

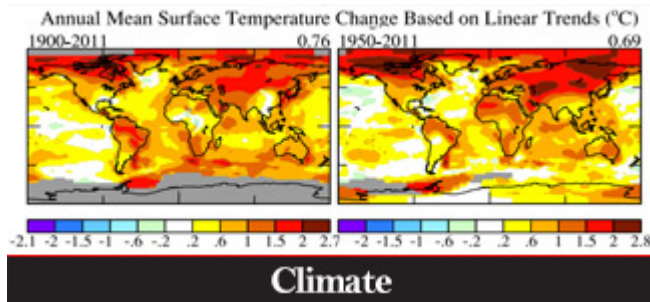
Cultura Progressista | Studiosi e ladri

Scienza del clima per tutti: correlazione e causalità (aggiornato)



Brian Angiss

5 anni fa



La correlazione non prova la causalità, ma nemmeno la mancanza di correlazione *confuta* la causalità.

Per leggere altri articoli di questa serie, clicca [qui](#)

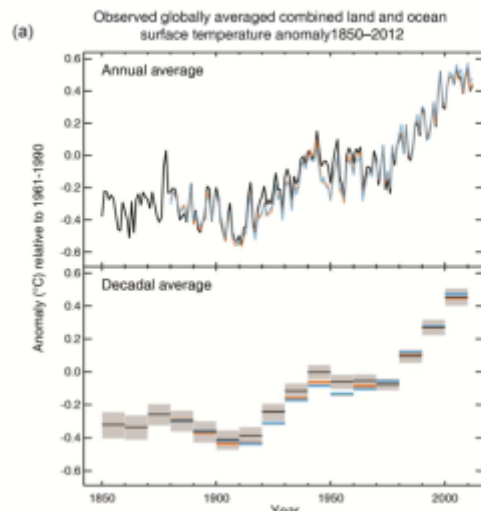
AGGIORNAMENTO: dopo alcuni feedback, ho aggiunto una serie di grafici agli elenchi di numeri seguenti per aiutare a chiarire.

Cosa significa quando qualcuno dice "la correlazione non prova la causalità?" È una frase comune pronunciata da individui che negano che il cambiamento climatico stia avvenendo, che sia dominato dalle emissioni industriali di gas serra come l'anidride carbonica (CO₂) e che i cambiamenti saranno dirompenti sia per gli ecosistemi che per la società umana (alias industriale interruzione del clima, riscaldamento globale o cambiamento climatico). Per capire perché questi negazionisti del cambiamento climatico industriale sono sbagliati, dobbiamo prima capire di cosa stanno parlando quando parlano di correlazione, causalità e relazioni tra i due.

Cos'è la correlazione?

Twelfth Session of Working Group I Approved Summaries for Policymakers

Figure SPM.1 [FIGURE SUBJECT TO FINAL COPYEDIT]

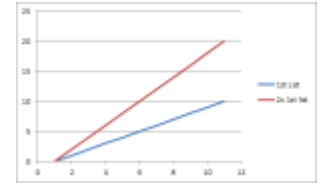


IPCC AR5 WG1 Variazione decennale della temperatura globale (IPCC)

La correlazione è una misura di quanto siano correlati tra loro due diversi insiemi di numeri. Ad esempio, supponiamo di avere due insiemi di numeri come segue:

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

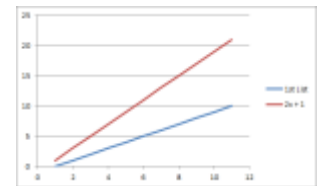
0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20



Possiamo dire guardando le due liste che il secondo insieme di numeri è esattamente due volte ($2x$) il primo insieme. In effetti, sono correlati al 100% perché ogni numero nel secondo elenco è **esattamente** 2 volte il primo set. Possiamo aggiungere uno a ogni numero, dandoci una terza serie di numeri:

1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21

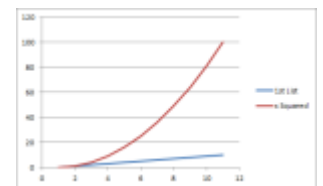
In questo caso, i secondi numeri sono ancora correlati al 100% perché sono **esattamente** $2x+1$ il primo insieme di numeri.



