

Richiamo: dato un campione $\{x_1, \dots, x_n\}$.

con X_i distribuito iid con distribuzione $p_x(x|\theta)$

$f_x(x|\theta)$ con parametro/i sconosciuto/i θ

Def: Uno stimatore (parametrico) $\hat{\theta}(x_1, \dots, x_n)$

è una funzione del campione che approssima il valore del parametro θ .

Il valore di $\hat{\theta}$ su un campione dato è la stima

Def: Lo stimatore di massima verosimiglianza è

$$\hat{\theta}(x_1, \dots, x_n) = \underset{\theta}{\operatorname{argmax}} \mathcal{L}(x_1, \dots, x_n | \theta)$$

dove

$$\mathcal{L}(x_1, \dots, x_n | \theta) = \begin{cases} p_x(x_1, \dots, x_n | \theta) = \prod_{i=1}^n p_x(x_i | \theta) \\ f_x(x_1, \dots, x_n | \theta) = \prod_{i=1}^n f_x(x_i | \theta) \end{cases}$$

↪ verosimiglianza

Metodo: 0) scrivere

$$\mathcal{L}(x_1, \dots, x_n | \theta) = \begin{cases} \prod_{i=1}^n p_x(x_i | \theta) \\ \prod_{i=1}^n f_x(x_i | \theta) \end{cases}$$

1) rimuovere da argmax tutto quello che non dipende da θ .

2) trovare θ che massimizza

$$\left(\frac{d}{d\theta} (\dots) = 0 \text{ se } L \text{ è differenziabile} \right).$$

Nota: a volte è più facile massimizzare l'esponente (vedi esempio Poisson) in altre parole

$$\arg \max_{\theta} L(x_1 \dots x_n | \theta) = \arg \max_{\theta} \log L(x_1 \dots x_n | \theta)$$

Esempio: $X_i \sim \text{Pois}(\theta)$ $\{x_1 \dots x_n\}$ $\hat{\theta}(x_1 \dots x_n) = \frac{1}{n} \sum_{j=1}^n x_j$

Esempio: $X_i \sim \text{Ber}(\theta)$ $\{x_1 \dots x_n\}$

esercizio

$$\Rightarrow \hat{\theta}(x_1 \dots x_n) = \frac{1}{n} \sum_{j=1}^n x_j$$

miglior scelta
del parametro
 θ

Esempio: $X_i \sim N(\mu, \sigma^2)$ $\theta = \mu$
sconosciute
conosciute

troviamo $\hat{\mu}(x_1 \dots x_n)$

$$\begin{aligned} \bullet) L(x_1 \dots x_n | \mu) &= f_X(x_1 \dots x_n | \mu) = \prod_{j=1}^n f_X(x_j | \mu) \\ &= \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x_j - \mu)^2}{2\sigma^2}} \end{aligned}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x_1 - \mu)^2}{2\sigma^2}} \cdot \dots \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_n - \mu)^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \cdot \exp \left(- \sum_{j=1}^n \frac{(x_j - \mu)^2}{2\sigma^2} \right)$$

só questo dipende da μ

$$1) \operatorname{argmax}_{\mu} p(x_1, \dots, x_n | \mu) = \operatorname{argmax}_{\mu} \exp \left(- \sum_{j=1}^n \frac{(x_j - \mu)^2}{2\sigma^2} \right)$$

$$= \operatorname{argmax} \cancel{\log \exp} \left(- \sum_{j=1}^n \frac{(x_j - \mu)^2}{2\sigma^2} \right)$$

$$2) \frac{d}{d\mu} \left(- \sum_{j=1}^n \frac{(x_j - \mu)^2}{2\sigma^2} \right) = + \frac{1}{\sigma^2} \sum_{j=1}^n \cancel{2} (x_j - \mu) \cancel{2} = 0$$

$$\Rightarrow \sum_{j=1}^n (x_j - \mu) = 0 \Rightarrow \sum_{j=1}^n x_j - n\mu = 0$$

$$\Rightarrow \hat{\mu} = \frac{1}{n} \sum_{j=1}^n x_j = \bar{X}_n$$

Exemplo: Sia $X_i \stackrel{iid}{\sim} N(\mu, \sigma^2)$ $\theta = (\mu, \sigma)$ sconosciute

