LEZIONI DI STATISTICA MEDICA

Prof. SIMONE ACCORDINI

Lezione n.12

- Test statistico

Sezione di Epidemiologia & Statistica Medica
Università degli Studi di Verona

IPOTESI SCIENTIFICA

Affermazione che si può sottoporre a verifica ovvero che si può tentare di falsificare. Con una procedura che comporta delle misurazioni si può cercare di dimostrare che l'ipotesi non è vera.

Un'ipotesi scientifica viene ritenuta vera finché non si dimostra il contrario.

IPOTESI STATISTICA

Affermazione su una determinata caratteristica (*valore di uno o più parametri, forma della distribuzione, modello probabilistico*) di una o più popolazioni, che si cerca di supportare o di rifiutare sulla base delle informazioni campionarie.

TEST D'IPOTESI

- una delle tecniche inferenziali più utilizzate in medicina
- utile in situazioni nelle quali si è interessati a prendere decisioni tra due o più alternative possibili sulla base delle osservazioni campionarie

esempi:

- *valutare l'efficacia di un nuovo farmaco rispetto ad un farmaco standard*
- *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 ad 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*).

Confronto della media tra due popolazioni

Esempio:

Cockburn et al (1980) riportano una ricerca clinica per la prevenzione dell'ipocalcemia infantile, nella quale donne in gravidanza che ricevevano un supplemento di vitamina D venivano messe a confronto con donne non trattate.

Calcemia del bambino misurata 6 giorni dopo la nascita:

	Numero di pazienti	Media (mg /100 ml)	DS (mg /100 ml)
Vitamina D **(D_1)**	**233**	**9.36**	**1.15**
Controllo **(D_0)**	**394**	**9.01**	**1.33**

La differenza osservata è dovuta al caso oppure alla vitamina D?

Test per il confronto della media tra due popolazioni (campioni indipendenti)

TEST Z e TEST T DI STUDENT

GRUPPO NON ESPOSTO (D_0):

$x_{01}, x_{02}, \dots, x_{0n_0}$ → n_0 **determinazioni indipendenti della v.c.** $X_0 \sim Norm(\mu_0, \sigma_0)$

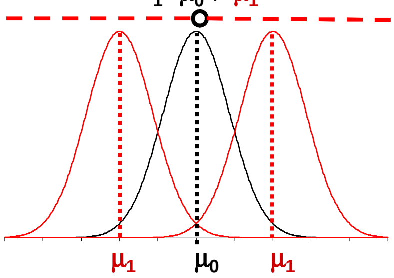
GRUPPO ESPOSTO (D_1):

$x_{11}, x_{12}, \dots, x_{1n_1}$ → n_1 **determinazioni indipendenti della v.c.** $X_1 \sim Norm(\mu_1, \sigma_1)$

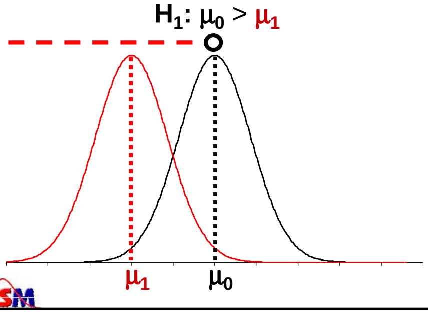
I° STEP: definire il sistema di ipotesi da verificare

$$\begin{cases} H_0: \mu_0 = \mu_1 & \text{(ipotesi nulla)} \\ H_1: \dots & \text{(ipotesi alternativa)} \end{cases}$$

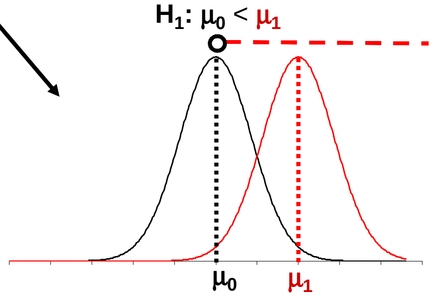
$H_1: \mu_0 \neq \mu_1$



$H_1: \mu_0 > \mu_1$



$H_1: \mu_0 < \mu_1$



$$\begin{cases} \mathbf{H_0: \mu_0 = \mu_1} & \text{(ipotesi nulla)} \\ \mathbf{H_1: \mu_0 \neq \mu_1} & \text{(ipotesi alternativa)} \end{cases}$$

