Fitting 2 - ML Parte 3

Wednesday, November 29, 2023 8:30 AM

Invarianza per trasf.^ni del parametro

$L(x_1 \dots x_n ; \vartheta)$

Supp.^mo di voler stimare non ϑ, ma una sua funzione assegnata $\tau(\vartheta) \Rightarrow$ stima di τ

Se τ è: monotona

suff.^te smooth (derivabile almeno 2 volte)

Riparametrizzo L come f.^ne di τ : $L = L(\tau)$

$$\left.\frac{d \ln L}{d\tau}\right|_{\hat{\tau}} = 0 \qquad \left.\frac{d^2 \ln L}{d\tau^2}\right|_{\hat{\tau}} < 0$$

Wednesday, November 29, 2023 8:30 PM

Ma:
$$\frac{d\ln L}{d\tau} = \frac{d\ln L}{d\theta}\frac{d\theta}{d\tau} = \frac{\frac{d\ln L}{d\theta}}{\frac{d\tau}{d\theta}}$$

$$\left.\frac{d\ln L}{d\tau}\right|_{\hat{\tau}} = 0 \iff \left.\frac{d\ln L}{d\theta}\right|_{\hat{\theta}} = 0$$

$\hookrightarrow \hat{\tau} = \tau(\hat{\theta})$

Inoltre:
$$\frac{d^2\ln L}{d\tau^2} = \frac{d^2\ln L}{d\theta^2}\left(\frac{d\theta}{d\tau}\right)^2 + \cancel{\frac{d\ln L}{d\theta}\left(\frac{d^2\theta}{d\tau^2}\right)}$$

$\hookrightarrow$ per $\theta = \hat{\theta}$

$$\left.\frac{d^2\ln L}{d\tau^2}\right|_{\hat{\tau}} = \frac{\left.\frac{d^2\ln L}{d\theta^2}\right|_{\hat{\theta}}}{\left(\frac{d\tau}{d\theta}\right)^2}$$

Hanno lo stesso segno

Se $\hat{\theta}$ è l'estimatore ML di θ $\Rightarrow$ $\hat{\tau} = \tau(\hat{\theta})$ è l'estimatore ML di τ

Wednesday, November 29, 2023 8:30 AM

$$E[\hat{\tau}] = E[\tau(\hat{\theta})] \neq \tau(E[\hat{\theta}])$$

$\hookrightarrow$ A meno che τ non sia lineare in θ

$\Downarrow$

<u>In generale</u> (per una generica f.^{ne} $\tau(\theta)$) la stima ML di τ <u>non è</u> unbiased

<u>Proprietà dell'estimatore ML</u>

*) Invarianza per trasf.ⁿⁱ del parametro

*) Consistenza $\longrightarrow$ $\hat{\theta}_n$ converge in probabilità al "valore vero" θ che vuole stimare

*) Efficienza
*) Normalità
} asintotiche ($n \to \infty$)

*) Sufficienza

Wednesday, November 29, 2023 8:25 AM

Stima della varianza del MLE

*) Metodo analitico

↳ Sfrutta l'inv. per trasf.^{ni} del parametro
Applicabile quando la stima di ML è effettuabile in maniera analitica

↳ Es.: distribuzione esponenziale

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

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

*) Metodo MC

*) Metodo basato sul Minimum Variance Bound (MVB)

$$V_\theta\left[\hat{\theta}_n\right] \geq \frac{\left(1+\frac{db_n}{d\theta}\right)^2}{\underbrace{E_\theta\left[-\left.\frac{\partial^2 \ln L(x_1 \ldots x_n;\theta)}{\partial\theta^2}\right|_\theta\right]}_{\text{Informazione di Fisher}}} = MVB_n(\theta)$$

$\hat{\theta}_n \to$ Generico estimatore di θ basato su n misure

$b_n(\theta) =$ bias

Nel caso unbiased, o almeno $\frac{db_n}{d\theta} = 0$ o trascurabile rispetto a 1:

$$\frac{1}{MVB_n(\theta)} = E_\theta\left[-\left.\frac{\partial^2 \ln L(x_1 \ldots x_n;\theta)}{\partial\theta^2}\right|_\theta\right] =$$

$$= -\int_{-\infty}^{+\infty} dx_1 \ldots \int_{-\infty}^{+\infty} dx_n \left.\frac{\partial^2 \ln L(x_1 \ldots x_n;\theta)}{\partial\theta^2}\right|_\theta L(x_1 \ldots x_n;\theta)$$

