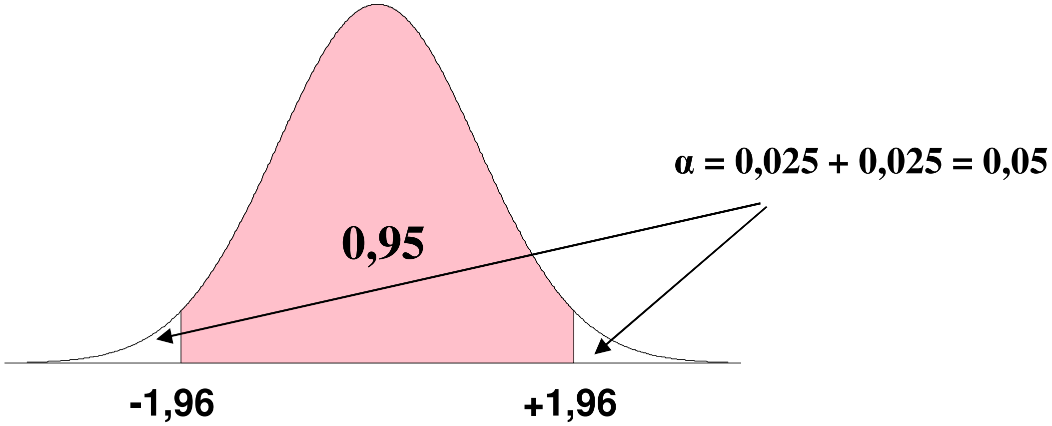
UNIVERSITÁ
DEGLI STUDI
DI TERAMO

Dal campione alla popolazione: stima puntuale e per intervalli; cenni sulla verifica delle ipotesi statistiche

Introduzione

- Introdurre il concetto di intervallo di confidenza
- Stima di parametri per piccoli e grandi campioni
- Stimare la proporzione di individui che possiedono un determinato attributo

Media Campionaria e Curva Normale



$$E(\overline{X}) = \mu$$

$$s.q.m.(\overline{X}) = \frac{\sigma}{\sqrt{n}}$$

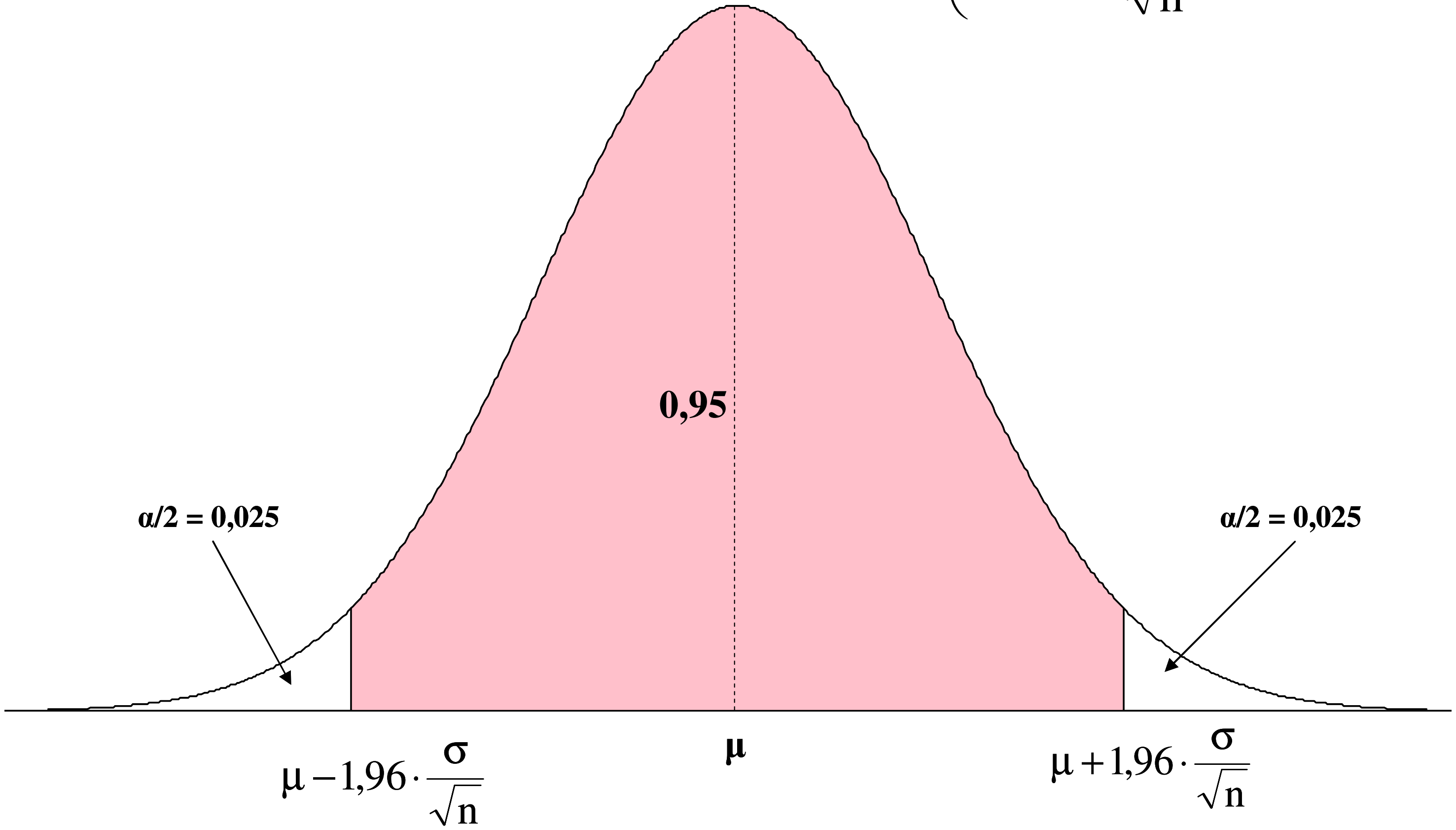
$$P(-1,96 \leq z \leq +1,96) = P\left(-1,96 \leq \frac{x-\mu}{\sigma} \leq +1,96\right) = 0,95$$