Ma la correlazione standard presuppone che la relazione tra due insiemi di numeri sia lineare, della forma $Y = mX + b$ (dove ' m ' è la pendenza e ' b ' è l'offset – 2 e 1 rispettivamente nell'esempio sopra). Se interrompiamo quella relazione lineare, la correlazione diminuisce. Nel caso in cui abbiamo un nuovo insieme di numeri in cui ogni voce è uguale al primo insieme moltiplicato per se stessa (x^2), la correlazione scende dal 100% (normalmente visualizzato dai programmi di matematica come "1") a 0,963.

0, 1, 4, 9, 16, 25, 36, 49, 64, 81, 100

Se la relazione lineare ha una pendenza negativa (nel caso in cui abbiamo numeri negativi da 0 a -10, per esempio), la relazione è -1 (-100%). E se non esiste una relazione lineare tra due insiemi di numeri, la correlazione sarà esattamente 0. In pratica, solo elenchi molto grandi di numeri casuali hanno una correlazione 0.



Cos'è la causalità?

"Causa" è solo una parola matematico/scientifica di fantasia per una relazione di causa-effetto tra due cose. Nel nostro primo esempio sopra, il secondo elenco di numeri era esattamente 2 volte il primo elenco di numeri perché li ho definiti in questo modo. Se cambiassi il primo set in modo che fosse qualcosa di diverso dai numeri interi compresi tra 0 e 10 inclusi, anche il secondo set cambierebbe. E poiché anche il terzo e il quarto insieme stanno cambiando con il primo,

possiamo dire che sono tutti **causati** dal primo insieme di numeri e dalla relazione matematica che ho definito per ciascuno di essi.

Quindi da dove viene l'intera cosa "correlazione non è uguale a causalità"?

Diciamo di avere due elenchi di numeri in cui un elenco è il doppio dell'altro, in questo modo:

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20

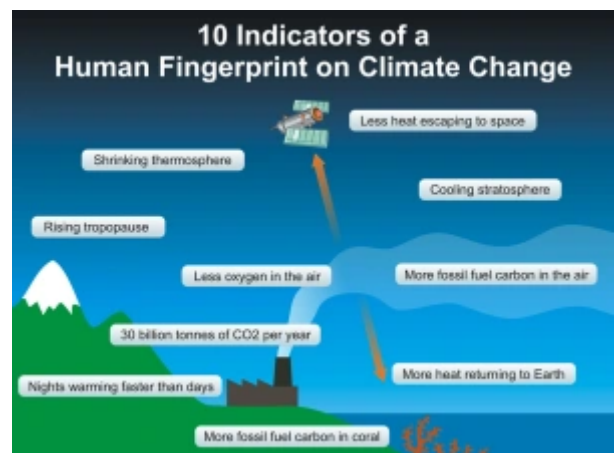
Sappiamo dall'alto che saranno correlati al 100% tra loro. Ma diciamo anche che il primo sono i giorni del mese mentre il secondo è il numero di persone sedute in sala d'attesa. Chiaramente non possiamo semplicemente moltiplicare i giorni della settimana per 2 e convertire magicamente un giorno in una persona (o 10 giorni in 20 persone). Quindi in questo esempio i due insiemi di numeri sono correlati al 100%, ma questa è solo una coincidenza dei due insiemi di numeri.

Questo è uno dei problemi più complicati della scienza, e in particolare della scienza del clima. C'è un alto grado di correlazione tra l'aumento dei livelli di CO₂ e l'aumento delle temperature globali, ma potrebbe essere solo una coincidenza dei numeri. Per fortuna, c'è un gruppo di scienziati che si sono impegnati a capire esattamente come determinare se la relazione tra CO₂ e temperatura

è semplicemente una coincidenza o se l'una è causata dall'altra. I risultati pubblicati di questa ricerca sono chiamati "studi di attribuzione" e

l'aumento della temperatura globale è stato [attribuito alla CO₂ entro il 2000](#) ed è stato [verificato](#) e [ricontrollato](#) molte volte da allora. A questo punto gli studi sull'attribuzione sono diventati così forti che non c'è praticamente dubbio che la CO₂ sia il principale motore del riscaldamento globale.

Che dire dei periodi in cui non c'è molta correlazione tra due serie di numeri?



