

4.1 LEMMA DI BOREL E CANTELLI (SECONDA PARTE). *Se gli eventi della successione A_n sono due a due indipendenti e la serie $\sum P(A_n)$ diverge, allora l'evento $\limsup A_n$ ha probabilità uno.*

Dimostrazione. Posto $a_k = P(A_k)$, e $s_n = a_1 + \dots + a_n$, poniamo $S_n = I_{A_1} + \dots + I_{A_n}$ e $S = \sup_n S_n$. Grazie all'indipendenza, le variabili $Y_i = I_{A_i} - a_i$ sono due a due ortogonali in L^2 e si ha $P(Y_i^2) = P(I_{A_i} + a_i^2 - 2a_i I_{A_i}) \leq a_i$. Ne segue $P[(S_n - s_n)^2] \leq s_n$.

Fissato q , da un certo n in poi si avrà $s_n > q$ e l'evento $\{S \leq q\}$ sarà incluso in $\{S_n \leq q\}$, a sua volta incluso in $\{|S_n - s_n| \geq s_n - q\}$. Dalla disuguaglianza di Markov scritta per la variabile $(S_n - s_n)^2$ e per i numeri $(s_n - q)^2$ si ha allora $P(S \leq q) \leq s_n(s_n - q)^{-2}$, che tende a zero al crescere di n . Facendo la riunione sui q razionali, si conclude che S è quasi certamente uguale a $+\infty$. $\square$

Indipendenza di variabili aleatorie

Due variabili aleatorie X e Y (definite sullo stesso spazio di probabilità) saranno dette indipendenti se lo sono gli eventi $\{X \in A\}$ e $\{Y \in B\}$ per ogni scelta di A e B misurabili negli spazi d'arrivo di X e Y rispettivamente. In altre parole, se ogni elemento della tribù generata dall'una è indipendente da ciascun elemento di quella generata dall'altra (con un ragionevole abuso di linguaggio, diremo indipendenti le tribù).

L'indipendenza di X e Y esprime il fatto che per ogni H di probabilità non nulla appartenente alla tribù generata da Y e per ogni A si ha l'uguaglianza $P_H(X \in A) = P(X \in A)$, e cioè che per ogni H siffatto la legge di X secondo P_H coincide con quella secondo P . In altre parole, un'informazione sul valore di Y non altera la legge di X ; analogo enunciato ove si scambi X con Y .

Dalla definizione appare chiaro che l'indipendenza di una coppia di variabili aleatorie è una proprietà della legge della coppia, di modo che una coppia di variabili isonoma a una indipendente è pure indipendente.

4.2 TEOREMA. *Due variabili sono indipendenti se e solo se la legge della coppia è la misura prodotto delle leggi delle singole variabili.*

Dimostrazione. L'indipendenza delle variabili X e Y equivale all'uguaglianza tra $P(X \in A, Y \in B)$, che è il valore della legge della coppia $[X, Y]$ sul rettangolo $A \times B$, e il prodotto $P(X \in A)P(Y \in B)$ dei valori della legge di X in A e di quella di Y in B , per ogni scelta del rettangolo. Equivale a dire che la legge della coppia e prodotto delle singole leggi (entrambe misure di probabilità) coincidono.

4.3 TEOREMA. *Il prodotto di due variabili aleatorie reali indipendenti è integrabile e ha speranza uguale al prodotto delle speranze.*

Dimostrazione. Se μ è la legge di X e ν quella di Y , la coppia avrà legge $\mu \otimes \nu$; per integrazione rispetto alla misura immagine si ha $P(X) = \int x \mu(dx)$, $P(Y) = \int y \nu(dy)$ e $P(XY) = \int xy \mu \otimes \nu(dx, dy)$ e ci si segue dal teorema di Fubini, una volta riconosciuta l'integrabilità con il medesimo ragionamento applicato ai valori assoluti delle funzioni.

Ancora un ricorso al teorema di Fubini porta al seguente lemma, certo molto frequentemente in futuro.

4.4 LEMMA DI RIDUZIONE. *Sia $\mathcal{F}$ una tribù indipendente da Y che X misurabile, e sia f una funzione reale definita sul prodotto degli spazi d'arrivo di X e di Y tale che $Z = f(X, Y)$ risulti integrabile. $g(x) = P[f(x, Y)]$, la variabile $g \circ X$ equivale allora a Z su $\mathcal{F}$.*

Dimostrazione. Sia J l'indicatrice di un elemento di $\mathcal{F}$; la coppia $[J, \text{indipendente da } Y]$, dette μ e ν le leggi rispettive, integrando rispetto alla misura immagine e usando il teorema di Fubini, la speranza di JZ si esprime nella forma $\int u (\int f(u, w) \nu(dw)) \mu(du, dv)$, ove l'integrale interno altro che $g(v)$, come si vede ancora per integrazione rispetto alla misura immagine. In conclusione si ottiene $P(JZ) = \int u g(v) \mu(du, dv) = P(Jg \circ X)$. $\square$

Sovente accade di applicare il lemma quando sono date X e Y indipendenti e $\mathcal{F}$ è la tribù generata da X . In particolare si ha $P[f(X, Y)] = P[\text{detto in modo informale, per calcolare la speranza di } f(X, Y) \text{ si tratta prima } X \text{ come una costante e si determina la funzione } g(x) = P[f(x, Y) \text{ si calcola in } X \text{ tale funzione e se ne prende la speranza; questo procedimento si fa in due fasi dà nome al lemma.}]$

Anche la proposizione seguente sarà utile in seguito.