$$P\left(-1,96 \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq +1,96\right) = 0,95$$

$$P\left(\mu - 1,96\frac{\sigma}{\sqrt{n}} \leq \overline{X} \leq \mu + 1,96\frac{\sigma}{\sqrt{n}}\right) = 0,95$$

Intervallo di probabilità

$$P\left(\mu - 1{,}96\frac{\sigma}{\sqrt{n}} \leq \overline{X} \leq \mu + 1{,}96\frac{\sigma}{\sqrt{n}}\right) = 0{,}95$$



Media Campionaria e Curva Normale - 2

X_i	n_i	$P(X_i)$
120	1	0,04
130	2	0,08
135	2	0,08
140	3	0,12
145	2	0,08
150	5	0,20
155	2	0,08
160	3	0,12
165	2	0,08
170	2	0,08
180	1	0,04
	25	1,00

$$\mu = 150$$

$$\sigma = 20$$

$$P\left(\mu - 1{,}96\frac{\sigma}{\sqrt{n}} \le \overline{X} \le \mu + 1{,}96\frac{\sigma}{\sqrt{n}}\right) = 0{,}95$$

$$P\left(150 - 1{,}96\frac{20}{\sqrt{2}} \le \overline{X} \le 150 + 1{,}96\frac{20}{\sqrt{2}}\right) = 0{,}95$$

$$P(150 - 27{,}72 \le \overline{X} \le 150 + 27{,}72) = 0{,}95$$

$$P(122{,}28 \le \overline{X} \le 177{,}72) = 0{,}95$$

$$P(122{,}28 \le \overline{X} \le 177{,}72) = \frac{23}{25} = 0{,}92$$

Intervalli di probabilità e di confidenza

$$P\left(\mu - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \overline{X} \le \mu + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

- ☐ **0,90 (90%) → al quale corrisponde un valore di z pari a ± 1,65**
- ☐ **0,95 (95%) → al quale corrisponde un valore di z pari a ± 1,96**
- ☐ **0,99 (99%) → al quale corrisponde un valore di z pari a ± 2,58**

Intervallo di probabilità non utile nella realtà (non conosciamo μ!)

