La probabilità di sbagliare tutto

Umberto Cerruti
Università di Torino

Quanto scommettereste?

Sul tavolo davanti a me c'è un mazzo di 50 carte numerate da 1 a 50, accuratamente mescolate, con il numero coperto. Sulla superficie di legno sono disegnati 50 rettangoli numerati anch'essi da 1 a 50. Prendo la prima carta del mazzo e la metto nel rettangolo numero 1, prendo la seconda e la metto nel rettangolo numero 2, e così via fino alla fine del mazzo. Se anche in un sol posto il numero impresso sulla carta coincide con quello del rettangolo sul quale è stata posata, io perdo. Vinco se e solo se nessuna carta si trova al posto giusto, se gli accoppiamenti sono tutti sbagliati.

Se perdo vi do 100 euro, se vinco dovete darmene 300. Ci state? Sareste disposti a puntare di più? Qual è la scommessa equa? E se le carte fossero 20, o 10?

Per rispondere a queste domande ci serve il principio di inclusione-esclusione.

Il principio di inclusione-esclusione

Siamo 20 amici, andiamo a cena in un posto alla moda, dove (a nostra insaputa) servono solo Pizza e Carne. Nel nostro gruppo 16 amano la pizza, 9 amano la carne, e 6 sia l'una che l'altra. Quanti di noi rimarranno delusi?

La domanda equivale a chiedersi quanti tra i 20 non amano né la carne né la pizza. Siamo 20, leviamo i 16 che amano la pizza e i 9 che amano la carne e otteniamo $20 - 16 - 9 = -5$. Qualcosa non va!

In effetti abbiamo sottratto due volte i fortunati che apprezzano entrambi gli alimenti. Dobbiamo riaggiungerli. La risposta esatta è quindi

$$20 - (16 + 9) + 6 = 1$$

Ampliando questo tipo di ragionamento non è difficile convincersi della validità del seguente Teorema

Teorema 1 *Sia assegnato un insieme X con N elementi. Su X sono definite m proprietà P_i. Poniamo*

$N_i =$ *numero degli elementi di X che godono della proprietà P_i.*

$N_{i,j} =$ *numero degli elementi di X che godono delle proprietà P_i e P_j.*

$N_{i,j,k} =$ *numero degli elementi di X che godono delle proprietà P_i, P_j e P_k.*

......

Allora il numero V degli elementi di X che non godono di alcuna delle proprietà definite è

$$V = N - \sum_{1\le i\le m} N_i + \sum_{1\le i<j\le m} N_{i,j} - \sum_{1\le i<j<k\le m} N_{i,j,k} \cdots + (-1)^m N_{1,2,\dots,m} \quad (1)$$

Vediamo una applicazione del teorema ai numeri primi.

Ricordiamo che ogni numero composto ha un divisore primo minore o uguale alla sua radice quadrata. Siano $p_1, p_2, \dots, p_m$ i numeri primi $\le n$. Sia $X = \{1, 2, \dots, n^2\}$. Sia $P_i(x)$ la proprietà: p_i divide x. Denotiamo con $\pi(x)$ in numero dei primi $\le x$. Se V è il numero degli interi in X che non soddisfano ad alcuna P_i, allora si ha

$$\pi(n^2) = V + m - 1 \quad (2)$$

Infatti, se un $1 \le x \le n^2$ non è divisibile per alcuno dei p_i, è certamente primo, con l'eccezione di 1, che non viene considerato primo. A questi $V - 1$ primi bisogna poi aggiungere gli m primi $p_1, p_2, \dots, p_m$, i quali, pur essendo primi, soddisfano rispettivamente a $P_1, P_2, \dots, P_m$ e pertanto non sono contati in V.

