le Rubriche
del gens - n. 78

Compleanni e non-compleanni

A cura di Silvano

> Nell'ufficio matematico di una compagnia di assicurazioni, un sondaggio ha casualmente rilevato che per tre dei 60 dipendenti il compleanno cadeva nello stesso giorno. Questo è sembrato essere un caso molto insolito e raro e sono stati fatti vari tentativi per quantificare la bassa probabilità di un tale evento. Mi sono stati presentati dei calcoli che hanno mostrato una probabilità di pochi millesimi! L'ipotesi di base era, naturalmente, che ciascuno dei k = 60 individui avesse la stessa probabilità n^{-1} di nascere in uno degli n = 365 giorni dell'anno.
> La teoria razionale della probabilità, che definisce una probabilità soltanto all'interno di un collettivo ben definito, porta a un risultato completamente diverso.
>
> [...]
>
> Il calcolo fornisce per $E(x_3)$ il valore di 0,22 con n = 365 e k = 60; cioè ogni circa 4 o 5 raggruppamenti di 60 persone ciascuno, in media 1 insieme (di persone dello stesso gruppo) dovrebbe avere tre giorni ripetuti.

Così scriveva Richard von Mises, lo studioso al quale è comunemente attribuita la presentazione del *paradosso dei compleanni*, in un articolo pubblicato nel 1939 a Istambul (dove si era trasferito dalla Germania dopo l'avvento di Hitler).[1]

Analogamente ai dipendenti di quella società di assicurazioni, anche a Delia è apparso molto strano che il 23 luglio ricorressero contemporaneamente **tre compleanni**, il suo, quello di Betty T. e quello di Marco (mio figlio che, pur non appartenendo al GENS, casualmente aveva partecipato due giorni prima all'escursione settimanale del gruppo). È proprio questo il cosiddetto "paradosso": il fatto che la probabilità di un simile evento non sia irrisoria appare del tutto controintuitivo!

Due compleanni

Per il vero, il paradosso dei compleanni è solitamente discusso per il caso di (almeno) **due compleanni coincidenti** in un gruppo di k persone, perché i calcoli necessari per tre, quattro o più coincidenze sono decisamente più complicati. Farò anch'io così, ma poi riporterò anche il risultato che interessa il GENS (con tre compleanni), senza mostrare i calcoli. Come indicato da von Mises, assumerò che vi sia la stessa probabilità di nascere in ciascuno dei 365 giorni di un anno (tralasciando il caso dei bisestili).[2]

---

[1] "Über Aufteilungs und Besetzungswahrscheinlichkeiten", *Revue de la Faculté des Sciences de l'Université d'Istanbul*, **4** (1939), 145-163. (http://alexander.shen.free.fr/vonMises_64_SelectedPapersVol2OCR.pdf#page=322)
Per una biografia di von Mises si può vedere la voce di Wikipedia (in inglese): https://it.wikipedia.org/wiki/Richard_von_Mises

[2] Se la probabilità di nascita non è la stessa in tutti i giorni dell'anno (come ci dicono le statistiche demografiche), la conseguenza è che la probabilità di coincidenze **cresce** leggermente.

Anziché cercare di determinare direttamente la probabilità $p(k)$ che vi siano due (o più) compleanni coincidenti in un gruppo di k persone, conviene calcolare la probabilità $p'(k)$ che ***non*** vi siano coincidenze. Naturalmente potremo poi ricavare $p(k) = 1 - p'(k)$, perché le due situazioni sono complementari.

Procediamo per gradi. Se vi fossero soltanto due persone, la probabilità di *non* avere due compleanni coincidenti sarebbe $p'(2) = 364/365$, perché la seconda persona avrebbe 364 giorni alternativi per il suo compleanno rispetto al giorno del compleanno dell'altra persona. Consideriamo ora tre persone: ammesso che le prime due abbiano il compleanno in due giorni diversi, la terza persona avrà 363 possibilità (su 365) per avere un compleanno differente. E, poiché la probabilità che le prime due non abbiano il compleanno in comune è $p'(2)$, sarà: $p'(3) = p'(2)\cdot 363/365 = (364/365)\cdot(363/365)$.