Dobbiamo rovesciare il ragionamento

$$\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

**INTERVALLO
DI CONFIDENZA**

Intervallo di Confidenza

$$\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\mu = 150 \qquad \sigma = 20 \qquad n = 2$$

$$(1-\alpha) = 0{,}95 \Rightarrow z_{\alpha/2} = 1{,}96$$

$$AC \Rightarrow \overline{x} = 135$$

$$135 - 1{,}96 \cdot \frac{20}{\sqrt{2}} \le \mu \le 135 + 1{,}96 \cdot \frac{20}{\sqrt{2}}$$

$$135 - 27{,}72 \le \mu \le 135 + 27{,}72$$

$$107{,}28 \le \mu \le 162{,}72$$

$$DD \Rightarrow \overline{x} = 160$$

$$160 - 1{,}96 \cdot \frac{20}{\sqrt{2}} \le \mu \le 160 + 1{,}96 \cdot \frac{20}{\sqrt{2}}$$

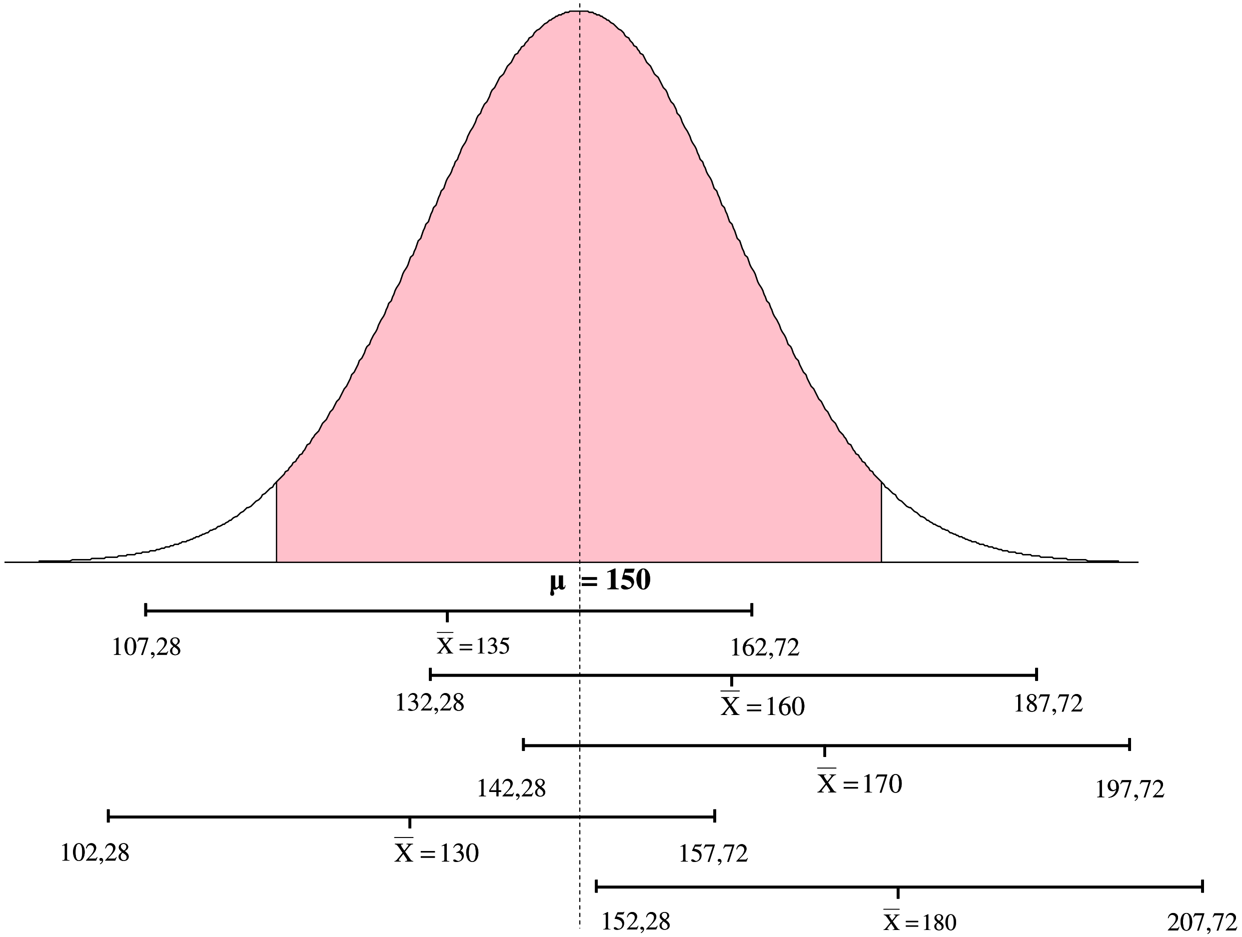
$$132{,}28 \le \mu \le 187{,}72$$

$$BD \Rightarrow \overline{x} = 170$$

$$170 - 1{,}96 \cdot \frac{20}{\sqrt{2}} \le \mu \le 170 + 1{,}96 \cdot \frac{20}{\sqrt{2}}$$

$$142{,}28 \le \mu \le 197{,}72$$

Intervallo di Confidenza - 2



Intervallo di Confidenza per Grandi Campioni

In realtà, il modello introdotto in precedenza è valido sotto opportune condizioni:

- il campione viene estratto da popolazioni nelle quali il carattere considerato si distribuisce come una Normale
- conosciamo lo scarto quadratico medio della popolazione (ricordate? Nella formula precedente abbiamo utilizzato σ…)

Teorema del limite centrale

n > 60-70

$$\overline{X} - z_{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}$$

Intervallo di Confidenza per Grandi Campioni - Esercizio

Vogliamo conoscere l'età media di un gruppo di individui; estraiamo un campione casuale semplice di 100 soggetti da tale popolazione, e calcoliamo un'età media del campione pari a 21,6 anni. Da una precedente indagine censuaria, sappiamo che lo scarto quadratico medio della popolazione è pari a 5,1. Si costruisca un intervallo di confidenza ad un livello di fiducia del 95% per la stima dell'età media della popolazione.

$$\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$n = 100 \qquad \overline{x} = 21{,}6 \qquad \sigma = 5{,}1$$

$$(1-\alpha) = 0{,}95 \Rightarrow z_{\alpha/2} = 1{,}96$$

$$21{,}6 - 1{,}96 \cdot \frac{5{,}1}{\sqrt{100}} \le \mu \le 21{,}6 + 1{,}96 \cdot \frac{5{,}1}{\sqrt{100}}$$

$$21{,}6 - 1 \le \mu \le 21{,}6 + 1$$

$$\boxed{20{,}6 \le \mu \le 22{,}6}$$

Intervallo di Confidenza per Grandi Campioni - Esercizio

$$(1-\alpha)=0{,}99 \Rightarrow z_{\alpha/2}=2{,}58$$

$$n=100 \qquad \overline{x}=21{,}6 \qquad \sigma=5{,}1$$

$$\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$21{,}6-2{,}58\cdot\frac{5{,}1}{\sqrt{100}} \le \mu \le 21{,}6+2{,}58\cdot\frac{5{,}1}{\sqrt{100}}$$

$$21{,}6-1{,}31 \le \mu \le 21{,}6+1{,}31$$

$$\boxed{20{,}29 \le \mu \le 22{,}91}$$

$$(1-\alpha)=0{,}20 \Rightarrow z_{\alpha/2}=0{,}254$$

$$\boxed{21{,}47 \le \mu \le 21{,}73}$$

95% → $\boxed{20{,}6 \le \mu \le 22{,}6}$

99% → $\boxed{20{,}29 \le \mu \le 22{,}91}$

20% → $\boxed{21{,}47 \le \mu \le 21{,}73}$

Esercizio 2

Si vuole stimare il tempo medio di risposta ad un test psico-attitudinale al quale sono stati sottoposti gli studenti di una certa Facoltà. Per tale motivo, viene estratto un campione casuale di 150 studenti, dai quali risulta un tempo medio di risposta pari a 24 minuti, con uno scarto quadratico medio di 3 minuti. Costruire un intervallo di confidenza, ad un livello di fiducia del 95% per il tempo medio di risposta al test.

$$\overline{X} - z_{\alpha/2}\frac{s}{\sqrt{n}} \le \mu \le \overline{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}$$

$$n = 150 \quad ; \quad \overline{x} = 24 \quad ; \quad s = 3$$

$$(1-\alpha) = 0,95 \Rightarrow z_{\alpha/2} = 1,96$$

$$24 - 1,96 \cdot \frac{3}{\sqrt{150}} \le \mu \le 24 + 1,96 \cdot \frac{3}{\sqrt{150}}$$

$$24 - 1,96 \cdot \frac{3}{12,25} \le \mu \le 24 + 1,96 \cdot \frac{3}{12,25}$$

$$24 - 0,48 \le \mu \le 24 + 0,48$$

$$\boxed{23,52 \le \mu \le 24,48}$$

$$n = 1000$$

$$24 - 1,96 \cdot \frac{3}{\sqrt{1.000}} \le \mu \le 24 + 1,96 \cdot \frac{3}{\sqrt{1.000}}$$

