



Intervallo di confidenza di una
media. Esercitazione

Es: IC 95% per la media

Si vuole stimare il valore medio della pressione sanguigna fra i soggetti sottoposti ad un certo trattamento farmacologico. Si dispone di un campione di 130 soggetti, con media campionaria pari a 160 e deviazione standard pari a 25.

Data l'ampiezza del campione, è possibile applicare la formula dell'intervallo di confidenza usando le tavole della Normale; scegliamo il livello 95%

$$\bar{x} = 160$$

$$s = 25, n = 130 \Rightarrow s.e.(\bar{X}) = 25 / \sqrt{130} = 2.19$$

$$z_{\frac{\alpha}{2}} = 1.96$$

$$\begin{aligned} 95\%CI &= (\bar{x} - 1.96 \cdot \sigma / \sqrt{n}, \bar{x} + 1.96 \cdot \sigma / \sqrt{n}) \\ &= (160 - 1.96 \cdot 2.19, 160 + 1.96 \cdot 2.19) \\ &= (155.7, 164.3) \end{aligned}$$

Esempio: calcolo dell'intervallo di confidenza nel caso di una media:

N 15

Media campionaria 149,133 mmHg (calcolo omissso)

$\sigma=2,53$ mmHg

Limite superiore dell' Int. conf. 95%= $149,133 + 1,960 * (2,53/\sqrt{15}) = 150,41$

Limite inferiore dell' Int. conf. 95%= $149,133 - 1,960 * (2,53/\sqrt{15}) = 147,85$

Intervallo di confidenza al 95%: $147,85 \leq \mu \leq 150,41$

Il modo comunemente usato per indicare media campionaria ed intervallo di confidenza è il seguente:

Media = 149,13 (IC 95%: 147,85 - 150,41)

Esempio: Calcolo dell'intervallo di confidenza della media di una popolazione

Problema: Qual è l'intervallo di confidenza al 95% della media del peso di una popolazione, se la media di un campione di 16 soggetti è pari a 75 Kg e la deviazione standard è pari a 12 Kg?

Dati: $\bar{x} = 75$ Kg $s = 12$ Kg $n = 16$ $1-\alpha = 95\%$ $t_{15, \alpha/2} = 2,131$

Formula da utilizzare: $I.C._{95\%} = \bar{x} \pm t_{\alpha/2} \cdot \sigma/\sqrt{n} = \bar{x} \pm t_{\alpha/2} \cdot E.S.$

I passo: calcolo l'errore standard

$$E.S. = s/\sqrt{n} = 12/\sqrt{16} = 12/4 = 3 \text{ Kg}$$

II passo: calcolo l'intervallo di confidenza

$$I.C._{95\%} = \bar{x} \pm t_{15, \alpha/2} \cdot E.S. = 75 \pm 2,131 \cdot 3 = \left[\begin{array}{l} 81,39 \text{ Kg} \\ 68,61 \text{ Kg} \end{array} \right.$$

L'intervallo che va da 68,61 Kg (limite inferiore) a 81,39 Kg (limite superiore) ha 95 probabilità su 100 di contenere la media vera della popolazione.

Esempio: calcolo dell'intervallo di confidenza usando la distribuzione t

Risultati:

N 15

Media campionaria 149,1 mmHg (calcolo omissivo)

s= 4,72 mmHg (calcolo omissivo)

Il valore della statistica t (errore 1° tipo $\leq 0,05$ e test a due code, 14 gl) = 2,145

Lim.superiore dell' Int. conf. 95% = $149,1 + 2,145 * (4,72/\sqrt{15}) = 151,75$ mmHg

Lim.inferiore dell' Int. conf. 95% = $149,1 - 2,145 * (4,72/\sqrt{15}) = 146,52$ mmHg

⌘ PROBLEMA n° 1

Per valutare l'efficacia della somministrazione di Ferro nella prevenzione dell'anemia sideropenica, si è trattato il 20% delle ragazze che frequentano la scuola secondarie con 1 compressa/die di solfato ferroso per 12 settimane consecutive.

