PRO MATHEMATICA: *Vol. X, Nos. 19-20, 1996*

UN ESTIMADOR DE SEUDO-MAXIMA VEROSIMILITUD, SU CONSISTENCIA FUERTE Y DISTRIBUCION ASINTOTICA

José **Flores**

1. Introducción

Uno de los métodos más conocidos para estimar parámetros es el de Máxima Verosimilitud, con él y bajo condiciones de "regularidad" se obtienen estimadores fuertemente consistentes y asintóticamente eficientes. Sin embargo, no siempre es posible obtener una fórmula explícita para tales estimadores, teniéndose que recurrir a métodos computacionales para obtener estimativas a partir de observaciones provenientes de una muestra aleatoria; aun así, puede presentarse el problema de una convergencia demasiado lenta. De ahí, tratando de superar estas dificultades, han aparecido modificaciones de este método, una de ellas es debida a Gong y Samaniego (1981) quienes definen un estimador de seudo-máxima verosimilitud y dan condiciones de regularidad bajo las cuales este estimador es débilmente consistente y asintóticamente normal.

➩ *Universidad de São Paulo, Brasil.*

El objetivo de este artículo es el de, bajo condiciones de regularidad similares, establecer la consistencia fuerte y derivar la distribución asintótica de tal estimador de seudo-máxima veosimilitud. Para ello seguiremos la idea dada en Parke (1983), pero con un sustento diferente, el Teorema de la función implícita. Consideraremos únicamente el caso uniparamétrico, aunque los resultados pueden extenderse fácilmente al caso multi-paramétrico.

2. Condiciones de Regularidad

Sea $\mathcal{F} = \{F_{\theta,\pi}\}_{(\theta,\pi)\in\Theta\times\Pi}$ una familia de distribuciones biparamétricas definidas sobre el conjunto de los números reales, $\mathbb{R}$, donde Θ y Π son dos subconjuntos de $\mathbb{R}$. Consideraremos a θ como parámetro estrucutral (o de interés) y a π como "nuissance" (o de incomodidad).

Las condiciones de regularidad que impondremos a la familia $\mathcal{F}$ son las siguientes:

($\mathcal{C}$ 1) Para cada par $(\theta,\pi) \in \Theta \times \Pi$, $F_{\theta,\pi}$ posee una función de densidad (o de probabilidad) que será denotada por $f(.|\,\theta,\pi)$ y cuyo soporte será el espacio muestral $\mathrm{X} \subseteq \mathbb{R}$.

($\mathcal{C}$ 2) Para cada $x \in \mathrm{X}$, la función $ln\, f(x\,|\,.\,,\,.)$, definida sobre $\Theta\times\Pi$, tiene derivada parcial con respecto a θ, $\frac{\partial}{\partial\theta}\, ln\, f(x\,|\,\theta,\pi)$, la cual es de la clase $\mathcal{C}_2$ sobre $\Theta\times\Pi$.

($\mathcal{C}$ 3) Para cada par $(\theta_1, \pi_1) \in \Theta\times\Pi$, existen funciones, no negativas y definidas sobre X, M_1, M_2 y M_3, tales que si $X : \Omega \to \mathbb{R}$ es una variable aleatoria con $X(\Omega) = \mathrm{X}$ y distribución F_{θ_1,π_1}; entonces $\forall i \in \{1,2,3\}$, $E_{\theta_1,\pi_1}[M_1(X)] \in \mathbb{R}_+$ y, además, existen bolas abiertas de radios finitos $A_{\theta_1} \subseteq \Theta$ y $B_{\pi_1} \subseteq \Pi$ de centros θ_1 y π_1, respectivamente, de manera que $\forall x \in \mathrm{X}$:

(i) $\forall(\theta,\pi) \in A_{\theta_1}\times B_{\pi_1}$, $\left|\frac{\partial^3}{\partial\theta^3}\, ln\, f(x\,|\,\theta,\pi)\right| \leq M_1(x)$;

(ii) $\forall \pi \in B_{\pi_1}$, $\left| \dfrac{\partial^3}{\partial \pi \, \partial \theta^3} \ln f(x \mid \theta_1, \pi) \right| \le M_2(x)$; y

(iii) $\forall \pi \in B_{\pi_1}$, $\left| \dfrac{\partial^3}{\partial \pi^2 \, \partial \theta} \ln f(x \mid \theta_1, \pi) \right| \le M_3(x)$.

($\mathcal{C}$4) Para cada par $(\theta, \pi) \in \Theta \times \Pi$, si $X: \Omega \to \mathbb{R}$ es una variable aleatoria con $X(\Omega) = \mathrm{X}$ y distribución $F_{\theta,\pi}$. Si definimos $I_\theta = E_{\theta,\pi}[\frac{\partial}{\partial \theta} \ln f(X \mid \theta, \pi)]^2$ e $I_{\theta,\pi} = E_{\theta,\pi}[\frac{\partial}{\partial \theta} \ln f(X \mid \theta, \pi)]$; entonces

(i) $E_{\theta,\pi}[\frac{\partial}{\partial \theta} \ln f(x \mid \theta, \pi)] = 0$;

(ii) $I_\theta \in \mathbb{R}_+$ e $I_\theta = - E_{\theta,\pi}[\frac{\partial}{\partial \theta^2} \ln f(x \mid \theta, \pi)]$ e

(iii) $I_{\theta,\pi} \in \mathbb{R}$ e $I_{\theta,\pi} = - E_{\theta,\pi}[\frac{\partial}{\partial \theta} \ln f(X \mid \theta, \pi)]$.

