Lezione 28

09/12/24

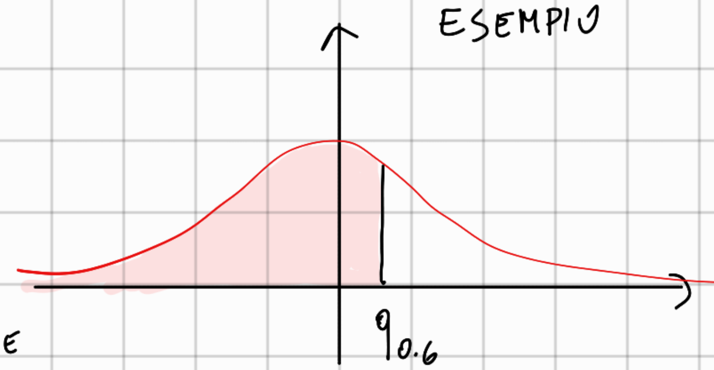
DEFINIZIONE

Sia X V.A. SU UNO SPAZIO DI PROBABILITÀ CONTINUO Ω

Dato $\alpha \in [0,1]$ il $\boxed{\text{QUANTILE DI ORDINE } \alpha}$ $q_\alpha \in \Omega$ è quel valore tale che

$$P(X \leq q_\alpha) = \alpha$$

ESEMPIO



NEL CASO DI VARIABILE NORMALE

$$P(X \leq q_\alpha) = \int_{-\infty}^{q_\alpha} f_{\mu,\sigma}(x)\,dx = \alpha$$

$N \sim (\mu, \sigma^2)$

$$q_{0.5} = \mu$$

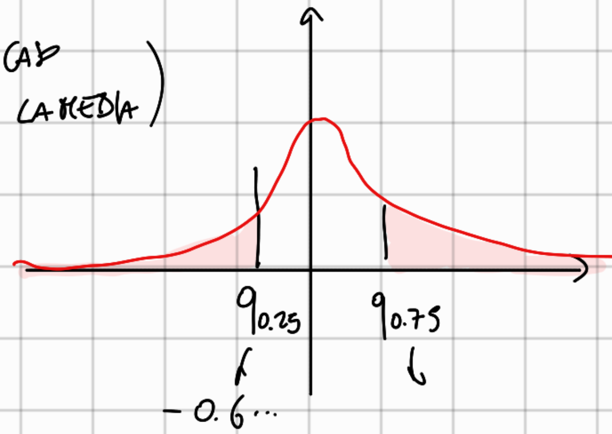
È LA MEDIANA (IN QUESTO CASO COINCIDE CON LA MEDIA)

$Z \sim N(0,1)$

$$q_{0.25} = \boxed{-0.674}$$

$$q_{0.75} = 0.674$$

$0.75 = 1 - 0.25$



$$P(q_{0.25} \leq Z \leq q_{0.75}) = \frac{1}{2}$$

$$P(Z \le q_{0.05}) = 0.05$$

$$q_{0.05} = -1.645$$

$$q_{0.95} = 1.645$$

$$P(-1.645 \le Z \le 1.645) = 0.9 = 1-2(0.05)$$

$-q_{0.95}$ $\qquad$ $\boxed{q_{0.95}} \rightarrow q_{1-\frac{\alpha}{2}}$ $\qquad$ $1-\alpha$

$\alpha = 0.1$

$$\boxed{P\left(-q_{1-\frac{\alpha}{2}} \le Z \le q_{1-\frac{\alpha}{2}}\right) = 1-\alpha}$$

$$\left[\begin{array}{lll} P(-1.645 \le Z \le 1.645) = 0.9 & & (90\,\%) \\ P(-1.960 \le Z \le 1.960) = 0.95 & & (95\,\%) \\ P(-2.576 \le Z \le 2.576) = 0.99 & & (99\,\%) \\ P(-3.291 \le Z \le 3.291) = 0.999 & & (99.9\,\%) \end{array}\right.$$

