

Laboratorio di metodi statistici per la psicologia dello sviluppo

3. Effect-size: La dimensione dell'effetto

Gianmarco Altoè

Dipartimento di Psicologia dello Sviluppo e della Socializzazione

Corso di Laurea Magistrale in Psicologia Clinica dello Sviluppo
Anno Accademico 2018 - 2019, Università di Padova

Sommario

- 1 **Introduzione**
- 2 **Il d di Cohen**
- 3 **Altre misure di effect size**
- 4 **Aspetti computazionali**
- 5 **Esercizio**

Cosa si diceva già parecchi anni fa

“[...] the emphasis given to formal tests of significance [...] has caused scientific research workers to pay undue attention to the **results of the tests of significance** they perform on their data, [...] and too little to the **estimates of the magnitude of the effects** they are investigating” (*Yates, 1951*)

Il problema: Significatività statistica e significatività clinica

- Ad esempio, il fatto che la differenza tra gruppo sperimentale e gruppo di controllo sia statisticamente significativa, è un informazione rilevante
- Questo però non ci dice nulla sulla grandezza della differenza. Un differenza significativamente diversa da 0, può non avere o avere poca rilevanza da un punto di vista pratico o clinico considerando la dimensione della differenza stimata tra i due gruppi.

Gli indici di effect-size

- Gli indici di effect-size servono a valutare la dimensione (magnitudine) dell'effetto di interesse indipendentemente dall'unità di misura delle variabili indagate
- Insieme al risultato inferenziale di una analisi, l'effect size deve essere attentamente valutato dal ricercatore per dare un'appropriata interpretazione del fenomeno di studio

Alcune raccomandazioni metodologiche

“APA stresses that *Null Hypothesis Statistical Significance Testing (NHST)* is but a starting point and that additional reporting elements such as **effect sizes**, confidence intervals, and extensive description are needed to convey the most complete meaning of the results”

(Publication Manual of the American Psychological Association (APA), 2010)

“[...] **when the results of an analysis are statistically significant, it is useful to consider the substantive interpretation of the effect.** For this purpose, the observed difference can be converted into an **effect size** to allow the interpretation of the size of the difference”

(National Center for Education Statistics (NCES), 2002)

“[...] è bene rilevare che pubblicare le dimensioni degli effetti [...] è sicuramente importante, ma altrettanto lo è **commentarle**”

Altoè, 2014

Il d di Cohen

- Uno degli indici più conosciuti per misurare la dimensione dell'effetto è il d di Cohen. A titolo esemplificativo studieremo il d per valutare la dimensione dell'effetto della differenza tra le medie di due gruppi indipendenti rispetto ad una variabile quantitativa
- La semplice differenza tra le medie di 2 gruppi dipende dall'unità di misura (e spesso le variabili utilizzate in psicologia non hanno una propria unità di misura: ad es. ansia, empatia, memoria, apprendimento) e inoltre non tiene conto della variabilità della variabile quantitativa nei due gruppi
- Cohen propose come dimensione dell'effetto l'indice d che può essere visto come la differenza tra le medie standardizzata. Dove come fattore di standardizzazione si utilizza la deviazione standard combinata (pooled) dei due gruppi
- Ad esempio: $d = .5$ indica che la differenza tra le due medie è pari alla metà della deviazione standard combinata. Come si può notare tale indicatore è indipendente dall'unità di misura del fenomeno indagato

Il d di Cohen: la formula

Stima del d di Cohen (Cohen, 1988):

$$d = \frac{\overline{X}_a - \overline{X}_b}{S_{pooled}}$$

dove:

- S_{pooled} è la deviazione standard combinata dei due campioni

$$S_{pooled} = \sqrt{\frac{(n_a - 1)s_a^2 + (n_b - 1)s_b^2}{n_a + n_b - 2}}$$

- $\overline{X}_a - \overline{X}_b$ è la differenza delle medie campionarie
- s_a^2 e s_b^2 sono le varianze campionarie dei due campioni
- n_a e n_b sono le numerosità dei due campioni