Observación: Estas condiciones son parte de unas condiciones de regularidad frecuentemente impuestas a una familia de distribuciones biparamétricas, para demostrar la consistencia y normalidad asintótica de los estimadores de máxima verosimilitud, véase, por ejemplo, Serfling (1980).

3. Notación A partir de ahora $\underline{x}_n = (X_1,\dots,X_n)$ con $n \in \mathbb{N}_+$, será una muestra aleatoria de alguna distribución perteneciente a la familia $\mathcal{F}$, esto es, $\forall i \in \{1,\dots,n\}$, $X_i : \Omega \to \mathbb{R}$ es una variable aleatoria con $X_i(\Omega) = \mathrm{X}$ y distribución F_{θ_0,π_0}, donde $(\theta_0,\pi_0) \in \Theta \times \Pi$ serán los verdaderos valores de los parámetros; y $X_1,\dots,X_n$ son independientes.

Para todo par $(\theta,\pi) \in \Theta \times \Pi$, definimos:

$\mathrm{L}_n(\theta,\pi; \underline{x}_n) = \sum_{j=1}^{n} \ln f(X_j / \theta, \pi)$, de modo que

$\forall \omega \in \Omega$, $Ln(\theta,\pi; \underline{x}_n(\omega)) = \sum_{j=1}^{n} \ln f(X_j(\omega) / \theta, \pi)$,

$x_n(\omega) = (X_1(\omega),\dots, X_n(\omega))$ y

$\overline{L}n(\theta, \pi; \underline{x}_n) = \dfrac{1}{n} Ln(\theta, \pi; \underline{x}_n)$.

4. Definición *Sea* $\underline{x}_n = (X_1,...,X_n)$ *una muestra aleatoria como en* 3 *y* $\hat{\pi}_n(\underline{x}_n)$ *un estimador del parámetro* π (*asumimos que* $\forall\omega\in\Omega, \hat{\pi}_n(\underline{x}_n(\omega))\in\prod$).

Si $\hat{\theta}_n(\hat{\pi}_n(\underline{x}_n)) = \hat{\theta}_n(\underline{x}_n) = \max_{\theta\in\Theta}\prod_{j=1}^{n} f(X_j|\theta, \hat{\pi}_n(\underline{x}_n))$; *entonces, decimos que* $\hat{\theta}_n(\underline{x}_n)$ *es un estimador de seudo-máxima verosimilitud.*

Obsérvese que $\hat{\theta}_n(\underline{x}_n) = \max_{\theta\in\Theta}\hat{\mathcal{L}}_n(\theta, \hat{\pi}_n(\underline{x}_n); \underline{X}_n)$.

5. Resultados Asintóticos

Lema 1. *Sea* $\underline{x}_n$ *una muestra aleatoria como en 3; entonces, bajo las condiciones de regularidad* ($\mathcal{C}$ 1), ($\mathcal{C}$ 2) y ($\mathcal{C}$ 4), *se tiene que con probabilidad* 1-(θ_0, π_0):

(i) $\dfrac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\theta_0, \pi_0; \underline{x}_n) \underset{n\uparrow\infty}{\rightarrow} 0$

(ii) $\dfrac{\partial^2}{\partial\theta^2}\overline{\mathcal{L}}_n(\theta_0, \pi_0; \underline{x}_n) \underset{n\uparrow\infty}{\rightarrow} I_{\theta_0}$ y

(iii) $\dfrac{\partial^2}{\partial\pi\partial\theta^2}\overline{\mathcal{L}}_n(\theta_0, \pi_0; \underline{x}_n) \underset{n\uparrow\infty}{\rightarrow} -I_{\theta_0,\pi_0}$.

Demostración:

(i) Por las condiciones de regularidad ($\mathcal{C}$ 1) y ($\mathcal{C}$ 2) $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\theta_0, \pi_0; \underline{x}_n) = \frac{1}{n}\sum_{j=1}^{n}\frac{\partial}{\partial\theta}\ln f(X_j|\theta_0, \pi_0)$, donde $\frac{\partial}{\partial\theta}\ln f(X_1|\theta_0, \pi_0),...,\frac{\partial}{\partial\theta}\ln f(X_n|\theta_0, \pi_0)$ son n variables aleatorias independientes e identicamente distribuídas y, según la condición de regularidad ($\mathcal{C}$4)-(i), con esperanza $E_{\theta_0,\pi_0}[\frac{\partial}{\partial\theta}\ln f(X_1|\theta_0, \pi_0)] = 0$. Por tanto, el resultado (i) sigue de la ley fuerte de los grandes números.

Análogamente se pueden demostrar (ii) y (iii).