STATISTICA INFERENZIALE

$N = \#$ DI INDIVIDUI NELLA POPOLAZIONE

Solitamente non abbiamo a disposizione tutte le misure sulla popolazione $x_1, \dots x_N$

Ma possiamo estrarre un campione di n unità statistiche

Facciamo un esperimento per conoscere

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

NON LA POSSIAMO CALCOLARE

$n \ll N$

SIGNIFICA

$(N + n \approx N)$

$\left(\frac{n}{N} \approx 0\right)$

VARIANZA DELLA POPOLAZIONE

$$\boxed{\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

PER FARLO

ESTRAIAMO UN CAMPIONE $x_1 \dots x_n$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \longleftarrow \text{È UNO STIMATORE PER } \mu$$

PER IL TEOREMA DEL LIMITE CENTRALE

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

PER UNA PROPRIETÀ DELLA NORMALE

$$\bar{x} - \mu \sim N\left(0, \frac{\sigma^2}{n}\right)$$

$$Z = \frac{\bar{x} - \mu}{\sigma}\sqrt{n} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

PER ESEMPIO È UTILE PER DIRE CHE

$$P\left(-q_{1-\frac{\alpha}{2}} < \frac{\bar{x} - \mu}{\sigma}\sqrt{n} < q_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$P\left(-q_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \bar{x}-\mu < q_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$P\left(\underbrace{\mu - q_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}} < \bar{x} < \underbrace{\mu + q_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}}\right) = 1-\alpha$$

$$P\left(x - \underbrace{q_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}} < \mu < \bar{x} + q_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

NON CONOSCO IL VALORE DI μ MA POSSO DIRE QUALCOSA USANDO IL CAMPIONE

$$P(\bar{x} \in [\ ,\ ]) = 1-\alpha$$

ESEMPIO: CAVIE DA LABORATORIO $N \sim (\mu, \sigma^2 = 6)$

$n = 15 \rightarrow$ PESARE 15 CAVIE

28 32 37 29 31 32 26 32 27 29 30 28 31 31 30

x_1 x_2 x_3

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 30.2$$

$\alpha = 0.1$ $\quad 1-\alpha = 0.9$

$$q_{0.95}\frac{\sigma}{\sqrt{n}} = 1.645 \cdot \frac{\sqrt{6}}{\sqrt{15}} \approx 1.0$$

$[30.2 - 1.0,\ 30.2+1]$

$[29.2, 31.2]$

POSSO DIRE CHE $\mu \in [29.2, 31.2]$ CON UNA CONFIDENZA DEL 90 %

$\boxed{\text{INTERVALLO DI CONFIDENZA}}$

Anziché dare una stima puntuale $\mu \approx 30.2$ fornisco un intervallo di valori.

PER ESERCIZIO

$\alpha = 0.05 \rightarrow [29.0, 31.4]$

$\alpha = 0.01 \rightarrow [28.6, 31.8]$

$\alpha = 0.001 \rightarrow [28.1, 32.3]$

COSA SUCCEDE SE NON CONOSCO σ

SI PUÒ DIMOSTRARE CHE UNO STIMATORE NON DISTORTO (UNBIASED) PER σ^2

è

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \qquad \boxed{\text{VARIANZA CAMPIONARIA}}$$

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2} = \boxed{\text{DEVIAZIONE STANDARD CAMPIONARIA}}$$

OSSERVAZIONE: Se n è molto grande $n \sim n-1$ ($n \geq 100$) possiamo dire che approssimativamente la distribuzione è ancora normale $\frac{n}{n+1} \simeq 1$

Se n piccolo ($n < 100$) $\boxed{\frac{\bar{X}-\mu}{\cancel{\sigma}\,S/\sqrt{n}}}$ NON SEGUE PIÙ UNA DISTRIBUZIONE NORMALE

$$\sim t_{n-1}$$

$\boxed{\text{T DI STUDENT A } n \text{ gradi di libertà}}$

↓ DIPENDE DA n

È SEMPRE UNA DISTRIBUZIONE A CAMPANA MA HA "LE CODE" PIÙ GROSSE

SI TROVANO I QUANTILI IN UNA TABELLA DIVERSA.

