

L'analisi statistica

Esempio (de Gans et al. NEJM 2002, 347: 1549-56)

| Esito | Trattamento | | Totale |
|---------------|--------------|------------|------------|
| | Desametazone | Placebo | |
| Sfavorevole | 23 | 36 | 59 |
| Favorevole | 134 | 108 | 242 |
| Totale | 157 | 144 | 301 |

RR = 0,60; RRR = 0,40; OR = 0,55; ARR = 0,10 NNT = 10

Possiamo interpretare questi risultati come dovuti ad una diversa efficacia dei trattamenti?

Il problema statistico

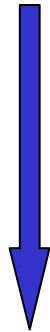
- ↪ Se non c'è un reale effetto del trattamento, è comunque possibile che per caso si osservino differenze tra i gruppi

Il problema statistico

- ↪ Se non c'è un reale effetto del trattamento, è comunque possibile che per caso si osservino differenze tra i gruppi
- ↪ Se c'è un reale effetto del trattamento, è comunque possibile che per caso non si osservi quella differenza fra i trattamenti (o, al contrario, che venga esagerata)

Il metodo statistico

Variabilità casuale degli effetti osservati



Quanto è probabile che la differenza
effettivamente osservata sia imputabile al caso?

Il test statistico: la significatività statistica

Esempio

La percentuale di pazienti con esito sfavorevole fu significativamente minore nel gruppo con desametazone rispetto al gruppo con placebo (15% vs 25%; rischio relativo 0,59; $p=0,03$)

(De Gans et al. NEJM 2002, 347: 1549-56)

Il test d'ipotesi

Fasi logiche

Ipotesi iniziale

Risultati attesi
sotto H_0

Confronto
osservati-attesi

Conclusioni

Fasi operative

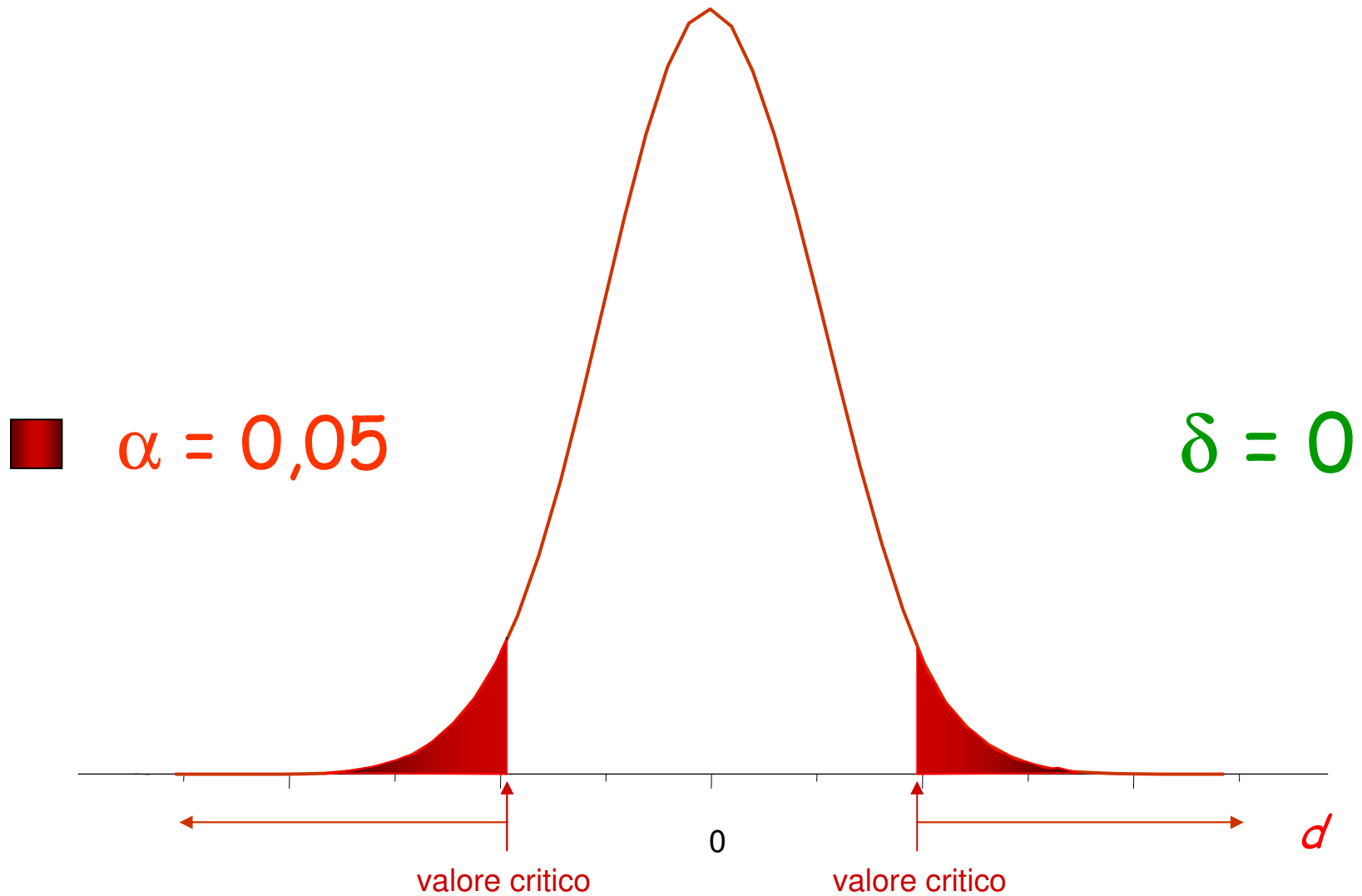
Ipotesi nulla (H_0)