Il numero degli interi $\le z$ e divisibili per y è ovviamente $\lfloor \frac{z}{y} \rfloor$ (dove $\lfloor w \rfloor$ è la parte intera di w). Quindi nel nostro caso $N_{i,j,\dots,h} = \lfloor \frac{n^2}{p_i p_j \cdots p_h} \rfloor$. Utilizzando la (1)e la (2) otteniamo che $\pi(n^2) =$

$$n^2 - \sum_{1\le i\le m} \lfloor \frac{n^2}{p_i} \rfloor + \sum_{1\le i<j\le m} \lfloor \frac{n^2}{p_i p_j} \rfloor + \cdots + (-1)^m \lfloor \frac{n^2}{p_1 p_2 \dots p_m} \rfloor + m - 1 \quad (3)$$

Esempio 2 *Prendiamo $n = 10$. Allora $m = 4$ e $p_1 = 2, p_2 = 3, p_3 = 5, p_4 = 7$.*

Pertanto

$\pi(100) = 100 - \left(\lfloor \frac{100}{2} \rfloor + \lfloor \frac{100}{3} \rfloor + \lfloor \frac{100}{5} \rfloor + \lfloor \frac{100}{7} \rfloor \right) + \left(\lfloor \frac{100}{2\times 3} \rfloor + \lfloor \frac{100}{2\times 5} \rfloor + \lfloor \frac{100}{2\times 7} \rfloor + \lfloor \frac{100}{3\times 5} \rfloor + \lfloor \frac{100}{3\times 7} \rfloor + \lfloor \frac{100}{5\times 7} \rfloor \right) - \left(\lfloor \frac{100}{2\times 3\times 5} \rfloor + \lfloor \frac{100}{2\times 3\times 7} \rfloor + \lfloor \frac{100}{2\times 5\times 7} \rfloor + \lfloor \frac{100}{3\times 5\times 7} \rfloor \right) + \lfloor \frac{100}{2\times 3\times 5\times 7} \rfloor +$

$4-1 = 100-(50+33+20+14)+(16+10+7+6+4+2)-(3+2+1+0)+0+4-1 = 25$

Ci sono dunque 25 primi ≤ 100. Si noti che abbiamo ottenuto questo risultato da una formula (un algoritmo effettivo) che riceve in input (come dati iniziali) i 4 primi minori di 10.

Questo fatto non è per nulla banale e dovrebbe stupire. Più in generale l'espressione (3) ci permette di calcolare *quanti* primi esistono nell'intervallo $[n+1, n^2]$ se noi sappiamo *quali* sono i primi nell' intervallo $[2, n]$. Per esempio se io ho un elenco dei 168 primi minori di 1000 mediante la (3) posso venire a sapere che tra 1001 e 1000000 ci sono esattamente 78330 primi!

La soluzione del problema

Ritorniamo alle nostre carte numerate e ai rettangoli sul tavolo. Supponiamo di avere n carte. Esse possono essere disposte negli n rettangoli in $n!$ modi diversi. (Ricordiamo che il fattoriale $n! = n(n-1)(n-2)\dots 1$ conta il numero delle permutazioni di n oggetti). Sia X l'insieme delle $n!$ disposizioni delle n carte negli n rettangoli. Sia $P_i(x)$ la proprietà: nella disposizione x la carta i va nel rettangolo i. Vogliamo calcolare il numero V delle disposizioni che non soddisfano ad alcuna di queste proprietà. Queste disposizioni (che sono i casi favorevoli a me, quelle che mi fanno vincere) sono dette spiazzamenti, o dismutazioni o dérangements. Calcoliamo N_i, il numero delle disposizioni per le quali P_i è vera. Se per esempio $i = 3$ la carta 3 va nel terzo rettangolo. Le altre $n-1$ carte possono andare dove vogliono, sono possibili tutte le permutazioni. Pertanto $N_i = (n-1)!$.

Calcoliamo $N_{i,j}$. Per esempio siano $i = 4$ e $j = 10$. Se $P_{4,10}$ è vera, la carta 4 va nel quarto rettangolo e la 10 va nel decimo. Come prima, le rimanenti $n-2$ sono libere, ci possono essere tutte le permutazioni. Dunque $N_{i,j} = (n-2)!$.

