Stima per intervalli - Intervalli di Confidenza(Affidabilità)

Leonardo Bizzoni

May 11, 2024

1 Stima per intervalli della Media

1.1 Stima per intervalli della Media con varianza nota

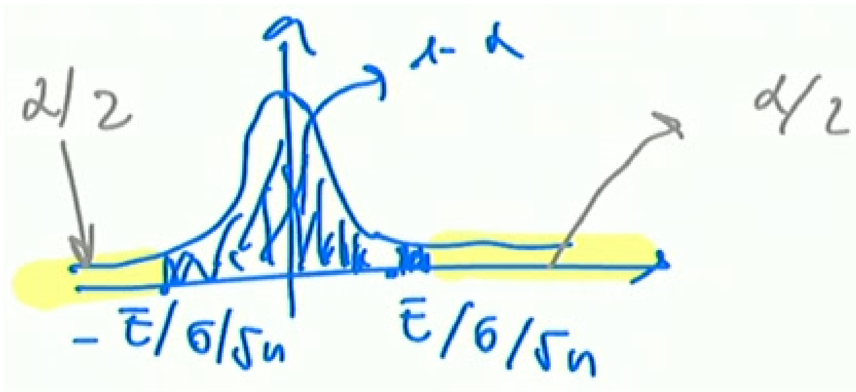
Sia $X_1, \dots, X_n$ un campione aleatorio con distribuzione normale $N(\mu, \sigma^2)$ con media μ incognitica e varianza σ^2 nota.

Sappiamo che la media campionaria è uno stimatore puntuale non distorto di μ e quindi al crescere di n l'errore E commesso nell'utilizzo dello stimatore diminuisce, il che rende la stima *più affidabile*. Andando a costruire un **intervallo centrato in $\mathbf{\overline{X_n}}$** possiamo aumentare notevolmente l'affidabilità della stima.

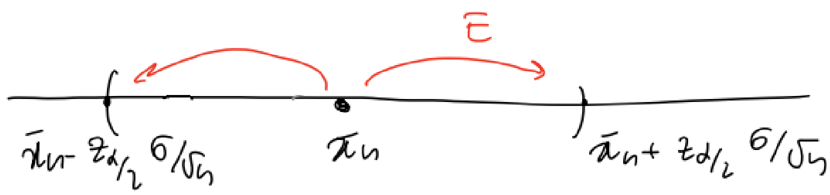
Si vuole costruire un intervallo centrato(*simmetrico*) in $\overline{X_n}$ a cui μ appartenga con probabilità $1 - \alpha$, con α piccolo:

$$P\left(\mu \in \left[\overline{X_n} - E,\ \overline{X_n} + E\right]\right) = 1 - \alpha$$

$$P\left(\left|\overline{X_n} - \mu\right| < E\right) = 1 - \alpha \ \rightarrow\ P\left(\frac{\left|\overline{X_n} - \mu\right|}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) < \frac{E}{\frac{\sigma}{\sqrt{n}}}\right) = 1 - \alpha \ \rightarrow$$
$$P\left(|Z| < \frac{E}{\frac{\sigma}{\sqrt{n}}}\right) = 1 - \alpha$$



Da cui possiamo concludere che $P\left(|Z| < \frac{E}{\frac{\sigma}{\sqrt{n}}}\right) = 1-\alpha \ \rightarrow \ P\left(Z > \frac{E}{\frac{\sigma}{\sqrt{n}}}\right) = \frac{\alpha}{2}$ quindi $E = z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ e l'**intervallo di confidenza** per μ di livello $100(1-\alpha)\%$ sarà $\left(\overline{X_n} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \ \overline{X_n} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$ ($z_{\frac{\alpha}{2}}$ *rappresenta* $l'\frac{\alpha}{2}-$*esimo percentile*).



La **bontà** della stima dipende:

- dal livello di confidenza: maggiore è, più **affidabile** è la stima (*vedi esempio*)
- dall'ampiezza dell'intervallo: minore è $2E$, più **precisa** è la stima (*vedi esempio*)

L'ampiezza dell'intervallo $2E = 2z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ cresce al diminuire di α, ovvero all'aumentare del livello di confidenza, e diminuisce al crescere del campione.

1.1.1 Esempio

Supponiamo che la statura dei piloti sia distribuita secondo una legge normale $N(\mu,^2)$. Si vuole stimare μ a partire dalle osservazioni della statura di 100 piloti. In prima approssimazione si suppone $\sigma = 6.1cm$. Fornire una stima puntuale di μ e calcolare un intervallo di confidenza di livello 95% nel caso in cui la statura media rilevata dalle 100 osservazioni sia pari a $178.5cm$.

```
n <- 100
sd <- 6.1
mc <- 178.5
alpha1 <- 0.05
alpha2 <- 0.01

err1 <- qnorm(1-alpha1) * sd / sqrt(n)
err2 <- qnorm(1-alpha2) * sd / sqrt(n)

c(c(mc-err1, mc+err1), c(mc-err2, mc+err2))
```

Aumentando il livello di confidenza da 95% a 99% l'intervallo di confidenza si ingrandisce da (177.5, 179.5) a (177.0, 179.9), quindi aumentando la probabilità che il vero valore μ sia all'interno dell'intervallo.

```
n1 <- 100
n2 <- 1000
sd <- 6.1
mc <- 178.5
alpha <- 0.05

err1 <- qnorm(1-alpha) * sd / sqrt(n1)
err2 <- qnorm(1-alpha) * sd / sqrt(n2)

c(c(mc-err1, mc+err1), c(mc-err2, mc+err2))
```

Aumentando la grandezza del campione(**osservazioni**), l'intervallo di confidenza diventa più preciso da (177.5, 179.5) a (178.2, 178.8).

