Fitting 2 - ML Parti 1 e 2

STIMA DI PARAMETRI DI PDF

Monday, November 27, 2023 8:37 AM

Stima di uno o più parametri incogniti $\underline{\theta}$ di una pdf $f(X;\underline{\theta})$ DENSITÀ di probab.

$\longmapsto \underline{\theta}^T = (\theta_1, \dots, \theta_p)$

n misure di X: $x_1, \dots, x_n$

Realizzazioni di n variabili casuali $X_1 \dots X_n$

i.i.d. $\longrightarrow$ identicamente distribuite

$\longmapsto$ indipendenti

$X_i \sim f(X_i;\underline{\theta}) \qquad i = 1, \dots n$

Monday, November 27, 2023 8:42 AM

moltiplico pdf $\Leftrightarrow$ moltipl la prob.

i.i.d. $\Rightarrow$ Probabilità congiunta di $X_1 \dots X_n$:

$$L(X_1 \dots X_n ; \underline{\theta}) = \prod_{i=1}^{n} f(X_i ; \underline{\theta})$$

$\forall \underline{\theta} \in \Theta$ spazio dei parametri

NB L, come probabilità congiunta (a $\underline{\theta}$ fissato) di $X_1 \dots X_n$ è normalizzata:

$$\int_{-\infty}^{+\infty} dx_1 \dots \int_{-\infty}^{+\infty} dx_n \; L(x_1 \dots x_n ; \underline{\theta}) = 1$$

fissato (arbitrario)

NB L non è viceversa normalizzata rispetto a $\underline{\theta}$

$\hookrightarrow$ In generale: $\displaystyle\int_{-\infty}^{+\infty} d\theta_1 \dots \int_{-\infty}^{+\infty} d\theta_p \; L(\underbrace{x_1 \dots x_n}_{\text{fissati}} ; \underline{\theta}) \neq 1$

$\Downarrow$

L non è interpretabile in alcun modo come una "pdf" per $\underline{\theta}$

DEF La funzione $L(x_1 \dots x_n ; \underline{\theta})$ pensata come f.ne di $\underline{\theta}$ a $x_1 \dots x_n$ fissati è detta Likelihood (Verosimiglianza)

NB $f_1 \cdot f_2 \cdot f_3 = L$

$\ln L = \ln(f_1 \cdot f_2 \cdot f_3) = \ln(f_1) + \ln(f_2) + \ln(f_3)$ passo da $\prod_i$ a $\sum_i \ln$

Monday, November 27, 2023 8:43 AM

Maximum likelihood ($p=1$) un parametro solo da stimare

Date n misure $x_1 .. x_n$, scelgo come migliore stima di θ il valore $\hat{\theta}$ per il quale:

$$L(\underbrace{x_1 .. x_n}_{\text{fissati}}; \hat{\theta}) \geq L(\underbrace{x_1 .. x_n}_{\text{fissati}}; \theta) \quad \forall \theta \in \Theta$$

Se f è sufficientemente "smooth" (derivabile almeno 2 volte in θ) e $\hat{\theta}$ è interno al dominio Θ, la stima di $\hat{\theta}$ si ottiene risolvendo:

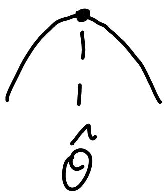
$$\frac{\partial L(x_1 \ldots x_n; \theta)}{\partial \theta}\bigg|_{\hat{\theta}} = 0$$

$\therefore$ $\hat{\theta}$ punto stazionario per L (a $x_1 \ldots x_n$ fissati)

sotto la condizione:

$$\frac{\partial^2 L(x_1 \ldots x_n; \theta)}{\partial \theta^2}\bigg|_{\hat{\theta}} < 0$$

$\Rightarrow$ il punto stazionario $\hat{\theta}$ è un massimo



ln monotona crescente

Monday, November 27, 2023 8:44 AM

nel caso i.i.d.

