

Nota in margine al congresso SIBioC 2000: sull'uso distorto del coefficiente di correlazione nel confronto fra metodiche

Marco B.L. Rocchi

Istituto di Biomatematica, Facoltà di Farmacia, Università di Urbino

Questa nota trae spunto dalla partecipazione di chi scrive alla sessione poster del giorno 14 settembre 2000, nell'ambito del Congresso SIBioC 2000 (32° Congresso Nazionale della Società Italiana di Biochimica Clinica e Biologia Molecolare Clinica) svoltosi a Rimini dal 12 al 15 settembre 2000. In quella circostanza, occupandosi l'autore di Biometria e Statistica Medica, il suo interesse era primariamente rivolto ai metodi statistici utilizzati. In questo contesto l'autore ha potuto riscontrare che numerosi poster contenevano il medesimo errore metodologico (dal punto di vista statistico, ovviamente): nel confronto fra due metodiche (da intendersi in senso lato: strumentazioni, kit diagnostici, ecc.) le conclusioni di uguale attendibilità delle stesse erano basate unicamente sul calcolo del coefficiente di correlazione r .

Incuriosito dal ripetersi di questo errore, lo scrivente si è dato la pena di valutarne la frequenza: ebbene, su 102 poster affissi nella sessione succitata, 29 erano affetti (alcuni ripetutamente) dallo stesso uso scorretto di questo coefficiente. Si tenga presente che la percentuale già di per sé elevata (28.4%) con tutta evidenza sottostima il fenomeno, dal momento che numerosi poster si ponevano tutt'altre questioni dal confronto di metodiche. Va peraltro aggiunto che dei 29 poster citati alcuni presentavano la valutazione di concordanza fra metodiche in margine a una ricerca di più ampio respiro, mentre numerosi altri la identificavano con la finalità stessa dell'indagine.

Il problema appare tanto più serio se si pensa che in questo modo si corre il rischio di avallare l'impiego di nuove metodiche, erroneamente giudicate di pari efficacia e affidabilità delle metodiche standard. E' una ben magra consolazione la considerazione che l'errore citato è uno tra i più diffusi all'interno della letteratura biomedica.

Il tutto nasce presumibilmente dalla diffusa incomprensione del significato di correlazione, e vale quindi forse la pena di rammentarlo in questa sede.

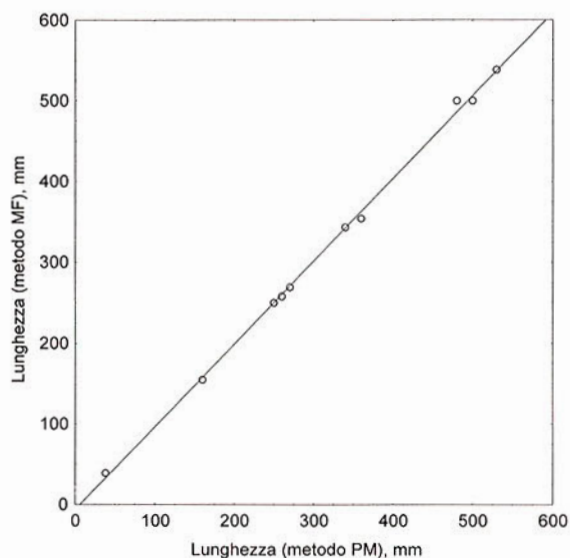
Il coefficiente di correlazione r di Bravais-Pearson (questa la completa dicitura) quantifica il grado del legame lineare tra una coppia di variabili. Quando $r = 1$ si ha una correlazione diretta perfetta; quando $r = 0$ le due variabili sono totalmente incorrelate; i valori intermedi hanno ovviamente significati altrettanto intermedi (visto il contesto, si trascura in questa sede il caso di $r = -1$, e più in generale di valori negativi, indici di correlazione inversa).

Traducendo in termini più generici, ma forse per ciò stesso più comprensibili, si può dire che il coefficiente di correlazione misura in che modo due variabili "si muovono insieme", e quindi, nel caso particolare di nostro interesse, come una variabile cresce al crescere dell'altra.

Il punto è proprio questo: nel caso di confronto fra due metodiche che misurino la stessa variabile, è assolutamente ovvio che - a meno di una enorme inaffidabilità di una delle due metodiche - il coefficiente di correlazione risulterà elevato, e con ogni probabilità prossimo a 1. Infatti, cosa c'è di più ovvio che misurare lo stesso fenomeno con due strumenti, e scoprire che quando uno strumento dà un valore elevato, anche l'altro

Figura 1

Confronto tra misure di lunghezza di 10 oggetti effettuate con il palmo della mano (metodo PM) e con il metro da falegname (metodo MF). L'equazione della retta di regressione è risultata: $y = -6,032 + 1,025 \cdot x$; il coefficiente di correlazione: $r = 0,99919$



fornisce un valore elevato? Questo risultato non rende affatto le due metodiche ugualmente affidabili; il coefficiente resterebbe elevato anche se uno dei due metodi desse luogo a fenomeni di inaccuratezza e/o di imprecisione.

L'equivalenza di due metodiche può dunque essere ricavata dallo studio delle differenze ottenute con le due tecniche per ciascuna osservazione (tecnicamente: analisi delle differenze fra misure entro soggetto).

Questo approccio consente di valutare - attraverso la media delle differenze - la accuratezza (ad esempio eventuali errori sistematici), e - attraverso la deviazione standard - la precisione delle metodiche. A questo proposito si suggerisce la lettura dei lavori citati in bibliografia.

Ma poiché a volte un esempio val più di mille parole, proponiamo anche un semplice esperimento realmente effettuato dall'autore e da chiunque facilmente riproducibile: mettiamo a confronto due "strumenti" per la misura di lunghezze, quali un metro da falegname e il palmo della mano dello sperimentatore.

Misuriamo con entrambi gli "strumenti" una serie di oggetti, e calcoliamo poi il coefficiente di correlazione fra le due serie di misure. La Figura 1 riporta il diagramma di dispersione dei dati ($n = 10$) dall'autore realmente ottenuti in questo esperimento. Il coefficiente di correlazione risulta, com'era prevedibile, molto elevato (addirittura $r = 0,999$).

Ogni commento a questo risultato, ancorché superfluo, è riassumibile in una semplice domanda: nonostante l'elevato valore di r , chi si farebbe costruire un armadio da un falegname che prende le misure a palmi?

RIFERIMENTI BIBLIOGRAFICI

1. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. *Statistician* 1983; 32: 307-320.
2. Altman DG. Statistica ed etica nella ricerca medica. In: Gore SM, Altman DG. *La statistica nella pratica medica*. Roma, Il Pensiero Scientifico Editore 1985; 34-35.
3. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurements. *Lancet* 1986; 1: 307-309.
4. Duca P. Validità delle osservazioni in medicina. *Aggiorn Med* 1989; 1: 26-32.