



WORKSHOP N. 3

La redazione di un protocollo di ricerca

La statistica: dalla numerosità del campione alle analisi statistiche previste

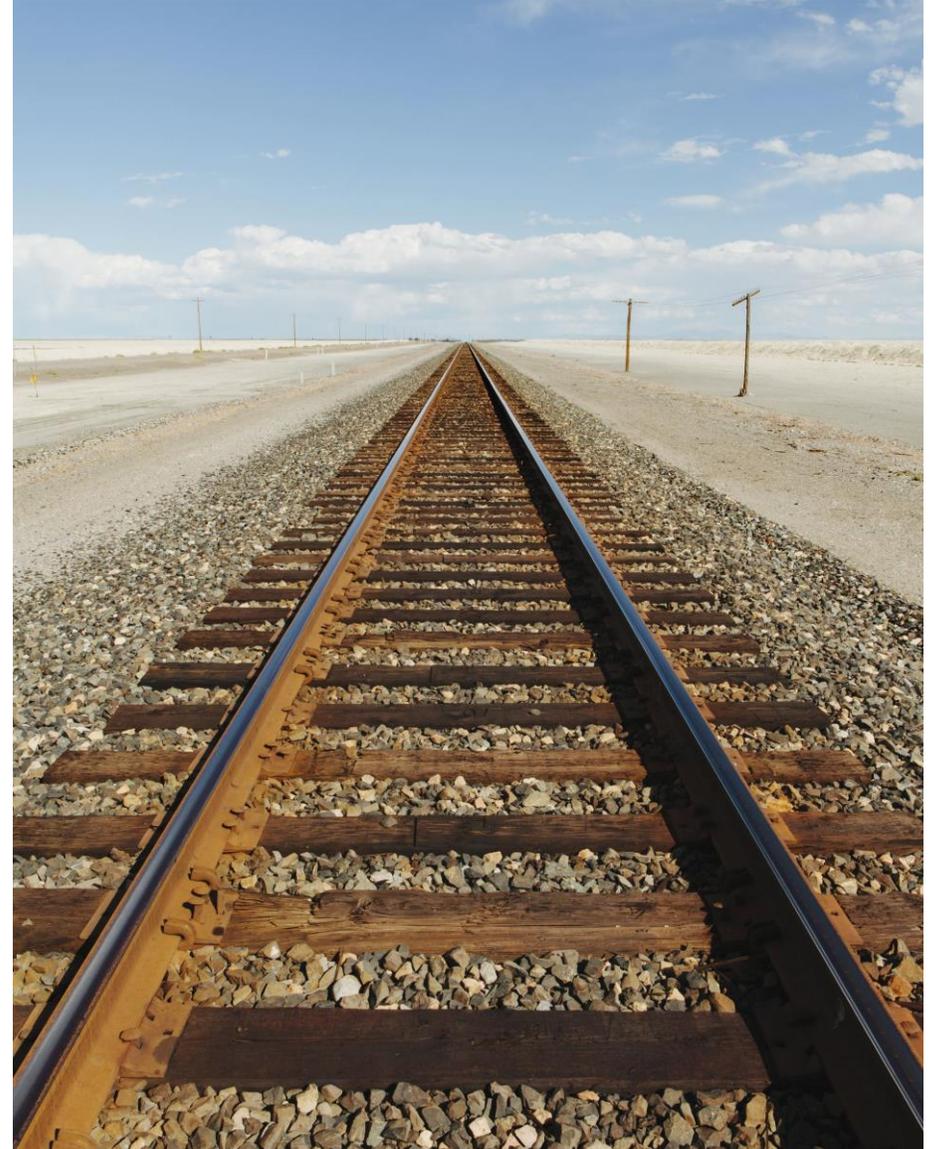
Dott.ssa Lea Godino, PhD

lea.godino2@unibo.it

IRCCS AOU di Bologna

INDICE

- Campionamento
- Statistica per uno studio
- Dimensione del campione in una indagine
- Dimensione o numerosità del campione





CAMPIONAMENTO

Popolazione: *un insieme N (ampiezza della popolazione) di unità (dette anche 'unità statistiche') che costituiscono l'oggetto del nostro studio.*

Campione: *un insieme delle n (ampiezza del campione) unità campionarie selezionate tra le N unità che compongono la popolazione, allo scopo di rappresentarla (da cui l'uso dell'espressione 'campione rappresentativo) ai fini del nostro studio. Deve essere **omogeneo e rappresentativo** per fare generalizzazioni sulla popolazione.*

Campionamento: *la procedura che seguiamo per scegliere le n unità campionarie dal complesso delle N unità della popolazione.*

UN BEL PIATTO DI CAVALLETTE E SALSA GUACAMOLA

Ci piaceranno ??



UN BEL PIATTO DI CAVALLETTE E SALSA GUACAMOLA

L'idea è di assaggiare il pasto

Prendiamo una cavalletta a caso e assumiamo che:

- una cavalletta vale l'altra
- una cavalletta (il campione) e' rappresentativa del resto del piatto (popolazione)
- una cavalletta e' efficiente in quanto minimizza il costo dell'indagine



IL PRELIEVO DEL SANGUE

Si ottiene un campione a caso del sangue in circolazione

- Sufficientemente ampio da rappresentare l'intero organo (popolazione)
- Sufficientemente piccolo da lasciare intatta la popolazione
- L'intera popolazione non conviene testarla!!!





CAMPIONAMENTO

Operare su un insieme ridotto n invece che sull'intera popolazione N , presenta degli evidenti vantaggi, ma anche degli svantaggi.

L'indagine totale fornisce il *valore esatto* del parametro che si vuole conoscere, l'indagine campionaria ne dà solo una *stima* (un valore approssimato).

