sono perciò stocasticamente indipendenti secondo de Finetti. Ma non secondo le *Definizione* 5.1, perché $\mathcal{A}_L(X)$ e $\mathcal{A}_L(Y)$ *non sono* logicamente indipendenti; basta infatti osservare per questo che se $x = r + s\sqrt{2}$, $y = r' + s'\sqrt{3}$ e $(r, s) \neq (r', s')$, r, s, r', s' razionali, gli eventi $X=x$ e $Y=y$ sono possibili e $(X=x) \wedge (Y=y) = \phi$.

Quella degli insiemi $\{X \le x : x \in \mathbb{R}\}$ e $\{Y \le y : y \in \mathbb{R}\}$ è anzi una indipendenza stocastica molto debole rispetto a quanto è richiesto per l'indipendenza stocastica di X, Y; com'è ovvio che sia, visto che i due numeri aleatori sono funzionalmente dipendenti. Si può estendere infatti l'indipendenza stocastica dei due insiemi alle loro algebre generate (ovvio), ma non agli insiemi (algebre $\mathcal{A}_P(X)$ e $\mathcal{A}_P(Y)$) di tutti gli eventi Peano-Jordan probabilizzabili, di interesse per le applicazioni. Infatti, se r, s sono razionali e scegliamo $F_X(\cdot)$ e $F_Y(\cdot)$ continue nei punti $x = r + s\sqrt{2}$ e $y = r + s\sqrt{3}$, allora x, y sono determinazioni di X, Y, $(X=x) = (Y=y)$ e $P(X=x) = P(Y=y) = 0 \neq 1 = P(X=x \mid Y=y) = P(Y=y \mid X=x)$.

Interessa osservare ancora che la fattorizzazione della funzione di ripartizione congiunta $F(\cdot, \cdot)$ – necessaria affinché gli insiemi di eventi $\{X \le x : x \in \mathbb{R}\}$ e $\{Y \le y : y \in \mathbb{R}\}$ siano stocasticamente indipendenti – implica la fattorizzazione della probabilità sull'insieme $\mathcal{A}_P(X) \wedge \mathcal{A}_P(Y) = \{E_X \wedge E_Y : E_X \in \mathcal{A}_P(X), E_Y \in \mathcal{A}_P(Y)\}$. Le algebre $\mathcal{A}_P(X)$ e $\mathcal{A}_P(Y)$ (dei rispettivi eventi Peano-Jordan probabilizzabili) sono dunque stocasticamente indipendenti secondo la definizione di de Finetti, ma non secondo la *Definizione* 4.1. Infatti, se $E_X \in \mathcal{A}_P(X)$ e $E_Y \in \mathcal{A}_P(Y)$, la fattorizzazione induce le condizioni $P(E_X \mid E_Y) = P(E_X \mid \bar{E}_Y) = P(E_X)$ e $P(E_Y \mid E_X) = P(E_Y \mid \bar{E}_X) = P(E_Y)$ solo se gli eventi ipotesi non sono di probabilità nulla (valgono le prime due uguaglianze se E_Y non è di probabilità estrema, le seconde due se non lo è E_X). Negli altri casi le condizioni dovrebbero essere imposte, ma non è coerente farlo perché le algebre $\mathcal{A}_P(X)$ e $\mathcal{A}_P(Y)$ non sono logicamente indipendenti.

Le considerazioni svolte qui sopra per i numeri aleatori X, Y si possono ripetere con qualche adattamento per i numeri Z, Y dell'*Esempio* 4 e raggiungere le medesime conclusioni. Si trova che è coerente supporre stocasticamente indipendenti gli insiemi $\{Z \leq z : z \in \mathbb{R}\}$ e $\{Y \leq y : y \in \mathbb{R}\}$ e le loro algebre generate, ma non le algebre $\mathcal{A}_P(Z)$ e $\mathcal{A}_P(Y)$. Nonostante la dipendenza logica tra Z e Y dell'*Esempio* 4 sia molto più debole di quella funzionale tra X e Y dell'*Esempio* 5, l'effetto sull'estendibilità dell'indipendenza stocastica resta fortemente restrittivo: neanche nel caso dei numeri aleatori Z e Y è coerente estendere l'indipendenza stocastica almeno alle algebre degli eventi Peano-Jordan probabilizzabili. Si badi che questi risultati valgono se si ragiona in termini di probabilità coerenti. Nella teoria classica X e Y possono infatti essere considerati tranquillamente stocasticamente indipendenti. Basta scegliere funzioni di ripartizione $F_X(\cdot)$, $F_Y(\cdot)$ continue (per assegnare probabilità nulla ai razionali) e prolungare la valutazione per fattorizzazione sulle σ-algebre degli eventi Lebesgue probabilizzabili usando probabilità σ-additive.

Riferimenti bibliografici

[1] CRISMA L. (2006), *Introduzione alla teoria delle probabilità coerenti,* Vol. 1, Ediz. EUT, Trieste.

[2] CRISMA L. (2006), *Introduzione alla teoria delle probabilità coerenti,* Vol. 2, in preparazione (versione provvisoria del cap.15, link: http://hdl.handle.net/10077/3497).

[3] DE FINETTI B. (1949), *Sull'impostazione assiomatica del calcolo delle probabilità,* Annali Triestini; 19, Sez. 2ª, 29-81.

[4] DE FINETTI B. (1970), *Teoria delle probabilità,* Vol. 1, Einaudi, Torino.

[5] HOLZER, S. (1984), *Sull'estensione delle probabilità subordinate,* Rendiconti Ist. Matem. Univ. Trieste, Vol. 16, 96-108.