Luigi Castoldi

Identità generalizzata di Wald nella teoria probabilistica delle sequenze casuali.

Bollettino dell'Unione Matematica Italiana, Serie 3, Vol. **11** (1956), n.1, p. 22–27.

Zanichelli

<http://www.bdim.eu/item?id=BUMI_1956_3_11_1_22_0>

Identità generalizzata di Wald nella teoria probabilistica delle sequenze casuali.

Nota di LUIGI CASTOLDI (a Cagliari)

Sunto. - *Si generalizza, utilizzando anche un procedimento deduttivo semplificato, una nota identità di* WALD *relativa a sequenze di variabili stocastiche identiche e indipendenti. A tale risultato si perviene contemplando il caso generale di variabili aleatorie* $X_i (i = 1, 2, ...)$ *indipendenti e dotate di leggi di probabilità a priori diverse. In relazione alla variabile cumulativa* $S_N = \sum_{j=1}^{N} X_j$, *si deduce poi la completa legge di probabilità per la* scadenza N *di raggiungimento di estremi fissati.*

1. Richiami ed ipotesi fondamentali.

Il classico problema probabilistico della "passeggiata a caso" con percorso finito ed esiziale agli estremi, nonchè quello delle vincite-perdite al giuoco di due contendenti con alternative di possibile rovina per ciascuno di essi, costituiscono come è noto, i più semplici esempi di un problema generale che può enunciarsi nel modo seguente.

Sia $\{X_i\}$ una successione illimitata di variabili aleatorie, con corrispondenti assegnate distribuzioni di probabilità a due a due indipendenti. Siano $F_i(x)$ le corrispondenti distribuzioni di frequenza, ed

$$M_i(\vartheta) = \int_{-\infty}^{\infty} e^{\vartheta x} dF_i(x) \tag{1}$$

le relative funzioni momento-generatrici.

Di quest'ultime supponiamo che sia, per ogni valore di i:

$$\boxed{M_i(\vartheta) \to \infty \quad \text{per} \quad \vartheta \to \pm\infty}, \tag{2}$$

conseguendone similmente, pur per ogni valore di i:

$$\overset{(i)}{M}(\vartheta) \equiv \prod_{j=1}^{i} M_i(\vartheta) \to \infty \quad \text{per} \quad \vartheta \to \pm\infty. \tag{3}$$

È ovviamente

$$M_i(0) - 1 = 0 \qquad (i = 1, 2, ...) \tag{4}$$

e pertanto

$$\overset{(i)}{M}(0) - 1 = 0 \qquad (i = 1, 2, ...), \tag{4'}$$

ed è pure, introdotte le medie m_i delle variabili X_i,

(5) $$m_i = M_i'(0) \equiv \left[\frac{d}{d\vartheta} M_i(\vartheta)\right]_{\vartheta=0}.$$

Di qui segue

(6) $$\left[\frac{d}{d\vartheta} \overset{(i)}{M}(\vartheta)\right]_{\vartheta=0} = \sum_{j=1}^{i} M_j'(0) = \sum_{1=j}^{i} m_j,$$

col risultato, che ove sia l'ultima somma diversa da zero, l'ipotesi (3) assicura esistente per l'equazione

(7) $$\overset{(i)}{M}(\vartheta) - 1 = 0$$

— oltre alla radice $\vartheta = 0$ — una seconda, ϑ_{0i}, dipendente generalmente da i e definita quale valore reale di ϑ per cui è insieme

(8) $$\overset{(i)}{M}(\vartheta_{0i}) - 1 = 0$$

e

(9) $$\underset{(i)}{M}(\vartheta) < 1 \quad \text{per } \vartheta \text{ interno all'intervallo } (0, \vartheta_{0i}).$$

Da quest'ultima condizione e da (6) si trae subito la disuguaglianza

(10) $$\left(\sum_{j=1}^{i} m_j\right)\vartheta_{0i} \leqq 0$$

atta a definire, per ogni valore di i per cui non sia $\vartheta_{0i} = 0$, il segno di questa grandezza.

In queste ipotesi (2), di ampia generalità per le variabili X_i, mi propongo di stabilire, attraverso un semplice procedimento deduttivo, una notevole generalizzazione della nominata identità di WALD (1), perseguendo anche l'intento di metterne in luce taluni, credo ancora nascosti, aspetti concettuali.