L'ipotesi nulla (H_0) è l'ipotesi che spiega la differenza osservata tra le medie come dovute al caso

$\Rightarrow$ <u>variazioni casuali</u>

L'ipotesi alternativa (H_1) è l'ipotesi che spiega la differenza osservata tra le medie come dovuta alla variabile in studio (trattamento, fattori di rischio, …)

$\Rightarrow$ <u>variazioni sistematiche</u>

II° STEP: calcolare la statistica test

$$\begin{cases} \mathbf{H_0: \mu_0 = \mu_1} & \text{(ipotesi nulla)} \\ \mathbf{H_1: \mu_0 \neq \mu_1} & \text{(ipotesi alternativa)} \end{cases}$$

differenza tra le medie = stima dell'effetto

$$\boxed{\frac{\overline{X}_1 - \overline{X}_0}{ES[\overline{X}_1 - \overline{X}_0]}}$$

errore standard della differenza = misura della precisione della stima

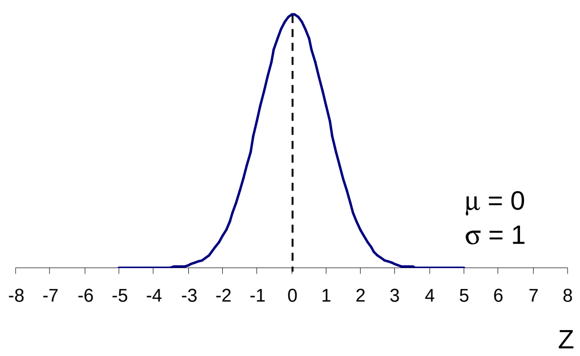
$$Z = \frac{\bar{X}_1 - \bar{X}_0}{\sigma \sqrt{1/n_1 + 1/n_0}}$$

$\sigma_1 = \sigma_0 = \sigma$ nota

$$Z = \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{\sigma_1^2/n_1 + \sigma_0^2/n_0}}$$

$\sigma_1 \neq \sigma_0$ note

Se è vera l'ipotesi nulla (H_0): $Z \sim N(0,1)$



$$T = \frac{\bar{X}_1 - \bar{X}_0}{S_p \sqrt{1/n_1 + 1/n_0}}$$

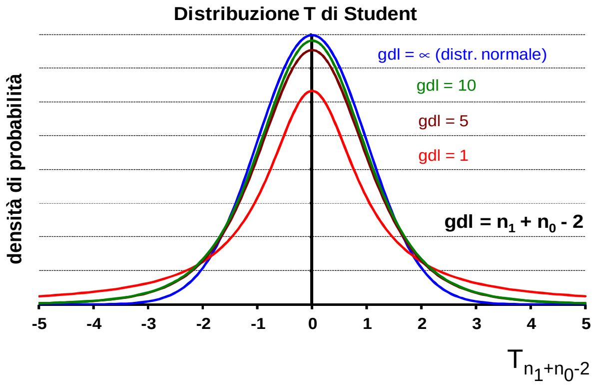
$\sigma_1 = \sigma_0 = \sigma$ ignota

$$T = \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{S_1^2/n_1 + S_0^2/n_0}}$$

problema di Behrens-Fisher

$\sigma_1 \neq \sigma_0$ ignote

Se è vera l'ipotesi nulla (H_0): $T \sim t_{n_1+n_0-2}$



$$T = \frac{\bar{X}_1 - \bar{X}_0}{S_p \sqrt{1/n_1 + 1/n_0}}$$

$\sigma_1 = \sigma_0 = \sigma$ ignota

STIMA

$$s_p = \sqrt{\frac{(n_1 - 1)\, s_1^2 + (n_0 - 1)\, s_0^2}{(n_1 - 1) + (n_0 - 1)}} = \sqrt{\frac{\Sigma\,(x_{1i} - \bar{x}_1)^2 + \Sigma\,(x_{0i} - \bar{x}_0)^2}{n_1 + n_0 - 2}}$$

deviazione standard 'pooled'