4.5 TEOREMA. *Se la somma di due variabili aleatorie reali indipendenti è integrabile, tale risultato ciascuna di esse.*

Dimostrazione. Siano X e Y le variabili; se $X + Y$ è quasi certamente costante c , la variabile $c - X$, quasi certamente uguale a Y , è indipendente da X , donde X risulta indipendente da se stessa; due eventi complementari della tribù generata da X sono indipendenti; si annulla il prodotto delle probabilità e dunque la probabilità di uno di essi, per cui la tribù è degenera. Dal lemma 3.2 segue che X è banale; lo stesso per Y .

18 3. VARIABILI ALEATORIE

$\|g_k \circ X_1 - X_1\|$ tende a zero e risulta minore di ϵ per k opportuno. Si prenda dunque $f = g_k$ e si noti che la legge di $f \circ X_n - X_n$ non dipende da n .

Detta A una costante che migliora il valor assoluto di f , poniamo $X_n^* = f \circ X_n$ e naturalmente $S_n^* = X_1^* + \dots + X_n^*$; si avrà evidentemente $|S_n^*/n| \leq A$; inoltre $\|(S_n - S_n^*)/n\| \leq \frac{1}{n} \sum_{k=1}^n \|X_k - X_k^*\| \leq \epsilon$.

Per la parte del teorema già dimostrata, la successione S_n^*/n tende quasi certamente a un limite L^* , e la convergenza, dominata da A , ha luogo anche in L^1 ; per il lemma di Fatou si ha poi $\|L - L^*\| \leq \epsilon$. Abbiamo allora

$$\begin{aligned} \|S_n/n - L\| &\leq \|(S_n - S_n^*)/n\| + \|S_n^*/n - L^*\| + \|L - L^*\| \\ &\leq 2\epsilon + \|S_n^*/n - L^*\| \end{aligned}$$

da cui $\limsup \|S_n/n - L\| \leq 2\epsilon$ e quindi S_n/n tende a L in L^1 per $n \rightarrow \infty$ dato che ϵ è arbitrario. $\square$

Con le ovvie modifiche, si vede facilmente che se le variabili X_n appartengono a L^p , la convergenza di S_n/n ha luogo pure in L^p .

L'integrabilità è tutt'altro che necessaria per la convergenza quasi certa delle medie: per una successione costante (dunque stazionaria) di variabili aleatorie tutte uguali a una non integrabile, le medie (ovviamente tutte uguali) convergono banalmente dappertutto. Addirittura esistono successioni stazionarie non integrabili le cui medie tendono quasi certamente a zero: basta prendere una variabile X non integrabile isonoma a $-X$ (impareremo in seguito che esistono) e definire una successione stazionaria ponendo $X_{2k+1} = X$ e $X_{2k} = -X$; si ha allora $|S_n/n| \leq \frac{1}{n}|X| \rightarrow 0$.

Concludiamo l'esposizione con un ulteriore risultato significativo.

3.14 TEOREMA. *Sia X_n una successione stazionaria di variabili integrabili positive e sia $L = \liminf (X_1 + \dots + X_n)/n$. Allora tutte le X_n sono nulle quasi certamente sull'insieme $\{L = 0\}$.*

Dimostrazione. Sia J l'indicatrice di $\{L = 0\}$; la successione JX_n è stazionaria, in particolare $P(JX_k) = P(JX_1)$ per ogni k ; facendo la media aritmetica su k da 1 a n si ottiene $P(JX_1) = P(J(X_1 + \dots + X_n)/n)$; facendo tendere n all'infinito e usando la convergenza in L^1 delle medie verso L se ne deduce $P(JX_k) = P(JX_1) = P(JL) = 0$, ossia $X_k = 0$ quasi certamente su $\{L = 0\}$. $\square$

CAPITOLO 4

Indipendenza

La nozione di indipendenza ha grande importanza nel calcolo delle probabilità: essa formalizza l'idea intuitiva di un sistema aleatorio nel quale la conoscenza di una parte non dà informazioni sugli aspetti probabilistici del comportamento del resto.

Per procedere occorre precisare meglio il concetto vago di probabilità relativa alla conoscenza di una parte.

Probabilità relativa, eventi indipendenti

La valutazione di probabilità degli eventi aleatori non è una caratteristica inmutabile, ma dipende dalle notizie che si hanno su di essi. Ovvio dunque che tale valutazione muti quando si aggiungono informazioni nuove.

Se ad esempio si conosce per certo l'evento H di probabilità non nulla, la probabilità P dovrà ragionevolmente essere sostituita con la nuova misura P_H definita da $P_H(A) = P(A \cap H)/P(H)$, che diremo *probabilità relativa all'evento H* .

La notazione $P[A | H]$ e il termine *probabilità condizionale* sono altrettanto usati.

