

**V CONGRESSO NAZIONALE SISMEC**

**REGRESSIONE QUANTILICA NON-LINEARE PER DATI  
LONGITUDINALI MEDIANTE LA DISTRIBUZIONE DI  
LAPLACE ASIMMETRICA**

Ferrante Luigi<sup>1</sup>, Gesuita Rosaria<sup>1</sup>, Cameriere Roberto<sup>2</sup>, Carle Flavia<sup>1</sup>

*<sup>1</sup>Sezione di Statistica Medica, Dip. di Medicina Clinica e Biotecnologie Applicate,  
Facoltà di Medicina e Chirurgia, Università Politecnica delle Marche*

*<sup>2</sup>Istituto di Medicina Legale, Università degli Studi di Macerata*

## dati analizzati nello studio

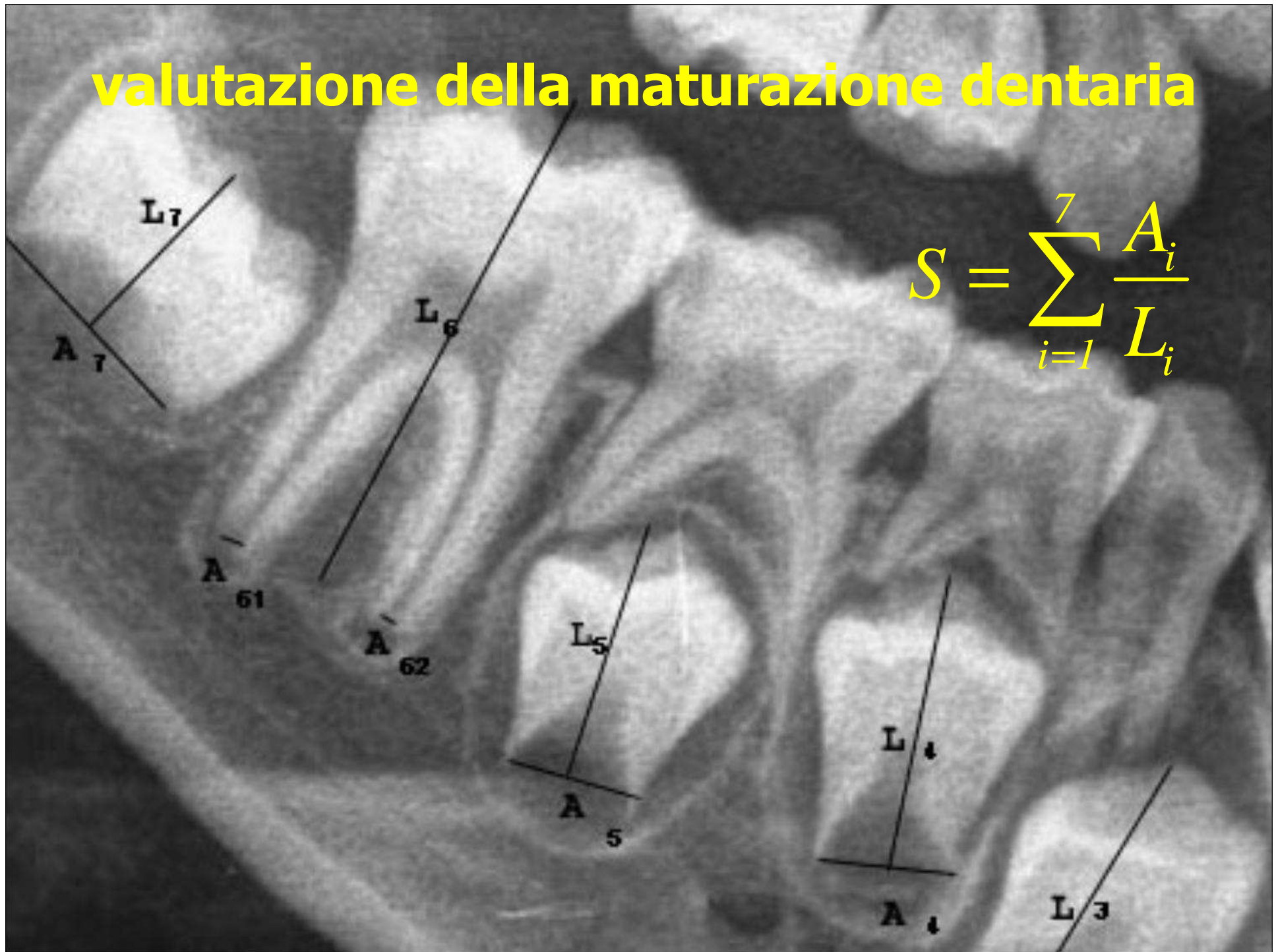
- **Studio di tipo longitudinale su N=30 individui maschi;**
- **Età: 8 – 12 anni;**
- **A ciascun ragazzo è stata valutata la maturazione dentaria in n=4 tempi diversi (circa 1 volta l'anno) mediante un indicatore, S, dell'ampiezza dell'apertura degli apici;**
- **In corrispondenza di ogni valutazione della maturazione dentaria è stata riportata l'età, E, del ragazzo.**

