Formulario stimatori

Media e varianza campionaria

In statistica, **una popolazione** è l'insieme degli elementi che sono oggetto di studio, come ad esempio:

- Il numero di componenti delle famiglie di una data area geografica.
- L'età dei cittadini di un certo paese...

Campione o campione aleatorio

Un insieme $X_1, X_2, \dots, X_n$ di variabili aleatorie **indipendenti**, tutte con la **stessa distribuzione**(i.i.d → indipendenti e identicamente distribuite), si chiama **campione** o **campione aleatorio** della distribuzione di F.

- $X_1, X_2, \dots, X_n$ hanno la stessa distribuzione
- Per ogni $i \in \{1, \dots, n\}$

$$\begin{cases} E(X_i) = \mu \\ Var(X_i) = \sigma^2 \end{cases}$$

μ e σ^2 si dicono **media** e **varianza** della popolazione

Statistica

Sia $X_1, \dots, X_n$ un campione. Una **statistica** è una funzione del campione.

$$T = g(X_1, \dots, X_n)$$

- Statistica → T
- Funzione → $g(X_1, \dots, X_n)$

Media campionaria

Sia $X_1, \dots, X_n$ un campione di dati estratti di una popolazione. La **media campionaria** del campione $X_1, \dots, X_n$ di dati estratti di una popolazione, è definita dalla media aritmetica:

$$\vec{X} := \frac{X_1 + \dots X_n}{n}$$

$\Rightarrow \vec{X}$ è una **statistica**.

$$E(\vec{X}) \to \text{media campionaria} = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu$$

▼ Dimostrazione

$$E(\vec{X}) = E\Big(\frac{X_1 + X_2 + \dots + X_n}{n}\Big) = \Big(\frac{E(X_1) + E(X_2) + \dots + E(X_n)}{n}\Big) = \frac{n\mu}{\mu} = \mu$$

$$E(\vec{X}) = \frac{1}{n}\sum_{i=1}^{n} x_i = g(X_1, \dots, X_n)$$

Valore atteso e varianza di un campione aleatorio

- $E(\vec{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu$
- $E(\vec{X}^2) = \frac{1}{n}\sum_{i=1}^{n} E(X_i^2)$

- $Var(\vec{X}) = \frac{\sigma^2}{\mu}$

Varianza campionaria

Sia $X_1, \dots, X_n$ un campione di media campionaria $\vec{X}$. La statistica S^2 è definita da:

$$S^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \vec{X})^2$$

si dice **varianza campionaria**. Essa è sia una **statistica**, in quanto funzione di un campione, che una **variabile aleatoria**, in quanto funzione di variabili aleatorie.

▼ Semplificazione formula

$$\begin{aligned}
\sum_{i=1}^{n}(X_i - \overline{X})^2 &= \sum_{i=1}^{n}(X_i^2 - 2\overline{X}X_i + \overline{X}^2) \\
&= \sum_{i=1}^{n} X_i^2 - 2\overline{X}\underbrace{\sum_{i=1}^{n} X_i}_{= n\overline{X}} + \underbrace{n\overline{X}^2}_{= \sum_{i=1}^{n} \overline{X}^2 = \overline{X}^2 \underbrace{\sum_{i=1}^{n} 1}_{n}} \\
&= \sum_{i=1}^{n} X_i^2 - 2n\overline{X}^2 + n\overline{X}^2 = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2
\end{aligned}$$

Varianza campionaria semplificata

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i^2 - \frac{n}{n-1}\vec{X}^2$$

Media della varianza campionaria

$$E(S^2) = \sigma^2$$

▼ Dimostrazione

$$
\begin{aligned}
E(S^2) &= \frac{1}{n-1} E\Big(\sum_{i=1}^{n}(X_i - \overline{X})^2\Big) \\
&= \frac{1}{n-1}\sum_{i=1}^{n} E(X_i^2) - \frac{n}{n-1}E(\overline{X}^2) \\
&= \frac{1}{n-1}\sum_{i=1}^{n}(\sigma^2 + \mu^2) - \frac{n}{n-1}(\frac{\sigma^2}{n} + \mu^2) \\
&= \frac{n}{n-1}(\sigma^2 + \mu^2) - \frac{n}{n-1}(\frac{\sigma^2}{n} + \mu^2) \\
&= \sigma^2
\end{aligned}
$$

Il **valore atteso** della **varianza campionaria** coincide con la **varianza della popolazione**.

Deviazione standard campionaria

$$S = \sqrt{S^2}$$

Valori attesi

- $E(S^2) = \frac{1}{n-1}\sum_{i=1}^{n} E(X_i^2) - \frac{n}{n-1}E(\vec{X}^2)$
- $Var(\vec{X}) = \frac{\sigma^2}{n} = E(\vec{X}^2) - (E(\vec{X}))^2 = E(\vec{X}^2) - \mu^2$

  $\Rightarrow E(\vec{X}^2) = \frac{\sigma^2}{n} + \mu^2$
- $E(X_i^2) = Var(X_i) + (E(X_i))^2 = \sigma^2 + \mu^2$

Quantità che dipendono dal campione

$$
\begin{cases}
\text{Media campionaria: } \vec{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \\
\text{Varianza campionaria: } S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \vec{X})^2
\end{cases}
$$

$\vec{X}$ e S^2 sono **variabili aleatorie** che possiamo studiare le loro distribuzioni $f_{\vec{X}}$ e f_{S^2}

$$
\begin{cases}
\text{Media campionaria: } \vec{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \\
\text{Varianza campionaria: } s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \vec{x})^2
\end{cases}
$$

$\vec{x}$ e s^2 sono numeri e sono i **valori assunti** di variabili aleatorie con distribuzioni $f_{\vec{X}}$ e f_{S^2}.