Si vede subito che P_H è portata da H , cioè si annulla su H^c ; rispetto a P ha densità $\frac{1}{P(H)}I_H$; per la speranza di X calcolata secondo P_H (speranza relativa a H) se ne ricava l'espressione $\int X dP_H = \frac{1}{P(H)} \int_H X dP$.

Due eventi A e B si dicono *indipendenti* (rispetto a P) se sussiste l'uguaglianza $P(A \cap B) = P(A)P(B)$; ovvia se uno dei due ha probabilità nulla, la relazione equivale a $P_B(A) = P(A)$ per $P(B) > 0$; esprime dunque l'idea che il sapere verificato B non altera la probabilità da attribuire a A e corrisponde all'idea intuitiva di indipendenza.

Si osservi che gli eventi indipendenti da un evento fissato formano una classe di Dynkin.

Un primo risultato completa il lemma di Borel e Cantelli.

4.1 LEMMA DI BOREL E CANTELLI (SECONDA PARTE). Se gli eventi della successione A_n sono due a due indipendenti e la serie $\sum P(A_n)$ diverge, allora l'evento $\limsup A_n$ ha probabilità uno.

Dimostrazione. Posto $a_k = P(A_k)$, e $s_n = a_1 + \dots + a_n$, poniamo $S_n = I_{A_1} + \dots + I_{A_n}$, e $S = \sup_n S_n$. Grazie all'indipendenza, le variabili $Y_i = I_{A_i} - a_i$ sono due a due ortogonali in L^2 e si ha $P(Y_i^2) = P(I_{A_i} + a_i^2 - 2a_i I_{A_i}) \leq a_i$. Ne segue $P[(S_n - s_n)^2] \leq s_n$.

Fissato q , da un certo n in poi si avrà $s_n > q$ e l'evento $\{S \leq q\}$ sarà incluso in $\{S_n \leq q\}$, a sua volta incluso in $\{|S_n - s_n| \geq s_n - q\}$. Dalla disuguaglianza di Markov scritta per la variabile $(S_n - s_n)^2$ e per i numeri $(s_n - q)^2$ si ha allora $P(S \leq q) \leq s_n(s_n - q)^{-2}$, che tende a zero al crescere di n . Facendo la riunione sui q razionali, si conclude che S è quasi certamente uguale a $+\infty$. $\square$

Indipendenza di variabili aleatorie

Due variabili aleatorie X e Y (definite sullo stesso spazio di probabilità) saranno dette indipendenti se lo sono gli eventi $\{X \in A\}$ e $\{Y \in B\}$ per ogni scelta di A e B misurabili negli spazi d'arrivo di X e Y rispettivamente. In altre parole, se ogni elemento della tribù generata dall'una è indipendente da ciascun elemento di quella generata dall'altra (con un ragionevole abuso di linguaggio, diremo indipendenti le tribù).

L'indipendenza di X e Y esprime il fatto che per ogni H di probabilità non nulla appartenente alla tribù generata da Y e per ogni A si ha l'uguaglianza $P_H(X \in A) = P(X \in A)$, e cioè che per ogni H siffatto la legge di X secondo P_H coincide con quella secondo P . In altre parole, un'informazione sul valore di Y non altera la legge di X ; analogo enunciato ove si scambi X con Y .

Dalla definizione appare chiaro che l'indipendenza di una coppia di variabili aleatorie è una proprietà della legge della coppia, di modo che una coppia di variabili isonoma a una indipendente è pure indipendente.

4.2 TEOREMA. Due variabili sono indipendenti se e solo se la legge della coppia è la misura prodotto delle leggi delle singole variabili.

Dimostrazione. L'indipendenza delle variabili X e Y equivale all'uguaglianza tra $P(X \in A, Y \in B)$, che è il valore della legge della coppia $[X, Y]$ sul rettangolo $A \times B$, e il prodotto $P(X \in A)P(Y \in B)$ dei valori della legge di X in A e di quella di Y in B , per ogni scelta del rettangolo. Equivale a dire che la legge della coppia e prodotto delle singole leggi (entrambe misure di probabilità definite sul prodotto degli spazi d'arrivo delle due variabili) coincidono.

4.3 TEOREMA. Il prodotto di due variabili aleatorie reali indipendenti tramite integrabili è integrabile e ha speranza uguale al prodotto speranze.

Dimostrazione. Se μ è la legge di X e ν quella di Y , la coppia $[X, Y]$ avrà legge $\mu \otimes \nu$; per integrazione rispetto alla misura immagine si ha $P(X) = \int x \mu(dx)$, $P(Y) = \int y \nu(dy)$ e $P(XY) = \int xy \mu \otimes \nu(dx, dy)$ e l'ciato segue dal teorema di Fubini, una volta riconosciuta l'integrabilità di xy con il medesimo ragionamento applicato ai valori assoluti delle funzioni.

Ancora un ricorso al teorema di Fubini porta al seguente lemma, che rimo molto frequentemente in futuro.

