

# Contents

<b>1</b>	<b>Introduzione</b>	<b>2</b>
----------	---------------------	----------

# 1 Introduzione

$$X_1 = 1.7$$

$$X_2 = 1.82$$

$$X_3 = 1.73$$

$$X_4 = 1.7$$

$$X_5 = 1.8$$

$\hat{\theta}$ ? Altezza della popolazione

**Possibile soluzione** :

$$\hat{\theta}_a = \frac{1}{n} \sum_4^5 x_i = \frac{1.7 + 1.82 + 1.73 + 1.7 + 1.8}{5} = \frac{8.75}{5} = 1.75$$

$$\hat{\theta}_b = \frac{\min(x_i) + \max(x_i)}{2} = \frac{3.52}{2} = 1.76$$

$$\hat{\theta}_c = \frac{1}{3} \sum_2^4 x_i = \frac{1}{3}(1.8 + 1.73 + 1.7) = \frac{5.23}{3} = 1.743$$

Scartiamo il più *piccolo* e il *massimo*, calcolando poi la **media** dei rimanenti

**Stima parametrica** (Point) Parametric Estimation

Ipotesi: - Esiste un parametro  $\theta$  incognito  $n$  dati a disposizione  $\{X_1, X_2, X_n\}$

**Legge di probabilità** che descrive il fenomeno che ha generato i dati

**Formula generica**: Bayes

$$P(\theta/X_1 \dots X_n) = \frac{P(X_1 \dots X_n/\theta)P(\theta)}{P(X_1 \dots X_n)}$$

## Verosomiglianza (likelihood)

**MLE** Maximum Likelihood Estimation (Stima a Massima Verosomiglianza)

$$\hat{\theta} = \operatorname{argmax} L(\theta) = \operatorname{argmax} [f(X_1 \dots X_n / \theta)]$$

**Esempio** (Legge -> Distribuzione di Poisson)

$$\begin{aligned} f(X_1, X_2 \dots X_n / \theta) &= f(X_1 / \theta) \cdot f(X_2 / \theta) \dots f(X_n / \theta) \\ &= \frac{1}{\theta} e^{-\frac{x_1}{\theta}} \cdot \frac{1}{\theta} e^{-\frac{x_2}{\theta}} \cdot \dots \cdot \frac{1}{\theta} e^{-\frac{x_n}{\theta}} \\ &= \frac{1}{\theta^n} e^{-\frac{1}{\theta} \sum_i x_i} \end{aligned}$$

**Esempio** (MLE Ipotesi di Bernoulli)

$$X_i = \begin{cases} 0 \\ 1 \end{cases}$$

$$P\{X_i = 1\} = 1 - P\{X_i = 0\}$$

$$P\{X_i = x\} = P^x (1 - P)^{1-x} \quad x \in \{0, 1\}$$

Dove **X** è una *variabile aleatoria* e **x** una *variabile sperimentale*

$$f(x_1 \dots x_n / P) = P^{x_1} (1 - P)^{1-x_1} \cdot P^{x_2} (1 - P)^{1-x_2} \dots P^{x_n} (1 - P)^{1-x_n} =$$

$P^{\sum_i x_i} (1 - P)^{n - \sum_i x_i} \longrightarrow$  Bisogna trovare il **massimo** della funzione

$$\begin{aligned}
\log(f(x_1 \dots x_n/P)) &= \sum_1^n x_i \log P - (n - \sum_i^n x_i) \log(1 - P) \\
&= \frac{d}{dP} [\log(f)] = 0 = \frac{1}{\hat{P}} \sum_i^n x_i - \frac{n - \sum_i^n x_i}{(1 - \hat{P})} \\
&= (1 - \hat{P}) \sum_i^n x_i = \hat{P} (n - \sum_i^n x_i) \\
&= \hat{P} = \frac{\sum_i^n x_i}{n} \quad \text{MLE}
\end{aligned}$$

**Esercizio 1** Probabilità che Oneto dia 30L (Lode)

$$n = 120$$

$$\sum_i^{120} x_i = 18$$

$$\hat{P} = \frac{18}{120} = 0.15 \rightarrow 15\%$$

**Esercizio 2** N studenti da 30 e lode

$$n_1 = 18 \leftarrow \text{Oneto}$$

$$n_2 = 20 \leftarrow \text{Anguilla}$$

$$n_{1,2} = 10 \leftarrow 30L \text{ sia con Oneto che con Anguilla}$$

$$N = ? \quad \text{Studenti da } \mathbf{30 \text{ e Lode}}$$

$$\hat{P}_1 \approx \frac{n_1}{n_2}$$

$$\hat{P}_1 \approx \frac{n_1}{N}$$

$$\frac{n_{1,2}}{n_2} = \frac{n_1}{N}$$

$$\Rightarrow N = \frac{n_1 \cdot n_2}{n_{1,2}} \rightarrow \frac{18 \cdot 20}{10} = 36$$