Al termine dello studio la distribuzione dei livelli di emoglobina (Hb) aveva media $\mu_T=140$ g/l e deviazione standard $\sigma_T=10$ g/l per le ragazze trattate, e media $\mu_C=131$ g/l e deviazione standard $\sigma_C=12$ g/l per le ragazze non trattate.

Nell'ipotesi che la distribuzione dell'emoglobina sia gaussiana, calcola:

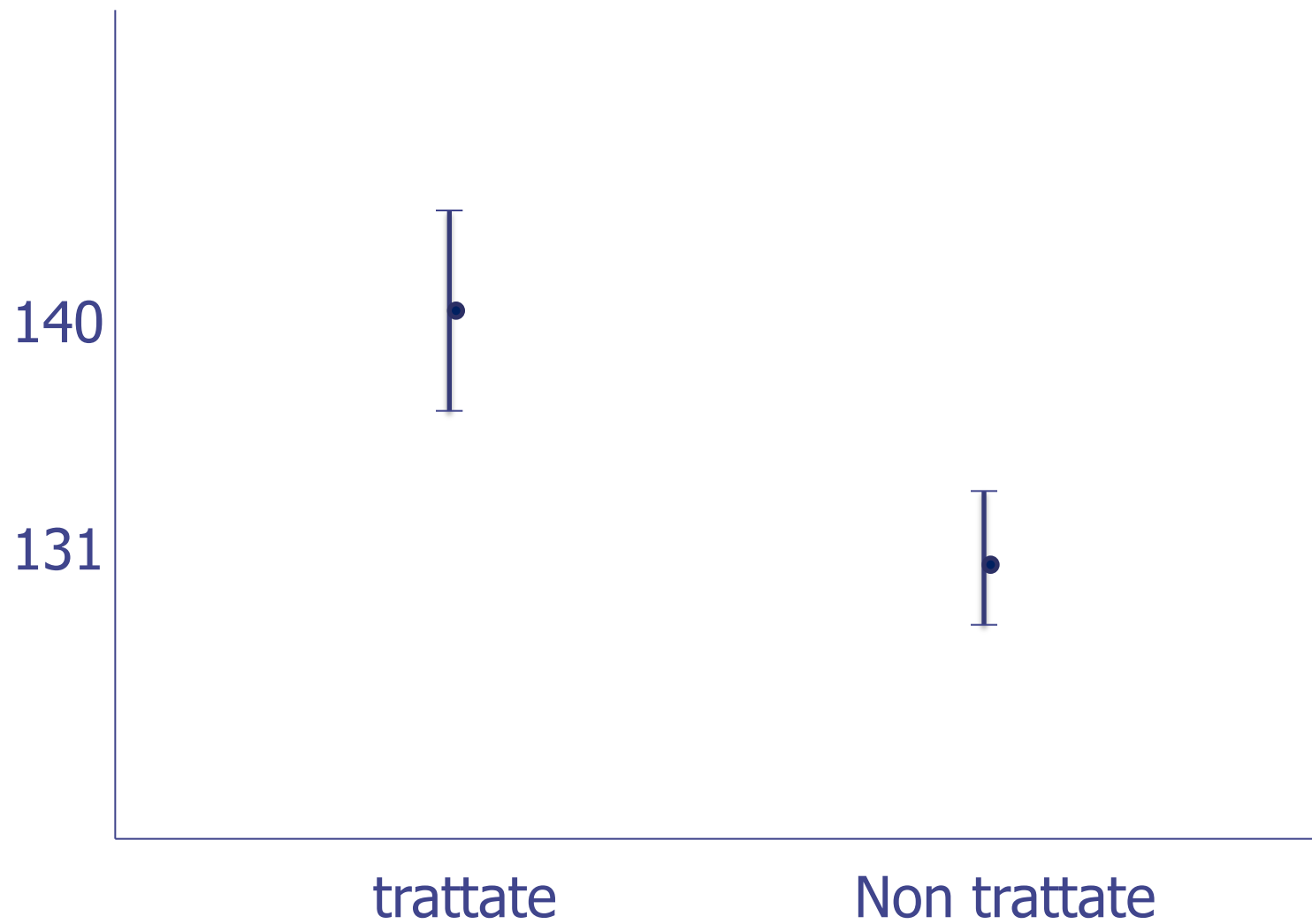
- l'intervallo di confidenza al 95% del livello di emoglobina per le ragazze trattate
- l'intervallo di confidenza al 95% del livello di emoglobina per le ragazze NON trattate