Lema 2. *Sea* $\underline{x}_n = (X_1,...,X_n)$ *una muestra aleatoria como en 3. Sean* $\hat{\theta}_n(\underline{x}_n)$ *y* $\hat{\pi}_n(\underline{x}_n)$ *tales que, con probabilidad* 1-(θ_0, π_0),

$(\hat{\theta}_n(\underline{x}_n),\hat{\pi}_n(\underline{x}_n)) \underset{n\uparrow\infty}{\rightarrow} (\theta_0,\pi_0)$. *Entonces, bajo las condiciones de regularidad* $(\mathscr{C}1)$ - $(\mathscr{C}4)$, *se tiene que con probabilidad* 1-(θ_0, π_0) :

(i) $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\hat{\theta}_n(\underline{x}_n),\hat{\pi}_n(\underline{x}_n);\underline{x}_n) \underset{n\uparrow\infty}{\rightarrow} 0$

(ii) $\frac{\partial^2}{\partial\theta^2}\overline{\mathcal{L}}_n(\hat{\theta}_n(\underline{x}_n),\hat{\pi}_n(\underline{x}_n);\underline{x}_n) \underset{n\uparrow\infty}{\rightarrow} -I_{\theta_0}$ y

(iii) $\frac{\partial^2}{\partial\pi\partial\theta^2}\overline{\mathcal{L}}_n(\hat{\theta}_n(\underline{x}_n),\hat{\pi}_n(\underline{x}_n);\underline{x}_n) \underset{n\uparrow\infty}{\rightarrow} -I_{\theta_0,\pi_0}$.

Demostración: Demostraremos (i), análogamente se pueden demostrar (ii) y (iii).

Sean A_{θ_0} y B_{π_0} como en la condición de regularidad $(\mathscr{C}3)$ y sea $C = \{\omega\in\Omega: ((\hat{\theta}_n(\underline{x}_n)),\hat{\pi}_n(\underline{x}_n(\omega)) \underset{n\uparrow\infty}{\rightarrow} -(\theta_0,\pi_0)\}$. Entonces, $P_{\theta_0,\pi_0}(C)=1$ y además, como $(\theta_0,\pi_0)\in A_{\theta_0}\times B_{\pi_0}$, tenemos que $\forall\omega\in C$, $\exists n_\omega \in \mathbb{N}_+$ tal que $\forall n \geq n_\omega$, $(\hat{\theta}_n(\underline{x}_n(\omega)),\hat{\pi}_n(\underline{x}_n(\omega)))\in A_{\theta_0}\times B_{\pi_0}$. Así, la condición de regularidad $(\mathscr{C}2)$ nos permite usar la fórmula de Taylor a la función $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\,\cdot\,,\hat{\pi}_n(\underline{x}_n(\omega));\underline{x}_n(\omega))$, primero, y después a las funciones $\frac{\partial}{\partial\theta^2}\overline{\mathcal{L}}_n(\theta_0,\,\cdot\,;\underline{x}_n(\omega))$ y $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\theta_0,\cdot;\underline{x}_n(\omega))$; de manera que $\forall\omega\in C$, $\forall n\geq n_\omega$:

$$
\begin{aligned}
&\tfrac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\hat{\theta}_n(\underline{x}_n(\omega)),\hat{\pi}_n(\underline{x}_n(\omega));\underline{x}_n(\omega)) = \tfrac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\theta_0,\pi_0;\underline{x}_n(\omega))+\\
&+(\hat{\pi}_n(\underline{x}_n(\omega))-\pi_0)\tfrac{\partial^2}{\partial\pi\partial\theta}\overline{\mathcal{L}}_n(\theta_0,\pi_0;\underline{x}_n(\omega))\\
&+\tfrac{1}{2}(\hat{\pi}_n(\underline{x}_n(\omega))-\pi_0)\tfrac{\partial^3}{\partial\pi^2\partial\theta}\overline{\mathcal{L}}_n(\theta_0,\tilde{\pi}_n(\underline{x}_n(\omega));\underline{x}_n(\omega))\\
&+(\hat{\theta}_n(\underline{x}_n(\omega))-\theta_0)[\tfrac{\partial^2}{\partial\theta^2}\overline{\mathcal{L}}_n(\theta_0,\pi_0;\underline{x}_n(\omega))+\tfrac{1}{2}(\hat{\pi}_n(\underline{x}_n(\omega))-\pi_0)\tfrac{\partial^3}{\partial\pi\partial\theta^2}\\
&\qquad\overline{\mathcal{L}}_n(\theta_0,\pi^*_n(\underline{x}_n(\omega));\underline{x}_n(\omega))]\\
&+\tfrac{1}{2}(\hat{\theta}_n(\underline{x}_n(\omega))-\theta_0)^2\tfrac{\partial^3}{\partial\theta^3}\overline{\mathcal{L}}_n(\tilde{\theta}_n(\underline{x}_n(\omega)),\hat{\pi}_n(\underline{x}_n(\omega)))
\end{aligned}
\qquad \textbf{(1)}
$$

donde $\tilde{\pi}_n(\underline{x}_n(\omega))$ y $\overset{*}{\pi}_n(\underline{x}_n(\omega))$ están entre $\hat{\pi}_n(\underline{x}_n(\omega))$ y π_0; y $\tilde{\theta}_n(\underline{x}_n(\omega))$ está entre $\hat{\theta}_n(\underline{x}_n(\omega))$ y θ_0.