2. L'identità generalizzata di Wald.

Colla posizione

(11) $$S_i = \sum_{j=1}^{i} X_j$$

è definita la *variabile cumulativa* S_i.

(1) L'identità originale trovasi in A. WALD, *Sequential Analysis*, New York (1947). Essa è riprodotta nella recente opera di M. S. BARTLETT, *An Introduction to Stochastic Processes*, Cambridge (1955).

L'indipendenza delle X_i consente di esprimere la funzione momento-generatrice della distribuzione di S_i come segue:

$$M_{S_i}(\mathfrak{z}_i) \equiv E_X \{ \exp(\mathfrak{z}_i S_i) \} = \prod_{j=0}^{i} M_j(\mathfrak{z}_i) = \overset{(i)}{M}(\mathfrak{z}_i), \tag{12}$$

ove è lecito, ovviamente, concepire distinte le variabili $\mathfrak{z}_i$ in corrispondenza ai diversi valori di i.

Per il seguito, converrà scrivere l'ultimo risultato nella forma seguente:

$$E_X \{ \exp(\mathfrak{z}_, S_i) \} \overset{(i)}{M}{}^{-1}(\mathfrak{z}_,) = 1, \tag{13}$$

Fino a questo punto, l'indice i ha avuto un ruolo puramente parametrico. Ma è possibile attribuire ad esso il significato di variabile aleatoria, con relativa distribuzione di probabilità, introducendo dapprima due estremi a e b ($a < b$ e, come sempre potremo supporre, $a < 0$, $b > 0$) alla variabilità di S_i, indi *caratterizzando il valore di* i *come quello per cui* Si *raggiunge o supera, per la prima volta, l'uno o l'altro degli estremi assegnati.*

In tal senso scriveremo N in luogo di i ma varrà pur sempre la relazione (13) nella forma

$$E_X \{ \exp(\mathfrak{z}_N S_N) \} \overset{(N)}{M}{}^{-1}(\mathfrak{z}_N) = 1 \qquad (N = 1,\ 2,\ \ldots). \tag{14}$$

Immaginata nota la distribuzione di N, prendendo in (14) il valor medio di entrambi i membri su questa variabile, e denotando con $P(\leq 1)$ la probabilità che S_i superi, quando che sia, l'uno o l'altro dei limiti imposti, otteniamo alla fine:

$$\boxed{E_N[E_X \{ \exp(\mathfrak{z}_N S_N) \} \overset{(N)}{M}{}^{-1}(\mathfrak{z}_N)] = P}\ . \tag{15}$$

È questa, concepita quale identità nelle variabili $\mathfrak{z}_N (N = 1, 2, \ldots)$, la annunciata generalizzazione dell'*identità di* WALD; stabilita originariamente da quell'Autore per variabili X_i tutte dotate di una medesima distribuzione, sotto ipotesi afferenti al valore 1 di P, e limitatamente a una scelta di variabili $\mathfrak{z}_N$ tutte coincidenti.

In tali circostanze, l'identità (15) diventa di fatto

$$\left[\begin{array}{l} E_N[E_X \{ \exp(\mathfrak{z} S_N) \} M^{-N}(\mathfrak{z})] = 1 \\ \qquad M(\mathfrak{z}) = M_N(\mathfrak{z}) \qquad\qquad (N = 1,\ 2,\ \ldots). \end{array}\right. \tag{15'}$$

3. Valori medi condizionali subordinati all'esatto raggiungimento di un estremo.

Dalla identità (15) può trarsene una analoga per valori medi *condizionali*, subordinata all'ipotesi del preventivo *esatto* raggiungimento di un estremo (a o b). Denotando con P_a e P_b le probabilità a priori che S_N si identifichi rispettivamente con a o con b, ed osservando che *ora* il valor medio su X delle variabili indicate non differisce da un corrispondente e noto valore esatto, la relazione nominata diventa ovviamente

$$(16) \qquad P_a E_{N/a}[\exp(a\vartheta_N)\overset{(N)}{M}{}^{-1}(\vartheta_N)] + P_b E_{N/b}[\exp(b\vartheta_N)\overset{(N)}{M}{}^{-1}(\vartheta_N)] = P.$$

Attribuiamo ora a ciascuna variabile ϑ_N il valor ϑ_{0N} definito da (8). Si ottiene in tal modo da (16) la relazione

$$(17) \qquad P_a E_{N/a}[\exp(a\vartheta_{0N})] + P_b E_{N/b}[\exp(b\vartheta_{0N})] = P.$$

Giova, a questo punto, associare alle distribuzioni condizionali $p_{N/a}$, $p_{N/b}$ della variabile stocastica N le analoghe *pseudo-distribuzioni*

$$(18) \qquad \tilde{p}_{N/a} = e^{a\vartheta_{0N}} p_{N/a} \quad , \quad \tilde{p}_{N/b} = e^{b\vartheta_{0N}} p_{N/b} ,$$

che dirò ϑ*-modificate*, della stessa variabile.

