CUMULANTI E INVERSIONI DI MOBIUS

Mauro CERASOLI
Università degli Studi dell'Aquila

Sunto. Mediante un'inversione di Möbius nel reticolo delle partizioni di un n-insieme può essere dimostrata la formula che esprime il cumulante n-esimo di una variabile aleatoria in funzione dei suoi momenti. Sono presentati infine alcuni problemi aperti.

1. MOMENTI E CUMULANTI

Sia X una variabile aleatoria su uno spazio di probabilità (Ω, F, P) avente momenti

$$m_k = E(X^k) = \int_\Omega X^k dP \qquad k = 0, 1, 2, \ldots$$

Associata ad m_k consideriamo la successione numerica $K_n(X)$, $n = 1, 2, \ldots$, definita dalla seguente identità:

$$(1) \qquad E(e^{tX}) = \Sigma_{j\geq 0}\, E(X^j)t^j/j! = \exp(\Sigma_{n\geq 1} K_n(X)t^n/n!)$$

dove t è una variabile formale. Il numero $K_n(X)$ viene chiamato <u>cumulante</u> n-esimo (o di ordine n) di X (o della distribuzione di probabilità di X). Introducendo l'operatore derivata D possiamo scrivere in forma più breve

$$K_n(X) = D^n \log E(e^{tX})_{|t=0}.$$

Si può controllare che i primi quattro cumulanti valgono

$K_1 = m_1 = E(X)$

$K_2 = m_2 - m_1^2 = Var(X)$

$K_3 = m_3 - 3m_2m_1 + 2m_1^3$

$K_4 = m_4 - 4m_3m_1 + 12m_2m_1^2 - 6m_1^4 + 3m_2^2$.

I cumulanti hanno notevoli applicazioni nella statistica perché generalizzano il concetto di varianza, ma la loro vera natura ci appare del tutto sconosciuta. Non è difficile dimostrare però che essi godono di tre particolari proprietà che riportiamo nel seguente teorema.

TEOREMA Il primo cumulante $K_1(X)$ coincide con la media m_1 di X e per ogni c reale si ha

a) $K_n(X+c) = K_n(X), \qquad n>1;$

b) $K_n(cX) = c^n K_n(X)$.

c) Se X e Y sono indipendenti allora

$$K_n(X+Y) = K_n(X) + K_n(Y).$$

Dim. Sostituendo X con X+c in (1) e sviluppando risulta

$$\exp(\Sigma_{n\geq 1}K_n(X+c)t^n/n!) = E(e^{t(X+c)}) = e^{tc}E(e^{tX})$$

$$= e^{tc}\exp(\Sigma_{n\geq 1}K_n(X)t^n/n!)$$

$$= \exp[t(c+K_1(X))+\Sigma_{n>1}K_n(X)t^n/n!]$$

quindi $K_1(X+c) = K_1(X)+c$, ovvero $K_1(X) = m_1$, e $K_n(X+c) = K_n(X)$ per $n>1$.

Anche la seconda proprietà è immediata perché

$$\exp(\Sigma_{n\geq 1}K_n(cX)t^n/n!) = E(e^{t(cX)}) = E(e^{(ct)X})$$

$$= \exp(\Sigma_{n\geq 1}K_n(X)c^n t^n/n!)$$

Infine la terza proprietà, da cui proviene il termine cumulante, è conseguenza del fatto che

$$E(e^{t(X+Y)}) = E(e^{tX})E(e^{tY})$$

se X e Y sono indipendenti.

Riportiamo qui di seguito, per meglio illustrare il concetto di cumulante, quanto è noto a proposito dei cumulanti di tre distribuzioni fondamentali della teoria della probabilità.

a) I primi quattro cumulanti della distribuzione binomiale di parametro p sono rispettivamente

$$K_1 = np, \qquad K_2 = np(1-p),$$

$$K_3 = np(1-p)(1-2p), \qquad K_4 = np(1-p)(1-6p+6p^2).$$

Più in generale si può dimostrare che

$$K_{n+1} = (p-p^2)\frac{dK_n}{dp}$$

e con questa formula si possono calcolare tutti i cumulanti della distribuzione binomiale;

b) la distribuzione di Poisson di parametro α ha tutti i cumulanti uguali ad α;

c) la distribuzione normale ha tutti i cumulanti nulli per $n>2$.