Posso rifare lo stesso lavoro di prima con due modifiche:

1. NON HO σ QUINDI LO STIMO CON S
2. SE $n < 100$ GUARDO LA TABELLA DI t_{n-1} ANZICHÉ DI Z PER CONOSCERE $q_{1-\frac{\alpha}{2}}$.

$$\mu \in \left[\bar{x} - \frac{q^{(t_{n-1})}_{1-\frac{\alpha}{2}}}{\sqrt{n}} s \, , \, \bar{x} + \frac{q^{(t_{n-1})}_{1-\frac{\alpha}{2}} \cdot s}{\sqrt{n}} \right]$$ CON CONFIDENZA $1-\alpha$

NELL' ESEMPIO

$\alpha = 0.01$

$$S = \sqrt{\frac{1}{14} \sum_{i=1}^{n} (x_i - \bar{x})^2} = \sqrt{7.03}$$

$$q_{1-\frac{\alpha}{2}} = 2.977$$

$$\frac{(2.977)\sqrt{7.03}}{\sqrt{15}} \simeq 2.0$$

$$[30.2 - 2.0, 30.2 + 2.0]$$

$$\boxed{\mu \in [28.2, 32.2] \text{ CON 99 \% DI CONFIDENZA}}$$

ESEMPIO DI ESERCIZIO

X CAMPIONE ESTRATTO DA UNA POPOLAZIONE CHE SEGUE UNA DISTRIBUZIONE NORMALE.

CALCOLARE L'INTERVALLO DI CONFIDENZA CON $\alpha = 0.001$

$$X = \left(\underset{x_1}{2.3}, \underset{x_2}{5.4}, \underset{x_3}{3.1}, \underset{x_4}{2.9}, \underset{x_5}{1.8} \right) \qquad n = 5$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{1}{5} (2.3 + 5.4 + 3.1 + 2.9 + 1.8) = 3.1$$

$1 - \frac{\alpha}{2}$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{4} \left[(2.3-3.1)^2 + (5.4-3.1)^2 + \dots + (1.8-3.1)^2 \right] =$$

$$= \frac{1}{4} \left[0.64 + 5.29 + 0 + 0.04 + 1.69 \right] =$$

$$\simeq 1.92$$

$$\frac{q_{1-\frac{\alpha}{2}}^{(t_{n-1})} S}{\sqrt{n}} = \frac{q_{1-\frac{\alpha}{2}}^{t_4} \cdot S}{\sqrt{n}} = \frac{8.610 \cdot \sqrt{1.92}}{\sqrt{5}} = 5.3$$

$$3.1 - 5.3 = -2.2$$
$$3.1 + 5.3 = 8.4$$

$\boxed{\mu \in [-2.2, 8.4]}$ COL 99 % DI CONFIDENZA

PER ESERCIZIO COL 90 % DI CONFIDENZA

$\alpha = 0,1 \rightarrow \quad q_{1-\frac{\alpha}{2}}^{(t_4)} = 2.132 \leftarrow$ DALLA TABELLA

$$\frac{2.132 \cdot \sqrt{1.92}}{\sqrt{5}} = 1.4 \quad , \quad \begin{aligned} \bar{x} - 1.4 &= 3.1 - 1.4 = 1.7 \\ \bar{x} + 1.4 &= 3.1 + 1.4 = 4.5 \end{aligned}$$

STIMO CHE $\boxed{\mu \in [1.7, 4.5]}$ COL 90 % DI CONFIDENZA

Questi intervalli non danno molta informazione perché i valori sono troppo pochi ($n = 5$ è troppo piccolo).