

Metodi statistici per la Neuropsicologia Forense

A.A. 2023/2024

Giorgio Arcara

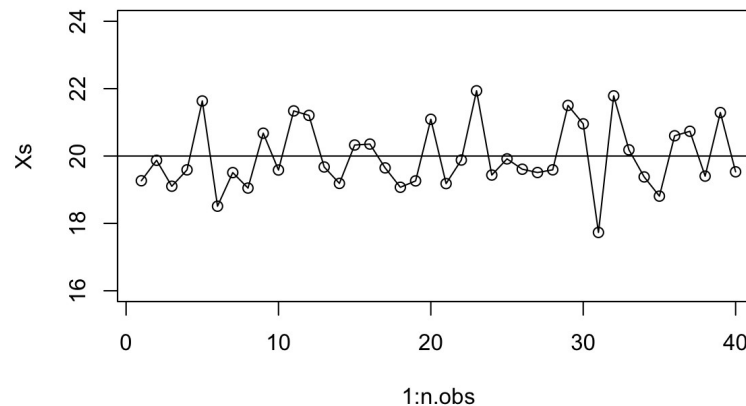
IRCCS San Camillo, Venezia
Università degli Studi di Padova





8. Altri utilizzi dei punteggi dei test

Supponiamo di testare un paziente tramite il test di Linguaggio Figurato
1. Il paziente ottiene un punteggio di 12. Dopo 6 mesi viene valutato nuovamente e ottiene un punteggio di 10. Cosa possiamo dire del paziente?



Ricordiamo (vedi slides affidabilità) che in caso di ripetizioni di una misurazione ci aspettiamo comunque delle oscillazioni nel punteggio osservato. Il punto è capire: la variazione che ho osservata è riconducibile a queste oscillazioni (dovute a imprecisione?) oppure ad un cambiamento “vero”.



Valutare i cambiamenti nel tempo

La valutazione dei cambiamenti nel tempo è una questione spinosa in neuropsicologia clinica ed è molto comune.

Perché valutare lo stesso soggetto/paziente più volte?

- per indagare se è cambiato in seguito ad un trattamento di neuroriabilitazione.
- per indagare se presenta un declino nel funzionamento cognitivo, possibilmente segnale di un deterioramento progressivo.
- Per indagare ha subito alterazioni cognitive in seguito ad un evento, come un'operazione neurochirurgica per rimuovere un tumore.



Valutare i cambiamenti nel tempo

Il problema della valutazione nel tempo

È assolutamente normale che un soggetto non ottenga esattamente lo stesso punteggio in due diverse valutazioni.

- *Effetti casuali*: legati alla precisione del test e catturati dall'affidabilità test-retest
- *Effetti sistematici*: il più comune è **l'effetto pratica**

Valutare i cambiamenti nel tempo

L'effetto pratica

È l'effetto per cui una seconda valutazione ha un punteggio migliore della precedente. Ci sono varie ragioni alla base dell'effetto pratica.

- Maggiore familiarità con la situazione testistica (meno ansia)
- Possibilità di utilizzare strategie scoperte nella prestazione precedente

L'utilizzo di *forme parallele* o la lunga distanza tra le due valutazioni non garantisce una protezione dall'effetto pratica (Mondini e Arcara, 2008)

Valutare i cambiamenti nel tempo

Come valutare i cambiamenti nel tempo

L'unico modo (rigoroso) è utilizzare **metodi statistici per valutare un cambiamento** (es. Chelune et al., 1993; Crawford & Garwhaite, 2006)

A mia conoscenza gli unici lavori in Italia che riportano tali dati non sono molti :
esempio

l'**ENB-2** (Mondini et al., 2011) e **APACS** (Arcara & Bambini, 2016), **NADL-F** (Arcara et al., 2019), **GEMS** (Mondini et al., 2023), **Tele-GEMS** (Montemurro et al. 2023).

Valutare i cambiamenti nel tempo

Come valutare i cambiamenti nel tempo

Supplementary Table 2.4. Thresholds for significant changes for the Figurative Language 1 task.

