Test d'ipotesi

TEST D'IPOTESI

In medicina una delle più utilizzate tecniche inferenziali è quella nota come *test d'ipotesi.*

Tale procedura è particolarmente utile in situazioni in cui noi siamo interessati a *prendere decisioni tra due o più alternative possibili*, piuttosto che alla stima del valore di uno o più parametri.

Ad esempio

- valutare l'efficacia di un nuovo farmaco rispetto al placebo

- valutare se il trattamento chirurgico di un particolare tumore in una data fase allunga la vita dei pazienti rispetto al trattamento chemioterapico

- valutare se l'esposizione a una determinata sostanza chimica è responsabile di un eccesso di tumori

In tali situazioni la valutazione dell'alternativa migliore è finalizzata a decidere quale intervento operare sulla realtà
(scelta del farmaco, tipo di terapia, tipo di intervento preventivo)

TEST D'IPOTESI

• Risultato ottenuto da un'indagine o un esperimento come uno dei possibili risultati di un modello probabilistico

variabilità biologica

variabilità campionaria

errore di misura

• Il test d'ipotesi ci fornisce un criterio per decidere se un campione qualsiasi appartiene alla classe dei 'PROBABILI' o degli 'IMPROBABILI', in base a quel modello probabilistico

probabilità P calcolata tramite un TEST STATISTICO

TEST D'IPOTESI

H_0: IPOTESI NULLA

Tutte le differenze osservate sono delle semplici fluttuazioni casuali

H_1: IPOTESI ALTERNATIVA

Le differenze riscontrate nelle statistiche campionarie rispecchiano una reale differenza nei parametri delle popolazioni corrispondenti

Esempio:

La glicemia dei diabetici italiani è uguale alla glicemia dei diabetici americani

La glicemia dei diabetici italiani è diversa dalla glicemia dei diabetici americani

Dati del campione

Test statistico

P>0,05 = la probabilità che le differenze osservate siano dovute al caso è superiore al 5%

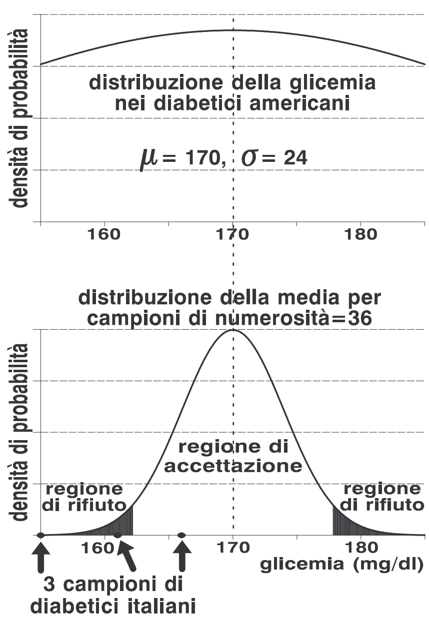
<u>Non</u> rifiuto l'ipotesi nulla (H_0)

tutte le differenze osservate tra i campioni possono essere attribuite al caso

P<0,05 = la probabilità che le differenze osservate siano dovute al caso è inferiore al 5%

Rifiuto l'ipotesi nulla (H_0)

le differenze osservate tra i campioni rispecchiano delle differenze reali tra le popolazioni



1. *Formulazione dell'ipotesi da verificare*

A.(H_0 ipotesi nulla)

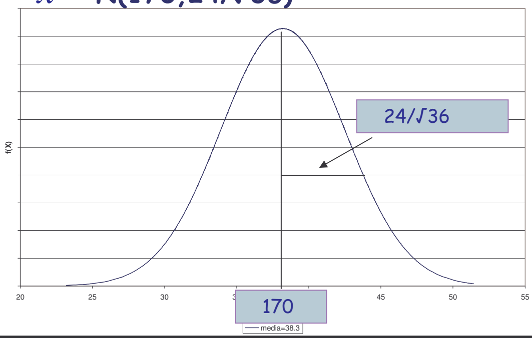
Spiega le differenze osservate come dovute al caso

