

Interpretazione dell'intervallo di confidenza

Interpretazione dell'intervallo di confidenza

- Per attribuire l'interpretazione corretta all'intervallo di confidenza per μ dobbiamo supporre di estrarre dalla popolazione tutti i possibili campioni aventi numerosità n e di costruire tutti i possibili intervalli di confidenza (uno per ciascun campione).

In queste circostanze,

- una frazione uguale a $1 - \alpha$ degli intervalli di confidenza conterrà il valore μ , e
- la rimanente frazione α non lo conterrà.

Interpretazione dell'intervallo di confidenza

- Per questa ragione assegnamo un livello di fiducia pari a $1 - \alpha$ all'affermazione secondo cui l'**intervallo di confidenza** $\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$ contiene il vero valore μ della media della popolazione.

Interpretazione dell'intervallo di confidenza

L'interpretazione corretta dell'intervallo di confidenza per μ al 95% potrebbe dunque essere formulata in questi termini:

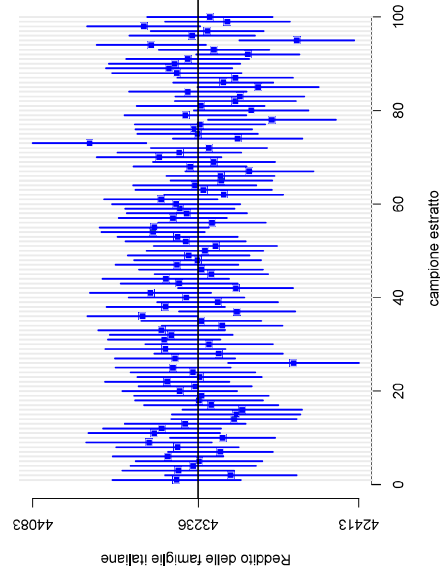
il ricercatore attribuisce un grado di fiducia al 95% all'affermazione secondo cui la media μ della popolazione è contenuta nell'intervallo compreso tra 38332 e 43868 nel senso che ha usato una procedura la quale produce la risposta corretta nel 95% dei campioni casuali di numerosità $n = 100$ e la risposta sbagliata nel restante 5% dei campioni;

il ricercatore però non può mai sapere se l'intervallo costruito utilizzando uno specifico campione contenga o meno il vero valore μ della media della popolazione.

Interpretazione dell'intervallo di confidenza

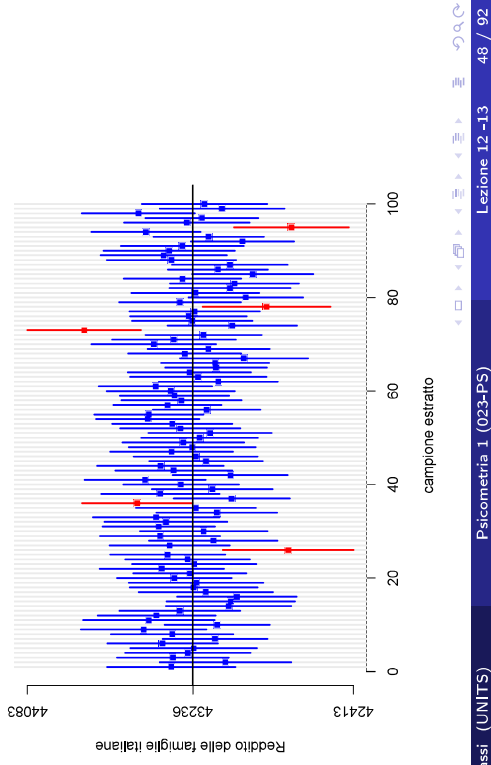
- Nella figura seguente sono presentati i risultati di una simulazione in cui 100 campioni casuali di numerosità $n = 100$ vengono estratti da una popolazione $\approx N(43236, 15500)$
- Per ciascun campione i – *esimo* vengono calcolate la media \bar{x}_i e la deviazione standard s_i .
- Usando queste informazioni, vengono calcolati 100 intervalli di confidenza al 95% per μ .
- Gli intervalli di confidenza sono rappresentati nella figura con dei segmenti verticali. La linea orizzontale rappresenta il reddito medio $\mu = 43236$
- I segmenti rossi rappresentano gli intervalli di confidenza al 95% che NON contengono la media della popolazione.

