



🏠 Correlazioni statistiche e causalità ← Biometria

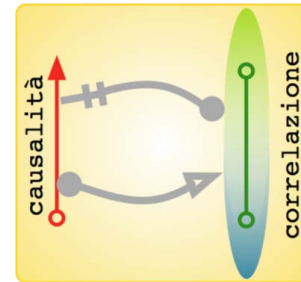
HOME ♦ .php ♦ .html ♦ .pdf

Biometria ♦ Epidemiologia ♦ Statistica

🏠 Correlazioni statistiche e causalità indice (sopprimi)

1. Correlazione verso causalità
 - 1.1 Correlazione ≠ causalità
 - 1.2 Dalla correlazione alle ipotetiche causalità
 - 1.3 Implicazioni della confusione
2. Analisi di regressione
3. Statistiche di paragone
4. Allegati
 - 4.1 Sitografia
 - 4.2 Commenti

Autore: Peter Forster

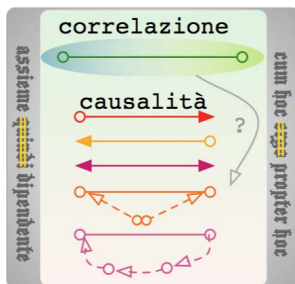


clic per accedere

MedPop Medicina popolare
a cura di D. Rüegg

Dei dati statistici possono per definizione illuminare **solo delle correlazioni** e mai delle dipendenze **causali**.

- **causale**: significa una dipendenza di uno o più fatti ← da uno o più altri (qualcosa implica qualcosa d'altro).
- **correlato**: significa la coincidenza (per intanto ignoti motivi) di due $\&\rightarrow$ fatti (due cose hanno una qualche relazione tra di loro).



È un insegnamento sin' dai tempi degli antichi filosofi greci, che il "cum hoc ~~ergo~~ propter hoc" (assieme ~~quindi~~ dipendente) è un grave errore di deduzione logica.

Malgrado sia banale questa differenza terminologica, si incontrano numerosissime "statistiche" che cortocircuitano correlato e causale tirando così molto spesso delle conclusioni completamente errate - non solo dei giornalisti, politici e responsabili di marketing, ma anche dei ricercatori stessi che si squalificano da brutto con simili mosse (forse perché non è più di moda l'insegnamento di logica formale). Personalmente smetto a

leggere un tale elaborato appena me ne accorgo, → cestino. Se mi capita più volte nella stessa rivista, in futuro non la guardo più e mi dedico a qualcosa di meno futile o più divertente. 8>)__ .

Sono trattati i seguenti argomenti:

Correlazione e causalità ♦ Analisi di regressione ♦ Statistiche di paragone

1. Correlazione verso causalità

[Correlazione \(statistica\) it.Wikipedia](#) ♦ [Causalità it.Wikipedia](#) ♦ [Sincronicità it.Wikipedia](#) ♦ [Cum hoc ergo propter hoc de.Wikipedia](#) ♦ [Correlation does not imply causation en.Wikipedia](#)

cum hoc ergo propter hoc (lat. ca. „assieme quindi dipendente“) indica un'errore logico con il quale due eventi coincidenti vengono dichiarati come causa ed effetto. Chi commette questo duplice errore:

- pretende in primo luogo (senza esame e/o motivazione), un nesso tra due qualsiasi eventi (probabilmente casuali),
- dichiara in secondo luogo, aleatoriamente un'evento come causa e l'altro evento quindi come effetto.

Si tratta di uno dei più divulgati errori logici dei nostri tempi (e forse anche di tempi passati), riportati ad infinitum dai mass media.

Esempio (secondo [Bobby Henderson](#)): La stima di **pirati** sui mari del globo, che ancora nel 1820 era di 35'000 si è ridotto fino al 2000 su 17. Nel medesimo lasso di tempo, la **temperatura aerea globale** è salita da 14.2°C a 15.8°C. Conoscendo i due "eventi":

- un giornalista scrive: *La diminuzione della pirateria fa aumentare il riscaldamento climatico!*
- un altro scrive: *L'aumento della temperatura globale fa diminuire gli atti di pirateria!*
- un terzo scrive: *Secondo affidabili fonti scientifiche: sospettato nesso tra pirateria e riscaldamento climatico!*

Un quarto non scrive niente in merito (perché lo ritiene una coincidenza completamente casuale) e viene poi seriamente chiamato all'ordine dal suo capo redattore, perché non prende posizione incirca *l'esistenziale discorso pubblico in merito alla pirateria e il riscaldamento globale*.