Distribuzione di
campionamento

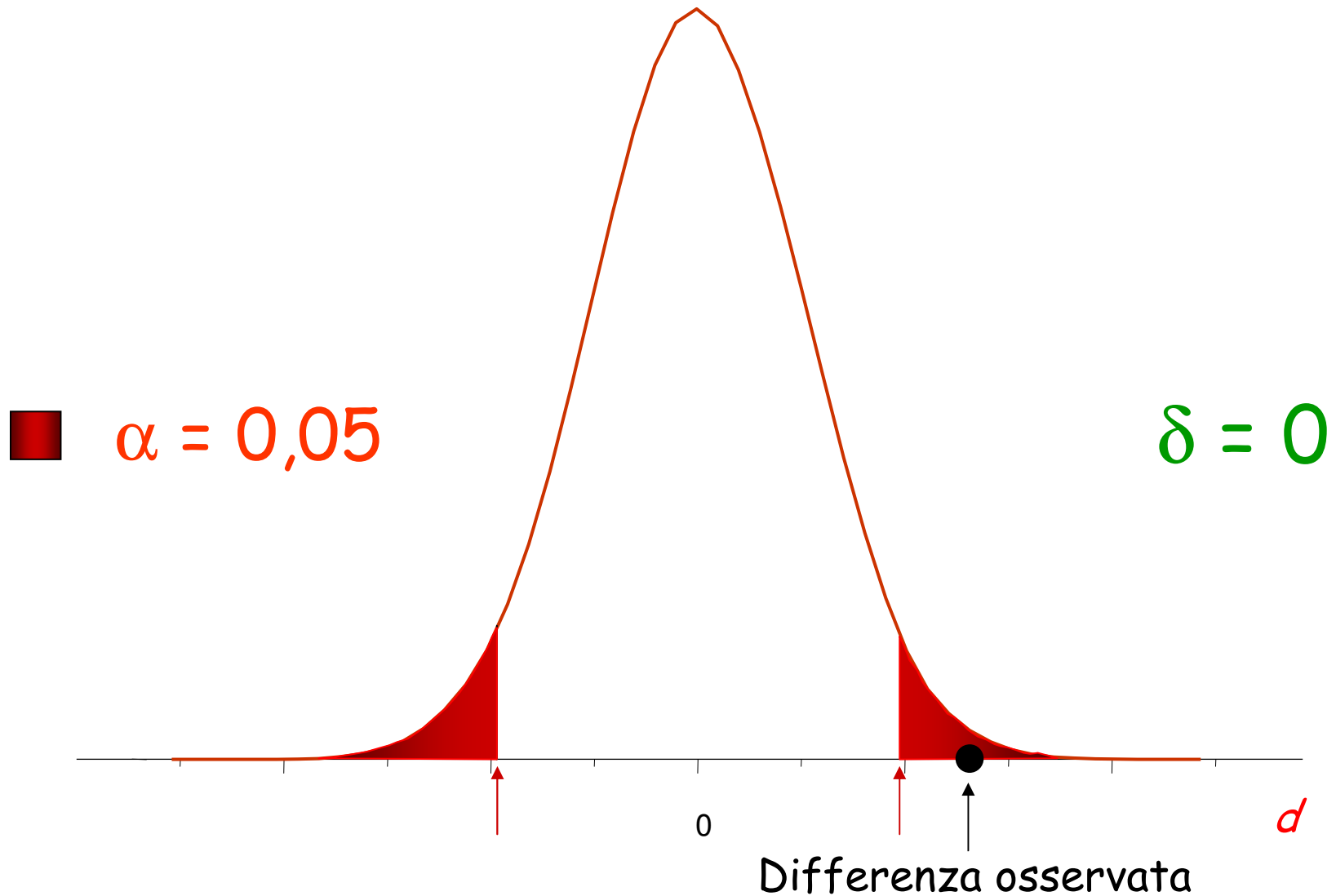
Test statistico

Rifiuto/non
rifiuto di H_0

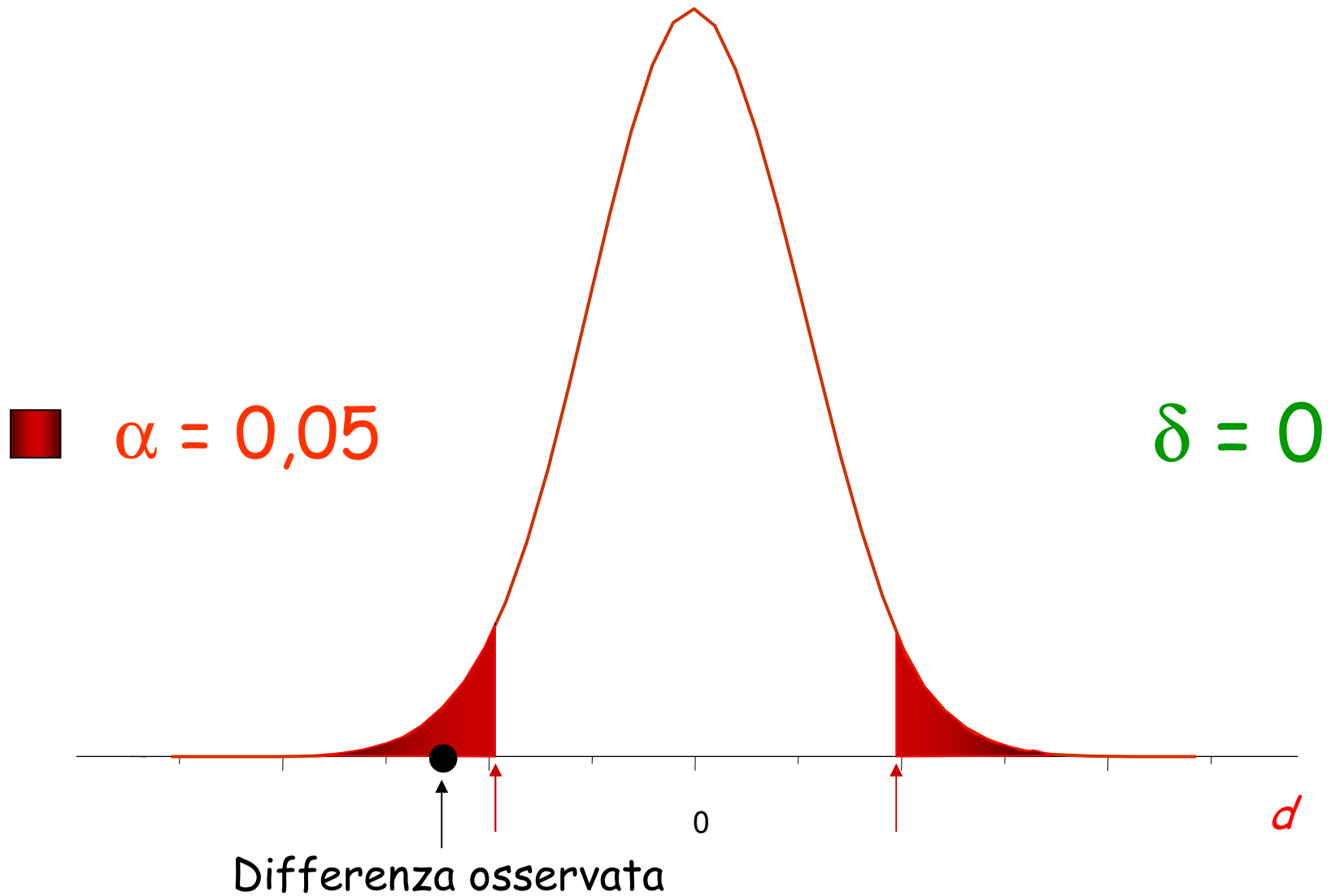
Il livello di significatività



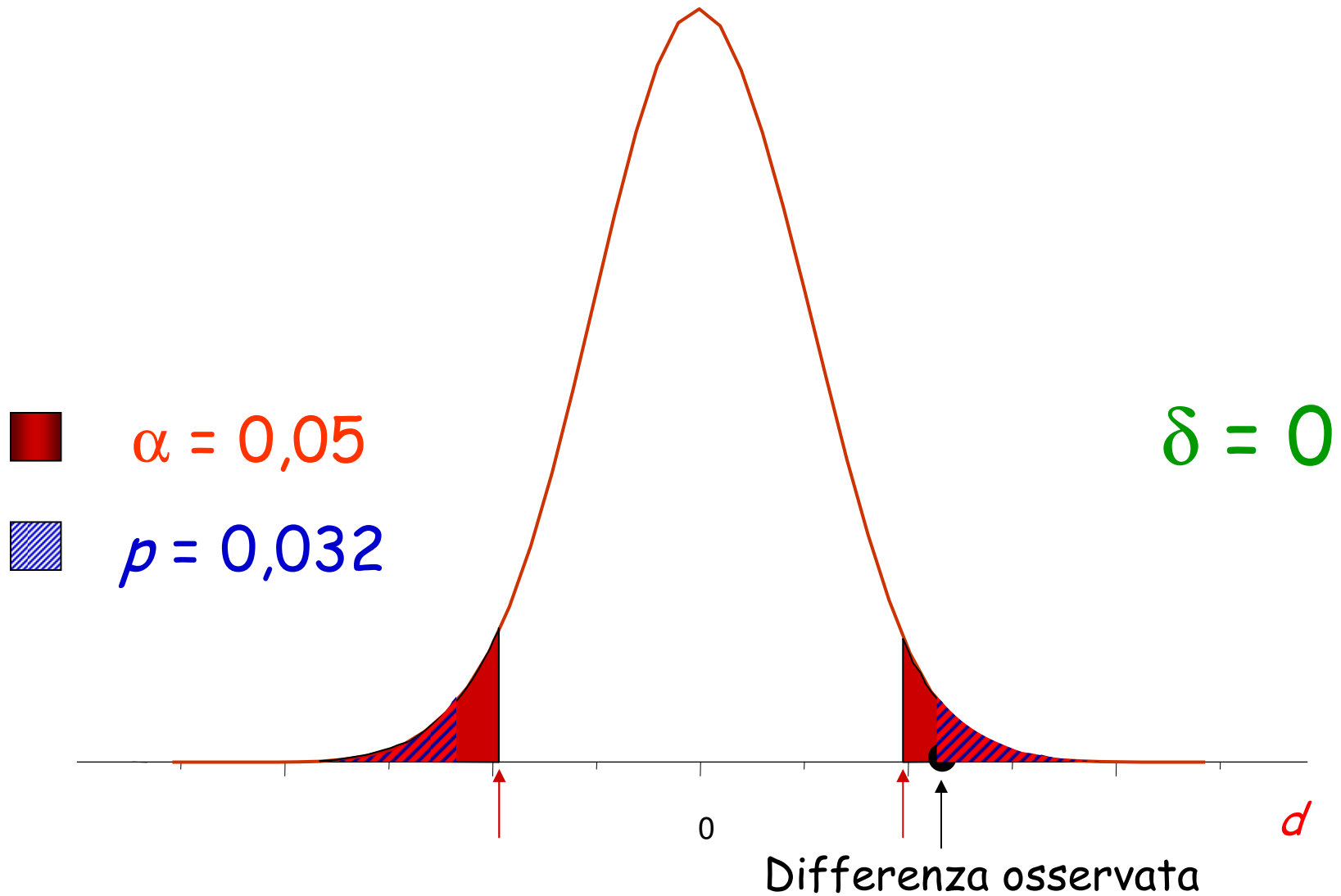
Statisticamente significativo



Statisticamente significativo



Statisticamente significativo



Esempio

(de Gans et al. NEJM 2002, 347: 1549-56)

| Esito | Trattamento | | Totale |
|---------------|--------------|------------|------------|
| | Desametazone | Placebo | |
| Sfavorevole | 23 | 36 | 59 |
| Favorevole | 134 | 108 | 242 |
| Totale | 157 | 144 | 301 |