$$0) p(x_1, \dots, x_n | \mu, \sigma) = \prod_{j=1}^n f_X(x_j | \mu, \sigma)$$

$$= \frac{1}{\underbrace{\sqrt{2\pi} \sigma^n}_{\text{dipende da } \mu, \sigma}} \cdot \exp \left(- \frac{\sum_{j=1}^n (x_j - \mu)^2}{2\sigma^2} \right)$$

$$1) \operatorname{argmax}_{\mu, \sigma} p(x_1, \dots, x_n | \mu, \sigma) = \operatorname{argmax}_{\mu, \sigma} \frac{1}{\sigma^n} \exp \left(- \frac{\sum_{j=1}^n (x_j - \mu)^2}{2\sigma^2} \right)$$

$$\log a \cdot b = \log a + \log b \Rightarrow \operatorname{argmax}_{\mu, \sigma} \log \frac{1}{\sigma^n} + \log \exp \left(- \frac{\sum_{j=1}^n (x_j - \mu)^2}{2\sigma^2} \right)$$

$$= \arg \max_{\mu, \sigma} \underbrace{-n \log \sigma - \frac{\sum_{j=1}^n (x_j - \mu)^2}{2\sigma^2}}_{e(x_1, \dots, x_n | \mu, \sigma)}$$

$$2) \begin{pmatrix} \frac{d}{d\mu} \\ \frac{d}{d\sigma} \end{pmatrix} e(x_1, \dots, x_n | \mu, \sigma) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} \frac{1}{\sigma^2} \sum_{j=1}^n (x_j - \mu) \\ -n \cdot \frac{1}{\sigma} - \sum_{j=1}^n (x_j - \mu)^2 \underbrace{\frac{d}{d\sigma} \frac{1}{2\sigma^2}}_{-\frac{1}{\sigma^3}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{cases} \hat{\mu} = \frac{1}{n} \sum_{j=1}^n x_j \\ \frac{n}{\hat{\sigma}^2} = \sum_{j=1}^n (x_j - \hat{\mu})^2 \cdot \frac{1}{\hat{\sigma}^4} \end{cases} \Leftrightarrow \begin{cases} \hat{\mu} = \frac{1}{n} \sum_{j=1}^n x_j = \bar{x} \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^n (x_j - \hat{\mu})^2 \end{cases}$$

⚠ questo valore di $\hat{\sigma}^2$ è diverso dalla varianza campionaria

$$s^2 = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2 \quad \hat{\mu} = \bar{x}$$

Domanda: Come possiamo valutare la qualità di uno stimatore?

Ci sono due criteri possibili:

1) Esempio: Assumiamo $X_1, \dots, X_n \sim N(\mu^*, \sigma^2)$

$$\begin{aligned}\mathbb{E}_{\mu^*}(\hat{\mu}(X_1, \dots, X_n)) &= \mathbb{E}_{\mu^*}\left(\frac{1}{n} \sum X_j\right) = \frac{1}{n} \sum \mathbb{E}_{\mu^*}(X_j) = \\ &= \frac{1}{n} \cdot n \cdot \mu^* = \mu^*\end{aligned}$$

Def: Sia $\{x_1, \dots, x_n\}$ un campione con $X_i \stackrel{iid}{\sim} f_X(x|\theta^*)$
uno stimatore $\hat{\theta}$ di θ^* è corretto se

$$\mathbb{E}_{\theta^*}(\hat{\theta}(X_1, \dots, X_n)) = \theta^* \quad \text{per ogni } \theta^*$$

2) Esempio: $X_j \sim N(\mu^*, \sigma^2)$ $\hat{\mu} = \frac{1}{n} \sum_{j=1}^n X_j$

Se avessimo $n = \infty$

Richiamo (LLN) $\lim_{n \rightarrow \infty} \mathbb{P}\left(\left| \frac{1}{n} \sum_{j=1}^n X_j - \underbrace{\mathbb{E}(X_j)}_{\mu^*} \right| > \varepsilon\right) = 0$

$$\Rightarrow \lim_{n \rightarrow \infty} \mathbb{P}_{\mu^*}\left(\left| \hat{\theta}(x_1, \dots, x_n) - \theta^* \right| > \varepsilon\right) = 0$$

Def: Sia $\{x_1, \dots, x_n\}$ un campione con $X_i \sim f_X(x|\theta)$
allora lo stimatore $\hat{\theta}(x_1, \dots, x_n)$ è consistente

se per ogni $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbb{P}_{\theta^*}\left(\left| \hat{\theta}(x_1, \dots, x_n) - \theta^* \right| > \varepsilon\right) = 0$$

Media e Varianza campionarie

A volte non ci interessa (non si può!) stimare i parametri della distribuzione di una popolazione, ma solo le sue medie/varianza (senza fare assunzioni sulla distribuzione stessa).