Por otra parte, por la condición de regularidad ($\mathscr{C}$3)-(i) y la ley fuerte de los grandes números, tenemos que si $D_1 = \{\omega \in \Omega: \frac{1}{n}\sum_{j=1}^{n} M_1(X_j(\omega)) \underset{n\uparrow\infty}{\rightarrow} E_{\theta_0,\pi_0}[M_1(X_1)]\}$, entonces $P_{\theta_0,\pi_0}(D_1) = 1$, y además, $\forall\omega\in C\cap D$, $\exists n_1(\omega) \in \mathbb{N}_+$ tal que $n_1(\omega) \geq n(\omega)$ y $\forall n \geq n_1(\omega)$:

$$\left|\frac{\partial^3}{\partial\theta^3}\overline{\mathscr{L}}_n(\tilde{\theta}_n(\underline{x}_n(\omega)), \hat{\pi}_n(\underline{x}_n(\omega)); \underline{x}_n(\omega))\right| \leq \frac{1}{n}\sum_{j=1}^{n} M_1(X_j(\omega)) < \tfrac{3}{2}E_{\theta_0,\pi_0}[M_1(X_1)] \qquad \textbf{(2)}$$

Análogamente, si $D_2 = \{\omega \in \Omega: \frac{1}{n}\sum_{j=1}^{n} M_2(X_j(\omega)) \underset{n\uparrow\infty}{\rightarrow} E_{\theta_0,\pi_0}[M_2(X_1)]\}$ y $D_3 = \{\omega \in \Omega: \frac{1}{n}\sum_{j=1}^{n} M_3(X_j(\omega)) \underset{n\uparrow\infty}{\rightarrow} E_{\theta_0,\pi_0}[M_3(X_1)]\}$; entonces $\forall\omega\in C\cap D_2$, $\exists n_2(\omega) \geq n(\omega)$ tal que $\forall n \geq n_2(\omega)$:

$$\left|\frac{\partial^3}{\partial\pi\partial\theta^2}\overline{\mathscr{L}}_n(\theta_0, \overset{*}{\pi}_n(\underline{x}_n(\omega)))\right| < \tfrac{3}{2}E_{\theta_0,\pi_0}[M_2(X_1)] \qquad \textbf{(3)}$$

y $\forall\omega\in C\cap D_3$, $\exists n_3(\omega) \geq n(\omega)$ tal que $\forall n \geq n_3(\omega)$:

$$\left|\frac{\partial^3}{\partial\pi^2\partial\theta}\overline{\mathscr{L}}_n(\tilde{\theta}_n(\underline{x}_n(\omega)), \hat{\pi}_n(\underline{x}_n(\omega)); \underline{x}_n(\omega))\right| < \tfrac{3}{2}E_{\theta_0,\pi_0}[M_3(X_1)] \qquad \textbf{(4)}$$

Luego, usando (1), (2), (3), (4), (i) del lema anterior y el hecho que $\forall\omega\in C\cap D_1 \cap D_2 \cap D_3$, $\hat{\theta}_n(\underline{x}_n(\omega)) - \theta_0 \underset{n\uparrow\infty}{\rightarrow} 0$ y $\hat{\pi}_n\underline{x}_n(\omega) - \pi_0 \underset{n\uparrow\infty}{\rightarrow} 0$, se deduce que $\forall\omega\in C\cap D_1 \cap D_2 \cap D_3$, $\frac{\partial}{\partial\theta}\overline{\mathscr{L}}_n(\hat{\theta}_n(\underline{x}_n(\omega)), \hat{\pi}_n(\underline{x}_n(\omega))) \underset{n\uparrow\infty}{\rightarrow} 0$. Por tanto, como $P_{\theta_0,\pi_0}(C\cap D_1\cap D_2\cap D_3) = 1$, se tiene demostrado el resultado.

Teorema 1. *Sean* $\underline{X}_n = (X_1,...,X_n)$ *una muestra aleatoria de* F_{θ_0,π_0}, *como en 3, y* $\hat{\pi}_n(\underline{x}_n)$ *un estimador fuertemente consistente de* π_0, *esto es* $P_{\theta_0,\pi_0}(\hat{\pi}_n(\underline{x}_n) \underset{n\uparrow\infty}{\to} \pi_0) = 1$. *Y supongamos que para cada* $\omega\in\Omega$ y $n\in\mathbb{N}_+$, *la ecuación en relación a* θ, $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\theta,\hat{\pi}_n(\underline{x}_n(\omega));\underline{x}_n(\omega)) = 0$ *tenga una sola raíz. Entonces, bajo las condiciones de regularidad* ($\mathcal{C}1$) - ($\mathcal{C}4$) *el estimador de seudo-máxima verosimilitud,* $\hat{\theta}_n(\underline{x}_n)$, *es fuertemente consistente para* θ_0.

Demostración: Las condiciones de regularidad nos permiten garantizar que existe $\tilde{\theta}_n(\underline{x}_n) = \tilde{\theta}_n(\underline{x}_n;\pi_0)$ tal que, con probabilidad 1-(θ_0,π_0), $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\tilde{\theta}_n(\underline{x}_n),\pi_0;\underline{x}_n) = 0$ y $\tilde{\theta}_n(\underline{x}_n) \underset{n\uparrow\infty}{\to} \theta_0$; vease, por ejemplo, Serfling (1980).

