**ESTRATTO DA "INDUZIONI", 32 (2001) pp 141-149**

# Significato geometrico del fattoriale, numeri pseudo-casuali e numero di Nepero

MARIO ABUNDO e LAURA SILVANI

È raro che ci si preoccupi di attribuire un significato geometrico al simbolo $n!$ (fattoriale di $n$). Siamo infatti abituati a considerare questo simbolo nell'ambito del calcolo combinatorio. Secondo la definizione, esso denota il prodotto dei primi $n$ numeri interi. Inoltre un semplice ragionamento mostra che $n!$ è anche il numero delle permutazioni di un insieme costituito da $n$ elementi. Il fattoriale è poi legato ad una delle più importanti funzioni dell'Analisi Matematica, la funzione esponenziale $x \to e^x$, e quindi al numero di Nepero $e$ = 2.7182818284…. ; infatti quest'ultimo può essere definito, non solo come limite della successione crescente $((1 + 1/n)^n)_{n\geq 1}$ , ma anche come somma di una serie, nel modo seguente:

$$e = \sum_{n=0}^{\infty} 1/n! \tag{1}$$

Infine anche il numero delle parti di un insieme di $n$ elementi costituite da $k$ elementi, ossia il cosiddetto coefficiente binomiale $C_{n,k}$ , può essere espresso mediante il fattoriale:

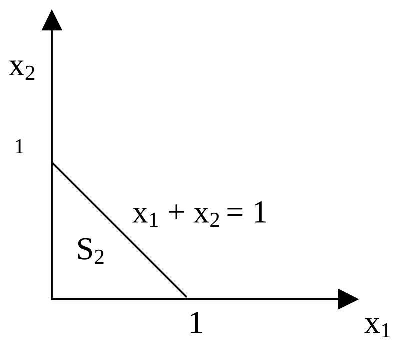
$$C_{n,k} = n! / [k!\,(n-k)!].$$

È dunque naturale che il fattoriale intervenga in ragionamenti di Probabilità discreta. Se poi si considera che al fattoriale è a sua volta legato, come sopra si è visto, il numero di Nepero, non dovrebbe stupire il fatto che anche quest'ultimo numero possa ricevere interessanti rappresentazioni in ambito probabilistico. Proprio di una di queste interpretazioni intendiamo occuparci in questa Nota, oltre che del significato geometrico del fattoriale.

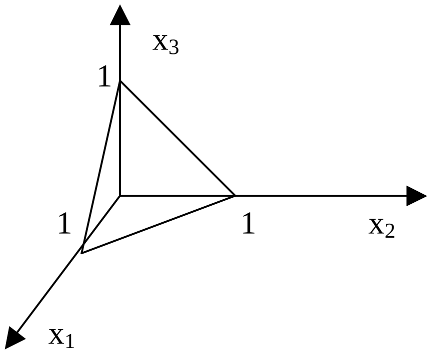
Per cominciare da quest'ultimo, consideriamo in $\mathbf{R}^n$ l'*n–simplesso canonico*, ossia l'insieme $\Sigma_n$ costituito dai punti di $\mathbf{R}^n$ le cui coordinate siano tutte maggiori di zero e abbiano somma minore di 1:

$$\Sigma_n = \{ (x_1,x_2,.....x_n) \in \mathbf{R}^n : x_i > 0,\ i = 1,...n \ \text{ e } \ 0 < x_1 + x_2 + .... + x_n < 1 \} \tag{2}$$

Come è facile vedere, $\Sigma_1$ è l' intervallo (0,1), $\Sigma_2$ non è altro che il triangolo delimitato dagli assi coordinati e dalla retta di equazione $x_1 + x_2 = 1$:



$\Sigma_3$ è la piramide avente base $\Sigma_2$ e delimitata dagli assi coordinati e dal piano di equazione $x_1 + x_2 + x_3 = 1$:




Aumentando la dimensione *n*, l'*n*–simplesso canonico non si può più disegnare. Quello che si può calcolare, per ogni valore di *n*, è invece il suo volume.
Il volume (più precisamente l'area) di $\Sigma_2$ è ½ , quello di $\Sigma_3$ si ottiene dalla formula del volume di una piramide, ovvero $V(\Sigma_3) = (1/2 \cdot 1) / 3 = 1 / ( 2\cdot 3 )$ .
Per calcolare il volume di $\Sigma_n$ in generale, occorre calcolare un integrale multiplo.
Precisamente, il volume di $\Sigma_n$ è dato da :

$$(3) \qquad V(\Sigma_n) = \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \int_0^{1-x_1-x_2} dx_3 \; ....... \int_0^{1-x_1-x_2-.....-x_{n-1}} dx_n$$

