

Metodi statistici per la Neuropsicologia Forense

A.A. 2023/2024

Giorgio Arcara

IRCCS San Camillo, Venezia
Università degli Studi di Padova





Valutare i deficit cognitivi

- 1. il concetto di soglie per percentuali o probabilità**
- 2. Punti z e probabilità**
- 3. Percentili**
- 4. Campione vs Popolazione**
- 5. test per il caso singolo (t di Crawford)**
- 6. Metodi di regressione per correzione punteggi**
- 7. Metodo dei Punteggi equivalenti**

Il cut-off clinico e valori soglia

Il concetto di cut-off clinico (o in generale di valori soglia o valori che delimitano certe percentuali) sono fondamentali nella statistica applicata alla neuropsicologia clinica e forense.

L'idea di fondo, nel loro utilizzo, è trovare un valore che mi aiuti a classificare la prestazione del paziente o, più in generale a determinare quanto è probabile/improbabile.

L'improbabilità è strettamente legata al concetto di *deficit* o presumibile *danno* cognitivo, che di fatto viene operazionalizzato come una prestazione insolita, improbabile, assumendo che si sia normali.



Valutare i deficit cognitivi

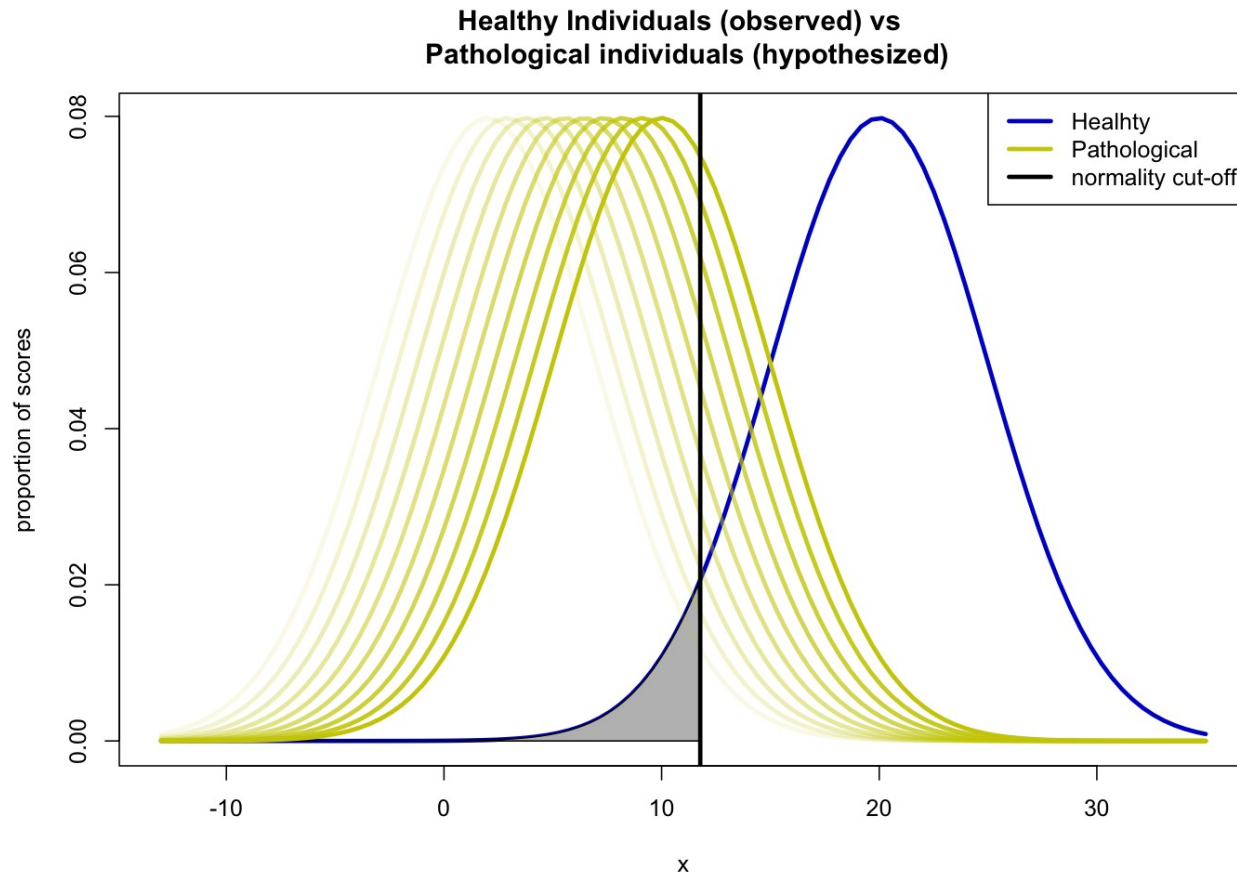
Il cut-off clinico

L'utilizzo delle probabilità che viene fatto in ambito clinico è però insidiosa perché è facile sbagliare nel loro utilizzo e nella sua interpretazione.

Formalizzare in maniera più rigorosa aiuta ad evitare gli effetti di una terminologia vaga.

Valutare i deficit cognitivi

Cut-off clinico



Non sapendo dove sono popolazioni patologiche (deficit o danno), mettiamo una soglia arbitrariamente che delimita 5% più basso (area grigia). Assumendo che deficit e patologia siano da qualche parte a sinistra.

Valutare i deficit cognitivi

I valori utili per la clinica

Definiamo in generale valori utili alla classificazione clinica quei punteggi x al di sotto del quale che c'è una probabilità nota k che un punteggio osservato o abbia un valore uguale o inferiore a x (nei test in cui un punteggio basso indica una performance peggiore).

$$x : P(o \leq x) = k$$

es. con il “cut-off di normalità”, $k = 0.05$ e il valore x è il valore è quello tale che $p(O \leq x) = 0.05$ (c'è il 5% di probabilità di osservare una prestazione inferiore a x , essendo basso, inferiamo un danno).
La soglia del 5% è completamente arbitraria.

Percentili

Questa definizione è sovrapponibile a quella che si usa in statistica di “percentili”, in riferimento a distribuzione

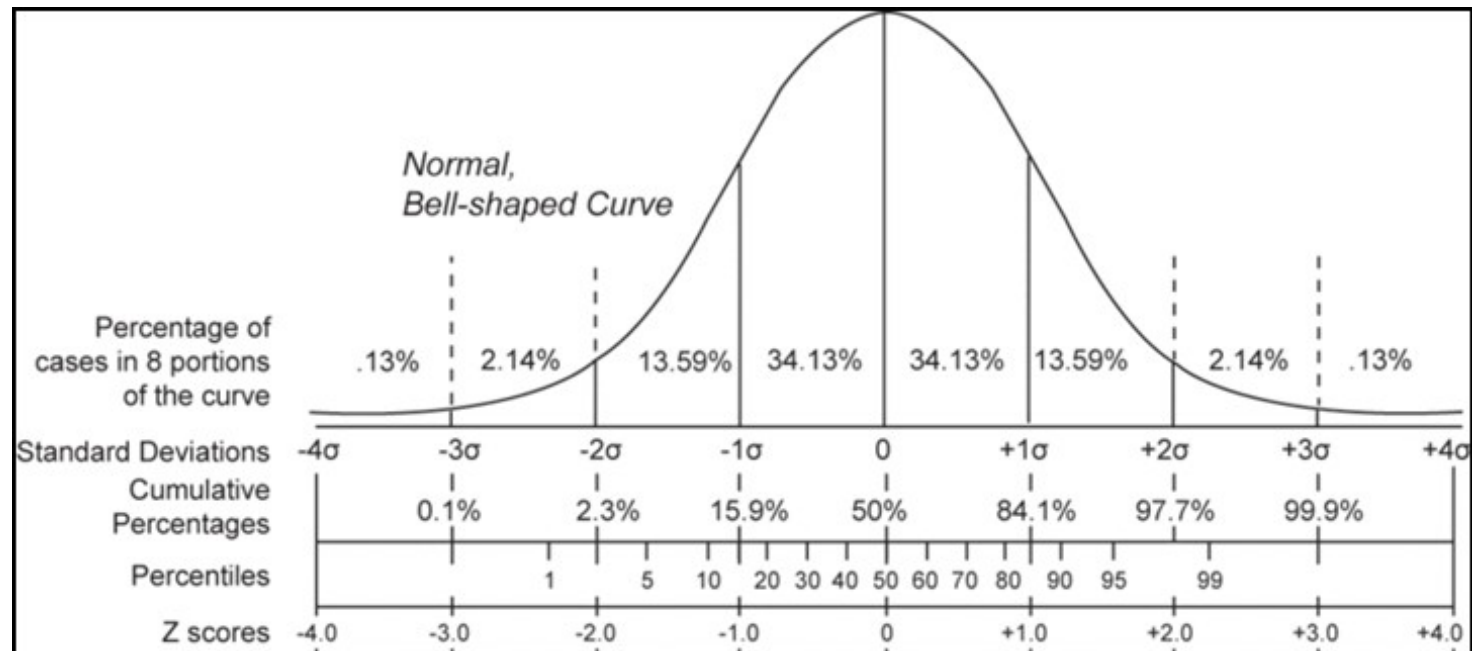
L'utilizzo della curva normale o di percentili non parametrici rappresentano diversi metodi per identificare queste percentuali.