Así tenemos que si $C = \{\omega\in\Omega : \frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\tilde{\theta}_n(\underline{x}_n(\omega)),\pi_0;\underline{x}_n(\omega)) = 0$ y $\tilde{\theta}_n(\underline{x}_n(\omega)) \underset{n\uparrow\infty}{\to} \theta_0\}$, entonces $P_{\theta_0,\pi_0}(C) = 1$; y de allí, por (ii) del Lema 2, si $D = \{\omega\in\Omega: \frac{\partial^2}{\partial\theta^2}\overline{\mathcal{L}}_n(\tilde{\theta}_n(\underline{x}_n(\omega)),\pi_0;\underline{x}_n(\omega)) \to -I_{\theta_0}\}$, entonces $P_{\theta_0,\pi_0}(D) = 1$.

Sea $\varepsilon > 0$, arbitrariamente elegido. Entonces, $\forall\omega\in C\cap D$, $\exists n_\omega\in\mathbb{N}_+$ tal que $\forall n \geq n_\omega$, $|\tilde{\theta}_n(\underline{x}_n(\omega)) - \theta_0| < \varepsilon/2$ y, recordando que $I_{\theta_0} > 0$, también $\frac{\partial^2}{\partial\theta^2}\overline{\mathcal{L}}_n(\tilde{\theta}_n(\underline{x}_n(\omega)),\pi_0;\underline{x}_n(\omega)) < 0$. Así tenemos que $\forall\omega\in C\cap D$, $\forall n \geq n_\omega$:
La función $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\cdot,\cdot,\underline{x}_n(\omega))$ es de clase $\mathcal{C}1$ sobre $B(\tilde{\theta}_n(\underline{x}_n(\omega)),\varepsilon/2)\times B(\pi_0,\varepsilon/2)$, $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\tilde{\theta}_n(\underline{x}_n(\omega)),\pi_0;\underline{x}_n(\omega)) = 0$ y $\frac{\partial^2}{\partial\theta^2}\overline{\mathcal{L}}_n(\tilde{\theta}_n(\underline{x}_n(\omega)),\pi_0;\underline{x}_n(\omega)) < 0$.

Por tanto, por el Teorema de la función implícita, para cada $\omega\in C\cap D$, se tiene para cada $n \geq n_\omega$ que:

Existen vecindades $U_{\omega,n}$ y $V_{\omega,n}$, tales que $\tilde{\theta}_n(\underline{x}_n(\omega)) \in U_{\omega,n} \subseteq B(\tilde{\theta}_n(\underline{x}_n(\omega)), \varepsilon/2)$ y $\pi_0 \in V_{\omega,n} \subseteq B(\pi_0,\varepsilon/2)$; y existe una única función $g_{\omega,n}$, de clase $\mathcal{C}1$ sobre $V_{\omega,n}$, tal que $g_{\omega,n}(\pi_0) = \tilde{\theta}_n(\underline{x}_n(\omega))$ y $\forall\pi \in V_{\omega,n}$, $g_{\omega,n}(\pi)\in U_{\omega,n}$ y $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(g_{\omega,n}(\pi),\pi;\underline{x}_n(\omega)) = 0$.

Y como también se tiene, por hipótesis, que si $E=\{\omega\in\Omega: \hat{\pi}_n(\underline{x}_n(\omega)) \underset{n\uparrow\infty}{\to} \pi_0\}$, entonces $P_{\theta_0,\pi_0}(E)=1$; se sigue entonces que, para cada $\omega\in C\cap D\cap E$, existe $m_\omega \in \mathbb{N}_+$ $m_\omega \geq n_\omega$ tal que para cada $n \geq m_\omega$, $g_{\omega,n}(\hat{\pi}_n(\underline{x}_n(\omega))) \in U_{\omega,n}$ y $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(g_{\omega,n}(\hat{\pi}_n(\underline{x}_n(\omega))),\underline{x}_n(\omega))=0$.

De allí, como $\forall\omega\in C\cap D\cap E$, $\forall n\in\mathbb{N}_+$, $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\hat{\theta}_n(\underline{x}_n(\omega)),\hat{\pi}_n(\underline{x}_n(\omega)); \underline{x}_n(\omega))=0$ y la ecuación, en relación a θ, $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\theta,\hat{\pi}_n(\underline{x}_n(\omega));\underline{x}_n(\omega))=0$ tiene una sola raíz, debe tenerse para cada $\omega\in C\cap D\cap E$ y para cada $n \geq m_\omega$ que $\hat{\theta}_n(\underline{x}_n(\omega))=g_{\omega,n}(\hat{\pi}_n(\underline{x}_n(\omega)))$, y por lo tanto también, $\hat{\theta}_n(\underline{x}_n(\omega))\in U_{\omega,n}\subseteq B(\tilde{\theta}_n(\underline{x}_n(\omega)),\varepsilon/2)$.

