Kai Lai Chung

Sul problema del ritorno all'equilibrio

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti Lincei. Matematica e Applicazioni, Serie 9, Vol. **10** (1999), n.3, p. 213–218.

Accademia Nazionale dei Lincei

<http://www.bdim.eu/item?id=RLIN_1999_9_10_3_213_0>

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti Lincei. Matematica e Applicazioni, Accademia Nazionale dei Lincei, 1999.

Rend. Mat. Acc. Lincei
s. 9, v. 10:213-218 (1999)

Calcolo delle probabilità. — *Sul problema del ritorno all'equilibrio.* Nota di Kai Lai Chung, presentata (*) dal Corrisp. G. Letta.

Abstract. — *On the return to equilibrium.* We consider, on the group of integers, a random walk starting from the origin and whose steps admit as possible values exactly two integers, a and b, with $a < 0 < b$. In the particular case $a = -1$, we give an explicite expression for the law of the first return time to the origin.

Key words: Random walk; Taboo probabilities; Ballot problem.

Riassunto. — Si considera, sul gruppo degli interi, una passeggiata aleatoria uscente dall'origine, i cui passi ammettano due soli possibili valori: uno strettamente negativo, l'altro strettamente positivo. Nel caso particolare in cui il primo di questi valori sia -1, si dà un'espressione esplicita per la legge del primo istante di ritorno nell'origine.

1. Introduzione ed enunciato del risultato principale

L'oggetto della presente *Nota* è un problema classico, che si riallaccia a temi ampiamente studiati sin dagli inizi del Calcolo delle probabilità, come quello tradizionalmente noto sotto il nome di «rovina del giocatore». Preferiamo enunciare il nostro problema in termini più moderni, ossia nel linguaggio delle passeggiate aleatorie.

Siano a, b due interi con $a < 0 < b$, e sia p un numero reale con $0 < p < 1$. Assegnata, su uno spazio probabilizzato $(\Omega, \mathcal{A}, P)$, una successione $(X_n)_{n\geq 1}$ di variabili aleatorie indipendenti e isonome, a valori in $\{a, b\}$, con

$$P\{X_1 = a\} = p, \qquad P\{X_1 = b\} = 1 - p = q,$$

poniamo $S_n = \sum_{k=1}^n X_k$ (in particolare, $S_0 = 0$). La successione $(S_n)_{n\geq 0}$ così definita è una passeggiata aleatoria (sul gruppo degli interi), uscente dall'origine. L'istante del suo primo ritorno nell'origine è rappresentato dalla variabile aleatoria T così definita:

$$T(\omega) = \inf\{n : n \geq 1, \ S_n(\omega) = 0\}$$

(con la convenzione $\inf \emptyset = +\infty$). Il problema del *ritorno all'equilibrio* (si veda [1, p. 48]) consiste nello studiare la legge di T. Una risposta molto semplice si ha nel caso *simmetrico*, cioè nel caso in cui sia $a = -1, \ b = 1, \ p = 1/2$. È noto infatti che in questo caso, nel quale si ha ovviamente $P\{T = 2n + 1\} = 0$ per ogni intero $n \geq 0$, una celebre formula di Désiré André fornisce, per $n \geq 1$,

$$P\{T = 2n\} = \frac{1 \cdot 3 \cdots (2n-3)}{2 \cdot 4 \cdots (2n-2)(2n)} \tag{1.1}$$

(in particolare, $P\{T = 2\} = 1/2, \ P\{T = 4\} = 1/8$).

(*) Nella seduta del 23 aprile 1999.

È curioso che nell'immenso archivio della probabilità classica, da De Moivre a Borel, non si trovi nessuna traccia d'una discussione del problema per altri valori dei parametri a, b, p: neanche, ad esempio, per $a = -1$, $b = 2$, $p = 2/3$. Allo scopo di colmare questa lacuna, annunziamo il seguente risultato parziale:

TEOREMA 1.2. *Si particolarizzino le ipotesi iniziali col supporre $a = -1$. Si ha allora, per ogni intero $n \geq 1$,*

$$P\{T = n(b+1)\} = \frac{b}{n(b+1)-1}\binom{n(b+1)}{nb} p^{nb} q^{n}. \tag{1.3}$$

(Il lettore potrà facilmente verificare che questa eguaglianza si riduce effettivamente alla (1.1) nel caso simmetrico).