Si osserverà che, ritenendosi nota la funzione ϑ_{0N} di N, la conoscenza delle pseudodistribuzioni (18), una volta acquisita, conduce immediatamente a quella delle effettive distribuzioni $p_{N/a}$, $p_{N/b}$ e, ove occorra a

$$(19) \qquad p_N = P_a p_{N/a} + P_b p_{N/b} .$$

Posto ora, per tentativo,

$$(20) \qquad \left[\begin{array}{l} \sum\limits_{N=0}^{\infty} e^{a\vartheta_{0N}} p_{N/a} \equiv E_{N/a}[\exp(a\vartheta_{0N})] = Ae^{a\vartheta_0} \\ \\ \sum\limits_{N=0}^{\infty} e^{b\vartheta_{0N}} p_{N/b} \equiv E_{N/b}[\exp(b\vartheta_{0N})] = Ae^{b\vartheta_0}, \end{array}\right.$$

definendo così univocamente $A\ (> 0)$ e ϑ_0, la relazione (17) assume la forma semplicissima

$$(21) \qquad P_a e^{a\vartheta_0} + P_b e^{b\vartheta_0} = PA^{-1} .$$

Di qui, ricordando che è

$$(22) \qquad P_a + P_b = P,$$

si traggono le soluzioni:

$$\boxed{P_a = P\frac{A^{-1} - e^{b\vartheta_0}}{e^{a\vartheta_0} - e^{b\vartheta_0}} \quad , \quad P_b = P\frac{e^{a\vartheta_0} - A^{-1}}{e^{a\vartheta_0} - e^{b\vartheta_0}}} \tag{23}$$

Escludendo dapprima che sia $\vartheta_0 = 0$, anzi ritenendo, per fissare le idee, $\vartheta_0 > 0$, il carattere a priori non negativo di P_a e P_b esige che sia $e^{-b\vartheta_0} < A < e^{-a\vartheta_0}$, e pertanto

$$Ae^{a\vartheta_0} < 1 \quad , \quad Ae^{b\vartheta_0} > 1 \qquad (\vartheta_0 > 0). \tag{24}$$

Contrariamente, si riconosce che, per $\vartheta_0 < 0$, dev'essere:

$$Ae^{a\vartheta_0} > 1 \quad , \quad Ae^{b\vartheta_0} < 1 \qquad (\vartheta_0 < 0). \tag{24'}$$

Si conclude che il tentativo fatto nello scrivere le equazioni (20) riesce giustificato soltanto ove i primi membri di queste posseggano valori oppostamente situati rispetto all'unità. Fortunatamente, questa circostanza si trova effettivamente verificata, come risulta dal confronto delle relazioni (17) e (22). È pertanto assicurata la consistenza, nel caso $\vartheta_0 \neq 0$, delle formule risolutive (23).

Con $\vartheta_0 = 0$, quest'ultime esigono $A = 1$, come altresì risulta da (17) e (22).

4. La distribuzione di N definita dalla corrispondente funzione generatrice.

Supponiamo ancor dapprima $\vartheta_0 \neq 0$ e riprendiamo l'identità (16) concependovi identificate in un'unica variabile ϑ tutte le $\vartheta_N \cdot A\vartheta$ consentiamo poi l'intervallo di variabilità definito da $|\vartheta| < |\vartheta_0|$.