Luego, recordando que también para cada $\omega\cap C\cap D\cap E$ y para cada $n\geq m_\omega$; $|\tilde{\theta}_n(\underline{x}_n(\omega))-\theta_0|<\varepsilon/2$; se sigue que $\forall\omega\in C\cap D\cap E$, $\forall n \geq m_\omega$: $|\hat{\theta}_n(\underline{x}_n(\omega))-\theta_0|<\varepsilon$. Así, como $\varepsilon > 0$ fue elegido arbitrariamente, se tiene que $\forall\omega\in C\cap D\cap E$, $\hat{\theta}_n(\underline{x}_n(\omega)) \underset{n\uparrow\infty}{\to} \theta_0$ y como $P_{\theta_0,\pi_0}(C\cap D\cap E)=1$, $\hat{\theta}_n(\underline{x}_n)$ es fuertemente consistente para θ_0.

Observación. Puede deducirse de la demostración anterior que bajo las condiciones de regularidad y si $\hat{\pi}_n(\underline{x}_n)$ es fuertemente consistente para π_0; entonces, con probabilidad 1-(θ_0,π_0) existe $\theta_n^*(\underline{x}_n)$ tal que

$$\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\theta_n^*(\underline{x}_n),\hat{\pi}_n(\underline{x}_n);\underline{x}_n)=0 \text{ y } \theta_n^*(\underline{x}_n)\underset{n\uparrow\infty}{\to}\theta_0.$$

Lema 3. *Sean* $\underline{x}_n=(X_1,\ldots,X_n)$ *una muestra aleatoria como en 3 y* $\hat{\pi}_n(\underline{x}_n)$ *tal que:*

$$\sqrt{n}\left(\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\theta_0,\pi_0;\underline{x}_n),\ \hat{\pi}_n(\underline{x}_n)-\pi_0\right)\xrightarrow{\mathcal{D}}\underline{W}=(W_1,W_2)\sim N_2(\underline{0},\Sigma) \tag{1}$$

donde $\sum=\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{12} & \sigma_{22}\end{pmatrix}$.

Entonces, bajo las condiciones de regularidad ($\mathcal{C}$1) - ($\mathcal{C}$4), tenemos:

(i) $\sigma_{11} = I_{\theta_0}$ e

(ii) $\sigma_{12} = 0$.

Demostración:

(i) Como

$\sqrt{n}(\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\theta_0,\pi_0;\underline{x}_n)) = \sqrt{n}[\frac{1}{n}\sum_{j=1}^{n}\frac{\partial}{\partial\theta}\ln f(X_j|\theta_0,\pi_0)]$ y $\frac{\partial}{\partial\theta}\ln f(X_1|\theta_0,\pi_0)$ $\ldots,\frac{\partial}{\partial\theta}\ln f(X_n|\theta_0,\pi_0)$ son variables aleatorias independientes e indénticamente distribuídas, como media cero variancia I_{θ_0}, según la condición de regularidad ($\mathcal{C}4$); entonces por el teorema central del límite, $\sqrt{n}(\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\theta_0,\pi_0;\underline{x}_n)) \xrightarrow{\mathcal{D}} W_1 \sim N(0,I_{\theta_0})$. Por lo tanto, $\sigma_{11} = I_{\theta_0}$.

(ii) Como en la demostración del Teorema 1, sea $\tilde{\theta}_n(\underline{x}_n)$ tal que con probabilidad 1-(θ_0, π_0) $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\tilde{\theta}_n(\underline{x}_n),\underline{x}_n) = 0$ y $\tilde{\theta}_n(\underline{x}_n) \underset{n\uparrow\infty}{\rightarrow} \theta_0$.

En forma análoga a lo hecho en la demostración de (ii) del Lema 2, y aplicando la fórmula de Taylor a la función $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\cdot,\pi_0;\underline{x}_n)$, podemos mostrar que con probabilidad 1-(θ_0, π_0) :

$$\sqrt{n}(\tilde{\theta}_n(\underline{x}) - \theta_0) + Y_n^{-1}(\underline{x}_n)\sqrt{n}\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\theta_0,\pi_0;\underline{x}_n) \underset{n\uparrow\infty}{\rightarrow} 0 \qquad \textbf{(2)}$$

donde $P_{\theta_0,\pi_0}(Y_n(\underline{x}_n) \underset{n\uparrow\infty}{\rightarrow} -I_{\theta_0}) = 1$.

Luego, usando (1), (2) y el teorema de Slutsky se puede concluir que:

$$\sqrt{n}(\tilde{\theta}_n(\underline{x}_n) - \theta_0, \hat{\pi}_n(\underline{x}_n) - \pi_0) \xrightarrow{\mathcal{D}} \underline{U} \sim N_2(\underline{0},\Sigma') \qquad \textbf{(3)}$$

donde $\sum' = \begin{pmatrix} I_{\theta_0}^{-1} & -I_{\theta_0}^{-1}\sigma_{12} \\ -I_{\theta_0}^{-1}\sigma_{12} & \sigma_{22} \end{pmatrix}$.

