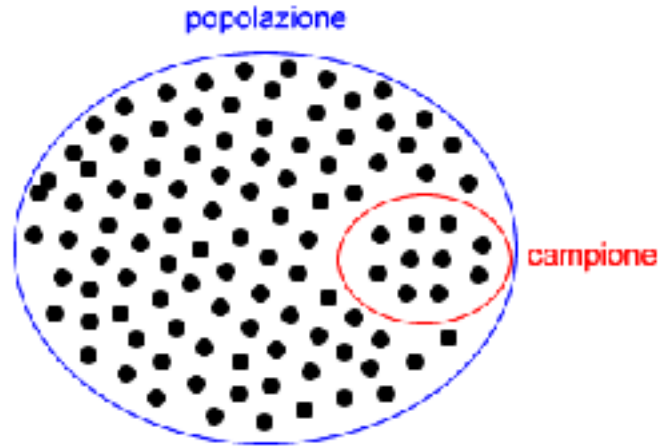


Dimensione campionaria



Nella ricerca clinica, la definizione della dimensione campionaria per uno studio clinico è uno degli aspetti più importanti per il successo dello studio stesso.

La dimensione del campione deve essere calcolata correttamente durante la fase di progettazione dello studio clinico, perché è una fase fondamentale da diversi punti di vista, tra cui quello metodologico, etico e finanziario.

Dimensione campionaria

La dimensione campionaria è il numero di soggetti da includere in uno studio clinico che consentirà di rilevare un effetto del trattamento clinicamente rilevante (cioè un risultato del trattamento che generalmente i medici considerano importante).

Poiché non è possibile reclutare l'intera popolazione, affetta, ad esempio, da una determinata malattia o che deve essere trattata con il trattamento sperimentale, sarà arruolato nello studio un sottogruppo di individui che rappresenti la popolazione complessiva.

Per calcolare la dimensione del campione più appropriata da arruolare in uno studio clinico è di fondamentale importanza la collaborazione tra biostatistici e clinici è importante.

Dimensione campionaria

In linea di massima si può affermare che il principio su cui si fonda il calcolo della dimensione campionaria è quello di rendere massima la probabilità di individuare la presenza di un effetto (protettivo o nocivo) di un trattamento e/o di un intervento con ragionevole certezza, cioè in maniera statisticamente significativa.

In generale, più grande è il campione, più precisi e attendibili saranno i risultati, a patto che il campionamento sia realizzato con un metodo corretto e sapendo che indagini su campioni di grandi dimensioni sono più costose e richiedono più tempo.



Armitage P, Berry G. Statistica Medica, metodi statistici per la ricerca in Medicina, 1996;.

Dimensione campionaria

L'arruolamento di un numero adeguato di soggetti ha anche una rilevanza etica che deve essere tenuta in considerazione quando si progetta una ricerca.

Infatti, uno studio sottodimensionato che, pertanto, molto probabilmente non consentirà un progresso nelle conoscenze scientifiche, potrebbe esporre inutilmente delle persone a interventi potenzialmente rischiosi; analogamente, uno studio sovradimensionato esporrebbe ai medesimi rischi un numero ancor più grande di persone al solo fine di conferire a un effetto osservato una significatività statistica che surroghi quasi artificialmente la difficile dimostrazione di una chiara significatività clinica.



Franco F, G Clin Nefrol Dial 2020;

Dimensione campionaria

Al fine del corretto dimensionamento di un campione è indispensabile che il clinico definisca in maniera esplicita, anche in termini quantitativi, quali sono gli obiettivi che lo studio si prefigge; in particolare, quant'è la differenza che si aspetta di osservare tra i gruppi a confronto, ad esempio quanto miglioramento ci si aspetta da una nuova terapia da introdurre rispetto a chi assume farmaci già correntemente utilizzati.

A tal fine è indispensabile avere una stima del parametro nel gruppo di controllo rispetto a quello al quale viene ad esempio somministrata la nuova terapia oggetto dello studio.



Franco F, G Clin Nefrol Dial 2020;

Dimensione campionaria

Il calcolo della dimensione del campione dipende dalle ipotesi statistiche prestabilite e dai seguenti presupposti:

Endpoint primario

Numero di gruppi di trattamento coinvolti

Errore di tipo I (alfa) e II (beta)

Potenza

Differenza minima significativa rilevabile

Variabilità della misura

Tasso di abbandono

Franco F, G Clin Nefrol Dial 2020;

Dimensione campionaria

Il calcolo della dimensione campionaria dipende dal tipo di studio epidemiologico che si vuole intraprendere: descrittivo, osservazionale o controllato randomizzato.