Esempio (calcemia):

	Numero di pazienti	Media (mg/100 ml)	DS (mg/100 ml)
Vitamina D **(D_1)**	**233**	**9.36**	**1.15**
Controllo **(D_0)**	**394**	**9.01**	**1.33**

$$\begin{cases} H_0: \mu_0 = \mu_1 \\ H_1: \mu_0 \neq \mu_1 \end{cases}$$

I° STEP

$$s_p = \sqrt{\frac{(n_1 - 1)\, s_1^2 + (n_0 - 1)\, s_0^2}{(n_1 - 1) + (n_0 - 1)}} = \sqrt{\frac{232 * 1.15^2 + 393 * 1.33^2}{625}} = 1.27$$

Esempio (calcemia):

	Numero di pazienti	Media (mg/100 ml)	DS (mg/100 ml)
Vitamina D (D_1)	233	9.36	1.15
Controllo (D_0)	394	9.01	1.33

$$\begin{cases} H_0: \mu_0 = \mu_1 \\ H_1: \mu_0 \neq \mu_1 \end{cases}$$

II° STEP

$$t_{oss} = \frac{\bar{x}_1 - \bar{x}_0}{s_p \sqrt{1/n_1 + 1/n_0}} = \frac{9.36 - 9.01}{1.27 \sqrt{1/233 + 1/394}} = 3.33$$

III° STEP: definire la regione di rifiuto dell'ipotesi nulla (H_0)

suddividere lo spazio campionario della STATISTICA TEST in due parti:

• regione che individua i valori della statistica test tale che H_0 viene rifiutata

→ **REGIONE DI RIFIUTO DI H_0**

• regione che individua i valori della statistica test tale che H_0 non viene rifiutata

→ **REGIONE DI NON RIFIUTO DI H_0**

IV° STEP: formulazione della decisione riguardante l'ipotesi nulla (H_0)

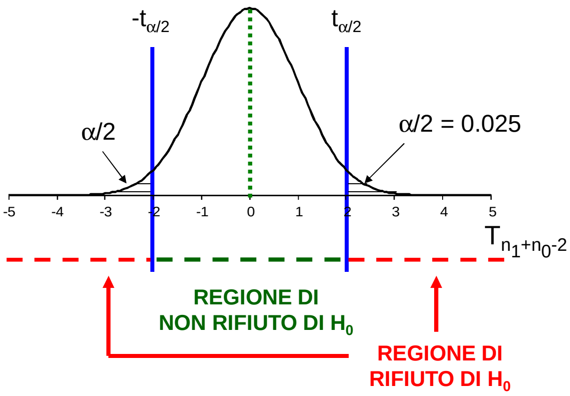
• se il valore della statistica test cade nella regione di rifiuto di H_0 si rifiuta H_0

• se il valore della statistica cade nella regione di non rifiuto di H_0 non si rifiuta H_0

$$\begin{cases} H_0: \mu_1 = \mu_0 \\ H_1: \mu_1 \neq \mu_0 \end{cases}$$

- si fissa il **livello di significatività α**
- si individua la **soglia critica $t_{\alpha/2}$** tale che:

$$\Pr(T_{n_1+n_0-2} < -t_{\alpha/2} \cup T_{n_1+n_0-2} > t_{\alpha/2} \mid H_0 \text{ è vera}) = \alpha$$