2. LA FORMULA DI CAUCHY

Tenendo presente la (1) consideriamo l'identità

$$(2) \qquad 1+\Sigma_{n\geq 1}a_n t^n/n! = \exp(\Sigma_{n\geq 1}b_n t^n/n!)$$

dove a_n e b_n sono variabili formali. Il problema che si presenta spontaneo è di esprimere a_n in funzione di $b_1, b_2, \ldots, b_n$ e, viceversa, di esprimere b_n in funzione di $a_1, a_2, \ldots, a_n$. Per determinare a_n basta osservare che il secondo membro della (2) si può scrivere nella forma

$$\exp(b_1t)\exp(b_2t^2/2!)\ldots\exp(b_nt^n/n!)\ldots =$$

$$= \prod_{j\geq 1} \Sigma_{k\geq 1}[(b_j t^j/j!)^k/k!]$$

$$= \prod_{j\geq 1}\Sigma_{k\geq 1}[b_j^k t^{jk}/(j!^k k!)]$$

$$= \Sigma_{n\geq 0} \sum_{i_1,i_2,\ldots,i_n\geq 0} \frac{b_1^{i_1} b_2^{i_2} \ldots b_n^{i_n} t^{i_1} t^{2i_2} \ldots t^{ni_n}}{1!^{i_1} 2!^{i_2} \ldots n!^{i_n} i_1! i_2! \ldots i_n!}$$

e quindi ricaviamo

$$a_n = \sum_{\substack{i_1,i_2,\ldots,i_n\geq 0 \\ i_1+2i_2+\ldots+ni_n=n}} \frac{n!}{1!^{i_1} \ldots n!^{i_n} i_1! \ldots i_n!} b_1^{i_1} \ldots b_n^{i_n}$$

In verità questo non è altro che un caso particolare della formula di Faà di Bruno che dà la derivata n-esima della funzione composta f(g(t)), con f uguale ad exp. I coefficenti a_n sono noti come polinomi di Bell e, ricorrendo ad artifizi di tipo umbrale, si potrebbe ricavare b_n in funzione di $a_1, a_2, \ldots, a_n$ (si veda ad esempio Riordan, pag. 36). Il coefficiente del monomio

$$b_1^{i_1} b_2^{i_2} \ldots b_n^{i_n}$$

ha un significato combinatorio che suggerisce però di risolvere diversamente il problema. Infatti sia S un insieme finito; ricordiamo che una partizione di S è una famiglia $\{A_1, A_2, \ldots, A_r\}$ di sue parti (dette classi o blocchi) non vuote, disgiunte, la cui unione è S. La partizione è detta di tipo

$$(i_1, i_2, \ldots, i_n)$$

se possiede i_k classi di cardinalità k, per ogni k = 1, 2,..., n. Allora il coefficiente del monomio su menzionato è uguale al numero di partizioni di un insieme di n elementi di tipo $(i_1, i_2, \ldots, i_n)$: un risultato che si fa risalire a Cauchy. Questa scoperta invita a ricavare i coefficienti b_n mediante una inversione di Möbius sul reticolo delle partizioni di un insieme finito. L'idea appare già nel lavoro di Doubilet-Rota-Stanley (p.101) ma, recentemente, studi più approfonditi sono stati affrontati da Speed a partire dal 1983. Vediamo pertanto come il problema proposto viene risolto in quest'altro modo.

3. UN'INVERSIONE DI MOBIUS NEL RETICOLO DELLE PARTIZIONI

Sia $S = \{1,2,\dots,n\}$; dati una partizione π di S, una classe B di π e le variabili formali $b_1, b_2, \dots, b_n$, poniamo:

a) $w(B) = b_k$ se $|B|=k$;

b) $g(\pi) = \prod_{B\in\pi} w(B) = b_1^{i_1} b_2^{i_2} \dots b_n^{i_n}$

se π è di tipo $(i_1, i_2, \dots, i_n)$. In particolare se $\hat{1}$ è l'unica partizione di S di tipo $(0,0,\dots,0,1)$ risulta ovviamente

$g(\hat{1}) = b_n$;

c) $f(\sigma) = \prod_{B\in\sigma} \sum_{\pi\in B^*} g(\pi)$

dove B^* è l'insieme delle partizioni di B. Questa identità equivale a

(3) $f(\sigma) = \sum_{\tau\leq\sigma} g(\tau)$

dove $\leq$ indica l'ordine parziale per raffinamento tra partizioni.

d) Se T è un insieme di k elementi, poniamo

$a_k = \sum_{\pi\in T^*} g(\pi)$, $|T| = k$.