E' chiaro a questo punto che il numero $N_{i,j\dots h}$ non dipende dai valori assunti dagli indici, ma soltanto dal numero degli indici. Se ci sono s indici, il numero vale $(n-s)!$. Quando calcoliamo

$$\sum_{1\leq i_1<i_2<\cdots<i_s\leq n} N_{i_1,i_2,\dots,i_s}$$

sommiamo tante volte uno stesso numero: $N_{i_1,i_2,\dots,i_s} = (n-s)!$. Quante volte lo sommiamo? Se $s = 1$ ci sono n N_i. Se $s = 2$ dobbiamo contare quante sono le coppie di posti. Abbiamo n carte, ci sono dunque $\binom{n}{2}$ coppie. In generale, per un qualsiasi s, ci sono $\binom{n}{s}$ sottoinsiemi di s carte. Ricordiamo che i numeri

$$\binom{n}{s} = \frac{n!}{(n-s)!s!} \tag{4}$$

si dicono binomiali, e sono i coefficienti che appaiono nella famosa e utilissima espressione (dovuta a Newton)

$$(a+b)^n = \sum_{s=0}^{s=n} \binom{n}{s} a^s b^{n-s} \tag{5}$$

Pertanto

$$\sum_{1 \leq i_1 < i_2 < \cdots < i_s \leq n} N_{i_1, i_2, \ldots, i_s} = \binom{n}{s}(n-s)! = \frac{n!}{s!} \tag{6}$$

Utilizzando la (6), la (1), e ricordando che nel nostro caso $N = n!$, otteniamo che il numero D_n degli spiazzamenti è

$$D_n = n! - \frac{n!}{1!} + \frac{n!}{2!} - \frac{n!}{3!} + \cdots + (-1)^n \frac{n!}{n!} \tag{7}$$

e infine

$$D_n = n! \sum_{s=0}^{s=n} \frac{(-1)^s}{s!} \tag{8}$$

Quindi, poiché D_n è per me il numero di casi favorevoli, e il numero complessivo dei casi (delle permutazioni) è $n!$, la mia probabilità $P(n)$ di vincere con n carte è esattamente $\frac{D_n}{n!}$, ovvero

$$P(n) = \sum_{s=0}^{s=n} \frac{(-1)^s}{s!} \tag{9}$$

Ecco l'elenco delle probabilità esatte per n che varia da 1 a 10

$$\left\{0, \frac{1}{2}, \frac{1}{3}, \frac{3}{8}, \frac{11}{30}, \frac{53}{144}, \frac{103}{280}, \frac{2119}{5760}, \frac{16687}{45360}, \frac{16481}{44800}\right\}$$

Lo 0 iniziale dipende dal fatto che con 1 carta sola non si può sbagliare!

Vediamo ora i valori approssimati a 6 cifre di queste probabilità

$$0.0, 0.5, 0.333333, 0.375, 0.366667, 0.368056, 0.367857, 0.367882, 0.367879, 0.367879$$

Sorpresa! Questi valori sono equivalenti (nella pratica) addirittura da 5 carte in su: la probabilità è circa 0.36, qualcosa più di un terzo. Da 9 carte in poi le probabilità coincidono fino alla sesta cifra decimale! Come si spiega?

Un risultato apparentemente paradossale

Come è chiaro dalla sua espressione (9) la mia probabilità di vincere, $P(n)$, aumenta (rispetto a $P(n-1)$) di $\frac{1}{n!}$ se n pari, e, viceversa, diminuisce della stessa quantità se n è dispari. Però il fattoriale cresce talmente in fretta che le differenze si avvicinano rapidamente a 0. Si intuisce subito che la sequenza dei $P(n)$ converge a qualcosa, e anche molto velocemente!

Con un minimo di analisi si vede subito che

$$\lim_{n\to\infty} P(n) = \lim_{n\to\infty} \left(\sum_{s=0}^{s=n} \frac{(-1)^s}{s!} \right) = \frac{1}{e}$$

dove $e = 2.71828182845905\ldots$ è la ben nota costante di Napier, la base del logaritmo naturale.