$p = 0,03$

Il confronto è statisticamente significativo

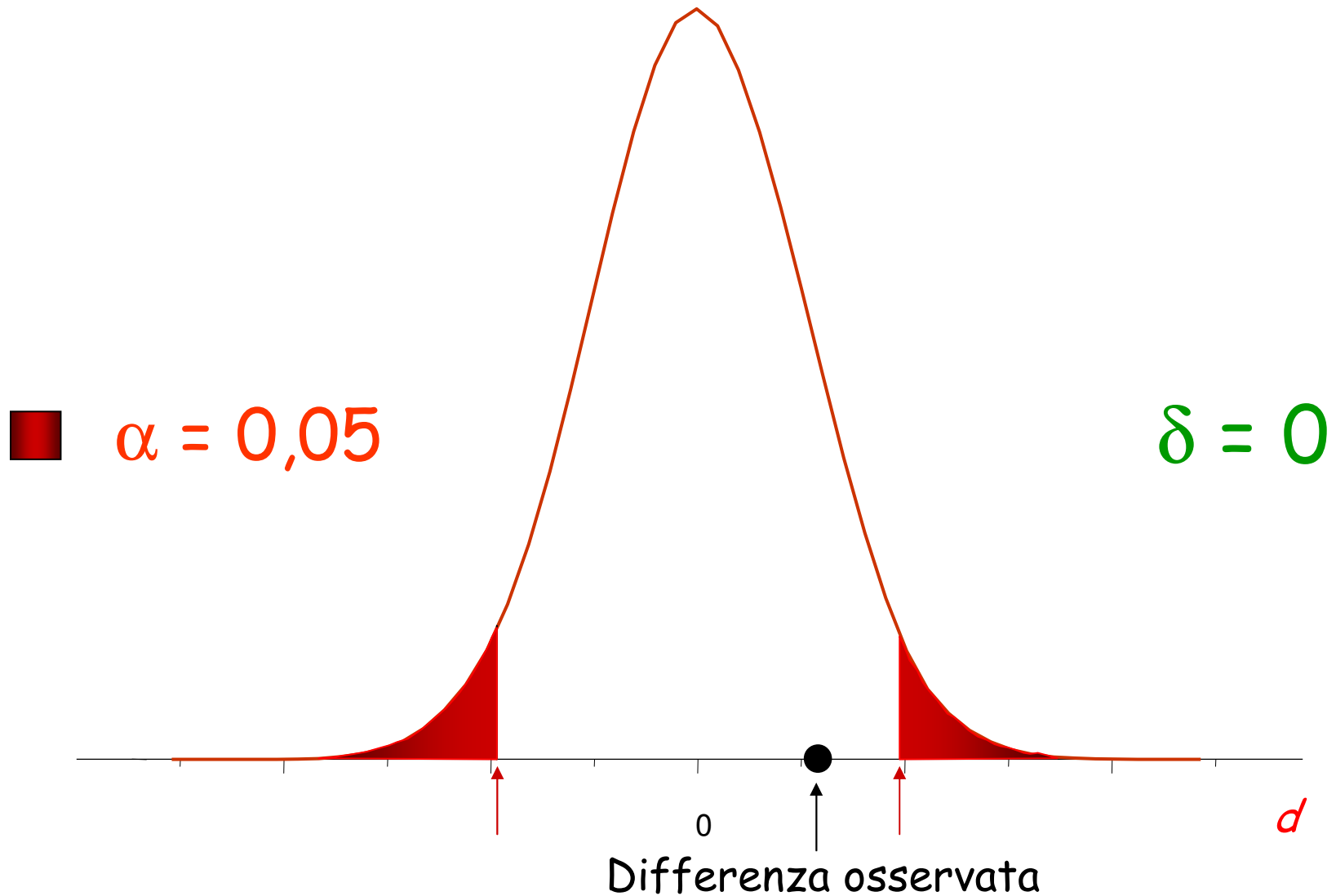
Esempio (Hammer et al. JAMA 2002, 288: 169-80)

| Esito | Trattamento | | Totale |
|-----------|-------------|---------|--------|
| | 2 IP | Placebo | |
| RNA < 200 | 112 | 36 | 148 |
| RNA > 200 | 212 | 121 | 333 |
| Totale | 324 | 157 | 481 |

