Riv. Mat. Univ. Parma (4) 8 (1982), 43-48

B. BARIGELLI e L. BASSI (*)

Probabilità disintegrabili e σ-additività (**)

1 - Una probabilità P su un insieme Ω si dice *disintegrabile* rispetto ad una partizione $\pi = \{A_i\}_{i\in J}$ di Ω (con J finito o numerabile) se per ogni evento E risulta

$$P(E) = \sum_{i\in J} P(A_i)\, P(E/A_i)\,. \tag{1}$$

Questa proprietà della probabilità si considera spesso (p. es., nei problemi di inferenza statistica bayesiana) valida (esplicitamente o tacitamente) rispetto a qualunque partizione di Ω. In realtà, se la partizione π non è finita, la (1) non sussiste incondizionatamente: una condizione necessaria e sufficiente è la σ-additività della probabilità P. Inoltre la disintegrabilità di P è collegata alla validità della proprietà conglomerativa (un concetto introdotto da B. de Finetti in [3]): in [4] Dubins ha provato, nel caso più generale di una *previsione*, definita su tutti i numeri aleatori, l'equivalenza dei due concetti analoghi ai precedenti, relativamente ad una arbitraria partizione.

In questo lavoro studiamo, nel caso di una *probabilità*, e con riferimento non soltanto ad assegnate partizioni (numerabili) di Ω, ma anche ad un fissato evento E, le condizioni di validità delle due proprietà suddette, esaminando in particolare i legami con la σ-additività di P. Si accenna poi, nella parte finale del lavoro, ai paradossi cui può dar luogo, nella statistica, l'uso acritico delle due proprietà: in particolare, il problema della prima cifra significativa dei dati statistici è oggetto di uno studio separato di R. Scozzafava in [6]$_2$. Un altro esempio interessante è discusso da B. M. Hill in [5].

(*) Indirizzo: Dipartimento di Matematica e Informatica, Università, 60100 Ancona, Italy.

(**) Lavoro eseguito nell'ambito del G.N.A.F.A. (C.N.R.). — Ricevuto: 20-III-1980.

2 – Una *massa* è una misura di probabilità finitamente additiva su un insieme Ω, cioè una funzione $P\colon \mathscr{P}(\Omega) \to \boldsymbol{R}$ tale che

$$\text{(i)}\ P(\Omega) = 1, \qquad \text{(ii)}\ P(E) \geqslant 0 \quad \text{per ogni } E \subseteq \Omega,$$

$$\text{(iii)}\ P(E_1 \cup E_2) = P(E_1) + P(E_2) \qquad \text{se } E_1 \cap E_2 = \emptyset.$$

In particolare, la massa P è σ-additiva se l'ultima relazione può essere estesa al caso di una infinità numerabile di eventi a due a due disgiunti.

Una massa può essere sempre prolungata come *probabilità condizionata completa* (cfr. [4]) su $\mathscr{P}(\Omega) \times (\mathscr{P}(\Omega) - \emptyset)$, cioè una funzione Q tale che

(i) $Q(\cdot/H)$ è una massa su $\mathscr{P}(\Omega)$ per ogni $H \neq \emptyset$; in particolare $Q(\cdot/\Omega) = P$;

(ii) $Q(H/H) = 1$ per ogni $H \subseteq \Omega$, con $H \neq \emptyset$;

(iii) $P(H \cap E) = P(H)Q(E/H)$ per ogni $H, E \subseteq \Omega$, con $H \neq \emptyset$.

Più in generale, si possono considerare probabilità condizionate complete su $\mathscr{E} \times \mathscr{A}^0$, con $\mathscr{A}^0 = \mathscr{A} - \emptyset$, dove $\mathscr{E}$, $\mathscr{A}$ sono *algebre* di eventi, con $\mathscr{E} \supseteq \mathscr{A}$.

Nel seguito indicheremo con la stessa lettera P anche il prolungamento Q di una massa P.