La detta identità diventa in tal modo

$$P_a e^{a\vartheta} E_{N/a}[\overset{(N)}{M}{}^{-1}(\vartheta)] + P_b e^{b\vartheta} E_{N/b}[\overset{(N)}{M}{}^{-1}(\vartheta)] = P. \tag{25}$$

Poniamo ancora, con φ reale,

$$P_a E_{N/a}[\overset{(N)}{M}{}^{-1}(\vartheta)] + P_b E_{N/b}[\overset{(N)}{M}{}^{-1}(\vartheta)] = P_a E_{N/a}(e^{i\varphi N}) + P_b E_{N/b}(e^{i\varphi N}), \tag{26}$$

definendo così — come risulta dal confronto di (25) con (22) e con (20), (21) — due funzioni complesse di φ, tendenti rispettivamente a 0 e a ϑ_0 allo svanire di φ:

$$\left[\begin{array}{l} \vartheta_1(\varphi) \to 0 \quad \text{per} \quad \varphi \to 0 \\ \vartheta_2(\varphi) \to \vartheta_0 \quad \text{per} \quad \varphi \to 0 . \end{array}\right. \tag{27}$$

Riscrivendo (25), una volta con $\vartheta = \vartheta_1(\varphi)$, un'altra con $\vartheta = \vartheta_2(\varphi)$, e ricordando (26), si ottengono, per le funzioni caratteristiche condizionali

$$C_{N/a}(\varphi) = E_{N/a}[e^{i\varphi N}] \quad , \quad C_{N/b} = E_{N/b} = [e^{i\varphi N}], \tag{28}$$

le seguenti due equazioni lineari

$$\text{(29)} \qquad \left|\begin{cases} P_a e^{a\vartheta_1(\varphi)} C_{N/a}(\varphi) + P_b e^{b\vartheta_1(\varphi)} C_{N/b}(\varphi) = P \\ P_a e^{a\vartheta_2(\varphi)} C_{N/a}(\varphi) + P_b e^{b\vartheta_2(\varphi)} C_{N/b}(\varphi) = P \end{cases}\right| .$$

Il complesso delle equazioni (20), (23) e (29), in cui si supponga provvisoriamente P assegnato, definisce univocamente le costanti A e ϑ_0, le probabilità a priori P_a e P_b, nonchè le funzioni caratteristiche condizionali $C_{N/a}$ e $C_{N/b}$. Tali equazioni consentono pertanto, almeno concettualmente, la risoluzione del problema della determinazione della distribuzione di N.

Va osservato che, una volta tratte da $C_{N/a}$ e $C_{N/b}$ le corrispondenti funzioni condizionali di frequenza mediante le formule di inversione

$$\text{(30)} \qquad \begin{cases} \Delta F(N/_b{}^a) = \lim\limits_{T\to\infty} \displaystyle\int_{-T}^{T} \frac{\Delta e^{-i\varphi N}}{-i\varphi} C_{N/_b{}^a}(\varphi) d\varphi \\ \Delta(\xi) \equiv \xi(x+h) - \xi(x), \end{cases}$$

la determinazione di P risulta dalla soluzione dell'equazione

$$\text{(31)} \qquad F(\infty/a) + F(\infty/b) = P.$$

Il caso $\vartheta_0 = 0$, finora escluso, comporta $\vartheta_1(\varphi) \equiv \vartheta_2(\varphi) = \vartheta(\varphi)$ e porge, in luogo di (29), il sistema

$$\text{(29')} \qquad \left|\begin{matrix} P_a e^{a\vartheta(\varphi)} C_{N/a}(\varphi) + P_b e^{b\vartheta(\varphi)} C_{N/b}(\varphi) = P \\ aP_a e^{a\vartheta(\varphi)} C_{N/a}(\varphi) + bP_b e^{b\vartheta(\varphi)} C_{N/b}(\varphi) = 0 \end{matrix}\right| .$$

Di qui si procede in modo del tutto simile a quello tenuto a partire da (29).

Va infine considerata la possibilità che l'uno o l'altro degli estremi a e b diverga, in modulo, all'infinito.

Consideriamo, per esempio, il caso in cui è $a = -\infty$, similmente procedendosi nell'eventualità rimanente.

In tale ipotesi è certo $P_a = 0$, $P_b = P$, il che è possibile, per (23), soltanto se è $\vartheta_0 < 0$. Aggiungasi che affinchè $M_i(\vartheta)$ risulti finito per ogni valore di i per cui non sia nulla la probabilità di qualche valore negativo di X_i, deve *ora* essere *non negativa* la parte reale di ϑ. A tale condizione soddisfa $\vartheta_1(\varphi)$, non $\vartheta_2(\varphi)$ che converge a $\vartheta_0 < 0$, allo svanire di φ.

Delle equazioni (29) rimane pertanto soltanto la prima a definire $C_{N/b}$ nella forma

$$\text{(32)} \qquad C_{N/b}(\varphi) = e^{-b\vartheta_1(\varphi)}.$$