# MatematicaMente

Pubblicazione mensile della sezione veronese della MATHESIS – Società Italiana di Scienze Matematiche e Fisiche – Fondata nel 1895 – Autorizzazione del Tribunale di Verona n. 1360 del 15 – 03 – 1999 – I diritti d'autore sono riservati. Direttore: Luciano Corso - Redazione: Luciano Corso, Elisabetta Capotosto, Arnaldo Vicentini - Via IV Novembre, 11/b – 37126 Verona – tel e fax (045) 8344785 – 338 6416432
e-mail: lcorso@iol.it – Stampa in proprio - Numero 91 – maggio 2005

## Introduzione al pensiero bayesiano

di Mauro Cerasoli [1]

[*Segue dal numero 90*]

**6. Un problema d'induzione scientifica**

Vediamo allora come stanno le cose da un punto di vista più propriamente tecnico.

Una scatola contiene $s$ palline di due tipi soltanto: rosso e blu. Ne vengono estratte $n$ senza rimessa ($0 < n \le s$). Sia $S_n$ la variabile aleatoria "numero di palline rosse nel campione estratto". È ben noto che, se nella scatola ci sono per ipotesi $r$ palline rosse, la probabilità che il campione contenga esattamente $k$ palline rosse (statistica di Maxwell-Boltzmann), vale

$$P\left(S_n = k\right) = \binom{r}{k} \cdot \binom{s-r}{n-k} \Big/ \binom{s}{n} \qquad (6.1)$$

Una formula più nota come *distribuzione ipergeometrica*. Essa è spesso utilizzata per risolvere problemi di giochi o di carte dove il numero di palline rosse è noto. Ad esempio, nel gioco del Lotto, $s$ = 90, $n$ = 5 ed $r$ è il totale di numeri giocati su una ruota.

Nella realtà di tutti i giorni invece, in una popolazione vera, il numero $r$ di palline rosse è quasi sempre sconosciuto: il numero di pazienti con una certa malattia, il numero di elettori favorevoli ad una certa legge, il numero di clienti che comprerebbero un dato prodotto ecc. Quindi nelle applicazioni vere la formula (6.1) è semplicemente inutile essendo sconosciuto $r$. E spesso anche $s$ è ignoto.

Ecco allora il pensiero bayesiano: considera $r$ una *variabile aleatoria* e non una *costante* come per ricavare la formula (6.1). Introduci pertanto una nuova variabile aleatoria ponendo $R_n$ = "numero di palline rosse nella scatola", e scrivi la formula (6.1) nella forma più corretta

$$P\left(S_n = k \middle| R_n = r\right) = \binom{r}{k} \cdot \binom{s-r}{n-k} \Big/ \binom{s}{n} \quad . \qquad (6.2)$$

L'esperienza consiste ora nel sapere che il campione ha $k$ palline rosse e quindi tutto si riduce a invertire il condizionamento nella (6.2) cioè a valutare $P(R_n = r \mid S_n = k)$. Tu non sai quante palline nella scatola sono rosse ma quante sono rosse nel campione! Questo sì, puoi saperlo! Basta uscire dal Dipartimento di Matematica e Statistica e andare a chiedere alla gente per la strada.

Ecco allora l'importanza della formula di Bayes estesa a più alternative $H_r$

$$P\left(H_r \middle| A\right) = P\left(H_r\right) \cdot P\left(A \middle| H_r\right) \Big/ \sum_r P\left(H_r\right) \cdot P\left(A \middle| H_r\right)$$

Pertanto si può affermare che

$$P\left(R_n = r \middle| S_n = k\right) = \frac{P\left(R_n = r\right) \cdot P\left(S_n = k \middle| R_n = r\right)}{\sum_{r=0}^{s} P\left(R_n = r\right) \cdot P\left(S_n = k \middle| R_n = r\right)}$$