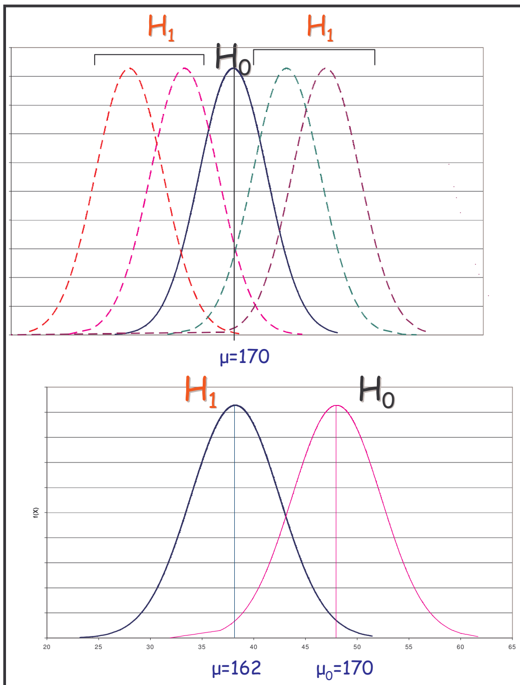
H_0: $\mu=\mu_0=170$ mg/dl

La media della popolazione italiana da cui proviene il campione (μ) è identica alla media della popolazione americana (μ_0)

L'ipotesi è <u>anche</u> una congettura sulla distribuzione campionaria, infatti

$\bar{x}$ **~N(170,24/√36)**





Ipotesi alternativa

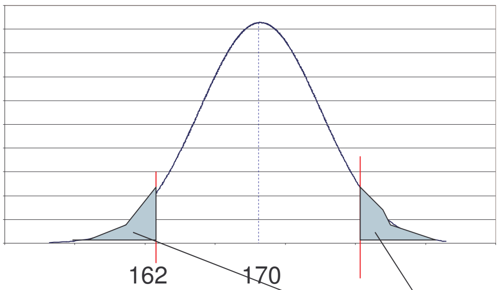
H_1: $\mu\neq\mu_0$

μ**=162**

TEST STATISTICO

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

$$z = \frac{162 - 170}{24 / \sqrt{36}} = -2$$



Rifiuto H_0, la media della glicemia nella popolazione italiana è diversa da quella della pop. americana

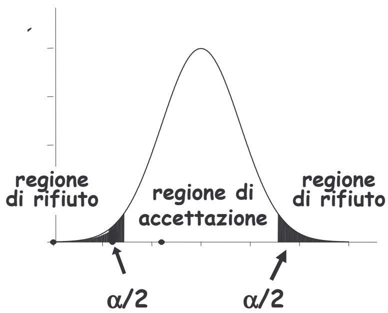
P=0,046 → P <0,05

ALTRE FORMULAZIONI DELL'IPOTESI NULLA

Test bidirezionale

$H_0: \mu = \mu_0$

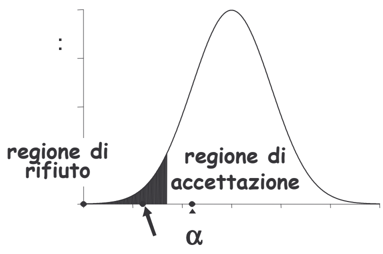
$H_1: \mu \neq \mu_0$



Test unidirezionale

$H_0: \mu = \mu_0$

$H_1: \mu < \mu_0$



Effettuazione di un TEST D'IPOTESI

1. Formulazione H_0 e H_1
2. Scelta del test statistico
3. Calcolo del test statistico

$$\text{test} = \frac{\hat{\theta} - \theta_0}{ES[\hat{\theta}]}$$

dove

$\hat{\theta}$ = stima del parametro di interesse calcolata sui dati campionari

θ_0 = valore del parametro di interesse sotto l'ipotesi nulla H_0

$ES[\hat{\theta}]$ = errore standard dello stimatore calcolato sotto H_0

Effettuazione di un TEST D'IPOTESI-*continua*

4. Test P probabilità di ottenere un risultato come quello osservato o più estremo per motivi casuali

5. Rifiuto (se P è bassa,<0,05) o non rifiuto H_0 (se P non è bassa, >0,05)

TEST D'IPOTESI- confronto media campionaria con media popolazione – campione grande