Impronte digitali dell'influenza umana sul clima terrestre (scienza scettica)

Ci sono alcune persone che rifiutano gli studi di attribuzione, tuttavia, anche se è praticamente necessario [infrangere le leggi della fisica](#) perché la

CO₂ non sia il principale motore del

riscaldamento globale. Una delle giustificazioni offerte per rifiutare la CO₂ come causa principale

è affermare (erroneamente) che non c'è stata alcuna correlazione tra CO₂ e aumento della

temperatura globale negli ultimi 15-20 anni. Se avessero ragione, la mancanza di correlazione non significherebbe che non può esserci causalità?

Affatto. In effetti, ci sono casi in cui la causalità è ben nota, ma c'è pochissima correlazione ovvia. Per vedere come funzionerebbe, diamo un'occhiata a quegli elenchi di numeri. Inizieremo dall'elenco originale di numeri:

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Ma al secondo elenco, aggiungiamo un numero casuale che è -1, 0 o 1:

-1, 3, 4, 6, 8, 11, 12, 14, 15, 19, 21

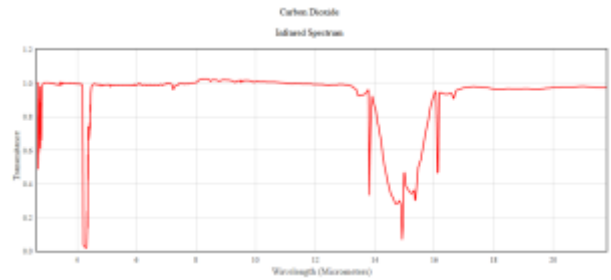
Il rumore aggiuntivo creato dai numeri casuali riduce la correlazione da 1 a 0,994. La correlazione non è più perfetta, ma è così vicina che possiamo ancora facilmente vedere che i due numeri sono ancora correlati e, poiché sappiamo che ho definito io stesso la relazione matematica, sappiamo che c'è ancora la causalità. Ma cosa succede se facciamo in modo che il numero casuale che aggiungiamo alla seconda lista sia compreso tra -10 e +10? Ora l'elenco diventa questo:

8, 9, -2, 14, 16, 0, 5, 24, 15, 12, 29

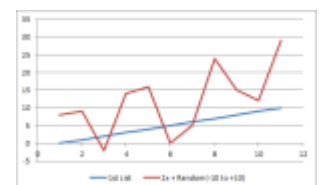
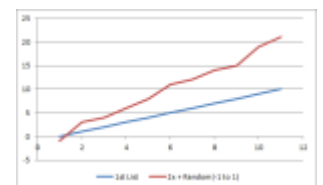
In questo caso, la correlazione è scesa da 1 a 0,57. È più correlato dei numeri casuali, ma chiaramente non è molto buono. Quindi, anche se

sappiamo che esiste una relazione causale tra il primo e il terzo elenco di numeri, non possiamo vederlo nella matematica. *O meglio, non possiamo ancora vederlo.*

Se aggiungessimo più dati, tuttavia, la correlazione risalirebbe lentamente. Aggiungo altri 10 numeri e la correlazione aumenta a circa 0,9. Aggiungo 20 numeri e la correlazione aumenta a 0,93 circa. Aggiungo 30 numeri e la correlazione aumenta a 0,97 circa. E così via.



Spettro misurato dal NIST delle bande di assorbimento dell'anidride carbonica

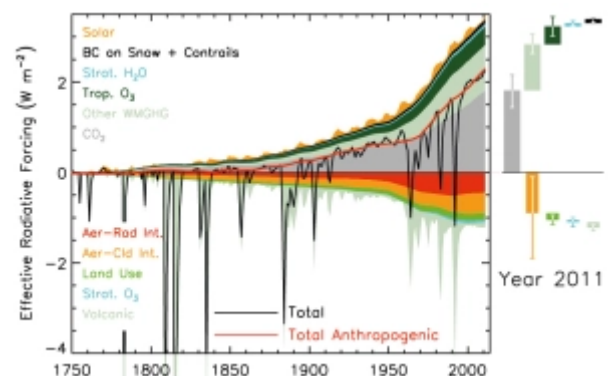


La ragione di ciò è che alla fine i numeri sono così grandi che la quantità di errore rispetto alla quantità di segnale diminuisce. Al numero 10, aggiungere un errore di +/- 10 numeri interi è un enorme errore del 100%. Ma al numero 100, +/-10 numeri interi di errore sono solo un errore del 10%. L'aggiunta di più dati alla fine rivelerà sempre una relazione di causa-effetto, ma potrebbe essere necessario un enorme elenco di numeri per vedere con precisione che il secondo set di dati è stato causato dal primo. Nel mondo reale, enormi elenchi di numeri possono richiedere molto tempo per essere raccolti: decenni o addirittura secoli nel caso della scienza del clima.