4.4 LEMMA DI RIDUZIONE. Sia $\mathcal{F}$ una tribù indipendente da Y che X misurabile, e sia f una funzione reale definita sul prodotto degli zzi d'arrivo di X e di Y tale che $Z = f(X, Y)$ risulti integrabile. $g(x) = P[f(x, Y)]$, la variabile $g \circ X$ equivale allora a Z su $\mathcal{F}$.

Dimostrazione. Sia J l'indicatrice di un elemento di $\mathcal{F}$; la coppia $[J, X]$ indipendente da Y ; dette μ e ν le leggi rispettive, integrando rispetto misura immagine e usando il teorema di Fubini, la speranza di JZ si es nella forma $\int u (\int f(u, w) \nu(dw)) \mu(du, dv)$, ove l'integrale interno altro i che $g(v)$, come si vede ancora per integrazione rispetto alla misura imma In conclusione si ottiene $P(JZ) = \int u g(v) \mu(du, dv) = P(Jg \circ X)$. $\square$

Sovente accade di applicare il lemma quando sono date X e Y indipen e $\mathcal{F}$ è la tribù generata da X . In particolare si ha $P[f(X, Y)] = P[g$ detto in modo informale, per calcolare la speranza di $f(X, Y)$ si tratta prima X come una costante e si determina la funzione $g(x) = P[f(x, Y)]$ si calcola in X tale funzione e se ne prende la speranza; questo procedim in due fasi dà nome al lemma.

Anche la proposizione seguente sarà utile in seguito.

4.5 TEOREMA. Se la somma di due variabili aleatorie reali indipende bande (rispettivamente: di L^p), tale risulta ciascuna di esse.

Dimostrazione. Siano X e Y le variabili; se $X + Y$ è quasi certamente ug alla costante c , la variabile $c - X$, quasi certamente uguale a Y , è indipen da X , donde X risulta indipendente da se stessa; due eventi compleme della tribù generata da X sono indipendenti; si annulla il prodotto delle probabilità e dunque la probabilità di uno di essi, per cui la tribù è dege. Dal lemma 3.2 segue che X è banale; lo stesso per Y .

è in L^p per ν -quasi tutti gli y , quindi per almeno uno, e X è in L^p ; lo stesso vale per Y . $\square$

Una successione di variabili aleatorie due a due indipendenti ha talvolta proprietà singolari, come nel seguente teorema di densità nel supporto, ove una simile successione mette quasi certamente in mostra il più ampio spettro di valori possibili.

4.6 TEOREMA. *Una successione di variabili aleatorie due a due indipendenti a valori in uno stesso spazio topologico a base numerabile di aperti e tra loro isonome è quasi certamente densa nel supporto della legge comune.*

Dimostrazione. Sia X_n la successione. Per ogni aperto G della base che incontra il supporto, gli eventi $\{X_n \in G\}$ hanno ugual probabilità strettamente positiva; si applica la seconda parte del lemma di Borel e Cantelli, e si trova in particolare che l'evento $H_G = \cap_n \{X_n \in G^c\}$ ha probabilità zero. L'insieme in cui la successione non è densa è d'altra parte incluso nella riunione degli eventi H_G al variare di G nella base. $\square$

Nello stesso spirito vale la seguente proposizione

4.7 LEMMA. *Siano X_n reali due a due indipendenti tra loro isonome non integrabili. Allora la successione $\{|X_n|/n\}$ è quasi certamente non limitata.*

Dimostrazione. La relazione $|Z| \leq 1 + \sum_{n=1}^{\infty} I_{\{|Z|>n\}}$ mostra che la serie $\sum P(|Z|>n)$ diverge quando Z non è integrabile.

Dato il naturale r , per le variabili non integrabili X_n/r si deduce la divergenza della serie $\sum_n P(|X_n|/n > r)$, che unita alla seconda parte del lemma di Borel e Cantelli permette di concludere che quasi certamente sussiste la relazione $|X_n|/n > r$ per infiniti indici n ; ma r è arbitrario, e la prova si conclude. $\square$

Famiglie indipendenti

Supponiamo che I abbia almeno due elementi; per ogni $i \in I$ sia X_i una variabile aleatoria su $(\Omega, \mathcal{A}, P)$ a valori in E_i ; la famiglia $\{X_i\}_{i \in I}$ si dice indipendente se un'informazione sui valori di alcuni suoi elementi non altera la legge degli altri; formalmente, se per ogni coppia J, L di parti disgiunte non vuote di I le due variabili aleatorie costituite dall'insieme delle X_j con $j \in J$ e dall'insieme delle X_l con $l \in L$ sono indipendenti, ossia lo sono le due tribù da esse generate.

Conviene poi definire indipendente per convenzione ogni famiglia formata

In particolare, le variabili di una famiglia indipendente sono indipendenti due a due, ma l'implicazione inversa non sussiste, come vedremo. È caso di osservare che una famiglia isonoma a una indipendente è indipendente anch'essa.

L'indipendenza di una famiglia comporta conseguenze assai importanti e estremamente significative la seguente

4.8 DICOTOMIA DI KOLMOGOROV. *La tribù terminale di una successione indipendente è degenere.*

Dimostrazione. Sia H un evento terminale per la successione indipendente X_n , avente probabilità non nulla; proviamo che ha probabilità nulla complementare.