$H_0: \mu_0 = \mu$

$H_1: \mu_0 \neq \mu$

Test statistico: $$z = \frac{\bar{x} - \mu_0}{e.s.(\bar{x})} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

(vedi esempio diabete)

TEST D'IPOTESI- confronto media campionaria con media popolazione – campione piccolo e σ ignota

$H_0: \mu_0 = \mu$

$H_1: \mu_0 \neq \mu$

Test statistico: $$t = \frac{\bar{x} - \mu_0}{e.s.(\bar{x})} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

t con (n-1) gradi di libertà

Assunzioni:

1- Il campione è stato selezionato casualmente dalla popolazione

2- La variabile è distribuita normalmente nella popolazione

TEST D'IPOTESI- per dati appaiati

H_0:δ=0

H_1:δ≠0

Test statistico: $$t = \frac{\bar{d}-0}{e.s.(d)} = \frac{\bar{d}-0}{\frac{s_d}{\sqrt{n}}}$$

con (n-1) gradi di libertà

Assunzioni:

1- I campioni appaiati sono selezionati casualmente dalla popolazione

2- La popolazione delle differenze è distribuita normalmente

TESTING FOR EQUALITY OF MEANS: PAIRED DATA

Table 7.3. Age at First Word in Months for 10 Children with Cyanotic Heart Disease and Their 10 Siblings

Pair Number	Cyanotic x_1	Sibling x_1	Difference d
1	11.8	9.8	2.0
2	20.8	16.5	4.3
3	14.5	14.5	0.0
4	9.5	15.2	- 5.7
5	13.5	11.8	1.7
6	22.6	12.2	10.4
7	11.1	15.2	- 4.1
8	14.9	15.6	- 0.7
9	16.5	17.2	- 0.7
10	16.5	10.5	6.0

$$\bar{d} = 1.32$$
$$s_d^2 = 22.488$$

Es. test t per dati appaiati

$H_0{:}\delta=0$

$H_1{:}\delta\neq 0$

Test statistico: $$t=\frac{\bar{d}-0}{\frac{s_d}{\sqrt{n}}}=\frac{1{,}32-0}{\frac{4{,}74}{\sqrt{10}}}=0{,}880$$

t con (n-1)=9 gradi di libertà

$t_{(\nu,\alpha/2)}=t_{(9,2.5\%)}=2{,}262$ soglia critica

0,880<2,262 → non rifiuto H_0

L'età alla prima parola per i bambini cianotici sembrerebbe essere la stessa dei loro fratelli

TEST D'IPOTESI- confronto tra medie – campioni grandi

$H_0{:}\mu_1-\mu_2=0$

$H1{:}\mu_1-\mu_2\neq 0$

Test statistico:

$$z=\frac{(\bar{x}_1-\bar{x}_2)-0}{e.s.(\bar{x}_1-\bar{x}_2)}=\frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

Assunzioni:

1- I campioni sono sufficientemente grandi: $n_1\geq 30$ e $n_2\geq 30$

2- I due campioni sono selezionati casualmente e indipendentemente dalla popolazione

TEST D'IPOTESI- confronto tra medie – piccoli campioni e σ ignota

$H_0: \mu_1 - \mu_2 = 0$

$H1: \mu_1 - \mu_2 \neq 0$

Test statistico: $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - 0}{e.s.(\bar{x}_1 - \bar{x}_2)} = \dfrac{\bar{x}_1 - \bar{x}_2}{s_{pooled}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$

$$s^2_{pooled} = \frac{(n_1 - 1) * s_1^2 + (n_2 - 1) * s_2^2}{n_1 + n_2 - 2}$$

t con (n1+n2-2) g.l.