Equivalentemente:

$$\frac{\partial}{\partial\theta} \ln L(x_1 \dots x_n ; \theta)\Big|_{\hat{\theta}} \equiv \frac{\partial}{\partial\theta} \sum_{i=1}^{n} \ln f(x_i ; \theta)\Big|_{\hat{\theta}} = 0$$

$$\frac{\partial^2}{\partial\theta^2} \ln L(x_1 \dots x_n ; \theta)\Big|_{\hat{\theta}} \equiv \frac{\partial^2}{\partial\theta^2} \sum_{i=1}^{n} \ln f(x_i ; \theta)\Big|_{\hat{\theta}} < 0$$

$(p=1)$

Nel caso generale $(p \geq 2)$

$$\frac{\partial}{\partial\theta_i} \ln L(x_1 \dots x_n ; \underline{\theta})\Big|_{\hat{\underline{\theta}}} = 0 \qquad i = 1, \dots p$$

$$\underbrace{\frac{\partial^2}{\partial\theta_i \partial\theta_j} \ln L(x_1 \dots x_n ; \underline{\theta})}_{\text{matrice } p \times p}\Big|_{\hat{\underline{\theta}}} \qquad \text{definita negativa}$$

Esempio 1: pdf esponenziale

$$X_i \sim f(X_i;\theta) = \frac{1}{\theta} e^{-\frac{X_i}{\theta}} \qquad \begin{cases} 0 \leq X_i < +\infty & \text{Supporto delle } X \\ 0 < \theta < +\infty & \Theta \end{cases}$$

$i = 1, \dots n$ iid

*) f è una pdf $\therefore$ $\int_0^{+\infty} f(x;\theta)\,dx = 1$

$$\int_0^{+\infty} \frac{1}{\theta} e^{-\frac{x}{\theta}} dx \underset{y = \frac{x}{\theta}}{=} \int_0^{+\infty} e^{-y} dy = -e^{-y}\Big|_0^{+\infty} = -(0-1) = 1 \quad \text{c.v.d.}$$

*) $E_\theta[X] = \theta$ $\therefore$ $\int_0^{+\infty} x \frac{1}{\theta} e^{-\frac{x}{\theta}} dx = \theta$

$$\int_0^{+\infty} \frac{x}{\theta} e^{-\frac{x}{\theta}} dx \underset{y = \frac{x}{\theta}}{=} \theta \int y\, e^{-y} dy = \theta \Big[\cancel{-y e^{-y}\Big|_0^{+\infty}} + \underbrace{\int_0^{+\infty} e^{-y} dy}_{1} \Big] = \theta$$

$$*)\quad V[X] = \theta^2$$

$$\overset{\shortparallel}{E}[X^2] - E^2[X] = E[X^2] - \theta^2 \qquad \therefore\ E[X^2] = 2\theta^2$$

$$\int_0^{+\infty} \frac{x^2}{\theta} e^{-\frac{x}{\theta}} dx \underset{\substack{\downarrow \\ y=\frac{x}{\theta}}}{=} \theta^2 \int_0^{+\infty} y^2 e^{-y} dy = \theta^2 \left[\cancel{-y^2 e^{-y} \Big|_0^{+\infty}} + 2 \underbrace{\int_0^{+\infty} y e^{-y} dy}_{\shortparallel \atop 1} \right] =$$

$$= 2\theta^2 \quad \text{c.v.d.}$$

Supponiamo di aver misurato $x_1 .. x_n$

Voglio stimare θ

$$L(x_1..x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \frac{1}{\theta^n}\prod_{i=1}^{n} e^{-\frac{x_i}{\theta}} = \frac{1}{\theta^n} e^{-\frac{\sum_{i=1}^{n} x_i}{\theta}}$$

$\hookrightarrow$ La tratto come f.ne di θ a $x_1..x_n$ fissati

$$\ell(x_1..x_n;\theta) \equiv \ln L(x_1..x_n;\theta) = -n\ln\theta - \frac{\sum_{i=1}^{n} x_i}{\theta}$$

$$\frac{d}{d\theta}\ell(\theta) = -\frac{n}{\theta} + \frac{\sum_{i=1}^{n} x_i}{\theta^2} = \frac{n}{\theta^2}\left(\frac{1}{n}\sum_{i=1}^{n} x_i - \theta\right) \overset{!}{=} 0$$

$\hat{\theta}_n(x_1..x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i \equiv \bar{x}_n$ Stima di ML di θ basata su $x_1..x_n$

$\hat{\theta}_n(X_1,..X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i$ Estimatore ML di θ

$\hookrightarrow$ variabile casuale

Monday, November 27, 2023 8:44 AM

$$\frac{d^2}{d\theta^2} \ell(\theta) = \frac{n}{\theta^2} - 2\,\frac{\sum_{i=1}^{n} x_i}{\theta^3} = -\frac{n}{\theta^3}\left(2\,\bar{x}_n - \theta\right) =$$

la calcoliamo per $\theta = \hat{\theta} = \bar{x}_n$

$$= -\frac{n}{\bar{x}_n^3}\left(2\,\bar{x}_n - \bar{x}_n\right) = -\frac{n}{\bar{x}_n^2} < 0$$