$\vec{x} = x_1, x_2, \ldots, x_n$ è una realizzazione del campione $\vec{X} = X_1, \ldots, X_n$

Teorema del limite centrale

Sia $X_1, \dots, X_n$ un campione e di conseguenza anche delle variabili aleatorie i.i.d (indipendenti e identicamente distribuite), tutte con **media** μ e **varianza** σ^2**.** Allora, per n molto grandi, la somma $X_1 + X_2 + \cdots + X_n$ è **approssimativamente** una **variabile aleatoria normale** con **media** $n\mu$ e **varianza** $n\sigma^2$

$$X_1 + X_2 + \cdots + X_n \sim N(n\mu, n\sigma^2)$$

La **somma normalizzata** (per poter utilizzare la tabella) **è:**

- $\frac{X_1+X_2+\cdots+X_n-n\mu}{\sigma\sqrt{n}} \sim N(0,1)$
- $\frac{n(\frac{X_1+X_2+\cdots+X_n}{n}-\mu)}{\sigma\sqrt{n}} \sim N(0,1)$
- $\sqrt{n}\left[\frac{\vec{X}-\mu}{\sigma}\right] \sim N(0,1)$
- $\frac{\vec{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$

Le 4 scritture sono equivalenti.

Nella formula utilizziamo $n\mu$ e σ, **non** $n\sigma^2$ →che sarebbe la varianza della variabile aleatoria normale alla quale abbiamo approssimato.

▼ Esempio limite centrale

Esempio. Si tirano 10 dadi non truccati. Determina approssimativamente quanto vale la probabilità che la somma dei loro punteggi sia compresa tra 30 e 40 inclusi.

Soluzione. Sia X_i il punteggio del $i-$esimo dado. Abbiamo

$$\begin{aligned} E(X_i) &= \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6} = \frac{7}{2} = \mu \\ E(X_i^2) &= \frac{1}{6}(1^2+2^2+3^2+4^2+5^2+6^2) = \frac{91}{6} \\ Var(X_i) &= \frac{91}{6} - \frac{49}{4} = \frac{70}{24} = \frac{35}{12} = \sigma^2 \end{aligned}$$

AMEUR DHAHRI 11 / 34

Dal Teormea del limite centrale la somma $Y = X_1 + \cdots + X_{10}$ è approssimativamente normale con media $10 \times \frac{7}{2} = \frac{70}{2} = 35$ e varianza $10\sigma^2 = 10 \times \frac{35}{12} = \frac{175}{6}$. Allora si ha

$N(35, \frac{175}{6})$

$$\begin{aligned} & P(30 \leq X_1 + \cdots + X_{10} \leq 40) \\ &= P\Big(\frac{30-35}{\sigma\sqrt{10}} \leq \frac{X_1 + \cdots + X_{10} - \overbrace{35}^{n\mu}}{\sigma\sqrt{10}} \leq \frac{40-35}{\sigma\sqrt{10}}\Big) = \Phi(b) - \Phi(a) \\ &= P\Big(\frac{-5}{\sigma\sqrt{10}} \leq \frac{X_1 + \cdots + X_{10} - 35}{\sigma\sqrt{10}} \leq \frac{5}{\sigma\sqrt{10}}\Big) \\ &\approx \Phi\Big(\frac{5}{\sigma\sqrt{10}}\Big) - \Phi\Big(\frac{-5}{\sigma\sqrt{10}}\Big), \quad \Phi \text{ è la funzione di ripartizione di } \mathcal{N}(0,1) \\ &\approx 2\Phi\Big(\frac{5}{\sigma\sqrt{10}}\Big) - 1 \simeq 2\Phi(0.92) - 1 = 2 \times 0.82 - 1 = 0.64 \end{aligned}$$

↳ vediamo i valori nella tabella

Distribuzione approssimativa della media campionaria

Siano $X_1, \ldots, X_n$ delle variabili aleatorie i.i.d (indipendenti e identicamente distribuite), tutte con media μ e varianza σ^2 e se n è grande, allora dal **Teorema del limite centrale:**

$$\vec{X} = \frac{1}{n}(X_1 + \cdots + X_n) \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$$