L'uso di punti z non è altro che un metodo che semplifica l'identificazione di questi percentili

Valutare i deficit cognitivi

Percentili

Ricordiamo che l'utilizzo della curva normale (e **della trasformazione in punti z**) ha senso in quanto ci aiuta ad identificare *percentili* e cioè valori che delimitano porzioni note della distribuzione.



Trasformazione in punti z

Quando si trasformano dei punteggi in punti z ci sono alcuni punti che risultano di importanza

Se i punteggi sono standardizzati (punti z con Media = 0 e SD = 1).

- 1.64 delimita il 5% dei punteggi più bassi
- 1.96 delimita il 2.5% dei punteggi più bassi

Se i punteggi non sono standardizzati.

Media - 1.64 * SD delimita il 5% dei punteggi più bassi
Media - 1.96 * SD delimita il 2.5% dei punteggi più bassi

A volte si trova “2 deviazioni standard”. Non è altro che un altro numero (sempre arbitrario) per delimitare performance bassa.



Valutare i deficit cognitivi

Percentili

E nel caso di distribuzione non normali?

Si possono utilizzare percentili non parametrici

(per esempi vedi codice R associato per potenziale impatto di non normalità, specie con un marcato effetto soffitto.)

La distribuzione di riferimento

Questa definizione può essere però approfondita concentrandosi su un punto critico su cui in precedenza abbiamo glissato e cioè che I valori che è relativa alla distribuzione di riferimento.

Possiamo quindi specificare meglio la formula così

$$x: P(o < x | N) = k$$

X è il valore per cui ci si attende un probabilità k che un punteggio O sia inferiore a x, *Assumendo che si provanga da una popolazione di persone normali nel senso di "sane" (cioè senza deficit cognitivo).*

Infatti se la persona venisse da altre popolazioni (es. Pazienti con Alzheimer) I valori attesi potrebbero essere diversi.

Analogia con ipotesi nulla in test statistico

Da notare che questa formulazione è del tutto analoga a quella del p-value in realazione dell'ipotesi nulla e all'identificazione dei valori critici della statistica test (es. Il t-value critico oltre il quale considero il risultato significativo).

$$X: P(t_{obs} < t_{crit} | H_0) = q^1$$

X è il valore per cui ci si attende un probabilità q che la statistica osservata (tobs in questo caso) sia inferiore al t critico, assumendo che H0 sia vera.
H0 dipende dal tipo di test, ma nel caso di un confronto fra campioni indica che I due campioni che sono stati confrontati con il t-test provengono dalla stessa popolazione.

¹Nota che questa formula è semplificata, e non sono considerati I gradi di libertà

Esperimento mentale

Esperimento.

Supponiamo di avere dei dati normativi e di sapere che la soglia che delimita il punto $Z -1.64$ (5%) è 4, oppure che il percentile 5% è 4.

Supponiamo di osservare una prestazione di 3.

La consideremmo come deficitaria?

Se vi dicessi che ho calcolato questa soglia da un campione con 20 soggetti
O con 200 soggetti. Ci fideremmo alla stessa maniera della soglia?

Considerare una soglia per inferenze di tipo clinico non può prescindere dal fare riflessioni sulla rappresentatività del mio campione rispetto alla popolazione e su quanto i risultati sono generalizzabili.



Campione o popolazione

Un aspetto cruciale è che sebbene percentili siano calcolabili per campione o popolazione, nel contesto della valutazione neuropsicologica l'interesse è sempre (implicitamente) alla popolazione

Non mi interessa classificare il mio paziente rispetto al mio specifico campione (es. Ai miei 10, 100, 1000 soggetti sani), ma come il mio paziente si classificherebbe *in generale* rispetto alla popolazione di sani.

Un altro modo di vedere il problema è che la mia classificazione non dovrebbe dipendere dallo specifico campione normativo, ma dovrebbe valere per *tutti* i possibili campioni normativi

Questo vale soprattutto se il mio campione normativo ha dimensioni limitate (ed è quindi poco rappresentativo della popolazione).



Percentile o probabilità?

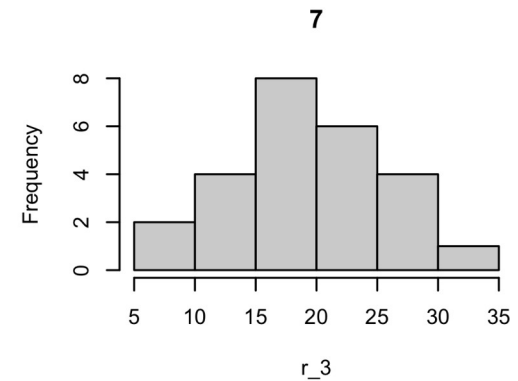
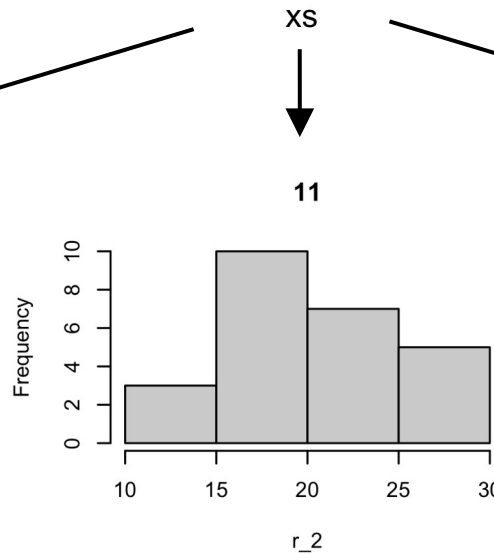
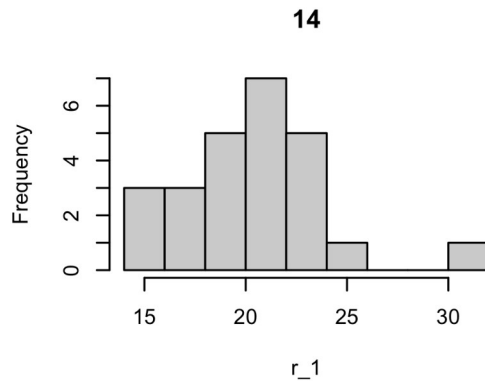
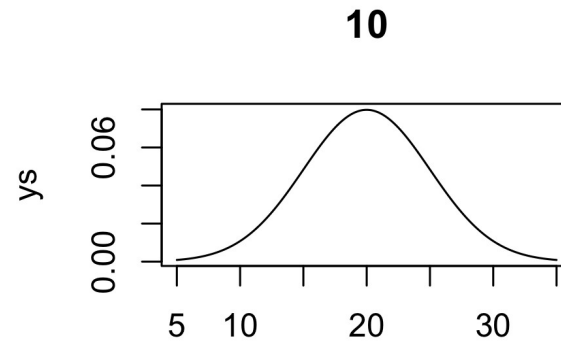
Notare che il percentile non mi dice nulla di probabilità, ma mi dice il valore che delimita una certa percentuale di osservazioni del campione.