Il lettore è invitato a verificare che si ha

$$(4) \qquad V(\Sigma_n) = 1 / n! \quad \text{ovvero} \quad n! = [V(\Sigma_n)]^{-1}$$

confermando i risultati ottenuti per $V(\Sigma_2)$ e $V(\Sigma_3)$ .
Dunque, *n*! è il reciproco del volume dell'*n*-simplesso canonico e questo è il significato geometrico al quale alludevamo. Osserviamo che, al crescere di *n*, $V(\Sigma_n)$ diminuisce e per $n \to \infty$, pur diventando l'insieme $\Sigma_n$ più complicato, il suo volume tende a zero (con velocità più che esponenziale). Ciò è dovuto sostanzialmente al fatto che nel calcolo del volume si moltiplicano numeri minori di 1.
Poiché il volume del cubo *n*-dimensionale

$$(5) \qquad Q_n = \{ (x_1,x_2,.....x_n) \in \mathbf{R}^n : 0 < x_i < 1, \quad i = 1,\dots n \}$$

è ovviamente 1, la probabilità che un punto scelto a caso in $Q_n$ cada in $\Sigma_n$ , è data da

$$(6) \qquad p = P\,(x_1 + x_2 + .... + x_n < 1) = V(\Sigma_n) / V(Q_n) = V(\Sigma_n) = 1 / n!$$

ovvero $n! = p^{-1}$ ; in questo modo abbiamo scritto il fattoriale come il reciproco di una probabilità, e questa è proprio il volume dell'*n*-simplesso canonico.

Ora andiamo a parlare dell'annunciata rappresentazione del numero di Nepero.
Supponiamo di generare dei numeri casuali, uniformemente distribuiti nell'intervallo [0,1]. Precisamente, siano $X_1$, $X_2$, …$X_n$,…. variabili aleatorie indipendenti e uniformemente distribuite in [0,1]. Con ciò intendiamo che tali variabili aleatorie sono stocasticamente indipendenti in blocco, e la densità di probabilità di ognuna delle $X_i$ , i = 1,2,…. , è data da:

(7) $$f(x) = \begin{cases} 1 & \text{se } x \in [0,1] \\ 0 & \text{altrimenti} \end{cases}$$

Poniamo ora:

(8) $$S_n = X_1 + X_2 + \ldots + X_n$$

e cerchiamo il primo intero $N \geq 1$ per cui si abbia $S_N \geq 1$. È chiaro che $N$ risulta una variabile aleatoria (*v.a.*) che può assumere solo valori interi; chiameremo *$N$ il tempo (discreto) di primo passaggio* di $S_n$ oltre la barriera $x = 1$.

| | | | | | |
|---|---|---|---|---|---|
| 0 | $X_1$ | $X_1+X_2$ | $X_1+X_2+X_3$ | 1 | $X_1+X_2+X_3+X_4$ |
| | $\parallel$ | $\parallel$ | $\parallel$ | | $\parallel$ |
| | $S_1$ | $S_2$ | $S_3$ | | $S_4$ |

Possiamo schematizzare la situazione come in figura, immaginando che ad ogni generazione di un numero casuale un punto, posto inizialmente in $x = 0$, si muova verso destra di una quantità pari al numero casuale generato. Nell'esempio risulta $N = 4$, essendo

$S_4 = X_1+X_2+X_3+X_4 > 1$.