2. RISULTATI PRELIMINARI

Tornando al caso generale, e indicando con c il massimo comun divisore dei due interi a, b, osserviamo che l'istante del primo ritorno nell'origine non muta se si sostituisce la passeggiata aleatoria $(S_n)_{n\geq 0}$ con la passeggiata aleatoria $(c^{-1}S_n)_{n\geq 0}$. Ci si può dunque limitare a studiare il problema del ritorno all'equilibrio nel caso particolare in cui i due interi a, b siano primi tra loro. In questa ipotesi, che è automaticamente verificata nel caso particolare del Teorema 1.2 (e che sarà sempre sottintesa nel seguito), se si pone

$$d = b - a = b + |a|,$$

si vede facilmente che, per ogni intero n non divisibile per d, si ha $P\{S_n = 0\} = 0$ (e quindi $P\{T = n\} = 0$). Si ha poi, per ogni intero $k \geq 1$,

$$P\{S_{kd} = 0\} = \binom{kd}{kb} p^{kb} q^{-ka}. \tag{2.1}$$

Inoltre, per ogni intero $n \geq 1$, dalla relazione

$$\{S_{nd} = 0\} = \bigcup_{k=1}^{n}\{T = kd,\ S_{nd} = 0\} = \bigcup_{k=1}^{n}\{T = kd,\ S_{nd} - S_{kd} = 0\}$$

discende

$$P\{S_{nd} = 0\} = \sum_{k=1}^{n} P\{T = kd\}\, P\{S_{(n-k)d} = 0\}. \tag{2.2}$$

Le formule (2.1) e (2.2) permettono di ricavare per ricorrenza il valore della probabilità $P\{T = nd\}$. (Basta per questo osservare che, per $n = 1$, si ha semplicemente $P\{T = d\} = P\{S_d = 0\}$). Il pregio della formula (1.3) è che essa fornisce invece, sia pure in un caso particolare, un'espressione *esplicita* della probabilità in questione.

Sempre rimanendo nel caso generale, vogliamo ora provare alcuni risultati preliminari, interessanti in sé, che si riveleranno utili per dimostrare il Teorema 1.2. Poniamo $H_n = \{S_{nd} = 0\}$, e osserviamo che, secondo la misura di probabilità $P_{H_n}(\cdot) = P(\cdot \mid H_n)$, la legge congiunta delle X_k con $1 \leq k \leq nd$ non dipende dal parametro p. Ne segue che la probabilità condizionale

$$P_{H_n}\{T = nd\} = P\{T = nd\}/P(H_n)$$

non dipende da p. D'altra parte, nelle ipotesi del Teorema 1.2, la relazione (1.3) da dimostrare si può scrivere nella forma equivalente

$$P\{T = nd\} = \frac{b}{nd-1}\binom{nd}{nb}p^{nb}q^n = \frac{b}{nd-1}P(H_n). \tag{2.3}$$

Si vede dunque che, se questa relazione è vera per un certo valore di p, essa è vera per *ogni* valore di p. Sarà perciò lecito limitarsi a dimostrarla per quel particolare valore di p che rende centrata la legge delle X_k. (Questa condizione di centratezza, nell'interpretazione tradizionale delle X_k come guadagni aleatori di un giocatore in una successione di partite, traduce l'*equità* del gioco).

Tornando al caso generale, osserviamo che l'evento $\{T = nd\}$ è la riunione dei due eventi (tra loro disgiunti)

$$\{T \geq nd,\ S_{nd-1} = -a,\ X_{nd} = a\}, \quad \{T \geq nd,\ S_{nd-1} = -b,\ X_{nd} = b\}.$$

Perciò la sua probabilità è data da

$$P\{T = nd\} = p\,P\{T \geq nd,\ S_{nd-1} = -a\} + q\,P\{T \geq nd,\ S_{nd-1} = -b\}. \tag{2.4}$$

Sussiste inoltre il lemma seguente:

LEMMA 2.5. *Si supponga soddisfatta la condizione di equità $ap + bq = 0$. Si ha allora*

$$P\{T = nd\} = P\{T \geq nd,\ S_{nd-1} = -a\} = P\{T \geq nd,\ S_{nd-1} = -b\}. \tag{2.6}$$