Assunzioni:

1- Le popolazioni da cui sono selezionati i campioni hanno una distribuzione approssimativamente normale della variabile studiata.

2- I campioni sono selezionati casualmente e indipendentemente dalle due popolazioni

3- Le varianze delle due popolazioni sono uguali.

Comparison of birth weights of children born to 15 non-smokers with those of children born to 14 heavy smokers.

Birth weight (Kg)	
Non-smokers	**Heavy smokers**
3.99	3.18
3.79	2.84
3.60	2.90
3.73	3.27
3.21	3.85
3.60	3.52
4.08	3.23
3.61	2.76
3.83	3.60
3.31	3.75
4.13	3.59
3.26	3.63
3.54	2.38
3.51	2.34
2.71	
$\bar{x}$ = 3.5933	$\bar{x}$ =3.2029
s =0.3707	***s*** =0.4927
n= 15	***n***= 14

$$s_{pooled} = \sqrt{\left[\frac{14\times 0.3707^2 + 13\times 0.4927^2}{(15+14-2)}\right]} = 0.4337kg$$

$$t = \frac{(3.5933-3.2029)}{0.4337\sqrt{(1/15+1/14)}} = \frac{0.3904}{0.1612} = 2.42, d.f. = 15+14-2 = 27$$

$t_{0.025;27} = 2.05$

2.42>2.05 ⇨ Rifiuto H0

i pesi dei bimbi delle forti fumatrici presentano una differenza statisticamente significativa rispetto a quelli delle non fumatrici

Con una variabile di tipo quantitativo, qual è il test statistico da effettuare?

Confronto fra soggetti diversi		Misure ripetute sugli stessi soggetti		Confronto fra variabili diverse
↓	↓	↓	↓	↓
2 gruppi	Più di 2 gruppi	2 misurazioni	Più di 2 misurazioni	
↓	↓	↓	↓	
t di Student	ANOVA a 1 criterio	t di Student per dati appaiati	ANOVA per misure ripetute	Regressione e Correlazione

ANOVA = ANalysis Of VAriance (Analisi della varianza)

SIGNIFICATIVITA' STATISTICA e RILEVANZA CLINICA

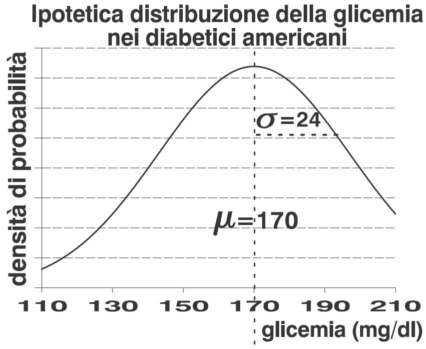
Un'indagine epidemiologica, condotta su un gran numero di persone, ha messo in luce che i fumatori dormono meno della popolazione generale.

La differenza aveva una significatività elevata ($P<0.001$), ovvero ben difficilmente poteva essere attribuita al caso.

La differenza consisteva in 3 minuti di sonno in meno nei fumatori rispetto ai non-fumatori.

L'intervallo di confidenza come test d'ipotesi

L'intervallo di confidenza è una stima intervallare, ma può anche essere considerato un vero e proprio test d'ipotesi.