**Quindi i dati raccolti sono della forma:**

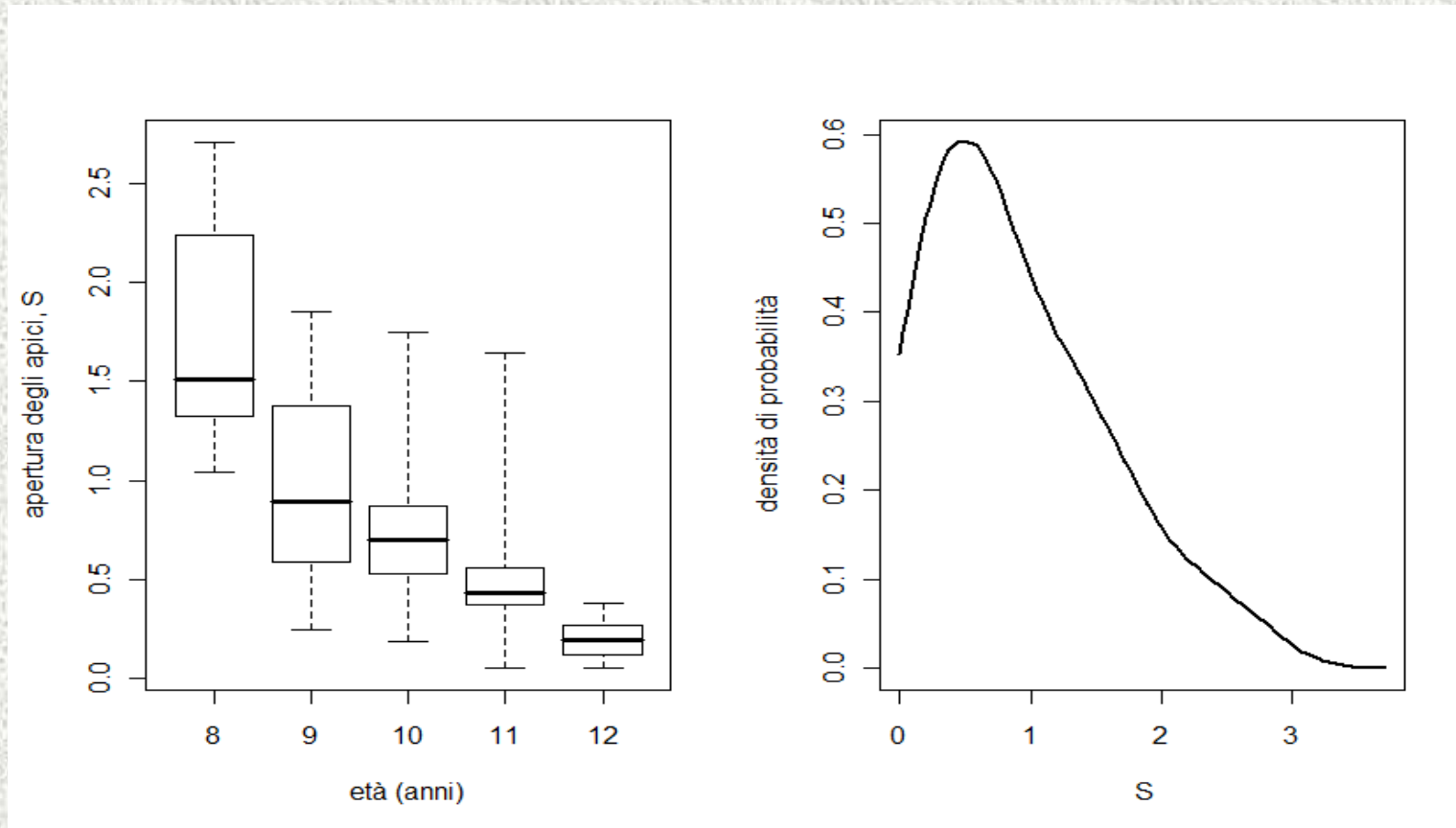
$$\{(E_{ij}, S_{ij}), j=1,\dots,n; i=1,\dots,N\}.$$