Avere una ercentuale non significa necessariamente che sto controllando la probabilità di errore, usare percentili o z-scores si approssima a probabilità (in termini di popolazione) solo quando I campioni sono grandi e sono rispettati assunti (Es. Normalità distribuzione della popolazione)

Valutare i deficit cognitivi

Campione o popolazione

La figura mostra tre ipotetici campioni ($n=25$) estratti dalla stessa popolazione. Il numero sopra indica il cut-off vero (5%) e quello ottenuto nei campioni



Valutare i deficit cognitivi

Campione o popolazione

Usare una soglia calcolata sul 5% del campione (es. tramite punti z) ha dunque dei limiti: non mi garantisce che delimiti *nella popolazione* la soglia del 5

Ma possiamo stimare un valore della popolazione a partire da ciò che abbiamo? (e cioè del campione)

Sì, è questa stima (associata ad errori) è alla base dei concetti di **statistica inferenziale**. La statistica inferenziale è (proprio per definizione) quella branca della statistica che ci permette di fare inferenze sulla popolazione a partire dai dati del campione, associando a certe percentuali di errori.



Confronti con valori soglia e statistica inferenziale

Confrontare un valore con un punto z ottenuto da un campione, non è strettamente parlando un *test statistico*.

Ricordate elementi di statistica di base e la differenza tra *statistica descrittiva* e *statistica inferenziale* (vedi [link](#))

Confrontare con un cut-off (ottenuto con punti z o percentili) e dire che c'è un danno, non è diverso da avere la media, confrontare e dire “è più basso/più alto della media”, ma non è un test-statistico.

Il paragone si può fare, ma solo a livello descrittivo e non inferenziale. Un confronto come quello sopra non ci dice che i valori sono *significativamente diversi*.



Confronti con valori soglia e statistica inferenziale

Una statistica inferenziale è caratterizzata dal poter fare dei precisi enunciati sulla probabilità di errore *riferiti* alla popolazione da cui il campione da cui è estratta e non al campione stesso.

In senso generale, non utilizzare un test statistico per fare un inferenza, può essere visto come una strategia qualitativa e meno robusta perchè non è nota la probabilità di commettere errori e la generalizzabilità dei risultati al di là dal campione usato per il confronto.

Più concretamente, Se calcolo il cut-off usando la soglia -1.64 del punteggio standardizzato:

- 1) non so quanto la mia classificazione si generalizzerebbe ad altri campioni.
- 2) non so quale è la probabilità di errore della mia classificazione (assumendo che la persona sia sana), immaginando infinite ripetizioni.

Valutare i deficit cognitivi

Confronti con valori soglia e statistica inferenziale

Nota statistica.

Usare il confronto con punti z come test statistico è noto come **test-z** e può essere un test statistico, ma assume

Di stare confrontando una media campionaria con una popolazione di riferimento

Di conoscere la *varianza della popolazione* (non del campione).

Con varianza non nota si dovrebbe usare il t-test (con la distribuzione t che si approssima a normale per campioni grandi)

Valutare i deficit cognitivi

Confronti con valori soglia e statistica inferenziale

Usare il confronto con punti z come test statistico implica dunque due errori:

1) sto trattando il mio singolo individuo come una media campionaria

2) sto trattando i dati normativi (che sono un campione) come se fossero una popolazione.

Al di là di correttezza formale e teorica, questo implica possibili differenze nella nostra possibilità di errore di classificare, che in assenza di statistica appropriata, non è nota.



Valutare i deficit cognitivi

Nota confortante

In realtà è da riconoscere che con campioni sufficientemente ampi (> 100) e se la distribuzione è realmente normale e con deviazione standard non troppo elevate, gli errori non sono drammatici

(Questa è una considerazione da prendere *cum granu salis*, a partire da simulazioni dello script 10 associato a queste slides)

t-test di Crawford & Howell

Il t-test di Crawford & Howell (1998)* è una modificazione del t-test che permette di confrontare un *caso singolo* rispetto con una media di un campione.

Il t-test di Crawford è una modificazione del t-test che permette di *confrontare un caso singolo rispetto ad un campione, ma con inferenza rispetto a una popolazione, con errore noto*.

Il test funziona anche con campioni piccoli (~10 soggetti).

* in realtà il test è di Sokhal e Rohlf (1995). Crawford & Howell hanno avuto il merito di diffonderlo per i single-case in contesto di neuropsicologia.

t-test di crawford

$$t = \frac{X_1 - \bar{X}_2}{s_2 \sqrt{\frac{N_2 + 1}{N_2}}}$$

Confrontando con distribuzione con $N_2 - 1$ gradi di libertà il risultato da
La probabilità (data ipotesi nulla vera) di ottenere valori di t uguali o inferiori

(vedi codice di esempio)

Valutare i deficit cognitivi

Confronto con dati normativi

Il test di Crawford, così come utilizzo di z-score e utilizzo di percentili, necessita però identificare un campione di riferimento

Abbiamo detto che i dati dovrebbero essere di soggetti *sani* (quindi normali) simili alla persona testa. Ma come definire questo?

Spesso per un confronto normativo, ci si riferisce a persone con stesso sesso e con età, scolarità simili.

La scelta di queste variabili è pragmatica: sono variabili che sappiamo influenzare performance cognitiva e sono facili da raccogliere (al momento di creazione del test).

Note su t-test di Crawford (1/2)

Il t-test di Crawford si comporta in maniera più appropriata dello z-test o di percentili non parametrici (vedi anche script con simulazioni)

Il t-test di Crawford è anche più robusto in caso di violazione di normalità. Questo significa che anche se una delle assunzioni è violata, tutto sommato il comportamento del test è adeguato.

Il test di Crawford funziona bene (a livello di Errore di 1° tipo) anche per campioni piccoli suggerendo un potenziale utilizzo anche per campioni raccolti ad-hoc (ma attenzione ad Affidabilità e Validità, dovrebbero essere disponibili).

Note su t-test di Crawford (1/2)

Il t-test di Crawford mantiene un errore di primo tipo costante al 5% anche per campioni molto piccoli ($n=5$).

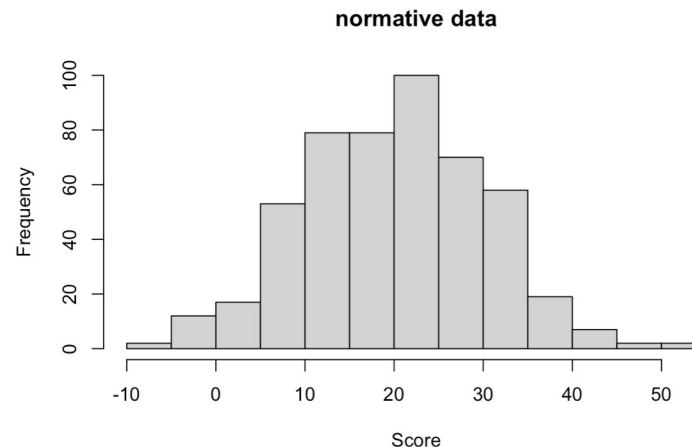
Questo non vuol dire che bastano campioni piccoli, in realtà è sempre meglio avere campioni numerosi, perchè porteranno cut-off più reali a quelli della popolazione

(Si veda simulazione in script 11_C&H_test)

La cosa importante è che il t-test di Crawford anche con campioni piccoli mantiene un errore di 1° tipo a livello del 5% (concretamente, per dare un esempio se una persona viene dalla popolazione dei sani, otterremo un valore sotto “cut-off” il 5% delle volte, proprio come è atteso. Nel caso del z-score la percentuale potrebbe essere 10% o più).