Quante generazioni $N$ occorrono in generale, affinché sia $S_N \geq 1$ ? È chiaro che non si può rispondere a questa domanda, essendo $N$ una variabile casuale, non deterministica. Nel caso più fortunato risulta $N = 1$ (cioè $X_1 = 1$ ); in quello più sfortunato $N = +\infty$. Il secondo caso si presenta, ad esempio, se per i numeri casuali successivamente generati si ha $X_i = 1/2^i$, avendosi allora:

$$S_N = \sum_{i=1}^{N} 1/2^i = (1 - (1/2)^{N+1}) / (1 - 1/2) - 1 = 1 - (1/2)^N$$

dunque, risulta $S_N = 1$ solo se $N = \infty$, ( più precisamente $\lim_{N\to\infty} S_N = 1$ ). Naturalmente, una simile situazione si verifica con probabilità zero.
Una quantità che si può calcolare agevolmente, però, è il valore di aspettazione, o valor medio di $N$, che indicheremo con $E(N)$. Quello che risulta, in maniera un po' insospettata, è che

(9) $$E(N) = e$$

ecco quindi spuntare la rappresentazione del numero di Nepero che avevamo annunciato.
Siccome $N$ è una *v.a.* che assume valori interi 1, 2,….., dalla definizione di valore di aspettazione, o valor medio probabilistico, si ha:

(10) $$E(N) = \sum_{n=1}^{\infty} n \cdot P(N = n)$$

Dunque, per calcolare $E(N)$, occorre conoscere $P(N = n)$, ovvero la probabilità che il tempo di primo passaggio attraverso la barriera $x = 1$, sia esattamente uguale ad $n$. Poiché è un po' difficile ottenere direttamente P($N$ = $n$), riordiniamo i termini della somma in (10) in una maniera più opportuna, ottenendo la seguente formula:

$$(11) \qquad \sum_{n=1}^{\infty} n \cdot P(N = n) = 1 + \sum_{n=1}^{\infty} P(N > n)$$

Infatti, il secondo membro di (11) è:

$1 + P(N > 1) + P(N > 2) + P(N > 3) + \ldots =$
$1 + \{ P(N = 2) + P(N = 3) + P(N = 4) + \ldots \} + \{ P(N = 3) + P(N = 4) + P(N = 5) + \ldots \} +$
$\{ P(N = 4) + P(N = 5) + P(N = 6) + \ldots \} + \ldots$

e, utilizzando il fatto che $P(N = 1) + P(N = 2) + P(N = 3) + P(N = 4) + \ldots = 1$, il secondo membro di (11) diviene infine:

$$P(N = 1) + 2\,P(N = 2) + 3\,P(N = 3) + 4\,P(N = 4) + \ldots = \sum_{n=1}^{\infty} n \cdot P(N = n) \;.$$

Osserviamo che la formula (11) vale per una qualsiasi *v.a.* a valori interi positivi.
Dunque, dalla (10) e dalla (11), si ha:

$$(12) \qquad E(N) = 1 + \sum_{n=1}^{\infty} P(N > n)$$

Ora, osserviamo che

$$(13) \qquad P(N > n) = P(S_n < 1) = P(X_1 + X_2 + \ldots X_n < 1)$$

(ovvero, $N > n$ equivale a dire che $S_n$ non supera 1).
Dunque, per la (6), si ha

$$(14) \qquad P(N > n) = 1 / n! \qquad \text{(il volume dell'}n\text{-simplesso)}$$

ed inserendo tale quantità nella (12), otteniamo infine

$$(15) \qquad E(N) = 1 + \sum_{n=1}^{\infty} 1/n! = \sum_{n=0}^{\infty} 1/n! = e$$

che prova la (9).
Una volta trovato $P(N > n)$, è facile ottenere $P(N = n)$. Infatti:

$$(16) \qquad P(N = n) = P(N \le n) - P(N \le n-1) = (1 - P(N > n)) - (1 - P(N > n-1)) =$$
$$= (1 - 1/n!) - (1 - 1/(n-1)!) = 1/(n-1)! - 1/n! = (n-1)/n!$$

È immediato verificare che inserendo la (16) nella (10) si riottiene il valore $e$.
Calcoliamo ora il valor medio di $N^2$, ovvero il cosiddetto momento di ordine 2 della *v.a.* $N$; esso è dato da:

$$(17) \qquad E(N^2) = \sum_{n=1}^{\infty} n^2 \cdot P(N = n)$$

Inserendo in (17) il valore di P ( $N = n$ ) fornito dalla (16), si ha:

$$(18) \qquad E(N^2) = \sum_{n=1}^{\infty} n^2 \cdot (n-1)/n! = \sum_{n=1}^{\infty} n \cdot n\,(n-1)/n\,(n-1)! = \sum_{n=2}^{\infty} n\,(n-1)/(n-1)!$$