Il d di Cohen: una prima interpretazione

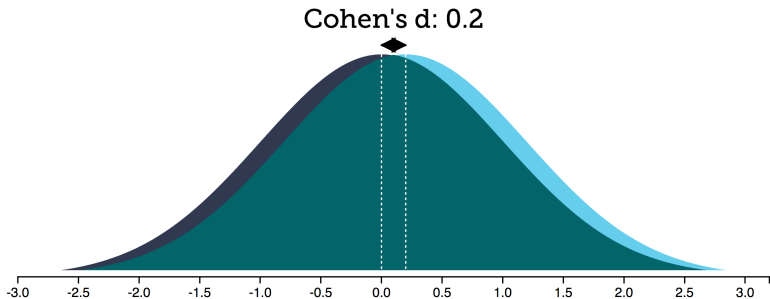
Per dare una **PRIMA IDEA** di come interpretare i valori di d , Cohen indicò i seguenti valori di riferimento:

- $d = .2$ effetto piccolo (small).
Example by Cohen (1977): The difference between the heights of 15 year old and 16 year old girls in the US
- $d = .5$ effetto medio (medium).
Example by Cohen (1977): The difference between the heights of 14 year old and 18 year old girls in the US
- $d = .8$ effetto grande (large).
Example by Cohen (1977): The difference between the heights of 13 year old and 18 year old girls in the US

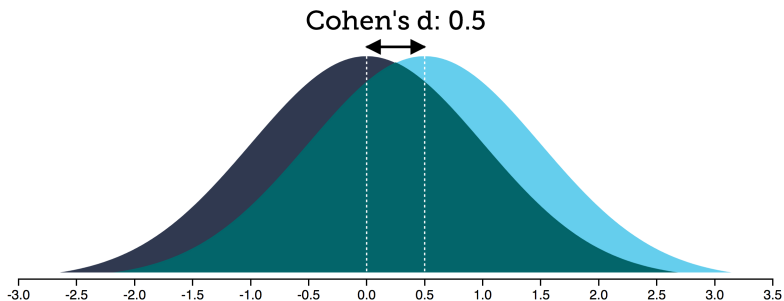
NOTA BENE. Lo stesso Cohen e altri autori in seguito specificarono che il valore di un effect-size deve essere valutato sulla base della teoria di riferimento, nell'ottica dei costi e benefici e in ottica comparativa rispetto a **studi simili** e non limitatamente ad una sorta di "t-shirt sizes classification"

Il d di Cohen: un'interpretazione grafica

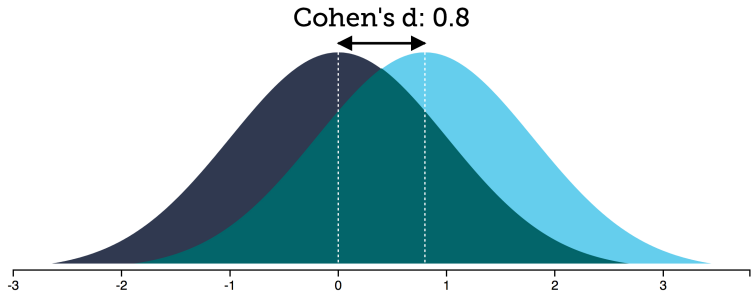
- Per chiarirci ulteriormente le idee utilizziamo la rappresentazione grafica
- Ipotizziamo che la variabile dipendente di interesse sia distribuita normalmente e con la stessa varianza nei due gruppi di interesse
- Utilizzeremo un tutorial free, molto utile e che si consiglia di visionare:
<http://rpsychologist.com/d3/cohend/>
- ...