Poiché nell'urna ci possono essere 0, 1, 2, …, $s$ palline rosse, sembra ragionevole supporre, per il *principio d'indifferenza di Bernoulli*, che sia uguale a 1 / ($s$ + 1) la probabilità a priori $P(R_n = r)$ di avere $r$ palline rosse nella scatola prima dell'estrazione. Tenendo presente l'identità combinatoria

$$\sum_{r=0}^{s} \binom{r}{k} \cdot \binom{s-r}{n-k} = \binom{s+1}{n+1}$$

di cui una dimostrazione combinatoria è sull'articolo dell'autore riportato nella bibliografia, si ottiene il risultato

$$P\left(R_n = r \middle| S_n = k\right) = \binom{r}{k} \cdot \binom{s-r}{n-k} \Big/ \binom{s+1}{n+1} \qquad (6.3)$$

Si comincia a capire ora cosa vuol dire *bayesiano*: significa ragionare nel modo in cui si è pervenuti alla (6,3) e non alla (6,1). Nella realtà nessuno conosce quante sono le palline rosse (elettori che voteranno SI in un referendum) e quante quelle nere (che votano NO) nella scatola. Ciò che *in realtà* (esperienza) si può conoscere è soltanto il numero $k$ di biglie rosse estratte (di elettori del campione che hanno votato SÍ). Si badi che nelle vere elezioni gli elettori che hanno votato costituiscono ancora un *campione* della popolazione degli elettori, se pure numeroso, in quanto, come spesso sta accadendo, una buona parte di cittadini non va più a votare.

Possiamo risolvere il problema anche nel caso in cui le palline vengono rimesse nell'urna dopo ogni estrazione (*caso binomiale*). Allora

$$P\left(S_n = k \middle| R_n = r\right) = \binom{n}{k} \cdot \left(\frac{r}{s}\right)^k \cdot \left(1 - \frac{r}{s}\right)^{n-k}$$

mentre, sempre nell'ipotesi che $P(R_n = r) = 1 / (s+1)$, si ha

$$P\left(R_n = r \middle| S_n = k\right) = \frac{r^k \cdot (s-r)^{n-k}}{\sum_{0 \le r \le s} r^k \cdot (s-r)^{n-k}}$$

La sommatoria al denominatore è la *funzione beta discreta* indicata spesso con $B(k,n,s)$.

**7. Una famosa formula di Laplace**

Gian Carlo Rota, a proposito della formula di Bayes, ha scritto inoltre: "*In order to see through the mistery one is forced to dig up the original instance of Bayes' law, due to Laplace. Unless one is privy to this example, Bayes'law will remain mysterious*".

Vediamo allora questo problema di Laplace. Sempre nell'ipotesi di avere una scatola con palline rosse e blu, e supponendo che il campione estratto di $n$ palline (senza rimessa) ne contenga $k$ rosse, qual è la probabilità che estraendo un'altra pallina questa sia rossa (evento $A$)?

Per la legge delle alternative, condizionando all'evento ($R_n = r \mid S_n = k$), tale probabilità è

$$P(A) = \sum_{0 \le r \le S} P\left(R_n = r \middle| S_n = k\right) \cdot P\left(A \middle| R_n = r,\, S_n = k\right)$$

$$\sum_{0 \le r \le s} \left[ \binom{r}{k} \cdot \binom{s-r}{n-k} \Big/ \binom{s+1}{n+1} \right] (r-k)/(s-n) =$$

$$\sum_{0 \le r \le s} \left[ (r-k) \cdot \binom{r}{k} \cdot \binom{s-r}{n-k} \right] \Big/ \left[ (s-n) \cdot \binom{s+1}{n+1} \right].$$

Ma vale l'identità combinatoria

$$\binom{n}{k} = \frac{n-k+1}{k} \cdot \binom{n}{k-1}$$

per cui con opportune sostituzioni si ricava

$$P(A)=(k+1)\cdot\frac{\sum_{0\le r\le s}\left[\binom{r}{k+1}\cdot\binom{s-r}{(n+1)-(k+1)}\right]}{\left[(n+2)\cdot\binom{s+1}{n+2}\right]}=$$

$= (k+1)/(n+2)$

Si noti che la probabilità ottenuta non dipende da $s$. La formula, dovuta a Laplace, apparve per la prima volta in Mém. de l'Acad. R. d. Sci., Parigi, 6, 1774, 621. Stranamente essa non viene riproposta nella *Théorie Analytique des Probabilités* del 1812.

