

## LA VERIFICA DELLE IPOTESI SULLA DIFFERENZA TRA LE MEDIE

Nella sezione dedicata alle fasi della verifica delle ipotesi, abbiamo esaminato un esempio di verifica delle ipotesi sulla media nel caso di un campione. In situazioni di quel tipo si vuole verificare se la media di un campione si discosta in maniera significativa da quella della popolazione di riferimento, per la variabile in oggetto. Il ricercatore può essere interessato, ad esempio, a verificare se i bambini sottoposti ad un metodo sperimentale di insegnamento della geografia hanno voti medi in geografia più alti rispetto al resto della popolazione di bambini di quell'età, oppure se nelle scuole di quartieri svantaggiati i punteggi medi in lettura e scrittura sono inferiori rispetto ai dati normativi della popolazione di riferimento, ecc.

Molto spesso l'ipotesi di ricerca riguarda invece il **confronto tra le medie di due o più campioni**. Le ipotesi che si sottopongono a verifica riguardano ovviamente non i campioni in esame, ma le **popolazioni** da cui i campioni sono estratti. La verifica delle ipotesi sulla differenza tra le medie può essere effettuata solo su **variabili su scala ad intervalli o rapporti equivalenti** (le uniche che consentono il calcolo della media) e consente di rispondere a domande del tipo (**esempi**):

- a) i maschi sono più abili delle femmine nelle prove visuo-spaziali?
- b) la memoria a breve termine dei giovani è migliore di quella degli anziani?
- c) i bambini con disturbo di lettura hanno punteggi più bassi nella consapevolezza fonologica rispetto ai bambini senza disturbo di lettura?
- d) mancini, destrimani e ambidestri, si differenziano nelle abilità verbali?

## Variabili indipendenti (V.I.) e dipendenti (V.D.).

La definizione di una variabile come indipendente o dipendente non è legata a caratteristiche peculiari della variabile, come è invece per il livello di misura, ma **al modo in cui la variabile viene trattata in un disegno sperimentale**.

Prendiamo l'esempio della **regressione lineare**. Se vogliamo prevedere il punteggio di comprensione nella lettura in base allo span di memoria, allora stiamo considerando lo **span di memoria** come **Variabile Indipendente (X)** e la **comprensione nella lettura** come **Variabile Dipendente (Y)**. Questo è ragionevole anche da un punto di vista teorico, infatti ci sono prove che lo span di memoria influenzi la comprensione della lettura nei bambini, quindi possiamo dire che quest'ultima "dipende" dallo span di memoria (ed anche, ovviamente, da tanti altri fattori). Tuttavia niente ci impedisce, da un punto di vista statistico e metodologico, di utilizzare la regressione lineare per prevedere lo span di memoria (Y) a partire dal punteggio di comprensione nella lettura (X), invertendo quindi Variabile Indipendente e Dipendente. **La definizione statistica di Variabile Indipendente e Variabile Dipendente può quindi non coincidere con quella teorica.**

Nella verifica delle ipotesi sulla differenza tra le medie viene definita **Variabile Indipendente**, in senso statistico, la variabile **"manipolata"** dallo sperimentatore, cioè quella **utilizzata per formare i gruppi da confrontare**. La **Variabile Dipendente** è invece **quella su cui viene effettuato il confronto tra le medie dei gruppi**. Se le differenze tra le medie, calcolate sulla Variabile Dipendente, risultano significative, si conclude che tali differenze "dipendono" dalle differenze tra i gruppi sulla Variabile Indipendente.

Riprendiamo il primo degli esempi formulati prima.

a) I maschi sono più abili delle femmine nelle prove visuo-spaziali?

Per verificare questa ipotesi è necessario selezionare un gruppo di maschi ed un gruppo di femmine, somministrare a tutti una prova visuospatiale e verificare se le medie dei due gruppi alla prova visuospatiale risultano significativamente diverse. La **Variabile Indipendente** in questo caso è il **“genere”** (valori: **maschio**, **femmina**) su **scala nominale**, la **Variabile Dipendente** è l'**abilità visuospatiale** (valori: **punteggi alla prova visuospatiale**) su **scala ad intervalli o rapporti equivalenti**\*. Se la differenza tra le medie dei due gruppi alla prova visuospatiale risulta significativa, si concluderà che tale differenza “dipende” dalla variabile genere (oppure che le differenze di genere influenzano le abilità visuospatiali, o che c'è un “effetto” del genere sull'abilità visuospatiale).

Come esercitazione si può provare ad effettuare lo stesso ragionamento sugli altri esempi di ipotesi precedentemente formulati:

b) la memoria a breve termine dei giovani è migliore di quella degli anziani?

c) i bambini con disturbo di lettura hanno punteggi più bassi nella consapevolezza fonologica rispetto ai bambini senza disturbo di lettura?

d) mancini, destrimani e ambidestri, si differenziano nelle abilità verbali?