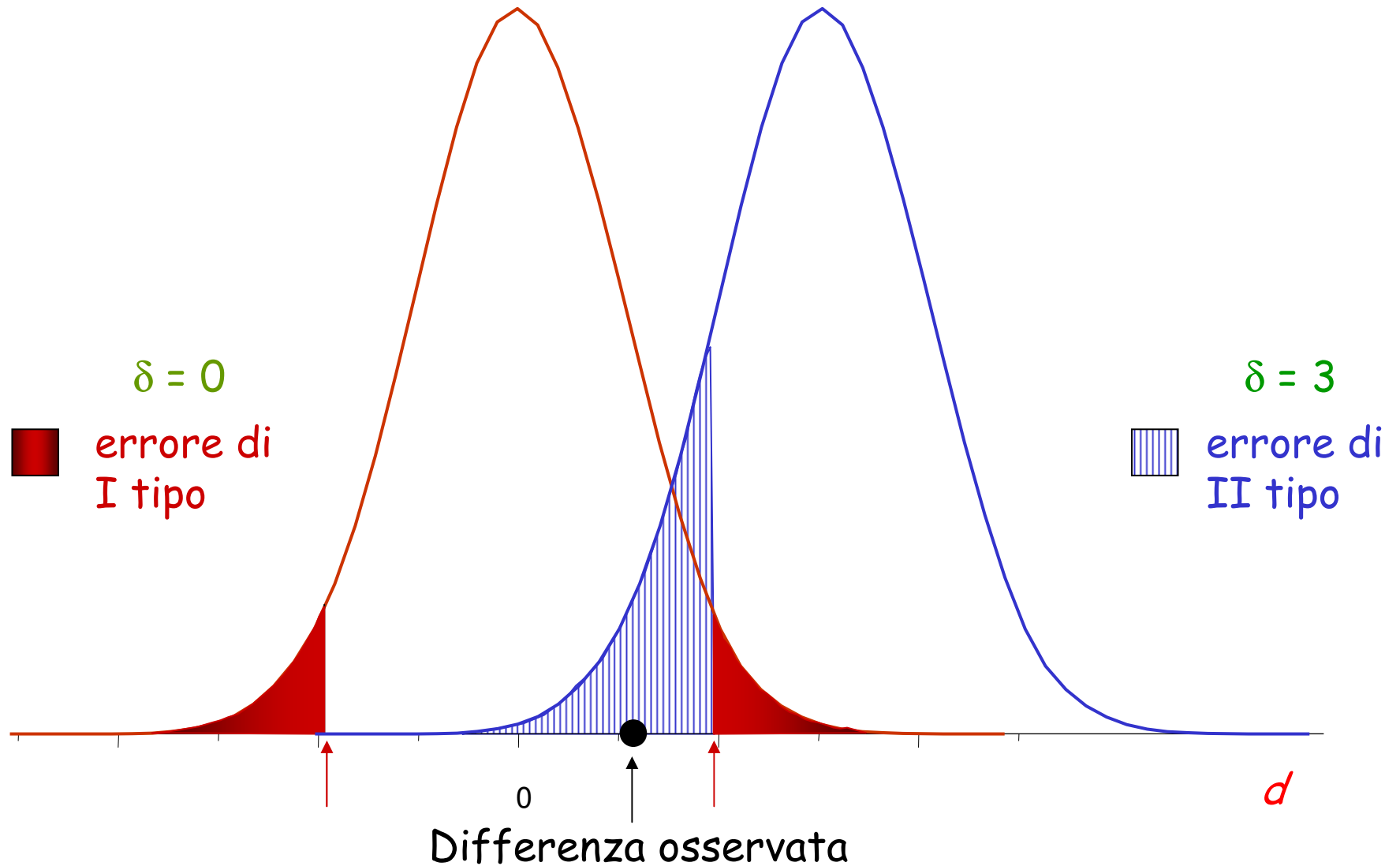
$p = 0,002$

Il confronto è statisticamente significativo

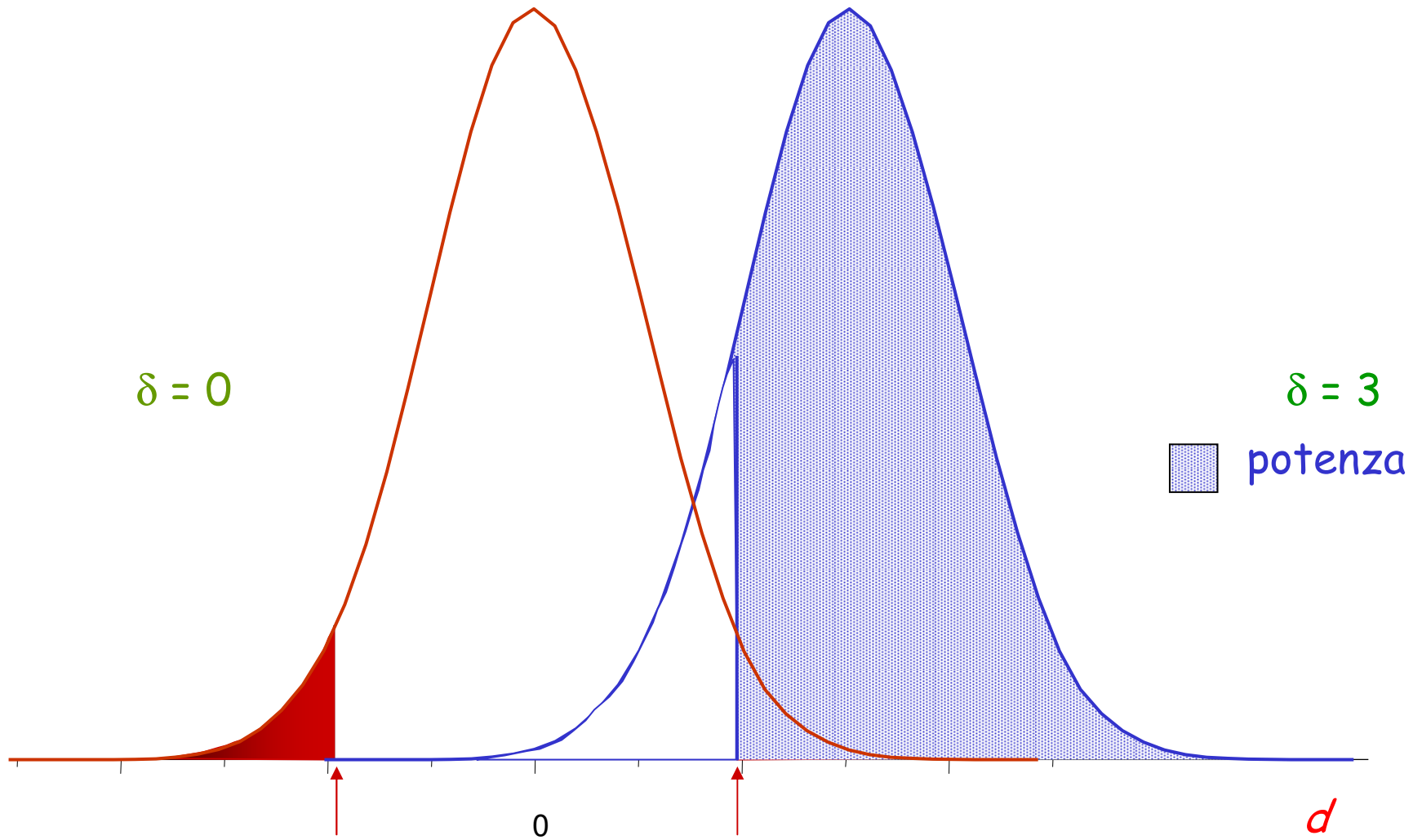
Statisticamente non significativo



L'errore di secondo tipo



La potenza del test



Significatività statistica e numerosità

- A parità di numerosità la p diminuisce con l'aumentare dell'effetto
- A parità di effetto la p diminuisce con l'aumentare della numerosità