# valutazione della maturazione dentaria

$$S = \sum_{i=1}^7 \frac{A_i}{L_i}$$



## Alcune osservazioni sui dati



**Asimmetria marcata della distribuzione di S; Skew = 0.87**

**Variabilità entro soggetto delle misure ripetute**

## Regressione quantilica non lineare (NLQR)

Nella QR un quantile  $q(\tau)$  risolve il problema:

$$\min_{\beta} \sum_{i=1}^N \sum_{j=1}^n \rho_{\tau} \left( S_{ij} - q \left( \tau / E_{ij}, \beta \right) \right) \quad \tau \in [0, 1]$$

$$j = 1, \dots, n = 4; i = 1, \dots, N = 30$$

$$\rho_{\tau}(x) = x \left( \tau - I_{(x < 0)} \right) \quad \tau \in [0, 1]$$

$$g(E_{ij}, \beta) = \frac{\beta_1}{E_{ij} - \beta_2} - \beta_3$$

$$q_{S_{ij}/u_i} \left( \tau / E_{ij}, u_i, \beta \right) = g(E_{ij}, \beta) + u_i$$

## Dalla QR alla QRME

$S_{ij} / u_i$  sono v.a. i.i.d con distribuzione AL

$$f(S_{ij} / \tau, \sigma, u_i, \beta) = \frac{\tau(1-\tau)}{\sigma} \exp \left\{ -\rho_\tau \left( \frac{S_{ij} - \mu_{ij}}{\sigma} \right) \right\}$$

Con predittore  $\mu_{ij} = q_{S_{ij}/u_i}(\tau | E_{ij}, u_i, \beta)$

$u_i$  sono v.a.  $\sim N(0, \varphi)$  indipendenti

## Dalla QR alla QRME

$$\min_{\beta} \sum_{i=1}^N \sum_{j=1}^n \rho_{\tau} \left( S_{ij} - q_{S_{ij}/u_i} \left( \tau / E_{ij}, u_i, \beta \right) \right) \quad \tau \in [0, 1]$$



$$\max_{\theta} L_c(\theta / \tau, S, u) = \max_{\theta} \prod_{i=1}^N \prod_{j=1}^n \frac{\tau(1-\tau)}{\sigma} \exp \left\{ -\rho_{\tau} \left( \frac{S_{ij} - \mu_{ij}}{\sigma} \right) \right\}$$

$$\mu_{ij} = q_{S_{ij}/u_i} \left( \tau / E_{ij}, u_i, \beta \right)$$

## Verosimiglianza completa e marginale

$$\begin{aligned} L_c(\theta / \tau, S, u) &= \prod_{i=1}^N \prod_{j=1}^n f(S_{ij}, u_i / \tau, \theta) = \\ &= \prod_{i=1}^N \prod_{j=1}^n f(S_{ij} / \tau, \sigma, u_i, \beta) f(u_i / \varphi) \end{aligned}$$

$$\begin{aligned} L_{oss}(\theta / \tau, S) &= \prod_{i=1}^N \int_{-\infty}^{+\infty} \prod_{j=1}^n f(S_{ij}, u_i / \tau, \theta) du_i = \\ &= \prod_{i=1}^N \int_{-\infty}^{+\infty} \prod_{j=1}^n f(S_{ij} / \tau, \sigma, u_i, \beta) f(u_i / \varphi) du_i \end{aligned}$$

con  $\theta = (\sigma, \varphi, \beta)$ .

# Algoritmo EM

dato  $\theta^{(t)}$  il valore del vettore di parametri all'iterazione t-ma

## Passo E

$$Q(\theta | \theta^{(t)}) = E_{\theta^{(m)}} \left[ \ln L_c(\theta | \tau, S, u) / S \right] =$$
$$= \sum_{i=1}^N \int_{-\infty}^{+\infty} \left\{ \ln \left[ \prod_{j=1}^n f(S_{ij} | \tau, \sigma, u_i, \beta) \cdot f(u_i | \varphi) \right] \right\} f(u_i | S_i, \theta^{(t)}) du_i$$

## Passo M

determino  $\theta^{(t+1)}$  in modo che

$$Q(\theta^{(t+1)} | \theta^{(t)}) = \max Q(\theta | \theta^{(t)})$$

Sostituisco  $\theta^{(t)}$  con  $\theta^{(t+1)}$  e torno al passo precedente fino alla convergenza (es. errore relativo <1%)