Approfondimenti vedi → Classificazione delle fallacie: [Fallacie di presupposizione it.Wikipedia](#).

Sono trattati i seguenti argomenti:

[Una correlazione rappresentata come causalità](#) ♦ [Dalla correlazione alla causalità](#) ♦ [Implicazioni della confusione](#) ♦

1.1 Correlazione ≠ causalità

(Una correlazione rappresentata come causalità)

Di seguito un esempio di apparenza scientifica pubblicato di associazione di pediatri Statunitensi, commentato poi da [Salman Khan](#) dalla [KHANacademy](#)

Collazione combatte l'obesità di adolescenti

In questo studio, pubblicato in Pediatrics dei ricercatori analizzarono i dati di 2'216 adolescenti durante un periodo di 5 anni dalle scuole pubbliche a Minneapolis-St. Paul, Minnesota.

I ricercatori scrivono, che gli adolescenti che fanno regolarmente collazione consumano meno grassi saturi ma più fibre vegetali e carboidrati che i loro compagni che saltano questo pasto. In più quelli che fanno regolarmente collazione sembrano più attivi fisicamente che i loro compagni saltapasti.

Nel tempo, costatarono i ricercatori, gli adolescenti che fanno regolarmente collazione aumentano meno di peso e hanno un'indice di peso corporeo inferiore che quelli "salta-

Eating Breakfast May Beat Teen Obesity

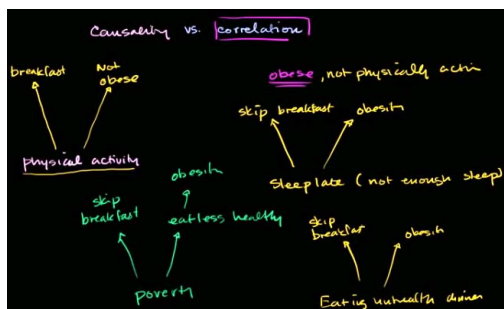
In the study, published in Pediatrics, researchers analyzed the dietary and weight patterns of a group of 2,216 adolescents over a five-year period (1998-1999 to 2003-2004) from public schools in Minneapolis-St. Paul, Minn.

The researchers write that teens who ate breakfast regularly had a lower percentage of total calories from saturated fat and ate more fiber and carbohydrates than those who skipped breakfast. In addition, regular breakfast eaters seemed more physically active than breakfast skippers.

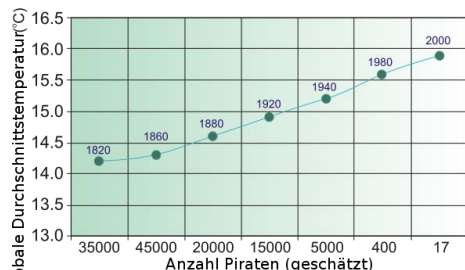
Over time, researchers found teens who regularly ate breakfast tended to gain less weight and had a lower body mass index than breakfast skippers.

cliccare per accedere

collazione".



Globale Durchschnittstemperatur vs. Anzahl Piraten



Pirati e riscaldamento climatico

Salman Khan commenta questo peccato mortale (anti)scientifico (subergia & acedia) ca. al seguente:

Collazione combatte l'obesità di adolescenti
 In questo studio, pubblicato in *Pediatrics* dei ricercatori analizzarono i dati di 2'216 adolescenti durante un periodo di 5 anni dalle scuole pubbliche a Minneapolis-St. Paul, Minnesota.

I ricercatori scrivono, che gli adolescenti che fanno regolarmente collazione consumano meno grassi saturi ma più fibre vegetali e carboidrati che i loro compagni che saltano questo pasto. In più quelli che fanno regolarmente collazione sembrano più attivi fisicamente che i loro compagni saltapasti.

Nel tempo, costatarono i ricercatori, gli adolescenti che fanno regolarmente collazione aumentano meno di peso e hanno un'indice di peso corporeo inferiore che quelli "salta-collazione".

