

I TEST NON PARAMETRICI

(Per I termini statistici di comprensione non immediata

si può far riferimento al glossario)

Sono metodi statistici parametrici tutti i metodi che si basano su distribuzioni probabilistiche quali la distribuzione normale o la distribuzione t di student. In contrapposizione le tecniche non parametriche sono quelle che non sono basate su alcuna distribuzione probabilistica. Quindi in generale i **test non parametrici** dovrebbero essere preferiti quando i dati non si distribuiscono secondo una normale, o comunque non si è in grado di dimostrarlo, ad esempio per numerosità ridotta.

La seguente tabella mostra alcuni dei più diffusi test non parametrici ed i loro equivalenti parametrici (basati sulla distribuzione normale).

UTILIZZO DEL TEST	TEST NON PARAMETRICI	TEST PARAMETRICI
<i>Test su un parametro di locazione (media o mediana)</i>	Test dei ranghi segnati di Wilcoxon	t-test
<i>Confronto di due campioni indipendenti</i>	Mann-Whitney U test Test dei ranghi segnati di Wilcoxon	t-test per campioni indipendenti
<i>Confronto di differenze</i>	Test dei ranghi segnati di Wilcoxon per dati appaiati	t-test per dati appaiati
<i>Bontà di adattamento dei dati ad un modello</i>	Test chi-quadro (Chi-square goodness of fit test) Kolmogorov-Smirnov	
<i>Determinazione di relazione lineare tra due variabili</i>	Coefficiente di correlazione dei ranghi si Spearman	Coefficiente di correlazione di Pearson
<i>Confronto di tre o più gruppi</i>	Analisi della varianza dei ranghi – Kruskal-Wallis	Analisi della varianza ad una via (F-test)
<i>Confronto di gruppi classificati secondo due fattori</i>	Analisi della varianza a due vie - Friedman	Analisi della varianza a due vie
<i>Analisi di tabelle di contingenza</i>	Test chi quadro per l'indipendenza Fisher exact test Test di Mantel-Haenszel	

Confronto tra <u>curve di sopravvivenza</u>	Test dei ranghi logaritmici (Log-rank test)	
---	---	--

Test basati sui ranghi

Occorre introdurre il concetto di rango: dato un insieme di **osservazioni** caratterizzate da misurazioni quantitative, le si ordinano in maniera crescente e ad ognuna si assegna un numero consecutivo crescente (1, 2, 3...), che viene detto rango. Quindi trasformando delle misurazioni in ranghi si perde l'informazione relativa alla misura mantenendo soltanto quella relativa all'ordinamento.

Test dei ranghi segnati di Wilcoxon: Si usa per testare se la **mediana** di una popolazione simmetrica è 0.

Test della somma dei ranghi di Wilcoxon (Mann-Whitney U test): Si usa per testare se due **campioni** provengono dalla stessa popolazione.

Test del Chi-quadrato per l'indipendenza

Questo test permette di studiare la relazione tra due **variabili categoriche**. In particolare il test chi-quadrato viene utilizzato per testare l'ipotesi nulla secondo cui le variabili, indicate con A e B, siano indipendenti, cioè tali che tutte le categorie della variabile B abbiano la stessa **frequenza** al variare delle categorie della variabile A

Coefficiente di correlazione tra ranghi di Spearman (Rho di Spearman)

E' l'analogo non parametrico del **coefficiente di correlazione di Pearson**. Si usa quindi per calcolare la **correlazione** tra due variabili. Si calcola convertendo ognuna delle due variabili in ranghi e calcolando quindi il **coefficiente di correlazione di Pearson** tra i due insiemi di ranghi.