Sia $\pi = \{A_1, A_2, \ldots, A_r, \ldots\}$ una *partizione* numerabile di Ω. Dato un evento $E \subseteq \Omega$, una massa P si dice *π-disintegrabile su E* se

$$P(E) = \sum_{r=1}^{\infty} P(A_r)P(E/A_r)\,, \tag{2}$$

e si dice *π-disintegrabile* se la (2) vale per ogni $E \subseteq \Omega$. Infine P si dice *disintegrabile* se la (2) vale per ogni E e per ogni partizione numerabile π.

La massa P si dice *π-conglomerativa su E* se

$$P(E/A_r) = p_0 \quad \forall r \in \boldsymbol{N} \Rightarrow P(E) = p_0\,, \tag{3}$$

e si dice *π-conglomerativa* se la (3) vale per ogni $E \subseteq \Omega$.

Dubins ha studiato in [4] i concetti analoghi di *previsione* π-disintegrabile e π-conglomerativa (per partizioni di cardinalità arbitraria), ed ha provato la loro equivalenza, nell'ipotesi che la previsione P sia definita sullo spazio lineare dei numeri aleatori su Ω (cioè delle funzioni reali *limitate* in Ω). Considerando la restrizione di P agli eventi (cioè ai numeri aleatori a valori 0 e 1), si ottiene una *probabilità*, per la quale non è significativa la definizione di Dubins analoga alla (3), in quanto essa risulta sempre banalmente verificata.

Quindi, nel caso da noi studiato, la definizione di disintegrabilità dà una condizione più forte di quella di Dubins.

Teorema 1. *Data una partizione* $\pi = \{A_r\}_{r\in N}$ *di* Ω, *la massa* P *è* σ-*additiva su* π *se e solo se* P *è* π-*disintegrabile.*

Dim. Osserviamo preliminarmente che, se P è σ-additiva su π e $\{B_r\}$ è una successione di eventi tali che $B_r \subseteq A_r$ per ogni $r \in \boldsymbol{N}$, allora P è σ-additiva anche sulla successione $\{B_r\}$; infatti, posto $A_r - B_r = C_r$, si ha

$$\sum_{r=1}^{\infty} P(B_r)$$

$$\leqslant P(\bigcup_{r=1}^{\infty} B_r) = P(\bigcup_{r=1}^{\infty}(A_r - C_r)) = P(\bigcup_{r=1}^{\infty} A_r - \bigcup_{r=1}^{\infty} C_r) = P(\bigcup_{r=1}^{\infty} A_r) - P(\bigcup_{r=1}^{\infty} C_r)$$

$$\leqslant \sum_{r=1}^{\infty} P(A_r) - \sum_{r=1}^{\infty} P(C_r) = \sum_{r=1}^{\infty}(P(A_r) - P(C_r)) = \sum_{r=1}^{\infty} P(B_r)\,.$$

Allora, qualunque sia $E \subseteq \Omega$, si ha

$$P(E) = P(E \cap \Omega) = P(E \cap (\bigcup_{r=1}^{\infty} A_r)) = P(\bigcup_{r=1}^{\infty}(E \cap A_r))$$

$$= \sum_{r=1}^{\infty} P(E \cap A_r) = \sum_{r=1}^{\infty} P(A_r)\, P(E/A_r)\,.$$

Viceversa, se P è π-disintegrabile, si ha

$$1 = P(\Omega) = \sum_{r=1}^{\infty} P(\Omega/A_r)\, P(A_r) = \sum_{r=1}^{\infty} P(A_r)\,,$$

cioè P è σ-additiva su π.

Corollario 1. *P è una massa σ-additiva (cioè, nella terminologia usuale, una misura di probabilità) se e solo se P è disintegrabile.*

Corollario 2. *Se una massa P soddisfa una qualunque delle seguenti tre condizioni*

(a) *fortemente non atomica* (cfr. [**6**]$_1$),

(b) *non atomica,*

(c) *atomica e pseudo assolutamente continua* (cfr. [**2**]) *rispetto ad una misura di probabilità non atomica,*

allora esiste una partizione π di Ω tale che P è π-disintegrabile.

Dim. Per i risultati stabiliti in [6]$_1$, [1], nei tre casi considerati esiste una partizione di Ω sulla quale P è σ-additiva.