Per ogni n , l'evento H appartiene alla tribù generata dalle variabili $i > n$, dunque è indipendente dalla famiglia finita $\{X_1, \dots, X_n\}$, la quale pertanto la stessa legge sia secondo P che secondo la probabilità relativa. Per il criterio delle sottofamiglie finite, l'intera successione X_n ha la legge secondo P e secondo P_H , cioè risulta indipendente da H .

Ora H^c appartiene alla tribù generata dalla successione, pertanto pendente da H ; il prodotto delle probabilità di H e di H^c è dunque zero. Dal teorema segue in particolare che ogni variabile aleatoria reale che sia terminale per una successione indipendente è banale.

Il teorema che segue caratterizza con maggior precisione le famiglie indipendenti. Per comodità tipografica, usiamo temporaneamente le abbreviazioni X_i per le $[X_i]_{i \in I}$ considerate in blocco e $\mathcal{F}_J$ per la tribù generata da esse.

4.9 TEOREMA. *Per la famiglia $\{X_i\}_{i \in I}$ sono equivalenti*

- la famiglia è indipendente;
- ogni sottofamiglia finita è indipendente;
- ogni parte finita F di I con più di un elemento è riunione di disgiunte non vuote F' e F'' tali che $X_{F'}$ e $X_{F''}$ sono indipendenti;
- per ogni parte finita non vuota F di I e arbitrari A_i misurabili spazio d'arrivo di X_i si ha $P(\cap_{i \in F} \{X_i \in A_i\}) = \prod_{i \in F} P(X_i \in A_i)$.

Dimostrazione. Ovviamente (a) implica (b) e (b) implica (c).

Che (c) implichi (d) è banale quando F ha un solo elemento; alt si usa l'esistenza della partizione e si trae la relazione $P(\cap_{i \in F} \{X_i \in A_i\}) = P(\cap_{i \in F'} \{X_i \in A_i\})P(\cap_{i \in F''} \{X_i \in A_i\})$; si continua poi ricorrendo a parti F' e F'' .

Infine (d) implica (a). Siano infatti I' e I'' parti disgiunte non vuote le intersezioni finite di elementi del tipo $\{X_i \in A_i\}$ con i in I' forma

ogni evento di K' è indipendente da un fissato elemento K'' di K'' , e perciò $\mathcal{F}_{I'}$ lo è da K'' grazie al lemma di Dynkin; scambiando i ruoli, si giunge alla voluta indipendenza di $X_{I'}$ e $X_{I''}$. $\square$

Il teorema rende evidenti alcune importanti conseguenze. In primo luogo, come si è visto nel caso due sole variabili, affinché una famiglia sia indipendente occorre e basta che la legge di ogni sua sottofamiglia finita sia il prodotto delle leggi individuali delle variabili che la costituiscono. Da questo segue

4.10 TEOREMA. *Perché due famiglie indipendenti di variabili aleatorie (con lo stesso insieme di indici a valori negli stessi spazi d'arrivo) siano isonome, basta che abbiano le stesse leggi individuali.*

Dimostrazione. Le sottofamiglie finite corrispondenti sono isonome, poiché hanno come legge il prodotto delle stesse leggi. $\square$

Come altra conseguenza del teorema di caratterizzazione va citato il

4.11 CRITERIO ITERATIVO D'INDIPENDENZA. *Sia I una parte della retta. Affinché la famiglia $\{X_i\}_{i \in I}$ sia indipendente è sufficiente che, per ogni $i \in I$ che non sia il primo elemento di I , la variabile aleatoria X_i sia indipendente da quella formata dall'insieme delle X_j con $j < i$.*

Dimostrazione. Se F è una parte finita di I con almeno due elementi e k l'ultimo elemento di F rispetto alla relazione $\leq$, allora $F' = \{k\}$ e $F'' = F \setminus F'$ formano una partizione che verifica 4.9 (c). $\square$

Legge dei grandi numeri

Una successione indipendente di variabili con ugual spazio d'arrivo tra loro isonome è stazionaria; è infatti isonoma alla sua traslata per il teorema 4.10, dato che si tratta di due successioni indipendenti con le stesse leggi individuali.

A una tale successione X_n di variabili reali integrabili si applica dunque il teorema delle medie del capitolo 3; posto $S_n = X_1 + \dots + X_n$, la successione $\{S_n/n\}$ converge quasi certamente e in L^1 verso la variabile aleatoria terminale $Z = \liminf (S_n/n)$, banale in questo caso per la dicotomia di Kolmogorov. La convergenza in L^1 prova poi che Z coincide quasi certamente con il comune valore della speranza delle variabili S_n/n , a sua volta uguale alla speranza di una qualsiasi delle X_n .