La stima del campione è affetta dall'**errore di campionamento**

Esempio:

<Il reddito medio della popolazione è di euro 1200>

<C'è il 95% di probabilità che il reddito medio sia di euro 1200 ± 50 >



CAMPIONAMENTO

L'errore di campionamento si colloca nel quadro più ampio degli

ERRORI DI SELEZIONE

che includono:

l'errore di copertura: dovuto alla mancata disponibilità della lista completa della popolazione oggetto di studio;

l'errore di campionamento: dovuto al fatto di operare su un particolare sottoinsieme della popolazione;

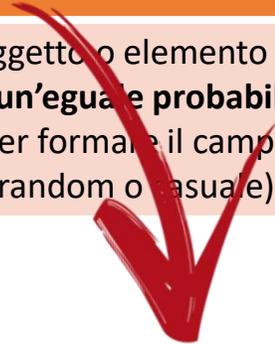
l'errore di non risposta: conseguente al fatto che, nel caso di ricerche condotte interrogando i soggetti, alcuni di questi possono non essere raggiunti dall'intervistatore oppure rifiutarsi di rispondere, pur facendo parte del campione.



CAMPIONAMENTO

Esistono vari tipi di campionamento riconducibili a due grandi categorie:

CAMPIONAMENTO PROBABILISTICO	CAMPIONAMENTO NON PROBABILISTICO
Ogni soggetto, oggetto o elemento in una popolazione ha un'eguale probabilità di essere scelto per formare il campione (campione random o casuale).	Il campione non è selezionato in maniera casuale e, conseguentemente, non tutti i soggetti, oggetti o elementi della popolazione hanno la stessa probabilità di essere scelti



Diverse tipologie di campionamento, alcuni sono:

Campionamento *casuale semplice*

Campionamento *sistematico*

Campionamento *stratificato*

Campionamento *a grappoli*



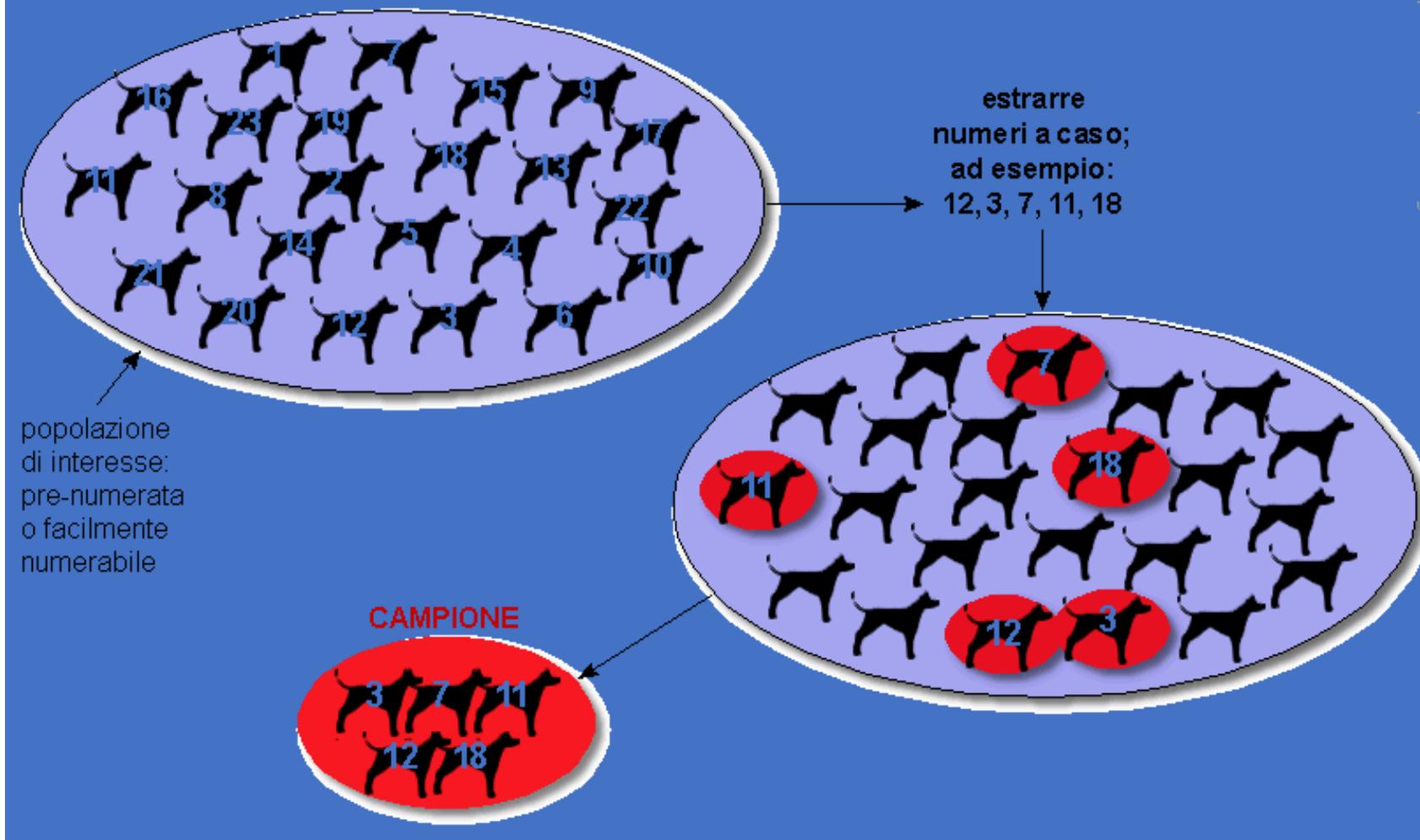
Campionamento casuale semplice

... tutte le unità hanno della popolazione hanno la stessa probabilità di essere incluse nel campione.

Come si costruisce questo campione?

Dopo aver definito una lista di campionamento in cui a ogni soggetto viene associato un numero d'ordine, ha inizio l'estrazione dei singoli individui (servendosi ad esempio di un generatore di numeri casuali) e, nel caso in cui ogni soggetto estratto non venga più reimpresso per le estrazioni successive, avremo un campionamento casuale semplice.