Teorema 2. *Se P è una massa π-disintegrabile, essa è anche π-conglomerativa.*

Dim. Se $P(E/A_r) = p_0$ per ogni $A_r \in \pi$, si ha

$$P(E) = \sum_{r=1}^{\infty} P(A_r)\, P(E/A_r) = p_0 \sum_{r=1}^{\infty} P(A_r) = p_0 \,.$$

3 – Se P *non* è σ-additiva su una partizione π, possono esistere eventi E tali che P è π-disintegrabile ma non π-conglomerativa su E, ed anche eventi E tali che P è π-conglomerativa ma non π-disintegrabile su E.

Esempio 1. Sia Ω l'insieme dei razionali di $(0, 1]$ (che indicheremo con lo stesso simbolo $(0, 1]$, e così faremo analogamente per i razionali di un intervallo $(a, b] \subseteq (0, 1]$, che indicheremo con $(a, b]$). Sull'algebra generata dagli insiemi $(a, b]$ definiamo una massa $P((a, b]) = b - a$, che poi prolunghiamo (teorema di Tarski, cfr. [7]) a tutto $\mathscr{P}(\Omega)$. Posto

$$E_k = \left(\frac{1}{k+1}, \frac{1}{2}\left(\frac{1}{k+1} + \frac{1}{k}\right)\right] \qquad \text{e} \qquad F_k = \left(\frac{1}{2}\left(\frac{1}{k+1} + \frac{1}{k}\right), \frac{1}{k}\,\right],$$

sia $E = \bigcup_{k=2}^{\infty} E_k$. Si ha

$$P\big((0, \tfrac{1}{2}]\big) = \tfrac{1}{2} = P(E) + P\big((0, \tfrac{1}{2}] - E\big)$$

$$\geqslant \sum_{k=2}^{\infty} P(E_k) + \sum_{k=2}^{\infty} P(F_k) = 2 \sum_{k=2}^{\infty} \frac{1}{2}\left(\frac{1}{k} - \frac{1}{k+1}\right) = \frac{1}{2}\,,$$

e quindi $P(E) = \frac{1}{4}$.

Consideriamo ora la seguente partizione di Ω: $\pi = \{A_2, A_3, \ldots, A_k, \ldots\}$ con $A_k = \left(\frac{1}{k+1}, \frac{1}{k}\right] \cup \{q_k\}$, essendo $q_2, q_3, \ldots$ una numerazione dei razionali di $(\frac{1}{2}, 1]$. Si ha $\bigcup_{k=2}^{\infty} A_k = (0, 1]$, mentre $\sum_{k=2}^{\infty} P(A_k) = \frac{1}{2}$, e quindi P non è σ-additiva su π.

Verifichiamo che $P(E)$ è π-disintegrabile:

$$1/4 = P(E) = P(\bigcup_{k=2}^{\infty}(E \cap A_k))$$

$$\geqslant \sum_{k=2}^{\infty} P(E \cap A_k) = \sum_{k=2}^{\infty} P(A_k)\, P(E/A_k) = \sum_{k=2}^{\infty} \left(\frac{1}{k} - \frac{1}{k+1}\right)\frac{1}{2} = \frac{1}{4}\,.$$

Osserviamo che $P(E)$ *non* è π-conglomerativa, perchè $P(E/A_k) = \frac{1}{2}$ per ogni $A_k \in \pi$.

Esempio 2. Sia $\Omega = \boldsymbol{N}$, e definiamo $P(E) = \lim_{m\to\infty} \gamma(m)/m$ (per gli E per cui tale limite esiste), con $\gamma(m) = \operatorname{card}(E \cap \{1, 2, ..., m\})$, prolungando poi P a tutto $\mathscr{P}(\Omega)$.

Considerando l'algebra $\mathscr{A}$ dei sottoinsiemi finiti e cofiniti H di Ω, estendiamo la P come probabilità condizionata su $\mathscr{P}(\Omega) \times \mathscr{A}^0$, ponendo $P(E/H) = \operatorname{card}(E \cap H)/\operatorname{card} H$ se H è finito, e $P(E/H) = P(E \cap H)/P(H)$ se H è cofinito. Consideriamo la partizione $A_k = \{2k-1, 2k\}_{k\in \boldsymbol{N}}$.

