

SERVIZIO SANITARIO REGIONALE
EMILIA-ROMAGNA



DIPARTIMENTO DI SCIENZE MEDICHE VETERinarie

ECM in corso di accreditamento

Corso teorico-pratico

Corso di formazione per il personale abilitato in materia di IMPIEGO DEGLI ANIMALI AI FINI SCIENTIFICI ED EDUCATIVI – Livello 2

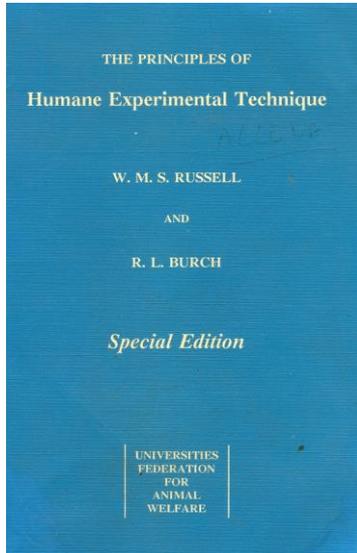
28-29-30 gennaio, 4-5-11-12 febbraio 2020

Dipartimento di Scienze Mediche Veterinaria dell'Alma Mater Studiorum - Università di Bologna, via Tolara di Sopra, 50, Ozzano dell'Emilia (BO)

Il disegno e il dimensionamento dello studio

F. Chiarotti, L. Ricceri
Istituto Superiore di Sanità

Modello delle “3R” (Russell e Burch, 1959)



- **Replacement** (Rimpiazzo, Sostituzione)
- **Reduction** (Riduzione)
- **Refinement** (Miglioramento, Raffinamento)

Uno strumento metodologico:

- forte validità generale normativa e teorica, come *forma mentis*;
- ha più di 60 anni: aspetti dinamici e di rinnovamento;
- British genius of compromise? (R&B si riferiscono con questo termine al Cruelty to animals Act del 1876)
- «*intimate relationship between humanity and efficiency in experimentation will recur constantly as a major theme in the present book*»

R di *Reduction*

Standard

Methods which minimise the number of animals used per experiment

Contemporary

Appropriately designed and analysed animal experiments that are robust and reproducible, and truly add to the knowledge base

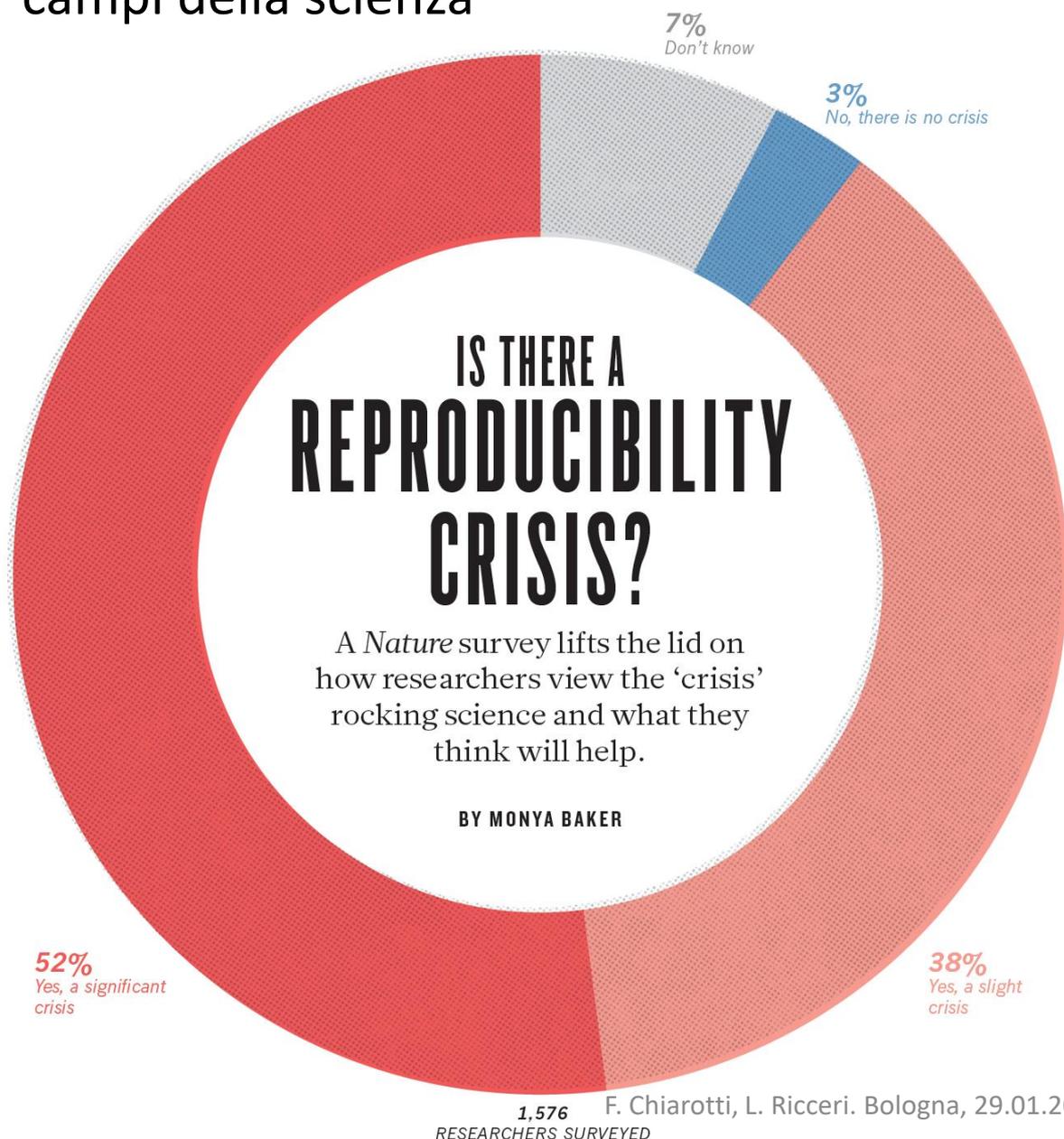
*... robust and reproducible ...
... truly add to the knowledge base*

→

mancanza di riproducibilità
dei risultati degli studi
sperimentali

<https://www.nc3rs.org.uk>

Molti scienziati segnalano la mancanza di riproducibilità in differenti campi della scienza



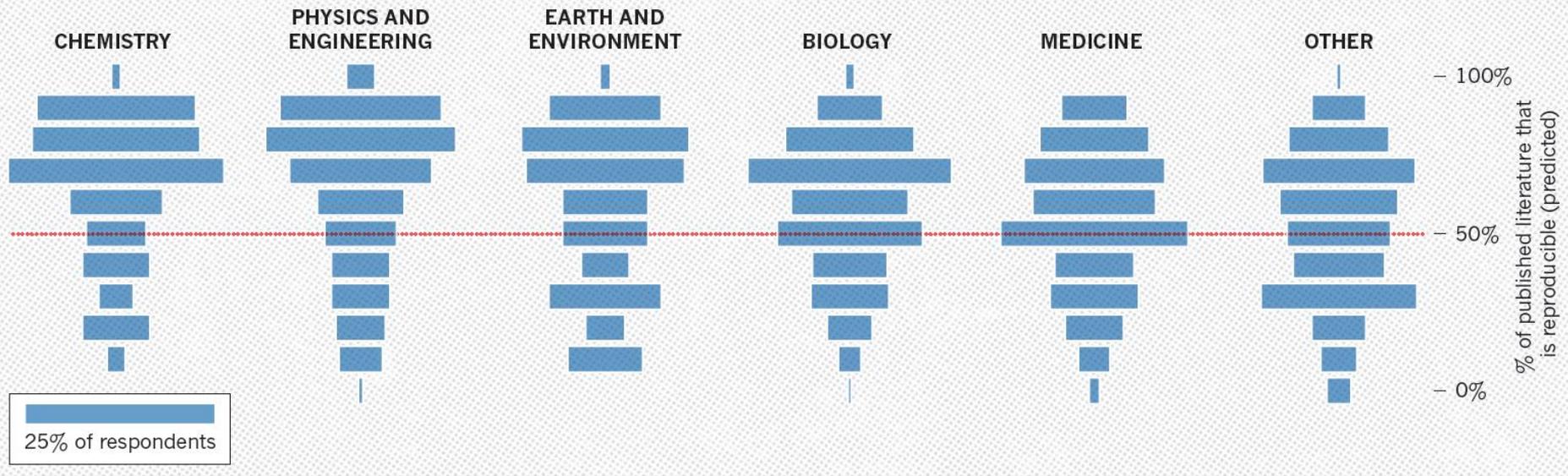
Baker, *Nature*, 2016

1576 ricercatori:

- Più del 70% dei ricercatori non sono riusciti a riprodurre i risultati di altri scienziati
- Più del 50% dei ricercatori non sono riusciti a riprodurre i propri risultati

HOW MUCH PUBLISHED WORK IN YOUR FIELD IS REPRODUCIBLE?

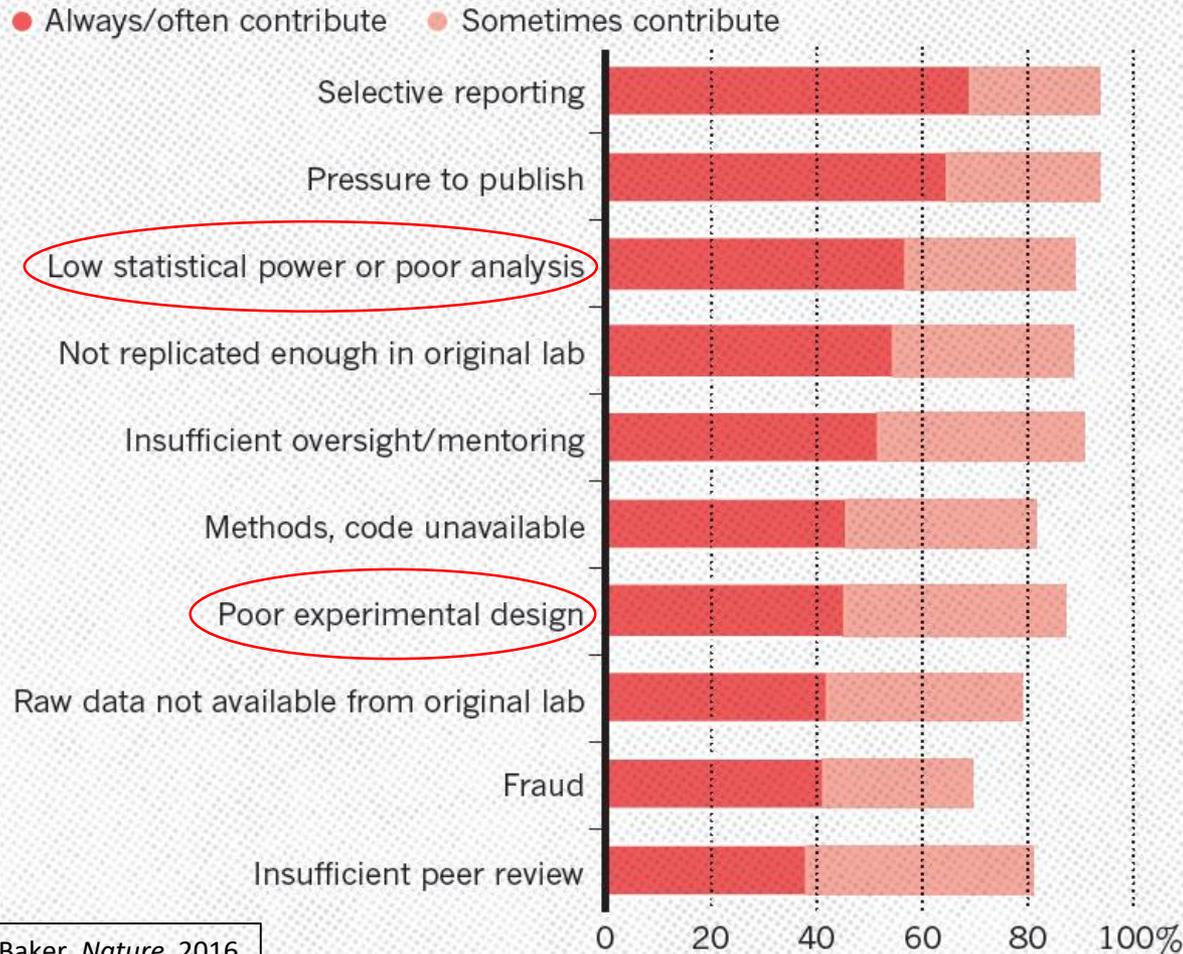
Physicists and chemists were most confident in the literature.



