Índice

Teoría asintótica de estimadores robustos en regresión no lineal

Tesis Doctoral

María Victoria Fasano

Director: Dr. Ricardo Antonio Maronna

Universidad Nacional de La Plata
Facultad de Ciencias Exactas
Departamento de Matemática

2009

1. Introducción

En el modelo de regresión no lineal con predictores aleatorios, observamos $\mathbf{z}_i = (y_i, \mathbf{x}_i)$, $1 \leq i \leq n$ vectores independientes idénticamente distribuidos (iid) $p+1$-dimensionales, con $y_i \in \mathbb{R}$, $\mathbf{x}_i \in \mathbb{R}^p$ que satisfacen

$$y_i = g\left(\mathbf{x}_i, \boldsymbol{\theta}_0\right) + u_i \quad 1 \leq i \leq n \tag{1}$$

siendo $\boldsymbol{\theta}_0 \in \Theta \subset \mathbb{R}^s$ el vector de parámetros de regresión donde Θ es el espacio de parámetros; u_i variables iid, independientes de las $\mathbf{x}_i$,y g es una función continua.

Algunos ejemplos de modelos no lineales con $x \in \mathbb{R}_+$son: el modelo de crecimiento exponencial,

$$g\left(x, \boldsymbol{\theta}\right) = \beta_0 + \sum_{j=1}^{p} \beta_j \exp\left(\alpha_j x\right), \quad \beta_j \in \mathbb{R},\ \alpha_j > 0\ \forall j; \tag{2}$$

y el de decrecimiento exponencial, igual al anterior pero con $\alpha_j < 0$; en ambos casos el vector de parámetros es $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta})$ donde $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_p)$ y $\boldsymbol{\beta} = \left(\beta_0, ..., \beta_p\right)$. Otro modelo es el de Michaelis-Menten

$$g\left(x, \boldsymbol{\theta}\right) = \frac{\beta x}{\alpha + x}, \quad \alpha, \beta > 0. \tag{3}$$

El estimador clásico para estos modelos es el de mínimos cuadrados definido como solución de

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta} \in \Theta} \frac{1}{n} \sum_{i=1}^{n} \left(y_i - g\left(\mathbf{x}_i, \boldsymbol{\theta}\right)\right)^2$$

Este estimador, cuando la distribución es normal, tienen propiedades óptimas; pero es muy sensible a observaciones atípicas: si la distribución no es exactamente normal, el ajuste puede variar mucho. Por esta razón, son necesarios estimadores robustos que sean más estables frente a distintas perturbaciones del modelo y que a su vez sean eficientes bajo normalidad.

En el caso lineal hay varios métodos para obtener estimadores robustos, entre ellos se encuentran los estimadores LMS y LTS propuestos por Rousseeuw (1984) y (1985), respectivamente, los estimadores de tipo M (o "M-estimadores") propuestos por Huber (1973), los MM- estimadores propuestos por Yohai (1987); los S-estimadores de Rousseeuw y Yohai (1987) que minimizan una escala robusta de los residuos, y los τ-estimadores de Yohai y Zamar (1988) que minimizan una escala robusta y eficiente de los residuos.

Para regresión no lineal también se han propuesto algunos estimadores robustos. Tiede y Pagano (1979) proponen un algoritmo para el cálculo de un M-estimador que aplican al análisis de radioinmunoensayos; Fraiman (1983) presenta una propuesta de regresión robusta no lineal, usando estimadores de influencia acotada; Markatou y Manos (1996) consideran tests de hipótesis en regresión no lineal basados en M-estimadores; Sakata y White (2001) trabajan con S-estimadores para modelos de regresión no lineal con observaciones dependientes; Stromberg (1993) propone algoritmos para el cálculo de MM-estimadores para regresión no lineal y Tabatabai y Argyros (1993) extienden el τ-estimador al caso no lineal, construyen tests y proponen un algoritmo para su cálculo. Mukherjee (1996) propone estimadores de mínima distancia. Carroll y Ruppert (1987) estudian métodos robustos para transformaciones no lineales de los datos.

Existen numerosos trabajos sobre teoría asintótica para mínimos cuadrados en regresión no lineal. Amemiya (1983), Jennrich (1969) y Johansen (1984) prueban la consistencia suponiendo entre otras condiciones la compacidad del espacio de parámetros. Wu (1981), por su parte, la demuestra para espacios de parámetros con una cantidad finita de elementos.

Jennrich (1969), Amemiya (1983) y Malinvaud (1970b) mostraron la normalidad asintótica. Bunke y Schmidt (1980) y Zwanzig (1980) extendieron los trabajos de Jennrich y Malinvaud a los casos de mínimos cuadrados pesados y distribución no idéntica de los errores, respectivamente.

En particular: Richardson y Bhattacharyya (1986) proponen un estimador de mínimos cuadrados modificado para regresión no lineal y prueban su consistencia suponiendo que la función g es acotada en $\boldsymbol{\theta}$.

Shao (1992) trata la consistencia del estimador de mínimos cuadrados y del estimador jacknife de su varianza. Es más general que Richardson y Battacharyya, pues no incluye la condición de compacidad del espacio de parámetros ni que la función de regresión g sea acotada; pero requiere condiciones muy fuertes que no se cumplen en casos importantes, como luego se verá.

Todos los trabajos sobre la teoría asintótica de estimadores robustos requieren la compacidad de Θ. Vainer y Kukush (1998) y Liese y Vajda (2003, 2004) tratan de M-estimadores. Estos últimos estudian la consistencia $O\left(n^{-1/2}\right)$ y la normalidad asintótica de M-estimadores en modelos estadísticos más generales incluyendo los modelos de regresión lineales y no lineales con observaciones independientes. Stromberg (1995) probó la consistencia débil del estimador LMS y Čížek (2005) trató la consistencia y normalidad asintótica del estimador LTS bajo dependencia.

En cuanto a medidas de robustez, hemos considerado el punto de ruptura. Este concepto lo introdujo Hampel (1971) y desde entonces se han hecho varias extensiones. Donoho y Huber (1983) extendieron la definición de punto de ruptura para muestras finitas que por su simplicidad resulta aplicable en varias situaciones, pero tiene la desventaja que depende de la parametrización y si el espacio de parámetros es acotado el punto de ruptura es 1 sin importar de qué estimador se trate. Stromberg y Ruppert (1992) proponen una definición alternativa del punto de ruptura para modelos de regresión no lineales que resulta invariante respecto a reparametrizaciones. Es ésta la definición de punto de ruptura que utilizaremos en este trabajo.