Confronto con dati normativi

Finora abbiamo fatto riferimento a formule che confrontano con un singolo gruppo di dati, denominato “campione normativo” o “dati normativi” che ci permette di ottenere una distribuzione di dati.



Questo gruppo sarà fatto di individui con una certa età, una certa scolarità, etc. E abbiamo sempre sottolineato che i dati normativi dovrebbero essere simili all'individuo da valutare.

Ma come facciamo a definire se il campione è adeguatamente simile?

Confronto con dati normativi

Come facciamo a scegliere un campione di riferimento come simile?

Es. se il mio paziente ha 68 anni e 5 anni di scolarità (come GH)
Quale è un gruppo adeguato?

Esempi

- persone da 65 a 75 anni con scolarità tra 3 e 5 anni.
- persone con esattamente 68 anni e esattamente 5 anni di scolarità.

In genere i test, riportano tabelle predefinite e, a livello di raccolta di dati è possibile che alcune combinazioni siano poco rappresentate. Al di là di questo, spesso non è possibile avere accesso agli interi dati e “comporre” il dato normativo di riferimento, ma occorre basarsi sulle tabelle disponibili.

Appropriatezza campione di riferimento

È un aspetto metodologico importante e con impatto. Se nel mio campione di riferimento sto usando persone che non sono *simili* alla persona di cui sto osservando la performance le mie inferenze (su deficit o su danno) sarebbero distorte.

Es. supponiamo che il mio test fornisce un cut-off che si riferisce all'intervallo 70-80 anni di età.

Supponiamo debba testare un paziente con 70 anni. Vuol dire che nella distribuzione di dati ci saranno persone di 80 anni di età (la cui performance potrebbe essere molto più bassa).

La stratificazione arbitraria in campioni per le variabili rilevanti (età, scolarità, sesso), ha sempre problemi sulla composizione del campione normativo e la sua rappresentatività.

Metodi di regressione

In alternativa ai metodi che usano un semplice campione di riferimento, esistono metodi che invece usano metodi basati su regressione.

Il razionale di questi metodi è il seguente:

- 1) A partire da alcune variabili posso prevedere la performance attesa del partecipante date le variabili osservate, *secondo un modello di regressione lineare*
- 2) Una volta che ho questo modello di regressione posso stabilire se la performance osservata è sufficientemente improbabile dati i valori osservati.

Valutare i deficit cognitivi

Metodi di regressione

Una potenziale formalizzazione dei cut-off ottenuti tramite i metodi di regressione è la seguente. Il cut-off x è il punteggio tale che il paziente i -esimo (con specifici valori per le variabili aggiuntive V_1, V_2, \dots, V_n), tale che la probabilità di osservare quella performance è k (con k critico normalmente 0.05)

$$x: P(o_i < x | V_1, V_2, \dots, V_n, N) = k$$

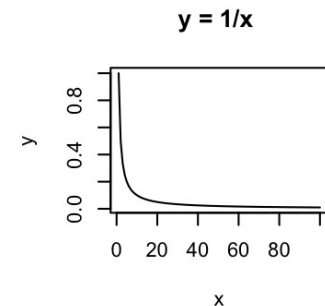
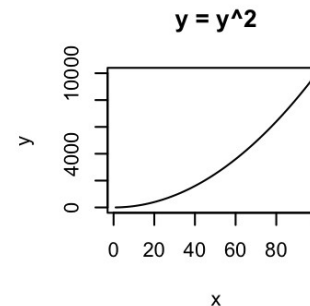
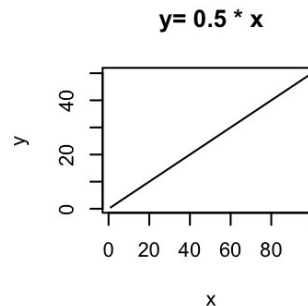
Classicamente i modelli di regressione includono età, scolarità e sesso come variabili predittore della performance

Metodi di regressione Variabili da considerare

Quali variabili da considerare nella regressione?

Esistono molti modi per selezionare le variabili da considerare e quali includere. Approfondire questo è oltre gli scopi di questo corso.

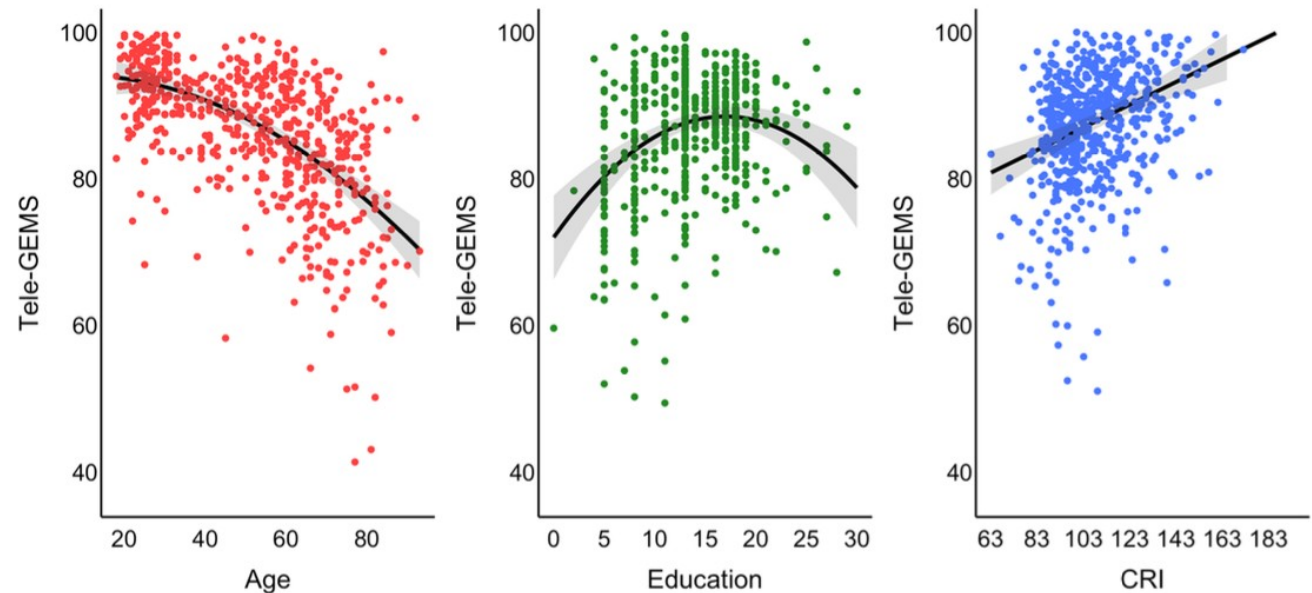
I metodi di regressione più raffinati includono anche metodi per tenere conto di eventuali “non-linearità” (relazioni descrivibili tra funzioni non lineari, es. Esponenziale, iperbole, e la variabile osservata).



