

1.2 Stimatori

Nella pratica spesso si assume che la forma di una qualche distribuzione di probabilità sia conosciuta a meno di un set di parametri incogniti $\theta_1, \dots, \theta_p$. Sia $x(u; \theta_1, \dots, \theta_p)$ la funzione dei quantili di una distribuzione con p parametri incogniti. In molte applicazioni tra i parametri incogniti si possono identificare un parametro di **posizione** ed un parametro di **scala**. Un parametro ξ di una distribuzione è un *parametro di posizione* se per la funzione dei quantili vale l'eguaglianza

$$(1.9) \quad x(u; \xi, \theta_2, \dots, \theta_p) = \xi + x(u; 0, \theta_2, \dots, \theta_p) .$$

Si dice, invece, che α è un *parametro di scala* della funzione dei quantili della distribuzione se

$$(1.10) \quad x(u; \alpha, \theta_2, \dots, \theta_p) = \alpha \times x(u; 1, \theta_2, \dots, \theta_p) .$$

Se per la distribuzione esistono entrambi questi parametri, allora vale l'eguaglianza

$$(1.11) \quad x(u; \xi, \alpha, \theta_3, \dots, \theta_p) = \xi + \alpha \times x(u; 0, 1, \theta_3, \dots, \theta_p) .$$

I parametri incogniti sono stimati a partire dai dati osservati. Dato un set di osservazioni, una funzione $\hat{\theta}$ di queste deve essere scelta come *stimatore* di θ . Lo stimatore $\hat{\theta}$ è a sua volta una variabile casuale ed ha una distribuzione di probabilità. La bontà di $\hat{\theta}$ come stimatore di θ tipicamente dipende da quanto $\hat{\theta}$ si avvicina a θ . La deviazione di $\hat{\theta}$ da θ può essere scomposta

in *distorsione* (tendenza di dare stime sistematicamente più alte, o più basse, del valore vero) e *variabilità* (deviazione casuale dal valore vero, che si verifica anche per gli stimatori che non presentano distorsione).

La performance di uno stimatore $\hat{\theta}$ può essere valutata con due misure, il *bias* (“distorsione”) e la *radice dell’errore quadratico medio* (RMSE), definite come

$$(1.12) \quad \text{bias}(\hat{\theta}) = E(\hat{\theta} - \theta) , \quad \text{RMSE}(\hat{\theta}) = \{E(\hat{\theta} - \theta)^2\}^{1/2} ,$$

e caratterizzate dall’aver la stessa unità di misura del parametro θ . Si dice che lo stimatore $\hat{\theta}$ è *indistorto* se $\text{bias}(\hat{\theta}) = 0$ ovvero se $E(\hat{\theta}) = \theta$. Diversi stimatori indistorti dello stesso parametro possono essere paragonati in termini della loro varianza: il rapporto $\text{var}(\hat{\theta}^{(1)})/\text{var}(\hat{\theta}^{(2)})$ si dice *efficienza* dello stimatore $\hat{\theta}^{(2)}$ rispetto allo stimatore $\hat{\theta}^{(1)}$. La radice dell’errore quadratico medio può essere anche scritta come

$$(1.13) \quad \text{RMSE}(\hat{\theta}) = [\{\text{bias}(\hat{\theta})\}^2 + \text{var}(\hat{\theta})]^{1/2} ,$$

da cui si vede come l’RMSE combina distorsione e variabilità di $\hat{\theta}$ e dà una misura globale dell’accuratezza della stima. Nei classici problemi di statistica in cui la stima dei parametri è basata su un campione di lunghezza n , sia il bias che la varianza di $\hat{\theta}$ sono asintoticamente proporzionali a n^{-1} per n grandi (v.es. ?), per cui l’RMSE di $\hat{\theta}$ è proporzionale a $n^{-1/2}$.

STIMA DEI PARAMETRI.

Esistono diversi metodi per la stima dei parametri:

- Metodo grafico di interpolazione su carta probabilistica.
- Metodo dei momenti.
- Metodo della massima verosimiglianza.
- Metodo degli L-Momenti.

METODO DEI MOMENTI.

Consiste semplicemente nell'eguagliare i momenti della popolazione ai momenti della serie.

