Modelli Statistici

Antonio Lucadamo

antonio.lucadamo@unisannio.it

La funzione di verosimiglianza

Consideriamo $\mathcal{F}$ un modello statistico parametrico per i dati y. Gli elementi di $\mathcal{F}$ sono o tutti funzioni di densità di probabilità (nel caso continuo) o tutti funzioni di probabilità (nel caso discreto). In entrambi i casi si può scrivere:

$$\mathcal{F} = \{p_Y(y;\theta), \theta \in \Theta \subseteq \mathbb{R}^d\}$$

dove θ parametro d-dimensionale con valori nello spazio parametrico Θ e y assume valori nel supporto di Y sotto θ.

In generale si indica con $p_Y(y;\theta)$ la funzione di densità e il caso continuo o discreto sarà ricostruibile dal supporto di Y.

La funzione di verosimiglianza

Consideriamo $\mathcal{F}$ un modello statistico parametrico per i dati y. Gli elementi di $\mathcal{F}$ sono o tutti funzioni di densità di probabilità (nel caso continuo) o tutti funzioni di probabilità (nel caso discreto). In entrambi i casi si può scrivere:

$$\mathcal{F} = \{p_Y(y;\theta), \theta \in \Theta \subseteq \mathbb{R}^d\}$$

dove θ parametro d-dimensionale con valori nello spazio parametrico Θ e y assume valori nel supporto di Y sotto θ.

In generale si indica con $p_Y(y;\theta)$ la funzione di densità e il caso continuo o discreto sarà ricostruibile dal supporto di Y.

La funzione di verosimiglianza

Si assume che θ sia identificabile cioè che la corrispondenza fra Θ e $\mathcal{F}$ sia biunivoca. Sia $p^0(y)$ la vera e ignota densità di Y. Il modello $\mathcal{F}$ è detto correttamente specificato se $p^0(y) \in \mathcal{F}$.

Se $\mathcal{F}$ è correttamente specificato il valore θ^0 tale che $p_Y(y;\theta^0) = p^0(y)$ è detto vero valore del parametro. La funzione $L : \Theta \to [0, +\infty)$ definita da

$$L(\theta) = c(y)p_Y(y;\theta)$$

con $c(y) > 0$ costante non dipendente da θ è detta funzione di verosimiglianza di θ basata sui dati y.

La funzione di verosimiglianza

Si assume che θ sia identificabile cioè che la corrispondenza fra Θ e $\mathcal{F}$ sia biunivoca. Sia $p^0(y)$ la vera e ignota densità di Y. Il modello $\mathcal{F}$ è detto correttamente specificato se $p^0(y) \in \mathcal{F}$.

Se $\mathcal{F}$ è correttamente specificato il valore θ^0 tale che $p_Y(y;\theta^0) = p^0(y)$ è detto vero valore del parametro. La funzione $L : \Theta \to [0, +\infty)$ definita da

$$L(\theta) = c(y)p_Y(y;\theta)$$

con $c(y) > 0$ costante non dipendente da θ è detta funzione di verosimiglianza di θ basata sui dati y.

Spesso si usa la log-verosimiglianza:

$$l(\theta) = \log L(\theta)$$

dove se $L(\theta) = 0$ si avrà $l(\theta) = -\infty$
Con $Y = (Y_1, \ldots, Y_n)^T$ e $Y_1, \ldots, Y_n$ indipendenti, con densità marginale $p_{Y_i}(y_i; \theta)$ si ha

$$l(\theta) = \sum_{i=1}^{n} \log p_{Y_i}(y_i; \theta)$$

La funzione di verosimiglianza sintetizza l'informazione disponibile su θ alla luce dei dati y. Permette di confrontare l'adeguatezza, alla luce dei dati, di coppie di valori parametrici $\theta^{'}$ e $\theta^{''}$ in Θ, tramite il rapporto di verosimiglianza.

Proprietà di invarianza

- $L(\theta)$ è invariante rispetto a trasformazioni biiettive dei dati y;
- $L(\theta)$ è invariante rispetto a riparametrizzazioni.

Proprietà campionarie

Proprietà esatte

In modelli con verosimiglianza regolare si ha:

- $E_\theta(l_*(\theta; Y)) = 0 \quad \forall \quad \theta \in \Theta$
- $E_\theta(l_*(\theta; Y)(l_*(\theta; Y))^T) = i(\theta) \quad \forall \quad \theta \in \Theta$

Proprietà campionarie

Proprietà asintotiche

Sotto alcune condizioni, fra cui è importante che la dimensione di Θ non dipenda da n, lo stimatore di massima verosimiglianza è consistente:

$$\hat{\theta}_n \xrightarrow{p} \theta$$

Inoltre

- $l_*(\theta) \dot{\sim} N_d(0, i(\theta))$
- $\hat{\theta} - \theta \dot{\sim} N_d(0, i(\theta)^{-1})$
- $\hat{\theta} - \theta \dot{\sim} N_d(0, j(\hat{\theta})^{-1})$

Proprietà campionarie

- $W_e(\theta) = (\hat{\theta} - \theta)^T j(\hat{\theta})(\hat{\theta} - \theta) \dot{\sim} \chi^2_d$
- $W_u(\theta) = l_*(\theta)^T i(\theta)^{-1} l_*(\theta) \dot{\sim} \chi^2_d$
- $W(\theta) = 2\{l(\hat{\theta}) - l(\theta)\} \dot{\sim} \chi^2_d$

Le tre quantità sono asintoticamente equivalenti e sono usate per costruire test e regioni di confidenza.