Funzione di ripartizione della somma normalizzata

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} = \frac{\vec{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Essendo $X_1, \ldots, X_n$ un campione con $X_i \sim \mathcal{N}(\mu, \sigma^2)$ e $i = 1, 2, \ldots, n$. Siccome la somma di variabili aleatorie normali e indipendenti ha ancora distribuzione normale, allora $\vec{X}$ è normale. La sua media e la sua varianza, come nel caso generale, sono μ e $\frac{\sigma^2}{n}$

$$\frac{\vec{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0,1)$$

è **approssimata** da Φ, dove Φ è la funzione di ripartizione della variabile normale standard $\mathcal{N}(0,1)$, cioè

$$P\Big(\frac{\vec{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq x\Big) \sim \Phi(x)$$

Il risultato rimane lo stesso se sostituiamo σ con la **deviazione standard campionaria S**, cioè sotto ipotesi del **Teorema del limite centrale**, la distribuzione di

$$\frac{\vec{X}-\mu}{\frac{S}{\sqrt{n}}} \sim \mathcal{N}(0,1)$$

$$P\Big(\frac{\vec{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq x \leq \frac{\vec{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\Big) = 2\Phi\Big(\frac{\vec{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\Big) - 1$$

▼ Esempio 2 limite centrale

Esempio. Un docente sa dall'esperienza passata che il punteggio all'esame finale degli studenti del suo corso è distribuito con media 77 e deviazione standard 15. Attualmente egli ha due classi diverse, una di 64 e una di 25 studenti.

$\overline{X}$

(a) Quanto vale circa la probabilità che la media aritmetica dei punteggio della classe di 25 studenti sia compresa tra 72 e 82?

(b) Quanto vale circa la probabilità che la media aritmetica dei punteggio della classe di 64 studenti sia compresa tra 72 e 82?

Soluzione.

(a) Con una ampiezza del campione di 25, Allora dal Teorema del limite centrale $\overline{X}$ sarà approssimativamente normale con media 77 e e deviazione standard $\frac{15}{\sqrt{25}} = 3$.

Ameur Dhahri 15 / 34

Allora si ha

$$\begin{aligned} P(72 < \overline{X} < 82) &= P\Big(\frac{72-77}{3} < \frac{\overline{X}-77}{3} < \frac{82-77}{3}\Big) \\ &= P\Big(-\frac{5}{3} < \frac{\overline{X}-77}{3} < \frac{5}{3}\Big) \\ &\simeq 2\Phi(\frac{5}{3}) - 1 \simeq 2\Phi(1.66) - 1 \\ &\simeq 2 \times 0.95 - 1 = 0.9 \end{aligned}$$

(b) Con una ampiezza del campione di 64, $\overline{X}$ sarà approssimativamente normale con media 77 e e deviazione standard $\frac{15}{\sqrt{64}} = \frac{15}{8} = 1.875$. Allora si ha

$$\begin{aligned} P(72 < \overline{X} < 82) &= P\Big(\frac{72-77}{1.875} < \frac{\overline{X}-77}{1.875} < \frac{82-77}{1.875}\Big) \\ &= P\Big(-\frac{5}{1.875} < \frac{\overline{X}-77}{1.875} < \frac{5}{1.875}\Big) \\ &\simeq 2\Phi(\frac{5}{1.875}) - 1 \simeq 2\Phi(2.66) - 1 \\ &\simeq 2 \times 0.996 - 1 = 0.992 \end{aligned}$$

▼ Esempio 3 limite centrale $P(\vec{X} > x)$

> 4. Domanda A.4 Si supponga che 100 abitazioni abbiano tutte il fabbisogno distribuito come X e che siano indipendenti.
> Calcolare (in modo approssimato) la probabilità che l'acqua consumata in una settimana dalle 100 abitazioni ecceda i 1010 ettolitri.

- $E(X_i) = 10$
- $Var(X_i) = 100 = \sigma^2$
- $\sigma = \sqrt{100} = 10$
- $X_1, \dots, X_{100} \sim N(100\mu, 100\sigma^2) \sim N(1000, 10000)$

$$P(\vec{X} \geq 1010) = 1 - P(\vec{X} \leq 1010) = 1 - P\Big(\vec{X} \leq \frac{1010 - n\mu}{\sigma\sqrt{100}}\Big) = 1 - P\Big(\vec{X} \leq \frac{1010-1000}{10\sqrt{100}}\Big) = 1 - \Phi(0,1$$

Il valore di $\Phi(0,1)$ è preso dalla tabella.

Stimatore puntuale

▼ Problema di introduzione

Consideriamo della variabili aleatorie esponenziali e indipendenti $X_1, X_2, \ldots, X_n$, tutte con media θ **incognita**. In questo caso la loro densità congiunta è

$$\begin{aligned} f(x_1, x_2, \ldots, x_n) &= f_{x_1}(x_1) \cdot f_{x_2}(x_2) \ldots \cdot f_{x_n}(x_n) \quad x_i > 0, i = 1, \ldots, n \\ &= \frac{1}{\theta} e^{-\frac{x_1}{\theta}} \cdot \frac{1}{\theta} e^{-\frac{x_2}{\theta}} \ldots \cdot \frac{1}{\theta} e^{-\frac{x_n}{\theta}} \\ \rightarrow f(x_1, x_2, \ldots, x_n) &= \frac{1}{\theta^n} \exp\left(-\frac{1}{\theta} \sum_{i=1}^{n} x_i\right) \end{aligned}$$

Un problema di interesse consiste nello **stimare** θ usando una qualunque statistica.