Ipotetiche catene causali

collazione → non obeso
 collazione → fisicamente attivo
 salta-collazione → obeso

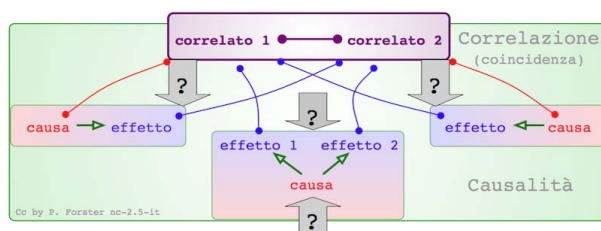
Deduzioni correlative

obeso ← fisicamente inattivo
 obeso ← dieta malsana
 obeso ← cena malsana
 obeso ← tardi a letto
 obeso ← mancanza di sonno
 obeso ← povertà
 obeso ← salta-collazione

In base ai ragionamenti si potrebbe rilevare altri dati per verificare oppure falsificare i rami delle catene causali ipotizzate (metodo scientifico).



1.2 Dalla correlazione alle ipotetiche causalità



Dai dati statistici, rilevati o misurati ed elaborati regola d'arte (come descritto nel capitolo sulla statistica descrittiva) si possono dedurre delle coincidenze (correlazioni) più o meno strette e con vari indicatori sulla distribuzione e l'affidibilità del campione. Ulteriori elaborazioni *inferenziali* riescono poi a chiarire variabilità e

significatività correlativa.

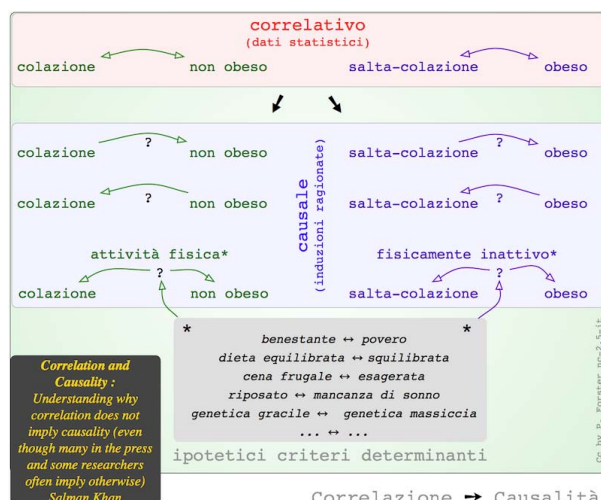
Non si riesce a dimostrare delle causalità con dei strumenti statistici.

Non si riesce p.es. a decifrare, se la correlazione tra colazione ↔ obesità dimostrata è dovuto al fatto:

- che l'obesità è un'effetto ← di mancante colazione (causa) ?,
- che la "salta-colazione" è un effetto ← dell'obesità (causa) ? oppure
- che c'è un terzo o vari altri fattori determinativi ? (p.es. attività fisica, sonno, genetica, televisione, ...) → causante sia la colazione → che l'obesità.

A partire da questo punto non sono più richieste delle conoscenze statistiche bensì di creatività mentale, professionalità nel tema trattato, una dose di scetticismo, abitudine di formulare delle ipotesi e di falsificarli. Alla fine rimangono (forse) poche supposizioni plausibili. Bisogna poi formularli, giustificarli e commentarli; cioè mettersi in discussione (che non è cosa di tutti).

Chi non si sente a fare tutto questo lavoro speculativo e di ragionamento pubblico per



Correlazione → Causalità

arrivare a una ipotesi (plausibile e stringente), meglio che ponga delle domande sugli punti non evidenti e non chiari. Partecipa meglio così alla progressione di sapere che sparando un'altro giudizio o un'ulteriore credenza (che non tiene) alla pattumiera esistente in merito.



1.3 Implicazioni della confusione

Per evitare delle brutte figure (come nell'esempio cui sopra), le riviste scientifiche serie usano il metodo del "[Peer Review](#)" (revisione paritaria): sottopongono un'articolo del genere (anonimizzato) a degli collaudati esperti scientifici, sia del tema trattato che di statistica, chiedendo una valutazione formalizzata. Certo che queste procedure sono impegnative, ma una rivista scientifica rinomata si mantiene la fiducia dei lettori solo in questo modo.



I politici, giornalisti e manager del marketing coinvolti usano comunque questi testi per i loro scopi, anche se raramente sono interessati (o capaci) di interpretare le "sfumature" ragionevoli (che sarebbero le causalità).

Gli effetti sociali possiamo poi ammirare nella pubblicità televisiva per i musli a colazione che si servono spudoratamente di queste oscenità "scientifiche".