Valutare i deficit cognitivi

Metodi di regressione Variabili da considerare

Fig. 2 Effect of age, education, and the Cognitive Reserve Index (CRI) on Tele-GEMS total score. Age, education, and CRI are reported on the x -axis, while Tele-GEMS total score is reported on the y -axis. Quadratic terms of age and education are included in this figure, according with the final regression model, which best fitted Tele-GEMS data



Esempio da Montemurro et al., 2023 (Tele-GEMS)

Metodi di regressione

Nelle slides che seguono viene mostrato un popolare metodo basato su regressioni
Che è alla base della maggior parte dei test neuropsicologici sviluppati in Italiano,
Il metodo dei **Punteggi Equivalenti**.

Questo metodo contiene due parti principali:

- una parte per ottenere punteggi “corretti” tramite regressioni.
- una parte per ottenere delle soglie che tengano conto della differenza campione/popolazione

I Punteggi Equivalenti

Introdotti da Erminio Capitani nella raccolta di taratura di test di Spinnler e Tognoni (1987)

Tre aspetti chiave:

- Correggono per età e scolarità tramite regressioni.
- Tengono conto di differenza campione/popolazione
- Permettono il confronto tra diversi test



Valutare i deficit cognitivi

Il calcolo dei Punteggi equivalenti è suddiviso in vari step

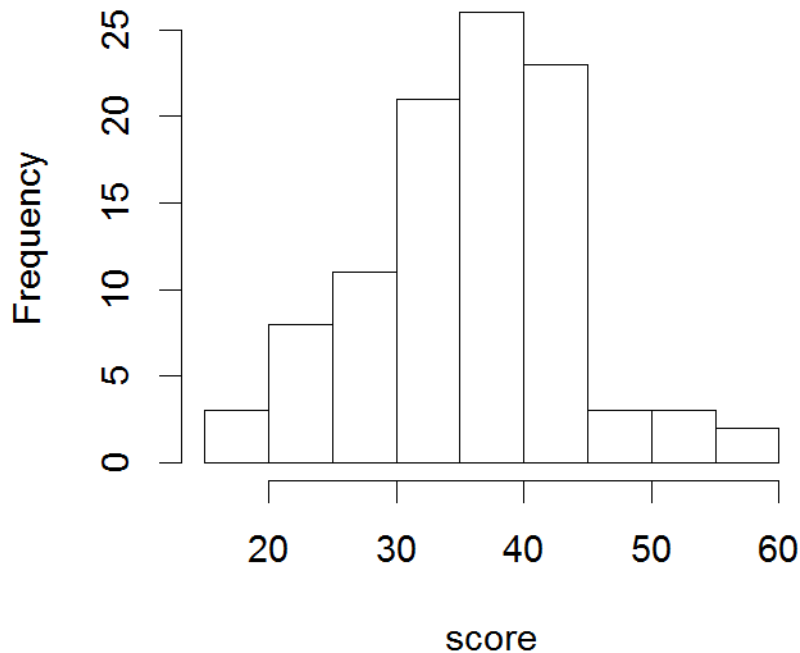
- 1) Calcolo di punteggi aggiustati, a cui è rimosso l'effetto di età e scolarità.
- 2) Calcolo limite di tolleranza ($PE = 0$).
- 3) Calcolo dei rimanenti PE

Valutare i deficit cognitivi

I Punteggi Equivalenti

- 1) Calcolo di punteggi aggiustati, a cui è rimosso l'effetto di età e scolarità (e sesso).

**Raw Scores distribution
in normative data**

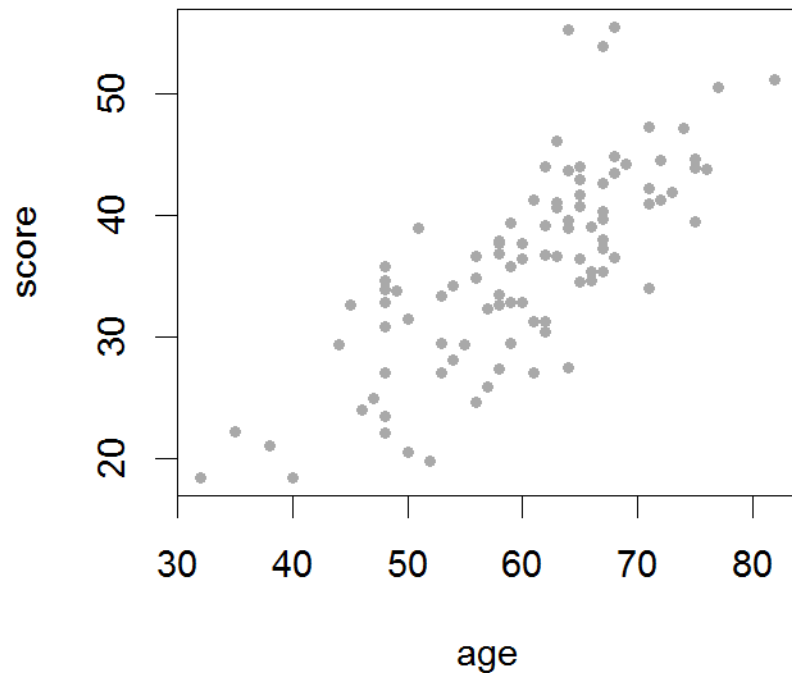


La distribuzione dei punteggi grezzi include persone di diversa età.

Valutare i deficit cognitivi

I Punteggi Equivalenti

- 1) Calcolo di punteggi aggiustati, a cui è rimosso l'effetto di età e scolarità.

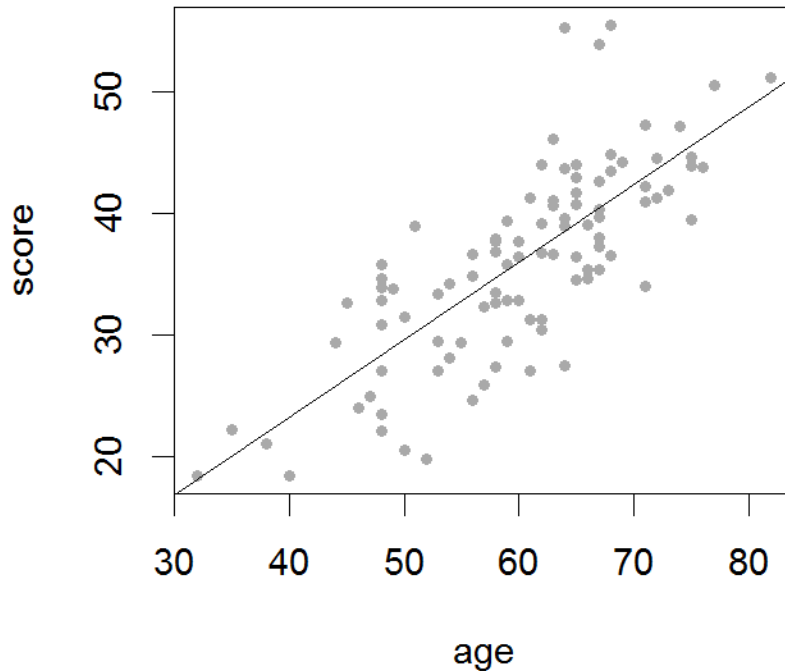


Effetto dell'età

Valutare i deficit cognitivi

I Punteggi Equivalenti

- 1) Calcolo di punteggi aggiustati, a cui è rimosso l'effetto di età e scolarità.