Tra le possibili motivazioni addotte per la mancanza di riproducibilità:
“To a large extent, this reproducibility crisis in basic and preclinical research may be as a result of failure to adhere to good scientific practice and the desperation to publish or perish.”

WHAT FACTORS CONTRIBUTE TO IRREPRODUCIBLE RESEARCH?

Many top-rated factors relate to intense competition and time pressure.



Possibili soluzioni?

“One of the best – publicized approaches to boosting reproducibility is **pre-registration**, where scientist **submit hypotheses and plans for data analysis to a third party before performing experiments**, to prevent cherry-picking statistically significant results later”

Argomenti da trattare nel protocollo sperimentale

Disegno sperimentale

- gruppi indipendenti
- misure ripetute

Caratteristiche delle unità statistiche

- ceppi degli animali
- demografia
- criteri di inclusione e di esclusione ...

Trattamenti

- tipo
- dose ...

Set di variabili da raccogliere sulle unità statistiche

- risposte (outcomes) primarie e secondarie (strumenti, test, apparati sperimentali)
- covariate (fattori di confondimento o di interazione)

Controllo per test multipli

- omics ?

Dimensione dell'effetto

- rilevanza rispetto alla variabilità attesa della risposta e all'impatto clinic/biologico

Livello di significatività e Potenza

- errore di I tipo vs errore di II tipo

Dimensione del campione

Disegno sperimentale (o dello studio)

Gruppi INDIPENDENTI (IND)	Gruppi APPAIATI (APP)
Caratteristiche	
unità statistiche tutte <u>indipendenti</u> <u>entro e tra</u> gruppi ciò che si osserva su una unità statistica non aiuta a prevedere ciò che si osserverà su un'altra unità statistica	unità statistiche <u>indipendenti</u> <u>entro</u> gruppo ma <u>appaiate</u> <u>tra</u> gruppi ciò che si osserva su una unità statistica aiuta a prevedere ciò che si osserverà sulla o sulle unità statistiche ad essa appaiate: <ul style="list-style-type: none">• misure sullo stesso soggetto in tempi diversi• misure sullo stesso soggetto in condizioni diverse• misure su soggetti appaiati naturalmente (es. consanguinei) o per una o più variabili selezionate (es. età, sesso, severità della malattia, luogo di residenza ...)
Vantaggi	
<ul style="list-style-type: none">• maggiore facilità di gestione	<ul style="list-style-type: none">• minore variabilità -> minore dimensione campionaria
Svantaggi	
<ul style="list-style-type: none">• maggiore variabilità -> maggiore dimensione campionaria	<ul style="list-style-type: none">• minore facilità di gestione: se si perde una unità si perde il blocco

Numero di gruppi a confronto

- **Due** gruppi
- **Tre o più** gruppi

DUE gruppi	TRE o PIU' gruppi
Vantaggi	
<ul style="list-style-type: none">• maggiore facilità di gestione del singolo esperimento	<ul style="list-style-type: none">• minore dimensione campionaria• possibilità di confrontare i gruppi sperimentali fra loro
Svantaggi	
<ul style="list-style-type: none">• maggior numero di esperimenti -> replicazione del gruppo di controllo• maggiore dimensione campionaria	<ul style="list-style-type: none">• maggiore difficoltà di gestione dell'esperimento• correzione per confronti multipli

Numero di fattori in esame

- **Un** fattore
- **Due o più** fattori (disegno fattoriale)

UN fattore	DUE o PIU' fattori
Vantaggi	
<ul style="list-style-type: none">• maggiore facilità di gestione del singolo esperimento	<ul style="list-style-type: none">• minore dimensione campionaria• possibilità di valutare l'interazione tra fattori
Svantaggi	
<ul style="list-style-type: none">• maggior numero di esperimenti• minore generalizzabilità del risultato• maggiore dimensione campionaria	<ul style="list-style-type: none">• maggiore difficoltà di gestione dell'esperimento• correzione per confronti multipli più complessa

Test statistici

I **test statistici inferenziali** servono per **verificare ipotesi** relative a fenomeni **nella popolazione utilizzando dati sperimentali ottenuti su un campione**

Affinché i risultati sperimentali possano essere estesi alla popolazione, il **campione** deve essere **rappresentativo** della popolazione (estratto casualmente dalla popolazione)

Il test di ipotesi è costruito **non per confermare l'ipotesi** in esame, **ma** se possibile **per confutarla**

Dobbiamo quindi stabilire **2 ipotesi**

- l'ipotesi da sottoporre al test = **ipotesi nulla H_0**
- l'ipotesi di interesse per lo sperimentatore = **ipotesi alternativa H_1**

Il test di ipotesi è costruito per rifiutare l'ipotesi in esame, non per confermarla, quindi il risultato del test è

- si può rifiutare H_0
- non si può rifiutare H_0 (H_0 non si accetta mai)

Rifiutando H_0 si conferma implicitamente (mai in maniera esplicita) H_1

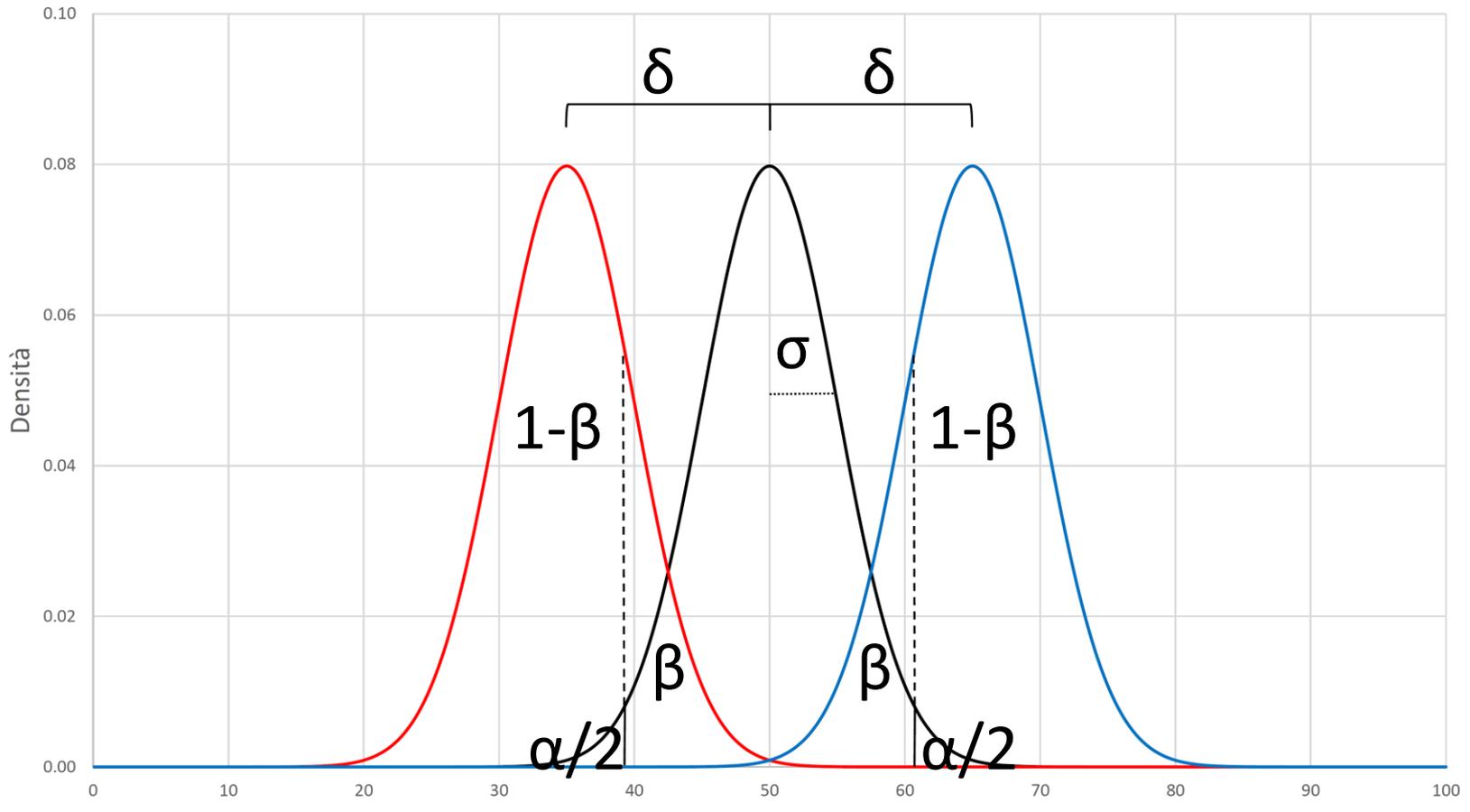
Con il test statistico si stima la probabilità di avere un valore della statistica del test pari a quello ottenuto nel campione, o uno ancora più "distante" dall'ipotesi nulla, nel caso in cui H_0 sia vera: se questa probabilità è bassa si rifiuta H_0 , altrimenti non si può rifiutare H_0 .