Un'alternativa è capire come trasformare i grandi errori in piccoli errori. Nel caso precedente, sarebbe come capire come trasformare +/-10 numeri interi di errore in +/-1 numeri interi di errore. Quando possiamo ridurre l'errore in qualche modo, possiamo rilevare la vera relazione causa-effetto molto più rapidamente. Esistono diversi modi per farlo nei sistemi reali, ma questi metodi esulano dallo scopo di questo particolare articolo.

Dagli esempi precedenti possiamo vedere che cambiamenti di breve durata possono nascondere una nota relazione causa-effetto e far *sembrare* che l'effetto non sia più determinato dalla causa.

Il clima terrestre è un sistema complicato e ci sono molte cose oltre alla CO₂ che influenzano la temperatura media della Terra. Alcune di queste cose includono se siamo in un [El Nino o La Nina](#), se c'è stata una [grande eruzione vulcanica di recente](#), in quale parte del [ciclo solare di 11 anni](#) ci troviamo, il periodo dell'anno e quanto [inquinamento viene emesso dalla combustione di carbone, petrolio e biomassa](#) (legno, sterco, ecc.). Quando gli scienziati osservano El Nino/La Nina, ad esempio, scoprono che esiste una correlazione molto elevata tra quegli episodi e le variazioni di breve durata (dell'ordine di alcuni anni) della temperatura terrestre. La Terra si riscalda durante un El Nino perché l'Oceano Pacifico tropicale sta rilasciando energia immagazzinata nell'atmosfera e la Terra si raffredda durante un La Nina perché il Pacifico tropicale sta assorbendo energia dall'atmosfera. Lo stesso vale per il modo in cui la temperatura terrestre si raffredda dopo un'eruzione vulcanica (a causa del rilascio di grandi quantità di anidride solforosa nell'atmosfera superiore). Ma nel corso dei decenni, la correlazione si interrompe perché qualcos'altro sta guidando una tendenza al riscaldamento a lungo termine.



Riscaldamento globale naturale e causato dall'uomo dal 1750 al 2011 (quinto rapporto di valutazione dell'IPCC, gruppo di lavoro 1, figura 8-18)

E gli studi di attribuzione hanno dimostrato che "qualcos'altro" è CO_2 .

Quindi la prossima volta che ti imbatti in qualcuno che afferma che "la correlazione non prova la causalità", ricordati di guardare i dati che stanno usando per fare tale affermazione. Se non ci sono *molti* dati, o se stanno usando un piccolo sottoinsieme dei dati disponibili (diciamo gli ultimi 15-20 anni di temperature globali invece del record completo di 136 anni), sii scettico. Perché la mancanza di correlazione non *smentisce* nemmeno la causalità.

NOTA: Tutti i numeri che utilizzo sopra possono essere verificati semplicemente inserendo gli elenchi di numeri che ho utilizzato in un foglio di calcolo Excel e utilizzando la funzione "correl". Se vuoi controllare come funzionano anche le cose con l'errore aggiunto, ho usato la funzione "randbetween" per quello. Nota che il generatore di numeri casuali significa che probabilmente non otterrai mai l'elenco esatto di numeri "casuali" che avevo sopra, ma sarai comunque in grado di vedere come aumentano le correlazioni mentre aggiungi i dati.

Condividi questo:



Imparentato

[Climate Illogic: l'interruzione del clima industriale non è un concorso di popolarità](#)

16 luglio 2013

In "Ambiente/Natura"

[Tom Harris - ipocrita venditore ambulante di editoriali ingannevoli sul cambiamento climatico \(corretto\)](#)

16 febbraio 2015

In "Ambiente/Natura"

[Roy Spencer attacca l'Anti-Defamation League per aver denunciato il suo uso dei "nazisti del riscaldamento globale"](#)

1 marzo 2014

In "Ambiente/Natura"

Categorie: [Istruzione](#) , [Scienza/Tecnologia](#)

Tag: [studi di attribuzione](#) , [anidride carbonica](#) , [causalità](#) , [clima](#) , [cambiamento climatico](#) , [Scienza del clima per tutti](#) , [correlazione](#) , [riscaldamento globale](#) , [interruzione del clima industriale](#) , [scienza](#)