Nel caso iid: $L(x_1 \ldots x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta)$

$\Downarrow$

$$\ln L(x_1 \ldots x_n;\theta) = \sum_{i=1}^{n} \ln f(x_i;\theta)$$

$\Downarrow$

$$\left.\frac{\partial^2 \ln L}{\partial\theta^2}\right|_{\theta} = \sum_{i=1}^{n} \left.\frac{\partial^2 \ln f(x_i;\theta)}{\partial\theta^2}\right|_{\theta}$$

Per cui:

$$\frac{1}{MVB_n(\theta)} = \int_{-\infty}^{+\infty} dx_1 \ldots \int_{-\infty}^{+\infty} dx_n \left(- \sum_{i=1}^{n} \left.\frac{\partial^2 \ln f(x_i;\theta)}{\partial\theta^2}\right|_{\theta} \right) \prod_{i=1}^{n} f(x_i;\theta) =$$

$$= n \int_{-\infty}^{+\infty} - \left.\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2}\right|_{\theta} f(x;\theta)\, dx = \frac{n}{MVB_1(\theta)}$$

$\Downarrow$

$$MVB_n(\theta) = \frac{MVB_1(\theta)}{n}$$

caso generale di più parametri $(p>1)$

$$U^{-1}_{ij} = E_{\underline{\theta}}\left[- \frac{\partial^2 \ln L(x_1 \ldots x_n ; \underline{\theta})}{\partial\theta_i \partial\theta_j}\bigg|_{\underline{\theta}} \right]$$

↳ Matrice di covarianza di $\hat{\theta}_1 \ldots \hat{\theta}_p$

Anche in questo caso: $U^{-1}_{ij} = n \int_{-\infty}^{+\infty} \left(- \frac{\partial^2 \ln f(x;\theta)}{\partial\theta_i \partial\theta_j}\bigg|_{\theta} \right) f(x;\theta)\,dx$

$$\downarrow$$

$$U_{ij} \propto \frac{1}{n}$$

Se n è suff.^{te} grande: $$V_\theta\left[\underset{\text{MLE}}{\hat{\theta}_n}\right] \simeq \frac{1}{E_\theta\left[-\frac{\partial^2 \ln L(x_1..x_n;\theta)}{\partial\theta^2}\bigg|_\theta\right]}$$

$$V_\theta\left[\hat{\theta}_n\right] \simeq MVB_n(\theta) = \frac{MVB_1(\theta)}{n}$$

Inv. per Trasf.^{ni} del param.

$$\hat{V}\left[\hat{\theta}_n\right] \simeq MVB_n\left(\hat{\theta}_n\right) = \frac{MVB_1(\hat{\theta}_n)}{n} =$$

$$= \frac{1}{n}\,\frac{1}{E_{\hat{\theta}_n}\left[-\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2}\bigg|_{\hat{\theta}_n}\right]}$$

Wednesday, November 29, 2023 8:25 AM

$$E[\ \cdot\ ] \simeq \frac{1}{n}\sum_{i=1}^{n} \xi_i$$

$$n\, E_{\hat{\theta}_n}\left[-\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2}\Big|_{\hat{\theta}_n} \right] \simeq$$

$$\simeq \cancel{n}\,\frac{1}{\cancel{n}}\sum_{i=1}^{n} \underbrace{\left(-\frac{\partial^2 \ln f(x_i;\theta)}{\partial\theta^2}\Big|_{\hat{\theta}_n} \right)}_{\xi_i} =$$

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

$$\Downarrow$$

$$\boxed{\hat{V}[\hat{\theta}_n] \simeq \frac{1}{-\frac{\partial^2 \ln L(x_1 \dots x_n;\theta)}{\partial\theta^2}\Big|_{\hat{\theta}_n}}}$$