Esempio di campionamento mediante randomizzazione semplice





Campionamento a grappoli

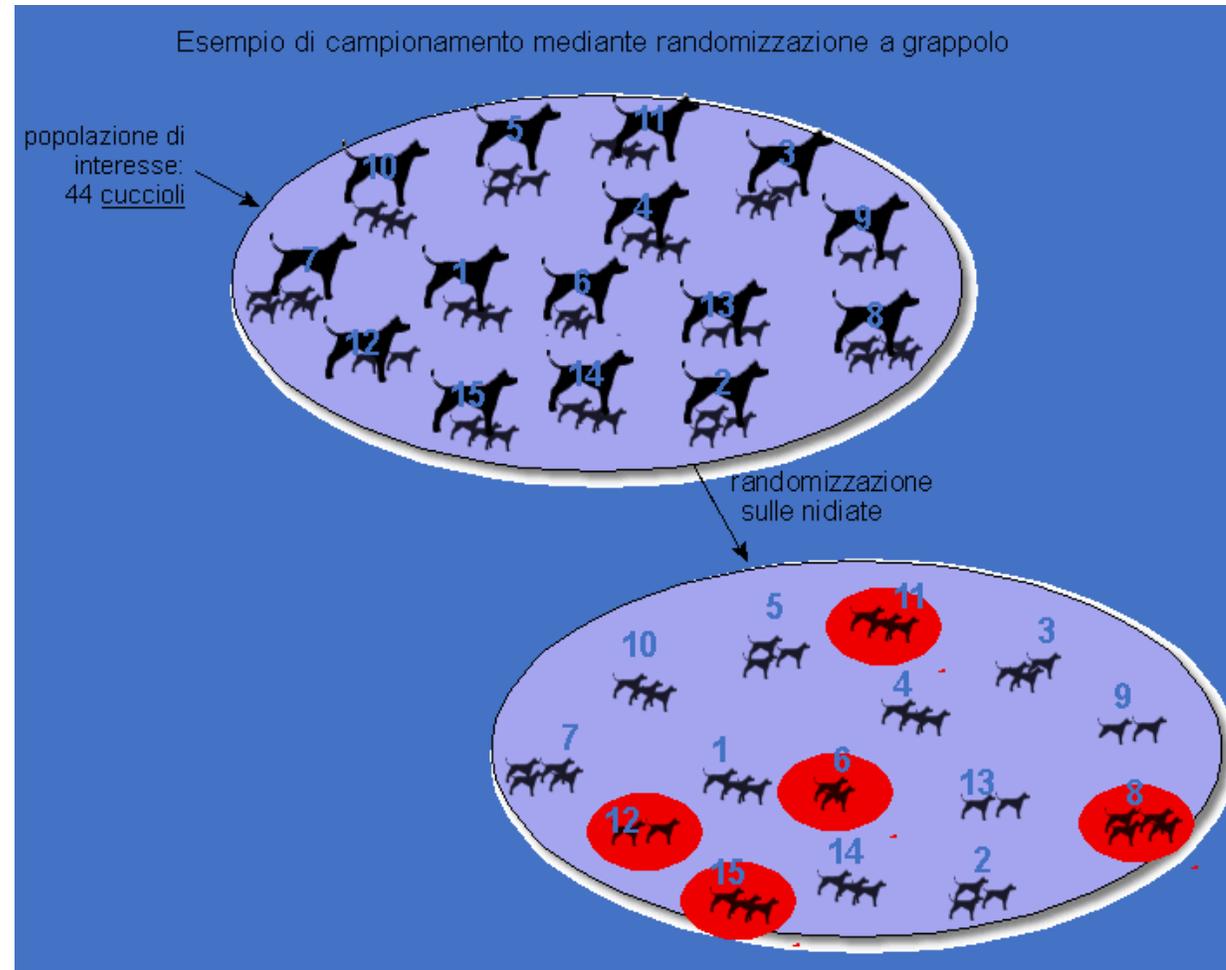
Il campionamento a grappolo consiste in un metodo in cui, invece di procedere alla selezione diretta delle singole unità di interesse, si selezionano dei gruppi (*grappoli* o *cluster*) di unità.

Nella pratica, spesso i *cluster* sono già preformati, e comprendono un numero limitato di unità (es. nidi di suinetti, box di vitelli, ecc.). In alternativa, i cluster possono essere individuati artificialmente (es. cluster di allevamenti raggruppati su base geografica).

CAMPIONAMENTO
A GRAPPOLO

- effettuato su una popolazione già spontaneamente suddivisa in gruppi (es. nidi, gabbie ecc.) di solito formati da poche unità
- può comportare un errore di campionamento superiore rispetto ad altri metodi randomizzati

Campionamento a grappoli



Autorizzato dal Prof. Ezio Bottarelli

STATISTICA

Statistica non è sinonimo di probabilità

Con statistica si intende invece la disciplina che, in senso lato, si interessa della raccolta e dell'analisi dei dati e dell'interpretazione dei risultati.

In particolare si parlerà di ...

STATISTICA DESCRITTIVA

- Descrivere i dati grezzi con pochi numeri o grafici significativi
- Si occupa di fotografare una data situazione e di sintetizzarne le caratteristiche salienti

STATISTICA INFERENZIALE

- Permette di conoscere se le relazioni osservate in un campione possano occorrere nella popolazione generale.
- Valuta la variabilità aleatoria e controlla i fattori di confusione

DIMENSIONE DEL CAMPIONE E INDAGINE

A quante persone è veramente
necessario inviare la mia indagine?

Ovvero ... quanto deve essere grande il
mio campione, al fine di stimare con
precisione la mia popolazione?



Fase 1: QUALE POPOLAZIONE?

Per popolazione si intende l'insieme di persone che vuoi analizzare.