RISULTATO DEL TEST	REALTÀ	
	H_0 vera	H_0 falsa
Non rifiuto H_0	$1-\alpha$	β
Rifiuto H_0	α	$1-\beta$

Test di ipotesi

- **Errore di I tipo**: probabilità indicata con α
- **Errore di II tipo**: probabilità indicata con β
- Potenza del test = $1-\beta$

Funzione di densità



H1

H0

H1

$\alpha=0.05$ a 2 code (0.025 per coda)

$\beta=0.18$; $1-\beta=0.82$

Test per la verifica di ipotesi statistiche per variabili quantitative

- Test **parametrici**

si basano su assunti (da verificare) relativi a:

- distribuzione dei dati (Normale)
- omogeneità delle varianze

- Test **non parametrici**

si basano sull'assunto di continuità della variabile risposta

Metodi **parametrici** vs metodi **non parametrici**

Metodi PARAMETRICI	Metodi NON PARAMETRICI
Vantaggi	
<ul style="list-style-type: none">• minore dimensione campionaria	<ul style="list-style-type: none">• assunti (condizioni di applicabilità) meno stringenti
Svantaggi	
<ul style="list-style-type: none">• assunti (condizioni di applicabilità) più stringenti e di difficile verifica	<ul style="list-style-type: none">• maggiore dimensione campionaria

Variabili quantitative, 1 fattore, 2 gruppi

I più comuni test statistici per il confronto fra **2 gruppi** relativamente a **variabili quantitative** sono:

- **Test parametrici** (distr Normale, σ omogenei)
 - IND** → t di Student per gruppi indipendenti
 - APP** → t di Student per gruppi appaiati
- **Test non parametrici** (distr non specificata)
 - IND** → U di Mann-Whitney
 - APP** → Wilcoxon matched-pair test

Variabili quantitative, 1 fattore, 3 o + gruppi

I più comuni test statistici per il confronto fra 3 o + gruppi con 1 solo fattore relativamente a variabili quantitative sono:

- Test parametrici (distr Normale, σ omogenei)

 - IND → ANOVA per gruppi indipendenti

 - APP → ANOVA per gruppi appaiati

- Test non parametrici (distr non specificata)

 - IND → ANOVA di Kruskal-Wallis

 - APP → ANOVA di Friedman

Variabili quantitative, 2 o + fattori

I più comuni test statistici per il confronto fra 2 o + gruppi con 2 o + fattori relativamente a variabili quantitative sono:

- Test parametrici (distr Normale, σ omogenei)
 - IND → ANOVA fattoriale per disegni completamente randomizzati
 - APP → ANOVA fattoriale per disegni a blocchi randomizzati
 - IND+APP → ANOVA fattoriale per disegni misti (split-plot)

Variabili qualitative, 1 fattore, 2 o + livelli

I più comuni test statistici per il confronto fra 2 o + gruppi indipendenti relativamente a variabili qualitative (categoriche o quantitative categorizzate in 2 o + livelli) sono:

- Test del Chi quadrato

(frequenze attese ≥ 5 in $\geq 80\%$ delle celle)

- Test della probabilità esatta di Fisher

Argomenti da trattare nel protocollo sperimentale

Disegno sperimentale

- gruppi indipendenti
- misure ripetute

Caratteristiche delle unità statistiche

- ceppi degli animali
- demografia
- criteri di inclusione e di esclusione ...

Trattamenti

- tipo
- dose ...

Set di variabili da raccogliere sulle unità statistiche

- risposte (outcomes) primarie e secondarie (strumenti, test, apparati sperimentali)
- covariate (fattori di confondimento o di interazione)

Controllo per test multipli

- omics ?

Dimensione dell'effetto

- rilevanza rispetto alla variabilità attesa della risposta e all'impatto clinic/biologico

Livello di significatività e Potenza

- errore di I tipo vs errore di II tipo

Dimensione del campione

Ma questi sono proprio gli elementi che la legge prevede vengano indicati nell'Allegato VI dal presentatore di una richiesta di autorizzazione

In particolare, per quanto riguarda la dimensione campionaria:

ARTICOLO 31, punto 4**ALLEGATO VI**

i) minor numero di animali per il raggiungimento delle finalità del progetto

**Sezione 10 Numero animali
Sezione 11 Numero animali
Sezione 20 b) Riduzione
Sezione 21.2 Cons. Statistiche
Sez. 32 Riutilizzo**

10. Si prevede di utilizzare il seguente numero complessivo di animali (specificare se animali geneticamente modificati)

Specie Animale	Ceppo	Numero / anno

11. Specificare se Animali Geneticamente modificati ed eventuali numeri parziali riferiti a diverse specie e/o ceppi animali

Specie Animale	Ceppo	Numero / anno

ARTICOLO 31, punto 4	ALLEGATO VI
i) minor numero di animali per il raggiungimento delle finalità del progetto	Sezione 10 Numero animali Sezione 11 Numero animali Sezione 20 b) Riduzione Sezione 21.2 Cons. Statistiche Sez. 32 Riutilizzo

20. Dichiarazioni riferite all'Articolo 13, comma 2.

- a. **Riduzione** (Massima riduzione del numero di animali utilizzati, compatibile con gli obiettivi del progetto di ricerca)

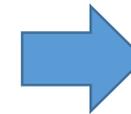
21. Metodologia e tecnica dell'esperimento

21.2 Considerazioni statistiche (*descrivere come è stato determinato il numero di animali necessari per lo studio*)

- Studio di dimensionamento. Come si è giunti all' "N" dei singoli esperimenti.

NB "20 a" non concide con "21.2"

- Allegato VI 15. Durata del progetto di ricerca [**massimo (non obbligatoriamente)** 60 mesi]
Pianificare 5 anni di esperimenti può essere complesso: 3 + 2? Spesso i numeri possono diventare più realistici (si eliminano opzioni – e.g. modelli, trattamenti, dosi – che si sono rivelate non utili)
- Il revisore fa “anche” **i conti** dei gruppi sperimentali per ogni esperimento previsto con relativa numerosità: se il disegno sperimentale non è chiaro, valutare la congruità della richiesta non è possibile
- Il disegno sperimentale ha anche a che vedere con la **R di Riduzione del numero di soggetti** (e.g. campioni appaiati quando possibile invece di campioni indipendenti)



Quando
possibile
TABELLE

Prima di condurre l'esperimento, si può determinare:

- la dimensione campionaria n , data la dimensione di effetto che si vuole essere in grado di evidenziare e la potenza desiderata consente di valutare la dimensione campionaria necessaria a rilevare effetti di dimensione desiderata (perché ritenuta clinicamente rilevante) con una potenza data
- la potenza, data la dimensione campionaria n che si può ragionevolmente gestire nell'esperimento e la dimensione di effetto che si vuole essere in grado di evidenziare consente di valutare se lo studio, "vincolato" a un numero di unità prestabilito, abbia una potenza sufficiente a valutare un effetto di dimensione data
- la dimensione di effetto, data la dimensione campionaria n e la potenza desiderata consente di stimare la dimensione di effetto che si è in grado di evidenziare con una potenza data in uno studio, "vincolato" a un numero di unità prestabilito

Tutte e tre le opzioni sono utili, perché aiutano a ragionare sull'esperimento in termini sia di potenziale impatto sulle conoscenze scientifiche sia di fattibilità

Per motivi etici occorre infatti **evitare di sovradimensionare gli esperimenti**, ma è altrettanto importante **non sottodimensionarli**, perché i risultati non sarebbero conclusivi e affidabili, e gli animali risulterebbero essere stati sacrificati inutilmente

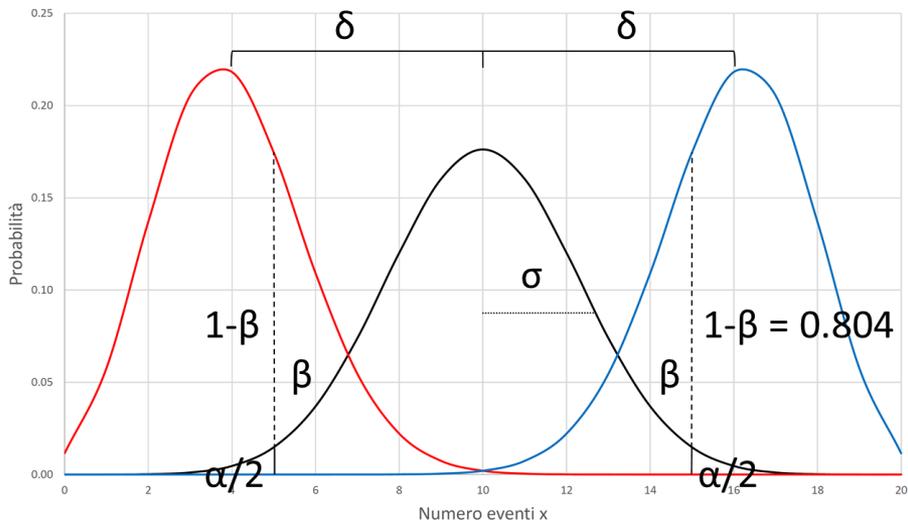
Bisogna perciò prestare molta attenzione alla **dimensione di effetto** che si vuole essere in grado di evidenziare, che deve essere clinicamente rilevante ma anche realisticamente raggiungibile, e alla **potenza** che NON PUO' e NON DEVE scendere al di sotto dell'80%

Dimensionamento (o studio di potenza)

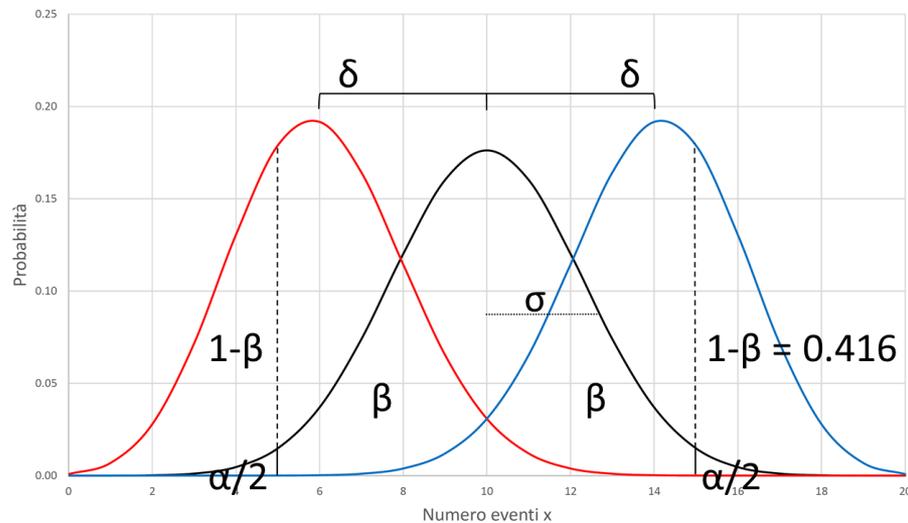
Gli elementi da tenere in conto, e che sono interdipendenti, sono:

- 1) α = il **livello di significatività** considerato accettabile per rifiutare H_0 (a 1 o 2 code) **valore prestabilito**
- 2) $1-\beta$ = la **potenza del test**, cioè la probabilità di evidenziare una differenza significativa **valore prestabilito**
- 3) σ = la **variabilità delle misure** che si confrontano (per variabili quantitative)
- 4) δ = la **minima differenza** che si vuole essere in grado di evidenziare tra i gruppi a confronto, perché **clinicamente/biologicamente rilevante**
- 5) n = la **dimensione campionaria** per ciascun gruppo

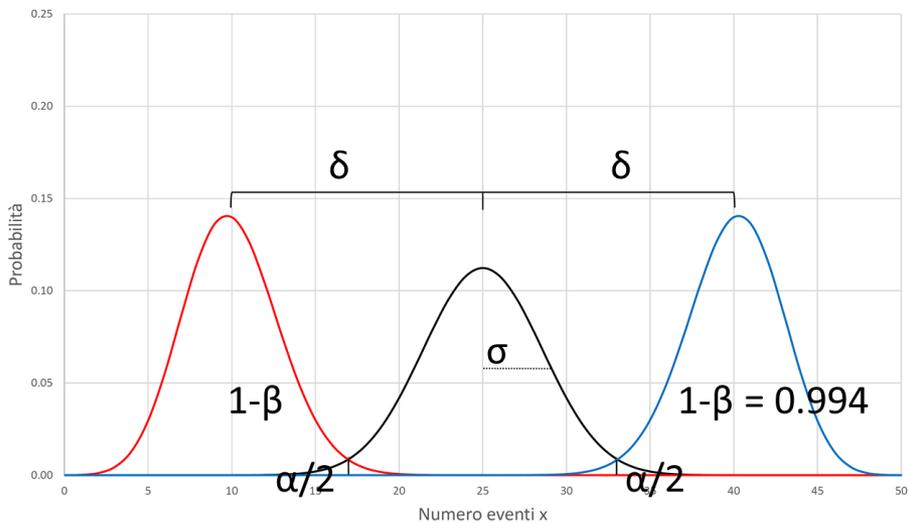
Distribuzione di probabilità $\alpha=0.05$
 $p_0=0.5$, $p_1=0.2$ o $p_1=0.8$; numero prove $n=20$