Por su parte, Sakata y White (1995) dan una definición de punto de ruptura finito para modelos de regresión que requiere un criterio de evaluación de la bondad del ajuste; Genton y Lucas (2003) introducen una definición de punto de ruptura más general que se aplica tanto para observaciones independientes como para observaciones dependientes.

En el tema de métodos numéricos para estimadores robustos no lineales, Dutter y Huber (1981) y Edlund et al. (1997) se ocupan de M estimadores. Stromberg (1993, 1994, 1997) desarrolla software para diversos estimadores de alto punto de ruptura, incluyendo LMS, LTS y MM. Vankeerberghen et al. (1995) proponen algoritmos genéticos para el cálculo de estimadores robustos.

2. Definición de los estimadores

Definimos un *M-estimador* como

$$\hat{\boldsymbol{\theta}}_n = \arg \min_{\boldsymbol{\theta} \in \Theta} \frac{1}{n} \sum_{i=1}^{n} \rho \left(\frac{y_i - g(\mathbf{x}_i, \boldsymbol{\theta})}{\widehat{\sigma}} \right) \tag{4}$$

donde $\widehat{\sigma}$ es un estimador de la escala σ de los errores, y ρ es una "*función-ρ*" en el sentido de Maronna et. al. (2006), esto es:

Definición 2.1 *Una* función-ρ *es una función real* ρ *que satisface las siguientes condiciones:*

i) $\rho(0) = 0$

ii) $\rho(-u) = \rho(u)$

iii) $0 \leq u \leq v \Rightarrow \rho(u) \leq \rho(v)$

iv) ρ *es continua*

v) Si $\rho(u) < \sup_u \rho(u)$ *y* $0 \leq u < v$ *entonces* $\rho(u) < \rho(v)$.

Si ρ es acotada, se supone además que $\sup_u \rho(u) = 1$.

A menos que se especifique lo contrario, las funciones ρ que utilizaremos en este trabajo serán funciones-ρ en este sentido.

Dada una muestra de tamaño n, $\mathbf{u} = (u_1, ..., u_n)$ y ρ una función-ρ acotada, se define el *M-estimador de escala* $s(\mathbf{u})$ como el valor de s que es solución de

$$\frac{1}{n} \sum_{i=1}^{n} \rho \left(\frac{u_i}{s} \right) = b \tag{5}$$

donde b es una constante entre 0 y 1.

Yohai (1987) definió los MM-estimadores para el caso de regresión lineal; aquí lo extenderemos al modelo no lineal.

Primera etapa: Tomamos un estimador $\mathbf{T}_{0,n}$ de $\boldsymbol{\theta}_0$ con alto punto de ruptura.

Segunda etapa: Calculamos los residuos

$$r_i\left(\mathbf{T}_{0,n}\right) = y_i - g\left(\mathbf{x}_i, \mathbf{T}_{0,n}\right) \qquad 1 \leq i \leq n$$

y calculamos el M-estimador de escala $s_n{=}s(\mathbf{r}\left(\mathbf{T}_{0,n}\right))$ definido por (5) usando una función-ρ acotada ρ_0 y

$$b = 0{,}5 \tag{6}$$

Tercera etapa: Sea ρ_1 otra función-ρ acotada que cumple

$$\rho_1(u) \leq \rho_0(u) \tag{7}$$

Sea $\psi_1 = \rho_1'$ y $\dot{\mathbf{g}}$ el gradiente de g, entonces el *MM-estimador* $\mathbf{T}_{1,n}$ se define como cualquier solución de

$$\sum_{i=1}^{n} \psi_1\left(\frac{r_i\left(\boldsymbol{\theta}\right)}{s_n}\right)\dot{\mathbf{g}}\left(\mathbf{x}_i, \boldsymbol{\theta}\right) = \mathbf{0} \tag{8}$$

que satisface

$$S(\mathbf{T}_{1,n}) \leq S(\mathbf{T}_{0,n}) \tag{9}$$

donde

$$S(\boldsymbol{\theta}) = \sum_{i=1}^{n} \rho_1\left(\frac{r_i\left(\boldsymbol{\theta}\right)}{s_n}\right) \tag{10}$$

Los τ-estimadores en regresión lineal fueron introducidos por Yohai y Zamar (1988). Sean ρ_1 y ρ_2 dos funciones-ρ acotadas y sea s_n el M-estimador de escala basado en ρ_1 definido en (5). Entonces dada la muestra $\mathbf{u} = (u_1, ..., u_n)$, el *estimador de escala* τ_n se define mediante

$$\tau_n^2\left(\mathbf{u}\right) = s_n^2\left(\mathbf{u}\right)\sum_{i=1}^{n} \rho_2\left(\frac{u_i}{s_n\left(\mathbf{u}\right)}\right) \tag{11}$$

El $\tau-$*estimador* para regresión se define como el valor $\hat{\boldsymbol{\theta}}$ tal que

$$\tau\left(\mathbf{r}\left(\hat{\boldsymbol{\theta}}\right)\right) = \min_{\boldsymbol{\theta}} \tau\left(\mathbf{r}\left(\boldsymbol{\theta}\right)\right) \tag{12}$$

3. Descripción de los resultados.

En esta Sección se dan las suposiciones principales y un resumen de los resultados más importantes. Los detalles de definiciones y demostraciones se darán en secciones posteriores.

Se supone que las $\mathbf{x}_i$ son aleatorias e independientes de u_i.

Se supone también que la función g satisface

$$\boldsymbol{\theta} \neq \boldsymbol{\theta}_0 \Rightarrow P\left\{g\left(\mathbf{x}, \boldsymbol{\theta}\right) = g\left(\mathbf{x}, \boldsymbol{\theta}_0\right)\right\} < 1. \tag{13}$$

Esta condición es necesaria para la identificación de los parámetros.