📌 La **stima puntuale** è lo specifico valore assunto da una statistica, calcolata in corrispondenza dei dati campionari e che viene utilizzata per stimare il vero valore non noto di un parametro di una popolazione

Consideriamo un **campione normale** $X_1, X_2, \ldots, X_n$ di media μ **incognita**. La **statistica**

$$\vec{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

è uno **stimatore** di μ.

θ : **è un parametro** di una **popolazione** (θ può essere parametro di una distribuzione, media, varianza... $\theta = \mu, \sigma^2$...).

T : Si definisce **stimatore** del parametro θ della popolazione, la **statistica** $T = g(X_1, \ldots, X_n)$ utilizzata per stimare θ.

- T è una variabile aleatoria
- T è una statistica
- T è uno stimatore di θ

Errore di stima

Si chiama errore di stima, la differenza fra lo stimatore e il parametro:

$$T - \theta$$

Proprietà degli stimatori

Stimatore corretto

La funzione $T = g(X_1, \ldots, X_n)$ dei dati campionari è uno stimatore **corretto** del parametro θ se il valore atteso dello stimatore è uguale al parametro da stimare:

$$E(T) = \theta$$

per ogni possibile valore del parametro.

Stimatore distorto

Se uno stimatore non è corretto, si dice **distorto**, e la sua **distorsione** è data dalla differenza

$$B(T) = E(T) - \theta$$

- $B(T)$ → è detta distorsione.

Se uno stimatore è **corretto**, la sua **distorsione è uguale a** 0.

▼ Esempio 1

La media campionaria è uno stimatore corretto della media della popolazione.

$$E(\vec{X}) = \mu$$
$$\Rightarrow \theta = \mu$$

▼ Esempio 2

La varianza campionaria è uno stimatore corretto della varianza della popolazione σ^2

$$E(S^2) = \sigma^2$$
$$\Rightarrow \theta = \sigma^2$$

Efficienza - Errore quadratico medio

Una misura della variabilità delle stime fornite da uno stimatore, è fornita dal suo **errore quadratico medio (***o momento secondo dell'errore di stima***)** che corrisponde a:

$$MSE(T) = E((T-\theta)^2)$$

In genere l’errore quadratico medio viene calcolato per confrontare l’efficienza di due diversi stimatori del parametro.

▼ Esempio

Considerati gli stimatori T_1 e T_2 di θ, se risulta

$$MSE(T_1) < MSE(T_2)$$

per ogni possibile valore del parametro θ, si conclude dicendo che T_1 è più efficiente di T_2.

L'errore quadratico medio può anche essere scritto come:

$$MSE(T) = Var(T) + (B(T))^2$$

▼ Dimostrazione

$$\begin{aligned}
MSE(T) &= E((T-\theta)^2) \\
&= E\left(\left[(T-E(T)) + (E(T)-\theta)\right]^2\right) \\
&= E((T-E(T))^2) + E((E(T)-\theta)^2) \\
&\quad + 2E(T-E(T))(E(T)-\theta) \\
&= \underbrace{E((T-E(T))^2)}_{Var(T)} + \underbrace{E((E(T)-\theta)^2)}_{(B(T))^2}, \\
&\quad \left(\text{perché } E(T-E(T)) = E(T)-E(T) = 0\right) \\
&= Var(T) + (B(T))^2
\end{aligned}$$

Determinare uno stimatore corretto tramite l' efficienza

Uno stimatore risulta **corretto** se il suo errore quadratico medio corrisponde alla sua varianza.

$$MSE(T) = Var(T)$$

Consistenza o Coerenza

Uno stimatore si dice **coerente** o **consistente** se, considerato un qualsiasi valore $\epsilon > 0$, risulta:

$$\lim_{n\to\infty} P(|T-\theta| \geq \epsilon) = 0$$

Per ogni possibile valore del parametro θ.

▼ Esempio

Data una popolazione di valore atteso μ e varianza unitaria (cioè $\sigma^2 = 1$), si estragga un campione casuale di 4 elementi e si considerino gli stimatori di μ :

$$T_1 = \frac{1}{2}X_1 + \frac{1}{8}X_2 + \frac{1}{4}X_3 + \frac{1}{8}X_4$$

$$T_2 = \frac{1}{2}\vec{T}$$
$$\vec{T} = \frac{T_1 + T_2 + T_3 + T_4}{4}$$

$$E(T_1) = \frac{1}{2}E(X_1) + \frac{1}{8}E(X_2) + \frac{1}{4}E(X_3) + \frac{1}{8}E(X_4) = \Big(\frac{1}{2} + \frac{1}{8} + \frac{1}{4} + \frac{1}{8}\Big)\mu = \mu$$

$$E(T_2) = \frac{1}{2}\mu$$

Lo stimatore T_1 è **quindi corretto**, mentre T_2 è **distorto**.