ove abbiamo fatto partire l'indice dell'ultima somma da 2, poiché per $n = 1$ si ha contributo nullo. Allora:

$$(19) \qquad E(N^2) = \sum_{n=2}^{\infty} n\,(n-1)/(n-1)! = \sum_{n=2}^{\infty} n\,(n-1)/(n-1)(n-2)! = \sum_{n=2}^{\infty} n\,/(n-2)!$$

$$= \sum_{n=2}^{\infty} [(n-2)/(n-2)! + 2/(n-2)!] = \sum_{n=3}^{\infty} 1/(n-3)! + 2\sum_{n=2}^{\infty} 1/(n-2)!$$

$$= \sum_{i=0}^{\infty} 1/i! + 2\sum_{j=0}^{\infty} 1/j! = e + 2e = 3e$$

Dunque, per la varianza di $N$, definita da

$$(20) \qquad \mathrm{Var}(N) = E(N - E(N))^2 = E(N^2) - (E(N))^2$$

si ottiene

$$(21) \qquad \mathrm{Var}(N) = 3e - e^2 = e\,(3 - e) \cong 0.7658$$

In conclusione, abbiamo ricavato che

(i) $e = E(N)$;

(ii) $e$ è la soluzione maggiore di 1 dell'equazione $x^2 - 3x + \mathrm{Var}(N) = 0$

(l'altra soluzione è minore di 1, e va quindi scartata, essendo per la (12), $E(N) > 1$ ).

La (i) e la (ii) possono essere prese come definizioni alternative del numero di Nepero; esse fanno intervenire $N$ che è una quantità non deterministica.
È istruttivo far realizzare agli studenti (forse è fattibile anche nella scuola secondaria, negli indirizzi sperimentali) un programma di poche righe (in BASIC, PASCAL, FORTRAN, o qualunque altro linguaggio di programmazione) per "stimare" numericamente $\mu = E(N)$, e quindi il numero $e$.
Basta effettuare un numero L abbastanza grande di cicli (lunghi, per esempio 100) di generazione di numeri pseudo-casuali, uniformemente distribuiti in [0,1] (questi hanno una distribuzione che simula quella di una *v.a.* uniformemente distribuita in [0,1]), e per ognuno dei cicli, i =1, 2 ,…. L, stimare $N_i$ (cioé cercando per ogni ciclo il numero $N_i$ di generazioni necessario, affinché $X_1+X_2+\ldots X_{N_i} > 1$ ). Allora, per la legge dei grandi numeri, risulta:

$$(22) \qquad \mu = E(N) \cong (N_1+N_2+\ldots N_L)/L = \hat{\mu}$$

Più precisamente, essendo per ogni realizzazione $E(N_i) = \mu$ (quantità teorica da stimare), per la disuguaglianza di Chebicev, si ha per lo stimatore $\hat{\mu}$ :

(23) $P(|\hat{\mu} - \mu| > \varepsilon) = P(|(N_1+N_2+\dots N_L)/L - E(N)| > \varepsilon) \leq$

$$\leq \mathrm{Var}\,[(N_1+N_2+\dots N_L)/L]/\varepsilon^2 = L \cdot \mathrm{Var}\,(N)/(L^2\varepsilon^2) = \mathrm{Var}\,(N)/(L\,\varepsilon^2)$$

e dunque $\hat{\mu} - \mu$ tende a zero (in probabilità ) per $L \to \infty$, come $1/L$ .
Per esempio, se scegliamo $\varepsilon = 1/100$ e richiediamo una confidenza del 95 %, cioè vogliamo che

$$P(|\hat{\mu} - \mu| \leq 10^{-2}) \geq 0.95 \quad \text{ovvero} \quad P(|\hat{\mu} - \mu| > 10^{-2}) \leq 0.05$$