- Con numerosità piccole è facile non riconoscere differenze clinicamente rilevanti
- Con numerosità elevate possono diventare 'statisticamente significative' anche differenze clinicamente irrilevanti

Esempio (Hammer et al. JAMA 2002, 288: 169-80)

| <u>Trattamento</u> | <u>HIV-RNA < 200</u> | p |
|--------------------|-------------------------|---|
| Saquinavir | 40/116 (34%) | |
| Indinavir | 25/69 (36%) | |
| Nelfinavir | 47/139 (34%) | |
| Placebo | 36/157 (23%) | |
| | | 0,25 0,29 0,004 |

Gli intervalli di confidenza

Esempio

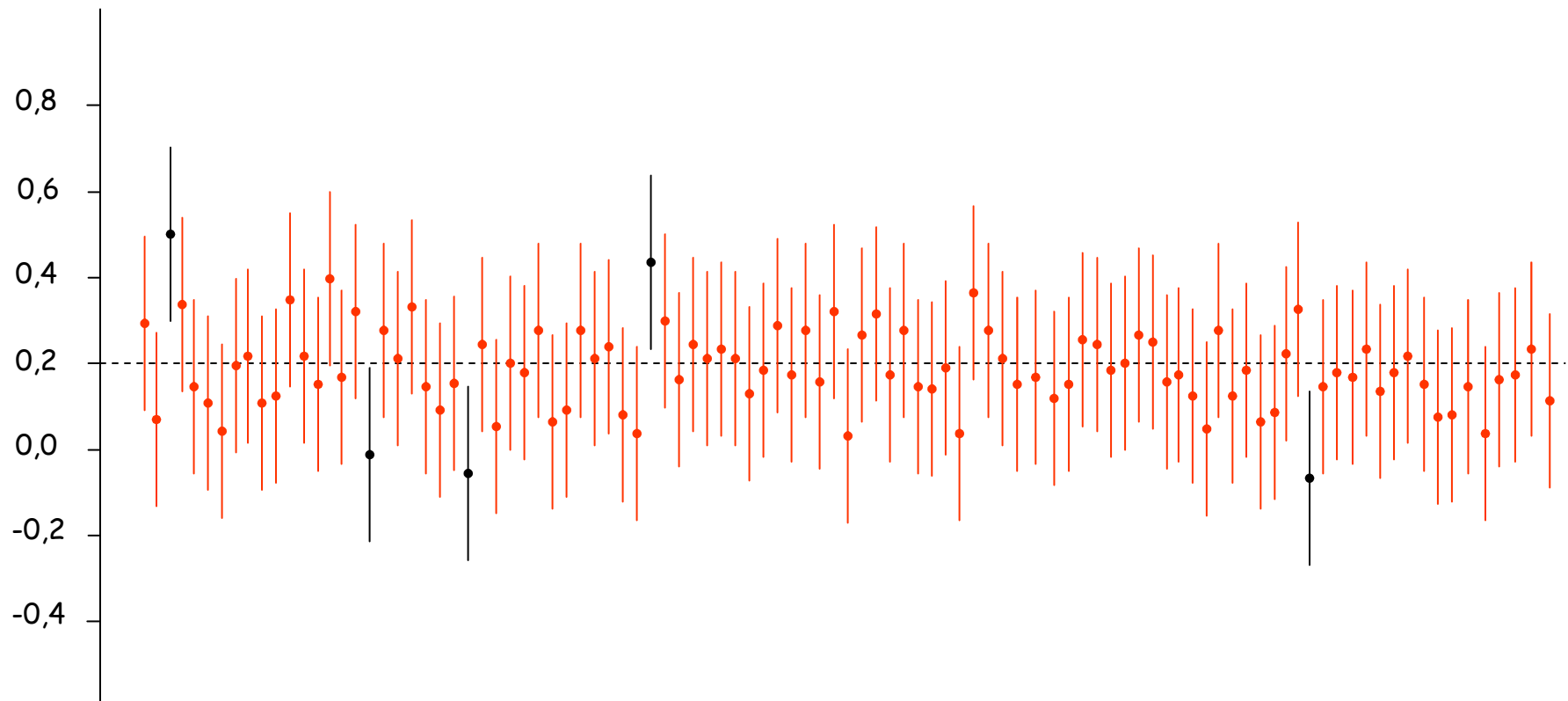
La percentuale di pazienti con esito sfavorevole fu significativamente minore nel gruppo con desametazone rispetto al gruppo con placebo (15% vs 25%; rischio relativo 0,59; intervallo di confidenza al 95% da 0,37 a 0,94; $p=0,03$)