$d = .2$ 

Nota. La probabilità che un soggetto preso a caso dal gruppo azzurro (media maggiore) abbia un punteggio maggiore rispetto a un soggetto preso a caso dal gruppo verde (media minore) è pari al **56%** (probability of superiority)

$d = .5$ 

Nota. La probabilità che un soggetto preso a caso dal gruppo azzurro (media maggiore) abbia un punteggio maggiore rispetto a un soggetto preso a caso dal gruppo verde (media minore) è pari al **64%** (probability of superiority)

$d = .8$ 

Nota. La probabilità che un soggetto preso a caso dal gruppo azzurro (media maggiore) abbia un punteggio maggiore rispetto a un soggetto preso a caso dal gruppo verde (media minore) è pari al **71%** (probability of superiority)

Differenze di autostima tra maschi e femmine

- Riprendiamo i dati relativi all'esercizio 5.5 visto nel corso di Analisi dei Dati.
- Nel file `d1ab.rda` sono contenuti i punteggi di Autostima (Rosenberg, 1965) ottenuti da maschi e femmine statunitensi di età compresa tra i 18 e i 30 anni. La numerosità campionaria complessiva è pari a 7735.
- Valutare se esiste differenza significativa tra maschi e femmine e calcolare la dimensione di tale effetto attraverso l'indice d di Cohen.

I dati

```

> rm(list=ls())
> load("dlab.rda")
> ls()

[1] "dlab"

> str(dlab)

'data.frame': 7735 obs. of  3 variables:
 $ id          : int  1 2 3 4 5 6 7 8 9 10 ...
 $ genere      : Factor w/ 2 levels "M","F": 2 1 1 2 1 1 1 1 2 1 ...
 $ autostima: num  2.4 1.1 2.1 2.3 4 3 2.8 3.5 2.5 1.5 ...

> # maschi e femmine
> table(dlab$genere)

```

```

      M      F
2689 5046

```

```

> # autostima (punteggi elevati indicano elevata autostima)
> summary(dlab$autostima)

```

```

      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 1.000   2.100   2.600   2.585   3.100   4.000

```

Statistiche descrittive

```

> # vedi ?tapply
> ## Medie per genere
> tapply(dlab$autostima,dlab$genere,mean)

      M      F
2.658981 2.545422

> ## Deviazioni standard per genere
> tapply(dlab$autostima,dlab$genere,sd)

      M      F
0.7204097 0.6728148

> #
> # COMMENTI?

```


Verifica di ipotesi

```
> # Utilizziamo un t.test per campioni indipendenti
> t.test(dlab$autostima~dlab$genere,paired=FALSE,var.equal=TRUE)
```

Two Sample t-test

```
data:  dlab$autostima by dlab$genere
t = 6.8957, df = 7733, p-value = 5.781e-12
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.08127716 0.14584067
sample estimates:
mean in group M mean in group F
    2.658981      2.545422
```

Dall'analisi emerge che i maschi hanno una media significativamente maggiore rispetto alle femmine ($t(7733) = 6.896$, $p < .001$; $M_{maschi} = 2.66$, $M_{femmine} = 2.55$)

Calcolo dell'effect size

```
> # Utilizziamo uno tra i tanti pacchetti di R
> # contenenti funzioni per il calcolo di effect size
> library(lsr)
> cohensD(dlab$autostima[dlab$genere=="M"],
+         dlab$autostima[dlab$genere=="F"])
[1] 0.1646422
```

L'effect size ($d = .16$) indica che la differenza tra maschi e femmine può essere ritenuta piccola (sulla base dei valori di riferimento di Cohen, 1988). Tuttavia tale valore andrebbe comparato con altri studi simili e interpretato sulla base della teoria del fenomeno indagato.

Una possibile classificazione

Gli indici di effect-size possono essere suddivisi in:

- differenze tra medie (ad es.: d di Choen, g di Hedges)
- varianza spiegata(ad es.: R^2 , η^2 , partial η^2)
- associazione tra variabili (ad es.: r , Odds Ratio, Rischio Relativo)

Per una rassegna completa si veda:

Borenstein, Hedges, Higgins, & Rothstein. (2011)

Introduction to meta-analysis. John Wiley & Sons, Ltd.