↳ il pto stazionario trovato è effettivamente un max

Monday, November 27, 2023 8:44 AM

stima basata su $x_1 \ldots x_n$

$$E_\theta\left[\hat{\theta}_n\right] = \int_0^{+\infty} dx_1 \ldots \int_0^{+\infty} dx_n \; \hat{\theta}_n(x_1 \ldots x_n) \underbrace{L(x_1 \ldots x_n; \theta)}_{\text{pdf congiunta di } X_1 \ldots X_n \text{ a } \theta \text{ fissato}} =$$

↳ Estimatore (variabile casuale)

$$= \int_0^{+\infty} dx_1 \cdots \int_0^{+\infty} dx_n \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{x_i}{\theta}} =$$

$$= \frac{1}{n}\sum_{i=1}^{n} \underbrace{\int_0^{+\infty} \frac{x_i}{\theta} e^{-\frac{x_i}{\theta}} dx_i}_{\theta} \; \underbrace{\prod_{\substack{j=1 \\ j \neq i}}^{n} \underbrace{\int_0^{+\infty} \frac{1}{\theta} e^{-\frac{x_j}{\theta}} dx_j}_{1}}_{1} = \frac{1}{n} n\theta = \theta$$

$\hat{\theta}_n$ è un estimatore unbiased di θ (come ci si aspettava che fosse)

Con un calcolo simile, si verifica che:

$$V_\theta\left[\hat{\theta}_n\right] \equiv E_\theta\left[\hat{\theta}_n^2\right] - E_\theta^2\left[\hat{\theta}_n\right] = \ldots = \frac{\theta^2}{n}$$

Monday, November 27, 2023 8:45 AM

θ = parametro da stimare (location parameter)

σ = parametro noto (shape parameter)

Esempio 2 pdf gaussiana (con σ fissata)

$$X_i \sim \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x_i-\theta)^2}{2\sigma}}$$

$X_1, \dots X_n$ iid

$$L(\theta) = \frac{1}{(2\pi)^{n/2}} \frac{1}{\sigma^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2}$$

Termini che non dipendono da θ

$$\ell(\theta) \equiv \ln L(\theta) = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2 + \dots$$

$$\frac{d}{d\theta}\ell(\theta) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\theta) = \frac{n}{\sigma^2}\left(\frac{1}{n}\sum_{i=1}^{n}x_i - \theta\right) = 0 \Rightarrow \hat{\theta}_n = \bar{x}_n = \frac{1}{n}\sum_{i=1}^{n}x_i$$

stima di θ basata su $x_1 \dots x_n$

$$\frac{d^2}{d\theta^2}\ell(\theta) = -\frac{n}{\sigma^2} < 0$$

$$\hat{\theta}_n(X_1 \dots X_n) = \frac{1}{n}\sum_{i=1}^{n}X_i$$

$$E_\theta\left[\hat{\theta}_n\right] = \theta$$

unbiased

Estimatore (variabile casuale)

Monday, November 27, 2023 8:45 AM

NB X_i $i=1,\dots,n$ **non** i.i.d.

indip.: OK

i.d. → NO

Esempio 3

$$X_i \sim \frac{1}{\sqrt{2\pi}\,\sigma_i} e^{-\frac{(x_i-\theta)^2}{2\sigma_i^2}}$$

$\sigma_1 \dots \sigma_n$: parametri noti (shape)
↳ diversi tra loro

θ : parametro da stimare
↳ comune a tutte le X_i

$$L(\theta) = \frac{1}{(2\pi)^n} \frac{1}{\prod_{i=1}^{n} \sigma_i} e^{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\theta}{\sigma_i}\right)^2}$$

$$\ell(\theta) \equiv \ln L = -\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\theta}{\sigma_i}\right)^2 + \dots$$

termini indip. da θ

$$\frac{d\ell}{d\theta} = \sum_{i=1}^{n} \frac{x_i-\theta}{\sigma_i^2} = \sum_{i=1}^{n} \frac{x_i}{\sigma_i^2} - \theta \sum_{i=1}^{n} \frac{1}{\sigma_i^2} = 0$$

$$\boxed{\hat{\theta}_n(x_1 \dots x_n) = \frac{\sum_{i=1}^{n} \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}}$$