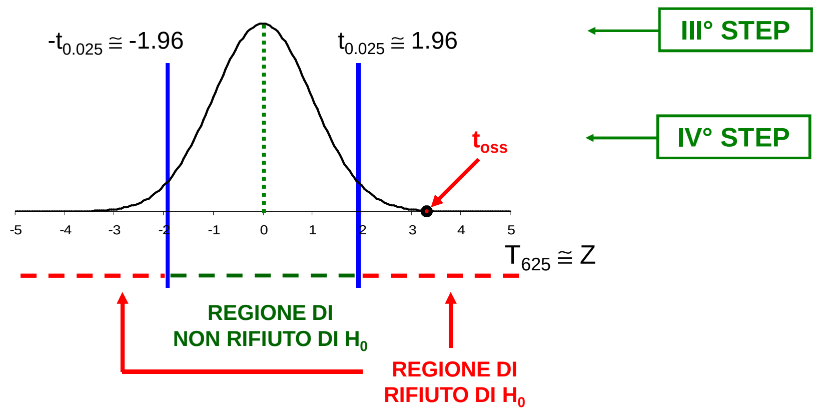


Soglia critica $t_{0.025}$ della distribuzione t di Student

gdl	$t_{0.025}$	gdl	$t_{0.025}$	gdl	$t_{0.025}$	gdl	$t_{0.025}$
1	12,706	26	2,056	51	2,008	76	1,992
2	4,303	27	2,052	52	2,007	77	1,991
3	3,182	28	2,048	53	2,006	78	1,991
4	2,776	29	2,045	54	2,005	79	1,990
5	2,571	30	2,042	55	2,004	80	1,990
6	2,447	31	2,040	56	2,003	81	1,990
7	2,365	32	2,037	57	2,002	82	1,989
8	2,306	33	2,035	58	2,002	83	1,989
9	2,262	34	2,032	59	2,001	84	1,989
10	2,228	35	2,030	60	2,000	85	1,988
11	2,201	36	2,028	61	2,000	86	1,988
12	2,179	37	2,026	62	1,999	87	1,988
13	2,160	38	2,024	63	1,998	88	1,987
14	2,145	39	2,023	64	1,998	89	1,987
15	2,131	40	2,021	65	1,997	90	1,987
16	2,120	41	2,020	66	1,997	91	1,986
17	2,110	42	2,018	67	1,996	92	1,986
18	2,101	43	2,017	68	1,995	93	1,986
19	2,093	44	2,015	69	1,995	94	1,986
20	2,086	45	2,014	70	1,994	95	1,985
21	2,080	46	2,013	71	1,994	96	1,985
22	2,074	47	2,012	72	1,993	97	1,985
23	2,069	48	2,011	73	1,993	98	1,984
24	2,064	49	2,010	74	1,993	99	1,984
25	2,060	50	2,009	75	1,992	100	1,984

se gdl >100 $\Rightarrow$ distribuzione approssimativamente normale

Esempio (calcemia):



si rifiuta l'ipotesi nulla (H_0)

→ la differenza osservata sembra essere dovuta alla vitamina D

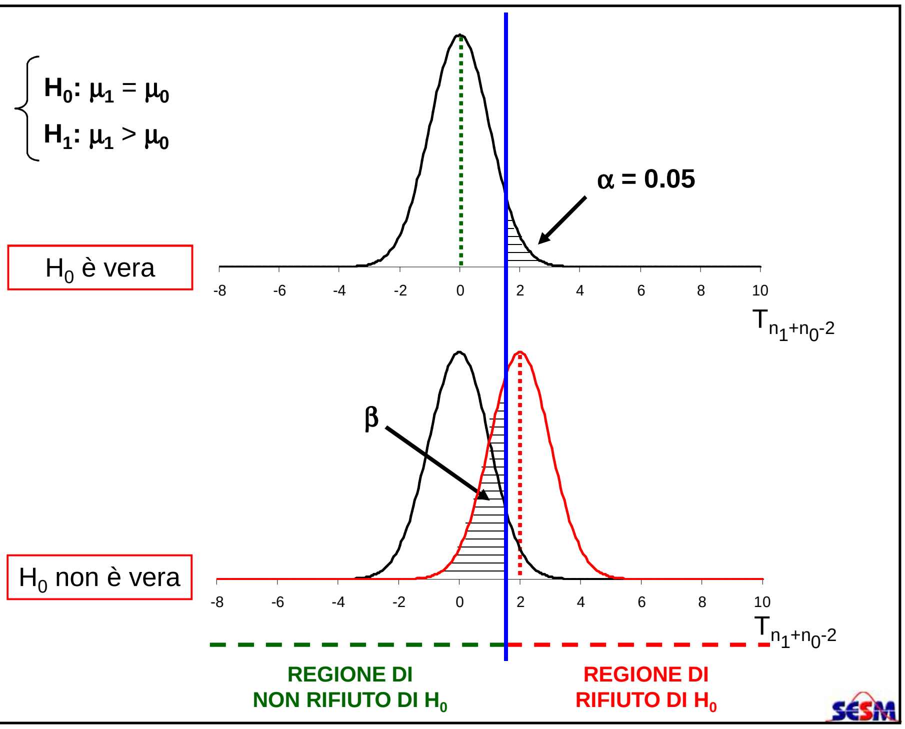
Test d'ipotesi ed errori della logica decisionale