Poiché i momenti della popolazione sono funzione dei parametri, basta considerare tanti momenti quanti sono i parametri da stimare e risolvere un sistema di r equazioni in r incognite (r : numero dei parametri da stimare) dando ovviamente preferenza ai momenti di ordine minore.

1. DISTRIBUZIONE NORMALE.

Due parametri da stimare μ e σ .

$$E[x] = \mu$$

$$\text{var}[x] = \sigma^2$$

A partire dai valori $(X_1, X_2, \dots, X_n)$ della serie si calcola:

$$\blacksquare \quad \bar{x} = \frac{\sum x_i}{n}$$

$$\blacksquare \quad s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Sistema di due equazioni in due incognite:

$$\hat{\mu} = \bar{x} \qquad \hat{\sigma}^2 = s^2$$

2. DISTRIBUZIONE LOG - NORMALE.

$$Y = \ln X$$

Due parametri: μ_y e σ_y

Due momenti della popolazione: μ_X e σ_X^2 :

$$\mu_Y = \ln \mu_X - \frac{1}{2} \sigma_Y^2$$

$$\sigma_Y^2 = \ln \left(1 + \frac{\sigma_X^2}{\mu_X^2} \right)$$

Dalla serie di dati $(X_1, X_2, \dots, X_n)$:

$$\bar{x} = \frac{\sum x_i}{n} = \hat{\mu}_X$$

$$s_X^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \hat{\sigma}_x^2$$

Sistema di due equazioni in due incognite:

$$\hat{\mu}_Y = \ln \bar{x} - \frac{1}{2} \hat{\sigma}_Y^2$$

$$\hat{\sigma}_y^2 = \ln \left(1 + \frac{s_x^2}{\bar{x}^2} \right)$$

DISTRIBUZIONE EV1 o DI GUMBEL.

Due parametri da stimare: ε ed α

Due momenti della popolazione:

$$E[x] = \mu = \varepsilon + \frac{0.57722}{\alpha}$$

$$\text{var}[x] = \sigma^2 = \frac{\pi^2}{6\alpha^2}$$

Dal sistema di 2 equazioni in 2 incognite:

$$\varepsilon = \mu - 0.450\sigma$$

$$\frac{1}{\alpha} = \frac{\sigma\sqrt{6}}{\pi}$$

METODO DELLA MASSIMA VEROSIMIGLIANZA.

Consiste nel determinare i valori dei parametri in modo da massimizzare la funzione di verosimiglianza:

$$V(x_1, x_2, \dots, x_n; \vartheta_1, \vartheta_2, \dots, \vartheta_r) = f_x(x_1)f_x(x_2) \dots f_x(x_n)$$

che è la funzione di densità congiunta del campione per osservazioni indipendenti.

La stima dei parametri $\vartheta_1, \vartheta_2, \dots, \vartheta_r$ si ottiene dal sistema di r equazioni in r incognite.

$$\frac{\partial V}{\partial \vartheta_i}(x_1, x_2, \dots, x_n; \vartheta_1, \vartheta_2, \dots, \vartheta_r) = 0 \quad i = 1, 2, \dots, r$$

Stima di Massima Verosimiglianza: Distribuzione normale



$$V(x_1, x_2, \dots, x_n; \theta_1, \theta_2) = \frac{1}{(2\pi\theta_2)^{\frac{n}{2}}} \exp \left[-\frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \theta_1}{\sqrt{\theta_2}} \right)^2 \right]$$

$$\ln V = -\ln [2\pi\theta_2]^{\frac{n}{2}} - \frac{1}{2\theta_2} \sum_{i=1}^n (x_i - \theta_1)^2$$

$$\max(V) \Rightarrow \max(\ln V) \Rightarrow \min(-\ln V)$$

$$\frac{\delta}{\delta \theta_1} \left[\frac{1}{2\theta_2} \sum_{i=1}^n (x_i - \theta_1)^2 + \ln(2\pi\theta_2)^{\frac{n}{2}} \right] = - \left[\frac{1}{\theta_2} \sum_{i=1}^n (x_i - \theta_1) \right] = 0 \quad \text{verificata se } \theta_1 = \mu$$

$$\frac{\delta}{\delta \theta_2} \left[\frac{1}{2\theta_2} \sum_{i=1}^n (x_i - \theta_1)^2 + \ln(2\pi\theta_2)^{\frac{n}{2}} \right] = -\frac{1}{2\theta_2^2} \sum_{i=1}^n (x_i - \theta_1)^2 + \frac{n}{2\theta_2} = 0 \quad \text{verificata se } \theta_2 = \sigma^2$$