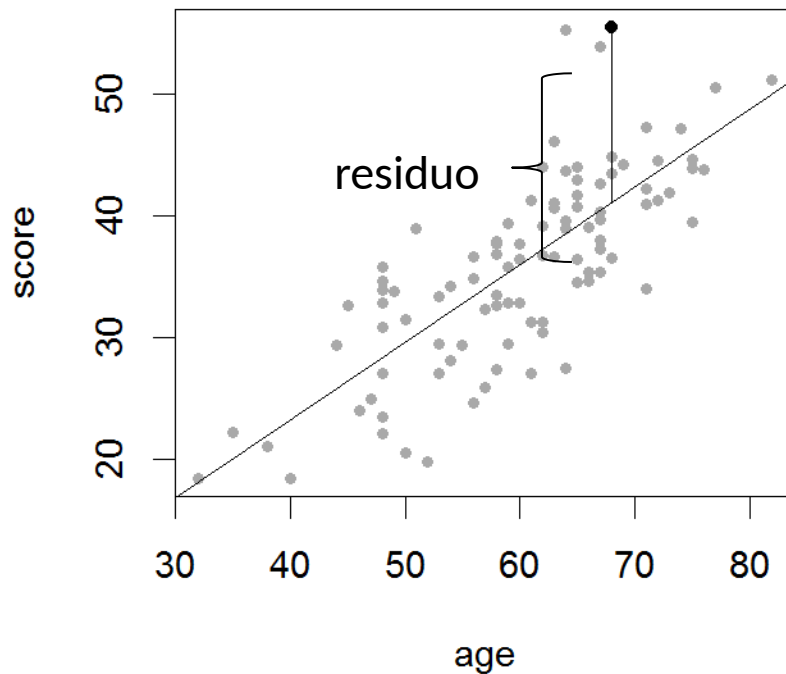


Effetto dell'età
e retta di regressione

Valutare i deficit cognitivi

I Punteggi Equivalenti

- 1) Calcolo di punteggi aggiustati, a cui è rimosso l'effetto di età e scolarità.

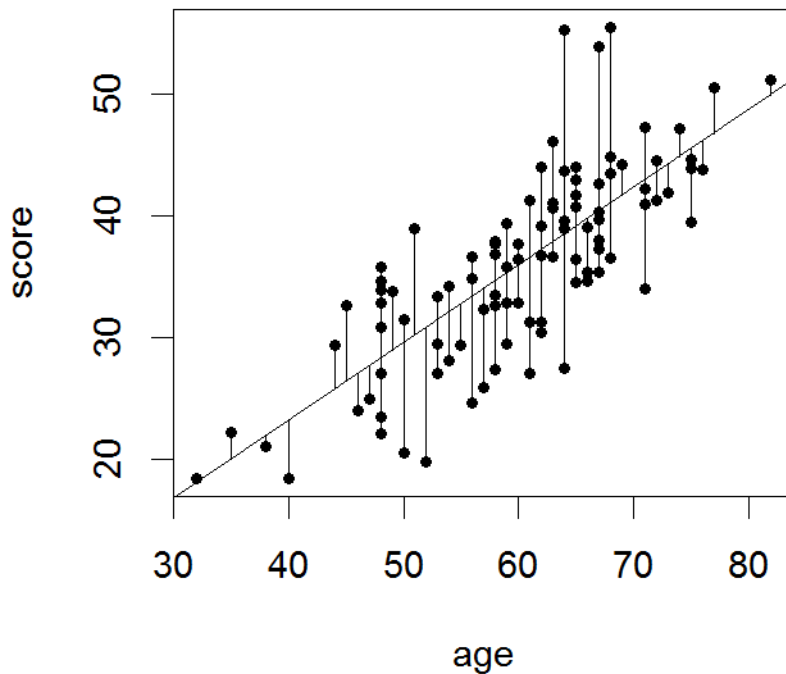


Differenza tra punteggio previsto e punteggio osservato.

Valutare i deficit cognitivi

I Punteggi Equivalenti

- 1) Calcolo di punteggi aggiustati, a cui è rimosso l'effetto di età e scolarità.

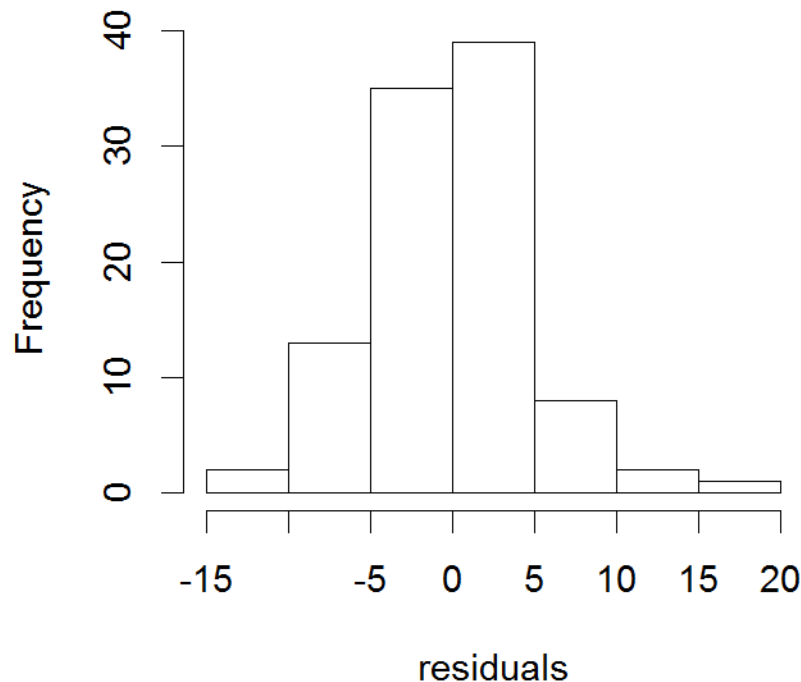


Calcolo residui di tutti i soggetti dei dati normativi

I Punteggi Equivalenti

- 1) Calcolo di punteggi aggiustati, a cui è rimosso l'effetto di età e scolarità.

**Residuals distribution
in normative data**



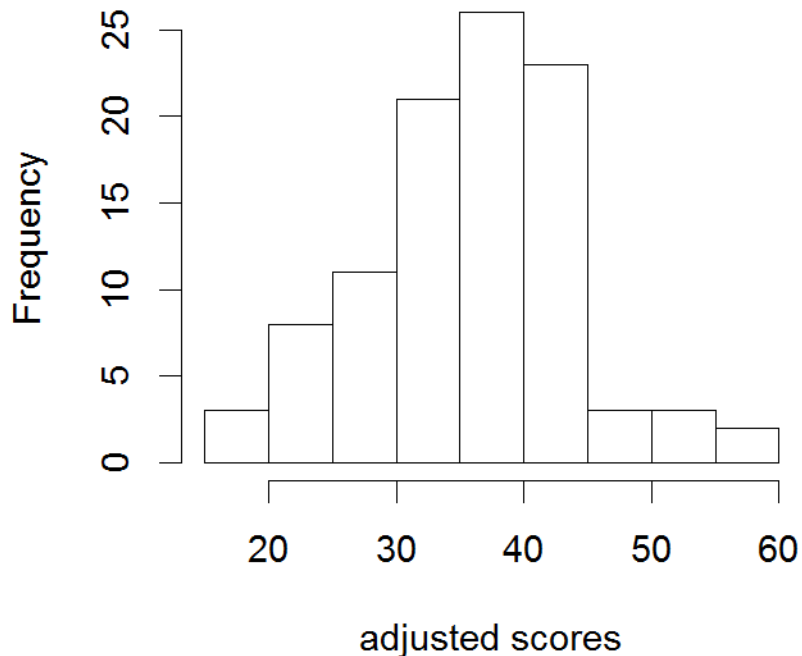
I residui potrebbero essere considerati già punteggi «aggiustati» perché sono ciò che rimane del punteggio tolta l'influenza dei predittori (in questo caso ripulito dell'effetto di età)

Valutare i deficit cognitivi

I Punteggi Equivalenti

- 1) Calcolo di punteggi aggiustati, a cui è rimosso l'effetto di età e scolarità.