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

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

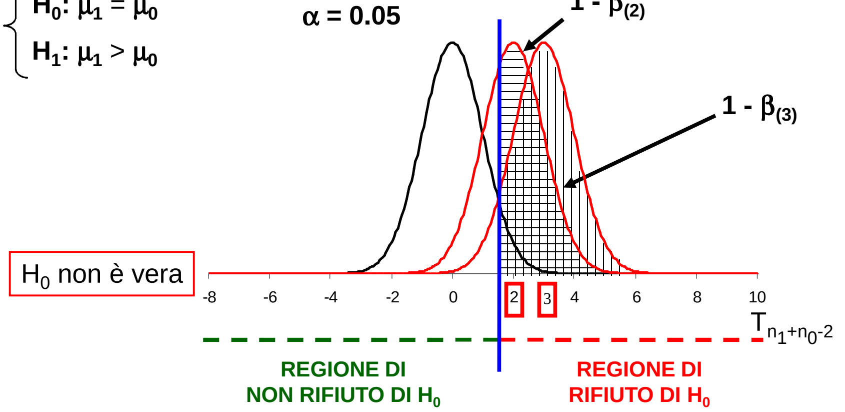
decisione sulla base del test	vero stato di natura: H_0 vera	vero stato di natura: H_0 non vera
non rifiuto H_0	**decisione corretta ($1 - \alpha$)**	**errore di II° tipo (β)**
rifiuto H_0	**errore di I° tipo (α)**	**decisione corretta ($1 - \beta$ = potenza)**

$$\begin{cases} H_0: \mu_1 = \mu_0 \\ H_1: \mu_1 > \mu_0 \end{cases}$$

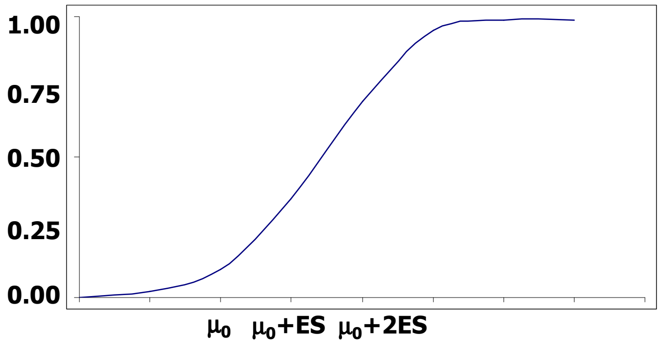


Potenza di un test d'ipotesi

$$\begin{cases} H_0: \mu_1 = \mu_0 \\ H_1: \mu_1 > \mu_0 \end{cases}$$



La funzione POTENZA assume valori tanto maggiori quanto più il parametro specificato dall'ipotesi alternativa è lontano da quello specificato dall'ipotesi nulla