Sea $G_0(\mathbf{x})$ la distribución de las $\mathbf{x}_i$ y $F_0(u)$ la distribución de los errores u_i. Entonces la distribución de $\mathbf{z}_i = (\mathbf{x}_i, y_i)$ está dada por

$$H_0(\mathbf{z}) = G_0(\mathbf{x}) F_0(y - g(\mathbf{x}, \boldsymbol{\theta}_0)) \tag{14}$$

Supondremos las siguientes dos condiciones:

A. $E\left|\rho\left(y - g\left(\boldsymbol{\theta}\right)\right)\right| < \infty$ para todo $\boldsymbol{\theta} \in \Theta$

B. La distribución F_0 tiene una densidad f_0 con las siguientes propiedades:

- **i)** f_0 es par.
- **ii)** $f_0(|u|)$ es monótona decreciente
- **iii)** $f_0(|u|)$ es estrictamente decreciente en un entorno de 0.

Una situación interesante es la que hemos llamado *componentes lineales,* en la cual la función g tiene una forma particular. Sea $p = p_1 + p_2$, $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta})$ con $\boldsymbol{\alpha} \in A \subset \mathbb{R}^{p_1}$, $\boldsymbol{\beta} \in \mathbb{R}^{p_2}$, $\Theta = A \times \mathbb{R}^{p_2}$, y

$$g\left(\mathbf{x}, \boldsymbol{\theta}\right) = g\left(\mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}\right) = \sum_{j=1}^{p_2} \beta_j h_j\left(\mathbf{x}, \alpha\right) = \boldsymbol{\beta}' \mathbf{h}\left(\mathbf{x}, \boldsymbol{\alpha}\right) \tag{15}$$

con $\mathbf{h}\left(\mathbf{x}, \boldsymbol{\alpha}\right) = \left(h_1\left(\mathbf{x}, \boldsymbol{\alpha}\right), ..., h_{p_2}\left(\mathbf{x}, \boldsymbol{\alpha}\right)\right)'$ donde las h_j son continuas en $\boldsymbol{\alpha}$. Dos ejemplos donde la función g tiene esta forma son el modelo de crecimiento o decrecimiento exponencial (2) con $h_j\left(x, \boldsymbol{\alpha}\right) = \exp\left(\alpha_j x\right)$ y el de Michaelis-Menten (3) con $h\left(x, \alpha\right) = \dfrac{x}{\alpha + x}$.

Para poder identificar los parámetros en estos modelos supondremos las siguientes condiciones sobre la función h:

$$\boldsymbol{\alpha} \neq \boldsymbol{\alpha}_0 \Rightarrow P\left\{\boldsymbol{\beta}'\mathbf{h}\left(\mathbf{x},\boldsymbol{\alpha}\right)=\boldsymbol{\beta}_0'\mathbf{h}\left(\mathbf{x},\boldsymbol{\alpha}_0\right)\right\}<1\ \forall\boldsymbol{\beta} \tag{16}$$

$$\text{y } P\left(\boldsymbol{\beta}'\mathbf{h}\left(\mathbf{x},\boldsymbol{\alpha}\right)=0\right)\leq\delta<1\ \forall\boldsymbol{\beta}\neq\mathbf{0},\ \forall\alpha \tag{17}$$

Para el desarrollo de la teoría asintótica trataremos los siguientes tres casos:

1. g general, Θ compacto
2. g acotada
3. g con componentes lineales

3.1. M -estimadores.

En esta sección se tratarán estimadores de la forma (4) suponiendo que la escala σ de los errores es conocida. Sin pérdida de generalidad se puede tomar $\sigma=1$.

Si llamamos $\dot{\mathbf{g}}$ al gradiente de g y $\psi=\rho'$, se tiene que $\hat{\boldsymbol{\theta}}_n$ es solución de la ecuación

$$\sum_{i=1}^{n}\psi\left(y_i-g\left(\mathbf{x}_i,\boldsymbol{\theta}\right)\right)\dot{\mathbf{g}}\left(\mathbf{x}_i,\boldsymbol{\theta}\right)=\mathbf{0}$$

Sea $\lambda\left(\boldsymbol{\theta}\right)=E\rho\left(y-g\left(\mathbf{x},\boldsymbol{\theta}\right)\right)$. A continuación analizamos la consistencia.

Teorema 3.1 *Supongamos que el espacio de parámetros Θ es compacto y se cumplen (A) y (B). Entonces el M-estimador $\hat{\boldsymbol{\theta}}_n$ definido en (4) es consistente para $\boldsymbol{\theta}_0$.*

Para el siguiente resultado no se requiere nada sobre Θ.

Teorema 3.2 *Supongamos que*

$$P\left\{\sup_{\boldsymbol{\theta}}\left|g\left(\mathbf{x},\boldsymbol{\theta}\right)\right|<\infty\right\}=1 \tag{18}$$

y se cumplen (A) y (B). Entonces el M-estimador $\hat{\boldsymbol{\theta}}_n$ definido en (4) es consistente para $\boldsymbol{\theta}_0$.

Para el caso de componentes lineales, $\lambda(\boldsymbol{\theta}) = \lambda(\alpha, \boldsymbol{\beta}) = E\rho(\mathbf{y} - g(\mathbf{x}, \alpha, \boldsymbol{\beta}))$. En este caso el estimador resulta consistente bajo las siguientes condiciones:

Teorema 3.3 *Supongamos el modelo (15) con las condiciones (16), (17), (A), (B) y* $\sup_{\boldsymbol{\alpha}\in A} \|\mathbf{h}(\mathbf{x}, \boldsymbol{\alpha})\| < \infty$ *c.s. Entonces el M-estimador* $\hat{\boldsymbol{\theta}}_n = \left(\hat{\boldsymbol{\alpha}}_n, \hat{\boldsymbol{\beta}}_n\right)$ *definido en (4) es fuertemente consistente para* $\boldsymbol{\theta}_0 = (\boldsymbol{\alpha}_0, \boldsymbol{\beta}_0)$.