Test della mediana

Si usa per testare se due **campioni** sono estratti da **popolazioni** con la stessa **mediana**. Si calcola la mediana sui due campioni uniti e quindi si classifica ogni osservazione originale secondo il campione di provenienza ed il fatto che preceda o sia successiva alla mediana complessiva. Si ottiene così una tabella 2x2 su cui si applica il **test chi-quadro** per l'omogeneità delle proporzioni, per testare se le due popolazioni abbiano la stessa mediana

Vantaggi delle procedure non parametriche

I test non parametrici richiedono assunzioni meno stringenti sui dati rispetto a quelle necessarie per i consueti test parametrici, e questo soprattutto per piccoli campioni. Ad esempio il **t-test** per un campione richiede che i dati siano distribuiti secondo la **distribuzione normale**. Per il t-test per due campioni indipendenti si richiede inoltre che le **deviazioni standard** calcolate sui due campioni siano uguali. Se tali assunzioni/condizioni vengono violate, il **P-value** e gli **intervalli di confidenza** possono risultare inattendibili. I test non parametrici come i test di Wilcoxon non richiedono tali assunzioni/condizioni. Bisogna però dire che talvolta trasformazioni dei dati (ad es. una trasformazione logaritmica) potrebbe riportare alla normalità o all'uguaglianza delle **varianze**.

I test non parametrici sono talvolta molto semplici da applicare in quanto non implicano una forte complessità computazionale, come i test di Wilcoxon.

I test non parametrici forniscono risultati più obiettivi dei parametrici nel caso in cui i dati siano misurati con scale ordinali "soggettive" del tipo *assolutamente non confortevole, poco confortevole, neutrale, confortevole, molto confortevole*. I test parametrici richiedono la conversione di questo tipo di scale in scale numeriche e producono dei risultati a volte troppo dipendenti dalla qualità di tale conversione.

Permettono il calcolo esatto della significatività statistica più agevolmente dei test parametrici.

Talvolta i dati a disposizione non sono un campione casuale della popolazione, ma la totalità della popolazione. Nel caso in cui la numerosità sia ridotta le procedure parametriche che sono basate su campionamento da popolazioni di dimensioni maggiori del campione, non sono più appropriate. In questi casi invece le procedure non parametriche possono essere utilizzate ugualmente attraverso alcuni artifici (modelli di randomizzazione).

Svantaggi delle procedure non parametriche

Il maggior svantaggio delle procedure non parametriche è contenuto nel nome stesso: il termine “non-parametrico” indica che tali test non coinvolgono la descrizione e quindi la stima di alcun parametro. E’ così difficile formulare ipotesi quantitative circa le reali differenze tra le popolazioni. Ad esempio parlare di differenze tra i segni di due trattamenti non dice molto sulla reale differenza tra i trattamenti, e non permette il calcolo di intervalli di confidenza per tali differenze. Tuttavia l’uso di adeguati software statistici riesce talvolta ad aggirare in parte tale problema, permettendo ad esempio di calcolare **intervalli di confidenza** per mediane o differenze tra mediane.

I test non parametrici “gettano via” informazione. I test dei segni, ad esempio, usa soltanto il segno delle osservazioni. I ranghi conservano le informazioni circa l’ordine dei dati ma trascurano i veri valori. A causa di tale perdita di informazione le procedure non-parametriche non possono avere la stessa capacità di rilevare le differenze esistenti delle corrispondenti procedure parametriche quando queste sono applicabili.

LA METANALISI

“È causa di grande preoccupazione constatare come la professione medica non abbia saputo organizzare un sistema in grado di rendere disponibili, e costantemente aggiornate, revisioni critiche sugli effetti dell’assistenza sanitaria”
Cochrane A.

*Effectiveness and efficiency.
Random reflections on health service.
London: Nuffield Provincial Hospital Trust,
1972.*

Revisione della letteratura
Meta analisi da risultati pubblicati
Pooled-analysis da dati originali

METANALISI

- **tecnica che ha lo scopo di analizzare una serie di studi clinici condotti sullo stesso argomento;**
- **come in un trial clinico si ricerca un’informazione complessiva a partire dall’osservazione di singoli pazienti e utilizzando uno specifico protocollo con relativi criteri di inclusione e di esclusione dei pazienti, così nella **metanalisi** si cerca di ottenere un’informazione complessiva a partire**

dall'analisi di singoli studi clinici ed utilizzando un protocollo basato su criteri di inclusione e di esclusione degli studi (e non dei pazienti).