(De Gans et al. NEJM 2002, 347: 1549-56)

Gli intervalli di confidenza

- Indicano un intervallo di valori all'interno del quale il ricercatore conclude, con una specificata probabilità, che sia compreso l'effetto vero del trattamento
- Forniscono informazioni sull'entità dell'effetto e sulla precisione della stima con un livello di 'confidenza' prestabilito (tipicamente 95%)

Gli intervalli di confidenza



Nel 95% circa dei campioni possibili l'intervallo di confidenza al 95% comprende l'effetto 'vero' ($\delta = 0,2$)

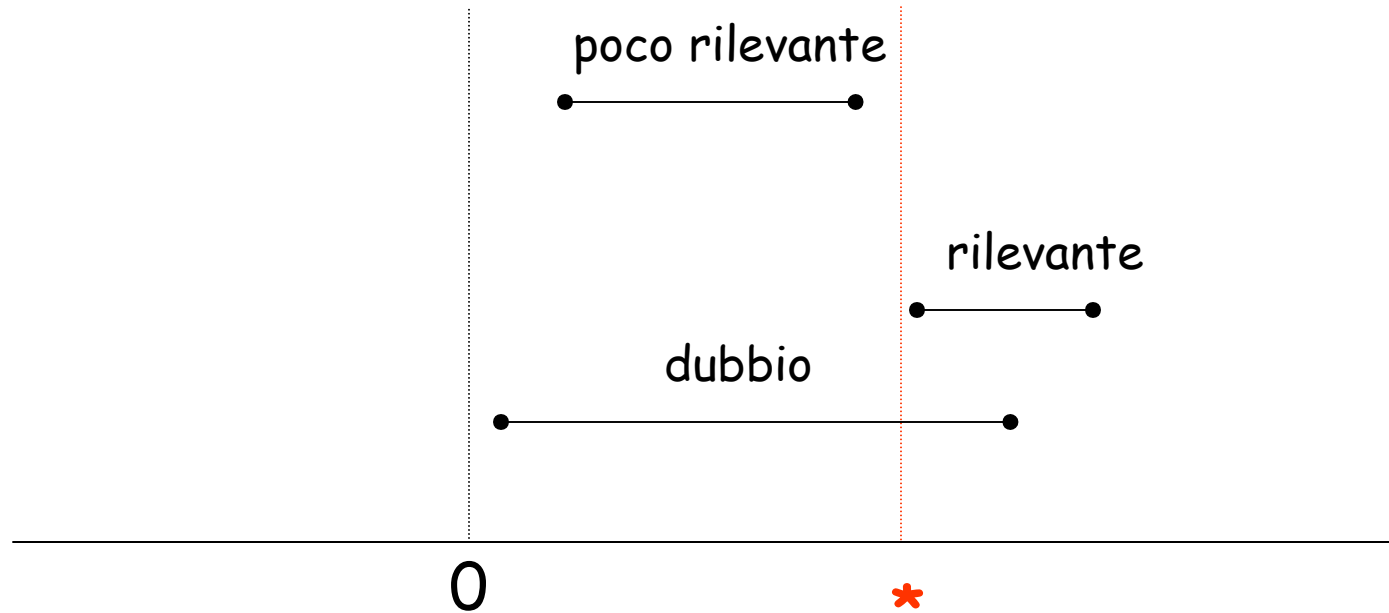
Gli intervalli di confidenza

- Maggiore è l'ampiezza dell' I.C. minore è la precisione della stima
- L'ampiezza dell' I.C., e quindi la precisione della stima, varia con la numerosità dello studio e il grado di confidenza desiderato
 - All'aumentare della numerosità l'ampiezza dell' I.C. diminuisce e la precisione aumenta
 - All'aumentare del grado di confidenza (es. 99% invece di 95%) l'ampiezza dell' I.C. aumenta e la precisione diminuisce

Intervalli di confidenza e significatività clinica

Intervalli di confidenza e significatività clinica

Studio di superiorità



- intervalli di confidenza dell'effetto osservato
- * minima differenza clinicamente rilevante

Le domande di efficacia: Non inferiorità