Shao (1992) prueba la consistencia del estimador de mínimos cuadrados sin suponer la compacidad de Θ, pero utiliza suposiciones sobre g que en el caso de componentes lineales excluyen a los modelos más usuales. En efecto, requiere (pg. 417) que para algún c_0 tal que $P(\|\mathbf{x}\| \leq c_0) > 0$ se cumpla

"Condición 1": $\lim_{\boldsymbol{\theta}\to\infty} |g(\mathbf{x}, \boldsymbol{\theta})| = \infty$ uniformemente en $\|\mathbf{x}\| \leq c_0$

o bien

"Condición 2": Existen a, c tales que $\lim_{\boldsymbol{\theta}\to\infty} |g(\mathbf{x}, \boldsymbol{\theta})| = c$ uniformemente en $\|\mathbf{x}\| \leq c_0$, y

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left(g(\mathbf{x}_i, \boldsymbol{\theta}_0) - a\right)^2 > 0.$$

Los modelos (2) y (3) evidentemente no cumplen la primera parte de la Condición 2; y tampoco cumplen la Condición 1. Por ejemplo, para $g(x, \boldsymbol{\theta}) = \beta e^{\alpha x}$ se puede hacer $g(x_0, \boldsymbol{\theta})$ =constante con $\alpha \to -\infty$ y $\beta \to 0$. Sin embargo, curiosamente Shao (1992) afirma en su "Example 2" (pg. 427) que aquí la condición se cumple.

Finalmente, probaremos la normalidad asintótica del M$-$estimador, para lo cual son necesarias algunas condiciones adicionales.

Teorema 3.4 *Sea* $\hat{\boldsymbol{\theta}}_n$ *definido en (4). Supongamos que se cumplen (A) y (B), que* g *satisface (18) y sus dos primeras derivadas son* F_0*-integrables. Además supondremos que la matriz* $V = E\left(\dot{\mathbf{g}}(\mathbf{x}, \boldsymbol{\theta})\, \dot{\mathbf{g}}(\mathbf{x}, \boldsymbol{\theta})'\right)$ *es definida positiva con probabilidad uno y que la función* ψ *es continua, acotada y tiene derivada acotada. Entonces*

$$\sqrt{n}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right) \xrightarrow{d} N\left(0, \frac{E\psi^2(u)}{\left(E\psi'(u)\right)^2} V^{-1}\right)$$

Denotamos por $\xrightarrow{d}$ *la convergencia en distribución, y* $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *denota la distribución normal multivariada con media* $\boldsymbol{\mu}$ *y matriz de covarianzas* $\boldsymbol{\Sigma}$.

3.2. MM-estimadores.

En esta sección trataremos los MM-estimadores definidos en (8). Extenderemos los resultados sobre consistencia, normalidad, robustez cualitativa y punto de ruptura, obtenidos por Yohai (1987) para el modelo de regresión lineal, al caso de regresión no lineal.

En esta sección y en la siguiente las funciones ρ que aparezcan serán funciones-ρ *acotadas.*

3.2.1. Consistencia

Se probará la consistencia del estimador de escala s_n definido en la segunda etapa y de la sucesión de MM-estimadores $\{\mathbf{T}_{1,n}\}_{n\geq p}$ de $\boldsymbol{\theta}_0$.

Consideramos $\boldsymbol{\theta}_0 \in \Theta$, con Θ compacto.

Teorema 3.5 *Supongamos que* $\{\mathbf{T}_{0,n}\}_{n\geq p}$ *es una sucesión de estimadores que es fuertemente consistente para* $\boldsymbol{\theta}_0$*. Entonces* s_n *es fuertemente consistente para* σ_0 *definido por*

$$E_{F_0}\rho_0\left(\frac{u}{\sigma_0}\right) = b \tag{19}$$

Para probar la consistencia del estimador será necesaria la siguiente suposición adicional:

C. $\forall\varepsilon > 0,\ P_{G_0}\left\{\inf_{\|\theta-\boldsymbol{\theta_0}\|>\epsilon} |g(\mathbf{x},\boldsymbol{\theta}) - g(\mathbf{x},\boldsymbol{\theta}_0)| > 0\right\} \geq 0{,}5$

El siguiente resultado muestra la consistencia de $\mathbf{T}_{1,n}$.

Teorema 3.6 *Supongamos que valen (6) y (7) ; que* F_0 *satisface (B) y que* G_0 *satisface (C). Supongamos también que la sucesión* $\{\mathbf{T}_{0,n}\}_{n\geq p}$ *es fuertemente consistente para* θ_0*. Entonces cualquier otra sucesión* $\{\mathbf{T}_{1,n}\}_{n\geq p}$ *que satisfaga (9) es fuertemente consistente.*

3.2.2. Espacio no compacto.

Los resultados anteriores han sido probados bajo la suposición que el espacio de parámetros Θ sea compacto. El teorema 3.6 vale también cuando el espacio Θ no es compacto pero la función $g\,(\mathbf{x},\boldsymbol{\theta})$ es acotada en $\boldsymbol{\theta}$, es decir cumple la condición (18).

3.2.3. Normalidad

Para probar la normalidad asintótica de los MM-estimadores necesitamos algunas suposiciones adicionales:

D. ρ_1 es dos veces diferenciable con continuidad y $\exists m$ tal que $|u| \geq m \Rightarrow \rho_1(u) = 1$

E. G_0 tiene segundo momento y la matriz

$$V = E_{G_0}\left(\dot{\mathbf{g}}(\mathbf{x}, \boldsymbol{\theta})\dot{\mathbf{g}}(\mathbf{x}, \boldsymbol{\theta})'\right) \tag{20}$$

es no singular.

Teorema 3.7 *Supongamos que ρ_1 satisface (D) y G_0 satisface (E). Sea s_n un estimador de la escala del error que converge fuertemente a σ_0. Sea $\mathbf{T}_n$ una sucesión de estimadores que satisface (8) y que es fuertemente consistente para el valor verdadero $\boldsymbol{\theta}_0$. Entonces,*

$$\sqrt{n}\left(\mathbf{T}_n - \boldsymbol{\theta}_0\right) \xrightarrow{d} N\left(0, \sigma_0^2 \frac{A(\psi_1, F_0)}{B(\psi_1, F_0)} V^{-1}\right) \tag{21}$$

donde

$$A(\psi, F) = E_F \psi^2\left(\frac{u}{\sigma_0}\right)$$

$$B(\psi, F) = \left(E_F \psi'\left(\frac{u}{\sigma_0}\right)\right)^2$$

3.2.4. Robustez cualitativa

Supongamos que la sucesión inicial $\{\mathbf{T}_{0,n}\}_{n\geq p}$ está dada por una funcional $\mathbf{T}_0$, es decir, $\mathbf{T}_{0,n} = \mathbf{T}_0\left(H_n\right)$, donde H_n es la distribución empírica de $\mathbf{z}_1, ..., \mathbf{z}_n$. El siguiente teorema muestra que la funcional que define el MM-estimador es continua.