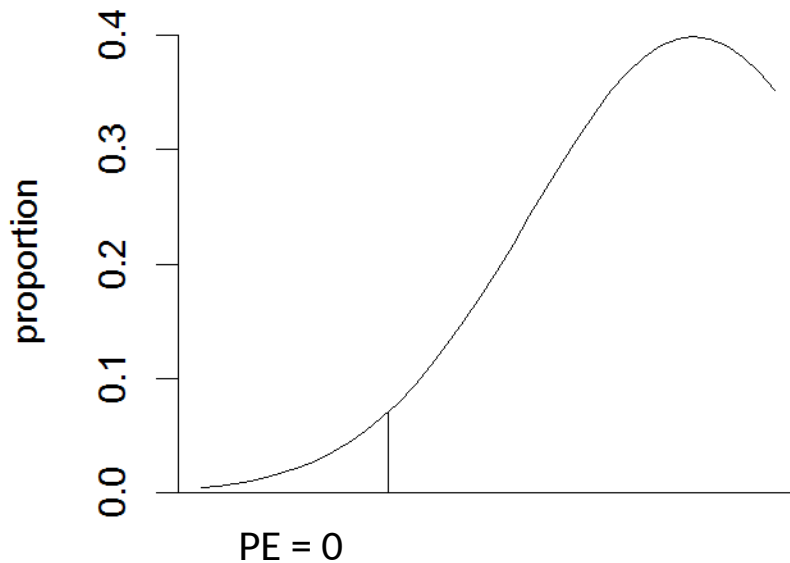
Adjusted Scores distribution in normative data



Nel metodo dei PE, i punteggi aggiustati sono riportati in una scala simile a quella originale.
(tecnicamente, aggiungendo l'intercetta al residuo)

I Punteggi Equivalenti

2) Calcolo limite di tolleranza (PE = 0)



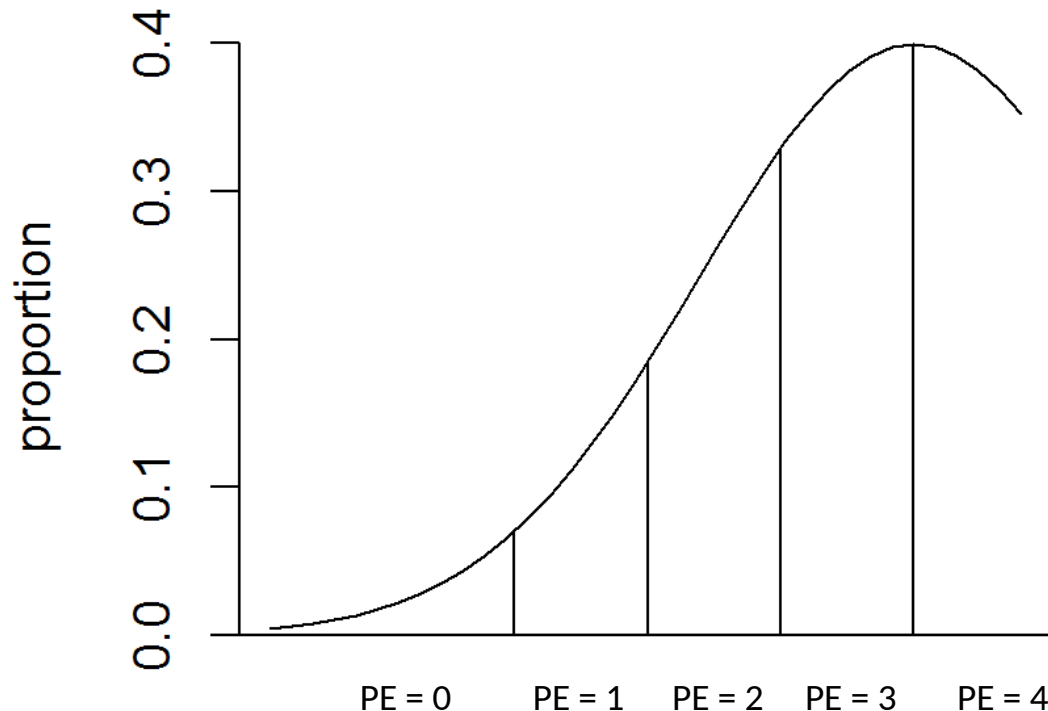
Si calcola il PE = 0 come:

**Limite di tolleranza interno
unidirezionale non parametrico
(5%)**

Indica il punteggio che non verrà ottenuto da più del 5% della popolazione. (quindi il cut-off, a livello di popolazione)

I Punteggi Equivalenti

2) Calcolo degli altri PE



In riferimento
all'ipotetica
distribuzione
dell'abilità sottostante,

Si trova anche il PE=4 e
poi si divide lo spazio
tra PE = 0 e PE = 4 in
intervalli equispaziati.
(equispaziati rispetto
all'abilità sottostante)

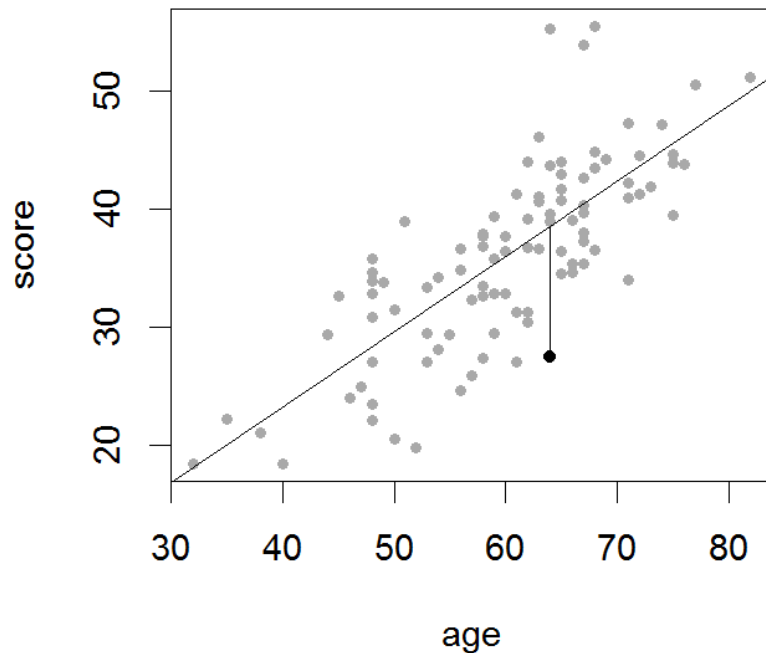
Valutare i deficit cognitivi

2) Quando valutiamo il nostro soggetto con un test con PE

- 1) Calcoliamo il **punteggio aggiustato**
Tramite confronto con tabella
o applicazione formula regressione
- 2) Confronto con tabella di
corrispondenza punteggio
aggiustato - Punteggio equivalente

Non si correggono risultati che ottengono
il punteggio massimo o minimo.

**Se punteggio corretto è $<$ soglia PE = 0,
allora prestazione deficitaria.**





Valutare i deficit cognitivi

Esempio di Punteggi Equivalenti

Proviamo insieme a fare un ipotetico utilizzo di punteggi equivalenti.

(vedere da articolo)

Considerazioni su metodi di regressione

I metodi di regressioni possono sembrare una soluzione perfetta al problema sollevato per i test più semplici di confronto con dati normativi (es. T-test di Crawford). In particolare **evitano gradienti arbitrari** e permettono una stima più corretta del paziente.

Essi non sono però privi di problemi. In particolare una stima non corretta del modello di regressione che porta ai punteggi finali può portare a stime distorte.