Le varianze di due stimatori sono

$$Var(T_1) = \frac{1}{4}Var(X_1) + \frac{1}{64}Var(X_2) + \frac{1}{16}Var(X_3) + \frac{1}{64}Var(X_4) = \frac{16+1+4+1}{64}\sigma^2 = \frac{22}{64}\sigma^2 = \frac{11}{32}, \text{per}$$

$$Var(T_2) = \frac{1}{4}Var(\vec{T}) = \frac{1}{4}\frac{\sigma^2}{4} = \frac{1}{16}$$

Di conseguenza i due errori quadratici medi sono

$$MSE(T_1) = Var(T_1) = \frac{11}{32}$$
$$MSE(T_2) = Var(T_2) + (B(T_2))^2 = \frac{1}{16} + \Big(\frac{1}{2}\mu - \mu\Big)^2 = \frac{1}{16} + \frac{1}{4}\mu^2$$

Intervalli di confidenza

La **stima puntuale** è lo specifico valore assunto da una statistica, calcolata in corrispondenza dei dati campionari e che viene utilizzata per stimare il vero valore non noto di un parametro di una popolazione

Uno **stimatore per intervallo** è un intervallo costruito attorno allo stimatore puntuale, in modo tale che sia nota e fissata la probabilità che il parametro appartenga all'intervallo stesso. Tale probabilità è detta **livello di confidenza** ed è in generale indicato con $(1-\alpha)$ dove $\alpha \in (0,1)$.

Dati una variabile aleatoria $Z \sim \mathcal{N}(0,1)$ e $\alpha \in (0,1)$. Vogliamo trovare $t_{\frac{\alpha}{2}}$ tale che

$$P(-t_{\frac{\alpha}{2}} < Z < t_{\frac{\alpha}{2}}) = 1-\alpha$$
$$2\Phi(t_{\frac{\alpha}{2}}) - 1 = 1-\alpha$$
$$\text{dove } \Phi \text{ è la funzione di ripartizione di } \mathcal{N}(0,1)$$
$$\Phi(t_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$$

Intervalli di confidenza per la media della distribuzione normale (con varianza nota):

$$\vec{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

$$\vec{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Sia $X_1, X_2, \ldots, X_n$ un campione estratto da una popolazione normale di **media incognita** μ e **varianza nota** σ^2**.** Allora

$$Z = \frac{\vec{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0,1) \quad \text{e} \, \forall \alpha \in (0,1)$$

Abbiamo:

$$\begin{aligned}
1 - \alpha &= P(-t_{\alpha/2} < Z < t_{\alpha/2}) \\
&= P(-t_{\alpha/2} < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < t_{\alpha/2}) \\
&= P(-t_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \overline{X} - \mu < t_{\alpha/2}\frac{\sigma}{\sqrt{n}}) \\
&= P(-t_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu - \overline{X} < t_{\alpha/2}\frac{\sigma}{\sqrt{n}}) \\
&= P(\overline{X} - t_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2}\frac{\sigma}{\sqrt{n}})
\end{aligned}$$

Se osserviamo il campione, e registriamo che $\vec{X} = \vec{x}$, allora possiamo dire che con il $(1 - \alpha)$ di confidenza

$$\vec{x} - t_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \vec{x} + t_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

Intervallo di confidenza bilaterale ad un livello di confidenza $1 - \alpha$ per la media μ, è:

$$\left(\vec{x} - t_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \vec{x} + t_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$$

Dove

$$\Phi(t_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$$

e $\vec{x}$ è il valore che si osserva per la media campionaria.

Errore standard della stima

La lunghezza dell'intervallo dipende sia dal livello di copertura desiderato (da cui dipende il quantile $t_{\frac{\alpha}{2}}$), sia dal grado di **precisione dello stimatore** misurato dalla quantità:

$$\frac{\sigma}{\sqrt{n}} \rightarrow \text{Errore standard della stima}$$

▼ Esempio stima puntuale → media, varianza e intervallo di confidenza al 90%

L'azienda PoliPharma ha 80000 dipendenti di sesso femminile e 20000 dipendenti di sesso maschile. Al rientro dalle ferie estive vengono selezionati casualmente tra i dipendenti 100 donne e 100 uomini ai quali viene eseguito un test rapido per la positività al Covid-19. Di questi 10 donne e 15 uomini sono risultati positivi.

D.1 Si stimi la probabilità che una generica dipendente di sesso femminile risulti positiva al test.

$n = 100$

$\bar{x} = \frac{10}{100} = 0{,}1$

$\theta_1 = E(\bar{x}) = \frac{1}{100} \cdot \sum_{i=1}^{100} x_i \to \frac{1}{100} \cdot 10 = 0{,}1$ femmine

D.2 Si stimi la probabilità che un generico dipendente di sesso maschile risulti positivo al test.