Non importa se si tratta di un intero paese o di un'azienda, individuare la popolazione che vuoi analizzare è un passo fondamentale.

Una volta trovata la tua popolazione, devi capire di quante persone (approssimativamente) dovrà essere composta.

Trovato il numero? Ok, procediamo...



Fase 2:

QUANTO DEVI ESSERE PRECISO?

Pensa a questa fase come a una valutazione di quanto sei disposto a rischiare che le risposte alla tua indagine siano un po' imprecise a causa del fatto che non stai conducendo l'indagine sull'intera popolazione.

Ecco le **due** domande a cui devi rispondere:

- 1. Quanto devi essere sicuro che le risposte riflettono le opinioni della tua popolazione?*

Questo è il margine di errore. Quindi, se, per esempio, al 90% del tuo campione piace la gomma da masticare al gusto di uva. Un margine di errore del 5% aggiungerebbe il 5% su entrambi i lati di quel numero. Questo significa che in realtà al 85-95% del tuo campione piace la gomma da masticare al gusto di uva. Il 5% è il margine di errore più comunemente utilizzato. Non è consigliato di aumentare il margine di errore al di sopra del 10%.



Fase 2:

QUANTO DEVI ESSERE PRECISO?

2. *Quanto devi essere sicuro che il campione rispecchi fedelmente la tua popolazione target?*

Questo è il tuo livello di confidenza.

Un livello di confidenza è *la probabilità che il campione da te scelto sia importante per i risultati ottenuti.*

Solitamente il calcolo viene eseguito nel modo seguente. Se hai scelto altri 30 campioni casuali dalla tua popolazione, con quale frequenza i risultati ottenuti con un campione differiscono in modo significativo da questi altri 30 campioni? Un livello di confidenza del 95% significa che si dovrebbero ottenere gli stessi risultati il 95% delle volte. 95% è il livello di confidenza più comunemente utilizzato.



Fase 3:

QUANTO DEVE ESSERE GRANDE IL CAMPIONE?

Utilizzando la tabella seguente, seleziona la popolazione target approssimativa, poi scegli il margine di errore per stimare il numero di indagini compilate di cui ne hai bisogno.

Popolazione	Margine di errore		Livello di confidenza	
	5%	1%	95%	99%
100	80	99	80	88
500	218	476	218	286
1000	278	906	278	400
10000	370	4900	370	623
100000	383	8763	383	660
oltre 1000000	384	9513	384	664



Fase 4:

QUANTO SARANNO REATTIVE LE PERSONE?

Siamo spiacenti di doverti dire che non tutti coloro a cui invii l'indagine la compileranno.

La percentuale di persone che effettivamente risponde all'indagine ricevuta viene definita "*percentuale di risposta*".

La valutazione della tua percentuale di risposta ti aiuterà a determinare il numero totale di indagini che dovrai inviare per poter ottenere il numero richiesto di indagini completate.

Le percentuali di risposta variano sensibilmente in base a diversi fattori, quali il rapporto con il tuo pubblico di destinazione, lunghezza e complessità dell'indagine, incentivi e argomento trattato.

Una percentuale di risposta tra il 10 e il 15% è un'ipotesi cauta e più sicura se non hai mai condotto prima un'indagine sulla tua popolazione.



Fase 5:

A QUANTE PERSONE DEVO INVIARLA?

**Determinare una appropriata numerosità campionaria
già in fase di disegno dello studio
è molto **importante****

**Studi basati su campioni troppo piccoli non hanno
la sensibilità adeguata per poter individuare un
effetto**

Ad esempio: in uno studio di confronto tra due gruppi anche quando una differenza esiste realmente può accadere che una numerosità campionaria troppo esigua comporti un livello di significatività alto e un intervallo di confidenza ampio e ciò potrebbe portare il ricercatore alla conclusione errata di assenza di differenza tra i due gruppi





LA NUMEROSITÀ CAMPIONARIA

A parità di altre condizioni, in genere uno studio è tanto più affidabile quante più persone arruola.

la numerosità del campione riduce la probabilità di due errori:

- il **falso positivo** ovvero l'attribuzione al trattamento di un'efficacia che in realtà non ha
- Il **falso negativo**, ovvero il mancato riconoscimento dell'efficacia che il trattamento ha ma che lo studio non documenta



LA NUMEROSITÀ CAMPIONARIA

Falso positivo e livello di significatività

- Per convenzione in genere si accetta che il caso abbia non più di 1 o 5 probabilità su 100 di determinare un risultato **falso positivo**: $p=0,01$ oppure $p=0,05$, significa che ci sono 1 e rispettivamente 5 probabilità su 100 che il risultato osservato sia dovuto al caso anziché al trattamento e la probabilità che sia stato proprio il trattamento a determinare il beneficio è del 99% o, rispettivamente del 95%:

solo in questi casi il risultato è considerato statisticamente significativo

LIVELLO DI SIGNIFICATIVITÀ E VALORI DELLA NORMALE (GAUSSIANA)

Ad ogni livello di significatività corrisponde un valore della Normale standardizzata: Al 95% di significatività corrisponde uno z critico pari a 1,96

Valori della Normale standardizzata corrispondente ai diversi livelli di significatività

p-value	significatività	Zcrit value
0,01	99%	2,58
0,02	98%	2,33
0,05	95%	1,96
0,1	99%	1,64



LA NUMEROSITÀ CAMPIONARIA: FALSO NEGATIVO E POTENZA DELLO STUDIO

Sempre per convenzione si accetta che il caso abbia non più di 10 o 20 probabilità su 100 di non far emergere un vantaggio relativo al trattamento **falso negativo**. In questo caso si dice che **lo studio ha una POTENZA rispettivamente del 90% o dell'80%**.

QUANTI SOGGETTI DEVONO ESSERE SELEZIONATI?

