LEZIONI DI STATISTICA MEDICA

Test d'ipotesi

-Test Z

Sezione di Epidemiologia & Statistica Medica
Università degli Studi di Verona

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)

ESEMPIO

Si è stabilito sperimentalmente su un gran numero di pazienti affetti da una determinata malattia che il tempo medio di sopravvivenza dalla diagnosi è di 38.3 mesi con d.s.= 43.3.

Un campione casuale di 100 pazienti con prima diagnosi viene trattato con una nuova tecnica terapeutica. Alla fine della sperimentazione il tempo medio di sopravvivenza per questo gruppo di pazienti risulta essere 46.9 mesi.

1. Formulazione dell'ipotesi da verificare

A. (H_0 – ipotesi nulla)

Spiega le differenze osservate come dovute al caso

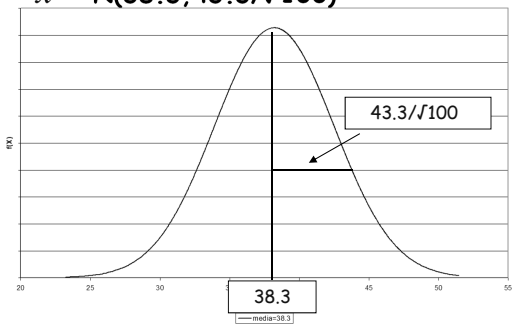
H_0: $\mu=\mu_0$=38.3 mesi

La media della popolazione da cui proviene il campione trattato con il nuovo farmaco (μ) è identica alla media della popolazione dei soggetti trattati in modo tradizionale (μ_0)

Ipotesi semplice: specifica tutti i parametri della distribuzione in modo univoco

L'ipotesi è anche una congettura sulla distribuzione campionaria dello stimatore d'interesse

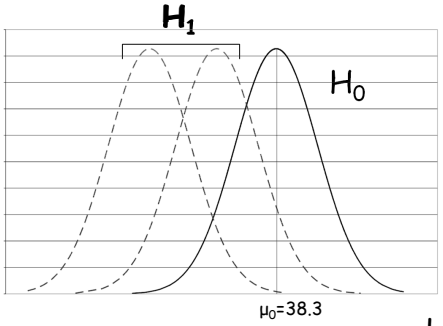
$\bar{x}$ **~N(38.3,43.3/√100)**



2. Formulazione dell'ipotesi alternativa – H_1

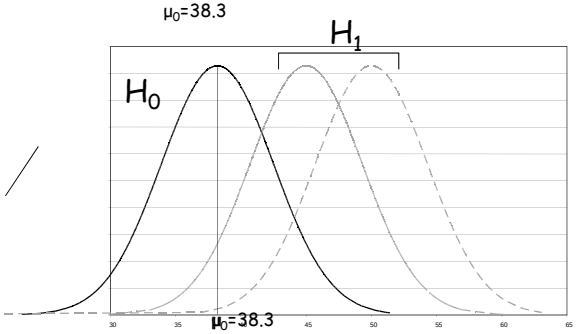
Casi possibili

1. H_1: $\mu < \mu_0$;

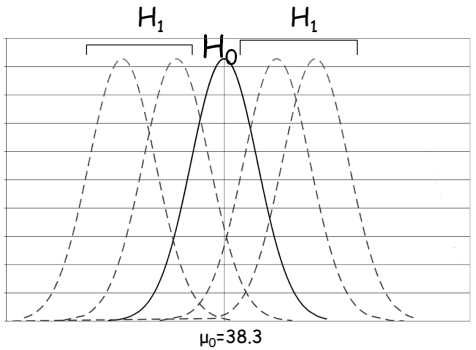


Spiega le differenze osservate come dovute alla variabile in studio (trattamento, esposizione a un fattore di rischio,….)

2. H_1: $\mu > \mu_0$;

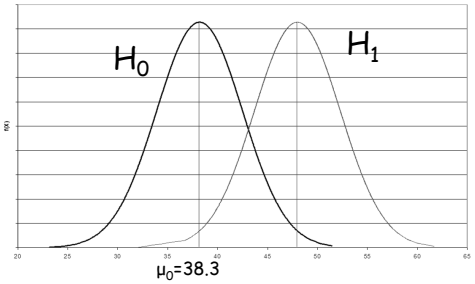


Ipotesi composta: non specifica in modo univoco i parametri della distribuzione



3. H_1: $\mu \neq \mu_0$

Le ipotesi possono essere:
- ***unidirezionali*** (1,2,4)
- ***bidirezionali*** (3)



4. H_1: μ = k

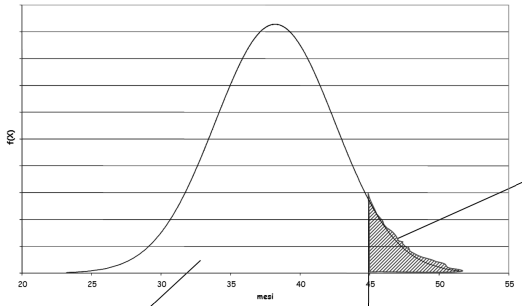
Per semplicità supponiamo che

H_1: μ **=48**

3 Definizione della regione critica

- Lo spazio sotto l'ipotesi nulla viene ripartito in 2 regioni
  1. **Regione critica (α):** individua i valori della stima (del test statistico) per cui verrà rifiutata l'ipotesi nulla
  2. **Regione di accettazione (1-α):** valori della media per cui sarà accettata H_0