IPOTESI NULLA: Nei diabetici italiani la distribuzione della glicemia è uguale a quella dei diabetici americani. Pertanto la media per campioni di numerosità 36 si distribuisce con:

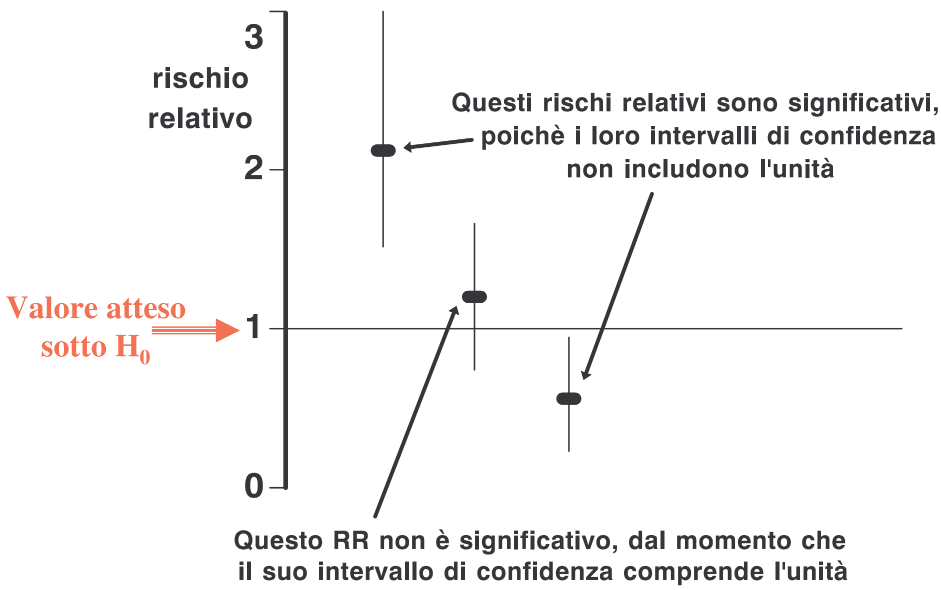
$\mu = 170$

$\sigma/\sqrt{n} = 24/\sqrt{36} = 4$

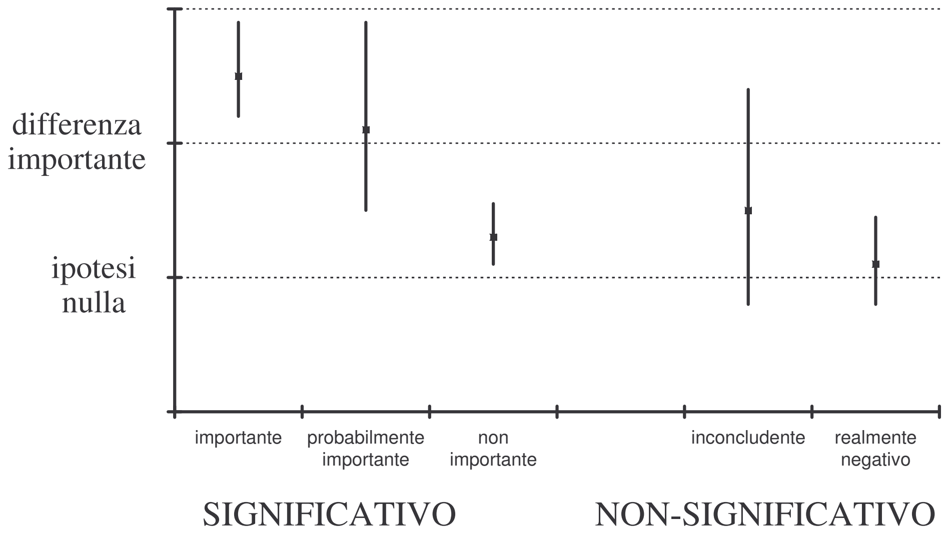
I primi due intervalli di confidenza sono significativi perché non contengono 170 mg/dl (valore atteso sotto H0)

stime intervallari di μ

155±1,96*4	147,2 —— 162,8
161±1,96*4	153,2 —— 168,8
166±1,96*4	158,2 —— 173,8



Valutazione clinico-statistica di un intervallo di confidenza



“Overemphasis on hypothesis testing - and the use of P values to dichotomise significant or non-significant results - has detracted from more useful approaches to interpreting study results, such as estimation and confidence intervals.