Pertanto la probabilità che cerchiamo è $\frac{1}{e} = 0.3678794411714442\ldots$.

Possiamo ora rispondere alle domande poste all'inizio. La probabilità che voi vinciate è $Q(n) = 1 - P(n)$. Nel caso di 50 carte $P(n)$ si può assumere uguale a $\frac{1}{e}$. Dunque, se siete stati al patto, la mia speranza di guadagno è $P(50) \times 300 = \frac{1}{e} \times 300$ euro, pari a circa 110.364 euro. La vostra è $Q(50) \times 100 = \frac{e-1}{e} \times 100$ euro, pari a circa 63.2121 euro.

Come si è visto, giocare con 10, 20, 50 o anche 100 carte non cambia sostanzialmente nulla.

Se i 60000 posti dello stadio olimpico sono stati tutti prenotati, e tutte le 60000 persone vengono, e poi si siedono a caso, la probabilità che nessuno si sieda al posto che gli è stato assegnato è sempre all'incirca = 0.36.

La scommessa equa si ottiene eguagliando la mia speranza di guadagno con la vosta, cioè rispettando la uguaglianza

$$\frac{1}{e} \times X = \frac{e-1}{e} \times Y$$

dove X e Y sono rispettivamente il mio e il vostro guadagno in caso di vittoria. Segue che

$$\frac{X}{Y} = e - 1 = 1.71828\ldots$$

Pertanto una scommessa (quasi :) equa sarebbe stata; se perdo vi do 100 euro, se vinco me ne date 172.

Spiazzamenti e probabilità

La sequenza degli interi D_n soddisfa a molte interessanti identità. Tra quelle più importanti ricordiamo

$$D_n = nD_{n-1} + (-1)^n \tag{10}$$

$$D_n = (n-1)(D_{n-1} + D_{n-2}) \tag{11}$$

$$n! = \sum_{k=0}^{k=n} \binom{n}{k} D_{n-k} \tag{12}$$

Soffermiamoci in particolare sul significato della (12).

Supponiamo che una permutazione fissi due elementi i e j. Allora i rimanti $n-2$ devono essere spiazzati. Poiché il numero delle coppie è $\binom{n}{2}$ ci sono esattamente $\binom{n}{2}D_{n-2}$ permutazioni che fissano due elementi. Più in generale ci sono

$$\binom{n}{k} D_{n-k} \tag{13}$$

permutazioni che fissano k elementi su n. La (12) esprime il fatto che che se sommiamo il numero delle permutazioni che fissano $0, 1, 2, \ldots, n$ elementi, troviamo tutte le $n!$ permutazioni.

Esempio 3 *Teniamo bene a mente che quando diciamo che una permutazione σ fissa k elementi intendiamo dire che σ fissa* esattamente *k elementi, né più, né meno. Per esempio, il numero di permutazioni che fissano $n-1$ elementi è 0: se σ fissa $n-1$ elementi deve fissare anche il rimanente.*

Se $n = 6$ la distribuzione è la seguente

$$\{265, 264, 135, 40, 15, 0, 1\}$$

Ci sono 265 permutazioni che non fissano nulla (gli spiazzamenti), 264 che fissano 1 elemento, 135 che fissano due elementi, $\ldots$, 0 fissano 5 elementi e 1 (quella identica) li fissa tutti.

La probabilità che qualcosa non cambi di posto

Da quanto detto segue che la probabilità che una permutazione, presa a caso, fissi k elementi su n è

$$\frac{\binom{n}{k} D_{n-k}}{n!}$$

Ricordando le definizioni di D_n (8) e di binomiale (4) otteniamo

$$\frac{\binom{n}{k} D_{n-k}}{n!} = \frac{1}{k!} \sum_{s=0}^{s=n-k} \frac{(-1)^s}{s!}$$

Poiché k è costante, il limite per $n \to \infty$ di questa epressione è

$$\frac{1}{ek!} \tag{14}$$

La convergenza è velocissima e, ai fini pratici, la (14) rappresenta la probabilità che una permutazione fissi k elementi, indipendentemente da n (ovviamente si deve avere $n > k$). Se $k = 0$ ($0! = 1$)ritroviamo la probabilità $\frac{1}{e}$ di uno spiazzamento. Ponendo $k = 1$ scopriamo che la probabilità che una permutazione fissi 1 elemento è ancora $\frac{1}{e}$.

