

# Test statistici non parametrici

Prof.ssa Paola Borrelli

# METODI NON PARAMETRICI



**Non formulano assunzioni sulla distribuzione delle variabili**

Quando le assunzioni dei test parametrici sono violate, sono disponibili almeno tre soluzioni alternative:

- 1) Ignorare la violazione, se questa è piccola;
- 2) Trasformare i dati (trasformazione logaritmica) in modo che soddisfino meglio le assunzioni;
- 3) Usare metodi non parametrici

I metodi per valutare l'adattamento della distribuzione normale a un insieme di dati comprendono principalmente le rappresentazioni grafiche e il test di Shapiro-Wilk

## QUANDO RICORRERE A UN TEST NON PARAMETRICO?

Spesso è possibile disporre solo di pochi dati, che sono assolutamente insufficienti per dimostrare la normalità della distribuzione o l'omogeneità delle varianze, in particolare quando il fenomeno studiato è nuovo e non è possibile citare dati di altre esperienze

**Gli aspetti da valutare** nella scelta del tipo di test (parametrico o non) in sintesi sono:

- **livello di misurazione della variabile di risposta.**
  - **assunti di applicabilità del test parametrico.**
  - **dimensioni campionarie.**
  - **campioni a confronto con un numero di dati molto differente**
  - **valori fortemente anomali (outliers) o indefiniti (perdita di valore della media e quindi delle misure da essa derivate, come la varianza).**
-

# VANTAGGI DEI TEST NON PARAMETRICI



- permettono di analizzare campioni di osservazioni provenienti da popolazioni diverse (**varianze disomogenee**),
- non richiedono un livello di misurazione tanto preciso come i test parametrici; alcuni si applicano a **variabili su scala ordinale, altri su scala nominale**,
- sono test basati su modelli che specificano soltanto condizioni molto generiche, meno ferree e più elastiche rispetto a quelle richieste dai test parametrici; **richiedono pochi assunti o nessun assunto sulle caratteristiche della popolazione obiettivo**,
- per la maggior parte, questi metodi sono fondati **sulle statistiche di rango**; non utilizzano la media, ma la **mediana** come misura della tendenza centrale; possono essere applicati indifferentemente a variabili casuali discrete, continue e ordinali

I test non parametrici si basano generalmente sui **ranghi** dei dati invece che sui loro valori effettivi.

I valori osservati vengono ordinati, dal più piccolo al più grande, e viene quindi registrato il rango di ognuno di essi, ossia la posizione rispetto alla lista ordinata.

Se **l'ipotesi nulla è vera**, i diversi ranghi saranno distribuiti circa nello stesso modo dei due gruppi.

Se invece uno dei due gruppi include la maggior parte delle misure più piccole e l'altro la maggior parte delle misure più grandi, saremo portati a ritenere che ci siano evidenze per rifiutare l'ipotesi nulla

| UTILIZZO DEL TEST                                   | TEST NON PARAMETRICI   | TEST PARAMETRICI                        |
|---|--|---|
| Test su un parametro di posizione (media o mediana) | <ul style="list-style-type: none"> <li>• <i>Test dei segni</i></li> <li>• <i>Test dei ranghi di Wilcoxon</i></li> </ul>              | T-test per campione singolo             |
| Confronto di due campioni indipendenti              | <ul style="list-style-type: none"> <li>• <i>Mann-Whitney U test</i></li> <li>• <i>Test dei ranghi segnati di Wilcoxon</i></li> </ul> | T-test per campioni indipendenti        |
| Confronto di due campioni appaiati                  | <ul style="list-style-type: none"> <li>• <i>Test dei ranghi di Wilcoxon per campioni appaiati</i></li> </ul>                         | T-test per campioni appaiati            |
| Normalità della distribuzione                       | <ul style="list-style-type: none"> <li>• <i>Test di Shapiro-Wilk</i></li> </ul>  |   |
| Relazione lineare tra due variabili                 | <ul style="list-style-type: none"> <li>• <i>Coefficiente di correlazione dei ranghi di Spearman</i></li> </ul>                       | Coefficiente di correlazione di Pearson |
| Analisi di tabelle di contingenza                   | <ul style="list-style-type: none"> <li>• <i>Test chi quadro per l'indipendenza</i></li> <li>• <i>Fisher exact test</i></li> </ul>    |   |

# II TEST DEI SEGNI

Il **test dei segni** è un test non parametrico che si può usare in alternativa al **test  $t$**  per un campione o al test  $t$  per dati appaiati quando non è soddisfatta l'assunzione di normalità

Il test dei segni valuta se la **mediana** di una popolazione sia uguale a un valore specificato dall'ipotesi nulla

Le misure che si trovano al di sopra del valore specificato dall'ipotesi vengono classificate con il simbolo «+» e quelle al di sotto con il simbolo «-»

al posto dell'ipotesi nulla  $H_0: \mu = \mu_0$  contro l'alternativa  $H_1: \mu \neq \mu_0$   
utilizza  **$H_0: m_e = m_{e_0}$  contro l'alternativa  $H_1: m_e \neq m_{e_0}$**

# IL TEST DEI SEGNI PER RANGHI DI WILCOXON

A differenza del test dei segni, il test di Wilcoxon conserva le informazioni sui valori assoluti, ossia tiene conto di quanto ogni valore si trova al disopra o al di sotto della mediana ipotizzata

Utilizza non solo il segno, ma anche i ranghi, è quindi più potente di quello dei segni poiché utilizza una maggior quantità di informazione dei dati

MA????

Assunto: la distribuzione delle misure nella popolazione sia simmetrica intorno alla mediana.....in altre parole, assume che non sia presente asimmetria

Limitazione del test

# IL TEST U di Mann-Whitney

Questo test può essere usato al posto del test  $t$  per due campioni quando l'assunzione di normalità non è soddisfatta.

Questo test considera i ranghi delle misure per valutare se le distribuzioni di frequenza di due gruppi siano uguali.

Se le distribuzioni dei due gruppi hanno la stessa forma, allora il test U di M-W si può utilizzare per confrontare le popolazioni centrale (mediana) nei due gruppi

**$H_0: me_1 = me_2$**  (la differenza delle mediane è uguale a zero)

**$H_1: me_1 \neq me_2$**  (la differenza delle mediane è diversa da zero)

<https://www.jamovi.org/download.html>