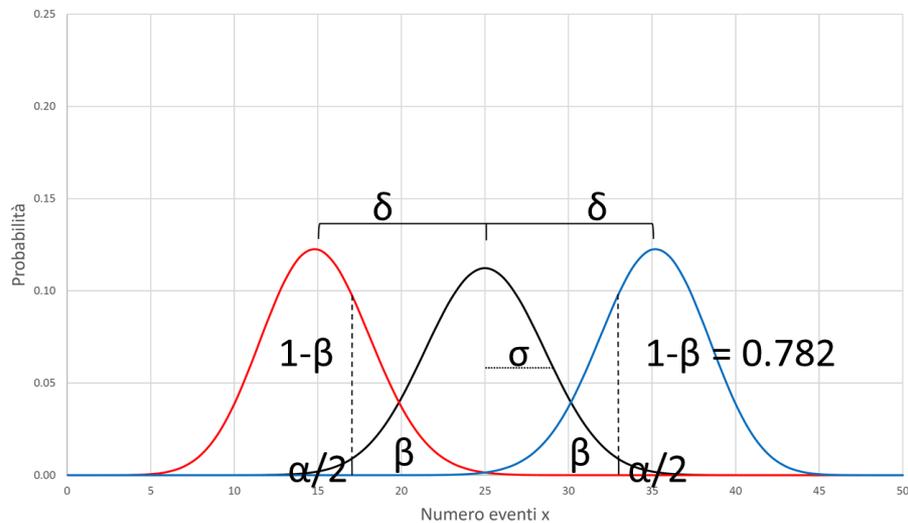
Distribuzione di probabilità $\alpha=0.05$
 $p_0=0.5$, $p_1=0.3$ o $p_1=0.7$; numero prove $n=20$



Distribuzione di probabilità $\alpha=0.05$
 $p_0=0.5$, $p_1=0.2$ o $p_1=0.8$; numero prove $n=50$



Distribuzione di probabilità $\alpha=0.05$
 $p_0=0.5$, $p_1=0.3$ o $p_1=0.7$; numero prove $n=50$



Da dove possiamo ricavare i dati necessari al dimensionamento?

α , $1-\beta$: valori standard

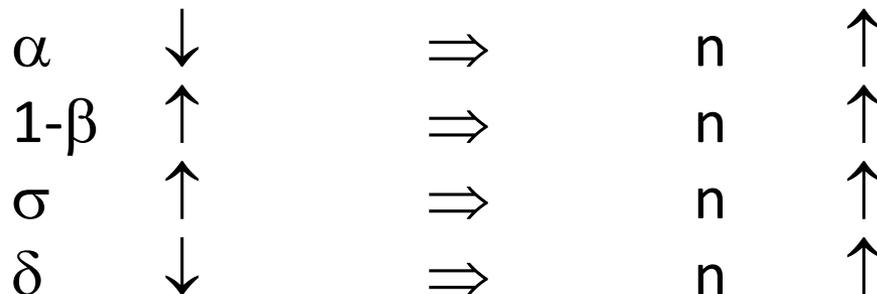
$\alpha \leq 0.05$ (**0.05**, 0.01; usualmente a 2 code)

$1-\beta \geq 0.80$ (**0.80**, 0.85, 0.90)

σ : dalla letteratura, da studi precedenti del proponente o da appositi studi (studio pilota o di fattibilità)

δ : sulla base di ipotesi di lavoro (cliniche, di rilevanza biologica)

NB: a parità di altre condizioni, la dimensione campionaria è legata a α , $1-\beta$, σ , δ



Studio pilota o di fattibilità

Utile quando non vi è sufficiente esperienza sugli effetti del trattamento (soprattutto sulla variabilità)

Obiettivi:

- selezionare la variabile risposta appropriata tra le potenziali
- determinare la precisione (variabilità) della variabile risposta per il dimensionamento
- valutare le possibili fonti di errore (bias)

Caratteristiche:

- lo studio non deve essere troppo ampio
- appropriato per analisi dati esplorative
- inappropriato per variabili risposta misurabili a lungo termine

MEMO: maggiore è la variabilità della risposta, maggiore è la dimensione campionaria necessaria

La **riduzione della variabilità** nei gruppi si può ottenere:

- maggiore omogeneità delle unità statistiche
 - restrizione delle unità basata su variabili biologiche (es: sesso, fascia d'età, ...)
 - uso di ceppi inbred (sperimentazione animale)
 - Pro:* maggiore omogeneità
 - Cons:* minore generalizzabilità
- controllo ripetibilità delle misure
 - replicabilità inter- ed intra-saggio
 - concordanza tra operatori
 - centralizzazione delle misure
 - controllo errori di trascrizione/immissione dati

Punto cruciale: scelta della dimensione dell'effetto δ

δ cosa rappresenta?

SI

minima differenza che si vuole essere in grado di evidenziare tra i gruppi a confronto, perché «**cl clinicamente**» rilevante

minima differenza che si vorrebbe osservare per ritenere che il fattore in studio produca una variazione «cl clinicamente» rilevante sulla risposta esaminata

NO

differenza che ci si aspetta di osservare tra i gruppi a confronto

(sebbene la dimensione δ dell'effetto debba essere comunque tale che si possa ragionevolmente osservare, cioè non ci si deve porre come obiettivo un effetto che non sia verosimilmente raggiungibile)

Come si può scegliere δ ?

In assenza di ipotesi specifiche, possiamo scegliere δ in funzione di σ . Per esempio, per la differenza fra medie Cohen* propone come misura *relativa* della dimensione dell'effetto $d = \delta/SD$

effetto **piccolo**: $d = 0.2 \Rightarrow \delta = 0.2 \times SD$ attesa

effetto **medio**: $d = 0.5 \Rightarrow \delta = 0.5 \times SD$ attesa

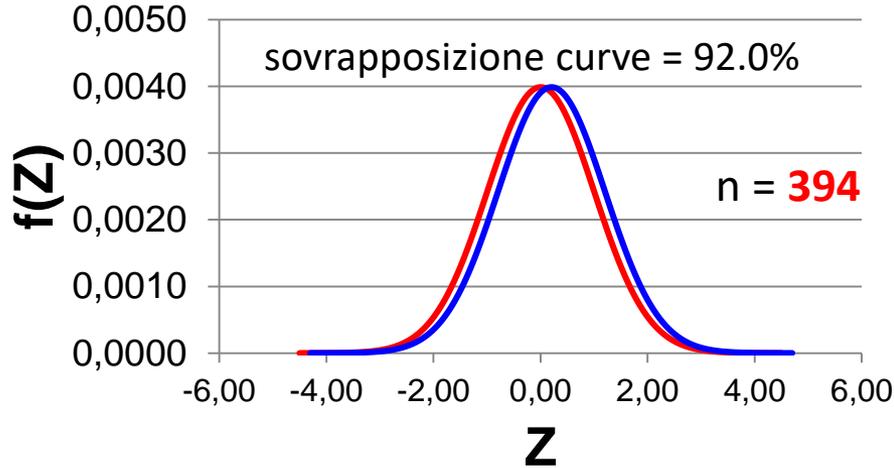
effetto **grande**: $d = 0.8 \Rightarrow \delta = 0.8 \times SD$ attesa

NB: minore è la dimensione dell'effetto che si vuole essere in grado di individuare, maggiore è la dimensione campionaria necessaria

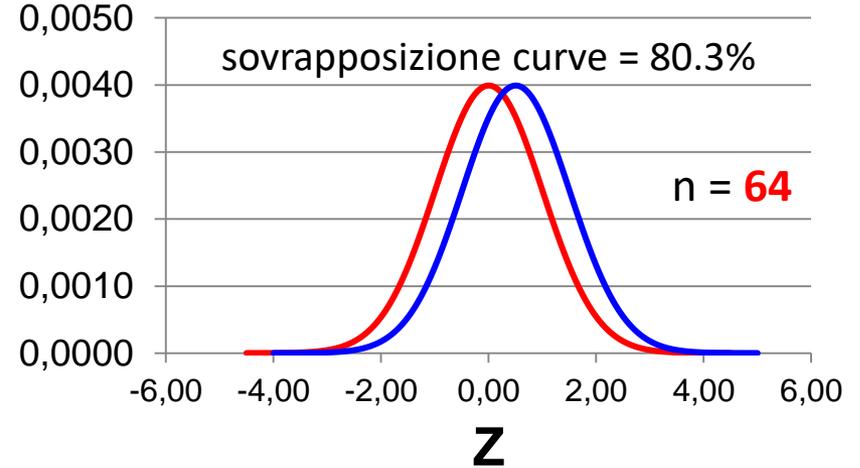
* J. Cohen. A power primer. Psychol Bull 1992 Jul;112(1):155-9

D. Wahlsten. Chapter 5 - Sample Size. In: WAHLSTEN, D. (ed.) *Mouse Behavioral Testing*. London: Academic Press, 2011 => Cohen's d : effetto piccolo = 0.5, effetto medio = 1.0, effetto grande = 1.5.