Un caso particolare notevole si ha per $k = n$. Allora, se tutte le palline estratte sono rosse, la prossima estratta sarà rossa con probabilità $(n+1) / (n+2)$. Naturalmente, questa tende ad 1 quando $n$ tende all'infinito. Per una errata applicazione di questo risultato (espresso in forma diversa) alla probabilità che domani sorga il sole, si rimanda al trattato del Feller, sempre a pag. 124. [*Segue al numero* 92]

[1] Presidente dell'ADT, e-mail: mceraso@tin.it

## WWW…Mathesis Verona

La sezione Mathesis di Verona è on line!

Grazie all'intercessione del prof. Ruggero Ferro e alla preziosa collaborazione del prof. Giuseppe Scollo, il sito della sezione è attivo presso lo spazio web dell'Università di Verona all'indirizzo http://amarena.sci.univr.it/mathesis .

Le pagine, realizzate da ABS Computers di Verona su mio progetto, riportano il calendario delle attività della sezione per il 2005, danno la possibilità di contattare il direttivo via mail e riportano un elenco di *links* utili.

I soci Mathesis di Verona e gli abbonati alla rivista Matematicamente, dietro registrazione al sistema BSCW che ospita il sito *mathesisvr*, potranno anche accedere ad un'area riservata in cui si potrà scaricare i numeri della rivista in formato PDF e partecipare al forum di discussione. Buona navigazione! ( *di Elisabetta Capotosto*)

# Sul caso: prevedibilità e computabilità

di Luciano Corso

[*Segue dal numero 91*]

Per quanto concerne π, pur essendo considerato statisticamente normale, esso pare che *non sia del tutto imprevedibile* nella sua sequenza di cifre, nel senso che un giocatore che potesse puntare una certa somma sull'uscita sequenziale delle sue cifre avrebbe una probabilità di vincere più alta (> ½) di quella di perdere [B.1]. L'*imprevedibilità* è un'altra caratteristica che contraddistingue la casualità (essa è concettualmente intrinseca all'assioma del disordine e non deve essere confusa con l'imprevedibilità dei sistemi dinamici caotici). Ci si chiede se esiste un numero con una sequenza di cifre assolutamente imprevedibile. Tale numero, se esiste, non dovrebbe essere computabile, per definizione; perché, se lo fosse, si potrebbe sempre trovare una procedura vincente nel prevederlo, con una probabilità maggiore di 0,5. Chiamiamo questi numeri, se esistono, *casuali in senso forte* (*o puri*).

Le considerazione fin qui svolte portano a un concetto forte per definire la casualità di un evento (nella fattispecie numero), legato alle idee di *comprimibilità* e di *computabilità*.

N. Kolmogorov e G. J. Chaitin, in modo autonomo, hanno dato la seguente definizione di casualità: *una successione di cifre che compongono un numero è casuale (e quindi il numero stesso risulta tale) se il più piccolo algoritmo in grado di comunicarla a un calcolatore consta circa dello stesso numero di bit d'informazione della successione stessa*.

Questa definizione vale solo per definire la casualità forte (o pura) di un numero. Per estensione, potremmo affermare che un evento $A$ risulta essere casuale se, tradotto in una parola binaria che lo identifica univocamente, è costituito di una sequenza di cifre binarie in ogni caso maggiore o uguale a quella necessaria a formare un algoritmo in grado di "rappresentarlo" con un computer.

Dal concetto di numero senza schema a quello di evento senza regole il passo cruciale dovrebbe essere fatto attraverso la traduzione di tutto in linguaggio binario e da un confronto tra bit di algoritmi che servono a calcolarlo e bit che lo rappresentano.