Stima di Massima Verosimiglianza: Distribuzione di Gumbel

$$\theta_2 = \bar{x} - \frac{\sum_{i=1}^n x_i \cdot e^{-\frac{1}{\theta_2} x_i}}{\sum_{i=1}^n e^{-\frac{1}{\theta_2} x_i}}$$

$$e^{-\frac{\theta_1}{\theta_2}} = \frac{1}{n} \cdot \sum_{i=1}^n e^{-\frac{1}{\theta_2} x_i}$$

con n =numero di dati del campione.

Il valore a primo membro è determinato sulla base di un valore di tentativo assegnato al secondo membro. Il procedimento si arresta quando i due valori differiscono di una quantità inferiore alla tolleranza assegnata (per es. 0.001 per θ_2).

Il primo valore di tentativo potrà essere quello derivante dalla stima dei momenti.

La correzione che si può apportare ad ogni tentativo per accelerare la convergenza può essere quella suggerita da Gumbel: $\theta_2(3) = \theta_2(2) + (\theta_2(1) - \theta_2(2))/3$

L-momenti

Gli **L-momenti** sono un modo alternativo di descrivere la forma delle distribuzioni di probabilità.

Derivano dai *probability weighted moments* (PWM) o momenti pesati in probabilità.

Per una variabile casuale x con funzione di probabilità cumulata $F(x)$ si ha

$$M_{p,r,s} = E [x^p \{F(x)\}^r \{1 - F(x)\}^s]$$

L-momenti vs momenti

$$\beta_r = M_{1,r,0} = E [x \{F(x)\}^r]$$

$$M_r = E [(x - \mu(x))^r]$$

$$\int_{-\inf}^{\inf} x \{F(x)\}^r \cdot f(x) dx$$

$$\int_{-\inf}^{\inf} (x - \mu(x))^r \cdot f(x) dx$$

PWM $\rightarrow$ lineari nella variabile casuale (la x non è mai elevata a potenza, ma lo è solo $F(x)$)

Momenti $\rightarrow$ la x è elevata a potenza per $r > 1$

L-momenti

I PWM possono essere usati per stimare i parametri di una distribuzione, ma sono difficili da interpretare direttamente come misure di dispersione, forma, ...

Tali informazioni sono però presenti in particolari combinazioni di PWM

$$\lambda_1 = \beta_0 \rightarrow \text{media}$$

$$\lambda_2 = 2\beta_1 - \beta_0 \rightarrow \text{dispersione}$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \rightarrow \text{asimmetria}$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$$

Coefficienti adimensionali

$$\tau = \frac{\lambda_2}{\lambda_1} \rightarrow \text{L-CV}$$

$$\tau_3 = \frac{\lambda_3}{\lambda_2} \rightarrow \text{L-skewness}$$

$$\tau_4 = \frac{\lambda_4}{\lambda_2} \rightarrow \text{L-kurtosis}$$

in generale $\tau_r = \frac{\lambda_r}{\lambda_1}$

Proprietà

esistenza se la media della distribuzione esiste, allora esistono tutti gli L-momenti

unicità se la media della distribuzione esiste, allora gli L-momenti definiscono in maniera univoca la distribuzione (non ci sono due distribuzioni con gli stessi L-momenti)

- ▶ $\lambda_2 \geq 0$
- ▶ $|\tau_r| < 1$ per tutti $r \geq 3$
- ▶ $\frac{1}{4}(5\tau_3^2 - 1) \leq \tau_4 < 1$
- ▶ per distribuzioni che assumono solo valori positivi τ_3 è limitato in funzione di τ $2\tau - 1 \leq \tau_3 < 1$



L-momenti

Notazione:

$x_1, x_2, \dots, x_i \dots x_n$ campione

$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(i)} \leq \dots \leq x_{(n)}$ campione ordinato in
senso crescente

$x_{1:n} \leq x_{2:n} \leq \dots \leq x_{i:n} \leq \dots \leq x_{n:n}$ campione ordinato in
senso crescente