Si ha $P(A_k) = 0$ per ogni $k \in \boldsymbol{N}$, e quindi P non è σ-additiva su π. Se $E = \{1, 3, 5, ..., 2k-1, ...\}$, si ha $P(E) = \frac{1}{2}$; inoltre, per ogni $k \in \boldsymbol{N}$, $P(E/A_k) = \frac{1}{2}$ e quindi P è π-conglomerativa su E. Invece, come è facile verificare, P non è π-disintegrabile su E.

In questo esempio, gli unici eventi E tali che $P(E/A_k)$ ha lo stesso valore per ogni $k \in \boldsymbol{N}$ sono, oltre a quello già considerato, l'insieme dei numeri pari (1) e gli insiemi $\emptyset$ e Ω, ed anche per essi vale la proprietà π-conglomerativa. Questo prova che il Teorema 2 non è invertibile.

D'altra parte, esistono anche partizioni di Ω rispetto alle quali P non è conglomerativa (cfr. [**3**]).

4 – Nella statistica matematica, per trarre conclusioni su possibili ipotesi, si fa spesso ricorso alle probabilità condizionate. Per esempio, se $P(E/H)$ è piccolo (secondo un criterio convenzionale prestabilito) ed E si verifica, questo si ritiene un motivo valido per giudicare poco plausibile H (si rigetta l'ipotesi H). In altre impostazioni, si fa un confronto tra due (o più) ipotesi alternative, decidendo per l'una o l'altra sulla base di valutazioni riguardanti $P(E/H)$ e $P(E/H^c)$, dove l'evento E rappresenta il risultato di un certo esperimento, e con H^c indichiamo l'evento contrario di H.

Lo studio delle proprietà conglomerativa e disintegrativa mette in luce come sia necessaria un'estrema cautela nel manipolare le probabilità condizionate. Un caso concreto è stato studiato in [**6**]$_2$: le osservazioni statistiche

(1) Più precisamente, si ha $P(E/A_k) = 1/2$ per ogni E tale che $\operatorname{cond}(E \cap A_k) = 1$.

sarebbero (secondo gli autori che si sono occupati di quel problema in precedenza) in contrasto con l'intuizione, ma ciò avviene solo perchè si considera valida (tacitamente) la proprietà conglomerativa (che invece in quel caso non sussiste).

Bibliografia

[1] E. BARONE, A. GIANNONE and R. SCOZZAFAVA, *On some aspects of the theory and applications of finitely additive probability measures*, Pubbl. Ist. Mat. Appl. Univ. Roma **221**, quad. **16** (1980), 43-53.

[2] G. COLETTI e G. REGOLI, *Sulla funzione di distribuzione di una misura di probabilità finitamente additiva*, Rend. Mat. (VII) **1** (1981), 319-329.

[3] B. DE FINETTI, *Sulla proprietà conglomerativa delle probabilità subordinate*, Ist. Lombardo Accad. Sci. Lett. Rend. **63** (1930), 414-418.

[4] L. E. DUBINS, *Finitely additive conditional probabilities, conglomerability and disintegrations*, Ann. Probability **3** (1975), 89-99.

[5] B. M. HILL, *On some statistical paradoxes and non-conglomerability*, International Meeting on Bayesian Statistics, Valencia 1979, 39-49.

[6] R. SCOZZAFAVA: $[\bullet]_1$ *Completa additività su opportune successioni di insiemi di una misura di probabilità semplicemente additiva e fortemente non atomica*, Boll. Un. Mat. Ital. (5) **16** B (1979), 639-648; $[\bullet]_2$ *Un esempio concreto di probabilità non σ-additiva: la distribuzione della prima cifra significativa dei dati statistici*, Boll. Un. Mat. Ital. (5) **18** A (1981), 403-410.

[7] A. TARSKI, *Algebraische Fassung des Massproblems*, Fund. Math. **31** (1938), 47-66.

Summary

We study disintegrations of conditional probabilities and their relationship to conglomerability and countable additivity. Disintegration is equivalent to countable additivity, and it implies conglomerability, but the converse is not true, as we show by a counter-example.

* * *