Por otro lado, Bahadur (1964) muestra que,sobre parte de las condiciones de regularidad ($\mathcal{C}1$) - ($\mathcal{C}4$), $\tilde{\theta}_n(\underline{x}_n)$ es asintóticamente eficiente para θ_0, esto es:

Si $\delta_n(\underline{x}_n)$ satisface $\sqrt{n}(\delta_n(\underline{x}_n)-\theta_0)\xrightarrow{\mathcal{D}} V \sim N(0,v)$, entonces

$$v \geq I_{\theta_0}^{-1} \tag{4}$$

Luego, de (3) y (4) se concluye que σ_{12} = 0, ya que en caso contrario, tomando $\delta_n(\underline{x}_n)=\tilde{\theta}_n(\underline{x}_n)+c\,(\hat{\pi}_n(\underline{x}_n)-\pi_0)$, con c cualquier valor entre 0 y $-2\,I_{\theta_0}^{-1}\sigma_{22}$, tendríamos por (3) que, $\sqrt{n}(\delta_n(\underline{x}_n)-\theta_0)\xrightarrow{\mathcal{D}} V_1 \sim N(0,v_c)$ y sin embargo $v_c < I_{\theta_0}^{-1}$ lo cual contradice a (4).

Teorema 2. *Bajo las hipótesis del Teorema 1 y si:*

$$\sqrt{n}\left(\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\theta_0,\pi_0;\underline{x}_n),\, \hat{\pi}_n(\underline{x}_n)-\pi_0\right)\xrightarrow{\mathcal{D}} \underline{W}=(W_1,W_2)\sim N_2(\underline{0},\Sigma) \tag{1}$$

donde $\sum=\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{12} & \sigma_{22}\end{pmatrix}$.

Entonces

$$\sqrt{n}(\hat{\theta}_n(\underline{x}_n)-\theta_0,\, \hat{\pi}_n(\underline{x}_n)\pi_0)-\xrightarrow{\mathcal{D}} \underline{V}=(V_1,V_2)\sim N_2(\underline{0},\Sigma')$$

donde

$$\sum=\begin{pmatrix} I_{\theta_0}^{-1}+I_{\theta_0}^{-2}I_{\theta_0\pi_0}^{-2}\sigma_{22} & -I_{\theta_0}^{-1}I_{\theta_0,\pi_0}\sigma_{22}\\ -I_{\theta_0}^{-1}I_{\theta_0,\pi_0}\sigma_{22} & \sigma_{22}\end{pmatrix} \tag{2}$$

En particular, $\sqrt{n}(\hat{\theta}_n(\underline{x}_n)-\theta_0)\xrightarrow{\mathcal{D}} V_1 \sim N(0, I_{\theta_0}^{-1}+I_{\theta_0}^{-2}I_{\theta_0,\pi_0}^{2}\sigma_{22})$.

Demostración: Primeramente, σ_{11} y σ_{12}, satisfacen la tesis del Lema 3.

Como en las demostraciones, del Teorema1 y del Lema 3, sea $\tilde{\theta}_n(\underline{x}_n)$ tal que, con probabilidad 1-(θ_0,π_0), $\frac{\partial}{\partial\theta}\overline{\mathcal{L}}_n(\tilde{\theta}_n(\underline{x}_n),\pi_0;\underline{x}_n)=0$ y $\tilde{\theta}(\underline{x}_n)\underset{n\uparrow\infty}{\rightarrow}\theta_0$.

Tenemos, como en (3) de la demostración del Lema 3 que:

$$\sqrt{n}(\tilde{\theta}_n(\underline{x}_n)-\theta_0,\hat{\pi}_n(\underline{x}_n)-\pi_0)\xrightarrow{\mathcal{D}} \underline{V}=(V_1,V_2)\sim N_2(\underline{0},\Sigma')$$

donde

$$\sum{}' = \begin{pmatrix} I_{\theta_0}^{-1} & -I_{\theta_0}^{-1}\sigma_{12} \\ -I_{\theta_0}^{-1}\sigma_{12} & \sigma_{22} \end{pmatrix} \quad \textbf{(3)}$$

Por otra parte, como en la demostración del Teorema 1, para cada $\omega \in C \cap D \cap E$, se tiene que para cada $n \geq m_{\omega}$:

- Existe una vecindad $V_{\omega,n} \supseteq \{\pi_0, \hat{\pi}_n(\underline{x}_n(\omega))\}$ **(4)**

- Existe una función $g_{\omega,n}$ tal que

$$g_{\omega,n}(\pi_0) = \tilde{\theta}_n(\underline{x}_n(\omega)), g_{\omega,n}(\hat{\pi}_n(\underline{x}_n(\omega))) = \hat{\theta}_n(\underline{x}_n(\omega))$$

$$\text{y} \qquad \forall \pi \in V_{\omega,n}, \frac{\partial}{\partial \theta}\overline{\mathcal{L}}_n(g_{\omega,n}(\pi);\underline{x}_n(\omega)) = 0 \quad \textbf{(5)}$$

y además, por el teorema de la función implícita, $\forall \pi \in V_{\omega,n}$:

$$g'_{\omega,n}(\pi) = \frac{-\frac{\partial^2}{\partial \pi \partial \theta}\overline{\mathcal{L}}_n(g_{\omega,n}(\pi),\pi;\underline{x}_n(\omega))}{\frac{\partial^2}{\partial \theta^2}\overline{\mathcal{L}}_n(g_{\omega,n}(\pi),\pi;\underline{x}_n(\omega))} \quad \textbf{(6)}$$