$\theta_2 = E(\bar{y}) = \frac{1}{100} \sum_{i=1}^{100} y_i \to \frac{1}{100} \cdot 15 = 0{,}15$ uomini

D.3 Si fornisca un intervallo di confidenza di livello 90% per la probabilità che una generica dipendente di sesso femminile risulti positiva al test

$S^2 =$ Varianza $\to \frac{1}{100-1} \cdot \sum_{i=1}^{10} (1-0{,}1)^2 \to \frac{1}{99} \cdot \frac{81}{10} \to \frac{9}{110} = 0{,}081$

$S =$ deviazione $= \sqrt{0{,}081} = 0{,}28$

errore standard della stima $\to \frac{0{,}28}{\sqrt{100}} = 0{,}028$

Intervallo di confidenza al 90%

$1 - \alpha = 0{,}9 \to \alpha = 0{,}1$

$\Phi(z_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2} \to \Phi(z_{0{,}05}) = 1 - 0{,}05 \to \Phi(z_{0{,}05}) = 0{,}95$

$\hookrightarrow$ dalla Tabella 1,645

$(0{,}1 - 1{,}645 \cdot 0{,}028 \,,\, 0{,}1 + 1{,}645 \cdot 0{,}028)$

$\downarrow$

$(0{,}053 \,,\, 0{,}14)$

D.4 Si stimi il numero atteso di dipendenti che risulterebbero positivi ad un eventuale screening di massa.

$80000 \cdot \underset{\theta_1}{0{,}10} + 20000 \cdot \underset{\theta_2}{0{,}15} = 11\,000$

▼ Esempio stima puntuale → media, varianza e intervallo di confidenza al 95%

- $\sigma^2 = 16$ → varianza nota
- $1 - \alpha = 0.95 = 95\% \to \alpha = 1 - 0.95 = 0.05$
- $n = 10$

- $X_1, \ldots, X_{10}$ è un campione normale $\mathcal{N}(\mu, 16)$
- $x_1, \ldots, x_{10}$ è una realizzazione del campione
- $\vec{X} = \frac{X_1+X_2+\cdots+X_{10}}{10}$
- $\vec{x} = \frac{x_1+x_2+\cdots+x_{10}}{10} = 9.766$ → sono i numeri relativi alle variabili aleatorie (il valore assunto dalle variabili aleatorie).
- $\frac{\sigma}{\sqrt{n}} = \frac{\sqrt{16}}{\sqrt{10}} = 1.265$
- $\Phi(t_{\frac{\alpha}{2}}) = 1 - \frac{0.05}{2} \to \Phi(t_{\frac{\alpha}{2}}) = 1 - 0.025 = 0.975$

$$\Phi(t_{0.025}) = 0.975 \leftarrow \text{trovo nella tabella}$$
$$t_{0.025} = 1.96$$

$$(9.755 - 1.96 * 1.265, 9.755 + 1.96 * 1.265) = (7.287, 12.245)$$

Esempio. Da informazioni derivanti da una precedente analisi, si sa che la durata delle telefonate che arrivano ad un call center si distribuisce con una distribuzione normale di media μ e varianza $\sigma^2 = 16$ minuti quadrati. Si vuole calcolare un intervallo di confidenza al livello $1 - \alpha = 0.95$ per la durata media delle telefonate. A tale scopo, si estrae un campione di $n = 10$ telefonate che fornisce le seguenti durate :

$$7.36, 11.91, 12.91, 9.77, 5.99, 10.91, 9.57, 11.01, 6.11, 12.12$$

Il calcolo dell'intervallo desiderato è a questo punto piuttosto semplice : si calcola dapprima la media campionaria ed il suo errore standard

$$\bar{x} = 9.766; \quad \frac{\sigma}{\sqrt{n}} = \frac{\sigma}{\sqrt{10}} = \frac{4}{\sqrt{10}} = 1.265$$

Se inoltre $1 - \alpha = 0.95$, $(\frac{\alpha}{2} = 0.025)$, il quantile desiderato è dato da $t_{0.025} = 1.96$ $(\Phi(t_{0.025}) = 0.975)$. L'intervallo è dunque dato da

$$(\bar{x} - t_{0.025}\frac{\sigma}{\sqrt{10}}, \bar{x} + t_{0.025}\frac{\sigma}{\sqrt{10}}) = (7.287, 12.245)$$

▼ Esempio stima puntuale → media e varianza note, intervallo di confidenza al 95% e 99%

L'azienda Polidrink produce la bibita energetica Polipower venduta in piccole lattine da 100 ml. Il direttore della produzione sospetta che il valore medio di Polipower inserita nelle lattine non corrisponda al valore nominale. Vengono raccolte 100 lattine a caso e misurato il loro contenuto ottenendo una media campionaria di 101.669 ml. Si assuma deviazione standard nota σ = 2 ml.

$E(\bar{x}) = 101{,}669\ ml \qquad \sigma = 2\ ml \qquad n = 100$

D.1 Si costruisca un intervallo di confidenza 95% per il contenuto medio della singola lattina.

$1-\alpha = 0{,}95 \qquad \alpha = 0{,}05 \qquad \Phi(t_{\alpha/2}) = 1 - 0{,}025 = 0{,}975$ ← dalla Tabella

$\Phi(t_{\alpha/2}) = 1{,}96$

Intervallo $\left(101{,}669 - 1{,}96 \cdot \frac{2}{\sqrt{100}} , 101{,}669 + 1{,}96 \cdot \frac{2}{\sqrt{100}}\right) \to (101{,}277 , 102{,}061)$

D.2 Un intervallo di confidenza al 99% risulterà più ampio o meno ampio di quello costruito al punto D.1?