# Algoritmo EM stocastico

dato  $\theta^{(t)}$  il valore del vettore di parametri all'iterazione t-ma

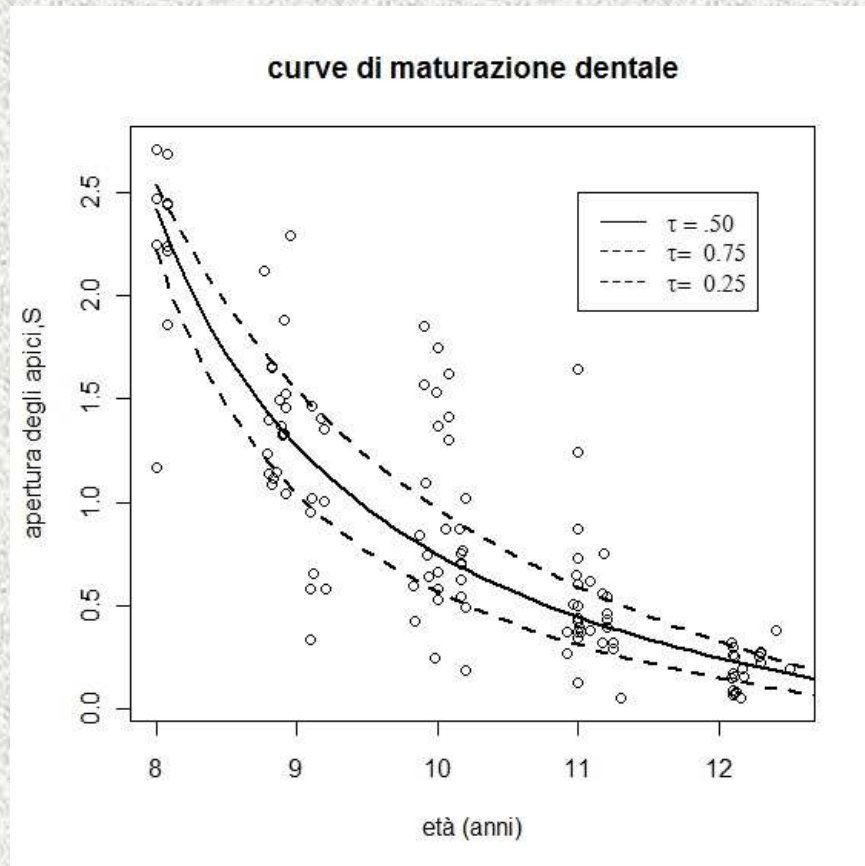
## Passo E

$$\begin{aligned}\tilde{Q}(\theta / \theta^{(t)}) &= \sum_{i=1}^N \frac{1}{m} \sum_{k=1}^m \ln L_c(\theta / \tau, S_i, v_{ik}) = \\ &= \sum_{i=1}^N \frac{1}{m} \sum_{k=1}^m \left[ \sum_{j=1}^n \ln f(S_{ij} / \tau, \sigma, v_{ik}, \beta) + \ln f(v_{ik} / \varphi) \right]\end{aligned}$$

I valori  $v_{ik}$  sono stati campionati mediante un algoritmo MCMC dalla distribuzione “a posteriori” delle  $u_i$ :

$$f(u_i / \tau, S_i, \theta)$$

# Risultati (curve quantiliche)



Modello:

$$g(E_{ij}, \beta) = \frac{\beta_1}{E_{ij} - \beta_2} - \beta_3$$

## Risultati: stima dei parametri

$\tau$		$\beta_1$	$\beta_2$	$\beta_3$	$\sigma$	$\varphi$
0.25	param	3.769	6.641	0.556	0.089	0.009
	se	1.168	0.534	0.173	5.51E-04	6.06E-04
0.5	param	5.316	6.288	0.685	0.121	0.007
	se	1.225	0.349	0.162	3.00E-04	1.01E-04
0.75	param	10.460	5.211	1.216	0.106	0.016
	se	4.088	0.821	0.420	3.16E-04	2.37E-04

---

# Conclusioni

- Per studiare la maturazione dentaria nei giovani è stato proposto un nuovo modello **NLQR con effetti casuali** che sicuramente è più flessibile del corrispondente NLME.
- Le stime dei parametri del modello di regressione quantilica sono state ottenute mediante la massimizzazione della verosimiglianza di v.a. con distribuzione di Laplace asimmetrica (AL).
- Il modello QRRE sembra adeguato a descrivere il processo di maturazione dentaria con l'età, anche se....
- **dati ancora insufficienti**
- **modello può essere migliorato**
- **pesante carico computazionale**

**Grazie per l'attenzione**