

Come scoprire una frode?

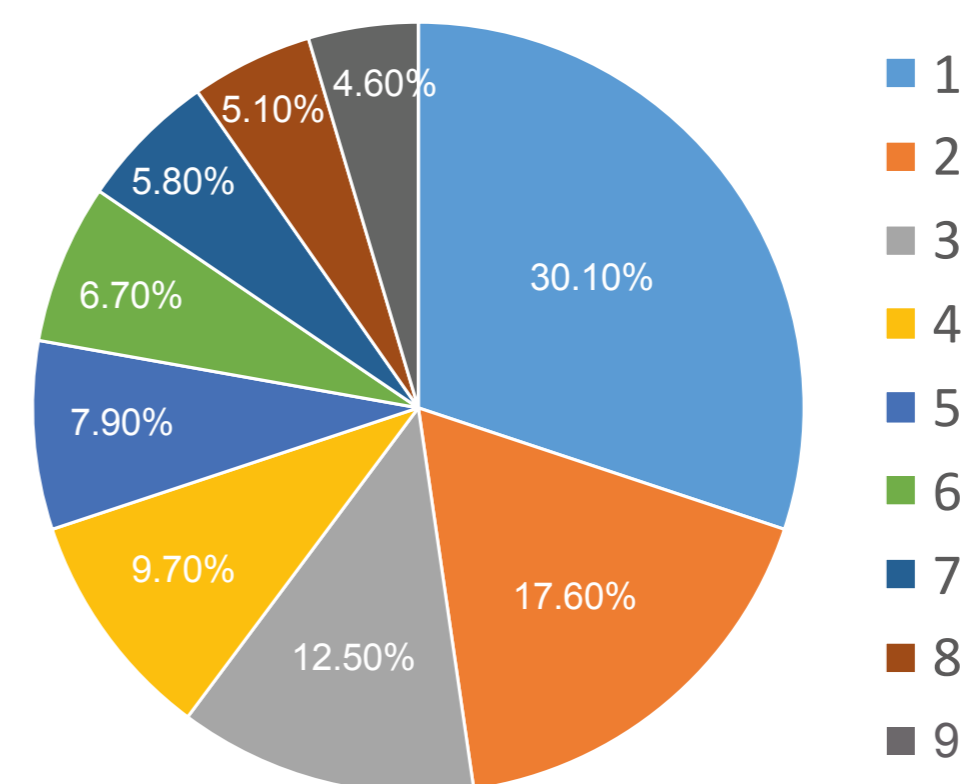
La matematica può essere d'aiuto per distinguere dati reali da dati inventati.

Tantissimi numeri casuali

Immaginate un elenco di numeri ottenuti da tanti dati reali, come il numero degli abitanti di ogni comune della Svizzera. Per elenchi di numeri di questo tipo, ingenuamente si potrebbe pensare a un comportamento casuale, per cui i numeri che iniziano con la cifra 1 siano circa tanti quanti quelli che iniziano con una qualsiasi altra cifra. E invece non è così!

Una regola empirica detta *Legge di Benford* prevede che siano di più i numeri che iniziano con l'1, rispetto a quelli che iniziano con il 2, con il 3 ecc.

Questa legge può essere applicata in qualsiasi contesto in cui ci siano elenchi di numeri tratti da dati reali, sempre che tali dati non siano influenzati da altri parametri. Ad esempio, nel caso delle altezze delle persone non è applicabile, dato che la prima cifra è molto raramente diversa da 1 e 2. Nei casi in cui è applicabile, la legge di Benford può essere estremamente utile: il matematico statunitense Mark Nigrini la utilizza da diversi anni, insieme ad altri strumenti statistici, per analizzare i bilanci delle aziende e individuare anomalie e possibili casi di frodi fiscali!



Le frequenze previste dalla Legge di Benford per numeri che iniziano con una determinata cifra.



La legge di Benford nelle ricerche online

Scegliete quattro cifre a caso e provate a fare una ricerca su internet scrivendo le quattro cifre precedute dall'1. Quanti risultati appaiono?

Provate ora a far precedere le quattro cifre dal 2, quanti sono i risultati?

Potete provare antepoendo alle quattro cifre casuali ogni volta una cifra diversa e fare così ulteriori verifiche empiriche della legge di Benford.



Testa o croce?

La matematica è ricca di risultati contrari alla nostra intuizione, come la legge di Benford. Soprattutto nell'ambito della probabilità è molto difficile avere una percezione corretta degli eventi casuali e non venire tratti in inganno.

Ad esempio, provate a realizzare il seguente esperimento. Inventate una sequenza di Testa (T) e Croce (C) in modo che secondo voi possa rappresentare gli esiti dei lanci di una moneta; poi lanciate davvero una moneta per ottenere un'altra sequenza.

Ci sono differenze tra le due sequenze? Quali?



Ecco a titolo d'esempio le due sequenze da noi generate:

TCTTCCTCCTTCTCCTTCCCTCTTCT

TTCCCCCTTCTCCTCCTCTTTTCC

Sapreste stabilire quale è delle due quella inventata rispetto a quella ottenuta registrando i 26 esiti del lancio di una moneta?

Nel nostro esempio, la prima sequenza è quella inventata, mentre la seconda sequenza è stata ottenuta lanciando realmente una moneta.

Alla maggior parte delle persone sembra invece il contrario: la seconda sequenza appare come quella inventata (possibile che siano davvero uscite tutte quelle C di fila dopo il terzo lancio?), mentre la prima appare come casuale (perché c'è molta alternanza tra T e C).

A livello intuitivo siamo portati a credere che una sequenza casuale debba contenere tante variazioni, ma non è necessariamente così: lanciando una moneta, ogni esito è indipendente dal precedente e non è così improbabile ottenere sequenze, che si scontrano con la nostra idea di "caso", in cui un determinato esito è ripetuto più volte.

Fonti

Codogno, M. (2011). La legge di Benford. *Il Post*. Disponibile in: <https://www.ilpost.it/mauriziocodogno/2011/04/11/la-legge-di-benford/> (consultato il 21.11.2020).

Icona *coin flipping* realizzata da Oleksandr Panasovskyi, scaricata da <https://thenounproject.com/>.

Icona *lente d'ingrandimento* e *punto di domanda* realizzate da Freepik e Vectors Market, scaricate da www.flaticon.com