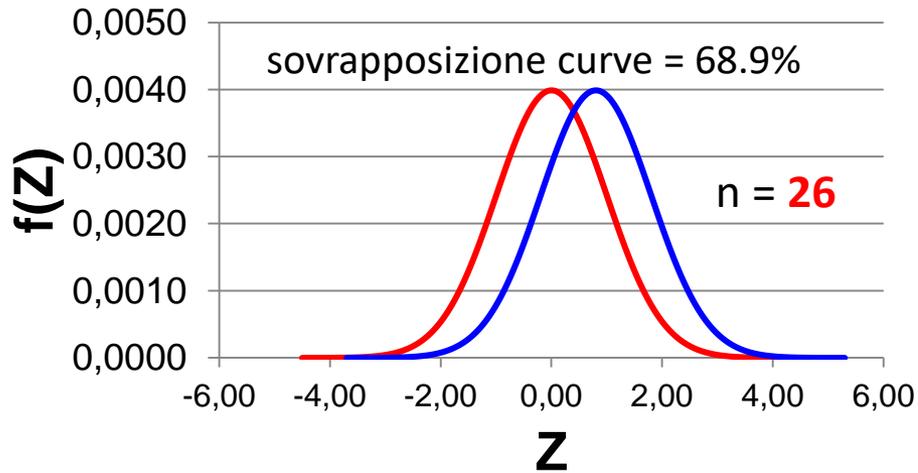
Z(0.0;1) vs Z(0.2;1) - effetto piccolo



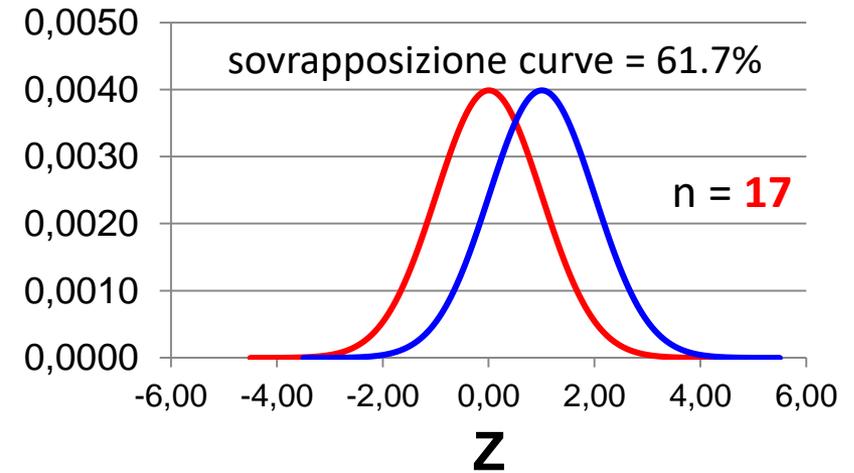
Z(0.0;1) vs Z(0.5;1) - effetto medio



Z(0.0;1) vs Z(0.8;1) - effetto grande



Z(0.0;1) vs Z(1.0;1) - effetto grande



Famiglia del test	Indice di dimensione di effetto		Effetto PICCOLO	Effetto MEDIO	Effetto GRANDE
Differenza fra 2 medie	Cohen d	$\frac{\bar{x}_1 - \bar{x}_2}{SD_{pooled}}$	0.20 0.50	0.50 1.00	0.80 1.50
Differenza fra 2 proporzioni	Cohen h	$2 * (\arcsin\sqrt{p_1} - \arcsin\sqrt{p_2})$	0.20	0.50	0.80
Differenza fra ≥ 3 medie	Cohen f	$f^2 = \frac{\eta^2}{1-\eta^2} = \frac{\omega^2}{1-\omega^2} = \frac{R^2}{1-R^2}$	0.10 0.25	0.25 0.50	0.40 0.75
Differenza fra ≥ 3 medie	η^2	$\frac{SS_{treatment}}{SS_{total}}$			
Differenza fra ≥ 3 medie	$\eta_{Partial}^2$	$\frac{SS_{treatment}}{SS_{treatment} + SS_{error}}$			
Differenza fra ≥ 3 medie	ω^2	$\frac{SS_{treatment} - df_{treatment} * MS_{error}}{SS_{total} + MS_{error}}$			

Famiglia del test	Indice di dimensione di effetto		Effetto PICCOLO	Effetto MEDIO	Effetto GRANDE
Coefficiente di determinazione	R^2 <i>ESS = Effect SS</i> <i>RSS = Residual SS</i> <i>TSS = Total SS</i>	$\frac{ESS}{TSS} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$ $\frac{1 - RSS}{TSS} = \frac{1 - \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$	Valore basso <0.30	Valore medio 0.30 † 0.60	Valore buono ≥0.60
Correlazione	r di Pearson		0.10 0.20	0.30 0.40	0.50 0.60

Una dimensione di effetto che può essere rilevante in uno studio può essere irrilevante in uno studio che abbia finalità diverse

Per esempio, un valore **d di Cohen = 0.20** potrebbe essere:

- **del tutto irrilevante** in uno studio preclinico che punti a testare in un modello animale molti trattamenti per individuare quelli potenzialmente efficaci da sperimentare successivamente sull'uomo
- **rilevante** in uno studio clinico che punti a individuare fattori prognostici da utilizzarsi in uno screening, quando la patologia da prevenire sia grave e l'intervento per prevenirla non abbia effetti collaterali tali da sconsigliarla in pazienti «falsi positivi»

La **scelta** della **dimensione di effetto** e dei **gruppi sperimentali** da includere nello studio è **cruciale** per il dimensionamento dello studio

Il revisore difficilmente può interferire in tale scelta

Il revisore può solo segnalare un dubbio sulla effettiva rilevanza della dimensione di effetto stabilita e sull'opportunità dell'inclusione dei vari gruppi sperimentali, la cui **scelta** rimane comunque **sotto la responsabilità del proponente** del protocollo sperimentale

Che la scelta della dimensione di effetto e dei gruppi sperimentali (per esempio, delle sostanze da testare e relative dosi) sia cruciale e non sempre appropriata lo si può desumere anche dalle valutazioni retrospettive

Tra le motivazioni più frequenti dell'impiego di un minor numero di animali rispetto a quanto preventivato vi sono:

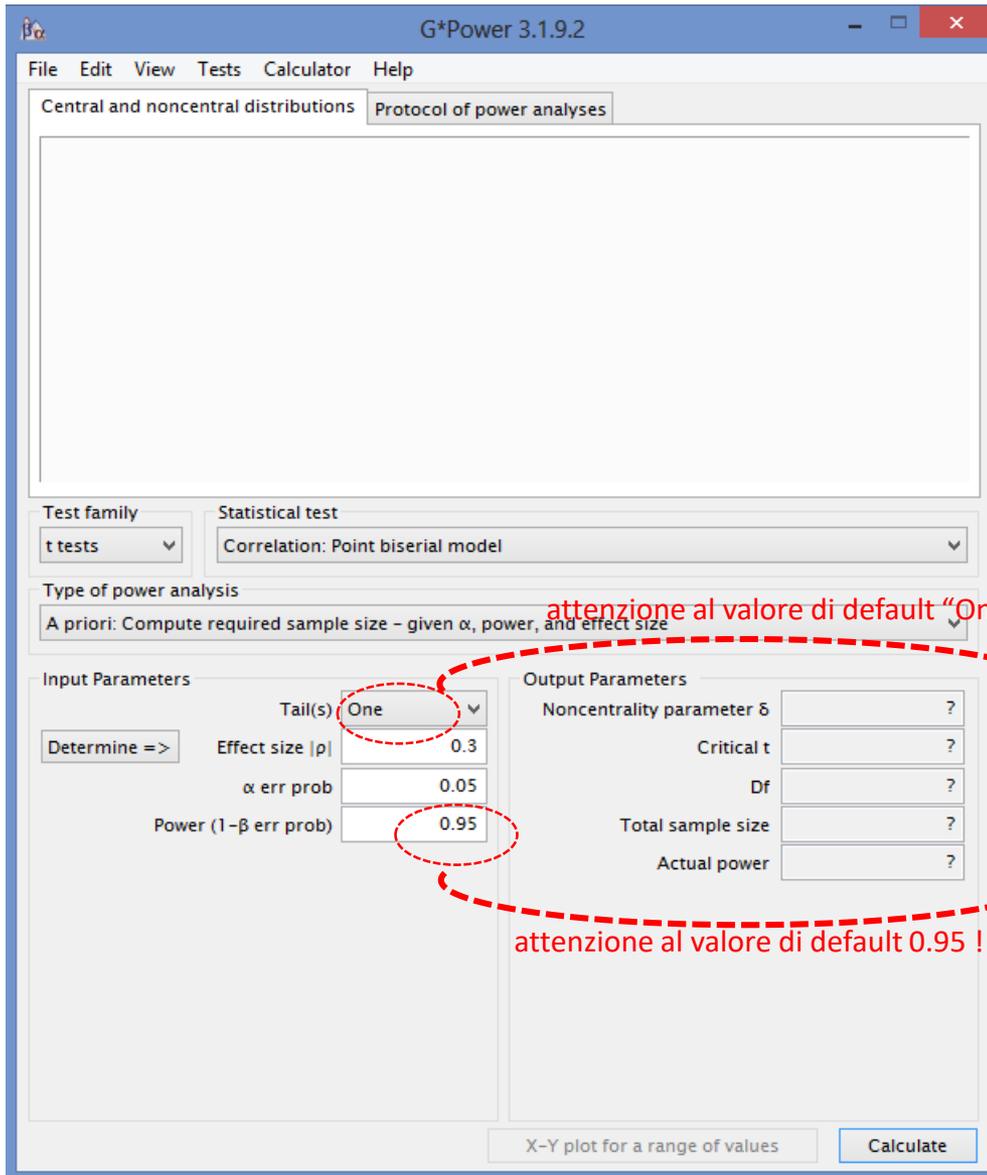
1. I risultati degli studi in vitro o degli studi preliminari, che comportano la riduzione o la cancellazione in toto della fase sperimentale successiva ⇒ **importanza di «spezzare» il protocollo**
2. La riduzione o modificazione dei gruppi sperimentali o l'individuazione di effetti di dimensione maggiore rispetto al preventivato, che risultano significativi anche con un numero minore di unità sperimentali ⇒ **importanza di scegliere razionalmente il disegno (gruppi, variabili, ...) e la dimensione di effetto rilevante**

G*Power

software gratuito per:

- il dimensionamento del campione (dato δ , $1-\beta$ e α)
- l'analisi di potenza (dato n , δ e α)
- la valutazione della dimensione dell'effetto (dato n , $1-\beta$ e α)

<http://www.psych.uni-duesseldorf.de/abteilungen/aap/gpower3/>



G*Power

Schermata di apertura

Bisogna scegliere:

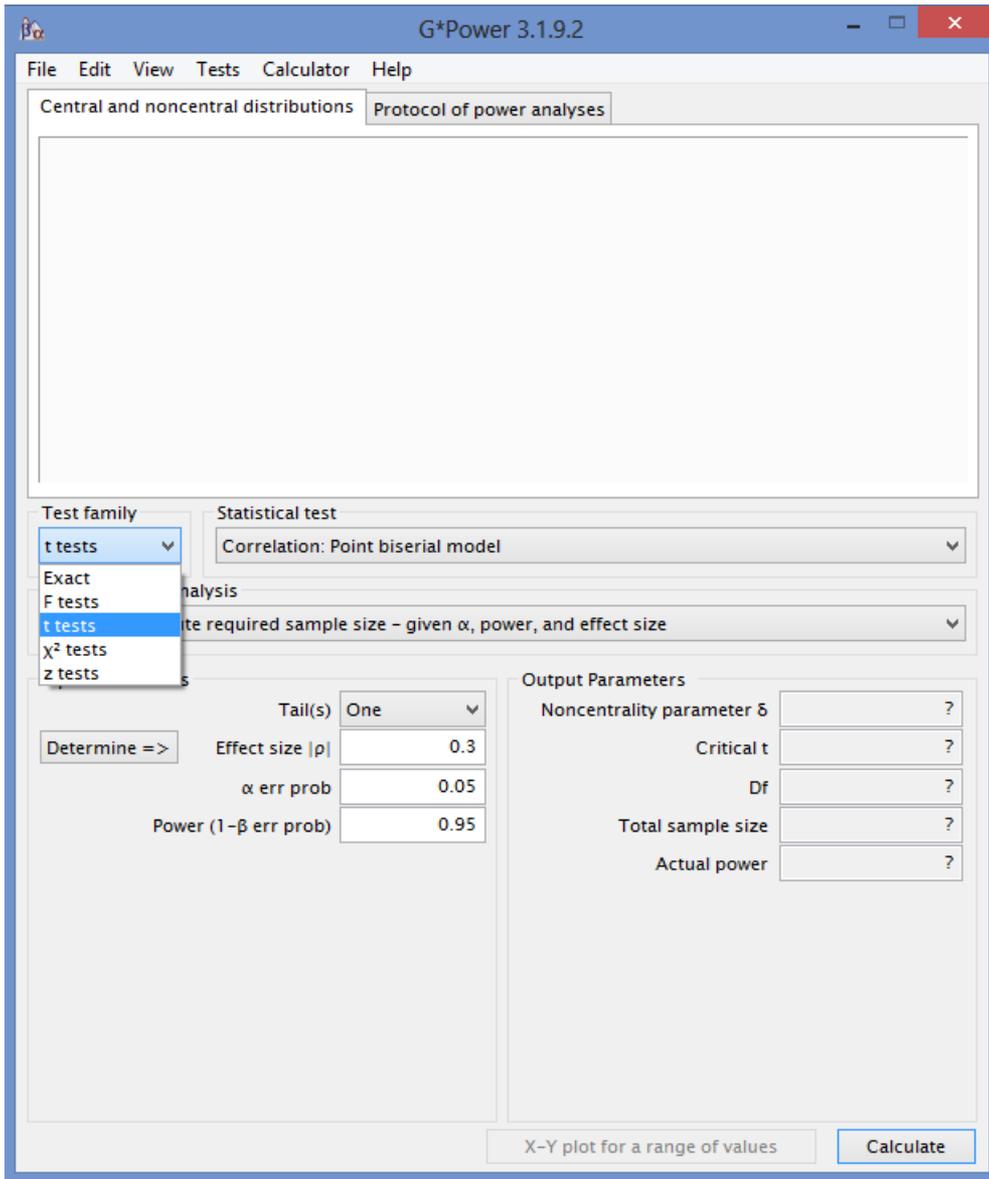
- Il tipo di presentazione
- La famiglia dei test
- Il test statistico dentro la famiglia
- Il tipo di analisi di potenza