Figurative Language 1							
Upper Threshold (2 nd meas.)	12	12	13	14	14	-	-
Observed value (1st meas.)	9	10	11	12	13	14	15
Lower Threshold (2 nd meas.)	11	12	12	13	14	14	15

Valutare i cambiamenti nel tempo

Esame Neuropsicologico Breve 2

Raffaello Cortina Editore

di S. Mondini, D. Mapelli, A. Vestri, G. Arcara, P.S. Bisiacchi

PROTOCOLLO MATERIALI MODULO CONSENSO VALUTAZIONE TEST - RETEST

▶ GIOVANI (16-30 anni) ▶ ANZIANI (58-82 anni)

NOME DATA ETÀ

Elenco Test	Test	Retest	Esito
Digit span	5	4	STABILE
Test di Memoria di prosa - immediata			
Test di Memoria di prosa - differita			
Test di Memoria con interferenza- 10 sec.			
Test di Memoria con interferenza- 30 sec.			
Trail making test - A			
Trail making test - B			
Test di Fluenza fonemica			
Test delle Figure Aggrovigliate			
Test dell'Orologio			

Valutare i cambiamenti nel tempo

Metodi per valutare i cambiamenti nel tempo

Esistono numerosi metodi per valutare i cambiamenti significativi nel tempo (si veda Collie 2002). In tutti questi metodi l'obiettivo è costruire delle *soglie* (una superiore e una inferiore) superate le quali inferiamo che la variazione non è riconducibile al caso e, pertanto, è da considerare come significativa.

Per farlo quasi tutti i metodi si basano su questa metodologia: si prende un gruppo di individui e li si testa due volte (t_0 e t_1) in un lasso di tempo entro cui si assume che non sia avvenuto un cambiamento. Ad esempio soggetti sani testati a t_0 e a t_1 dopo un mese. Questi soggetti non devono avere fatto training o altro che potrebbe spiegare differenze tra i punteggi. Ogni differenza tra t_0 e t_1 sarà dunque riconducibile a imprecisioni del test. A partire da questo (tramite metodi statistici diversi) si costruisce un'intervallo con un limite superiore o inferiore di cambiamento significativo.

Valutare i cambiamenti nel tempo

Metodi per valutare i cambiamenti nel tempo

Il principio di questi metodi è spesso analogo a quanto già visto per identificazione di deficit/danno. Si costruisce una sorta di distribuzione che ci dice i valori attesi assumendo che non c'è cambiamento (H_0) e si va a vedere se il punteggio osservato (o la variazione di punteggio osservata) è così estrema da non essere riconducibile al caso. Se è così allora si inferisce cambiamento (peggioramento/miglioramento) significativo.

Valutare i cambiamenti nel tempo

Reliable Change Index (RCI)

L'RCI è il metodo più famoso per calcolo di differenze significative nel tempo.

$$RCI = \frac{X_{i1} - X_{i0}}{SE_{diff}}$$

$$SE_{diff} = \sqrt{2 SE^2}$$

$$SE = s_0 \sqrt{1 - r_{01}}$$

$$s_0 = sd(X_0) \quad r_{01} = cor(X_0, X_1)$$

Una volta ottenuto l'RCI può essere trattato come uno z-score.
Se > 1.64 , indica un miglioramento significativo (5%), mentre se < -1.64 , indica un peggioramento significativo (5%).

Valutare i cambiamenti nel tempo

Modified Reliable Change Index (RCI)

Il mRCI è una variazione del RCI, che tiene conto di effetto pratica.

$$mRCI = \frac{(X_{i1} - X_{i0}) - m}{SE_{diff}}$$

$$m = \bar{X}_1 - \bar{X}_0$$

$$SE_{diff} = \sqrt{2 SE^2}$$

$$SE = s_0 \sqrt{1 - r_{01}}$$

$$s_0 = sd(X_0) r_{01} = cor(X_0, X_1)$$

Una volta ottenuto il mRCI può essere trattato come uno z-score.