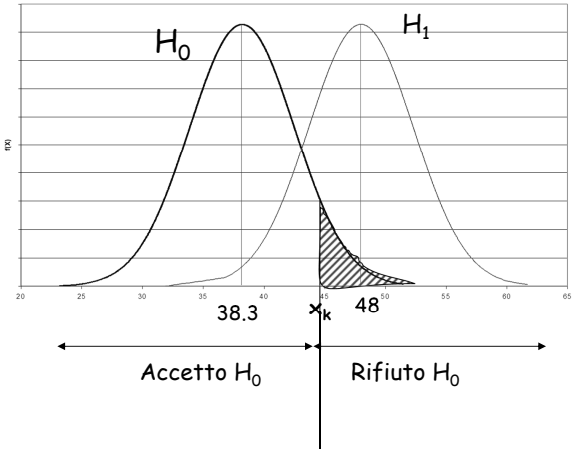
Per convenzione:
α= 0.05



α

$1-\alpha$

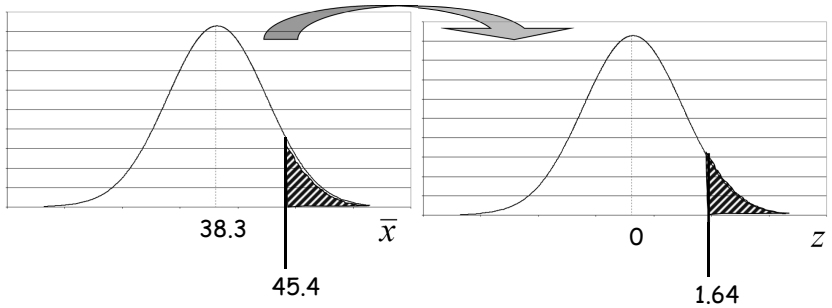
Soglia critica



• Il valore della soglia critica può essere immediatamente determinato in termini di media di surv. o in termini di *Z*

La soglia critica nell'esempio è tale per cui: $P(\overline{X} > \overline{x}_k \mid \mu = 38.3) = 0.05 \rightarrow 1.64 = \dfrac{\overline{x}_k - 38.3}{4.33} \rightarrow \overline{x}_k = 45.4$

TEST STATISTICO $\Rightarrow$ $z = \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$



$media = 38.3$ $\Rightarrow$ $media = 0$

$soglia = 45.4$ $\qquad$ $soglia = 1.64$

4. Regola di decisione

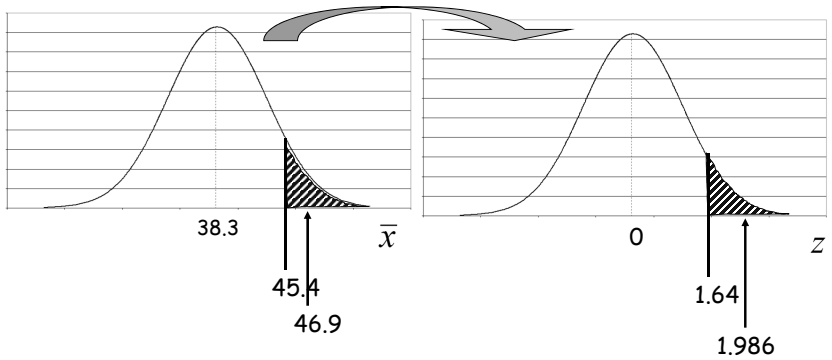
- Se il valore del "*test statistico z*" (media campionaria) <u>supera</u> il valore della <u>soglia critica si rifiuta H_0</u> a favore dell'ipotesi alternativa
- Se il valore del test statistico cade nella regione di accettazione di <u>H_0</u> <u>non si hanno elementi per rifiutare H_0</u>

Nel nostro esempio: $z = \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} = \dfrac{46.9 - 38.3}{4.33} = 1.986$

$\Rightarrow$ Il valore di Z calcolato nel campione (=1.986) supera il valore della soglia critica (=1.64) $\Rightarrow$ Si rifiuta H_0

equivalentemente

La media campionaria (=46.9) supera il valore della soglia critica (45.4) $\Rightarrow$ Si rifiuta H_0