PWM campionari

$$b_0 = \frac{1}{n} \sum_{j=1}^n x_{j:n}$$

$$b_1 = \frac{1}{n} \sum_{j=2}^n \frac{(j-1)}{(n-1)} x_{j:n}$$

$$b_2 = \frac{1}{n} \sum_{j=3}^n \frac{(j-1)(j-2)}{(n-1)(n-2)} x_{j:n}$$

$$b_r = \frac{1}{n} \sum_{j=r+1}^n \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} x_{j:n}$$

L-momenti campionari

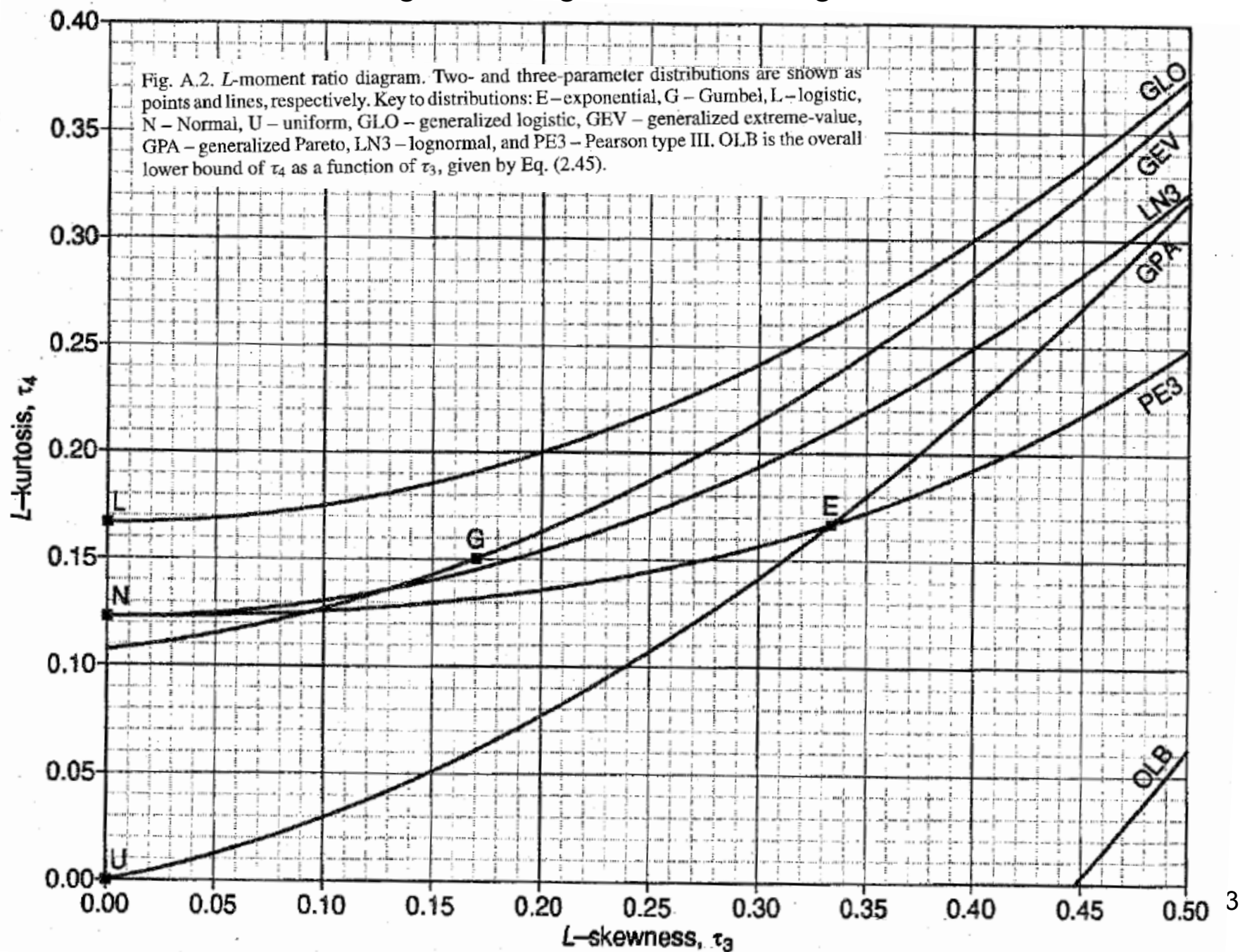
$$l_1 = b_0$$

$$l_2 = 2b_1 - b_0$$

$$l_3 = 6b_2 - 6b_1 + b_0$$

$$l_4 = 20b_3 - 30b_2 + 12b_1 - b_0$$

Diagramma diagnostico di Hosking e Wallis



Stima dei parametri di alcune distribuzioni mediante gli L-momenti

La procedura per la stima dei parametri è analoga a quella utilizzata nel metodo dei momenti: gli L-momenti teorici vengono equiparati a quelli campionari. Invertendo la relazione tra parametri e L-momenti si ottengono i parametri.