Un intervallo di confidenza maggiore, è più ampio

$1-\alpha = 0{,}99 \qquad \alpha = 0{,}01$

$\Phi(t_{\alpha/2}) = 1 - 0{,}005 = 0{,}995 \qquad \Phi(t_{\alpha/2}) = 2{,}575$

Ha ampiezza maggiore

$\left(101{,}669 - 2{,}575 \cdot \frac{2}{\sqrt{100}} , 101{,}669 + 2{,}575 \cdot \frac{2}{\sqrt{100}}\right) \to (101{,}154 , 102{,}184)$

Un comitato di garanzia vuole verificare che il contenuto nominale dichiarato dall'azienda corrisponda al reale contenuto medio delle lattine. Specificare ipotesi nulla e alternativa e trarre una conclusione in merito con α = 5%

Soluzione: Poiché 100 non è contenuto nell'intervallo di confidenza costruito al punto 1 si rifiuta l'ipotesi nulla H0 : μ = 100 in favore dell'ipotesi alternativa H1 : μ ≠ 100

Intervallo di confidenza unilaterale (destro e sinistro) ad un livello $1-\alpha$ per μ

$$\left(\vec{x} - t_\alpha \frac{\sigma}{\sqrt{n}}, \infty\right) \to \text{destro}$$

$$\text{sinistro} \to \left(-\infty, \vec{x} + t_\alpha \frac{\sigma}{\sqrt{n}}\right)$$

▼ Dimostrazione

$$\begin{aligned} 1-\alpha &= P(Z < t_\alpha) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < t_\alpha\right) \\ &= P\left(\frac{\mu-\overline{X}}{\sigma/\sqrt{n}} > -t_\alpha\right) \\ &= P\left(\mu > \overline{X} - t_\alpha \frac{\sigma}{\sqrt{n}}\right) \end{aligned}$$

$$\Rightarrow 1-\alpha = P\left(\mu < \vec{X} + t_\alpha \frac{\sigma}{\sqrt{n}}\right)$$

Dall'identità

$$1-\alpha = P(Z < t\alpha)$$
$$\Rightarrow 1-\alpha = \Phi(t\alpha)$$

Analogamente abbiamo

$$1-\alpha = P(Z > -t\alpha)$$
$$\Rightarrow 1-\alpha = \Phi(t\alpha)$$

▼ Dimostrazione

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \rightarrow P(Z > -t\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-t\alpha}^{\infty} e^{-\frac{x^2}{2}} dx$$

$$\boxed{y = -x \rightarrow dy = -dx \rightarrow \ dx = -dy}$$

$$\frac{1}{\sqrt{2\pi}} \int_{t\alpha}^{-\infty} e^{-\frac{-y^2}{2}} (-dy)$$

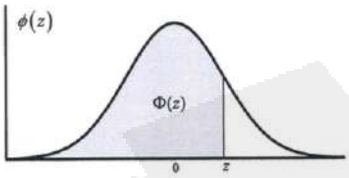
$$\rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t\alpha} e^{-\frac{-y^2}{2}} dy = \Phi(t\alpha)$$

Tabella funzione di ripartizione della variabile normale standardizzata

Dove le righe indicano il punto del quale vogliamo conoscere il valore della funzione di ripartizione e le colonne indicano $\Phi(x + colonna)$ → 2° colonna = $\Phi(x + 0.02)$

$\Phi(z) =$ area sotto la curva fino a $z = 1$

Tavola 1 – Funzione di ripartizione della variabile casuale normale standardizzata $N(0,1)$.



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Esempio lettura dati dalla tabella

- il valore di $\Phi(0.2) = 0.579$
- Il valore di $\Phi(2.52) = 0.9941$

Intervalli di confidenza per la media della distribuzione normale (con varianza NON nota)

S^2 è uno stimatore corretto di σ^2

Teorema di Student

Siano $X_1, \dots, X_n$, n variabili aleatorie i.i.d, aventi ciascuna una distribuzione normale con media μ e varianza σ^2. Allora

$$T = \frac{\vec{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

ha una distribuzione t di Student con $n-1$ gradi di libertà ($T \sim t_{n-1}$), dove $\vec{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ e $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \vec{X})^2$.

Consideriamo ora un campione $X_1, \dots, X_n$ estratto da una popolazione normale di **media incognita** μ e **varianza non nota** σ^2. Vogliamo costruire un intervallo di confidenza per μ ad un livello di $1-\alpha$. Ricordiamo che la distribuzione di

$$T = \frac{\vec{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

Quindi abbiamo per ogni $\alpha \in (0, 1)$,

$$\begin{aligned} 1-\alpha &= P(-t_{\alpha/2,n-1} < \frac{\overline{X} - \mu}{S/\sqrt{n}} < t_{\alpha/2,n-1}) \\ &= P(-t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} < \overline{X} - \mu < t_{\alpha/2,n-1}\frac{\sigma}{\sqrt{n}}) \\ &= P(\overline{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}) \end{aligned}$$