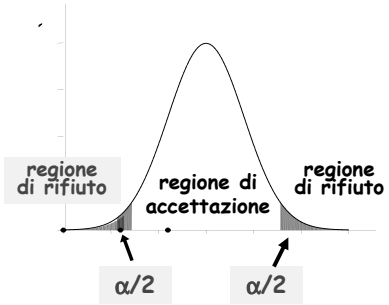


RELAZIONE TRA IPOTESI ALTERNATIVA E REGIONE DI ACCETTAZIONE

Test bidirezionale

$H_0: \mu = \mu_0$

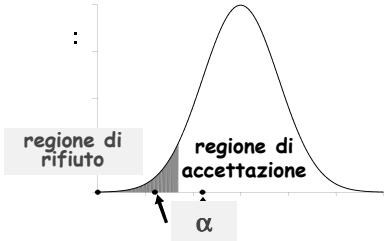
$H_1: \mu \neq \mu_0$



Test unidirezionale

$H_0: \mu = \mu_0$

$H_1: \mu < \mu_0$



summary: TEST PER IL CONFRONTO DI UNA MEDIA CON UN PARAMETRO

1. Ipotesi nulla: $H_0: \mu = \mu_0$

2. Ipotesi alternativa: $H_1: \begin{cases} \mu \neq \mu_0 \\ \mu > \mu_0 \\ \mu < \mu_0 \end{cases}$

3. Test: $z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

4. Soglia critica: test a una coda $\Longrightarrow$ z_α oppure $-z_\alpha$

test a due code $\Longrightarrow$ $\pm z_{\alpha/2}$

5. Regola di decisione test a una coda $\Longrightarrow$ $z_{oss} > z_\alpha$ oppure $z_{oss} < -z_\alpha$

test a due code $\Longrightarrow$ $z_{oss} > |z_{\alpha/2}|$

Esercizio

- **L'emoglobina è un componente dei globuli rossi che porta l'ossigeno dai polmoni ai vari tessuti. Valori inferiori ai 12 g/dl sono suggestivi di anemia.Un ricercatore sospetta che i bambini che vivono in una nazione in via di sviluppo abbiano livelli di emoglobina <12. Il ricercatore campiona 35 bambini che hanno una media di emoglobina =11.2 g/dl. Valutate l'ipotesi del ricercatore sapendo che l'emoglobina ha una distribuzione normale con media=12 g/dl e d.s=1.6 g/dl**

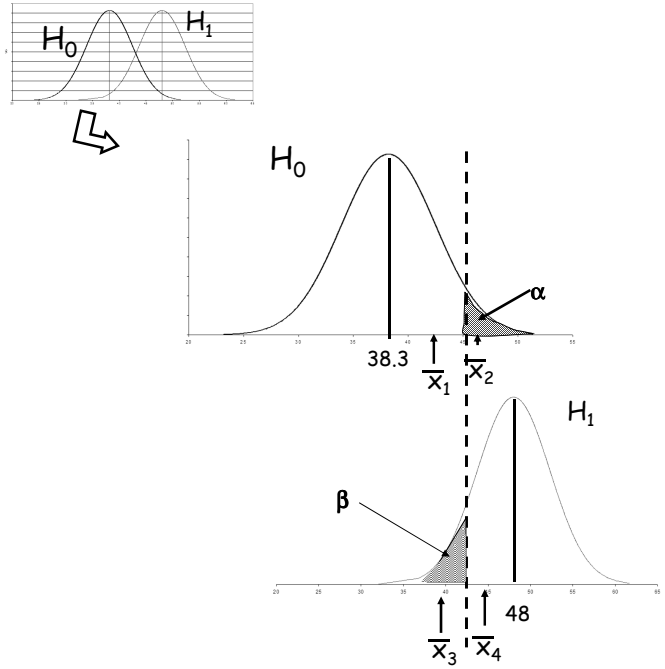
- **Il valore di colesterolo serico dei ragazzi degli USA è μ = 175 mg /ml con s = 50. Su un CCS di 39 ragazzi il cui padre aveva avuto una storia di infarto, la media di colesterolo era di 195 mg/ml. Utilizzando un test a due code, valutate se il campione studito è signficativamente diffferente dall'atteso.**

- **La prevalenza di asma è del 5%. Su un campione di 1000 bambini atopici, 98 avevano l'asma. Secondo voi esiste un associazione tra asma e atopia?**

Nell'effettuare un test, possiamo compiere 2 tipi di errori:

1. Errore di primo tipo: rifiutare H_0 quando è vera. La probabilità di commettere un errore del primo tipo è α
2. Errore di secondo tipo: accettare H_0 quando è vera H_1. La probabilità di commettere un errore del secondo tipo è β

	H_0 vera	H_0 falsa
Non rifiuto H_0	Decisione corretta $1-\alpha$	Errore di II tipo β
Rifiuto H_0	Errore di I tipo α	Decisione corretta $1-\beta$



SIGNIFICATIVITA' STATISTICA

Confronto tra il valore osservato del p-value e il cut-off di significatività (0.05):

p-value ≥ 0.05 → la differenza osservata NON è statisticamente significativa ⇒ NON rifiuto l'ipotesi nulla (H0) (la differenza osservata sembra essere dovuta al caso)

p-value < 0.05 → la differenza osservata è statisticamente significativa ⇒ rifiuto l'ipotesi nulla (H0) (la differenza osservata sembra essere dovuta ad altri fattori: trattamento, fattori di rischio, …)