In medical studies investigators are usually interested in determining the size of difference of a measured outcome between groups, rather than a simple indication of whether or not it is statistically significant ...

Confidence intervals, if appropriate to the type of study, should be used for major findings in both the main text of a paper and its abstract.”

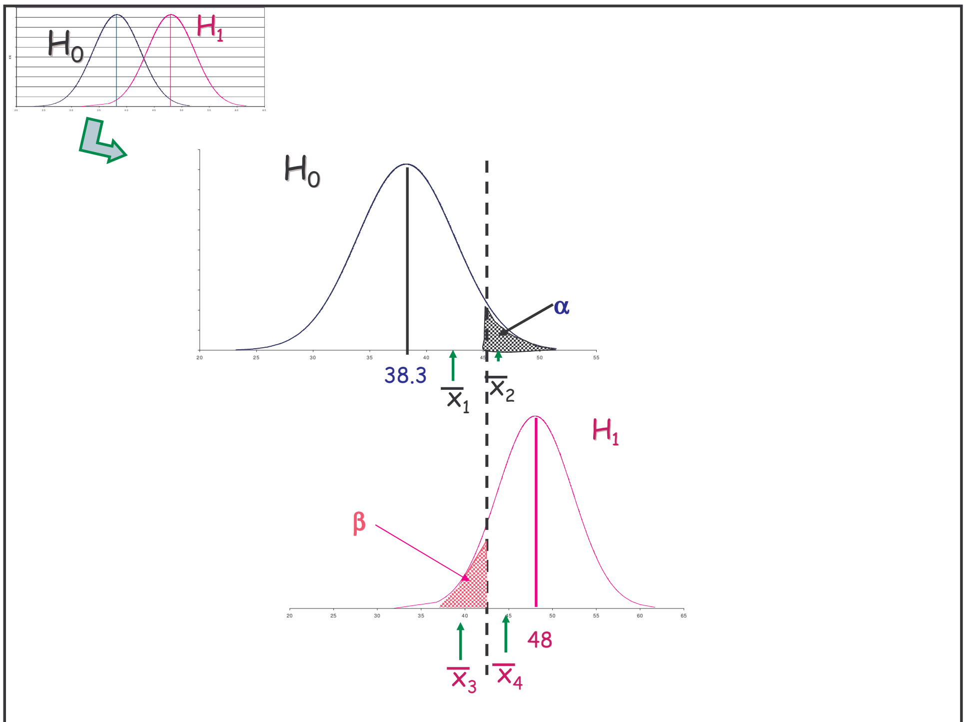
Gardner MJ, Altman DG (1986) Confidence intervals rather than P values: estimation rather than hypothesis testing. British Medical Journal, 292: 746-750

TEST D'IPOTESI ED ERRORI DELLA LOGICA DECISIONALE

Quando una decisione è presa sulla base del risultato di un test di ipotesi, si possono commettere <u>due tipi di errore</u>:

- ERRORE DI I° TIPO (α): rifiutare l'ipotesi nulla H_0 quando essa è vera
- ERRORE DI II° TIPO (β): non rifiutare l'ipotesi nulla H_0 quando essa è falsa

		vero stato di natura	
		H_0 vera	H_0 non vera
decisione sulla base del test	non rifiuto H_0	**decisione corretta (1 - α)**	**errore di II° tipo (β)**
	rifiuto H_0	**errore di I° tipo (α)**	**decisione corretta (1 - β)**



Potenza di un test

La potenza di un test è la probabilità di rifiutare H_0 quando essa è falsa ($1-\beta$)

Nel nostro esempio:

$$\beta = P(\bar{x} < 45.4 \mid \mu = 48) \rightarrow z = \frac{45.4 - 48}{4.33} = -0.60$$

$$\beta = P(Z < -0.60) = 0.27$$

POTENZA DEL TEST=($1-\beta$)=0.73

POTENZA di un test = 1- beta = 1 - P(errore del II tipo)

E' la probabilità che un test statistico ha di falsificare l'ipotesi nulla quando l'ipotesi nulla è effettivamente falsa.

