Esercitazione 13

Statistica

Alfonso Iodice D'Enza

iodicede@unina.it

Università degli studi di Cassino

Outline

Outline

Ex.1: IC sulla media

Si suppponga di dover comunicare un segnale da una sorgente A ad una destinazione B. Il segnale emesso ha intensità μ. L'intensità del segnale viene percepito in B secondo una distribuzione Normale con media μ e deviazione standard $\sigma = 3$. Quindi, per effetto dei disturbi alla tramissione, l'intensità del segnale in B differisce da quella in A con media $\mu = 0$ e scarto quadratico medio $\sigma = 3$. Si consideri di aver effettuato $n = 10$ trasmissioni tra A e B registrando l'intensità a destinazione.

$$\{17, 21, 20, 18, 19, 22, 20, 21, 16, 19\}$$

- costruire l'intervallo di confidenza all 95%
- costruire l'intervallo di confidenza all 90%
- costruire l'intervallo di confidenza all 99%

Ex.1: IC sulla media

Esercitazione 13

A. Iodice

Svolgimento

Si consideri lo stimatore media campionaria $\bar{X}$, la cui media è $\mu_{\bar{X}} = \mu$ e il cui scarto quadratico medio è $\sigma_{\bar{X}} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$. Il valore standardizzato di $\bar{X}$ è

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \sqrt{n}\frac{\bar{X} - \mu}{\sigma}$$

dunque, il livello di confidenza $(1 - \alpha)$ corrisponde a

$$P(\frac{\sqrt{\sigma}}{n} \mid \bar{X} - \mu \mid \leq Z_{\alpha/2}) = 1 - \alpha$$

equivalentemente

$$P(\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

Ex.1: IC sulla media

Svolgimento

- costruire l'intervallo di confidenza all 95%

il livello di confidenza $(1-\alpha)=0.95$ corrisponde a

$$P(\bar{X}-Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}<\mu<\bar{X}+Z_{\alpha/2}\frac{\sigma}{\sqrt{n}})=0.95$$

da cui, gli estremi dell'intervallo di confidenza sono $\bar{X}\pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

$$\bar{X}=\frac{17+21+20+18+19+22+20+21+16+19}{10}=19.3$$

Inoltre $\sigma=3$, $n=10$ e $Z_{\alpha/2}=Z_{0.025}=1.96$. E' ora possibile calcolare gli estremi

$$\bar{X}\pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}=19.3\pm 1.96\frac{3}{\sqrt{10}}=19.3\pm 1.86$$

La stima intervallare è dunque $[17.44,21.16]$

Ex.1: IC sulla media

Svolgimento

- costruire l'intervallo di confidenza all 90%

il livello di confidenza $(1-\alpha) = 0.90$ corrisponde a
Gli elementi per costruire l'intervallo sono: $\bar{X} = 19.3$, $\sigma = 3$, $n = 10$ e $Z_{\alpha/2} = Z_{0.05} = 1.645$. E' ora possibile calcolare gli estremi

$$\bar{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 19.3 \pm 1.645\frac{3}{\sqrt{10}} = 19.3 \pm 1.56$$

La stima intervallare è dunque $[17.74, 20.86]$

- costruire l'intervallo di confidenza all 90%

In questo caso $Z_{\alpha/2} = Z_{0.005} = 2.576$

$$\bar{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 19.3 \pm 2.576\frac{3}{\sqrt{10}} = 19.3 \pm 2.44$$

La stima intervallare è dunque $[16.86, 21.74]$

Ex.2: IC sulla media; determinazione numerosità campionaria

Dato un intervallo di confidenza, fissata l'ampiezza dell'intervallo a b, si determini la numerosità n del campione. Ricordando che gli estremi dell'intervallo sono dati da $\bar{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ pertanto l'ampiezza dell'intervallo è $2 \times Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

$$2 \times Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = b$$

dunque la numerosità è data da

$$2 \times Z_{\alpha/2}\frac{\sigma}{b} = \sqrt{n} \Longrightarrow \left(2 \times Z_{\alpha/2}\frac{\sigma}{b}\right)^2 = n$$

Quale deve essere la numerosità per ottenere un intervallo di confidenza sulla media μ al 95% di ampiezza $b = 0.01$, tenuto conto che $\sigma = 2$?