Bisogna inserire:

- Code
- Dimensione dell'effetto
- Probabilità dell'errore di I tipo
- Potenza del test
- Rapporto tra le dimensioni dei campioni

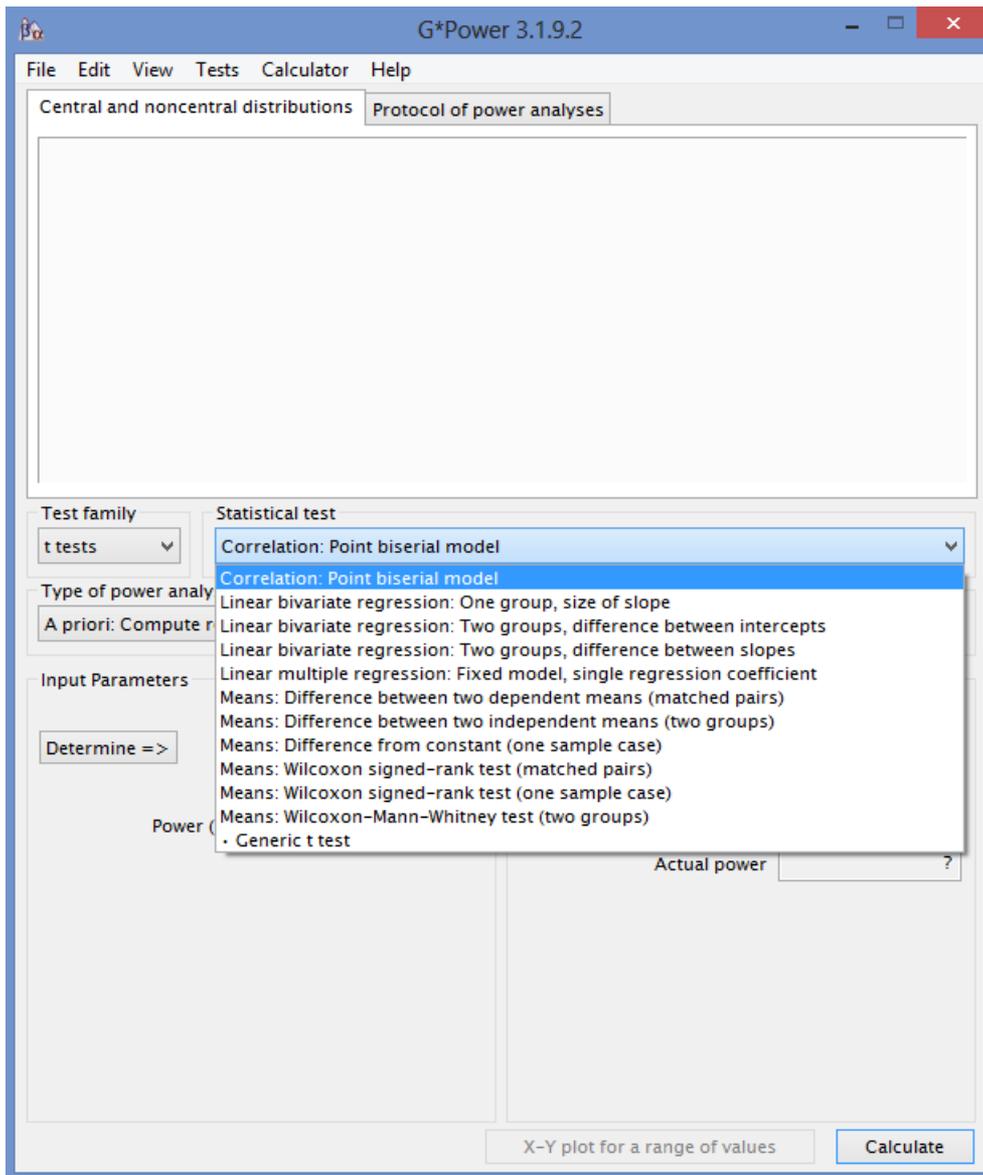
attenzione al valore di default "One" !

attenzione al valore di default 0.95 !