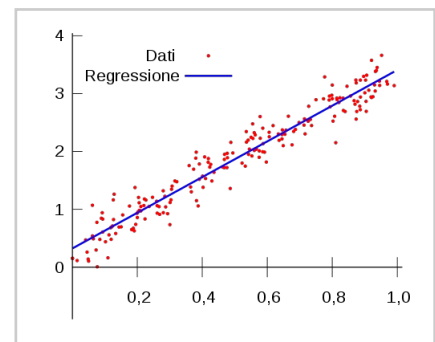
2. Analisi di regressione

[Regressione](#) *it.Wikipedia* [Regression analysis](#) *en.Wikipedia*

Delle ipotesi formulate (come descritto sopra) sono al solito controllate con la "curva intercetta dei minimi quadrati". Il metodo più divulgato è chiamata *regressione lineare*, con la quale si tenta di stimare la "bontà di adattamento" della ipotesi e la significatività statistica dei parametri stimati. Si riesce a farsi un'idea, se un'ipotesi può essere plausibile o se è da scartare.

Visto che il metodo richiede specifiche conoscenze statistiche (e non solo) non intendo di approfondire il tema. Di seguito le spiegazioni di *it.Wikipedia*.

L'analisi della regressione è una tecnica usata per modellare ed analizzare una serie di dati che consistono in una variabile dipendente e una o più variabili indipendenti. La variabile dipendente nella equazione di regressione è modellata come una funzione delle variabili indipendenti più un termine d'errore. Quest'ultimo è una variabile casuale e rappresenta una variazione non controllabile e imprevedibile nella variabile dipendente. I parametri sono stimati in modo da descrivere al meglio i dati. Il metodo più comunemente utilizzato per ottenere le migliori stime è il metodo dei "minimi quadrati" (OLS), ma sono utilizzati anche altri metodi.

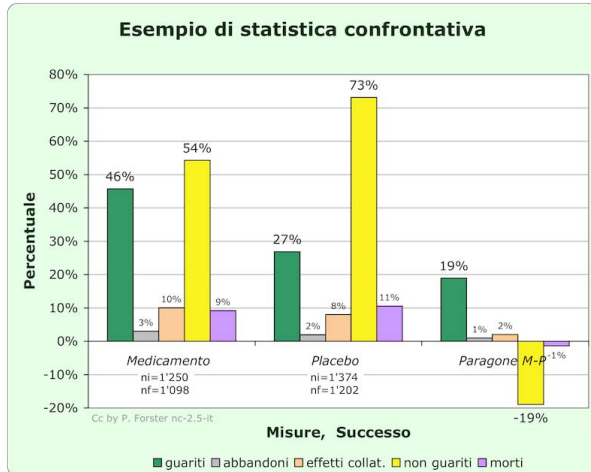


L'analisi della regressione può essere usata per effettuare previsioni (ad esempio per prevedere dati futuri di una serie temporale), inferenza statistica, per testare ipotesi o per modellare delle relazioni di dipendenza. Questi usi della regressione dipendono fortemente dal fatto che le assunzioni di partenza siano verificate. L'uso dell'analisi della regressione è stato

criticato in diversi casi in cui le ipotesi di partenza non possono essere verificate. Un fattore che contribuisce all'uso improprio della regressione è che richiede più competenze per criticare un modello che per adattarlo. it.Wikipedia

3. Statistiche di paragone

Placebo it.Wikipedia ♦ *Nocebo it.Wikipedia* ♦ *Trattato su placebo e nocebo MmP*



Il gratticapo di correlazione → causalità **non** si pone in delle **statistiche di paragone** (confrontative): si paragonano direttamente i correlati di due statistiche (p.es. due medicinali per la stessa malattia o di un medicinale e un **placebo**).

La difficoltà in questo caso è di garantire delle condizioni paragonabili (più simili possibile) per le "due statistiche". Questo richiede una meticolosa pianificazione del progetto e la garanzia, che i partecipanti ai test siano all'ignoto (p.es. della somministrazione di medicinale o placebo in un indagine a doppio cieco).

Arrivando al dunque, si trae la differenza tra le due correlazioni per poter giudicare l'effetto delle due misure.

Nell'esempio (finto) cui sopra, i medici seri parleranno di un successo del medicinale del ca. 20%, mentre i responsabili del commercio parlerebbero probabilmente di un successo del ca. 50% (dimenticandosi che l'effetto placebo è del ca. 30%).

N.B. In gergo medico si chiamano "placebo" e "nocebo" questo che statisticamente sono delle variabili casuali e parametri ignoti, che creano degli effetti al di là del medicinale. Un classico esempio di arroganza medica per non dover studiare degli effetti oltre farmacologici che influiscono evidentemente anche il percorso di una malattia.