Teorema 3.8 *Sea H_0 dada por (14). Supongamos que $\mathbf{T}_0$ es continua en H_0 y $\mathbf{T}_0\left(H_0\right) = \theta_0$. Supongamos también que se cumplen (6) y (7), F_0 satisface (B) y G_0 satisface (C). Entonces la funcional $\mathbf{T}_1$, correspondiente a la sucesión de MM-estimadores $\{\mathbf{T}_{1,n}\}_{n\geq p}$ es también continua en H_0.*

Como corolario de este teorema resulta que el estimador es asintóticamente cualitativamente robusto.

3.2.5. Punto de ruptura

Definimos

$$M=\min_{i_0}\min_{A}\max_{B}\min_{\boldsymbol{\theta}\in U(B,i_0)}\#\left\{i:g\left(\mathbf{x}_i,\boldsymbol{\theta}\right)>A\right\} \tag{22}$$

siendo $U\left(B,i_0\right)=\left\{\boldsymbol{\theta}:g\left(\mathbf{x}_{i_0},\boldsymbol{\theta}\right)>B\right\}$.

Sea $c_n=1-\dfrac{M}{n}$. Así si $c_n<0{,}5$ entonces $M>\frac{n}{2}$. En el siguiente resultado veremos que el punto de ruptura del MM-estimador es no menor que el punto de ruptura del estimador inicial.

Teorema 3.9 *Supongamos que se cumplen (6) y (7) y que* $c_n<0{,}5$. *Entonces si* $\mathbf{T}_0=\left\{T_{0,n}\right\}_{n\geq p}$, $\mathbf{T}_1=\left\{T_{1,n}\right\}_{n\geq p}$ *son dos sucesiones de estimadores que satisfecen (9), se tiene que para cualquier muestra* $\mathbf{Z}_n=(z_1,...,z_n)$

$$\varepsilon_{n,+}^{*}(g,\mathbf{T}_1,\mathbf{Z}_n)\geq\min\left\{\varepsilon_{n,+}^{*}(g,\mathbf{T}_0,\mathbf{Z}_n),\frac{1-2c_n}{2-2c_n}\right\}$$

donde $\varepsilon_{n,+}^{*}$ *es el punto de ruptura definido por Stromberg y Ruppert (1992), cuya definición se da en la Sección 3.2.*

3.3. Tau-estimadores.

En esta sección analizaremos el comportamiento de los $\tau-$estimadores definidos en (12).

Nótese que τ_n es equivariante por escala, esto es $\tau_n\left(\lambda\mathbf{u}\right)=\left|\lambda\right|\tau_n\left(\mathbf{u}\right)$ para todo $\lambda\in\mathbb{R}$.

Si $\rho_1=\rho_2$, se tiene $\tau_n=\sqrt{b}s_n$. Si $\rho_2\left(u\right)=u^2$, se obtiene el desvío standard muestral.

Sean $\psi_j=\rho_j'$, para $j=1,2$, y sea para $i=1,...,n$

$$t_i=\frac{r_i\left(\boldsymbol{\theta}\right)}{s_n\left(\mathbf{r}\left(\boldsymbol{\theta}\right)\right)};$$

entonces derivando τ_n respecto de $\boldsymbol{\theta}$ en (11) se obtiene el siguiente sistema de ecuaciones para el $\tau-$estimador

$$\begin{aligned}
&\frac{\partial \tau^2 (\mathbf{r}(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \\
&= \frac{2s_n (\mathbf{r}(\boldsymbol{\theta}))}{n} \frac{\partial s_n (\mathbf{r}(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \sum_{i=1}^{n} \rho_2 (t_i) \\
&+ \frac{1}{n} \sum_{i=1}^{n} \psi_2 (t_i) \left(-\dot{\mathbf{g}} (x_i, \boldsymbol{\theta}) s_n (\mathbf{r}(\boldsymbol{\theta})) - r_i (\boldsymbol{\theta}) \frac{\partial s_n (\mathbf{r}(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \right) = 0
\end{aligned} \tag{23}$$

Por otro lado, derivando (5)

$$\frac{\partial s_n (\mathbf{r}(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} = -s_n (\mathbf{r}(\boldsymbol{\theta})) \frac{\sum_{i=1}^{n} \psi_1 (t_i) \dot{\mathbf{g}} (x_i, \boldsymbol{\theta})}{\sum_{i=1}^{n} \psi_1 (t_i) r_i (\boldsymbol{\theta})} \tag{24}$$

Reemplazando (24) en (23) se obtiene la ecuación del $\tau-$estimador

$$\sum_{i=1}^{n} \left[W_n (\boldsymbol{\theta}) \psi_1 (t_i) + \psi_2 (t_i) \dot{\mathbf{g}} (x_i, \boldsymbol{\theta}) \right] = 0 \tag{25}$$

siendo

$$W_n (\boldsymbol{\theta}) = \frac{\sum_{i=1}^{n} \left[2\rho_2 (t_i) - \psi_2 (t_i) t_i \right]}{\sum_{i=1}^{n} \psi_1 (t_i) t_i} \tag{26}$$

y $s_n (\mathbf{r}(\boldsymbol{\theta}))$ que satisface

$$\frac{1}{n} \sum_{i=1}^{n} \rho_2 \left(\frac{r_i (\boldsymbol{\theta})}{s_n (\mathbf{r}(\boldsymbol{\theta}))} \right) = b \tag{27}$$