Dati: 100 ragazze (20 trattate 80 non trattate)

l'intervallo di confidenza al 95% del livello di emoglobina per le ragazze trattate	$C.I. = [140 \pm (1.96)(10)/\sqrt{20}]$ $= [135.6 ; 144.4]$
l'intervallo di confidenza al 95% del livello di emoglobina per le ragazze NON trattate	$C.I. = [131 \pm (1.96)(12)/\sqrt{80}]$ $= [128.4 ; 133.6]$

Il test $t = 140 - 131 / [es (140 - 131)] = 3.093$

$P = 0.003$



⌘PROBLEMA n° 2

Nel Framingham Study, sono stati misurati i livelli di colesterolo totale sierico in una coorte di 475 soggetti maschi di età 40-49 anni ed in buona salute.

La coorte è stata poi seguita per sedici anni.

Alla fine di questo periodo, i soggetti sono stati divisi in due gruppi:

- 1) coloro che **non avevano sviluppato** malattie coronariche 354 (normali).
- 2) coloro che **avevano sviluppato** malattie coronariche 121 (malati).

La distribuzione dei livelli di colesterolo sierico nei due gruppi (all'inizio dello studio) era stata trovata approssimativamente gaussiana

- con media $\mu_1 = 219$ mg/dl e dev.stand. $\sigma_1 = 41$ mg/dl nei soggetti che NON hanno poi sviluppato malattia
- con media $\mu_2 = 244$ mg/dl e dev.stand. $\sigma_2 = 51$ mg/dl nei soggetti che hanno poi sviluppato la malattia e

(MacMahon SW, McDonald GJ: *A population at risk: prevalence of high cholesterol levels in hypertensive patients in Framingham Study*, (1986) Am.J.Med.Sup.:80:40-47)

determinare

l'intervallo di confidenza al 95% del livello di colesterolo sierico per coloro che hanno sviluppato la malattia coronaria

l'intervallo di confidenza al 95% del livello di colesterolo sierico per coloro che NON hanno sviluppato la malattia coronaria

l'intervallo di confidenza al 95% del livello di colesterolo sierico per coloro che hanno sviluppato la malattia coronaria	C.I.=[244 ± (1.96)(51)/√121] =[234.9 ; 253.1]
l'intervallo di confidenza al 95% del livello di colesterolo sierico per coloro che NON hanno sviluppato la malattia coronaria	C.I.=[219 ± (1.96)(41)/√354] =[214.8 ; 223.2]

$$t = (244-219)/ES(244-219) = 25/4.068 = 5.426$$

con 473 gradi di libertà; P = 0.000

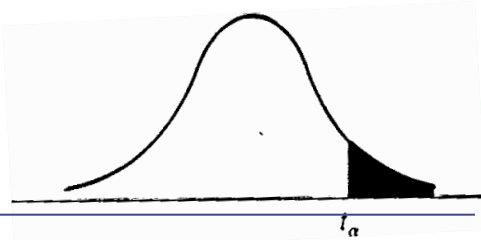
Deviata gaussiana Standard: Aree per $z > +z^*$ (o per $z < -z^*$)

z^*	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414
0.1	.46017	.45620	.45224	.44828	.44433	.44038	.43644	.43251	.42858	.42465
0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
1.0	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08691	.08534	.08379	.08226
1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
2.0	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831

Devziata gaussiana Standard: Aree Per $z > +z^*$ (o per $z < -z^*$)

z^*	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
2-4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
3.0	.00135	.00097	.00069	.00048	.00034	.00023	.00016	.00011	.00007	.00005
4.0	.00003	.00002	.00001	.00001	.00001	.00000	.00000	.00000	.00000	.00000

Tavola dei valori della funzione t di Student in funzione dei gradi di libertà e della probabilità in una coda della distribuzione (.100, .050, .025, .010, .005)



DEGREES OF FREEDOM	t_{.100}	t_{.050}	t_{.025}	t_{.010}	t_{.005}
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947