DIMOSTRAZIONE. Grazie alla (2.4), basta provare la seconda delle due eguaglianze (2.6). A questo scopo, cominciamo con l'osservare che, sfruttando l'ipotesi di equità, si ottiene, con facili conti,

$$P\{S_{kd-1} = -a\} = P\{S_{kd-1} = -b\} \quad \text{per} \quad k \geq 1. \tag{2.7}$$

Si ha inoltre

$$\begin{aligned} P\{T \geq nd,\ S_{nd-1} = -a\} &= P\{S_{nd-1} = -a\} - \sum_{k=1}^{n-1} P\{T = kd,\ S_{nd-1} = -a\} = \\ &= P\{S_{nd-1} = -a\} - \sum_{k=1}^{n-1} P\{T = kd\}\,P\{S_{(n-k)d-1} = -a\}, \end{aligned} \tag{2.8}$$

e una relazione analoga, con b in luogo di a. Grazie alla (2.7), ne segue la conclusione.

(2.9) OSSERVAZIONE. La formula (2.8) non è che un caso particolare di una formula riguardante le cosiddette *probabilità tabù*, valida per una catena di Markov (si veda [2, p. 46]).

3. Dimostrazione del risultato principale

Siamo ora in grado di dimostrare il Teorema 1.2. Supponiamo dunque $a = -1$. Supponiamo inoltre verificata la condizione di equità (che si scriverà ora nella forma $p = bq$, ovvero $q = 1/d$). Dal Lemma 2.5 discende allora

$$P\{T=nd\}=P\{S_{nd-1}=-b,\ T\geq nd\}=P\{S_{nd-1}=-b,\ S_k<0 \text{ per } 1\leq k\leq nd-1\}, \tag{3.1}$$

dove l'eguaglianza finale è dovuta al fatto che la nostra passeggiata aleatoria (i cui soli passi discendenti sono di ampiezza unitaria) non può, dopo aver visitato uno stato positivo, raggiungere lo stato $-b$ senza prima incontrare l'origine. Poniamo

$$X_j' = I_{\{X_j=-1\}}, \qquad S_k' = \sum_{j=1}^{k} X_j'.$$

Si ha allora

$$S_k = -S_k' + b\,(k - S_k') = kb - d\,S_k'.$$

Perciò la relazione (3.1) si può così riscrivere:

$$P\{T = nd\} = P\{S_{nd-1}' = nb,\ S_k' > b\,(k - S_k') \text{ per } 1\leq k\leq nd-1\}. \tag{3.2}$$

Fortunatamente, il secondo membro di quest'ultima relazione si può calcolare in modo esplicito applicando un classico risultato combinatorio riguardante il cosiddetto *problema dello scrutinio*. Per comodità del lettore, richiameremo brevemente questo risultato. Consideriamo un ballottaggio tra due candidati, i quali abbiano ottenuto rispettivamente R e N voti. Ognuno dei possibili modi di eseguire lo scrutinio si può rappresentare mediante una sequenza finita di elementi di $\{0, 1\}$, della forma

$$(x_k)_{1\leq k\leq R+N}, \quad \text{con} \quad \sum_{k=1}^{R+N} x_k = R. \tag{3.3}$$

(Naturalmente è da intendere che, per ogni k, l'elemento x_k valga 1 o 0 secondo che il k-esimo voto scrutinato sia a favore del primo o del secondo dei due candidati). Si supponga ora $R > \alpha N$, con α intero strettamente positivo, e si consideri l'insieme degli scrutini nei quali, a ciascun istante, il punteggio parziale del primo candidato risulti superiore ad α volte quello dell'avversario. Formalmente, si tratta dell'insieme costituito da tutte le sequenze di elementi di $\{0, 1\}$ della forma (3.3) per le quali è verificata l'ulteriore condizione seguente:

$$\sum_{j=1}^{k} x_j > \alpha \left(k - \sum_{j=1}^{k} x_j \right) \text{ per } 1\leq k\leq R+N.$$

Il cardinale di questo insieme, come ha dimostrato Aeppli nel 1924, è eguale a

$$\frac{R-\alpha N}{R+N}\binom{R+N}{R}.$$

(Per la dimostrazione, e per qualche notizia storica, si veda [4]).