$$\boxed{23,81 \le \mu \le 24,19}$$

Esercizio 3

Il manager che si occupa del controllo di qualità di un'azienda che produce lampadine intende stimare la durata media delle lampadine facenti parte di un determinato lotto di produzione. Si sa che lo scarto quadratico medio delle lampadine prodotte è pari a 100 ore. Viene estratto un campione di 64 lampadine, che forniscono una durata media pari a 350 ore. Rispondere alle seguenti domande:

1) calcolare un intervallo di confidenza al livello di fiducia del 95% per la durata media delle lampadine;

2) può il produttore affermare che la durata media delle lampadine è pari a 400 ore?

$$\overline{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$n = 64 \quad ; \quad \overline{x} = 350 \quad ; \quad \sigma = 100$$

$$(1-\alpha) = 0{,}95 \Rightarrow z_{\alpha/2} = 1{,}96$$

$$350 - 1{,}96 \cdot \frac{100}{\sqrt{64}} \leq \mu \leq 350 + 1{,}96 \cdot \frac{100}{\sqrt{64}}$$

$$\boxed{325{,}5 \leq \mu \leq 374{,}5}$$

Intervallo di Confidenza per Piccoli Campioni

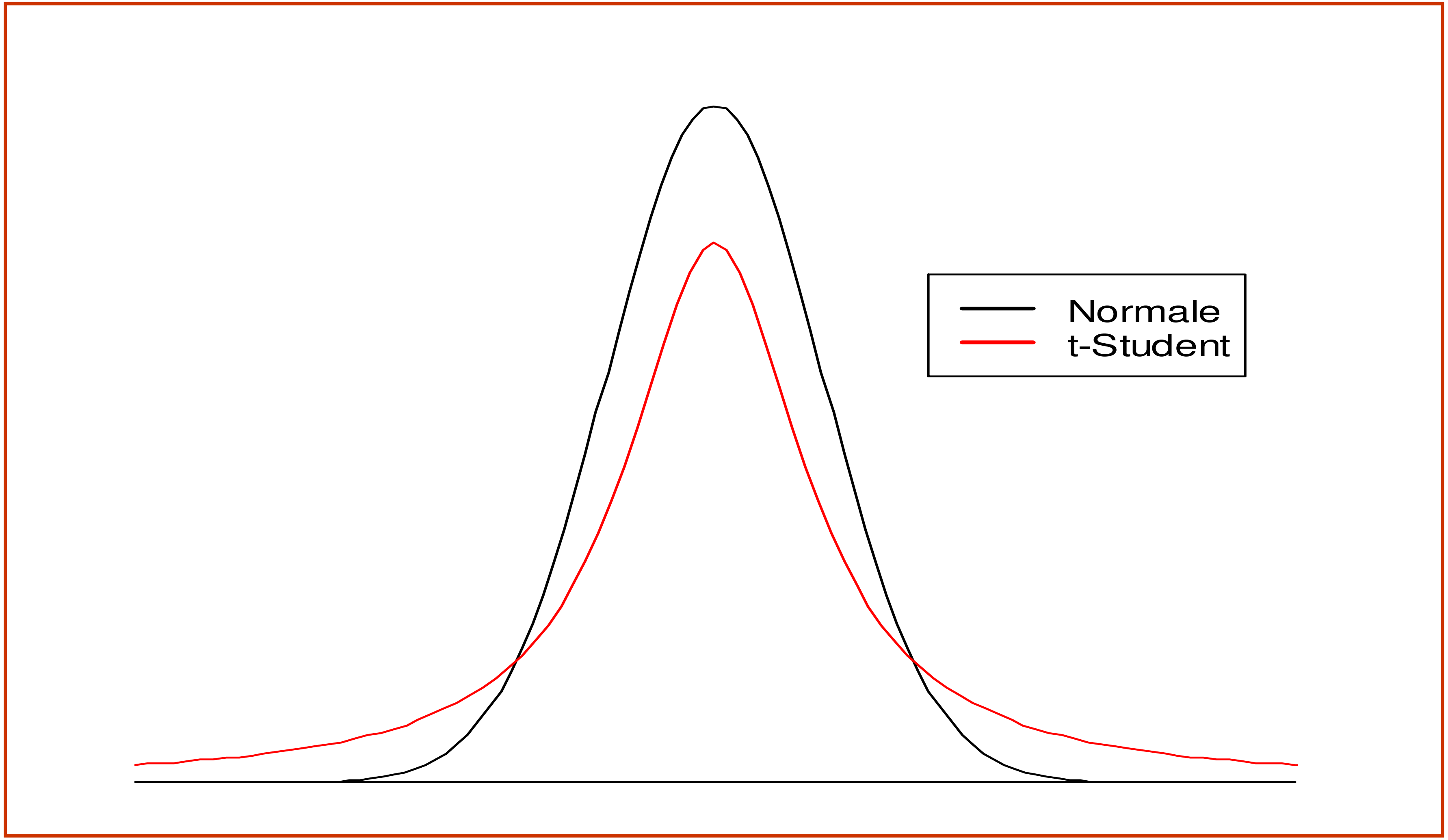
Quando abbiamo a che fare con campioni di bassa numerosità, e dal momento che, nella maggior parte dei casi, la varianza della popolazione è un parametro sconosciuto, non sarà possibile fare riferimento alla curva normale come modello generatore del fenomeno che stiamo considerando (ossia, la stima per intervallo della media aritmetica della popolazione)