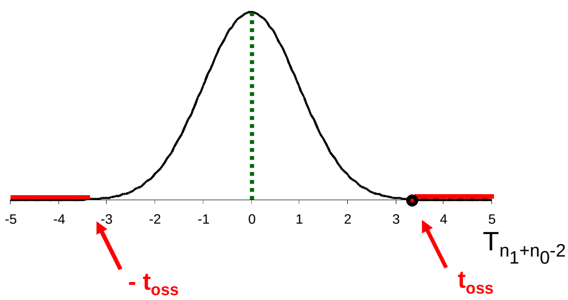


TEST DI SIGNIFICATIVITA' STATISTICA

Il risultato inferenziale del test viene rappresentato tramite una misura continua del livello di consistenza dei dati con l'ipotesi nulla, ovvero di quanto sono plausibili i dati sotto H_0 $\rightarrow$ **P-VALUE = livello di significatività osservato**

ipotesi alternativa bilaterale

$$\begin{cases} H_0: \mu_0 = \mu_1 \\ H_1: \mu_0 \neq \mu_1 \end{cases}$$



$$\boxed{p\text{-}value = 2 * \Pr(T_{n_1+n_0-2} \geq t_{oss} \mid H_0 \text{ è vera})}$$

Esempio (calcemia):

	Numero di pazienti	Media (mg/100 ml)	DS (mg/100 ml)
Vitamina D (D_1)	233	9.36	1.15
Controllo (D_0)	394	9.01	1.33

$$\begin{cases} H_0: \mu_0 = \mu_1 \\ H_1: \mu_0 \neq \mu_1 \end{cases}$$

$$t_{oss} = \frac{\bar{x}_1 - \bar{x}_0}{s_p \sqrt{1/n_1 + 1/n_0}} = 3.33$$

p-value = $2 * Pr(T_{625} \cong Z \geq 3.33 \mid H_0 \text{ è vera}) = 2 * 0.000434 = 0.0009$

i dati non supportano l'ipotesi nulla (H_0)

→ la differenza osservata sembra essere dovuta alla vitamina D

Test di significatività statistica e test d'ipotesi

- Il p-value è una probabilità (valore tra 0 e 1) = livello di significatività osservato

  → p-value piccolo ⇒ scarso livello di compatibilità dei dati con H_0

  → p-value elevato ⇒ grande livello di compatibilità dei dati con H_0

- Il livello di significatività α (i.e. errore di I° tipo) e l'errore di II° tipo (β) derivano dall'uso del test come criterio di decisione

  ⇒ non riguardano il test di significatività statistica ma solo il test d'ipotesi

- Se si utilizza un cut-off prefissato per il livello di significatività α (usualmente 0.05)

  ⇒ <u>si effettua un test d'ipotesi e non di significatività statistica</u>

  → p-value < α ⇒ risultato '*significativo*' e si rifiuta H_0

  → p-value ≥ α ⇒ risultato '*non significativo*' e non si rifiuta H_0

Esempio (calcemia):

	Numero di pazienti	Media (mg/100 ml)	DS (mg/100 ml)
Vitamina D **(D_1)**	**233**	**9.36**	**1.15**
Controllo **(D_0)**	**394**	**9.01**	**1.33**

p-value = 0.0009 < 0.05 $\Leftrightarrow$ **t_{oss} si trova nella regione di rifiuto di H_0**

la differenza osservata è **statisticamente significativa**

$\Rightarrow$ **rifiuto l'ipotesi nulla (H_0)**

(la differenza osservata sembra essere dovuta alla vitamina D)

SESM

Esercizio:

E' stato condotto uno studio per valutare l'effetto del fumo materno in gravidanza sulla densità ossea dei neonati.

Il contenuto medio di minerali in 77 neonati da madri fumatrici era di 0.098 g/cm^3 (d.s. = 0.026 g/cm^3).

In un campione di 161 neonati da madri non fumatrici, il contenuto medio era di 0.095 g/cm^3 (d.s. = 0.025 g/cm^3).