Simulazione su 100 campioni

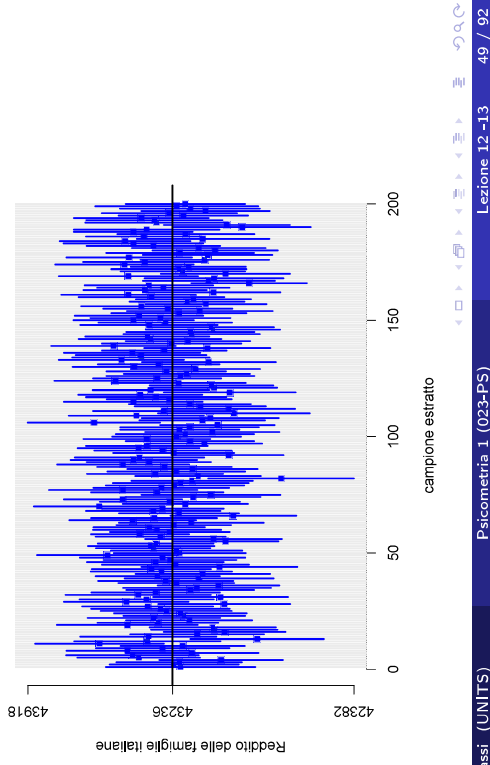


Simulazione su 100 campioni

5 (5%) non contengono la media $\mu = 43236$

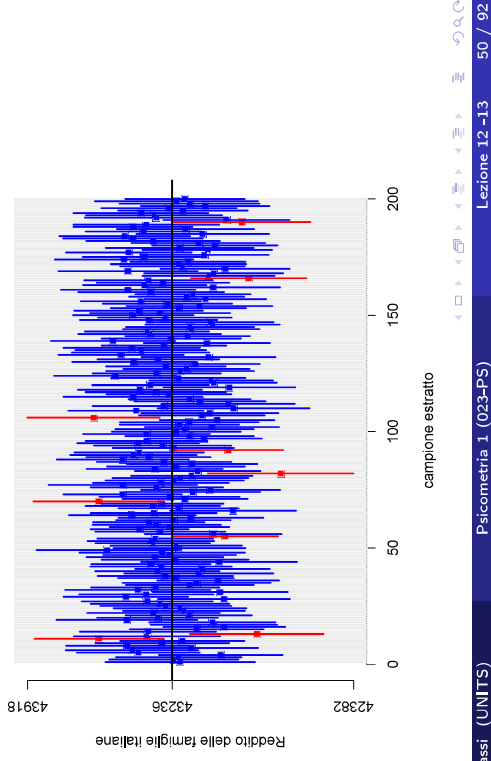


Simulazione su 200 campioni



Simulazione su 200 campioni

9 ($\approx 5\%$) non contengono la media $\mu = 43236$



Probabilità d'errore

- La probabilità che un intervallo di confidenza NON contenga il parametro è chiamata **probabilità d'errore**.
- La probabilità d'errore è denotata da α .
- Il coefficiente di confidenza C è uguale a $(1 - \alpha)$

Per un intervallo di confidenza al 95%, per esempio, $C = 0.95$ e $\alpha = 0.05$.

Interpretazione errata

Le seguenti interpretazioni dell'intervallo di confidenza NON sono corrette.

- La media μ della popolazione ha una probabilità 0.95 di essere contenuta nell'intervallo

$$\bar{x} \pm t_{(\alpha/2, n-1)} \hat{\sigma}_{\bar{x}} = 41100 \pm 2768$$

La media della popolazione è contenuta nell'intervallo (nel qual caso la probabilità che sia compresa nell'intervallo è 1), oppure non è contenuta nell'intervallo (nel qual caso la probabilità che sia contenuta nell'intervallo è 0).