Asintotica gaussianità

Per dimostrare che per lo stimatore di ML valga l'asintotica gaussianità introduciamo

Def Condizioni di regolarità

1. Le condizioni di consistenza → legge dei grandi numeri uniforme → Convessità
2. Le osservazioni $X_1, \dots, X_n$ sono iid.
3. La log verosimiglianza ammette due derivate continue $\frac{\partial^2 \log L_n}{\partial^2 \theta} \in C^2$
4. La derivata seconda ha momento finito $E\left[\left|\frac{\partial^2 \log L_n}{\partial^2 \theta}\right|\right] < \infty$
5. Il vero parametro θ_0 appartiene all'interno dell'insieme dei valori ammissibili, $\theta_0 \in Int(\Theta)$
6. È possibile scambiare due volte la derivata con l'integrale del valor medio della log-verosimiglianza

Oss

Queste condizioni sono più forti di quelle che ci servono, ma le definiamo in questo modo così che vengano definite solo una volta

Def Funzione punteggio

La funzione punteggio $S_n : \mathbb{R}^p \longrightarrow \mathbb{R}^p$ è definita da

$$S_n(\theta; X_1, \dots, X_n) = \frac{\partial}{\partial \theta} \log L(\theta; X_1, \dots, X_n)$$

Assumendo che esista la derivata esista

Oss

È possibile notare come S_n sia una variabile aleatoria

Lemma

Sotto le condizioni di regolarità vale che

$$E_{\theta_0}\left[S_n(\theta_0; X_1, \dots, X_n)\right] = 0$$

$\hookrightarrow \boxed{E_{\theta_0}[\cdot]\Big|_{\theta=\theta_0}}$

Dim $p=1$

Per definizione abbiamo che

Prendendo in considerazione X_i

$$E_{\theta_0}\left[S_n(\theta_0; X_1,\dots,X_n)\right] = \sum_{i=1}^{n} E\left[\frac{\partial}{\partial\theta}\log\left(f(\theta,x)\right)\right] = \int \frac{\partial}{\partial\theta}\log(f(\theta;x_i))\, f(\theta,x_i)\, dx_i = \int \frac{\frac{\partial}{\partial\theta} f(\theta,x_i)}{\cancel{f(\theta,x_i)}}\, \cancel{f(\theta,x_i)}\, dx_i =$$

applichiamo la condizione 6

$$= \int \frac{\partial}{\partial\theta} f(\theta,x_i) = \frac{\partial}{\partial\theta}\underbrace{\int f(\theta,x_i)}_{\text{densità}} = \frac{\partial}{\partial\theta} 1 = 0$$

Es

Sia $\{X_n\}_{n\in\mathbb{N}}$ una sequenza di v.a. con distribuzione $X \sim N(\mu,1)$

e quindi legge $f(x_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2}$

Sappiamo che

$$L(\mu; x_1,\dots,x_n) = \prod_{i=1}^{n} f(x_i;\mu) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Quindi abbiamo

$$\log L(\mu; x_1,\dots,x_n) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}(X_i-\mu)^2$$

Derivando rispetto a μ abbiamo che

$$\frac{\partial}{\partial\mu}\log L(\mu; X_1,\dots,X_n) = S_n(\mu; X_1,\dots,X_n) = \sum_{i=1}^{n}(X_i-\mu)$$

Applicando poi il valor medio abbiamo che

$$E\left[S_n(\mu; X_1,\dots,X_n)\right] = n(\mu-\mu_0)\Big|_{\mu=\mu_0} = 0$$

Lemma

Sotto le condizioni di regolarità vale che

per $p=1$

$$E\left[\left(S_n(\theta, x_1,\dots,x_n)\right)^2\right] = -E\left[\frac{\partial}{\partial\theta} S_n(\theta; x_1,\dots,x_n)\right] \rightsquigarrow -E\left[\frac{\partial^2}{\partial^2\theta}\log L(\theta; x_1,\dots,x_n)\right]$$

per $p>1$

$$E\left[\underbrace{(\nabla \log L)}_{p\times 1}\underbrace{(\nabla \log L)^T}_{1\times p}\right] = -E\left[\underbrace{H \log L}_{p\times p}\right]$$

gradiente $\rightsquigarrow$ vettore delle derivate parziali rispetto un solo parametro