Possiamo continuare nello stesso modo. La probabilità che quattro persone abbiano compleanni diversi sarà: $p'(4) = (364/365)\cdot(363/365)\cdot(362/365)$, e così via.
Per comodità, possiamo scrivere anche: $p'(4) = (365/365)\cdot(364/365)\cdot(363/365)\cdot(362/365)$ e, quindi:

$$p'(4) = 365\cdot 364\cdot 363\cdot 362/365^4.$$

A questo punto è facile generalizzare, scrivendo:

$$p'(k) = 365\cdot 364\cdot 363\cdot 362\cdot \cdots \cdot(365 - k + 1)/365^k.$$

Ne viene allora, che:

$$p(k) = 1 - 365\cdot 364\cdot 363\cdot 362\cdot \cdots \cdot(365 - k + 1)/365^k.$$

Già, ma… quanto vale k per il caso del GENS? Consultando l'elenco dei nomi e includendo anche coloro che, come Delia, partecipano di rado alle escursioni, troviamo $k = 71$. Per trovare la probabilità di avere due persone del gruppo con lo stesso compleanno dobbiamo quindi calcolare $p(71)$: [3]

$$p(71) = 1 - \frac{365\cdot 364\cdot 363\cdot \cdots \cdot 296\cdot 295}{365^{71}} =$$

$$= 1 - \frac{365!}{294!\cdot 365^{71}} \cong 0{,}999.$$

In altre parole, che in un gruppo di 71 persone ve ne siano due che compiono gli anni lo stesso giorno è un evento praticamente certo!
E, in effetti, il calendario GENS riporta più coincidenze: il 10 gennaio (Enzo e Lisa), il 29 maggio (Fosca e Mario), il 23 luglio (Betty T. e Delia) e l'8 settembre (Arturo e Piera Z.).[4]

In un altro modo

C'è un altro modo per arrivare al medesimo risultato. Bisogna considerare in quante maniere differenti le 365 date possono essere assegnate alle $k = 71$ persone, in un caso escludendo che ci possano essere ripetizioni (coincidenze di compleanni), nell'altro ammettendolo. Qui ci fornisce le risposte il calcolo combinatorio[5]: il primo caso è quello delle *disposizioni semplici*:

$$D_{365,71} = \frac{365!}{(365\text{-}71)!}$$

mentre il secondo caso è quello delle *disposizioni con ripetizione*:

$$D'_{365,71} = 365^{71}.$$

Il rapporto tra D (casi senza coincidenze) e D' (casi con coincidenze) ci dà la probabilità $p'(71)$ di *non* avere compleanni coincidenti. Si vede subito che questo risultato coincide con quello che avevamo ricavato sopra.

---

[3] Nel calcolo che segue, il simbolo (!) indica il fattoriale, ovvero: $n! = n\cdot(n\text{-}1)\cdot(n\text{-}2)\cdot \cdots \cdot 2\cdot 1$.

[4] Anzi, potrebbero essere di più perché il valore atteso per il numero di giorni con coincidenze è 6 (https://math.stackexchange.com/q/1366028).

[5] Gli argomenti di calcolo combinatorio fanno parte del programma di matematica dei licei. Si possono, quindi, trovare in uno qualsiasi dei manuali scolastici. In alternativa, c'è sempre Wikipedia: https://it.wikipedia.org/wiki/Calcolo_combinatorio

Tre compleanni

E per tre persone? Come accennato sopra, il calcolo è assai più complicato[6]. Tuttavia il programma di statistica R fornisce una funzione apposita, `pbirthday()`, che calcola i valori richiesti e ci permette di ottenere la probabilità di tre coincidenze in un gruppo di 72 persone[7]. Il risultato è:

$$P(72) \cong 0{,}33$$

o, in altre parole, ci si può attendere che in media, ogni tre gruppi di 72 persone, uno abbia tre coincidenze nei compleanni. Decisamente una probabilità non trascurabile.[8]

Il numero minimo

Si può fare anche un calcolo “inverso” e, cioè, ricavare il numero minimo di persone per il quale la probabilità di avere *k* coincidenze è maggiore del 50%. Questi i risultati, sempre ottenuti con R:

k	*n*
2	23
3	88
4	187

I non-compleanni

Nel film *Alice nel paese delle meraviglie* (*Alice in Wonderland*, 1951) il Cappellaio Matto, la Lepre Marzolina e il Ghiro festeggiano il loro non-compleanno. Anche Alice si rende conto che quel giorno è pure per lei un non-compleanno e, quindi, riceve una bella fetta di “torta di non-compleanno” dal Cappellaio Matto.

È un caso raro che quattro personaggi non compiano gli anni in un giorno fissato? No di certo.
La probabilità che un certo giorno sia un non-compleanno è pari a 364/365 per ciascuno dei personaggi; dunque la probabilità che nessuno dei quattro abbia il compleanno in quel giorno è $(364/365)^4 \cong 0{,}989$.
Insomma, ci si può proprio scommettere!

---

[6] E. H. Mckinney. “Generalized Birthday Problem”, *The American Mathematical Monthly*, **73** (1966), n. 4, 385-387 (https://www.jstor.org/stable/2315408)

[7] Ai 71 gensini dobbiamo aggiungere Marco. La funzione di R usa un metodo di calcolo approssimato, ma che è del tutto sufficiente per i nostri scopi.

[8] Per chi volesse cimentarsi, il problema è risolto anche dal programma computazionale WolframAlpha: https://www.wolframalpha.com/input/?i=birthday+paradox+72+people