1.2 Stima per intervalli della Media con varianza incognita

Sia $X_1, \ldots, X_n$ un campione aleatorio con distribuzione normale $N(\mu, \sigma^2)$ con media μ incognitica e varianza σ^2 incognita.

Sappiamo che la media campionaria è uno stimatore puntuale non distorto di μ e utilizziamo S_n^2 come stimatore della varianza. Andando a costruire un **intervallo centrato in $\overline{\mathbf{X}_\mathbf{n}}$** possiamo aumentare notevolmente l'affidabilità della stima. Ora si procede analogamente al caso di varianza nota

$$P\left(\left|\overline{X_n}-\mu\right|<E\right)=1-\alpha \;\rightarrow\; P\left(\frac{\left|\overline{X_n}-\mu\right|}{\frac{\sqrt{S_n^2}}{\sqrt{n}}}\sim t(n-1)<\frac{E}{\frac{\sqrt{S_n^2}}{\sqrt{n}}}\right)=1-\alpha \;\rightarrow$$
$$P\left(|T|<\frac{E}{\frac{\sqrt{S_n^2}}{\sqrt{n}}}\right)=1-\alpha$$

Da cui possiamo concludere che $P\left(|T|<\frac{E}{\frac{\sqrt{S_n^2}}{\sqrt{n}}}\right)=1-\sqrt{S_n^2} \;\rightarrow\; P\left(T>\frac{E}{\frac{\sqrt{S_n^2}}{\sqrt{n}}}\right)=\frac{\sqrt{S_n^2}}{2}$ quindi $E=t_{n-1,\frac{\alpha}{2}}\frac{\sqrt{S_n^2}}{\sqrt{n}}$ e l'**intervallo di confidenza** per μ di livello $100(1-\alpha)\%$ sarà $\left(\overline{X_n}-t_{n-1,\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\ \overline{X_n}+t_{n-1,\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$ ($t_{n-1,\frac{\alpha}{2}}$ *rappresenta l'$\frac{\alpha}{2}$−esimo percentile*).

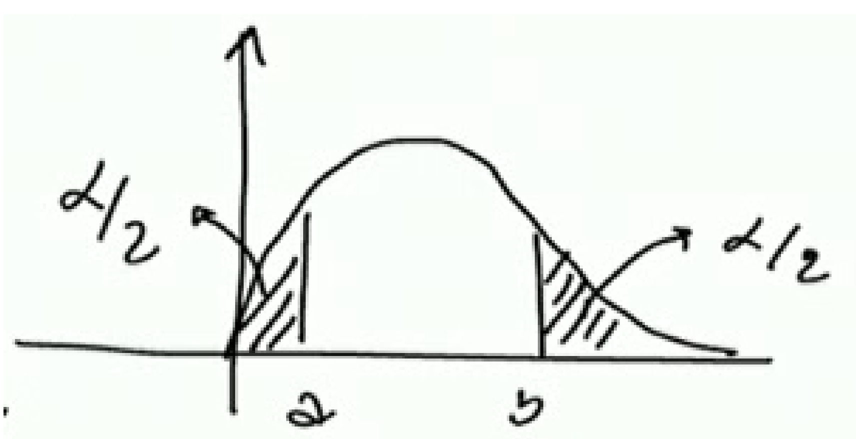
2 Stima per intervalli della Varianza

2.1 Stima per intervalli della Varianza con media incognita

Sia $X_1,\ldots,X_n$ un campione aleatorio con distribuzione normale $N(\mu,\sigma^2)$ con media μ incognitica e varianza σ^2 nota. Sappiamo che la media campionaria è uno stimatore puntuale non distorto di μ e utilizziamo S_n^2 come stimatore della varianza.

Si vuole costruire un intervallo tale per cui la coda sinistra è uguale alla coda destra con probabilità $1-\alpha$, con α piccolo:

$$P\left(a<\frac{(n-1)S_n^2}{\sigma^2}\sim\chi^2(n-1)<b\right)=1-\alpha$$



Da cui possiamo ricavare che $P\left(Y \sim \chi^2(n-1) > b\right) = P\left(0 < Y \sim \chi^2(n-1) < a\right) = \frac{\alpha}{2}$, quindi $b = \chi^2_{n-1,\frac{\alpha}{2}}$ e $a = \chi^2_{n-1,1-\frac{\alpha}{2}}$, a e b sono l'uno il percentile opposto dell'altro. L'intervallo di confidenza al $100(1-\alpha)\%$ sarà:

$$P\left(\chi^2_{n-1,1-\frac{\alpha}{2}} < \frac{(n-1)S_n^2}{\sigma^2} < \chi^2_{n-1,\frac{\alpha}{2}}\right) = 1-\alpha \;\rightarrow$$

$$P\left(\chi^2_{n-1,1-\frac{\alpha}{2}} * \sigma^2 < (n-1)S_n^2 < \chi^2_{n-1,\frac{\alpha}{2}} * \sigma^2\right) = 1-\alpha \;\rightarrow$$

$$\begin{cases} P\left(\chi^2_{n-1,1-\frac{\alpha}{2}} * \sigma^2 < (n-1)S_n^2\right) \\ P\left((n-1)S_n^2 < \chi^2_{n-1,\frac{\alpha}{2}} * \sigma^2\right) \end{cases} \rightarrow \begin{cases} P\left(\sigma^2 < \frac{(n-1)S_n^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}\right) \\ P\left(\sigma^2 > \frac{(n-1)S_n^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right) \end{cases} \rightarrow$$

$$\left(\frac{(n-1)S_n^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{(n-1)S_n^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right)$$

2.2 Stima per intervalli della Varianza con media nota

Identico al caso con media incognita ma i gradi di libertà aumentano da $n-1$ ad n.