Se i valori osservati sono $\vec{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ e $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \vec{x})^2$, allora

Intervallo di confidenza bilaterale ad un livello di confidenza di $1-\alpha$ per la media μ.

$$\left(\vec{x} - t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}, \vec{x} + t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}}\right)$$

Si ha:

$$1-\alpha = P(-t_{\frac{\alpha}{2},n-1} < T < t_{\frac{\alpha}{2},n-1}) = 2\Phi t_{n-1}(t_{\frac{\alpha}{2},n-1}) - 1$$
$$\Rightarrow \Phi t_{n-1}(t_{\frac{\alpha}{2},n-1}) = 1 - \frac{\alpha}{2}$$

Analogamente

Intervallo di confidenza unilaterale (destro e sinistro) ad un livello $1-\alpha$ per μ

$$\left(\vec{x} - t_\alpha\frac{s}{\sqrt{n}}, \infty\right) \rightarrow \text{destro}$$

$$\text{sinistro} \rightarrow \left(-\infty, \vec{x} + t_\alpha\frac{s}{\sqrt{n}}\right)$$

▼ Esempio

$$1-\alpha = 0.95 \rightarrow \alpha = 0.05$$
$$\frac{\alpha}{2} = 0.025$$

$$\Phi t_{24}(t_{0.025,24}) = 1 - 0.025$$
$$\Phi t_{24}(t_{0.025,24}) = 0.975 \leftarrow \text{vedo nella tabella t di student}$$
$$t_{0.025} = 2.064$$

Esempio. Si vuole stimare l'età media degli utenti di una biblioteca civica. A questo scopo si seleziona un campione di una popolazione normale composto da $n = 25$ persone avente media $\bar{x} = 29$ anni e deviazione standard $s = 8$ anni. Siccome

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{24}$$

allora un intervallo di confidenza per l'età media μ ad un livello di confidenza 95% è dato da

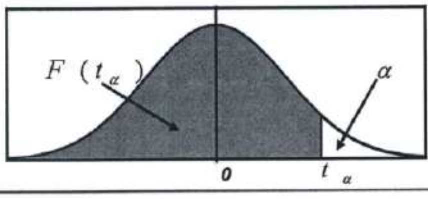
$$\begin{aligned}
&(\bar{x} - t_{0.025,24}\frac{s}{\sqrt{n}}, \bar{x} + t_{0.025,24}\frac{s}{\sqrt{n}}) \\
&= (29 - 2.064.\frac{8}{\sqrt{25}}, 29 + 2.064.\frac{8}{\sqrt{25}}) \\
&= (25.69, 32.30),
\end{aligned}$$

dove $\phi_{t_{24}}(t_{0.025,24}) = 0.975$, cioè $t_{0.0.25,24} = 2.064$

Tabella t di Student

$$F(t_\alpha) = 1 - \alpha$$

t di Student



$F(t_\alpha)$ / gr.lib.	0,700	0,750	0,800	0,850	0,900	0,950	0,975	0,990	0,995	0,999	0,9995
1	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	318,309	636,619
2	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	22,327	31,599
3	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	10,215	12,924
4	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	5,893	6,869
6	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	4,785	5,408
8	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,485	3,768
24	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,396	3,659
30	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,385	3,646
31	0,530	0,682	0,853	1,054	1,309	1,696	2,040	2,453	2,744	3,375	3,633
32	0,530	0,682	0,853	1,054	1,309	1,694	2,037	2,449	2,738	3,365	3,622
33	0,530	0,682	0,853	1,053	1,308	1,692	2,035	2,445	2,733	3,356	3,611
34	0,529	0,682	0,852	1,052	1,307	1,691	2,032	2,441	2,728	3,348	3,601
35	0,529	0,682	0,852	1,052	1,306	1,690	2,030	2,438	2,724	3,340	3,591
36	0,529	0,681	0,852	1,052	1,306	1,688	2,028	2,434	2,719	3,333	3,582
37	0,529	0,681	0,851	1,051	1,305	1,687	2,026	2,431	2,715	3,326	3,574
38	0,529	0,681	0,851	1,051	1,304	1,686	2,024	2,429	2,712	3,319	3,566
39	0,529	0,681	0,851	1,050	1,304	1,685	2,023	2,426	2,708	3,313	3,558
40	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,307	3,551
80	0,526	0,678	0,846	1,043	1,292	1,664	1,990	2,374	2,639	3,195	3,416
120	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,160	3,373
160	0,525	0,676	0,844	1,040	1,287	1,654	1,975	2,350	2,607	3,142	3,352
200	0,525	0,676	0,843	1,039	1,286	1,653	1,972	2,345	2,601	3,131	3,340
∞	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,090	3,291

Esempio lettura dati dalla tabella

- $F_{T_3}(0.765) = 0.750$ → dove **3 sono i gradi di libertà** e 0.765 è il valore del quale vogliamo conoscere il valore della **funzione t di Student**.
- $F_{T_2}(0.816) = 0.750$ → dove **2 sono i gradi di libertà** e 0.816 è il valore del quale vogliamo conoscere il valore della **funzione t di Student.**