A differenza di quanto accade per una generica successione stazionaria, per le successioni indipendenti non può aversi convergenza quasi certamente delle medie se manca l'integrabilità. Più precisamente, si ha la seguente *legge dei grandi numeri*:[†]

[†] Fatto abbastanza sorprendente, il teorema resta valido sotto la semplice ipotesi che le

4.12 TEOREMA. *Siano X_n reali indipendenti, con ugual legge, e per $S_n = X_1 + \dots + X_n$. Se le X_n sono integrabili con comune speranza allora $S_n/n \rightarrow m$ quasi certamente e in L^1 . Se sono non integrabili, la successione S_n/n è quasi certamente non limitata. Se sono infine positive integrabili, allora S_n/n tende quasi certamente a $+\infty$.*

Dimostrazione. Della prima asserzione si è detto. Quanto alla seconda, che le variabili siano due a due indipendenti, tra loro isonome e non integrate, grazie al lemma 4.7, la successione $\{X_n/n\}$ risulti quasi certamente non limitata. A maggior ragione sarà quasi certamente non limitata la successione $\{S_n/n\}$, come si vede dalla semplice relazione

$$\frac{X_n}{n} = \frac{S_n}{n} - \frac{n-1}{n} \frac{S_{n-1}}{n-1}.$$

Per la terza, si applichi la prima parte alle variabili $X_n \wedge k$, indipendenti e con la stessa legge, per dedurne che $\liminf_n S_n/n$ maggiori speranza di $X_1 \wedge k$; dal momento che k è arbitrario e $P(X_1 \wedge k)$ tende al crescere di k , la prova è conclusa. $\square$

Costruzione di variabili aleatorie indipendenti

Il problema consiste nel trovare, su un opportuno spazio di probabilità partenzia, una famiglia indipendente di variabili aleatorie con spazi d'assegnati e leggi individuali prefissate.

La soluzione è banale se la famiglia è finita: basta prendere come spazio di probabilità il prodotto degli spazi d'arrivo, munito del prodotto delle leggi individuali assegnate, e considerarvi le proiezioni coordinate (famiglia canonica). In analogia a 4.2, si constata agevolmente che la famiglia, le cui variabili sono per costruzione quelle volute, è indipendente.

Un teorema classico di DANIELL prova che l'idea funziona per una famiglia qualsiasi; sebbene non troppo difficile, il teorema presenta tuttavia una complicazione tecnica, e non lo dimostreremo. Di volta in volta, allestirei famiglie desiderate "a mano", in modo non canonico; di solito i procedimenti adottati si rivelerebbero istruttivi.

Accade ad esempio di voler schematizzare una successione di esperimenti aleatori in ciascuno dei quali può verificarsi un certo fenomeno, supponendo che sepperlo avverato o meno in alcuni esperimenti non alteri la probabilità suo manifestarsi in altri (schema delle *prove ripetute*); tecnicamente, si costruisce una successione indipendente di indicatori.

Ne tratteremo subito una, definita sullo spazio discreto numerabile fornito dai naturali.

Si noti che questo facilissimo risultato di convergenza in media quadratica è sufficiente a dedurre il lemma di equivalenza delle ridotte e quindi tutti i successivi sviluppi del capitolo, senza dover utilizzare la legge dei grandi numeri.

Appendice

Vale anche la pena di riportare una dimostrazione elementare diretta della dicotomia di Hewitt e Savage, che non richiede il lemma delle ridotte né la legge dei grandi numeri.

Sia X_n una successione indipendente di variabili aleatorie a valori nel comune spazio d'arrivo E , aventi tutte la stessa legge; per f misurabile limitata su E poniamo $c(f) = P[f(X_n)]$.

Se un evento simmetrico S ha probabilità non nulla, facciamo vedere che si ha $P(S^c) = 0$. Siano f, g misurabili limitate su E , E^{k-1} rispettivamente; dal momento che la legge della successione non si altera per permutazione e che l'evento S è simmetrico, sussiste la relazione

$$P[Isq(X_1, \dots, X_{k-1})f(X_k)] = P[Isq(X_1, \dots, X_{k-1})f(X_{k+i})]$$

e il secondo membro si scrive $P[Isq(X_1, \dots, X_{k-1})\{\frac{1}{n} \sum_{i=1}^n f(X_{k+i})\}]$, pur di sommare e dividere per n .

Le funzioni in parentesi graffa tendono a $c(f)$ in L^2 per $n \rightarrow \infty$, dato che le variabili aleatorie $U_i = f(X_{k+i}) - c(f)$ sono tra loro ortogonali per effetto dell'indipendenza delle X_n ; e $P(U_i^2)$ non dipende da i ; dunque si ha $P[(U_1 + \dots + U_n)^2] = nP(U_1^2)$, sicché $\frac{1}{n}(U_1 + \dots + U_n)$ tende a zero in norma quadratica. Dato che Isq è limitata, se ne trae

$$P[Isq(X_1, \dots, X_{k-1})f(X_k)] = c(f)P[Isq(X_1, \dots, X_{k-1})]$$

Per induzione si stabilisce la relazione

$$P[Isf_1(X_1) \dots f_k(X_k)] = c(f_1) \dots c(f_k)P(S)$$

per $f_1, \dots, f_k$ limitate arbitrarie. Di conseguenza, la speranza rispetto alla probabilità relativa P_S di $f_1(X_1) \dots f_k(X_k)$ non dipende dalla scelta di S . Prendendo per f_i indicatorici opportune, si trova che le probabilità P e P_S coincidono sugli eventi della forma $\{X_1 \in A_1, \dots, X_k \in A_k\}$; dunque sulla tribù generata dalla successione X_n . Se ne conclude $P(S^c) = P_S(S^c) = 0$.