Se a w si dà il significato di funzione peso, possiamo interpretare a_n come l'enumeratore di S^* (vedi Cerasoli-Eugeni-Protasi, p.131). Dalla (3), con $\hat{1}$ al posto di σ, si ricava

$$f(\hat{1}) = \Sigma_{\tau\leq\hat{1}}\ g(\tau) = \Sigma_{\pi\in S^*} g(\pi) = a_n.$$

Applicando il teorema d'inversione di Möbius alla (3) otteniamo

$$(4) \qquad g(\sigma) = \Sigma_{\tau\leq\sigma}\ \mu(\tau,\sigma) f(\tau)$$

dove μ è la funzione di Mobius del reticolo di partizioni di S^*. Per i nostri scopi è necessario conoscere il valore $\mu(\tau,\hat{1})$ dove $\hat{1}$ è la partizione massima di S. Come è noto, Schützenberger ha dimostrato che

$$\mu(\tau,\hat{1}) = (-1)^{|\tau|-1}(|\tau|-1)!.$$

Ponendo quindi $\hat{1}$ al posto di σ nel primo membro della (4), si ottiene la formula che esprime b_n in funzione di $a_1, a_2, \dots, a_n$. Questa, quando scriviamo $K_n = K_n(X)$ per b_n ed $m_k = E(X^k)$ per a_k, diventa:

$$K_n = \sum_{\substack{k=1 \\ i_1+i_2+\dots+i_n=k \\ i_1+2i_2+\dots+ni_n=n}}^{n} \frac{(-1)^{k-1}(k-1)!n!}{1!^{i_1}\dots n!^{i_n}\ i_1!\dots i_n!} m_1^{i_1}\dots m_n^{i_n}$$

In definitiva questa formula esprime i cumulanti di X in funzione dei momenti. Con essa proviamo a ricavare i cumulanti $K_n(X)$ di una variabile aleatoria X quando n = 1, 2, 3, 4, dimostrando le formule date nel primo paragrafo.

a) Per n = 1 risulta ovviamente $K_1(X) = E(X)$;

b) per n = 2 le partizioni di S = {1,2} sono 1|2 e 12 di tipi rispettivamente (2,0) e (0,1). Pertanto il secondo cumulante vale

$$K_2(X) = E(X^2)-E(X)^2 = Var(X)$$

cioè la varianza di X.

c) Le partizioni di $\{1,2,3\}$ sono:

$1|2|3$ di tipo $(3,0,0)$,

$1|23$, $12|3$, $13|2$ di tipo $(1,1,0)$

123 di tipo $(0,0,1)$.
Quindi il terzo cumulante è

$$K_3(X) = E(X^3) - 3E(X)E(X^2) + 2E(X)^3.$$

Esso è noto anche in statistica come <u>curtosi</u>.
d) Si può infine dimostrare che

$$K_4(X)=E(X^4)-4E(X)E(X^3)+12E(X^2)E(X)^2-6E(X)^4+3E(X^2)^2$$

e con questo abbiamo ottenuto i quattro cumulanti presentati all'inizio.

Se X è costante, uguale ad 1, i suoi momenti valgono tutti 1 e i cumulanti sono nulli a partire dal secondo. Allora dalla formula che dà i cumulanti otteniamo la semplice identità combinatoria

$$\sum_{k=1}^{n} (-1)^{k-1}(k-1)!S(n,k) = \delta_{n1}$$

dove $S(n,k)$ è il numero di Stirling di seconda specie, cioè il numero di partizioni di un n-insieme in k classi.