Se agrega la condición

F. ρ_2 satisface la desigualdad

$$2\rho_2 (u) - \psi_2 (u) u \geq 0 \tag{28}$$

Esta condición implica que $W_n (\boldsymbol{\theta}) \geq 0$. Entonces se puede pensar al $\tau-$estimador como un M-estimador con función ψ "adaptiva"

$$\psi_n (u) = W_n (u) \psi_1 (u) + \psi_2 (u)$$

que es un promedio pesado de ψ_1 y ψ_2. Efectivamente se verá que, asintóticamente, el $\tau-$estimador se comporta como un M-estimador con función ψ

$$\psi_0\,(u) = W_0 \psi_1 \left(\frac{u}{\sigma_0} \right) + \psi_2 \left(\frac{u}{\sigma_0} \right) \tag{29}$$

donde

$$W_0 = \frac{2E_{F_0} \rho_2 \left(\frac{u}{\sigma_0} \right) - E_{F_0} \left\{ \psi_2 \left(\frac{u}{\sigma_0} \right) \frac{u}{\sigma_0} \right\}}{E_{F_0} \left\{ \psi_1 \left(\frac{u}{\sigma_0} \right) \frac{u}{\sigma_0} \right\}} \tag{30}$$

y donde σ_0 se define mediante la ecuación

$$E_{F_0} \rho_1 \left(\frac{u}{\sigma_0} \right) = b \tag{31}$$

Yohai y Zamar (1988) probaron la consistencia, normalidad y punto de ruptura para τ-estimadores en el modelo de regresión lineal. En las secciones que siguen extenderemos estos resultados para el modelo de regresión no lineal.

3.3.1. Consistencia

El teorema a continuación establece la consistencia del $\tau-$estimador, para lo cual será necesaria una suposición adicional:

G. ρ_2 es continuamente diferenciable.

Teorema 3.10 *Sean ρ_1 que satisface (6) y ρ_2 que satisface (F) y (G). Supongamos que F_0 satisface (B) y G_0 satisface (C). Sea $\{T_{0,n}\}_{n \geq p}$ una sucesión de estimadores que converge fuertemente a $\boldsymbol{\theta}_0$ y sea $\{T_{1,n}\}_{n \geq p}$ otra sucesión de estimadores tal que*

$$\tau_n\left(\mathbf{r}\left(T_{1,n}\right)\right) \leq \tau_n\left(\mathbf{r}\left(T_{0,n}\right)\right). \tag{32}$$

Entonces $T_{1,n}$ también es fuertemente consistente para $\boldsymbol{\theta}_0$.

Observación 3.1 *Según este resultado no es necesario hallar un mínimo absoluto para obtener consistencia. Será suficiente con encontrar un mínimo local que mejore la τ_n -escala de los residuos del estimador definido a partir de un estimador consistente.*

3.3.2. Espacio no compacto

También en este caso podemos extender los resultados de consistencia para el caso en que el espacio de parámetros Θ no sea compacto pero la función g sea acotada en $\boldsymbol{\theta}$ con probabilidad 1, de manera análoga a como se hizo con los M-estimadores en la sección (3.1).

3.3.3. Normalidad

Veremos ahora que un estimador fuertemente consistente que satisface (25), (26) y (27) es asintóticamente equivalente a un M-estimador con función ψ dada por (29).

Teorema 3.11 *Supongamos que ρ_1 y ρ_2 satisfacen (D) y G_0 satisface (E). Supongamos también que T_n es una sucesión de estimadores fuertemente consistentes bajo H_0 que satisfacen (25), (26) y (27). Entonces*

$$\sqrt{n}\left(\mathbf{T}_n-\boldsymbol{\theta}_0\right) \xrightarrow{d} N\left(\mathbf{0}, \boldsymbol{\sigma}_0^2 \frac{A\left(\psi_0, F_0\right)}{B\left(\psi_0, F_0\right)} V^{-1}\right) \tag{33}$$

donde

$$A(\psi, F)=E_F \psi^2(u)$$
$$B(\psi, F)=\left(E_F \psi^{\prime}(u)\right)^2,$$

V está dada por (20) y ψ_0 está definida en (29).

3.3.4. Punto de ruptura

El siguiente teorema muestra que no es necesario hallar un mínimo absoluto para tener punto de ruptura alto, sino que basta con un estimador $\{T_{1,n}\}$ que satisfaga (34), siendo $\{T_{0,n}\}$ un estimador inicial con punto de ruptura alto.

Teorema 3.12 *Supongamos que ρ_1 satisface (6) y ρ_2 satisface (F). Sean $\mathbf{T}_0=\{T_{0,n}\}_{n\geq p}$, $\mathbf{T}_1=\{T_{1,n}\}_{n\geq p}$ dos sucesiones de estimadores de $\boldsymbol{\theta}_0$ tales que*

$$\tau_n\left(\mathbf{r}\left(T_{1,n}\right)\right) \leq \tau_n\left(\mathbf{r}\left(T_{0,n}\right)\right). \tag{34}$$

Entonces para cualquier muestra $\mathbf{Z}_n=(z_1,...,z_n)$ se tiene

$$\varepsilon_{n,+}^*\left(g, \mathbf{T}_1, \mathbf{Z}_n\right) \geq \min\left\{\varepsilon_{n,+}^*\left(g, \mathbf{T}_0, \mathbf{Z}_n\right), \frac{1-2c_n}{2-2c_n}\right\} \tag{35}$$

siendo $\varepsilon^*_{n,+}$ *el punto de ruptura definido por Stromberg y Ruppert (1992).*

4. Demostraciones correspondientes a la Sección 3.1.

Demostración del teorema 3.1. Se aplica directamente el teorema 1 de Huber (1967). ■

Demostración del teorema 3.2. Por el lema 3.1 de Yohai (1987) $E\rho\left(u-t\right)$ alcanza un único mínimo en $t=0$, luego $E\rho\left(\mathbf{y}-g\left(\mathbf{x},\boldsymbol{\theta}\right)\right)$ alcanza el mínimo sólo cuando $g\left(\mathbf{x},\boldsymbol{\theta}\right)=g\left(\mathbf{x},\boldsymbol{\theta}_0\right)$ y por (13) esto es así cuando $\boldsymbol{\theta}=\boldsymbol{\theta}_0$ Por lo tanto,

$$\boldsymbol{\theta}\in\boldsymbol{\Theta},\boldsymbol{\theta}\neq\boldsymbol{\theta}_0\Rightarrow\lambda\left(\boldsymbol{\theta}\right)>\lambda\left(\boldsymbol{\theta}_0\right):=\lambda_0 \tag{36}$$

Si Θ compacto, el resultado se obtiene por el teorema 3.1. Si Θ no es compacto, pero g es acotada con probabilidad uno, sea Θ_c la compacitificación de Stone-Čech de Θ [Munkres 1975]. Entonces la función g puede extenderse de manera continua y acotada a una función $\tilde{g}$ definida en Θ_c. Como para $\boldsymbol{\theta}\in\Theta_c$ sigue valiendo (13), entonces también se tiene (36).