***Bibliografia**:* [B.1] A cura di Domenico Costantini, *Caso, Probabilità e statistica*, Quaderni de Le Scienze, n. 98, ottobre 1997; [B.2] Grégoire Nicolis Ilya Prigogine, *La complessità*, Einaudi, Torino, 1991; [B.3] Patrick Billingsley, *Probabilità and Measure*, J. Wiley & Sons, New York, 1995; [B.4] Luciano Corso, *La probabilità nel continuo: un approccio didattico*, pagg. 209 e segg., Atti del congresso nazionale Mathesis di Berga-mo, anno 2002, Bergamo; [B.5] B. B. Mandelbrot, *The fractal geometry of Nature*, W. H. Freeman and company, New York, 1983; [B.6] B. de Finetti, *Filosofia della probabilità*, Il Saggiatore, Milano, 1995; [B.7] Luciano Corso, *Probabilità frequentistica e collettivi di Von Mises*, MatematicaMente n. 22, ottobre 1999, Mathesis VR; [B.8] Peitgen, Jürgens, Saupe, *Chaos and Fractal*, New frontiers of Science, Springer-Verlag, New York, 1992; [B.9] Luciano Corso, *Caos deterministico e frattali stocastici*, Atti Mathesis di Verona 1996, pagg. 341-353.

# Problema su un noto teorema di Chasles

di Nazario Magnarelli

Vogliamo illustrare con un problema il seguente teorema di Chasles: *" Il piano tangente ad una superficie rigata in un punto P contiene la generatrice g passante per il punto stesso. Inoltre, mentre il punto descrive la generatrice, il piano tangente in tal punto varia nel fascio di piani di asse g, corrispondendo proiettivamente a P. Se poi la generatrice è a carattere sviluppabile, il piano tangente rimane fisso al variare del punto".*

Riferiamoci per semplicità ad una quadrica rigata dal punto di vista reale, per esempio al paraboloide a punti iperbolici (o a sella) di equazione

$$x^2-2xy-xz-2y+z=0 \qquad (1)$$

La (1) è un paraboloide perché si ha determinante $A$ > 0, discriminante $A_{44}$ = 0.

Il piano $z$ = 2$y$, tangente nell'origine $O$(0,0,0), interseca il paraboloide secondo una conica degenere le cui rette sono due generatrici di equazioni:

$$g:\begin{cases}x=0\\ z=2y\end{cases}\quad;\qquad h:\begin{cases}x=4y\\ z=2y\end{cases}\quad. \qquad (2)$$

Consideriamo la generatrice $g$. Dando a $y$ i valori $-1, 0, 1, 2, \dots$ otteniamo i punti

$A$(0, -1, -2), $O$(0, 0, 0), $B$(0, 1, 2), $C$(0, 2, 4). (3)

Troviamo i piani tangenti alla quadrica in questi punti. Ricordiamo che il piano tangente in un generico punto $P(x_0, y_0, z_0)$ ha l'equazione:

$$f_x'(P)\cdot(x-x_0)+f_y'(P)\cdot(y-y_0)+f_z'(P)\cdot(z-z_0)=0 \qquad (4)$$

Nel nostro caso si ha:

$$\begin{aligned}&f(x,y,z)=x^2-2xy-xz-2y+z=0,\\ &f_x'=2x-2y-z\,,\ f_y'=-2x-2\,,\ f_z'=-x+1\end{aligned} \qquad (*)$$

Si trova che i piani tangenti nei punti suddetti sono

$$\begin{aligned}&\alpha:\quad 4x-2y+z=0 &&;\quad \omega:\quad -2y+z=0\\ &\beta:\quad -4x-2y+z=0 &&\phantom{;}\quad \gamma:\quad -8x-2y+z=0\end{aligned} \qquad (*)$$

Tutti questi piani passano per i punti $A$, $O$, $B$, $C$ della generatrice $g$, nonostante che questi punti siano distinti fra di loro. [*Segue al n. 93*]