MLE

*) Metodo grafico

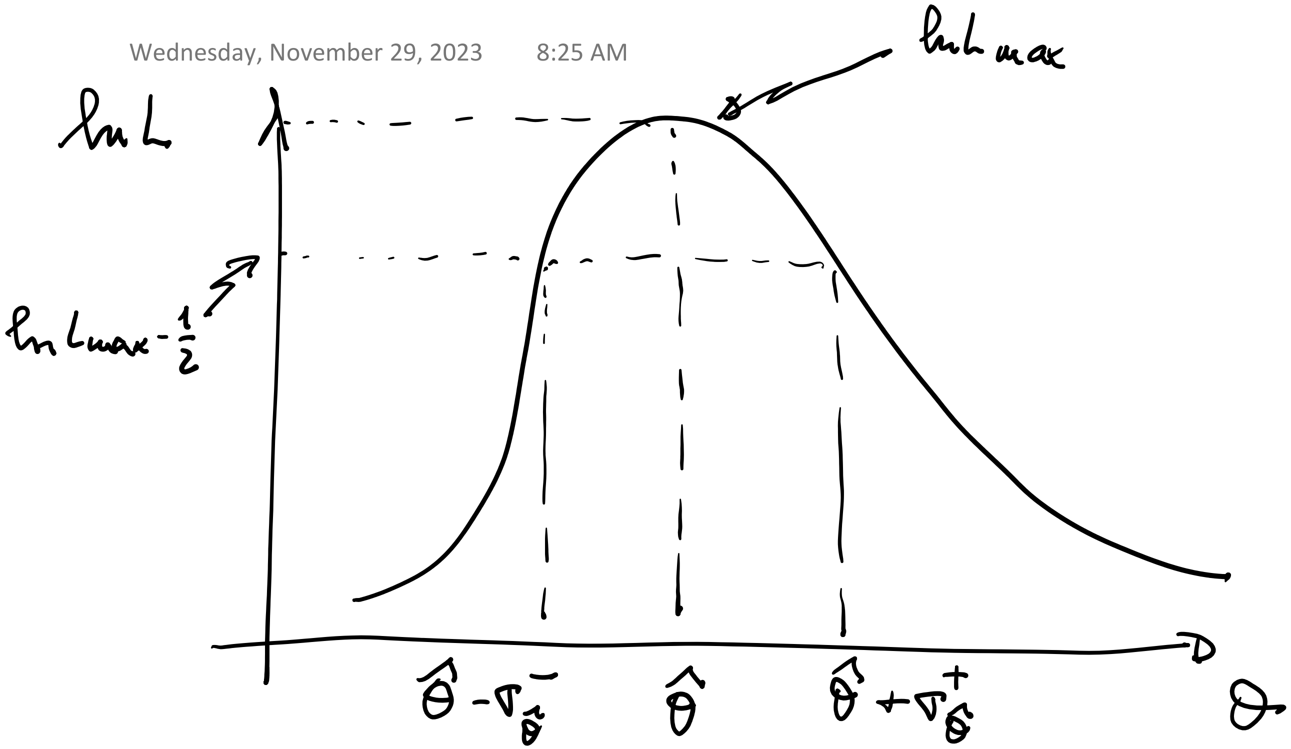
$$\ln L(\theta) = \underbrace{\ln L(\hat{\theta})}_{\ln L_{max}} + \cancel{\frac{\partial \ln L}{\partial \theta}\Big|_{\hat{\theta}}} (\theta - \hat{\theta}) + \frac{1}{2} \frac{\partial^2 \ln L}{\partial \theta^2}\Big|_{\hat{\theta}} (\theta - \hat{\theta})^2 + \dots$$

Termini trascurabili per n suff.te grande

$$\ln L(\theta) \simeq \ln L_{max} + \frac{1}{2} \frac{\partial^2 \ln L}{\partial \theta^2}\Big|_{\hat{\theta}} (\theta - \hat{\theta})^2 \simeq$$

$$\simeq \ln L_{max} - \frac{1}{2} \frac{(\theta - \hat{\theta})^2}{\hat{V}(\hat{\theta})}$$

$$\hat{V}[\hat{\theta}] \simeq - \frac{1}{\frac{\partial^2 \ln L}{\partial \theta^2}\Big|_{\hat{\theta}}}$$



$p = 1$

λ $x_1 \dots x_n$ fissati

Wednesday, November 29, 2023 8:25 AM

Extended ML

$X \sim f(X;\theta)$

$x_1 \dots x_n$ n misure di X

n realizzazione della misura di una variabile casuale (numero di conteggi)

Likelihood "estesa"

$$L(n, x_1 \dots x_n; \nu, \underline{\theta}) = \frac{\nu^n e^{-\nu}}{n!} \prod_{i=1}^{n} f(x_i; \underline{\theta})$$