A ogni tipo di studio corrisponde una specifica formula di calcolo della dimensione campionaria,



Dimensione campionaria

La power analysis è la tecnica statistica maggiormente utilizzata per determinare la numerosità campionaria minima necessaria da raccogliere nel momento in cui impostiamo uno studio scientifico, soprattutto se si ha la necessità di effettuare un trial o uno studio randomizzato (caso-controllo).

Negli studi retrospettivi, tipicamente, la dimensione del campione è predeterminata e non è possibile modificarla a posteriori, tuttavia tramite la power analysis a posteriori possiamo comunque comprendere se quella dimensione campionaria è adeguata o se corriamo il rischio che i nostri test ed i nostri modelli siano under-powered, ovvero non abbiano una sufficiente potenza statistica.

<https://www.analisi-statistiche.it/>

Dimensione campionaria

		Realtà	
		H_0 vera	H_0 falsa
Conclusione del test	H_0 vera	Decisione corretta	Errore di II tipo
	H_0 falsa	Errore di I tipo	Decisione corretta

Per capire cos'è la potenza statistica dobbiamo introdurre il concetto di errore.

In statistica esistono due tipi di errore: l'errore di I tipo (α) e l'errore di II tipo (β).

L'errore di I tipo è quello che si compie quando si rifiuta l'ipotesi nulla del test (H_0) quando questa è vera. L'errore di II tipo avviene se si accetta l'ipotesi nulla del test (H_0) quando questa è falsa.

La **potenza statistica** di un test per definizione è $1-\beta$ e possiamo interpretarla come la probabilità di non commettere l'errore di II tipo, ovvero la probabilità di rifiutare correttamente l'ipotesi nulla.

Dimensione campionaria

La numerosità campionaria minima il numero di unità statistiche necessario per il nostro studio e deve essere scelto in modo da ottenere una potenza statistica per i nostri test almeno pari, tipicamente, a 0.8.

Per il calcolo della numerosità campionaria ottimale abbiamo bisogno di:

- Scegliere il livello di significatività (tipicamente 0.05)
- Determinare la potenza statistica minima (tipicamente almeno 0.8)
- Identificare l'outcome primario per il nostro studio, ovvero la variabile di interesse primario
- Analizzare i dati preliminari o di letteratura

Infatti, non esiste una tecnica di power analysis unica e generalizzata, bensì dipende dall'outcome primario dello studio.

Dimensione campionaria

Se ho la necessità di valutare se esiste una differenza statisticamente significativa su un parametro di interesse quantitativo tra due gruppi, che sono stati tipicamente randomizzati inizialmente, il test parametrico di riferimento sarà il t-test per campioni indipendenti.

I dati che ci servono per la power analysis per campioni indipendenti sono dunque:

- Un valore atteso per la media dell'outcome primario nei due gruppi
- Un valore atteso per la deviazione standard dell'outcome primario nei due gruppi

Questi dati li possiamo trovare in letteratura oppure tramite dati preliminari già raccolti. Se non sono disponibili in nessun caso questi dati, li possiamo semplicemente ipotizzare in base al buon senso o in base alle nostre supposizioni di ricerca.

In alternativa posso fornire direttamente anche l'effect size che ci aspettiamo per il test. Quella più comunemente usata per due campioni indipendenti è la D di Cohen.

Dimensione campionaria

Ad esempio, per la nostra Power analysis possiamo utilizzare SPSS (ultime versioni):

1. Andiamo su Analizza
2. Selezioniamo Analisi di potenza
3. Medie
4. T-test a campione indipendente
5. Inseriamo i dati richiesti

Un altro ottimo tool gratuito per il calcolo della dimensione minima campionaria per due gruppi indipendenti è G*Power. Si tratta di un software gratuito, disponibile in rete e piuttosto semplice da utilizzare.

1. Mettere come famiglia dei test: t tests
2. Come Statistical test selezionare: Means: Difference between two independent means (two group)
3. Scegliere il tipo di power analysis (tipicamente a priori)
4. Compilare con i parametri di input

<https://www.analisi-statistiche.it/>

Dimensione campionaria

Il calcolo della dimensione del campione è il primo e uno dei passi più importanti nella pianificazione di uno studio clinico e qualsiasi negligenza nella sua stima può portare al rifiuto di un trattamento efficace e all'approvazione di un trattamento inefficace.

Sebbene le tecniche per il calcolo della dimensione del campione siano descritte in vari libri di statistica, l'esecuzione di questi calcoli può essere complicata ed è auspicabile consultare un biostatistico esperto nella stima di questo parametro vitale dello studio.