Per il calcolo della numerosità del campione è necessario considerare:

- **p-value** la probabilità di interpretare erroneamente come vero un risultato casuale (errore di I° tipo)
- **Potenza**: la probabilità di identificare correttamente una differenza tra due gruppi in studio quando questa differenza esiste nelle popolazioni da cui i campioni sono stati estratti (1-errore di II° tipo)
- **Effetto**: la differenza minima attesa tra i due gruppi (in termini di esiti in studio) . E' una quantità soggettiva, basata sul giudizio clinico.

- **Deviazione standard** ovvero variabilità casuale del fenomeno in studio. Si considerano dati di letteratura, indagini precedenti ecc.

QUANTI SOGGETTI DEVONO ESSERE SELEZIONATI?

IL CALCOLO DELLA NUMEROSITÀ CAMPIONARIA

$$N = \frac{4\sigma^2(Z_{\text{crit}} + Z_{\text{pwr}})^2}{D^2}$$

- σ^2 è la varianza (il quadrato della deviazione standard)
- Z_{crit} è il valore della normale in corrispondenza del p-value fissato
- Z_{pwr} è il valore della normale in corrispondenza della potenza fissata
- D^2 è il quadrato della differenza minima attesa



IL CALCOLO DELLA NUMEROSITÀ CAMPIONARIA

Studio di confronto tra due gruppi

Calcolare la numerosità campionaria fissato un livello di significatività del 5% e una potenza dell'90%.

La differenza attesa è 3 e la varianza 20

$$N = \frac{4\sigma^2(Z_{\text{crit}} + Z_{\text{pwr}})^2}{D^2} = \frac{4 \times 20 \times (1,96 + 1,28)^2}{3^2} = 93$$

Valori della Normale standardizzata corrispondente ai diversi livelli di significatività

p-value	significatività	Z _{crit} Value
0,01	99%	2,58
0,02	98%	2,33
0,05	95%	1,96
0,1	90%	1,64

Valori della Normale standardizzata corrispondenti a diversi valori di potenza statistica dello studio

Potenza	Z _{pwr} value
80%	0,84
85%	1,04
90%	1,28
95%	1,64



IL CALCOLO DELLA NUMEROSITÀ CAMPIONARIA PER STUDI DESCRITTIVI

Studi descrittivi

Nel caso di studi descrittivi in cui non interessa il confronto tra gruppi ad esempio negli studi di prevalenza, in generale negli studi in cui si voglia stimare una proporzione di un fenomeno, il calcolo della numerosità campionaria si basa sulla seguente formula:

$$N = \frac{4p(1-p) Z_{\text{crit}}^2}{D^2}$$

$p(1-p)$ = è la varianza della proporzione p . Il valore di p deriva da studi precedenti o dati di letteratura e nel caso non si conosca viene posto pari a 0,5

D^2 = è la precisione della stima che si desidera ottenere, ovvero l'ampiezza dell'intervallo di confidenza



IL CALCOLO DELLA NUMEROSITÀ CAMPIONARIA PER STUDI DESCRITTIVI

L'obiettivo di uno studio sull'alimentazione degli adolescenti, è di stimare il loro consumo proteico medio giornaliero, misurato in grammi. Quanti adolescenti è necessario reclutare nello studio, in modo da stimare il valore medio entro un intervallo di confidenza di 10 grammi, ad un livello di significatività del 5% e sapendo che la varianza (ovvero la variabilità intorno al valor medio), come risulta da precedenti studi è di 20 grammi

$$p(1-p) = 20$$

$$D^2 = 10^2$$

$$Z_{\text{crit}}^2 = 1,96^2$$

$$N = \frac{4p(1-p) Z_{\text{crit}}^2}{D^2} = \frac{4 \times 20 \times 1,96^2}{10^2} = 61,47$$

Quindi per ottenere stime ad una significatività del 95% e intervallo di confidenza di ampiezza 10 grammi
il campione deve essere composto da **non meno di 62 adolescenti**

DIMENSIONE O NUMEROSITÀ DEL CAMPIONE

Questa è una delle parti più delicate nella pianificazione di una indagine.

Ovviamente, più grande sarà il campione e più precisi e attendibili saranno i risultati, a patto che il campione sia stato selezionato con un metodo corretto.

Tuttavia, indagini su campioni di grandi dimensioni sono più costose e richiedono più tempo. Quindi, bisogna accettare un compromesso.



DIMENSIONE O NUMEROSITÀ DEL CAMPIONE

Il calcolo della dimensione del campione, più propriamente detta **numerosità**, è abbastanza complicato e, soprattutto, richiede la conoscenza di informazioni diverse.

I principali fattori che da considerare nell'individuazione della numerosità del campione sono: la **varianza** e l'ampiezza desiderata dell'**intervallo di confidenza**.

FATTORI DA CONSIDERARE
nella scelta della
DIMENSIONE DEL CAMPIONE

VARIANZA nella popolazione
del carattere in studio

ampiezza desiderata dell'
INTERVALLO DI CONFIDENZA



DIMENSIONE O NUMEROSITÀ DEL CAMPIONE

La **varianza** è una misura del grado di variazioni o oscillazioni presenti, relativamente al parametro che vogliamo stimare, nella popolazione.

L'entità di queste variazioni può essere derivata, almeno approssimativamente, dall'esperienza, o dai risultati di altre analoghe indagini effettuate in precedenza, o dalla conoscenza della storia naturale della malattia, o da altri fattori.

Una popolazione in cui il parametro da misurare presenta ampie oscillazioni ha una varianza elevata; una popolazione in cui le oscillazioni sono scarse ha una varianza bassa.

È intuitivo che la precisione di un campione è maggiore quando la popolazione da cui è stato estratto è tendenzialmente omogenea, mentre è minore quando la popolazione è eterogenea..