↳ variabile casuale Poissoniana con $E[n] = \nu$

$n \sim \mathcal{P}(\nu)$ $\qquad P(n;\nu) = \frac{\nu^n e^{-\nu}}{n!}$

*) Se ν non dip. da θ

$$\hat{\nu} : \left.\frac{\partial \ln L}{\partial \nu}\right|_{\hat{\nu}} = 0$$

$$\Downarrow$$

$$\frac{n}{\hat{\nu}} - 1 = 0 \quad \rightarrow \quad \hat{\nu} = n$$

*) $\nu = \nu(\underline{\theta})$

Termini che non dip. da $\underline{\theta}$

$$\ln L(n, x_1 \ldots x_n; \underline{\theta}) = n \ln \nu(\underline{\theta}) - \nu(\underline{\theta}) + \sum_{i=1}^{n} \ln f(x_i; \underline{\theta}) + \ldots$$

$$= -\nu(\underline{\theta}) + \sum_{i=1}^{n} \ln \left(\nu(\underline{\theta}) f(x_i; \underline{\theta})\right) + \ldots$$

ML con dati "binnati"

$x_1 \ldots x_n \qquad f(X;\theta)$

n_i = num di misure nell'i-mo bin di un istogramma con un range "pari" al supporto di X e con N bin

$(x_i^{min}, x_i^{max}) \rightarrow$ range i-mo bin

$\underline{n}^T = (n_1, \ldots, n_N)$

sotto l'hp che $X \sim f(X;\theta)$ i valori di aspettaz. delle n_i, che chiameremo $\underline{\nu}^T(\theta) = (\nu_1(\theta), \ldots, \nu_N(\theta))$

$$\nu_i(\theta) = n \int_{x_i^{min}}^{x_i^{max}} f(x;\theta)\,dx \qquad \therefore \quad p_i(\theta) = \int_{x_i^{min}}^{x_i^{max}} f(x;\theta)\,dx = \frac{\nu_i(\theta)}{n}$$

Num tot di misure : $n = \sum_{i=1}^{N} n_i$

Pdf congiunta di $n_1 \dots n_N$ $\rightarrow$ <u>distribuzione multinomiale</u>

(generalizzazione per $N \geq 2$ della distribuzione binomiale)

$$f_{cong}\left(\underline{n} ; \underline{\nu}(\theta)\right) = \frac{n!}{n_1! \dots n_N!} p_1^{n_1}(\theta) \dots p_N^{n_N}(\theta) =$$

$$= \frac{n!}{n_1! \dots n_N!} \left(\frac{\nu_1(\theta)}{n}\right)^{n_1} \dots \left(\frac{\nu_N(\theta)}{n}\right)^{n_N}$$

$$\ln L(n_1 \dots n_N ; \theta) = \sum_{i=1}^{N} n_i \ln \nu_i(\theta) + \dots$$

termini indip. da θ

ML binnata ed estesa

$$f_{cong}(n, \underline{n}; \nu, \theta) = \frac{\nu^n e^{-\nu}}{n!} \frac{n!}{n_1! \ldots n_N!} p_1^{n_1}(\theta) \ldots p_N^{n_N}(\theta) =$$

$\underline{n}$ ↳ $\underline{n}^T = (n_1 \ldots n_N)$

n → Num. di conteggi (var. poissoniana con $E[n] = \nu$)

$$p_i = \frac{\nu_i}{\nu}$$

$$e^{-\sum_{i=1}^{N} \nu_i} = \prod_{i=1}^{N} e^{-\nu_i}$$

$$= \frac{\cancel{\nu^n} e^{-\nu}}{\cancel{n!}} \frac{\cancel{n!}}{n_1! \ldots n_N!} \underbrace{\left(\frac{\nu_1(\theta)}{\nu}\right)^{n_1} \ldots \left(\frac{\nu_N(\theta)}{\nu}\right)^{n_N}}_{\frac{\nu_1^{n_1}(\theta) \ldots \nu_N^{n_N}(\theta)}{\cancel{\nu^n}}} =$$

$$\nu = \sum_{i=1}^{N} \nu_i$$

$$n = \sum_{i=1}^{N} n_i$$

$$= \boxed{\prod_{i=1}^{N} \frac{e^{-\nu_i} \nu_i^{n_i}}{n_i!} = L(\nu_i; n_i)}$$

serve per test hypo

$$\ln L(n, \underline{n}; \nu, \theta) = -\nu + \sum_{i=1}^{N} n_i \ln \nu_i(\nu, \theta) + \dots$$

Termini indip. da θ