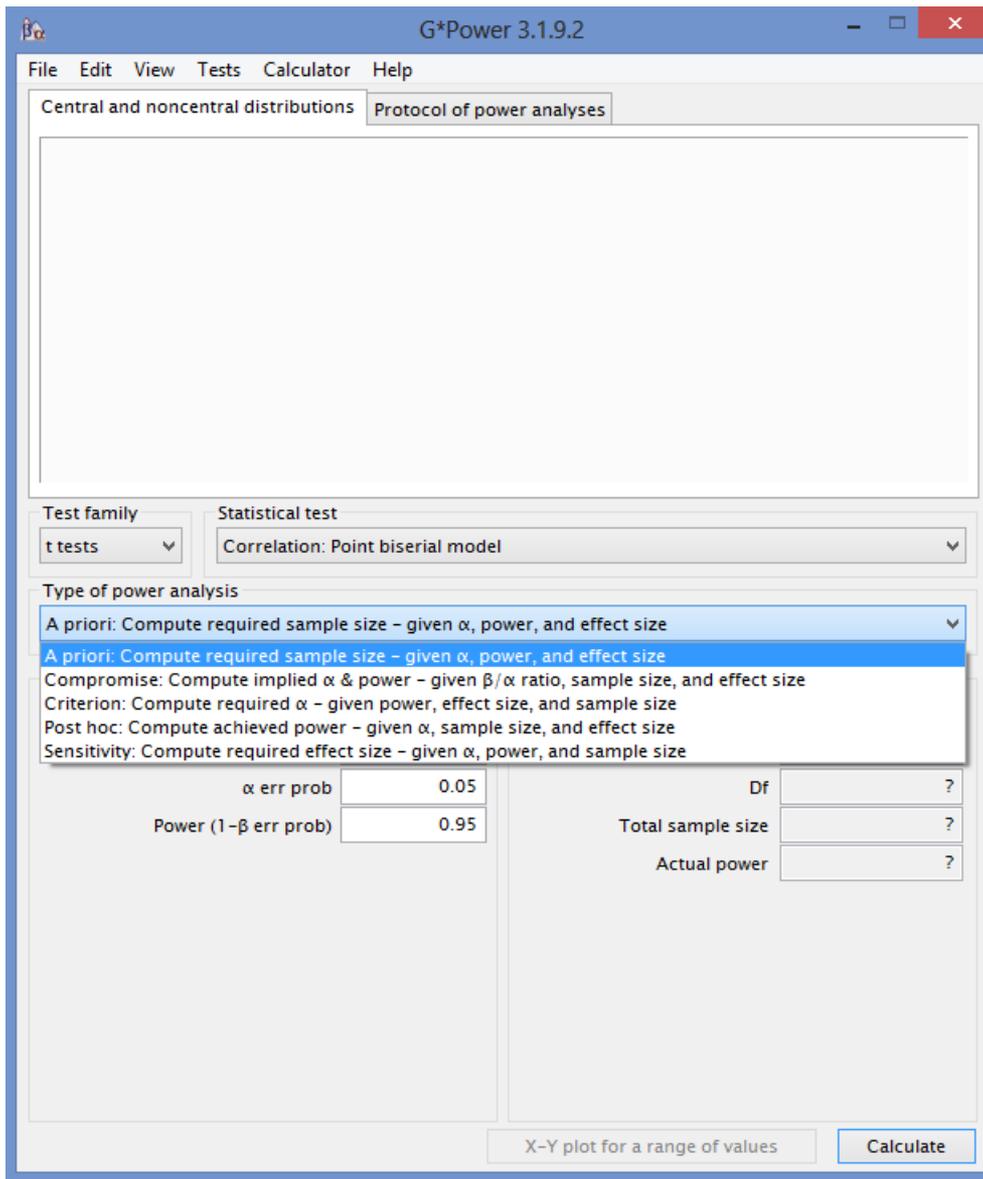
G*Power

Scegliamo la famiglia di test



G*Power

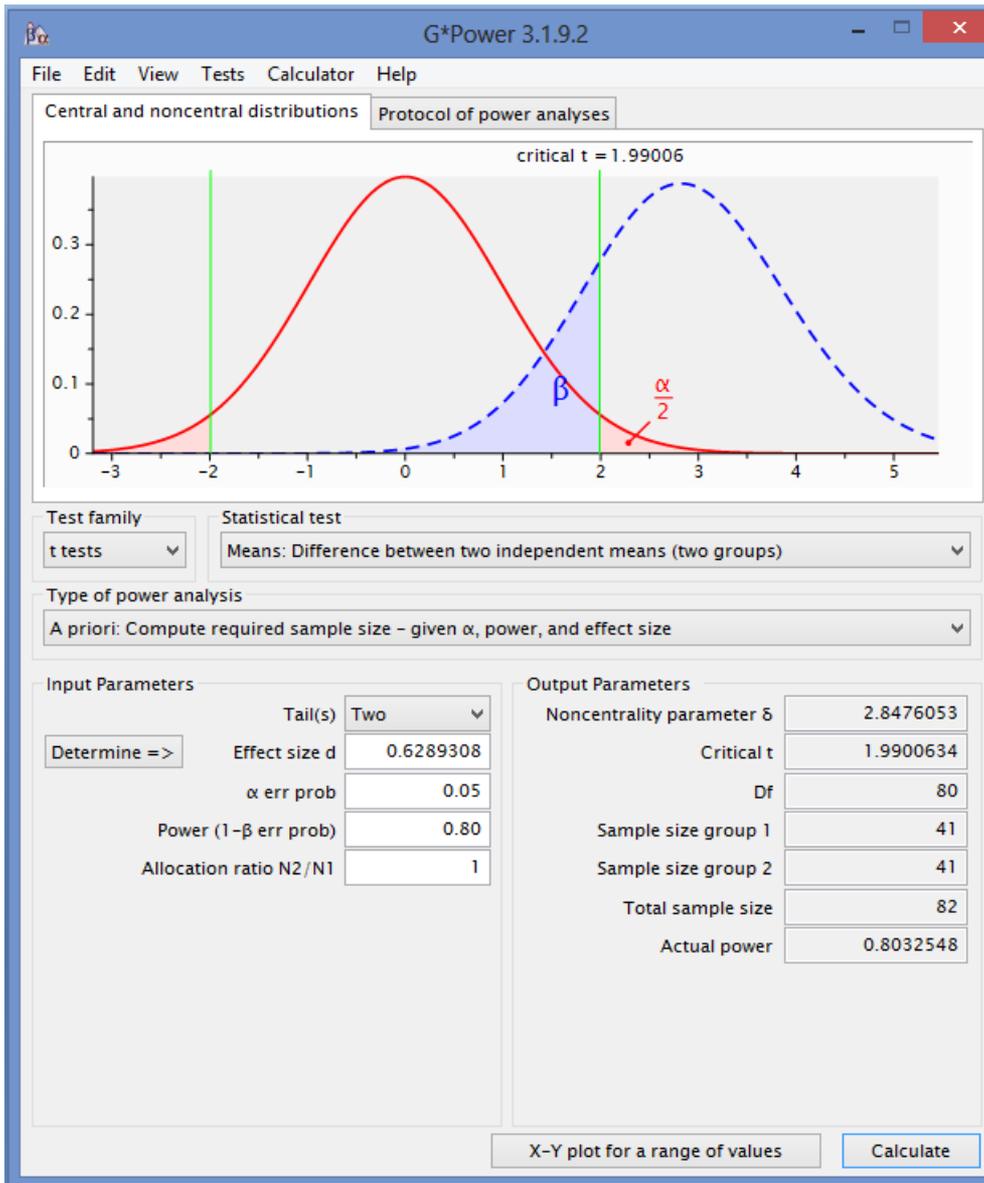
Scegliamo il test dentro la famiglia



G*Power

Scegliamo il tipo di power analysis

NB: siamo interessati a calcolare la dimensione campionaria



G*Power

t test per gruppi indipendenti

Inseriamo i dati richiesti per il calcolo della dimensione campionaria

Incremento del 12.5% (da 80 a 90)

Effect size $d = 0.63$ (effetto medio)

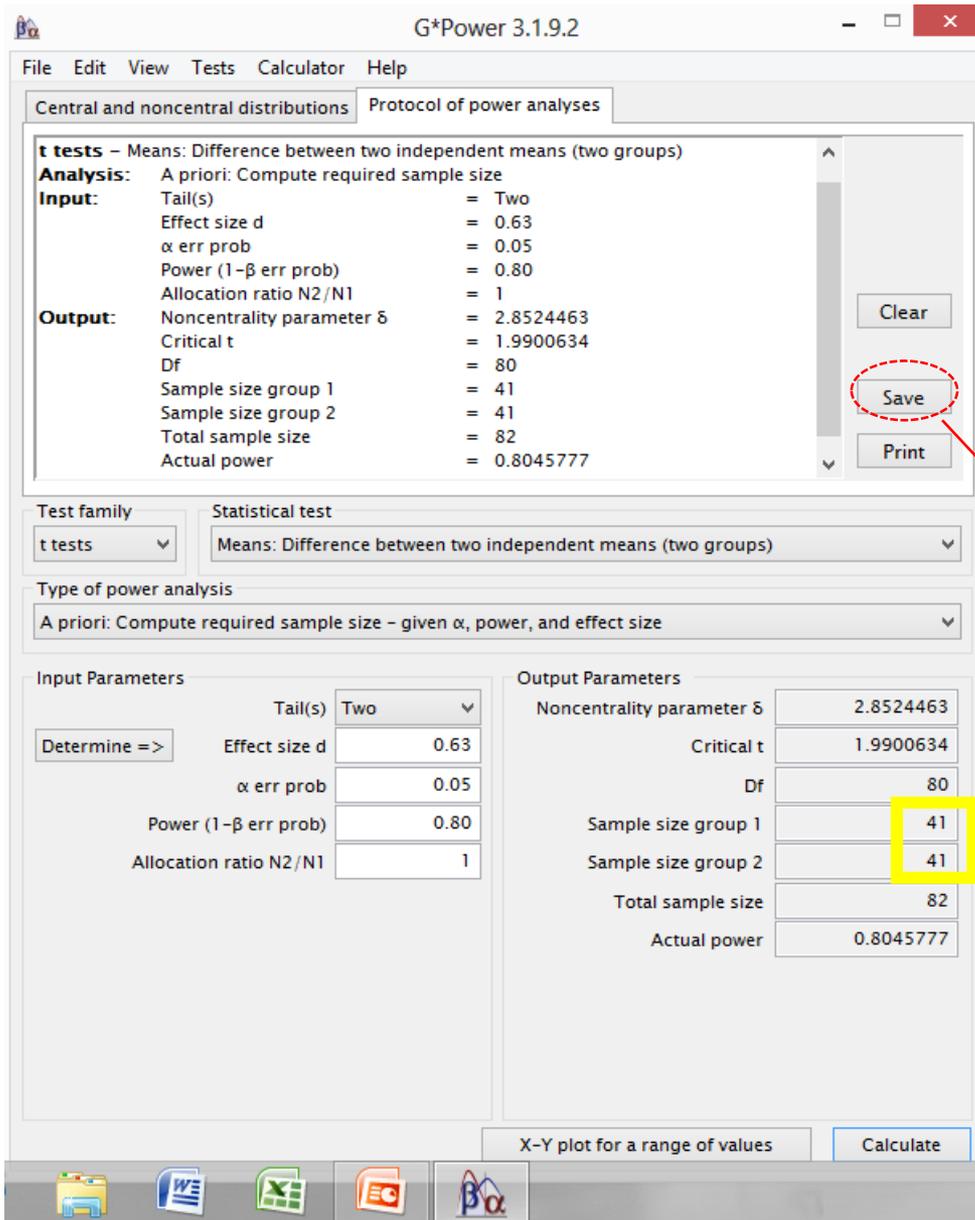
$n1 \neq n2$
 Mean group 1: 0
 Mean group 2: 1
 SD σ within each group: 0.5

$n1 = n2$
 Mean group 1: 80
 Mean group 2: 90
 SD σ group 1: 15.9
 SD σ group 2: 15.9

Calculate Effect size d: 0.6289308

Calculate and transfer to main window

Close



G*Power

t test per gruppi indipendenti

Inseriamo i dati richiesti per il calcolo della dimensione campionaria

Effect size $d = 0.63$ (effetto medio)

Si può salvare il protocollo, per conservarlo come documentazione

G*Power 3.1.9.2

File Edit View Tests Calculator Help

Central and noncentral distributions Protocol of power analyses

t tests – Means: Difference between two independent means (two groups)

Analysis: A priori: Compute required sample size

Input:

Tail(s)	=	Two
Effect size d	=	1.0062893
α err prob	=	0.05
Power (1- β err prob)	=	0.80
Allocation ratio N2/N1	=	1

Output:

Noncentrality parameter δ	=	2.9338123
Critical t	=	2.0369333
Df	=	32
Sample size group 1	=	17
Sample size group 2	=	17
Total sample size	=	34
Actual power	=	0.8118655

Clear Save Print

Test family: t tests

Statistical test: Means: Difference between two independent means (two groups)

Type of power analysis: A priori: Compute required sample size – given α , power, and effect size

Input Parameters

Tail(s)	Two	
Determine =>	Effect size d	1.0062893
	α err prob	0.05
	Power (1- β err prob)	0.80
	Allocation ratio N2/N1	1

Output Parameters

Noncentrality parameter δ	2.9338123
Critical t	2.0369333
Df	32
Sample size group 1	17
Sample size group 2	17
Total sample size	34
Actual power	0.8118655

X-Y plot for a range of values Calculate

G*Power

t test per gruppi indipendenti

Inseriamo i dati richiesti per il calcolo della dimensione campionaria

Incremento del 20% (da 80 a 96)

Effect size $d = 1.0$ (effetto grande)