Dove H e' la matrice Hessiana $\rightsquigarrow$ matrice contenente tutte le derivate 2° parziali

Dim $p=1$

$$0 = \frac{\partial}{\partial\theta}\int \log' L(\theta; x_1,\dots,x_n)\, f(x_1,\dots,x_n;\theta)\, dx_1,\dots,dx_n = \int \log'' L(\theta;x_1,\dots,x_n)\, f(x_1,\dots,x_n;\theta)\, dx_1,\dots,dx_n + \int \log' L(\theta;x_1,\dots,x_n)\, f'(x_1,\dots,x_n;\theta)\, dx_1,\dots,dx_n =$$

$$= E\left[\frac{\partial^2}{\partial^2\theta}\log L(\theta;x_1,\dots,x_n)\right] + \int \log' L(\theta;x_1,\dots,x_n)\underbrace{\frac{f'(x_1,\dots,x_n;\theta)}{f(x_1,\dots,x_n;\theta)}}_{\log' L(x_1,\dots,x_n)} f(x_1,\dots,x_n;\theta)\, dx_1,\dots,dx_n =$$

$$= E\left[\frac{\partial^2}{\partial^2\theta}\log L(\theta;x_1,\dots,x_n)\right] + \int \left(\log' L(\theta;x_1,\dots,x_n)\right)^2 f(x_1,\dots,x_n;\theta)\, dx_1,\dots,dx_n =$$

$$= E\left[\frac{\partial^2}{\partial^2\theta}\log L(\theta;x_1,\dots,x_n)\right] + E\left[\left(\frac{\partial}{\partial\theta}\log L(\theta;x_1,\dots,x_n)\right)^2\right]$$

Oss

Abbiamo dimostrato che $E\left[S_n(\theta_0)\right]=0$

Ponendo $\theta=\theta_0$ abbiamo che

$$-E\left[\frac{\partial^2}{\partial\theta}\log L(\theta;x_1,\dots,x_n)\right] = E\left[\left(S_n(\theta_0;x_1,\dots,x_n)\right)^2\right] = Var\left(S_n(\theta_0)\right)$$

Quindi la varianza di $S_n(\theta)$ e' uguale al reciproco della derivata seconda della $\log L$ nello stesso punto

Def Informazione di Fisher

La varianza della funzione punteggio e' nota come informazione di Fisher

$$Var(S_n(\theta_0)) = I_n(\theta_0)$$

d: dimensione $p\times p$

Teo Asintotica gaussianita' dello stimatore di ML

Sotto le condizioni di regolarita' vale che

$$\sqrt{n}\left(\tilde{\theta}_{ML}-\theta_0\right) \xrightarrow[d]{} N\left(0; I_1^{-1}(\theta_0)\right)$$

Dim $p=1$

Per il teo del valor medio di Lagrange

$$\exists\ \bar{\theta}_n \in (\theta_0;\tilde{\theta}_{ML}) \text{ tale che}$$

$$0 = \log' L(\tilde{\theta}_{ML}) = \log' L(\theta_0) + \log''(\bar{\theta}_n)(\tilde{\theta}_{ML}-\theta_0)$$

Possiamo riscrivere la relazione come

$$\sqrt{n}\left(\tilde{\theta}_{ML}-\theta_0\right) = -\frac{\log' L(\theta_0)/\sqrt{n}}{\log'' L(\bar{\theta}_{ML})/n}$$

Studiamo ora il comportamento di numeratore e denominatore

a) Numeratore

$$\frac{\log' L(\theta_0)}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial \log f(X_i;\theta)}{\partial \theta} \xrightarrow[d]{} N(0; I_1(\theta_0))$$

Somma di v.a. i.i.d. ⤳ applichiamo il teo del limite centrale

b) Denominatore

$$\frac{\log'' L(\bar{\theta}_n)}{n} = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial^2 \theta_n} \log f(x_i;\theta) \xrightarrow[p]{} \bar{E}\left[\frac{\partial^2}{\partial \theta^2} \log L(\theta; X_1, \dots, X_n)\right] = Var(S_n(\theta_0)) = \tilde{I}_1(\theta_0)$$

Somma di v.a. iid. con $\bar{E}\left[\frac{\partial^2}{\partial \theta^2} \log f(x_i;\theta)\right] < \infty$ ⤳ legge dei grandi numeri U.I.

Per Slutzky abbiamo quindi l'enunciato del teorema