Esempio: Temperature media in un anno

$\Rightarrow$ formuliamo stimatori per queste quantità.

Richiamo: per un campione $\{x_1, \dots, x_n\}$ le

- media campionaria

$$\bar{X}(x_1, \dots, x_n) = \frac{1}{n} \sum_{j=1}^n x_j \longrightarrow \text{stimatore di } E(X_i)$$

- varianza campionaria

$$S^2(x_1, \dots, x_n) = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{X})^2 \longrightarrow \text{stimatore di } \text{Var}(X_i)$$

Teorema: $\bar{X}, S^2$ sono stimatori corretti e consistenti di $E(X_i), \text{Var}(X_i)$ per qualsiasi distribuzione di X_i .

quindi
$$\left\{ \begin{aligned} \mathbb{E}(\bar{X}(x_1, \dots, x_n)) &= \mathbb{E}(X_i) \\ \lim_{n \rightarrow \infty} P(|\bar{X}(x_1, \dots, x_n) - \mathbb{E}(X_i)| > \varepsilon) &= 0 \end{aligned} \right.$$

e rispettivamente per $\underline{S^2}$, $\underline{\text{Var}(X_i)}$

Infatti:

Per $\bar{X}$:

- $\mathbb{E}(\bar{X}(x_1, \dots, x_n)) = \mathbb{E}\left(\frac{1}{n} \sum X_j\right) = \frac{1}{n} \cdot n \mathbb{E}(X_j) = \mathbb{E}(X_j)$

- LLN : $\lim_{n \rightarrow \infty} P\left(|\underbrace{\frac{1}{n} \sum X_j}_{\bar{X}} - \mathbb{E}(X_j)| > \varepsilon\right) = 0$

Per S^2 : Vogliamo che $\mathbb{E}(\cancel{(n-1)} S^2) = \cancel{(n-1)} \text{Var}(X_j)$

$$\begin{aligned} \mathbb{E}((n-1) S^2) &= \mathbb{E}\left(\sum_{j=1}^n (X_j - \bar{X})^2\right) = \\ &= \mathbb{E}\left(\sum (X_j^2 - 2X_j \bar{X} + \bar{X}^2)\right) = \\ &= \mathbb{E}\left(\sum X_j^2\right) - 2n \bar{X} \cdot \bar{X} - n \bar{X}^2 \\ &= \mathbb{E}\left(\sum X_j^2\right) - n \bar{X}^2 = \\ &= n \left(\mathbb{E}(X_j^2) - \bar{X}^2\right) \end{aligned}$$

$$\begin{aligned}
&= n \left(\underbrace{\mathbb{E}(X_j^2) - \mathbb{E}(X_j)^2}_{\text{Var}(X_j)} - \underbrace{(\bar{X}^2 - \mathbb{E}(X_j)^2)}_{(\dots)} \right) \\
&= n \text{Var}(X_j) - \text{Var}(\bar{X}) \\
&= (n-1) \text{Var}(X_j)
\end{aligned}$$

Quindi ~~$(n-1) \mathbb{E}(S^2) = (n-1) \text{Var}(X_j)$~~ $\square$

$\Rightarrow$ lo stimatore di massima verosimiglianza

$$\mathbb{E}(\hat{\sigma}^2) = \mathbb{E}\left(\frac{n-1}{n} \frac{1}{n-1} \sum (X_j - \bar{X})^2\right) =$$

$$= \frac{n-1}{n} \mathbb{E}\left(\frac{1}{n-1} \sum (X_j - \bar{X})^2\right)$$

$$= \frac{n-1}{n} \text{Var}(X_j) = \frac{n-1}{n} \sigma^2$$