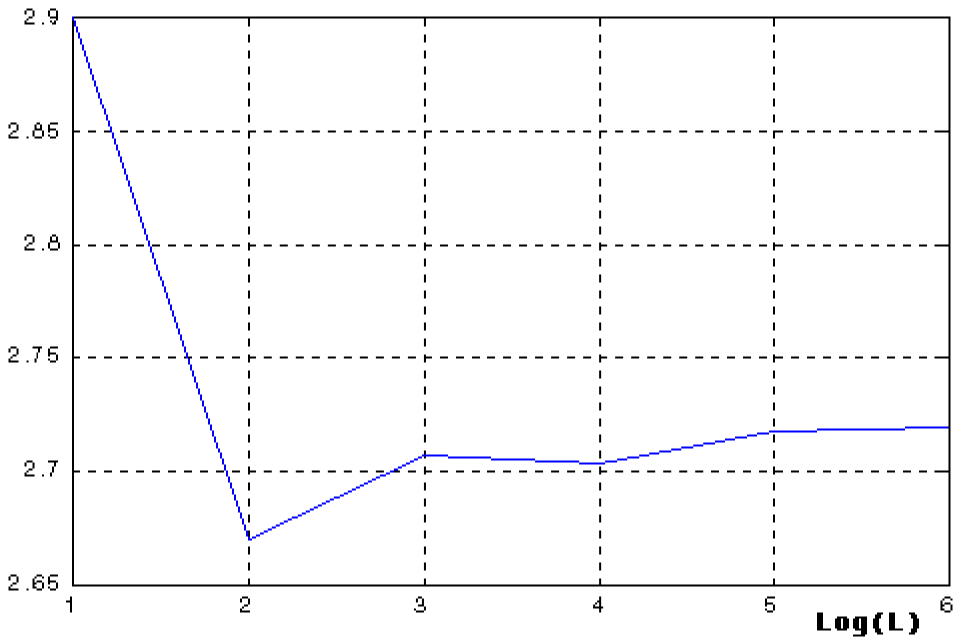
basta che sia, per la (23) e per la (21), $0.7658/(L \cdot 10^{-4}) \leq 0.05$, da cui
$L \geq 0.7658 \cdot 10^6/5 \cong 1.52 \cdot 10^5$
Dunque, simulando 152 000 cicli di numeri pseudo-casuali uniformemente distribuiti

in [0,1], si ottiene, mediante lo stimatore $\hat{\mu}$ dato dalla (22), una valutazione di *e* al 95 % di confidenza, con precisione $10^{-2}$ , nel senso che con probabilità alta (≥ del 95 %) risulta

$$|\hat{\mu} - \mu| \leq 10^{-2}$$

Per avere approssimazioni sensibilmente migliori, occorre naturalmente considerare valori di L molto più grandi.
Abbiamo scritto un programma di poche righe per l'algoritmo non deterministico che calcola approssimativamente il numero di Nepero, e l'abbiamo implementato su un PC, utilizzando una routine per la generazione dei numeri pseudo-casuali uniformemente distribuiti in [0,1].

Nel grafico qui sotto è riportato l'andamento della stima $\hat{\mu} = \hat{\mu}(L)$ di *e* in funzione di L (sull'asse orizzontale è riportato il logaritmo decimale di L, su quello verticale la stima di *e*). Come si vede, per L grande, il valore trovato approssima sensibilmente il valore *vero* di *e*.



**Conclusione**

In questa breve nota, abbiamo mostrato una rappresentazione (non deterministica) del numero di Nepero, utilizzando il "significato geometrico di *n*!" secondo il quale esso è uguale al reciproco del volume dell'*n*–simplesso canonico.

Abbiamo poi esaminato degli esperimenti numerici per *calcolare* il numero *e*, basati sulla generazione di numeri pseudo-casuali.

Tutto ciò ha un certo interesse dal punto di vista didattico, poiché si mostra che una costante universale, quale il numero *e*, può essere scritta come valor medio di una certa variabile aleatoria, ed essa può essere stimata con esperimenti numerici (o di simulazione) mediante un computer. Naturalmente, se si utilizza una delle classiche rappresentazioni deterministiche del numero di Nepero, gli errori che si commettono, sia stimando *e* per troncatura della serie

$\sum_{n=0}^{\infty} 1/n!$ al termine L (cioè calcolando $\sum_{n=0}^{L} 1/n!$ ), sia approssimando *e* con $a_L = (1 + 1/L)^L$, sono molto minori di quello che si ottiene con l'algoritmo stocastico descritto sopra, per lo stesso valore di L.

Mario Abundo
Università " Tor Vergata", Roma

Laura Silvani
IPIA "Europa", Roma