Tornando ora alla nostra dimostrazione, osserviamo che il secondo membro di (3.2) è eguale alla somma dei termini della forma

$$P\left(\bigcap_{k=1}^{nd-1}\{X'_k = x_k\}\right), \tag{3.4}$$

corrispondenti a tutte le possibili sequenze finite di elementi di $\{0\,,1\}$, della forma $(x_k)_{1\le k\le nd-1}$, verificanti le condizioni

$$\sum_{j=1}^{nd-1} x_j = nb\,, \quad \sum_{j=1}^{k} x_j > b\left(k - \sum_{j=1}^{k} x_j\right) \quad \text{per} \quad 1 \le k \le nd-1.$$

Il numero m di queste sequenze si calcola applicando il risultato combinatorio sopra ricordato, nel quale si prenda $R=nb\,,\ N=n-1\,,\ \alpha=b$. Si trova così:

$$m = \frac{nb - b(n-1)}{nb + n-1}\binom{nb + n-1}{nb} = \frac{b}{nd-1}\binom{nd-1}{nb} = \frac{b}{nd-1}\binom{nd}{nb}\frac{1}{d}\,.$$

È poi chiaro che, per ognuna delle sequenze in questione, la probabilità (3.4) è eguale a $p^{nb}q^{n-1}$. Dalla relazione (3.2) si deduce dunque (ricordando l'eguaglianza $q = 1/d$):

$$P\{T = nd\} = m\cdot p^{nb}q^{n-1} = \frac{b}{nd-1}\binom{nd}{nb}p^{nb}q^{n}.$$

La relazione (2.3) è così dimostrata.

4. Notizie storiche

La mia scoperta della formula (1.3) avvenne in Cina, e precisamente a Kunming, intorno al 1940, cioè nel periodo assai duro dell'invasione straniera della Cina. Essa fu fortemente influenzata dalla lettura del celebre libro di É. Borel [1]. Intuii quella formula esaminando i casi particolari $b = 2$ e $b = 3$, cioè mediante una sorta di «induzione incompleta». La annunciai poi nel 1946, al momento del mio ingresso come studente all'Università di Princeton, anche sotto la forma della seguente identità combinatoria:

$$\binom{bn}{n} = \sum_{k=1}^{n}\frac{b-1}{bk-1}\binom{bk}{k}\binom{b(n-k)}{n-k}\,.$$

Appunto sotto questa forma essa fu verificata da H. W. Gould nel 1956, con l'ausilio di un'altra identità combinatoria (associata ai nomi di Rothe, Schläffi e Hagen). Per i particolari tecnici e per i riferimenti bibliografici, il lettore potrà consultare [3, 4].

In seguito mi dimenticai della questione e persi ogni traccia dei miei giovanili appunti di Kunming. Soltanto nel 1993 tornai ad occuparmi del problema. Studiando numericamente la questione, fui colpito da alcune singolari simmetrie, le quali m'indussero a sospettare che valesse l'affermazione del Lemma 2.5. Dopo che ulteriori esperimenti numerici ebbero rafforzato il mio sospetto, trasformandolo in vero e proprio convincimento, mi misi alla ricerca di una dimostrazione. Sono contento di constatare che, a farmi raggiungere lo scopo, ha contribuito in modo determinante l'idea delle *probabilità tabù*: cioè la stessa idea che, in anni lontani, mi era stata di grande aiuto nello studio delle catene di Markov.

Vorrei da ultimo aggiungere che, se mi sono deciso a pubblicare qui il risultato parziale espresso dal Teorema 1.2, è soprattutto perché spero che ciò possa spronare qualcuno a cimentarsi col caso generale, ossia a ricercare, per la legge di T, un'espressione esplicita che sia valida anche quando a sia un arbitrario intero minore di zero.

Bibliografia

[1] É. Borel, *Valeur pratique et philosophie des probabilités*. Gauthier-Villars, Paris 1939.
[2] Kai Lai Chung, *Markov chains with stationary transition probabilities*. Second edition, Springer-Verlag, Berlin - Heidelberg - New York 1967.
[3] R. Guy, *A pentagonal pot-pourri of perplexing problems, primarily probabilistic*. American Mathematical Monthly, 91, 1984, 559-563.
[4] L. Takács, *On the ballot problem*. Advances in combinatorial methods and applications. Birkhäuser, Boston 1997.

Pervenuta il 26 marzo 1999,
in forma definitiva il 19 aprile1999.

Department of Mathematics
Stanford University
Stanford, California 94305-2125 (U.S.A.)