Svolgimento

$$n = \left(2 \times 1.96\frac{2}{0.01}\right)^2 = (78.4)^2 = 6146.4$$

Il che vuol dire che la numerosità deve essere $n = 6147$

Ex.3: IC sulla media; varianza non nota

...prima di cominciare

Se lo scarto quadratico medio σ della popolazione non è noto occorre stimarlo attraverso lo scarto quadratico medio s della media campionaria $\bar{X}$. Dunque,

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

La quantità standardizzata è

$$T_{n-1} = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

e si distribuisce secondo una distribuzione t di Student con $n-1$ gradi di libertà.

Ex.3: IC sulla media; varianza non nota

Un'agenzia deve valutare il grado di concentrazione di una sostanza tossica, il PCB, nel latte materno. Per fare questo, viene considerato un campione di 20 madri e studiato il grado di concentrazione della sostanza nel latte. I valori risultanti dalle analisi sono, in parti per milione,

$$\{16, 0, 0, 2, 3, 6, 8, 2, 5, 0, 12, 10, 5, 7, 2, 3, 8, 17, 9, 1\}$$

media e scarto quadratico medio campionari sono

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 5.8 \text{ e } s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}} = 5.085$$

costruire intervalli di confidenza al 95% e al 99%.

Ex.3: IC sulla media; varianza non nota

Svolgimento

Essendo la varianza non nota

$$T_{n-1} = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

Dunque, valgono le seguenti relazioni

$$P(\frac{s}{\sqrt{n}} \mid \bar{X} - \mu \mid \leq t_{\alpha/2}) = 1 - \alpha$$

equivalentemente

$$P(\bar{X} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}) = 1 - \alpha$$

da cui, gli estremi dell'intervallo di confidenza sono $\bar{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$

- gli estremi dell'intervallo di confidenza al 95%

Poichè risulta $\bar{X} = 5.8$,$s = 5.085$,$n = 20$ e $t_{\alpha/2} = 2.093$

$$\bar{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} = 5.8 \pm 2.093\frac{5.085}{\sqrt{20}} = 5.8 \pm 2.38$$

le stime ad intervallo sono dunque $[3.42, 8.18]$

- gli estremi dell'intervallo di confidenza al 99%

Poichè risulta $\bar{X} = 5.8$,$s = 5.085$,$n = 20$ e $t_{\alpha/2} = 2.861$

$$\bar{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} = 5.8 \pm 2.861\frac{5.085}{\sqrt{20}} = 5.8 \pm 3.25$$

le stime ad intervallo sono dunque $[2.55, 9.05]$

Ex.4: IC sulla proporzione

Sia p la proporzione di unità statistiche della popolazione presentano una certa caratteristica. La statistica campionaria corrispondente è la proporzione campionaria $\hat{p} = \frac{x}{n}$, dove x rappresenta il numero di unità nel campione che presentano una determinata caratteristica. $\hat{p}$ è tale che $\mu_{\hat{p}} = p$ e $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$. Ricorrendo all'approssimazione della binomiale alla normale, gli estremi dell'intervallo sono dati da

$$\hat{p} \pm Z_{\alpha/2}\sigma_{\hat{p}}$$

essendo p incognito si stima lo scarto quadratico medio $\sigma_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$; dunque

$$\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Ex.4: IC sulla proporzione

Su un campione di 100 studenti universitari, 82 hanno terminato gli esami del primo anno negli appelli previsti.

- Si costruisca un intervallo di confidenza per p al 99%.

Svolgimento

In base ai dati del problema, $\hat{p} = \frac{82}{100}$

$$\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.82 \pm 2.576\sqrt{\frac{0.82(1-0.82)}{100}} = 0.82 \pm 0.099$$

dunque le stime intervallari sono $[0.721, 0.919]$.