DIMENSIONE O NUMEROSITÀ DEL CAMPIONE

L'**intervallo di confidenza** rappresenta una misura della bontà di una *stima*. Un intervallo di confidenza molto ampio suggerisce che non siamo molto sicuri del punto in cui si trova il «vero» valore. Viceversa, un intervallo ristretto indica che siamo abbastanza sicuri che il valore trovato è piuttosto vicino al valore vero della popolazione; in questo caso la stima sarà, quindi, più precisa.

Il **livello di confidenza** è una misura della sicurezza della stima: ad esempio, con un livello di confidenza 95% siamo sicuri al 95% che il valore vero cade nell'intervallo trovato. Cioè, se ripetessimo lo studio 20 volte, in media sbaglieremmo 1 volta ma saremmo nel giusto 19 volte.

<https://clincalc.com/stats/samplesize.aspx>

<http://ebcp.it/evidencebasednursing/index.html>



INDICE DEI CONTENUTI

SINOSI	3
SCHEMA DELLO STUDIO	5
INTRODUZIONE E RAZIONALE	6
OBIETTIVI	7
PRIMARIO	7
SECONDARI.....	7
ENDPOINT	7
PRIMARIO	7
SECONDARI.....	7
DISEGNO DELLO STUDIO	7
CRITERI DI ELEGGIBILITÀ	7
CRITERI DI INCLUSIONE:.....	7
CRITERI DI ESCLUSIONE:	8
SETTING	8
INTERVENTI	8
SPERIMENTALE.....	8
STANDARD	9
CONSIDERAZIONI STATISTICHE	9
RANDOMIZZAZIONE E ALLOCATION CONCEALMENT	9
ANALISI STATISTICA	9
DATA MANAGEMENT	10
MONITORAGGIO DELLO STUDIO	11
CONSERVAZIONE DELLA DOCUMENTAZIONE	12
ASPETTI ETICI	13
DICHIARAZIONE DI HELSINKI	13
ASSICURAZIONE	14
DISSEMINAZIONE	14
ACCORDO DI SEGRETEZZA	14
RENDICONTAZIONE E PUBBLICAZIONE DEI RISULTATI.....	15
ASPETTI ORGANIZZATIVI	15
IMPLICAZIONI CLINICO-ORGANIZZATIVE	15
UTILIZZO DEI FONDI E DELLE RISORSE	17
RIFERIMENTI BIBLIOGRAFICI	17

Piano statistico

Descrivere tutti i metodi statistici inclusi quelli usati per il controllo dei fattori confondenti, e il livello di significatività.

Descrivere modalità di trattamento dei dati mancanti (missing).

Descrivere i metodi usati per l'analisi dei sottogruppi e per l'identificazione di eventuali interazioni.

Per studi di coorte: descrivere la modalità di gestione delle perdite al follow-up.

Per studi caso controllo: descrivere la modalità di gestione dei dati appaiati.

Per studi trasversali: descrivere la modalità di analisi coerente con la strategia di campionamento (casuale, stratificato, cluster etc).

Descrivere l'analisi di sensitività, cioè una verifica della persistenza dei risultati in funzione di diverse assunzioni e/o strategie analitiche.

Protocollo di studio

Simultaneous Care e Qualità di Vita in Oncologia

Versione# 2.0 del 25/03/2014

I

PROTOCOLLO DI STUDIO

Titolo del progetto:	Valutazione dell'efficacia del modello assistenziale "Simultaneous Care" nel miglioramento della qualità di vita dei pazienti affetti da neoplasie maligne. Studio clinico controllato e randomizzato.
Titolo breve	Simultaneous Care e Qualità di Vita in Oncologia
Codice Unico di Progetto	E35E13000030002 con delibera della Giunta R. 1939 del 16 dicembre 2013

Numerosità dei soggetti

Sono necessari 243 pazienti arruolati per avere una potenza dell'80% di scoprire una differenza statisticamente significativa tra i due gruppi nella modifica del FACT-G score di almeno 6.5 punti (SD=15). La stima comprende una quota del 30% di persi al follow-up.

Analisi statistica

La popolazione in studio sarà ITT, quindi tutti i dati dei pazienti randomizzati e con questionario valido verranno analizzati. Per descrivere le variabili dello studio saranno utilizzate frequenze, medie, mediane, intervalli interquartile e deviazioni standard. Le differenze tra i gruppi saranno valutate utilizzando il test chi-quadro per le variabili categoriche e il test t-Student per campioni indipendenti per le variabili continue. I modelli di regressione lineare multivariata saranno aggiustati per gli score al baseline e saranno utilizzati per esaminare l'effetto dell'intervento sulla qualità della vita. Per l'analisi di sopravvivenza verrà utilizzato il metodo di Kaplan-Meier ed il tempo di sopravvivenza verrà calcolato dalla data di arruolamento alla data di morte. I dati dei persi al FU verranno censurati. Il Proportional hazard model di Cox sarà utilizzato per valutare l'effetto delle cure palliative sulla sopravvivenza, aggiustata per le caratteristiche demografiche e per il performance status al baseline. Le analisi saranno effettuate utilizzando il package statistico SAS.

SINOSI

Protocollo di studio

Simultaneous Care e Qualità di Vita in Oncologia

Versione# 2.0 del 25/03/2014

I

PROTOCOLLO DI STUDIO

Titolo del progetto:	Valutazione dell'efficacia del modello assistenziale "Simultaneous Care" nel miglioramento della qualità di vita dei pazienti affetti da neoplasie maligne. Studio clinico controllato e randomizzato.
Titolo breve	Simultaneous Care e Qualità di Vita in Oncologia
Codice Unico di Progetto	E35E13000030002 con delibera della Giunta Regionale N.

Considerazioni statistiche

Sample size

Per scoprire una differenza significativa, a 12 settimane tra i due gruppi nella modifica del FACT-G score di almeno 6.5 punti, nell'ipotesi che la sopravvivenza a 12 settimane sia pari all'80%, al fine di assicurare una potenza statistica di 0.80 con un $\alpha=0.05$, tramite un t-test a due code e una $SD=15$ [Temel 2010], si stima un numero necessario di pazienti da arruolare pari a 243. Il dato è comprensivo di una quota in più, pari al 30%, per compensare la perdita prevista di pazienti al follow-up.

