Dati:

$X_1 \ldots X_n$ var. casuali i.i.d

$\underline{\theta} = (\theta_1 \ldots \theta_p)'$ parametro

$\Rightarrow^{(*)} L_x(\underline{\theta}) = \prod_{i=1}^{n} f(x_i / \underline{\theta}) = f(\underline{x}, \underline{\theta})$

ovviamente

ogni $\underline{x} = (x_1 \ldots x_n)$ individua

un diverso valore per $L(\cdot)$

Il valore che massimizza (*)

è la "stima di massima verosi-

miglianza" $\hat{\underline{\theta}}$

NB: Talvolta risulta più conveniente

utilizzare il logaritmo della veros.

$l_x(\underline{\theta}) = \ln L_x(\underline{\theta})$

MLE: Il caso scalare

ϑ è uno scalare

il valore di ϑ: $\hat{\vartheta}$ che massimizza

la log-verosimiglianza soddisfa alla

condizione:

$$\frac{\partial \ell(\vartheta)}{\partial \vartheta} = 0 \quad (*)$$ LIKELIHOOD EQUATION

SCORE FUNCTION

Nell'ipotesi di distribuzione normale la $\ell(\cdot)$ è una fz quadratica $\Rightarrow$ ha un unico massimo.

Il problema di trovare una soluzione alla (*) è particolarmente semplice: la (*) è fz lineare di ϑ.

Se la $\ell(\cdot)$ non è quadratica $\Rightarrow$ l'equazione di verosimiglianza non è lineare.

OSS.

Sotto condizioni non molto restrittive

lo stimatore di massima verosimiglianza corretto

è | consistente, asintoticamente normale

con: $E(\hat{\theta}) = \theta$

$$Var(\hat{\theta}) = Var(\theta)/n$$

dove $$Var(\theta) = \frac{1}{E\left\{\left[\frac{\partial}{\partial\theta}\log l_x(\theta)\right]^2\right\}}$$

Si tratta di studiare tecniche numeriche per trovare le radici di equazioni non lineari

In questo caso bisogna analizzare la derivata seconda della $l(\cdot)$

Ricordo

$$I(\theta) = \left[- E_\theta \, \ddot{l}_X(\cdot) \right]\Big|_\theta = E_\theta \left[\dot{l}_X(\cdot) \right]^2 \Big|_\theta$$

FISHER INFORMATION

Il CASO DI UN UNICO PARAMETRO

è importante per 3 aspetti

1
È un caso che sorge frequentemente nelle teorie di investigazione statistica e negli studi Monte Carlo

2.
È importante nel caso di massimizzazione della verosimiglianza condizionata (un parametro per volta).

3° Molti metodi di ricerca delle radici nel caso di un problema multivariato sono estensioni di tecniche univariate.

$\exists \neq$ tecniche per risolvere un tale problema.

I metodi sono in generale iterativi
Essi sono validi anche fuori dal contesto statistico, si indicano nel seguito

$f(x) = 0$ una equazione non lineare in x, di cui vogliamo ricavare le radici

NEWTON - RAPHSON ITERATION

- Richiede alcune ipotesi sulla $f(x)$ che sembrano restrittive ma sono soddisfatte sovente nei problemi pratici.

- La convergenza è quadratica

$\Longrightarrow$ È uno dei metodi numerici più utilizzati.

OSS:

L'ordine di convergenza è così definito:

Sia s la soluzione esatta

$\varepsilon_i = |x_i - s|$ l'errore al passo i

$\Longrightarrow$ La sequenza $x_1, x_2, \ldots x_i \ldots$ si dice convergente di ordine β se:

$$\lim_{i \to \infty} \varepsilon_{i+1} = C \varepsilon_i^{\beta}$$

per qualche $C \neq 0$

$C \in \mathbb{R}$

Sia $f(x)$ due volte derivabile con derivata seconda continua

Si fissa un valore iniziale x_0

Si deriva il metodo espandendo in serie di Taylor la f e scrivendo:

$$0 = f(s) = f(x_i) + (s - x_i) f'(x_i) + \frac{(s-x_i)^2}{2} f''(x_i)$$

con x_i sufficientemente prossimo a s

$\Rightarrow$

$$s \simeq x_i - \frac{f(x_i)}{f'(x_i)}$$

ALGORITHMO DI NEWTON-RAPHSON

$x_0 :=$ valore iniziale

for : $i := 0$ to ∞ until convergence do

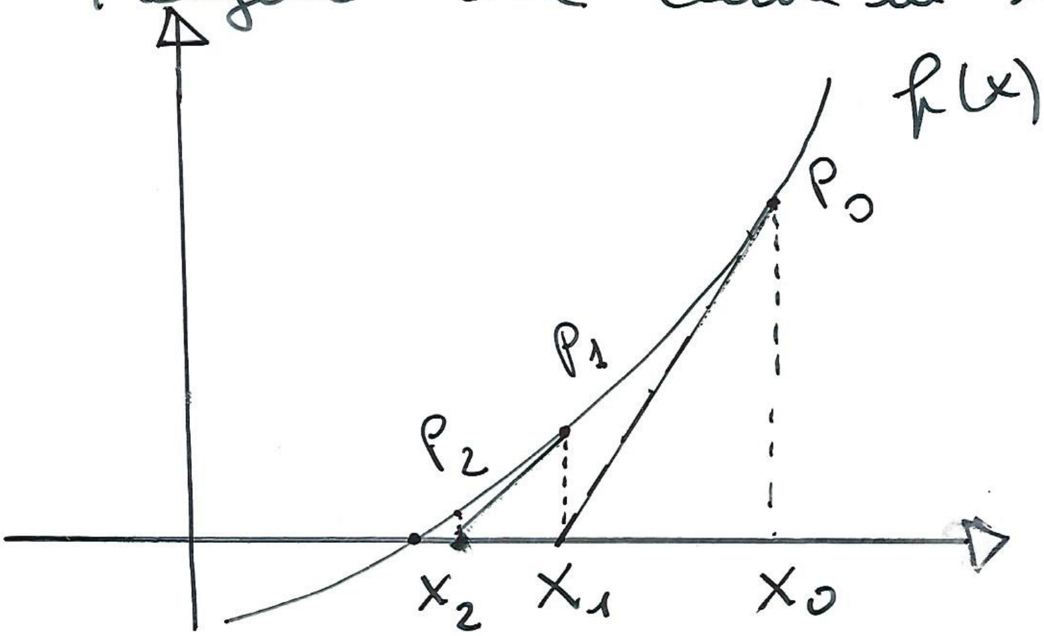
$x_{i+1} := x_i - f(x_i) / f'(x_i)$

— . —

Per garantire la convergenza il valore x_0 deve essere prossimo alla soluzione

L'algoritmo di Newton-Raphson ha una semplice interpretazione grafica

Al passo i-esimo
la funzione e' approssimata dalla
retta tangente la curva in x_i



Il punto in cui la tangente
taglia l'asse x e' il prossimo
valore dell'iterazione

$P_n : (x_n, f(x_n))$

NEWTON RAPHSON

Sotto le ipotesi che

- La log-likelihood sia due volte derivabile con derivate seconde continue

Si può utilizzare il metodo di Newton-Raphson

SIA:

$$\dot{\ell}(\theta_0) = \left. \frac{\partial \ell(\theta)}{\partial \theta} \right|_{\theta=\theta_0}$$

$$\ddot{\ell}(\theta_0) = \left. \frac{\partial^2 \ell(\theta)}{\partial \theta} \right|_{\theta=\theta_0}$$

SI CONSIDERI L'APPROSSIMAZIONE DELLA $\ell(\theta)$ ~~DATA~~ DALLA ESPANSIONE IN SERIE DI TAYLOR:

$$\ell(\theta) \cong \ell(\theta_0) + (\theta - \theta_0)\, \dot{\ell}(\theta_0) + \frac{(\theta - \theta_0)^2}{2} \ddot{\ell}(\theta_0)$$

(*)

L'idea alla base del metodo N-R è cercare θ in modo da massimizzare la (*) vale a dire

$$\frac{\partial l(\theta)}{\partial(\theta)} = \dot{l}(\theta_0) + (\theta-\theta_0)\,\ddot{l}(\theta_0) = 0$$

Sia θ_0 il valore iniziale $\Longrightarrow$

$$\dot{l}(\theta_0) = -(\theta-\theta_0)\,\ddot{l}(\theta_0)$$

ovvero

$$(\theta-\theta_0) = -\frac{\dot{l}(\theta_0)}{\ddot{l}(\theta_0)}$$

Ovvero:

$$\boxed{\theta_1 = \theta_0 - \frac{\dot{l}(\theta_0)}{\ddot{l}(\theta_0)}}$$

$$\theta_2 = \theta_1 - \frac{\dot{l}(\theta_1)}{\ddot{l}(\theta_1)}$$

$\vdots$