Stabilite se c'è una differenza statisticamente significativa tra i due campioni. Assumete che il contenuto di minerali abbia distribuzione normale e che la deviazione standard sia uguale nei due gruppi ma ignota.

SESM

Esercizio (soluzione):

Definire il sistema d'ipotesi:
$$\begin{cases} H_0: \mu_1 = \mu_0 \\ H_1: \mu_1 \neq \mu_0 \end{cases}$$

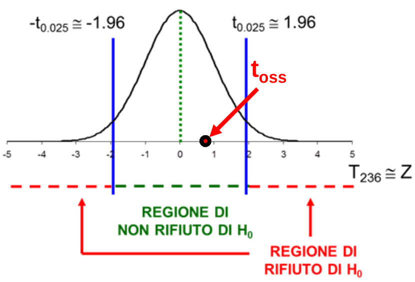
Calcolare la statistica test:

$$s_p = \sqrt{\frac{(n_1 - 1)\, s_1^2 + (n_0 - 1)\, s_0^2}{(n_1 - 1) + (n_0 - 1)}} = \sqrt{\frac{76 * 0.026^2 + 160 * 0.025^2}{236}} = 0.0253$$

$$t_{oss} = \frac{\overline{x}_1 - \overline{x}_0}{s_p \sqrt{1/n_1 + 1/n_0}} = \frac{0.098 - 0.095}{0.0253 \sqrt{1/77 + 1/161}} = 0.86$$

Esercizio (soluzione):

Definire la regione di rifiuto e formulare la decisione (test d'ipotesi):



NON si rifiuta l'ipotesi nulla (H_0)

→ la differenza osservata sembra essere dovuta al caso

IN ALTERNATIVA - Calcolare il p-value (test di significatività statistica):

p-value = $2 * \Pr(T_{236} \cong Z \geq 0.87 \mid H_0 \text{ è vera}) = 0.39$

RELAZIONE TRA INTERVALLO DI CONFIDENZA E TEST STATISTICO

Esempio (calcemia):

	Numero di pazienti	Media (mg/100 ml)	DS (mg/100 ml)
Vitamina D (D_1)	233	9.36	1.15
Controllo (D_0)	394	9.01	1.33

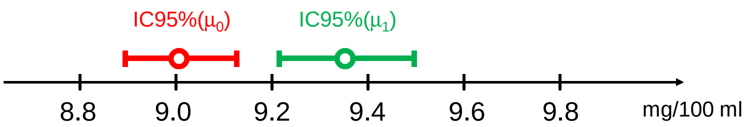
$$IC95\%(\mu_0) = \bar{x}_0 \pm 1.96\ s_0/\sqrt{n_0} = 9.01 \pm 1.96 * 1.33/\sqrt{394}$$

$$IC95\%(\mu_1) = \bar{x}_1 \pm 1.96\ s_1/\sqrt{n_1} = 9.36 \pm 1.96 * 1.15/\sqrt{233}$$

SESM

Esempio (calcemia):

	controllo (n = 394)	vitamina D (n = 233)	p-value
media [IC95%]	9.01 [8.88, 9.14]	9.36 [9.21, 9.51]	0.0009 (<0.05)



gli intervalli di confidenza non sono sovrapposti

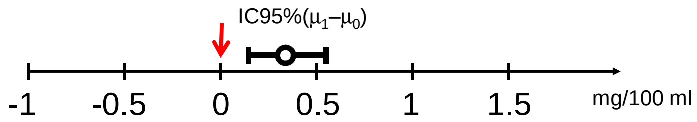
⇒ la differenza osservata è statisticamente significativa

SESM

Esempio (calcemia):

differenza delle medie [IC95%]	**0.35 [0.14, 0.56]**

mg/100 ml



l'intervallo di confidenza della differenza delle medie non contiene il valore nullo (0 mg/100ml)

⇒ la differenza osservata è statisticamente significativa ma NON è clinicamente rilevante

SIGNIFICATIVITÀ STATISTICA vs RILEVANZA CLINICA