Per la stima è stato utilizzato un approccio basato su una simulazione MonteCarlo [Landau 2013] per poter tenere conto dell'eterogeneità della casistica inclusa (varie tipologie di tumore). Non essendo disponibili in letteratura dati sul FACT-G al baseline suddivisi per tipologia di tumore, sono stati utilizzati i valori basali per performance status ECOG, nel dettaglio: ecog0 ($\mu=87.8$, $sd=14.2$), ecog1 ($\mu=78.8$, $sd=15.2$), ecog2 ($\mu=71.1$, $sd=-15.4$) come riportato in un articolo in fase di pubblicazione del Dr. Cella (ideatore della scala).

Randomizzazione e allocation concealment

La lista di randomizzazione verrà creata con un apposito programma in SAS. L'algoritmo di generazione della lista sarà protetto, a cura della UO Ricerca e Innovazione, in un documento separato. La randomizzazione verrà stratificata per centro, al fine di ridurre il rischio di sbilanciamento dovuto alla diversa expertise del personale coinvolto, e sarà centralizzata su sito web dedicato, per garantire il mascheramento della sequenza di allocazione al trattamento (allocation concealment). Verrà assicurata la cecità dello statistico mascherando il codice di randomizzazione relativo al trattamento.

PROTOCOLLO

Analisi statistica

La popolazione in studio sarà analizzata secondo l'intention to treat (ITT), quindi tutti i dati dei pazienti come randomizzati e con questionario valido verranno analizzati. Le procedure utilizzate per descrivere l'analisi statistica saranno documentate nello Statistical Analysis Plan (SAP), definito prima dell'arruolamento del primo paziente.

Per l'analisi descrittiva delle variabili in studio verranno utilizzate frequenze, medie, mediane, intervalli interquartili e deviazioni standard. Le differenze tra i gruppi saranno valutate utilizzando il test chi-quadro per le variabili categoriche e il test t-Student per campioni indipendenti per le variabili continue. Il modello di regressione lineare multivariata sarà utilizzato per esaminare l'effetto dell'intervento sulla qualità della vita, aggiustando per lo score al basale e per altri confondenti individuati in letteratura e mediante analisi univariata ($p<0.20$). Per l'analisi di sopravvivenza verrà utilizzato il metodo di Kaplan-Meier ed il tempo di sopravvivenza verrà calcolato dalla data di arruolamento alla data di morte. I dati dei persi al FU verranno censurati. Il Proportional hazard model di Cox sarà utilizzato per valutare l'effetto delle cure palliative sulla sopravvivenza, aggiustata per le caratteristiche demografiche e per il performance status al baseline. Le analisi saranno effettuate utilizzando il package statistico SAS nella versione più aggiornata disponibile.

1936: ELEZIONI PRESIDENZIALI USA

Candidati: F.D. Roosevelt e A. Landon



LO SAPEVATE CHE...??!

Indagine Literary Digest:

- 10 milioni di fac-simile di schede elettorali inviate a nominativi estratti dagli elenchi telefonici e dai registri automobilistici
- Risultato previsto:
Roosevelt 41 % e Landon 59 %



1936: ELEZIONI PRESIDENZIALI USA

Risultato delle elezioni: Roosevelt 61 %

Gli errori del Literary Digest:

- **ERRORE DI COPERTURA**
- le liste usate non erano complete
- gli elenchi usati non erano rappresentativi dell'intera popolazione ma solo dei ceti più abbienti che tendevano a votare repubblicano
- **AUTOSELEZIONE del CAMPIONE**
- le caratteristiche socio-demografiche dei cittadini che risposero al sondaggio erano presumibilmente diverse da quelle di chi non rispose (istruzione, reddito, etc.)

"All the News That's Fit to Print."

The New York Times

LATE CITY EDITION
City, much color today. Tomorrow, still mostly and...
Times Herald—Mon., 46, No. 12

VOL. XC, No. 30,237. PUBLISHED DAILY, EXCEPT ON SUNDAYS, HOLIDAYS AND WEDNESDAYS.
NEW YORK, WEDNESDAY, NOVEMBER 4, 1936.
THREE CENTS

ROOSEVELT ELECTED PRESIDENT; CERTAIN OF 429 ELECTORAL VOTES; DEMOCRATS KEEP HOUSE CONTROL

RETAIN HOUSE GRIP

Democrats, Holding 225 Seats, Gain at Least Ten From Rivals

65 ARE NOW IN DOUBT

Latest Figures Indicate Republican Gain of 1 to 3 Senators

THE VOTE FOR PRESIDENT

State	Electors	Dem.	Rep.	Other
Alabama	10	10	0	0
Arizona	8	8	0	0
Arkansas	12	12	0	0
California	21	21	0	0
Colorado	10	10	0	0
Connecticut	12	12	0	0
Delaware	6	6	0	0
District of Columbia	3	3	0	0
Florida	10	10	0	0
Georgia	15	15	0	0
Idaho	6	6	0	0
Illinois	27	27	0	0
Indiana	13	13	0	0
Iowa	12	12	0	0
Kansas	8	8	0	0
Kentucky	12	12	0	0
Louisiana	10	10	0	0
Maine	10	10	0	0
Massachusetts	14	14	0	0
Michigan	20	20	0	0
Minnesota	15	15	0	0
Mississippi	8	8	0	0
Missouri	16	16	0	0
Montana	6	6	0	0
Nebraska	8	8	0	0
Nevada	6	6	0	0
New Hampshire	10	10	0	0
New Jersey	14	14	0	0
New Mexico	6	6	0	0
New York	36	36	0	0
North Carolina	15	15	0	0
North Dakota	6	6	0	0
Ohio	25	25	0	0
Oklahoma	8	8	0	0
Oregon	6	6	0	0
Pennsylvania	23	23	0	0
Rhode Island	6	6	0	0
South Carolina	10	10	0	0
South Dakota	6	6	0	0
Tennessee	12	12	0	0
Texas	34	34	0	0
Vermont	6	6	0	0
Virginia	12	12	0	0
Washington	8	8	0	0
West Virginia	6	6	0	0
Wisconsin	12	12	0	0
Wyoming	6	6	0	0
Total	531	429	102	0