Distribuzione Normale

$$f(x) = \frac{1}{\theta_2 \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\theta_1}{\theta_2} \right)^2}$$

$$x(F) = \theta_1 + \theta_2 \Phi^{-1}(F)$$

L-momenti

Parametri

$$\lambda_1 = \theta_1$$

$$\lambda_2 = 0,5642\theta_2 \quad \theta_1 = \lambda_1$$

$$\tau_3 = 0 \quad \theta_2 = \pi^{1/2} \lambda_2$$

$$\tau_4 = 0,1226$$

Distribuzione di Gumbel

$$F(x) = e^{-e^{-\frac{x-\theta_1}{\theta_2}}}$$

$$f(x) = \frac{1}{\theta_2} e^{-\frac{x-\theta_1}{\theta_2}} e^{-e^{-\frac{x-\theta_1}{\theta_2}}}$$

$$x(F) = \theta_1 - \theta_2 \ln[-\ln(F)]$$

L-momenti

Parametri

$$\lambda_1 = \theta_1 + 0,5772\theta_2$$

$$\lambda_2 = \theta_2 \log 2 \quad \theta_1 = \lambda_1 - 0,5772(\lambda_2 / \log 2)$$

$$\tau_3 = 0,1699 \quad \theta_2 = \lambda_2 / \log 2$$

$$\tau_4 = 0,1504$$

Stima dei parametri di alcune distribuzioni mediante gli L-momenti

Distribuzione Log Normale

$$f_X(x) = \frac{1}{x\theta_2\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\ln x - \theta_1}{\theta_2} \right)^2 \right]$$

$$x(F) = \exp [\theta_1 + \theta_2 \Phi^{-1}(F)] \quad \theta_2 > 0$$

L-momenti

$$L_1 = \exp(\theta_1 + \theta_2^2/2)$$

$$L_2 = e^{\theta_1 + \theta_2^2/2} [2\Phi(\theta_2/\sqrt{2}) - 1]$$

Parametri

$$\hat{\theta}_1 = \ln l_1 - \theta_2^2/2$$

$$\hat{\theta}_2 = \sqrt{2}\Phi^{-1} \left(\frac{1+l_2/l_1}{2} \right)$$

Distribuzione GEV

$$F(x) = \exp \left[- \left(1 - \frac{\theta_3}{\theta_2} (x - \theta_1) \right)^{1/\theta_3} \right]$$

$$x(F) = \begin{cases} \theta_1 + \theta_2 [1 - (-\ln F)^{\theta_3}] / \theta_3, & \theta_3 \neq 0 \\ \theta_1 - \theta_2 \ln(-\ln F), & \theta_3 = 0 \end{cases}$$

L-momenti

$$L_1 = \theta_1 + \theta_2 [1 - \Gamma(1 + \theta_3)] / \theta_3$$

$$L_2 = \theta_2 (1 - 2^{-\theta_3}) \Gamma(1 + \theta_3) / \theta_3$$

$$\tau_3 = 2(1 - 3^{-\theta_3}) / (1 - 2^{-\theta_3}) - 3$$

$$\tau_4 = \frac{5(1-4^{-\theta_3})-10(1-3^{-\theta_3})+6(1-2^{-\theta_3})}{(1-2^{-\theta_3})}$$

Parametri

$$\hat{\theta}_3 \simeq 7.8590c + 2.9554c^2$$

$$c = \frac{2}{3+l_3/l_2} - \frac{\ln 2}{\ln 3}$$

$$\hat{\theta}_2 = \frac{l_2 \hat{\theta}_3}{(1-2^{-\hat{\theta}_3}) \Gamma(1+\hat{\theta}_3)}$$

$$\hat{\theta}_1 = l_1 - \frac{\hat{\theta}_2}{\hat{\theta}_3} \left(1 - \Gamma(1 + \hat{\theta}_3) \right)$$