

Power Analysis

Per stimare la numerosità del campione necessaria per riscontrare una differenza significativa tra due gruppi, occorrono i seguenti dati:

- l'*effect size*, cioè la differenza (standardizzata) tra le medie dei gruppi che si spera di rilevare;
- la *deviazione standard*. Di solito si usa lo stesso valore di σ per ciascun gruppo, ma se i valori sono significativamente diversi vanno utilizzati entrambi;
- α , cioè il *livello di significatività* (di solito 0.05);
- β (cioè la *potenza*), la probabilità di accettare l'ipotesi nulla quando è falsa (0.50, 0.80 e 0.90 sono i valori più comuni);
- il *rapporto* tra le dimensioni dei due gruppi. La massima potenza (statistica) si ha quando i due gruppi sono di uguale numerosità ($N_1/N_2 = 1$), ma questo non è ovviamente sempre possibile.

Per il calcolo della numerosità ottimale del campione in base alla potenza statistica richiesta occorre quindi fornire i suddetti parametri, che comunque corrispondono alla misurazione di una singola variabile (gene, nel nostro caso) in particolare per quanto attiene all'*effect size* ed alla *deviazione standard*.

Ciò naturalmente non è possibile ripeterlo per tutti i geni (alcune decine di migliaia!), e pertanto andrebbero preventivamente esaminati (almeno a campione per alcuni geni) i valori di espressione per determinarne l'espressione media μ su tutti i pazienti di ciascun gruppo, e la relativa deviazione standard σ .

Una volta ottenuta una stima di μ_1 , μ_2 , σ_1 e σ_2 si può determinare l'*effect size* e procedere all'analisi della potenza statistica per varie numerosità del campione.

A titolo di esempio allego due schermate ottenute utilizzando il programma G*Power (il riferimento accademico ufficiale per questo tipo di calcolo) in cui ho ipotizzato $\mu_1=100$, $\mu_2=103$, $\sigma_1=\sigma_2=5.26$, *effect size*=0.57, $\alpha=0.05$, $1-\beta=0.90$, $N_1/N_2 = 1$.

Nella seconda schermata si evince l'andamento del campione (*total sample size*), oscillante da 60 (data *power* = $1-\beta = 0.60$) a 160 (data *power* = $1-\beta = 0.95$).