Abolition Move Defeated by About 206,550—Simpson Is Elected

CITY MARGIN WIDE

Lead Totals 727,254—Queens, Richmond Won by Wilkie

P. R. SYSTEM UPHELD

By EEO KGAN

Franklin D. Roosevelt piled up a plurality of 727,254 in New York City yesterday as voters in record-breaking numbers went to the polls under clear blue skies to elect their President.

The President-elect's three most populous counties in the city—Queens, Richmond and Manhattan—won by Wilkie.

ROOSEVELT WINNER IN MASSACHUSETTS

Indicated Margin Is Below That of 1936—Gallonstall Ahead in a Close Race

NEW JERSEY VOTE GOES TO PRESIDENT

Willkie Margin Cut in Normal Republican Areas—Edison and Barbour Win

DEMOCRATS CARRY STATE BY 230,000

Mead, O'Day, Merritt and Donnelly Carry Seats in New York City Column

ROOSEVELT LOOKS TO DIFFICULT DAYS

But Tells Celebrators That He Will Carry On for the Country 'Just the Same'



WINNERS OF PRESIDENCY AND VICE PRESIDENCY
Franklin Delano Roosevelt Henry Agard Wallace

ROOSEVELT LOOKS TO DIFFICULT DAYS

But Tells Celebrators That He Will Carry On for the Country 'Just the Same'

PRESIDENT TAKES KEYSTONE STATE

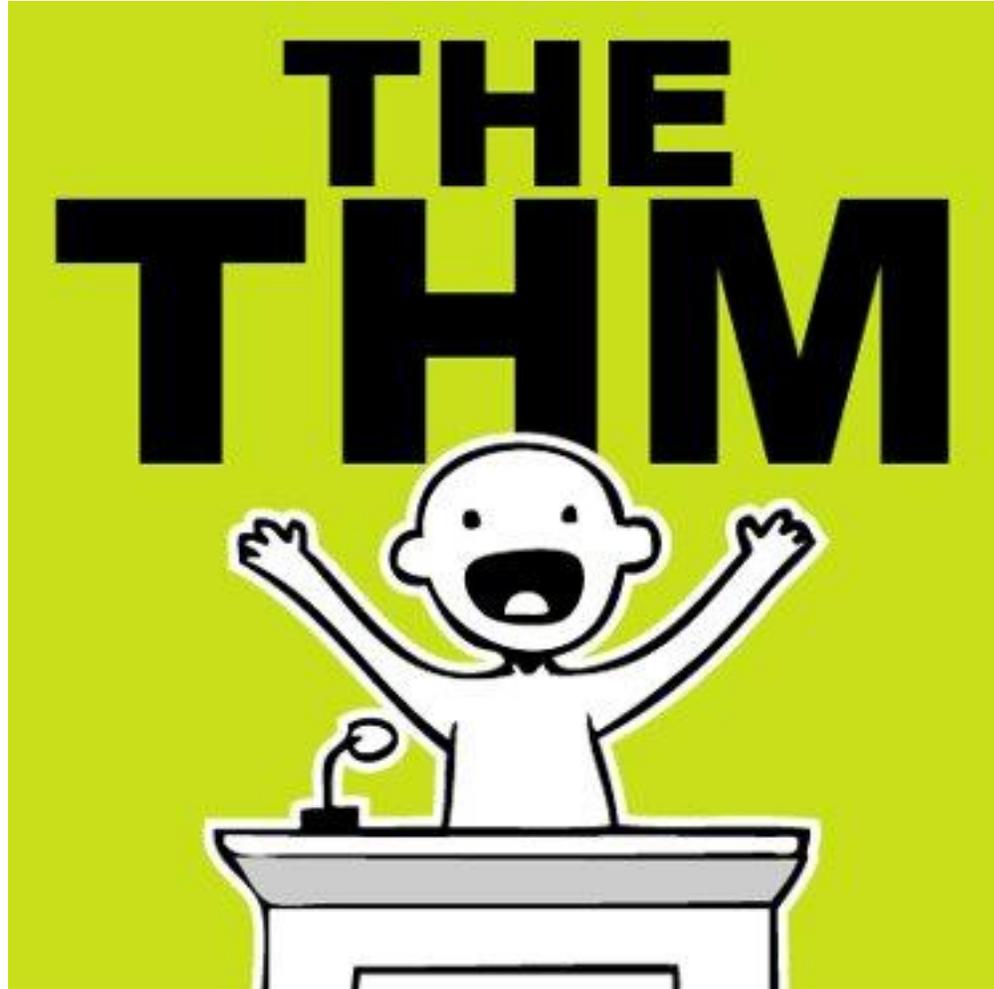
Republican Chairman Concedes Pennsylvania—Gaffney Ahead in Senate Race

ROOSEVELT LOOKS TO DIFFICULT DAYS

But Tells Celebrators That He Will Carry On for the Country 'Just the Same'

ROOSEVELT LOOKS TO DIFFICULT DAYS

But Tells Celebrators That He Will Carry On for the Country 'Just the Same'



**Un campione deve essere
rappresentativo
dell'universo da cui è tratto**



40° Congresso Nazionale SIAN

L'evoluzione professionale e la formazione: ieri, oggi e domani

9 - 11 maggio 2022 📍 Rimini

Società Italiana Area Nefrologica