$n1 \neq n2$

Mean group 1: 0

Mean group 2: 1

SD σ within each group: 0.5

$n1 = n2$

Mean group 1: 80

Mean group 2: 96

SD σ group 1: 15.9

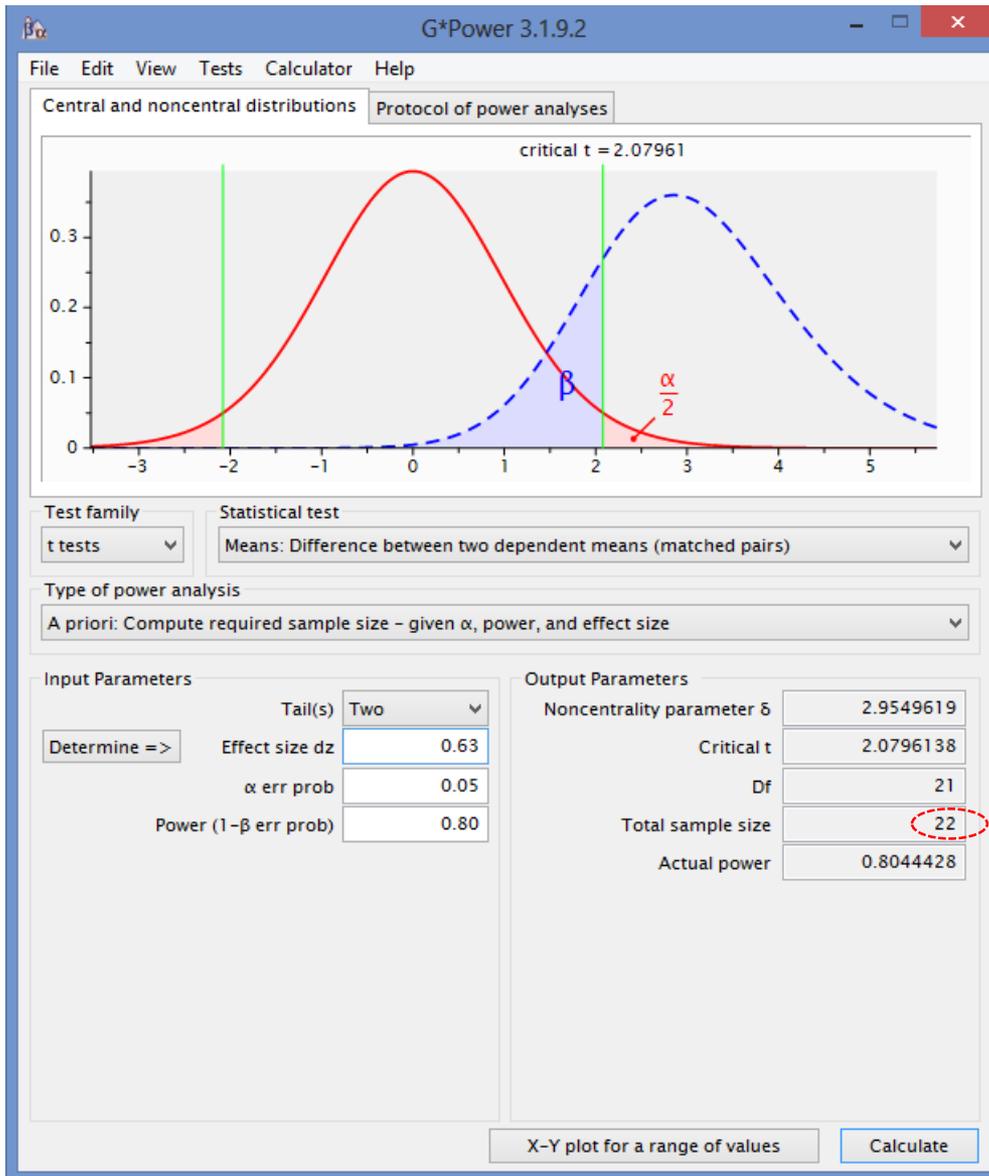
SD σ group 2: 15.9

Calculate Effect size d: 1.006289

Calculate and transfer to main window

Close

La dimensione campionaria passa da 41 a 17 unità per gruppo sperimentale



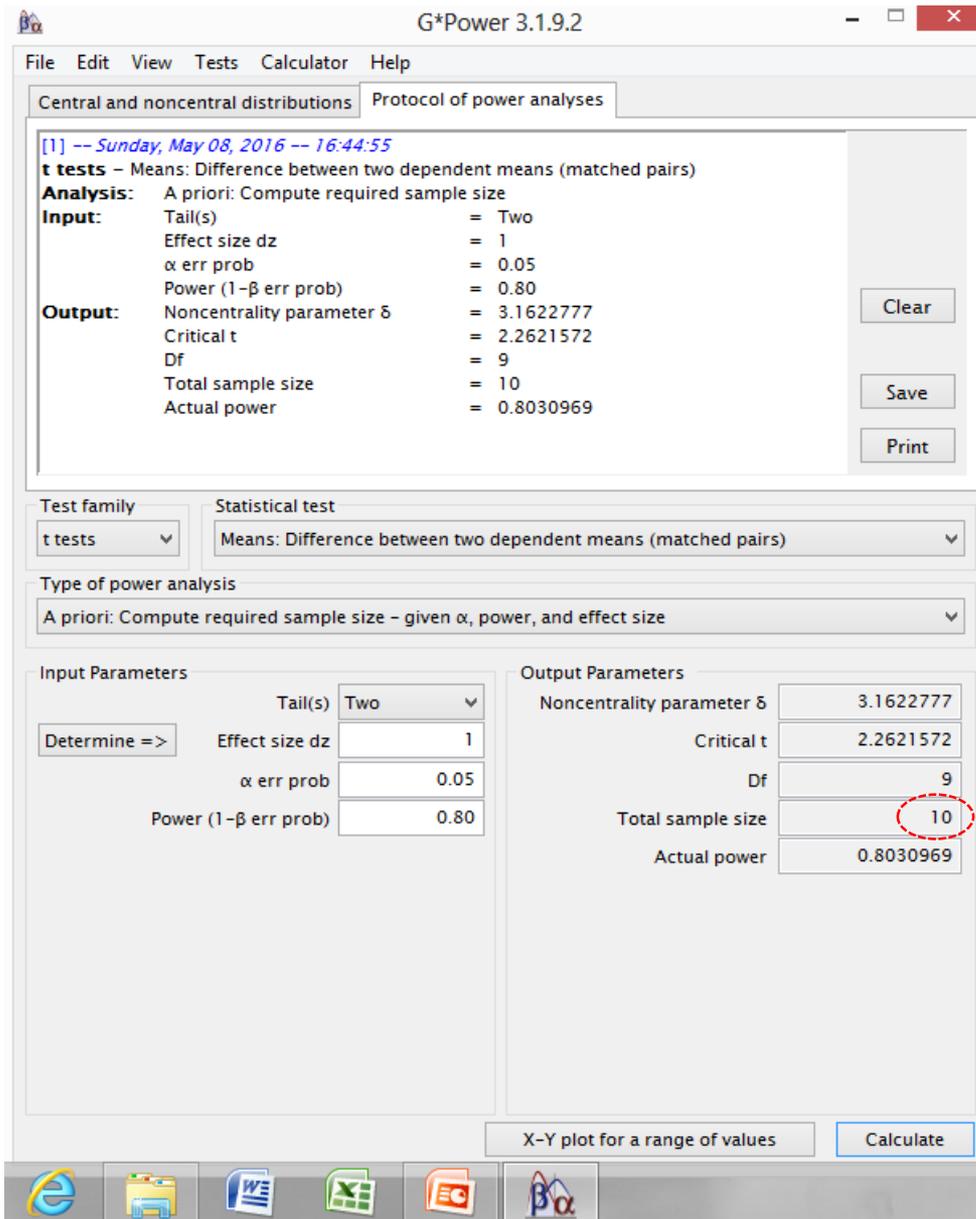
G*Power

t test per gruppi **appaiati**

Inseriamo i dati richiesti per il calcolo della dimensione campionaria

Effect size $d = 0.63$ (effetto medio)

La dimensione campionaria passa da 41 a 22 unità per gruppo sperimentale



G*Power

t test per gruppi **appaiati**

Inseriamo i dati richiesti per il calcolo della dimensione campionaria

Effect size $d = 1.0$ (effetto grande)

La dimensione campionaria passa da 22 a 10 unità per gruppo sperimentale

G*Power 3.1.9.2

File Edit View Tests Calculator Help

Central and noncentral distributions Protocol of power analyses

[1] -- Saturday, October 07, 2017 -- 13:47:09

F tests – ANOVA: Fixed effects, omnibus, one-way

Analysis: A priori: Compute required sample size

Input:

Effect size f	= 0.5
α err prob	= 0.05
Power (1-β err prob)	= 0.80
Number of groups	= 4

Output:

Noncentrality parameter λ	= 12.0000000
Critical F	= 2.8164658
Numerator df	= 3
Denominator df	= 44
Total sample size	= 48
Actual power	= 0.8029537

Clear Save Print

Test family: F tests Statistical test: ANOVA: Fixed effects, omnibus, one-way

Type of power analysis: A priori: Compute required sample size – given α, power, and effect size

Input Parameters: Determine => Effect size f: 0.5, α err prob: 0.05, Power (1-β err prob): 0.80, Number of groups: 4

Output Parameters: Noncentrality parameter λ: 12.0000000, Critical F: 2.8164658, Numerator df: 3, Denominator df: 44, Total sample size: 48, Actual power: 0.8029537

X-Y plot for a range of values Calculate

G*Power

ANOVA a un fattore, 4 gruppi **indipendenti** (1 controllo e 3 gruppi sperimentali)

Inseriamo i dati richiesti per il calcolo della dimensione campionaria

Effect size $f = 0.5$ (effetto grande, equivalente al valore 1.0 del caso di 2 gruppi indipendenti a confronto)

La dimensione campionaria totale scende da $(17+17)+(17+17)+(17+17) = 102$ a 48 unità statistiche

	t Student gruppi indipendenti		t Student gruppi appaiati		ANOVA gruppi indipendenti
Effetto medio $\delta = 0.63$					
A vs B	n = 41	N = 82	n = 22	N = 44	
Effetto grande $\delta = 1.00$					
A vs B	n = 17	N = 34	n = 10	N = 20	
Effetto grande $\delta = 1.00$ (f = 0.50)					
A vs B	n = 17	N = 34			A vs B vs C vs D
A vs C	n = 17	N = 34			n = 12 N = 48
A vs D	n = 17	N = 34			A vs B; A vs C; A vs D
		N = 102			n = 23 N = 92

Per tornare alla «crisi di riproducibilità» degli studi, la comunità scientifica sente il bisogno di una maggiore consapevolezza nella pianificazione degli esperimenti e di una maggiore trasparenza nel riportare il disegno dello studio e i relativi risultati

Figure 3. One of these labels is not mandatory

	<h3>Research Facts</h3> <table border="1"> <tr> <th colspan="2">Sample</th> </tr> <tr> <td>Size</td> <td>n=50</td> </tr> <tr> <td>Set</td> <td>in advance</td> </tr> <tr> <td>Dropped measures</td> <td>0%</td> </tr> <tr> <td>Dropped conditions</td> <td>0%</td> </tr> <tr> <td>Dropped observations</td> <td>2%</td> </tr> </table>	Sample		Size	n=50	Set	in advance	Dropped measures	0%	Dropped conditions	0%	Dropped observations	2%		<h3>Nutrition Facts</h3> <p>Serving Size: 13 rings (28g)</p> <table border="1"> <thead> <tr> <th colspan="2">Amount Per Serving</th> <th></th> </tr> <tr> <th colspan="2">Calories 140</th> <th>Calories from Fat 60</th> </tr> <tr> <th colspan="3">% Daily Value*</th> </tr> </thead> <tbody> <tr> <td>Total Fat</td> <td>7 g</td> <td>11%</td> </tr> <tr> <td>Saturated Fat</td> <td>1 g</td> <td>5%</td> </tr> <tr> <td>Trans Fat</td> <td>0 g</td> <td></td> </tr> <tr> <td>Cholesterol</td> <td>0 mg</td> <td>0%</td> </tr> <tr> <td>Sodium</td> <td>270 mg</td> <td>11%</td> </tr> <tr> <td>Potassium</td> <td></td> <td></td> </tr> <tr> <td>Total Carbohydrate</td> <td>18 g</td> <td>6%</td> </tr> <tr> <td>Dietary Fiber</td> <td>1 g</td> <td>4%</td> </tr> <tr> <td>Sugars</td> <td>1 g</td> <td></td> </tr> <tr> <td>Sugar Alcohols</td> <td></td> <td></td> </tr> <tr> <td>Protein</td> <td>2 g</td> <td></td> </tr> <tr> <td>Vitamin A</td> <td>0 IU</td> <td>0%</td> </tr> <tr> <td>Vitamin C</td> <td>0 mg</td> <td>0%</td> </tr> <tr> <td>Calcium</td> <td>0 mg</td> <td>0%</td> </tr> <tr> <td>Iron</td> <td>1.08 mg</td> <td>6%</td> </tr> </tbody> </table>	Amount Per Serving			Calories 140		Calories from Fat 60	% Daily Value*			Total Fat	7 g	11%	Saturated Fat	1 g	5%	Trans Fat	0 g		Cholesterol	0 mg	0%	Sodium	270 mg	11%	Potassium			Total Carbohydrate	18 g	6%	Dietary Fiber	1 g	4%	Sugars	1 g		Sugar Alcohols			Protein	2 g		Vitamin A	0 IU	0%	Vitamin C	0 mg	0%	Calcium	0 mg	0%	Iron	1.08 mg	6%
Sample																																																																					
Size	n=50																																																																				
Set	in advance																																																																				
Dropped measures	0%																																																																				
Dropped conditions	0%																																																																				
Dropped observations	2%																																																																				
Amount Per Serving																																																																					
Calories 140		Calories from Fat 60																																																																			
% Daily Value*																																																																					
Total Fat	7 g	11%																																																																			
Saturated Fat	1 g	5%																																																																			
Trans Fat	0 g																																																																				
Cholesterol	0 mg	0%																																																																			
Sodium	270 mg	11%																																																																			
Potassium																																																																					
Total Carbohydrate	18 g	6%																																																																			
Dietary Fiber	1 g	4%																																																																			
Sugars	1 g																																																																				
Sugar Alcohols																																																																					
Protein	2 g																																																																				
Vitamin A	0 IU	0%																																																																			
Vitamin C	0 mg	0%																																																																			
Calcium	0 mg	0%																																																																			
Iron	1.08 mg	6%																																																																			

“We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.”

From: Joe Simmons, Leif Nelson and Uri Simonsohn. A 21 Word Solution. In: Society for Personality and Social Psychology (SPSP) Dialogue Newsletter, vol 26 n.2 (Fall 2012)

Grazie per l'attenzione