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \longrightarrow$$

T-STUDENT
con (n-1) Gradi di Libertà

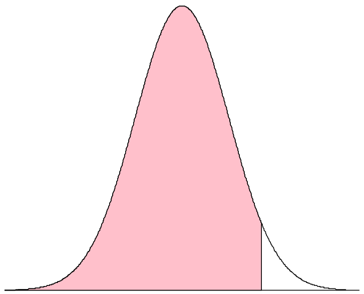
$$P\left(-t_{\alpha/2} \leq \frac{\overline{X}-\mu}{S/\sqrt{n}} \leq +t_{\alpha/2}\right) = 1-\alpha \longrightarrow \overline{X}-t_{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \overline{X}+t_{\alpha/2}\frac{s}{\sqrt{n}}$$

Curva Normale e Curva T-Student



Valori critici della T-Student

	Area nella coda destra					
Gradi di libertà	**0,250**	**0,100**	**0,050**	**0,025**	**0,010**	**0,005**
1	1,0000	3,0777	6,3137	12,7062	31,8210	63,6559
2	0,8165	1,8856	2,9200	4,3027	6,9645	9,9250
3	0,7649	1,6377	2,3534	3,1824	4,5407	5,8408
4	0,7407	1,5332	2,1318	2,7765	3,7469	4,6041
5	0,7267	1,4759	2,0150	2,5706	3,3649	4,0321
6	0,7176	1,4398	1,9432	2,4469	3,1427	3,7074
7	0,7111	1,4149	1,8946	2,3646	2,9979	3,4995
8	0,7064	1,3968	1,8595	2,3060	2,8965	3,3554
9	0,7027	1,3830	1,8331	2,2622	2,8214	3,2498
10	0,6998	1,3722	1,8125	2,2281	2,7638	3,1693
50	0,6794	1,2987	1,6759	2,0086	2,4033	2,6778
70	0,6780	1,2938	1,6669	1,9944	2,3808	2,6479
80	0,6776	1,2922	1,6641	1,9901	2,3739	2,6387
100	0,6770	1,2901	1,6602	1,9840	2,3642	2,6259
110	0,6767	1,2893	1,6588	1,9818	2,3607	2,6213
120	0,6765	1,2886	1,6576	1,9799	2,3578	2,6174
150	0,6761	1,2872	1,6551	1,9759	2,3515	2,6090
200	0,6757	1,2858	1,6525	1,9719	2,3451	2,6006
∞	0,6745	1,2816	1,6449	1,9600	2,3264	2,5758



$$(1-\alpha)=0{,}95$$

$$\alpha=0{,}05$$

$$\alpha/2=0{,}025$$

Intervallo di Confidenza per Piccoli Campioni - Esercizio

Si vuole conoscere il peso medio dei neonati venuti alla luce in un certo ospedale. Si estrae un campione casuale di 16 elementi, e si riscontra un peso medio di 3,42 kg, con una varianza campionaria pari a 0,4624. Costruire un intervallo di confidenza ad un livello di fiducia del 99% per la stima del peso medio della popolazione di neonati.