Aumentando il numero di punti fissi la probabilità tende rapidamente a 0. E' circa 0.18394 per $k = 2$, 0.061313 per $k = 3$ ed è vicina a 10^{-7} per $k = 10$. Se $k = 20$, è 1.51×10^{-19}.

Concludiamo con due esempi. I particolari e le dimostrazioni si trovano nella letteratura citata.

I vicini di tavola

In occasione del pranzo e della cena sociale di un convegno, i partecipanti vengono fatti sedere intorno ad un grande tavolo rotondo. Qual è la probabilità che a sera almeno uno si ritrovi accanto alla stessa persona?

Si prova che questa probabilità tende a $\frac{1}{e^2} = 0.135335283236612\ldots$.

Più in generale, la probabilità che, nelle stesse ipotesi, esattamente k coppie si ritrovino vicine è

$$\frac{2^k}{e^2 k!}$$

Per esempio, per $k = 2, 3, 4$ otteniamo, rispettivamente, le probabilità $0.270671, 0.180447, 0.090223$.

I regali di Natale

Nella classe di Silvia (quarta elementare), è tradizione, per Natale, che i bambini si facciano un piccolo dono reciproco. I nomi degli alunni vengono scritti su striscioline di carta, che sono poste in una grande scatola. La scatola viene scossa ben bene, e poi ogni bimbo prende uno dei biglietti, che contiene (questa è la regola) il nome del compagno al quale darà il regalino.

Se si estrae il proprio nome, si rimette il foglio nella scatola. Viene quindi prodotta una permutazione che è uno *spiazzamento.*

Silvia estrae il nome della sua amica Beatrice. E accade che Beatrice estrae il nome di Silvia! Silvia si stupisce e pensa che si tratti di un evento eccezionale. Si dimostra però che in questo caso la probabilità tende (sempre molto velocemente) a

$$\frac{1}{e^{\sqrt{2}}} = 0.606530659712633\ldots$$

La probabilità che si verifichi questo evento è quindi più vicina a $\frac{2}{3}$ che a $\frac{1}{2}$.

Quanto abbiamo visto ci dà almeno un insegnamento.

E' estremamente improbabile che le cose vadano a posto per caso!

Se i libri della nostra biblioteca vengono, per uno scherzo di pessimo gusto, rimescolati in modo totalmente casuale, la probabilità che almeno due di essi si ritrovino al loro posto è appena più grande di $\frac{1}{4}$. Infatti le probabilità di uno spiazzamento o di un singolo successo valgono entrambe $\frac{1}{e}$. Perciò almeno due libri si ritroveranno al loro posto con probabilità $1 - \frac{2}{e} = 0.26424\ldots$.

Un calcolo analogo ci dice che in meno di 6 casi su 10000 almeno 5 libri saranno a posto. E ritroveremo almeno 10 libri dove erano prima soltanto con probabilità 10^{-8} (cioè praticamente MAI).

Bibliografia

[1] BENGT ASPVALL, *The dinner table problem*, Stanford Department of Computer Science, Report No. STAN-(X-80-829), 1980.

[2] UMBERTO CERRUTI, *Il Paradosso del Compleanno*, Blog [**58**], 14 Gennaio 2008

[3] BRIAN CONREY AND TOM DAVIS, *Derangements*, http://www.geometer.org/mathcircles

[4] KELLY M. MCGUIRE, GEORGE MACKIW, CHRISTOPHER H. MORRELL, The Secret Santa Problem, The Mathematical Gazette, Vol. 83, No. 498 (Nov., 1999), pp. 467-472.

[5] DANIELA ROMAGNOLI, *Problemi di combinatorica*, Quaderno 43 - Giugno 2008