Statisticamente sarebbe desiderabile di fare non solo un paragone tra medicinale e placebo ma anche con "far nulla". Certo che in medicina questo non è sempre fattibile per motivi etici, ma ci sono anche delle malattie con le quali si potrebbe: mia nonna curatrice disse, che un'influenza trattato dal medico dura quindici giorni, curato in proprio due settimane.

4. Allegati

4.1 Sitografia

MedPop

- 🌟 **Biometria**
description: Biometria, Statistica, epidemiologica, descrittiva, inferenziale, Indici di peso corporeo, Rischio peso, Colesterolemia, Lipidemie, MedPop, Medicina popolare
- 🌟 **Statistica epidemiologica**
description: Insidie, statistiche, Visualizzazione grafica, Standardizzazione, Statistiche strane, Statistica descrittiva, Statistica inferenziale, MedPop, Medicina popolare
- 🌟 **Statistica descrittiva**
description: Dati, statistici, fondamentali, Media, moda, mediana, Deviazione standard, Indicatori, distribuzioni, Campo di variazione, Coefficiente di variazione, interpretazione, Lavoisier, somma, totale, Medie, mobili, Ovvietà, paradosso di Simpson, fattore d'inganno, Contesto, MedPop, Medicina popolare
- 🌟 **Statistica inferenziale**
description: campionamento, sondaggi, probabilistico, randomizzazione, Errori, caso, Variabilità, stima, Errore standard, limiti fiduciali, significatività, chi-quadro, Yates, Fisher, Meta-analisi, Eterogeneità, MedPop, Medicina popolare
- 🌟 **Indici di peso corporeo**
description: Formula di Broca, Body mass index, BMI, IMC, giovani, Indice di massa corporea, mortalità
- 🌟 **Rischio peso**
description: biometria, antropometria, medicina popolare, profana, laica, complementare, rischio peso, aspettativa vita, corporeo, Indici, IMC, indice di massa corporea, BMI, body mass index, Valutazione, biologiche, biologica, Implicazioni, mediche, sanitarie, Determinanti, peso, Valutazione peso, Misure, pazienti, sottopeso, sovrappeso, magrezza, obesità, Peter Forster, Daniela Rüegg
- 🌟 **Colesterolemia**
description: medicina popolare, colesterolo, colesterolemia, mortalità, stress, prevenzione, sanitocrazia, lipoproteine, Lipidemie, Valutazione, rischi, HDL, LDL, valori di referenza, variazioni, statine, resa, emboli, trombi, Correzioni, infarto, occlusioni vasali
- 🌟 **Lipidemie**
- **Trattato su placebo e nocebo**

Wikipedia

- **Statistica**: it.Wikipedia
- **Epidemiologia**: it.Wikipedia
- **Statistica descrittiva**: it.Wikipedia
- **Statistica inferenziale**: it.Wikipedia
- **Statistica esplorativa**: it.Wikipedia
- **Correlazione**: it.Wikipedia
- **Causalità**: it.Wikipedia
- **Sincronicità**: it.Wikipedia
- **Cum hoc ergo propter hoc**: de.Wikipedia
- **Correlation does not imply causation**: en.Wikipedia
- **Regressione**: it.Wikipedia
- **Regression analysis**: en.Wikipedia
- **Placebo**: it.Wikipedia
- **Nocebo**: it.Wikipedia

Vari

- **galenotech**: FACOLTA' di FARMACIA e MEDICINA corso di laurea in SCIENZE FARMACEUTICHE APPLICATE: *La Sapienza* Università di Roma
- **EPIDEMIOLOGIA STATISTICA**: : FACOLTA' di FARMACIA e MEDICINA corso di laurea in SCIENZE FARMACEUTICHE APPLICATE: *La Sapienza* Università di Roma
- **Tubes**: filmini didattici **KHAN**academy
- **Statistics Tubes**: filmini didattici **KHAN**academy

4.2 Commenti

alla pagina *Biometria* / 🏠 *Correlazioni statistiche e causalità*: ev. cliccare sul titolo per accedere.

Peter — 29 December 2011, 11:21

Test

Proveniente da Redazione: <http://pforster.no-ip.org/~admin/pmwiki/pmwiki.php/Biometria/CorrelazioneCausalita>
ultima modifica December 29, 2011, at 06:13 PM
Cc by P. Forster nc-2.5-it