$$n = 16 \qquad \overline{x} = 3{,}42 \qquad s^2 = 0{,}4624$$

	Area nella coda destra					
Gradi di libertà	**0,25**	**0,10**	**0,05**	**0,025**	**0,01**	**0,005**
1	1,0000	3,0777	6,3137	12,7062	31,8210	63,6559
2	0,8165	1,8856	2,9200	4,3027	6,9645	9,9250
3	0,7649	1,6377	2,3534	3,1824	4,5407	5,8408
4	0,7407	1,5332	2,1318	2,7765	3,7469	4,6041
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
12	0,6955	1,3562	1,7823	2,1788	2,6810	3,0545
13	0,6938	1,3502	1,7709	2,1604	2,6503	3,0123
14	0,6924	1,3450	1,7613	2,1448	2,6245	2,9768
15	0,6912	1,3406	1,7531	2,1315	2,6025	2,9467
16	0,6901	1,3368	1,7459	2,1199	2,5835	2,9208
17	0,6892	1,3334	1,7396	2,1098	2,5669	2,8982

$$(1-\alpha) = 0{,}99$$

$$\alpha / 2 = 0{,}005$$

$$\text{g.d.l.} = 16 - 1 = 15$$

Intervallo di Confidenza per Piccoli Campioni – Esercizio

$$n = 16 \qquad \overline{x} = 3{,}42 \qquad s^2 = 0{,}4624 \qquad s = \sqrt{s^2} = \sqrt{0{,}4624} = 0{,}68$$

$$(1 - \alpha) = 0{,}99 \Rightarrow t_{\alpha/2;15} = 2{,}947$$

$$\overline{X} - t_{\alpha/2} \frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$3{,}42 - 2{,}947 \cdot \frac{0{,}68}{\sqrt{16}} \leq \mu \leq 3{,}42 + 2{,}947 \cdot \frac{0{,}68}{\sqrt{16}}$$

$$3{,}42 - 0{,}5 \leq \mu \leq 3{,}42 + 0{,}5$$

$$\boxed{2{,}92 \leq \mu \leq 3{,}92}$$

Intervallo di Confidenza per la stima di una proporzione

Quando vogliamo stimare una proporzione (o, in altre parole, la percentuale di unità statistiche che presentano un determinato attributo) è possibile estrarre un campione dalla popolazione di riferimento e calcolare tale proporzione all'interno del campione

E' possibile dimostrare che, se potessimo calcolare tutte le proporzioni all'interno di tutti i campioni di una certa numerosità n estraibili da una determinata popolazione, tali proporzioni danno origine ad una variabile casuale che si distribuisce come una curva normale con valore atteso pari a π e varianza pari a $\pi(1-\pi)/n$

$$\frac{p-\pi}{\sqrt{\frac{\pi\cdot(1-\pi)}{n}}}$$

per n sufficientemente grande [almeno pari a 5 sia il prodotto $n\pi$, sia il prodotto $n(1-\pi)$] si distribuisce come una V.C Normale standardizzata

Stima di una proporzione - Esercizio

Si vuole conoscere la proporzione di individui che voteranno Si al prossimo referendum. A tale scopo, viene effettuata un'indagine su 80 soggetti, 60 dei quali rispondono che voteranno Si. Si determini l'intervallo di confidenza, ad un livello di fiducia del 95%, per la stima della proporzione di votanti in modo affermativo.

$$p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \leq \pi \leq p + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

$$0{,}75 - 1{,}96\sqrt{\frac{0{,}75 \cdot 0{,}25}{80}} \leq \pi \leq 0{,}75 + 1{,}96\sqrt{\frac{0{,}75 \cdot 0{,}25}{80}}$$

$$0{,}75 - 1{,}96 \cdot 0{,}048 \leq \pi \leq 0{,}75 + 1{,}96 \cdot 0{,}048$$

$$0{,}75 - 0{,}09 \leq \pi \leq 0{,}75 + 0{,}09$$

$$\boxed{0{,}66 \leq \pi \leq 0{,}84}$$

$$n = 80 \quad ; \quad p = \frac{60}{80} = 0{,}75$$

$$(1-p) = 1 - 0{,}75 = 0{,}25$$

$$np = 80 \cdot 0{,}75 = 60$$

$$n(1-p) = 80 \cdot 0{,}25 = 20$$

$$z_{0,025} = 1{,}96$$