Sea

$$\hat{\boldsymbol{\theta}}_{n,c}=\arg\min_{\boldsymbol{\theta}\in\Theta_c}\frac{1}{n}\sum_{i=1}^{n}\rho\left(y_i-g\left(\mathbf{x}_i,\boldsymbol{\theta}\right)\right)$$

Ahora estamos trabajando nuevamente en un compacto y se tiene como antes la consistencia de $\hat{\boldsymbol{\theta}}_{n,c}$. Luego, como $\hat{\boldsymbol{\theta}}_{n,c}\underset{n\to\infty}{\to}\boldsymbol{\theta}_0\in\Theta$ entonces $\exists N$ tal que $\forall n\geq N$, $\hat{\boldsymbol{\theta}}_{n,c}\in\Theta$, por lo tanto $\hat{\boldsymbol{\theta}}_{n,c}=\hat{\boldsymbol{\theta}}_n$, $\forall n\geq N$ y entonces $\hat{\boldsymbol{\theta}}_n\underset{n\to\infty}{\to}\boldsymbol{\theta}_0$ c.s. ■

Para la demostración del teorema 3.3 será necesario un lema previo.

Lema 4.1 *Supongamos el modelo (15) con las condiciones (16), (17), (A), (B) y A compacto. Entonces* $\left\|\hat{\boldsymbol{\beta}}_n\right\|$ *está acotado para* n *grande con probabilidad uno.*

Demostración del lema 4.1. Sea

$$\lambda\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)=E\rho\left(y-\boldsymbol{\beta}'\mathbf{h}\left(\mathbf{x},\boldsymbol{\alpha}\right)\right).$$

Por el lema 3.1 de Yohai, $\lambda\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)$ alcanza su mínimo sólo cuando $\boldsymbol{\beta}'\mathbf{h}\left(\mathbf{x},\boldsymbol{\alpha}\right)=\boldsymbol{\beta}_0'\mathbf{h}\left(\mathbf{x},\boldsymbol{\alpha}_0\right)$ c.s. y por (16) esto es así cuando $\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)=\left(\boldsymbol{\alpha}_0,\boldsymbol{\beta}_0\right)$. Por lo tanto,

$$\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)\neq\left(\boldsymbol{\alpha}_0,\boldsymbol{\beta}_0\right)\Rightarrow\lambda\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)>\lambda_{\text{mín}}=:\lambda\left(\boldsymbol{\alpha}_0,\boldsymbol{\beta}_0\right) \tag{37}$$

Sea $\Gamma = \{\boldsymbol{\gamma} \in \mathbb{R}^{p_2} : \|\boldsymbol{\gamma}\| = 1\}$. Entonces podemos escribir $\boldsymbol{\beta} = t\boldsymbol{\gamma}$ con $t = \|\boldsymbol{\beta}\| \in \mathbb{R}_+$, $\boldsymbol{\gamma} \in \Gamma$.

Supongamos primero que $S = \sup_u \rho(u) < \infty$. Para cada $(\boldsymbol{\alpha}, \boldsymbol{\gamma}) \in A \times \Gamma$ tenemos

$$\lim_{t\to\infty} E\rho\left(\mathbf{y} - t\boldsymbol{\gamma}'\mathbf{h}(\mathbf{x}, \boldsymbol{\alpha})\right) \geq S.(1-\delta) > \lambda_{\text{mín}},$$

donde δ está definido en .(17). Sea

$$\xi = S(1-\delta) - \lambda_{\text{mín}} > 0,\ \varepsilon = \frac{\xi}{4S} < \frac{1-\delta}{4}.$$

Como (17) implica que $P\left(|\boldsymbol{\beta}'\mathbf{h}(\mathbf{x}, \boldsymbol{\alpha})| > 0\right) \geq 1-\delta$, entonces para $(\boldsymbol{\alpha}, \boldsymbol{\gamma}) \in A \times \Gamma$, existen a, b positivos tales que

$$P\left(|y| \leq a, |\boldsymbol{\gamma}'\mathbf{h}(\mathbf{x}, \boldsymbol{\alpha})| \geq b\right) \geq 1 - \delta - \varepsilon \tag{38}$$

Luego, por (38), existe $T > 0$ tal que $t > T$ implica

$$E \inf_{t>T} \rho\left(\mathbf{y} - t\boldsymbol{\gamma}'\mathbf{h}(\mathbf{x}, \boldsymbol{\alpha})\right) > S(1 - \delta - 2\varepsilon) \tag{39}$$

Por lo tanto, (39) implica que para cada $(\boldsymbol{\alpha}, \boldsymbol{\gamma}) \in A \times \Gamma$ existen un entorno $U(\boldsymbol{\alpha}, \boldsymbol{\gamma}) \subset A \times \Gamma$ y un $T(\boldsymbol{\alpha}, \boldsymbol{\gamma}) \in \mathbb{R}_+$ tales que

$$E \inf_{(\boldsymbol{\alpha}_1, \boldsymbol{\gamma}_1) \in U(\boldsymbol{\alpha}, \boldsymbol{\gamma})} \inf_{t > T(\boldsymbol{\alpha}, \boldsymbol{\gamma})} \rho\left(\mathbf{y} - t\boldsymbol{\gamma}_1'\mathbf{h}(\mathbf{x}, \boldsymbol{\alpha}_1)\right) > S(1 - \delta - 3\varepsilon) = \lambda_{\text{mín}} + \frac{\xi}{4} \tag{40}$$

Cubrimos $A \times \Gamma$ con los entornos $U(\boldsymbol{\alpha}, \boldsymbol{\gamma})$, por ser compacto, existe un subcubrimiento finito $\left\{U_j = U\left(\alpha_j, \boldsymbol{\gamma}_j\right)\right\}_{j=1}^{N}$.