Sull'argomento delle famiglie scambiabili torneremo più avanti: nel capitolo sulle leggi reali mostreremo come costruire arbitrarie successioni scambiabili di indicatorici; in quello dedicato alla nozione di legge condizionale stabiliremo risultati che generalizzano il teorema di de Finetti.

CAPITOLO 6

Tempi d'arresto

Sia I una parte non vuota della retta estesa (sarà abitualmente un intero contenuto in $\mathbb{R}_+$ oppure l'insieme $\mathbb{N}$); una *filtrazione* (di indice in I) su una famiglia crescente $\{\mathcal{F}_t\}_{t \in I}$ di tribù su Ω , tale cioè che $\mathcal{F}_s \subset \mathcal{F}_t$ quando t sono elementi di I con $s \leq t$, alla filtrazione sarà associata la tribù $\mathcal{F}$ minima contenente $\mathcal{F}_t$ per ogni t in I . Porremo poi $I = I \cup \{+\infty\}$.

La nozione di filtrazione formalizza l'idea dell'asimmetria dello scorrere tempo: pensato I come l'insieme dei tempi di osservazione, la tribù $\mathcal{F}_t$ presenta gli eventi conoscibili entro l'istante t nel corso dell'osservazione. Un'applicazione $T: \Omega \rightarrow \bar{I}$ si dirà *opzionale* (ovvero che è un *tempo resto*) se, per ogni $t \in I$, si ha $\{T \leq t\} \in \mathcal{F}_t$; la nozione esprime l'ide tempo di irruzione di un fenomeno non anticipativo, tale cioè che ad istante sia possibile decidere se il fenomeno ha avuto luogo o meno sulla scorta dell'osservazione trascorsa.

Dalla relazione $\{S \wedge T \leq t\} = \{S \leq t\} \cup \{T \leq t\}$ viene che $S \wedge T$ è opzionale se tanto S che T lo sono. È chiaramente opzionale un'applicazione costante: per ogni $t \in I$ l'insieme $\{T \leq t\}$ è infatti totale oppure vuoto.

Sia $\mathcal{F}_T$ la classe degli elementi A di $\mathcal{F}^\infty$ per i quali, per ogni $t \in I$, l'insieme $A \cap \{T \leq t\}$ è in $\mathcal{F}_t$; se T è opzionale, è una tribù; è infatti non vuota (con Ω), visibilmente chiusa per riunione numerabile e altresì per complementazione, dato che $A^c \cap \{T \leq t\} = \{T \leq t\} \setminus (A \cap \{T \leq t\})$. Intuitivamente, tribù degli eventi anteriori a T .

Vanno sottolineate un paio di osservazioni.

La tribù $\mathcal{F}_T$ rende misurabile l'applicazione T : in effetti per a e t l'insieme $\{T \leq a\} \cap \{T \leq t\} = \{T \leq a \wedge t\}$ appartiene a $\mathcal{F}_{a \wedge t}$ e quindi a $\mathcal{F}_T$.

Inoltre, la monotonia delle tribù vale anche per tempi opzionali, senso che da $S \leq T$ segue $\mathcal{F}_S \subset \mathcal{F}_T$: per $A \in \mathcal{F}_S$ e $t \in I$ si ha in $A \cap \{T \leq t\} = (A \cap \{S \leq t\}) \cap \{T \leq t\}$.

Comode condizioni equivalenti sussistono quando l'insieme dei tempi è numerabile.

6.1 TEOREMA. *Sia I numerabile. Affinché T sia opzionale occorre e basta che sia $\{T = r\} \in \mathcal{F}_r$ per ogni $r \in I$; in tal caso un elemento A di $\mathcal{F}^\infty$ appartiene a $\mathcal{F}_T$ se e solo se $A \cap \{T = r\} \in \mathcal{F}_r$ per ogni $r \in I$.*

Dimostrazione. Risulta dalla relazione $\{T = r\} = \{T \leq r\} \setminus \bigcup_{t < r} \{T \leq t\}$, e dalla analoga $\{T \leq t\} = \bigcup_{r \leq t} \{T = r\}$; in entrambe intervengono riunioni numerabili. $\square$

Una famiglia $X = \{X_t\}_{t \in I}$ di variabili aleatorie definite su uno spazio misurabile $(\Omega, \mathcal{A})$ si dice *adattata* alla filtrazione $\{\mathcal{F}_t\}_{t \in I}$ su Ω se, per ogni $t \in I$, l'applicazione X_t è misurabile per $\mathcal{F}_t$. La minima filtrazione che rende adatta una famiglia è detta la sua *filtrazione naturale*.