y de aquí, por la condición de regularidad ($\mathcal{C}2$), también

$$\begin{aligned} g''_{\omega,n}(\pi) = [h_1(\pi)]^{-2} \cdot \{h_1(\pi)[h_2(\pi) \cdot g'_{\omega,n}(\pi) + h_3(\pi)] - \\ h_3(\pi)[h_4(\pi) \cdot g'_{\omega,n}(\pi) + h_2(\pi)\} \end{aligned} \quad \textbf{(7)}$$

donde

$$h_1(\pi) = \frac{\partial^2}{\partial \theta^2}\overline{\mathcal{L}}_n(g_{\omega,n}(\pi),\pi;\underline{x}_n(\omega)),\ h_2(\pi) = \frac{\partial^3}{\partial \pi \partial \theta^2}\overline{\mathcal{L}}_n(g_{\omega,n}(\pi),\pi;\underline{x}_n(\omega))$$

$$h_3(\pi) = \frac{\partial^2}{\partial \pi \partial \theta}\overline{\mathcal{L}}_n(g_{\omega,n}(\pi),\pi;\underline{x}_n(\omega)) \text{ y } h_4(\pi) = \frac{\partial^3}{\partial \partial \theta^3}\overline{\mathcal{L}}_n(g_{\omega,n}(\pi),\pi;\underline{x}_n(\omega))$$

Luego, aplicando la fórmula de Taylor a la función $g_{\omega,n}$, considerando los puntos π_0 y $\hat{\pi}_n(\underline{x}_n(\omega))$, y usando (3), (4), (5) y (6), podemos verificar, procediendo de manera análoga a lo hecho en la demostración de (i) del Lema (2); que con probabilidad 1-(θ_0, π_0) :

$$\sqrt{n}(\hat{\theta}_n(\underline{x}_n)-\theta_0)-\sqrt{n}(\hat{\theta}_n(\underline{x}_n)-\theta_0)-\sqrt{n}(\hat{\pi}_n(\underline{x}_n)-\pi_0)Z_n \underset{n\uparrow\infty}{\rightarrow} 0 \qquad (8)$$

con $Z_n \underset{n\to\infty}{\rightarrow} -I_{\theta_0}^{-1} I_{\theta_0\pi_0}$.

Luego el Teorema se sigue usando (1), (3), (8) y el teorema de Slutsky.

Observación 3. Bajo las condiciones de regularidad mencionadas en la Observación 1, si $\overset{*}{\theta}_n(\underline{x}_n)$ y $\overset{*}{\pi}_n(\underline{x}_n)$ son los estimadores de máxima verosimilitud de θ_0 y π_0, respectivamente, entonces:

$$\sqrt{n}(\overset{*}{\theta}_n(\underline{x}_n)-\theta_0,\overset{*}{\pi}_n(\underline{x}_n)-\pi_0)\xrightarrow{\mathcal{D}} N_2(\underline{0},\Sigma''),$$

donde

$$\sum{}'' = \begin{pmatrix} I_{\theta_0}^{-1}+I_{\theta_0}^{-2}I_{\theta_0\pi_0}^{2}\overset{*}{\sigma}_{22} & -I_{\theta_0}^{-1}I_{\theta_0,\pi_0}\overset{*}{\sigma}_{22} \\ -I_{\theta_0}^{-1}I_{\theta_0,\pi_0}\overset{*}{\sigma}_{22} & \overset{*}{\sigma}_{22} \end{pmatrix}$$

siendo $\overset{*}{\theta}_{22} = I_{\theta_0}(I_{\theta_0,\pi_0}I_{\pi_0} - I_{\theta_0,\pi_0}^{2})^{-1}$, con

$$I_{\pi_0} = -E_{\theta_0,\pi_0}[\frac{\partial^2}{\partial\pi^2}\ln f(X_1|\theta,\pi)].$$

Compárese esta última expresión con la dada en el Teorema 2. Queda claro que la eficiencia asintótica de $\hat{\theta}_n(\underline{x}_n)$ depende de la correspondiente a $\hat{\pi}_n(\underline{x}_n)$, en particular si $\hat{\pi}_n(\underline{x}_n)$ es asintóticamente eficiente, o sea si $\sigma_{22} = \overset{*}{\sigma}_{22}$, también lo será $\hat{\theta}_n(\underline{x}_n)$.

Referencias:

[1] Bahadur, R.R. (1964), *On Fisher's bound for asymptotic Variances.* Ann. Statist. 1545-1552.

[2] Gong, G. and F.J. Samaniego (1981). *Pseudo maximun likelihood estimation*: Theory and applications. Ann. Statist. 9 861-869.

[3] Parke, W.R. (1983). *Pseudo maximun likelihood estimation: The asymptotic distribution.* Ann. Statist. 1 355-457.

[4] Serfling, R.J. (1980). *Approximation theorems of mathematical statistics.* Willey, New York.

flores@ime.usp.br

Alumno del doctorado en Estadística del
Instituto de Matemática y Estadística
de la Universidad de São Paulo, Brasil.