In generale ci aspettiamo effetti di distorsione maggiormente per valori estremi nelle variabili predittori, esempio scolarità o età molto alte o molto basse (sono i punti in cui eventuali imprecisioni del modello darebbero effetti più accentuati).

I modelli di regressione possono inoltre dare l'illusione di avere delle buone predizioni per alcuni punti (nello spazio dei valori possibili) per cui in realtà potrebbero esserci poche osservazioni e stime imprecise.

Limiti di metodi di regressione (2/2)

I modelli di regressione possono inoltre dare l'illusione di avere delle buone predizioni per alcuni punti (nello spazio dei valori) per cui in realtà potrebbero esserci poche osservazioni e stime imprecise.

Ad esempio, supponiamo di avere dati normativi di 1000 soggetti con un metodo di regressione. Supponiamo però di dovere testare un soggetto particolare (es. Donna con 80 anni e 23 anni di scolarità).

È possibile che in realtà nel mio campione ci siano pochissimi (o addirittura nessuna) persona che rispecchi questa combinazione di valori.

Sarebbe più opportuno avere un piccolo campione (es. anche 15 persone) che però siano veramente rappresentative rispetto all'individuo che voglio valutare e un test appropriato (es. T-test di Crawford).

Limiti di metodi di regressione (2/2)

I modelli di regressione possono inoltre dare l'illusione di avere delle buone predizioni per alcuni punti (nello spazio dei valori) per cui in realtà potrebbero esserci poche osservazioni e stime imprecise.

Ad esempio, supponiamo di avere dati normativi di 1000 soggetti con un metodo di regressione. Supponiamo però di dovere testare un soggetto particolare (es. Donna con 80 anni e 23 anni di scolarità).

È possibile che in realtà nel mio campione ci siano pochissimi (o addirittura nessuna, in tal caso si parla di *estrapolazione*) persona che rispecchi questa combinazione di valori.

Sarebbe più opportuno avere un piccolo campione (es. anche 15 persone) che però siano veramente rappresentative rispetto all'individuo che voglio valutare e un test appropriato (es. t-test di Crawford).



Considerazione su errore di primo e secondo tipo

In generale abbiamo visto come usare metodi che tengono conto di differenza tra campione e popolazione sia meglio

Il metodo t di Crawford (e metodi di regressioni di Crawford) mostrano anche che sono in grado di mantenere l'errore di 1° tipo (cioè avere un soggetto sotto cut-off, quando è sano), a livello atteso (cioè 0.05 o 5%) quando ci sono un numero elevato (in realtà infinito) di ripetizioni.

Esiste anche un altro tipo di errore rilevante che è quello di 2° tipo, cioè non riuscire ad identificare correttamente un soggetto patologico. Per i test utilizzati per il deficit/danno, questo è difficile da calcolare, perchè noi *non sappiamo quale è la performance attesa per chi ha deficit/danno*. Per calcolare quindi errore di primo tipo, si studiano in genere diversi possibili scenari.



Considerazione su errore di primo e secondo tipo

Esiste anche un altro tipo di errore rilevante che è quello di 2° tipo, cioè non riuscire ad identificare correttamente un soggetto patologico. Per i test utilizzati per il deficit/danno, questo è difficile da calcolare, perchè noi *non sappiamo quale è la performance attesa per chi ha deficit/danno (vedi figura in slides iniziali)*. Per calcolare quindi errore di primo tipo, si studiano in genere diversi possibili scenari, ma questi sono sempre comunque ipotetici.

Errore di primo e secondo tipo, inoltre si applicano non ad uno specifico campione, ma immaginando *infiniti campioni*. Si veda simulazione per vedere quanto è più variabile la stima del cut-off con campioni piccoli.

Viste questi aspetti sarebbe sempre desiderabile (laddove possibile) avere campioni ampi.

Riassunto finale (1/2)

- Siamo partiti dal problema di definire deficit e danno.
 - il metodo più diffuso è basato sull'*improbabilità di osservare quella prestazione se si è sani* (ragionamento contorto, ma necessario, vedi slides). Spesso c'è necessità di una soglia, ed essa è arbitrariamente messa al 5% delle prestazioni più basse. Da notare che entrambe le scelte ricordano il rifiuto dell'ipotesi nulla che c'è per i test statistici.
 - A questo punto è comparso il problema di cosa si intende per 5%. Non avrebbe senso riferirsi alla distribuzione del campione, ma sarebbe meglio riferirsi alla popolazione.
 - In generale esistono vari metodi per calcolare soglie e probabilità, e sarebbe meglio:
 - a) usare quelli che tengono effettivamente conto di distinzione campione/popolazione
 - b) avere campioni numerosi (almeno $n > 100$)
- Anche usando metodi non ottimali, avere campioni ampi migliora.

Riassunto finale (2/2)

- I metodi che si basano su campioni hanno però problemi dei “gradini” che si vengono a creare per ogni gruppo in cui abbiamo trasformato una variabile continua (es. Età) in gruppi separati.
- I metodi che si basano su regressione risolvono il problema dei “gradini”, ma attenzione a rappresentatività nei metodi con regressione e attenzione a valori estremi nei predittori (es. Scolarità, età). Sono i valori in cui la stima potrebbe essere più imprecisa.

GH è un uomo di 68 anni, vedovo e citato in giudizio per aggressione. La sua storia personale non è degna di note. Ha abbandonato le scuole elementari per lavorare. In passato ha avuto qualche piccolo problema con la legge per delle risse in gioventù. All'età di 25 anni ha avuto un incidente a lavoro: cadendo ha battuto la testa. Non è mai stato seguito per questo, ma poco dopo si è separato dalla moglie.

Adesso GH è in giudizio per l'aggressione del genero. Prima dell'evento rilevante per il processo, la figlia di GH e il genero erano da mesi in una situazione relazionale molto tesa e prossima alla separazione. Dopo un incontro chiarificatore tra la figlia di GH e il genero (che non si sentivano da alcuni giorni), la figlia di GH scrive due messaggi a GH e le dice "papà vieni a prendermi, sono in piazza. Questa volta mi ha dato veramente un mazzata".

Dopo questo messaggio. GH non va da dalla figlia, ma va direttamente dal genero e lo aggredisce.

La neuropsicologia forense entra in gioco per due motivi. I periti di parte vogliono dimostrare che GH ha probabilmente capacità di giudizio alterate dal trauma cranico subito (e non trattato) anni prima. Inoltre, secondo la difesa, GH non ha capito esattamente quello che intendeva la figlia, fraintendendo il messaggio come se la figlia fosse stata aggredita fisicamente dal genero.

Per valutare GH possiamo utilizzare un test che abbia comprovate Validità e Affidabilità e che misuri capacità di valutare Linguaggio Figurato, un possibile esempio potrebbe essere Linguaggio Figurato 1 di APACS

Per questo test abbiamo una consistenza interna sufficiente (0.65), una buona affidabilità test-retest ad un mese $r = 0.94$ e basso effetto pratica. C'è qualche evidenza di validità di costrutto, da analisi fattoriale e una buona validità contenuto a partire da valutazione di alcuni esperti.

Il cut-off per una persona con 68 anni di età e 5 di scolarità è 12. Considerando che il metodo usato da APACS per i cut-offs considera la distinzione tra campione e popolazione questo significa che un punteggio al di sotto del cut-off è atteso da meno del 5% della popolazione (di sani).

Se GH ottenesse meno 12 potremmo concludere che è una prestazione *insolita* se fosse sano, ma non possiamo dire se è un deficit pre-esistente, o un danno legato al trauma cranico (I metodi per calcolare deficit o danno, sono analoghi).

Andando a vedere le tabelle di APACS, possiamo vedere che però nel gruppo di sani usati, sono 40 rientrano nel gruppo di età e scolarità di GH. Le stime sono quindi abbastanza solide.