Funzione di verosimiglianza

Osservato un determinato campione $x_1, \ldots, x_n$ estratto da una popolazione X la cui distribuzione dipende da un parametro θ

La funzione di verosimiglianza $L(\theta)$ indica la probabilità di osservare il campione al variare del parametro θ

Poiché le osservazioni campionarie sono indipendenti e identicamente distribuite possiamo scrivere la funzione di verosimiglianza come prodotto delle probabilità delle singole osservazioni campionarie:

$$L(\theta) = P(\text{dati osservati}; \theta) = P(X_1 = x_1, \ldots, X_n = x_n; \theta) =$$

$$= P(X_1 = x_1; \theta) \ldots P(X_n = x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

dove $f(x_i; \theta)$ corrisponde alla probabilità o alla densità a seconda se X è discreta o continua.

Metodo di massima verosimiglianza

Osservazione:

Nella funzione di verosimiglianza i dati campionari sono fissati, mentre il valore del parametro incognito può variare.

Quindi la $L(\theta)$ deve essere vista come una funzione del solo parametro θ che ci indica quanto è plausibile uno specifico valore del parametro ignoto, una volta che sia stato osservato il campione.

Se θ_1 e θ_2 sono due valori del parametro ignoto e se, dato il campione, $L(\theta_1) > L(\theta_2)$, diremo che il valore θ_1 è più verosimile di θ_2, alla luce del campione osservato.

Metodo di massima verosimiglianza

Il metodo di massima verosimiglianza per ottenere una stima del parametro ignoto, consiste nel prendere il valore di θ che massimizza la funzione di verosimiglianza $L(\theta)$.

Il valore che massimizza la $L(\theta)$ è la stima di massima verosimiglianza di θ ed è indicato con $\hat{\theta}$.

Per convenienza si utilizza massimizzare il logaritmo della $L(\theta)$ detta log-verosimiglianza, $\log L(\theta)$.

La stima di massima verosimiglianza di θ è il valore $\hat{\theta}$ che massimizza la $\log L(\theta)$, ossia

$$\log L(\hat{\theta}) = \sup_{\theta} \log L(\theta)$$

Stimatore di massima verosimiglianza

La stima di massima verosimiglianza di θ è la soluzione dell'equazione di verosimiglianza: $\partial \log L(\theta) / \partial\theta = 0$
che nel punto $\theta = \hat{\theta}$ soddisfa $\partial^2 \log L(\theta) / \partial\theta^2 = 0$

Al variare del campione osservato si avrà in generale una diversa stima di massima verosimiglianza del parametro. Si ottiene perciò una v.c. detta stimatore di massima verosimiglianza del parametro θ .

Se la popolazione è Normale con media μ e varianza σ^2
gli stimatori di massima verosimiglianza per μ e σ^2 sono:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad\qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

Esempio

Popolazione Bernoulliana con parametro π

Campione osservato $x_1, \dots, x_n$

$$L(\pi) = \prod_{i=1}^{n} p(X_i = x_i; \pi) == \prod_{i=1}^{n} \pi^{x_i} (1-\pi)^{1-x_i} =$$

$$= \pi^{x_1+\cdots+x_n} (1-\pi)^{(1-x_1)+\cdots+(1-x_n)} =$$

$$= \pi^{x_1+\cdots+x_n} (1-\pi)^{n-(x_1+\cdots+x_n)}$$

$$\partial \log L(\pi) / \partial \pi = \frac{\partial}{\partial \pi} \left[\log \left(\pi^{x_1+\cdots+x_n} (1-\pi)^{n-(x_1+\cdots+x_n)} \right) \right] =$$

$$= \frac{x_1 + \cdots + x_n - n\pi}{\pi(1-\pi)} = 0 \Longrightarrow x_1 + \cdots + x_n - n\pi = 0$$

$$\boxed{\hat{\pi} = \overline{x}}$$