4. ALCUNI PROBLEMI APERTI

Come si è detto in precedenza, poco si sa sui cumulanti; elenchiamo pertanto alcuni problemi aperti ed interessanti.
a) Data una successione k_n, $n = 1, 2, \dots$, di numeri reali, determinare sotto quali condizioni esiste una variabile aleatoria il cui n-esimo cumulante è k_n. Il problema è collegato al classico problema dei momenti.

b) Esiste una variabile aleatoria con cumulanti non nulli solo per $n>3$?

c) Tra i coefficienti dello sviluppo in polinomi di Hermite della distribuzione di probabilità di X e i cumulanti di X, esiste una strana relazione: se X è standard (ha media 0 e varianza 1), essi coincidono fino ad n = 5 (si veda Kendall-Stuart, p.158). Studi ulteriori su questo fatto potrebbero illuminare di più il concetto di cumulante.

d) Sia X_n una successione di variabili aleatorie tali che

$$E(X_n) = m, \qquad Var(X_n) = \sigma^2, \qquad n = 1, 2, \ldots$$

Se $K_j(X_n)$ converge a 0 quando $n \to \infty$, per ogni j, la successione X_n converge in legge ad una variabile aleatoria normale di media m e varianza σ^2. Sarebbe interessante conoscere altri teoremi di convergenza simili a questo.

e) Sia $p_n(x)$ una successione di polinomi, di grado n = 0, 1, 2, ... nella variabile reale x. Essa è detta di tipo binomiale, se

$$(5) \qquad p_n(x+y) = \sum_{k=0}^{n} \binom{n}{k} p_k(x) p_{n-k}(y)$$

per ogni n (si veda Rota-Kahaner-Odlyzko). Supponiamo poi che per una variabile aleatoria X siano definiti i <u>momenti binomiali</u>

$$b_n(X) = E(p_n(X)) = \int_\Omega p_n(X)\,dP$$

associati alla successione $p_n(x)$. Se X e Y sono indipendenti, allora vale un'identità simile alla (5) con b, X, Y rispettivamente al posto di p, x, y. Sia ora $g_X(t)$ la funzione generatrice esponenziale dei numeri $b_n(X)$,

$$g_X(t) = \Sigma_{n\geq 0} b_n(X) t^n / n!$$

Il problema è di studiare i <u>cumulanti binomiali</u> definiti per analogia dalla identità formale

$c_n(X) = D^n \log g_X(t)|_{t=0}, \qquad n = 1, 2, \ldots$

Se $p_n(x) = x^n$, i cumulanti binomiali si riducono a quelli ordinari; se $p_n(x) = x(x-1)(x-2)\ldots(x-n+1)$, si riducono invece a quelli fattoriali, ben noti in statistica. Un primo passo in questa direzione potrebbe essere quello di studiare i cumulanti relativi ai polinomi fattoriali crescenti

$p_n(x) = x(x+1)(x+2)\ldots(x+n-1)$.

BIBLIOGRAFIA

1. M.CERASOLI-F.EUGENI-M.PROTASI, "Elementi di Matematica Discreta", Zanichelli, Bologna, 1988

2. P.DOUBILET-G.C.ROTA-R.STANLEY, "Finite Operator Calculus", The Idea of Generating Function, pp. 83-134 Academic Press, New York, 1975

3. M.G.KENDALL-A.STUART, "The advanced Theory of Statistics", vol.I, Griffin, London, 1963

4. J.RIORDAN, "An Introduction to Combinatorial Analysis", Wiley, New York, 1958

5. G.C.ROTA-D.KAHANER-A.ODLYZKO, "Finite Operator Calculus", (vedi 2.) pp.7-82

6. T.P.SPEED, Cumulants and partitions lattices, Austral. J. Statistics, 25 (1983), no. 2, 378-388

7.-- II. Generalised k-statistics, J. Austral. Math. Soc. Ser. A 40 (1986), no. 1, 34-53

8.-- III. Multipy-indexed arrays, J. Austral. Math. Soc. Ser. A 40 (1986), no. 2, 161-182

9.-- IV. A.s. convergence of generalised k-statistics, J. Austral. Math. Soc. Ser. A 41 (1986), no. 1, 79-94.