Nella pratica:

- si definisce un protocollo che sia quanto più possibile “vicino” ai protocolli dei singoli trials.
- definisce criteri di inclusione e di esclusione;
- identifica gli end-points (che debbono essere stati valutati nei singoli trials).
- si confronta questo protocollo con il protocolli dei singoli trials.
- inclusione se il trial risponde a tutti i requisiti richiesti;
- esclusione se il trial non risponde anche ad uno solo dei requisiti.

Metanalisi

Una revisione sistematica e quantitativa delle evidenze relative ad un determinato problema clinico

La **metanalisi** si differenzia dalle tradizionali revisioni della letteratura per:

1. La ricerca sistematica ed esaustiva delle evidenze disponibili (pubblicate e non pubblicate)
2. L'esplicitazione dei criteri di inclusione degli studi considerati
3. L'analisi **statistica** dei risultati degli studi

Vantaggi:

consente un riassunto quantitativo dell'evidenza scientifica

minor peso delle valutazioni soggettive rispetto alle tradizionali rassegne

riduzione dell'errore non sistematico, quindi > potenza (> numero di eventi)

consentono analisi sui sottogruppi

La **metanalisi** è una tecnica di analisi **statistica** che combina i risultati provenienti da studi diversi consentendo una sintesi **quantitativa** dei risultati.

trials con risultati falsamente negativi per ridotta potenza

Esempio di trials su terapia fibrinolitica nell'IMA

fino al 1986 solo 4 trials su 24 avevano evidenziato un effetto positivo mentre era evidente il rischio emorragico

una **metanalisi** del 1985 ha evidenziato l'efficacia del trattamento,

confermata poi dai mega-trials (GISSI-1, ISIS-2)
ritardo nella applicazione di un trattamento efficace

l'analisi dei sottogruppi può mettere in evidenza differenze nell'efficacia del trattamento puramente dovute al caso

Esempio:

**Studio SOLVD a CONSENSUS II evidenziavano minor effetto protettivo degli ACE-inibitori in pazienti con scompenso cardiaco trattati con aspirina
eterogeneità reale o casuale?**

Metanalisi recente su 12.000 pazienti non ha evidenziato alcuna eterogeneità nei sottogruppi

- “i risultati di alcune **metanalisi** sono stati contraddetti da studi successivi screditando così la tecnica”
- attenzione alle possibili distorsioni (errore sistematico)
- Esempio: **metanalisi** di piccoli trials evidenziava effetto positivo di magnesio e.v in IMA
risultati smentiti da ISIS-4

Possibili cause di errore sistematico in una **metanalisi:**

Selection bias

- criteri di inclusione degli studi
- ampiezza degli studi: <qualità di studi piccoli
- qualità degli studi
(randomizzazione, doppio-cieco, follow-up..)

Eterogeneità degli studi

Publication bias

- maggior facilità di pubblicazione per studi con risultati positivi
- esclusione di articoli non in lingua inglese
- citation bias
- multiple publication bias

soluzione registrazione a livello internazionale di tutti i trial in fase di progettazione

Attenzione a risultati solo marginalmente positivi

- l'elevata numerosità delle meta-analisi fa sì che differenze anche modeste tra trattamento sperimentale e controlli raggiungano la significatività **statistica**
- la minor probabilità di pubblicazione di trials con risultati negativi amplifica il peso dei trial con esito positivo

Attenzione se la **metanalisi indica un risultato favorevole di modesta entità del nuovo trattamento derivanti da piccoli trials**

Se dalla rappresentazione grafica emerge eterogeneità dei risultati dei singoli studi,

la meta-analisi non è giustificata

**L'eterogeneità può essere formalmente valutata
con un test statistico**

Limiti delle meta-analisi

- **Non sono studi sperimentali**
- **Dipendono dalla esaustività della ricerca
degli studi rilevanti**
- **Dipendono dalla qualità degli studi
inclusi**
- **Sovrastimano l'effetto a causa del publication bias**