## MLE POISSON

$$\begin{aligned}f(x_1, x_2 \dots x_n / \lambda) &= \frac{e^{-\lambda} \lambda^{x_1}}{x_1!} \cdot \frac{e^{-\lambda} \lambda^{x_2}}{x_2!} \dots \frac{e^{-\lambda} \lambda^{x_n}}{x_n!} \\&= \frac{e^{-n\lambda} \lambda^{\sum_i x_i}}{x_1! x_2! \dots x_n!}\end{aligned}$$

**Formula generica:**  $\lambda = \frac{\sum_i x_i}{n}$  MLE

**Esercizio 3** Stima del numero di incidenti medio in auto n = 10

$$x_1 = \{4, 0, 6, 5, 2, 1, 2, 0, 4, 3\}$$

$$\hat{\lambda} = \frac{\sum_i x_i}{n} = \frac{27}{10} = 2.7$$

$$P\{x \leq 2\} = e^{-2.7} \left( \frac{2.7^0}{0!} + \frac{2.7^1}{1!} + \frac{2.7^2}{2!} \right) \approx .4936 \rightarrow 49.36\%$$

Probabilità che non ci siano più di **2 incidenti**

## MLE UNIFORME

$$f(x_1, x_2 \dots x_n / \theta) = \begin{cases} \frac{1}{\theta^n} & 0 < x_i < \theta \\ 0 & \text{altrimenti} \end{cases}$$

$$\hat{\theta} = \max\{x_i\}$$

$$\frac{\hat{\theta}}{2} = \frac{\max\{x_i\}}{2}$$

## MLE GAUSSIANA

$$f(x_1, x_2 \dots x_n / \mu, \sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x_i - \mu)^2}{2\sigma^2}}$$

$$\left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \frac{1}{\sigma^n} e^{\frac{-\sum_i (x_i - \mu)^2}{2\sigma^2}}$$

$$\log[f] = -\frac{n}{2} \log 2\pi - n \log \sigma - \frac{\sum_i (x_i - \mu)^2}{2\sigma^2}$$

$$\frac{d \log f}{d \mu} = 0 = \frac{\sum_i (x_i - \mu)^2}{\sigma^2} \longrightarrow \hat{\mu} = \frac{\sum_i x_i}{n}$$

$$\frac{d \log f}{d \sigma} = 0 = -\frac{n}{\sigma} + \frac{\sum_i (x_i - \mu)^2}{4\sigma^4} \longrightarrow \sigma = \sqrt{\frac{\sum_i (x_i - \mu)^2}{n}}$$

**Esercizio** primo

$$x_1 = 1.7$$

$$x_2 = 1.82$$

$$x_3 = 1.73$$

$$x_4 = 1.7$$

$$x_5 = 1.8$$

$$\hat{\mu} = \frac{\sum_i x_i}{n} = \frac{1.7 + 1.82 + 1.73 + 1.7 + 1.8}{5} = 1.75$$

$$\hat{\sigma} = \sqrt{\frac{0.05^2 + 0.07^2 + 0.02^2 + 0.05^2 + 0.05^2}{5}} \approx 0.051$$

**Intervalli di confidenza** normali TODO

**Intervalli di confidenza** gaussiani  $\sigma^2$  Nota

$$x_1, x_2, \dots, x_n$$

$$\hat{\mu} \longleftarrow \mu$$

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

$$P(-1.96 < \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} < +1.96) = 0.95$$

$$\longrightarrow P\left(-1.96 \frac{\sigma}{\sqrt{n}} < \bar{x} - \mu < 1.96 \frac{\sigma}{\sqrt{n}}\right)$$

$$P\left(\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}\right)$$

**Esempio:** Sistema di comunicazione  $\sigma^2 = 4$   $n = 9$

$$x_1 = \{5.85, 12, 15, 7, 9, 7.5, 6, 5, 10.5\}$$

$$\hat{\mu} = \frac{1}{n} \sum_i^n x_i = \frac{1}{9} \sum_i^n x_i = \frac{81}{9} = 9$$

$$P\left(9 - 1.96 \frac{\sigma}{\sqrt{m}} < \mu < 9 + 1.96 \frac{\sigma}{\sqrt{m}}\right) = 0.95$$

$$p\left(9 - 1.96 \frac{2}{3} < \mu < 9 + 1.96 \frac{2}{3}\right) = 0.95$$

$\longrightarrow [7.693, 10.31] \rightarrow \mu$  si trova tra 7.693 e 10.31

**In generale** Prob =  $1 - \alpha$

$$\left(\bar{x} - z_a \frac{\sigma}{\sqrt{n}}, \bar{x} + z_a \frac{\sigma}{\sqrt{n}}\right) \rightarrow \text{Si rileva dalle tavole}$$