STATISTICA DESCRITTIVA

Ω : popolazione

Y : carattere

$\underline{y} = (y_1, y_2, \dots, y_N)$

rilevazione dati

N : taglia di Ω

STATISTICA INFERENZIALE PARAMETRICA

Ω : spazio campione

X : genitrice

$\boxed{\underline{X} = (X_1, X_2, \dots, X_n)}$

campione casuale semplice

$\Leftrightarrow$ le v.a. $X_1, X_2, \dots, X_n$ sono indipendenti e somiglianti a X e si chiamano "osservazioni"

$\underline{x} = (x_1, x_2, \dots, x_n)$

realizzazione del campione

n : taglia del campione

una funzione di $\underline{X}$,

$$T = t(\underline{X})$$

è chiamata "statistica" e se essa viene usata a scopo di stima è chiamata "stimatore"

La disciplina è denominata "STATISTICA" in quanto traduzione del vocabolo inglese STATISTICS, ovvero è una disciplina che "tratta" le funzioni di un campione casuale.

<u>ESEMPIO</u>

Per illustrare la differenza tra punto campione, campione e realizzazione

campione, campione e realizzazione
del campione.

$$X \sim B(1, p), \quad n = 10$$

$$\begin{array}{c|cccccccccc}
\underline{\omega} & (T, & C, & C, & T, & C, & C, & T, & T, & T, & T) \\
\underline{X}(\underline{\omega}) & (1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1) \\
 & x_1, & x_2, & x_3, & x_4, & x_5, & x_6 & x_7 & x_8, & x_9, & x_{10} \\
 & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow
\end{array}$$

$$X \sim X_1 \sim X_2 \sim X_3 \sim X_4 \sim X_5 \sim X_6 \sim X_7 \sim X_8 \sim X_9 \sim X_{10}$$

indipendenti

$\square$

ESEMPI DI STATISTICHE

$$T = X_1 + X_1 + \cdots + X_n$$ la somma delle osservazioni

$$X = \frac{1}{n} = \frac{1}{n} \sum_{i=1} X_i$$ la media campionaria

$$k \in \mathbb{N}, \quad \bar{X}^{(k)} = \frac{1}{n} \sum_{i=1}^{n} X_i^{k}$$ il momento empirico di ordine k

ovviamente

$$\bar{X}^{(1)} = \bar{X}$$

e

$\bar{X}^{(1)}$ è l'alter ego "campionario" di

$$\mu_X = \mu_1' = \mathbb{E}(X)$$ media

$k \in \mathbb{N}$, $\bar{X}^{(k)}$ è l'alter ego "campionario" di

$$\mu_k' = \mathbb{E}(X^k)$$ momento teorico di ordine k

Risultato importante

Si ha che :

$$K \in \mathbb{N}, \qquad \bar{X}^{(k)} \xrightarrow[n \to \infty]{?} \mu'_k = \mathbb{E}(X^k).$$

In altri termini, al crescere della taglia n il momento campionario di ordine tende (in un senso che sarà specificato nelle prossime lezioni) a un valore deterministico che coincide con il momento teorico di ordine k (se esistente e finito).

$\square$

Questo risultato è alla base del "metodo dei momenti" che è utile per la costruzione degli stimatori.

METODO DEI MOMENTI

Si preferisce di far precedere l'esposizione del metodo da alcuni esempi.

ESEMPIO

a) $X \sim B(1, p)$, $p \in (0,1)$ incognito.

Fatto teorico

b)
$$\mu_1' = \mathbb{E}(X) = p. \qquad (*)$$

Fatto teorico (sperimentale)

c) $(X_1, X_2, \dots, X_n)$ e $X_1 \sim X$.
c.c.s.

<u>Metodo</u>

d) sostituire μ_1' con $\bar{X}^{(1)} = \bar{X}$.

L'equazione $(*)$ diventa

$$\bar{X} \simeq p$$

da cui scaturisce la seguente definizione

$$\hat{p}_{MM} := \underset{\text{v.a.}}{\bar{X}} = \underbrace{\frac{X_1 + X_2 + \dots + X_n}{n}}_{\text{media campionaria}} \qquad (\hat{p}_{MM} \neq p)$$

Conclusione: lo stimatore del parametro

di una v.a. di Bernoulli individuato con il metodo dei momenti è la media campionaria.

Quando si otterrà la realizzazione del campione $\underline{x} = (x_1, x_2, \dots, x_n)$ allora

$$\hat{p}_{MM} = \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$ media aritmetica

fornisce la stima di p.

$\square$

ESEMPIO 2

a) $X \sim G(p)$, $p \in (0,1)$ e p incognito.