\* **N.B. Ai fini della formazione dei gruppi, la Variabile Indipendente deve essere di tipo **categoriale**, quindi su scala ordinale o nominale, o ridotta a scala ordinale o nominale; la Variabile Dipendente invece, deve essere su scala ad intervalli o rapporti equivalenti, per consentire il calcolo delle medie dei gruppi.**

## Test statistici utilizzati per la verifica delle ipotesi sulla differenza tra le medie.

### Per il confronto tra due campioni:

per campioni di numerosità  $\geq 30$  ---> **test z.**

per campioni di numerosità  $< 30$  ---> **t di Student.**

### Per il confronto tra due o più campioni:

#### **Analisi della Varianza**

Il test dell'Analisi della Varianza può essere utilizzato sia nel caso di due campioni sia che si vogliano confrontare le medie di tre o più campioni, ed in effetti è il test più usato per la verifica delle ipotesi sulla differenza tra le medie, anche perché si adatta a disegni sperimentali diversi ed anche molto complessi. Vedremo quindi più in dettaglio questo test successivamente.

Relativamente al test z ed alla t di Student che, ripetiamo, consentono di verificare ipotesi sulla differenza tra le medie soltanto nel caso di due campioni, vedremo qui gli ambiti di applicazione e l'interpretazione dei risultati (per i presupposti teorici e le formule di calcolo si rimanda al testo *Statistica per la Psicologia*, vol II, pagg. 104-114).

## Esempi di applicazione del test z e della t di Student

L'unica differenza tra i due test riguarda la **numerosità dei campioni messi a confronto**, per il tipo di problemi ed il tipo di ipotesi da verificare i due test sono equivalenti. Per la procedura di verifica dell'ipotesi cambia soltanto la statistica calcolata e la relativa distribuzione campionaria.

### Esempio 1.

Due gruppi di 40 bambini, rispettivamente di 5 e 6 anni di età, vengono sottoposti al *Visual Form Discrimination Test* di Benton (Risposte Corrette) per la valutazione delle abilità percettive. Si vuole verificare se la prestazione dei due gruppi al test è significativamente diversa ( $\alpha = 0.05$ ). Dal confronto tra le medie dei due gruppi si ottiene:  $z = -1.28$ ;  $z$  critico =  $\pm 1.96$ .

#### Disegno sperimentale

**Variabile Indipendente:** **Età** Valori: **5 anni, 6 anni**

**Livello di misura:** **scala ordinale** (ai fini della formazione dei gruppi, i valori della V.I., che è su scala a rapporti equivalenti, sono trattati come categorie in quanto rappresentano classi di valori, per esempio la categoria “5 anni” comprende bambini di 5 anni ed 1 mese, 5 anni e 2 mesi e così via, e lo stesso accade per la categoria “6 anni”)

**Variabile Dipendente:** **Abilità percettiva**

**Valori:** **N. di risposte corrette al test di Benton**

**Livello di misura:** **scala a rapporti equivalenti**

#### Ipotesi statistiche

$H_0: \mu_1 = \mu_2$  (ovvero i due campioni provengono da popolazioni con medie uguali al test di Benton)

$H_1: \mu_1 \neq \mu_2$  (ovvero i due campioni provengono da popolazioni con medie diverse al test di Benton)

**Ipotesi bidirezionale**

**Risultati e commento:** è stato utilizzato il test  $z$  perché la numerosità dei campioni è superiore a 30. I valori critici di  $z$  ( $+1.96$  e  $-1.96$ ) delimitano le due “code” destra e sinistra della distribuzione normale, ognuna corrispondente al 2.5% della curva ( $\alpha/2$ ). Il valore di  $z$  calcolato ( $-1.28$ ) non supera il valore critico, quindi non si può rifiutare  $H_0$ . Non è quindi possibile affermare che la prestazione al test di Benton sia influenzata dalla differenza di età dei gruppi considerati.

### Esempio 2.

A due gruppi di ragazzi che praticano attività sportiva, il primo gruppo ( $n = 12$ ) di tipo agonistico (Gruppo A) ed il secondo gruppo ( $n = 16$ ) di tipo non agonistico (Gruppo NA), viene somministrato un test sull'ansia legata alla prestazione sportiva. Si vuole verificare se il punteggio medio di ansia sia più elevato per il gruppo che pratica attività agonistica ( $\alpha = 0.01$ ). Si ottengono i seguenti risultati: Media ansia Gruppo A = 42; Media ansia Gruppo NA = 36;  $t$  di Student = 2.64;  $t$  critico = 2.48

#### Disegno sperimentale

**Variabile Indipendente:** **Tipo di attività sportiva**

**Valori:** **Agonistica, Non Agonistica**

**Livello di misura:** **scala nominale**

**Variabile Dipendente:** **Ansia**

**Valori:** **Punteggi al test**

**Livello di misura:** **scala ad intervalli equivalenti**

#### Ipotesi statistiche

$H_0: \mu_1 = \mu_2$  (ovvero i due campioni provengono da popolazioni con medie uguali al test di ansia)

$H_1: \mu_1 > \mu_2$  (ovvero la popolazione di provenienza del gruppo A ha media superiore al test di ansia rispetto alla popolazione di provenienza del gruppo NA)

**Ipotesi monodirezionale destra**

**Risultati e commento:** è stato utilizzato il **test  $t$**  perché la numerosità dei campioni è inferiore a 30. Il valore critico di  $t$  (2.48) delimita la “coda” di destra della distribuzione della  $t$  di Student, che comprende l'1% dei risultati estremi positivi. Il valore di  $t$  calcolato (2.64) supera il valore critico, quindi si può rifiutare  $H_0$ . Si può concludere, con una probabilità di errore dell'1%, che i ragazzi che praticano attività sportiva di tipo agonistico vivono l'attività sportiva in modo più ansioso di quelli che praticano attività sportiva di tipo non agonistico.