Aspetti computazionali

- 1 Oltre che dai dati grezzi gli effect-size possono essere calcolati a partire da statistiche riassuntive (tipicamente disponibili nelle pubblicazioni)
Vedi ad esempio la funzione [escalc](#) del pacchetto [metafor](#). Tale funzione oltre all'effect size restituisce in output il valore della varianza di stima (il quadrato dello standard error) dell'effect size, che come vedremo ci servirà nella meta-analisi
- 2 Inoltre è possibile convertire un indice di effect size in un altro indice di effect size (con l'obiettivo di avere per più studi lo stesso indice di effect size)
Vedi ad esempio il pacchetto [compute.es](#)

Riferimenti utili per i lavori di gruppo

- <http://www.metafor-project.org/doku.php>
- <https://cran.r-project.org/web/packages/computes/index.html>

Esercizio

```
> # Consideriamo "d" una parte del dataset dat.gibson2002
> # contenuto nel pacchetto metafor
> rm(list=ls())
> library(metafor)
> d<-dat.gibson2002[1:13,1:8]
> #####
> # DESCRIPTION OF d
> # Results from 13 trials examining the effectiveness
> # of self-management education and regular medical review
> # for adults with asthma.
> #####
> # SOURCE
> # Gibson et al. (2002). Self-management education and regular
> # practitioner review for adults with asthma.
> # Cochrane Database of Systematic Reviews, 3
```

Esercizio

```
> #####  
> # VARIABLES  
> ## author: first author of study  
> ## year: publication year  
> ## n1i: number of participants in the intervention group  
> ## m1i: mean number of days off work/school  
> ##      in the intervention group  
> ## sd1i: standard deviation of the number of days off work/school  
> ##      in the intervention group  
> ## n2i: mean number of participants in the control group  
> ## m2i: mean number of days off work/school  
> ##      in the control group  
> ## sd2i: standard deviation of the number of days off work/school  
> ##      in the control group
```

Esercizio

```
> # I dati
```

```
> d
```

| | author | year | n1i | m1i | sd1i | n2i | m2i | sd2i |
|----|----------------|------|-----|-------|-------|-----|-------|-------|
| 1 | Cote | 1997 | 50 | 2.20 | 12.73 | 54 | 5.20 | 12.50 |
| 2 | Ghosh | 1998 | 140 | 17.60 | 24.20 | 136 | 34.10 | 38.80 |
| 3 | Hayward | 1996 | 23 | 0.38 | 0.56 | 19 | 0.23 | 0.29 |
| 4 | Heard | 1999 | 97 | 2.09 | 5.93 | 94 | 2.66 | 4.95 |
| 5 | Ignacio-Garcia | 1995 | 35 | 4.92 | 6.05 | 35 | 20.00 | 26.34 |
| 6 | Knoell | 1998 | 45 | 0.85 | 4.75 | 55 | 2.31 | 9.16 |
| 7 | Lahdensuo | 1996 | 56 | 2.80 | 9.00 | 59 | 4.80 | 7.20 |
| 8 | Sommaruga | 1995 | 20 | 24.10 | 11.80 | 20 | 31.80 | 17.90 |
| 9 | Zeiger | 1991 | 128 | 1.40 | 3.30 | 143 | 2.30 | 7.60 |
| 10 | Garret | 1994 | 119 | 6.23 | 12.20 | 100 | 5.71 | 8.57 |
| 11 | Neri | 1996 | 32 | 2.10 | 8.00 | 33 | 5.10 | 14.00 |
| 12 | Hilton | 1986 | 86 | 0.73 | 1.48 | 100 | 0.47 | 1.20 |
| 13 | Gallefoss | 1999 | 25 | 8.00 | 32.00 | 24 | 26.00 | 70.00 |

Esercizio

```
> # 1) Calcolare il d di Cohen e l'associata varianza di stima
> #     per ogni studio
> # ( Hint: escalc(measure="SMD", ...) )
> # 2) Analizzare a livello descrittivo i d ottenuti
> # 3) Valutare a livello descrittivo la relazione tra
> #     i d ottenuti e le rispettive varianze di stima
```