Se > 1.64 , indica un miglioramento significativo (5%), mentre se < -1.64 , indica un peggioramento significativo (5%).

Valutare i cambiamenti nel tempo

Regression Method Crawford & Garthwaithe, 2006

Anche se mRCI tiene conto di effetto pratica, non tiene conto di *regressione verso la media* (i.e., punteggi estremi a t_0 tenderanno a essere più vicini alla media a t_1).

Essi inoltre non tengono conto di differenze che ci potrebbero essere a seconda del punteggio alla baseline (a t_0), l'effetto pratica è corretto sempre nella stessa maniera.

Il metodo di Crawford & Garthwaite, tiene invece conto di tutti questi aspetti e inoltre fornisce una statistica che tiene conto della differenza tra campione e popolazione (e funziona anche con campioni piccoli $n \sim 5$).

RCI e MRCI invece non tengono conto di distinzione campione e popolazione (come z-score)

Approfondire la formula va al di là degli scopi di questo corso.



Valutare i cambiamenti nel tempo

Come valutare i cambiamenti nel tempo

Cosa fare se il test non ha valori di cambiamento significativo?

Fare **MOLTA** attenzione nell'utilizzo per inferire cambiamento
(evitarlo se possibile)

Vedere i valori di affidabilità test-retest e selezionare test in cui sono alti, ma ricordando che non è incluso l'effetto pratica.

Valutare i cambiamenti nel tempo

Da ricordare

È da ricordare che come per affidabilità test anche gli intervalli di cambiamento significativo si riferiscono allo specifico intervallo temporale con cui sono stati costruiti (es. 1 mese) ed esistono infinite distanze temporali (tra t_0 e t_1) per cui calcolare intervalli di cambiamento

In genere queste sono fatte a tempi relativamente corti (es. 1 mese) perché altrimenti diventa difficile assumere che non c'è stato un cambiamento.

In alternativa si usano gli stessi intervalli di rilevanza clinica. (es. 2 settimane per una operazione chirurgica, 6 mesi per Sclerosi Multipla).



Confrontare punteggi di test diversi



Confrontare punteggi di test diversi

In certi casi può essere utile confrontare test diversi.

Possono essere confrontati test diversi che misurano lo stesso costrutto (per approfondire se la prestazione deficitaria è legata ad un test specifici).

o test che misurano cose diverse (per inferire deficit selettivi)

Confrontare punteggi di test diversi

Per permettere il confronto tra test diversi bisogna che i punteggi siano trasformati in modo da essere confrontabili.

La strategia più diffusa è trasformare il punteggio sulla base della percentuale di campione o popolazione che ha ottenuto quel punteggio (o un punteggio inferiore). Ad esempio, il 20%, il 30%, il 40%, etc.

Questo è spesso fatto tramite:

- Conversione in **punti - z** (ma le distribuzioni dei test devono essere simili)
- Utilizzo dei **punteggi equivalenti** (Capitani e Laiacina, 2007)
- (esistono altri metodi, es. Crawford, Howell & Garthwaite, 1998)

Confrontare punteggi di test diversi

Limiti confronto di test diversi

È importante ricordarsi alcuni aspetti nell'utilizzo di strategie di confronto di test diversi tramite trasformazioni.

In particolare dobbiamo ricordarci che il confronto è spesso da considerarsi come descrittivo (e non legato ad una inferenza statistica)

Supponiamo di trasformare dei punteggi e di vedere che la persona si colloca al 70% nei test di memoria, mentre al 20% per i test di funzioni esecutive.

Si potrebbe interpretare che c'è un declino selettivo delle funzioni esecutive.

Questo è possibile (e ragionevole), ma è importante ricordare che non sappiamo quanto *improbabile* è osservare questo pattern nei sani. Molte delle inferenze che abbiamo fatto erano basate su improbabilità, ma in questo caso noi non sappiamo se è un pattern effettivamente molto improbabile da suggerire un danno selettivo (lo stiamo *assumendo implicitamente*).