In altre parole, la Potenza di un test è la sua capacità di cogliere delle differenze, quando queste differenze esistono.

Il test statistico è costruito in modo da mantenere costante il livello di significatività, indipendentemente dalla numerosità campionaria. Ma questo risultato viene raggiunto a spese della potenza del test, che aumenta all'aumentare della numerosità campionaria.

La POTENZA di un test dipende:

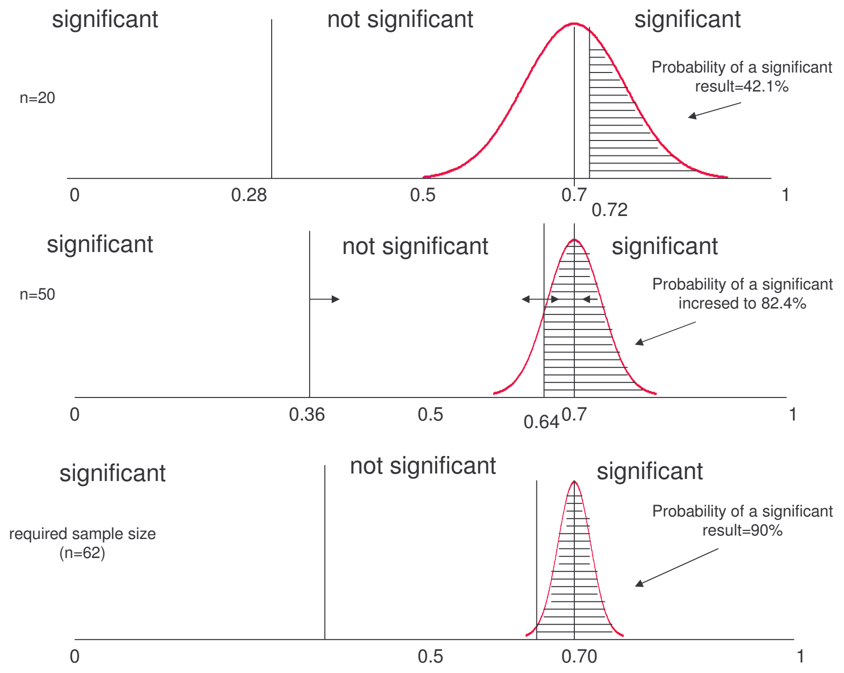
1) dalla numerosità del campione

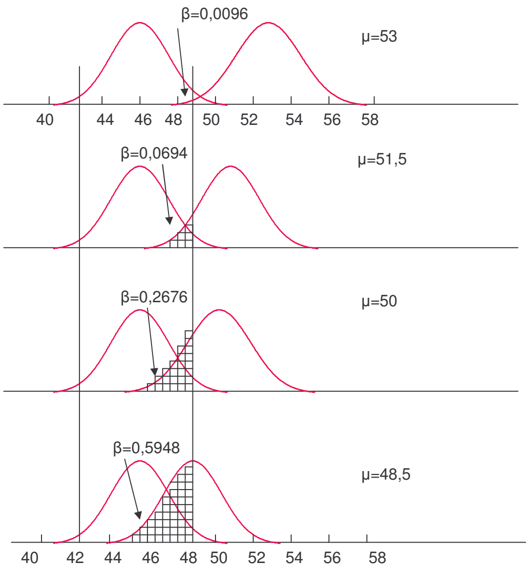
2) dalla variabilità del fenomeno in studio

3) dalla differenza minima che si vuole mettere in evidenza

4) dal livello di significatività adottato.

Il modo principale per raggiungere un'adeguata potenza è pianificare un'adeguata numerosità campionaria nel protocollo dello studio.





β=0,0096
μ=53
40
44
46
48
50
52
54
56
58
β=0,0694
μ=51,5
β=0,2676
μ=50
β=0,5948
μ=48,5
40
42
44
46
48
50
52
54
56
58