Fatto teorico

b) $$\mu_1' = \mathbb{E}(X) = \frac{1}{r}. \qquad (**)$$

<u>Fatto teorico (sperimentale)</u>

c) $(X_1, X_2, ..., X_n)$ c.c.s. e $X_1 \sim X$.

<u>Metodo</u>

d) sostituire μ_1' con $\overline{X}^{(1)} = \overline{X}$.

L'equazione $(**)$ diventa

$$\overline{X} \approx \frac{1}{r}$$

da cui scaturisce la seguente definizione

$$\hat{p}_{MM} := \frac{1}{\overline{X}} = (\overline{X})^{-1} \left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)^{-1} \qquad (\hat{p}_{MM} \neq p)$$

$\downarrow$ v.a. — reciproco della media campionaria

Conclusione: lo stimatore del parametro di una v.a. geometrica individuato con il metodo dei momenti è il reciproco della media campionaria.

Quando si otterrà $\underline{x} = (x_1, x_2, \ldots, x_n)$ allora

$$\hat{p}_{MM} = \frac{1}{\overline{x}} = \left(\frac{1}{n} \sum_{i=1}^{n} x_i\right)^{-1} \quad \text{media aritmetica}$$

fornisce la stima di p.

$\square$

ESEMPIO 3

a) $X \sim \Pi(\lambda)$, $\lambda > 0$ incognito,

b) $\mu'_1 = \lambda$

c) $\underline{X} = (X_1, X_2, \dots, X_n)$

d) sostituire μ'_1 con $\overline{X}$:

$$\overline{X} \approx \lambda \Rightarrow$$

stimatore $\hat{\lambda}_{MM} := \overline{X}$, $\quad (\hat{\lambda}_{MM} \neq \lambda)$

stima $\hat{\lambda} = \overline{x}$

stima $\hat{\lambda}_{MM} = \bar{x}$.

Conclusione: lo stimatore del parametro di una v.a. di Poisson individuato con il metodo dei momenti è la media campionaria.

$\square$

ESEMPIO 4

a) $X \sim N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$ incogniti

b) $\mathbb{E}(X) = \mu$

$$\mathbb{D}^2(X) = \sigma^2 \iff \mathbb{E}(X^2) - \mathbb{E}^2(X) = \sigma^2$$

$$\Leftrightarrow \quad \mathbb{E}(X^2) = \sigma^2 + \mu^2$$

$$\begin{cases} \mathbb{E}(X) = \mu \\ \mathbb{E}(X^2) = \sigma^2 + \mu^2 \end{cases} \Leftrightarrow \begin{cases} \mu = \mathbb{E}(X) \\ \sigma^2 = \mathbb{E}(X^2) - \mu^2 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} \mu = \mathbb{E}(X) \\ \sigma^2 = \mathbb{E}(X^2) - \mathbb{E}^2(X) \end{cases}$$

c) $\underline{X} = \underset{\text{c.c.s}}{(X_1, ---, X_n)}, \quad X_1 \sim X$

d) sostituire μ_1' con $\bar{X}$ e $\mu_2' = \mathbb{E}(X^2)$ con $\bar{X}^{(2)}$:

$$\Big(\mu \simeq \bar{X}$$

$$\begin{cases} \cdots \\ \sigma^2 \approx \overline{X}^{(2)} - (\overline{X})^2 \end{cases} \Longrightarrow$$

stimatore per μ $\qquad \hat{\mu}_{MM} := \overline{X}$

stimatore per σ^2 $\qquad \hat{\sigma}^2_{MM} := \overline{X}^{(2)} - (\overline{X})^2$

Conclusione: lo stimatore del parametro bidimensionale di una v.a. di Gauss individuato con il metodo dei momenti è la coppia costituita dalle statistiche media campionaria e differenza tra il momento campionario del secondo ordine e il quadrato della media campiomarzia

Se si desidera ottenere uno stimatore per σ basta osservare che

$$\sigma = \sqrt{\sigma^2} = f(\sigma^2)$$

la radice quadrata è funzione continua in tutto il suo dominio

e applicare la stesse funzione allo stimatore di σ^2 individuato dal metodo dei momenti:

$$\hat{\sigma}_{MM} = f(\hat{\sigma}^2_{MM}) = \sqrt{\hat{\sigma}^2_{MM}}$$

$$= \sqrt{\bar{X}^{(2)} - (\bar{X})^2} = g(\underline{X})$$

è una funzione del campione

è uno stimatore "buono" per la deviazione standard: aumentando la taglia del campione "aumenta" la precisione della stima. $\square$