Se gli elementi X_t della famiglia assumono valori in un medesimo spazio misurabile e se $T: \Omega \rightarrow I$ è dappertutto diversa da $+\infty$, è ben definita l'applicazione $\omega \mapsto X_{T(\omega)}(\omega)$, che scriveremo X_T . Non c'è ragione perché X_T goda di particolari proprietà di misurabilità; è pertanto significativo

6.2 TEOREMA. *Se X è adattata, T opzionale finita e I numerabile, allora X_T risulta misurabile per $\mathcal{F}_T$.*

Dimostrazione. Per ogni insieme misurabile B dello spazio in cui prendono valori le X_t e per ogni $r \in I$ l'insieme

$$\{X_T \in B\} \cap \{T = r\} = \{X_r \in B\} \cap \{T = r\}$$

appartiene a $\mathcal{F}_r$; poiché $T < +\infty$, si ha poi

$$\{X_T \in B\} = \bigcup_{r \in I} (\{X_T \in B\} \cap \{T = r\}),$$

il che prova l'appartenenza a $\mathcal{F}^\infty$. $\square$

L'ipotesi di numerabilità per I è essenziale: per $I = \mathbb{R}_+$ sia $\Omega = I$ e $\mathcal{A}$ la tribù dei boreliani di $\mathbb{R}_+$; prendiamo una funzione reale h non boreliana su Ω e sia T l'identità di Ω in I . Allora T è finita e opzionale per la filtrazione costante $\mathcal{F}_t = \mathcal{A}$, la famiglia $X_t = h(t)$ è adattata e tuttavia X_T è la funzione h , non misurabile rispetto a $\mathcal{A}$.

Data una famiglia $X = \{X_t\}_{t \in I}$ e data T opzionale, la famiglia definita da $Y = \{X_{T \wedge t}\}_{t \in I}$ si dice ottenuta *arrestando* X al tempo T , e si indica con il simbolo $X^{|T}$. Chiaramente X e $X^{|T}$ coincidono su $\{T = +\infty\}$.

Risultati importanti si ottengono poi quando $I = \mathbb{N}$.

6.3 TEOREMA. *Sia X_n una successione adattata di variabili aleatorie a valori in $(E, \mathcal{E})$, sia $H \in \mathcal{E}$ e S opzionale. Allora $\inf\{k \in \mathbb{N} : k \geq S, X_k \in H\}$ (tempo di primo ingresso di X in H dopo l'istante S) è opzionale.*

Dimostrazione. Si ha $\{T \leq t\} = \bigcup_{s \leq r \leq t} \{S = s\} \cap \{X_r \in H\}$, d'ogni parte non vuota di $\mathbb{N}$ ha un minimo. $\square$

Successioni scambiabili e tempi d'arresto

Se T è finita opzionale per la filtrazione naturale di una successione $\{X_n\}$ di variabili aleatorie a valori nello stesso spazio, la legge di X_T non è che modo semplice a quelle delle X_n (nemmeno quando queste ultime hanno la stessa legge).

Se la successione è scambiabile, valgono però conseguenze strutturali come il seguente teorema, detto di *isonomia per campionamento*.

Intuitivamente, esso esprime il fatto che se si osserva una successione di istanti successivi a tempi opzionali il risultato è una successione che ha la stessa legge di quella data in partenza.

Per dare sempre senso alla scrittura X_T , converremo di porla uguale a una costante arbitraria sull'insieme $\{T = \infty\}$; quando T sia quasi certamente finita, la scelta della costante sarà irrilevante per la legge di X_T .

6.4 TEOREMA. *Sia X_n una successione scambiabile, T_k una successione di tempi opzionali per la filtrazione naturale, quasi certamente finiti, e sia $T_k < T_{k+1}$ su $\{T_k < \infty\}$. Allora la successione X_{T_k+1} è isonoma a quella data da X_n .*

Dimostrazione. Introdotte le abbreviazioni $Y_n = X_{T_n+1}$ e $T = T_{n+1}$ lo spazio in cui prendono valori le X_n , fissiamo n e sia f l'indicatrice parte misurabile di E^n e g quella di una parte misurabile di E ; sia a un intero positivo e Z una variabile aleatoria terminale.

Dato che $Y_{n+1} = X_{T_{n+1}}$ su $\{T = t\}$, si ha

$$P(f(Y_1, \dots, Y_n)g(Y_{n+1})Z I_{\{T \leq p\}}) = \sum_{t=0}^p P(f(Y_1, \dots, Y_n)g(X_{t+1})Z I_{\{T=t\}})$$

Per le proprietà dei tempi opzionali e la monotonia stretta delle T_k , le variabili Y_i misurabili per $\mathcal{F}_{T_i+1}$, sono tutte misurabili per $\mathcal{F}_T$ per $i \leq n$, e $f(Y_1, \dots, Y_n)I_{\{T=t\}}$ è indicatrice dell'intersezione con $\{T = t\}$ di un elemento di $\mathcal{F}_T$, cioè indicatrice di un elemento di $\mathcal{F}_t$; pertanto è funzione misurabile di $X_1, \dots, X_t$; per $r > p$, la variabile Z si può d'altra parte esprimere come funzione misurabile delle sole X_i con indice maggiore di r .

Dato che per ogni $t \leq p$ la successione $\{X_1, \dots, X_t, X_{t+1}, \dots\}$ ha la legge della successione X_n , negli addendi della somma sull'indice t possiamo sostituire X_{t+1} con il termine X_T , che non dipende da t ; sommando si

$$P(f(Y_1, \dots, Y_n)g(X_T)Z I_{\{T \leq p\}}) =$$