Sea $T_0 = \text{máx}_j\, T\left(\alpha_j, \boldsymbol{\gamma}_j\right)$. Veremos que $\liminf_{n\to\infty} \left\|\hat{\boldsymbol{\beta}}_n\right\| \leq T_0$ c.s.

Sea

$$\lambda_n(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n} \rho\left(y_i - \boldsymbol{\beta}'\mathbf{h}(x_i, \boldsymbol{\alpha})\right)$$

Entonces,

$$\begin{aligned}
\inf_{\|\boldsymbol{\beta}\| > T_0} \inf_{\boldsymbol{\alpha} \in A} \lambda_n(\boldsymbol{\alpha}, \boldsymbol{\beta}) &\geq \frac{1}{n}\sum_{i=1}^{n} \inf_{\boldsymbol{\alpha} \in A, \boldsymbol{\gamma} \in \Gamma} \inf_{t > T_0} \rho\left(y_i - t\boldsymbol{\gamma}'\mathbf{h}(x_i, \boldsymbol{\alpha})\right) \\
&= \min_{j=1,\ldots,N} \frac{1}{n}\sum_{i=1}^{n} \inf_{(\boldsymbol{\alpha}, \boldsymbol{\gamma}) \in U_j} \inf_{t > T_0} \rho\left(y_i - t\boldsymbol{\gamma}'\mathbf{h}(x_i, \boldsymbol{\alpha})\right),
\end{aligned}$$

y por lo tanto, (40) y la Ley de Grandes Números implican

$$\liminf_{n\to\infty} \inf_{\|\boldsymbol{\beta}\|>T_0} \inf_{\boldsymbol{\alpha}\in A} \lambda_n(\boldsymbol{\alpha},\boldsymbol{\beta}) \geq \lambda_{\text{mín}} + \frac{\xi}{4} \text{ c.s.},$$

pero también,

$$\lambda_n\left(\hat{\boldsymbol{\alpha}}_n, \hat{\boldsymbol{\beta}}_n\right) = \inf_{\boldsymbol{\beta}} \inf_{\boldsymbol{\alpha}\in A} \lambda_n(\boldsymbol{\alpha},\boldsymbol{\beta}) \leq \lambda_n(\boldsymbol{\alpha}_0,\boldsymbol{\beta}_0) \to \lambda_{\text{mín}} \text{ c.s.}$$

con lo cual se concluye que para n grande, $\left\|\hat{\boldsymbol{\beta}}_n\right\| \leq T_0$ con probabilidad 1.

Supongamos ahora que $\sup_u \rho(u) = \infty$. Entonces un procedimiento similar, pero más simple, muestra la existencia de T_0 y entornos $U(\boldsymbol{\alpha},\boldsymbol{\gamma})$ tales que el miembro izquierdo de (40) es mayor que $2\lambda_{\text{mín}}$, y el resto de la prueba es similar. ■

Demostración del teorema 3.3. Si A no es compacto, usamos el mismo enfoque de (Richardson y Bhattacharyya, 1986): la compactificación de Stone- Čech nos da un conjunto compacto $\tilde{A} \supset A$ tal que toda función continua y acotada sobre A tiene una única extensión continua en $\tilde{A}$.

Por lo tanto, podemos aplicar el lema para concluir que $\left(\hat{\boldsymbol{\alpha}}_n, \hat{\boldsymbol{\beta}}_n\right)$ para n grande permanece en un compacto c.s.

El teorema se sigue del teorema 1 de Huber (1967). ■

Demostración del teorema 3.4. De acuerdo al teorema 3 y su corolario de Huber (1967) basta con ver que se cumplen las condiciones N1 a N4 siguientes para poder concluir que el estimador $\boldsymbol{\theta}_n$ es asintóticamente normal.

Sea $\Psi(\mathbf{z},\boldsymbol{\theta}) = \Psi(\mathbf{y},\mathbf{x},\boldsymbol{\theta}) = \psi(\mathbf{y} - g(\mathbf{x},\boldsymbol{\theta}))\,\dot{\mathbf{g}}(\mathbf{x},\boldsymbol{\theta})$

N1) Para cada $\boldsymbol{\theta} \in \Theta$, $\Psi(\mathbf{z},\boldsymbol{\theta})$ es $\mathcal{U}-$medible y $\Psi(\mathbf{z},\boldsymbol{\theta})$ es separable en el sentido de Doob: existe un conjunto N de H_0 medida nula y un subconjunto numerable $\Theta' \subset \Theta$ tales que para todo abierto $U \subset \Theta$ y todo intervalo cerrado B, los conjuntos

$$\{\mathbf{z}/\ \Psi(\mathbf{z},\boldsymbol{\theta}) \in B\ \forall\boldsymbol{\theta} \in U\} \quad \text{y} \quad \{\mathbf{z}/\ \Psi(\mathbf{z},\boldsymbol{\theta}) \in B\ \forall\boldsymbol{\theta} \in U \cap \Theta'\}$$

difieren a lo sumo en un subconjunto de N.

Sea $\phi(\boldsymbol{\theta}) = E\Psi(\mathbf{z},\boldsymbol{\theta})$

$$u(\mathbf{z},\boldsymbol{\theta},d) = \sup_{\|\boldsymbol{\tau}-\boldsymbol{\theta}\|\leq d} \|\Psi(\mathbf{z},\boldsymbol{\tau}) - \Psi(\mathbf{z},\boldsymbol{\theta})\|$$

N2) Existe $\boldsymbol{\theta}_0 \in \Theta$ tal que $\phi(\boldsymbol{\theta}_0) = 0$

N3) Existen a, b, c, $d_0 \in \mathbb{R}$ estrictamente positivas tales que:

- **i)** $\|\phi(\boldsymbol{\theta})\| \geq a.\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|$ para $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq d_0$
- **ii)** $Eu(\mathbf{z}, \boldsymbol{\theta}, d) \leq b.d$ para $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| + d \leq d_0, \ \ d \geq 0$
- **iii)** $Eu^2(\mathbf{z}, \boldsymbol{\theta}, d) \leq c.d$ para $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| + d \leq d_0, \ \ d \geq 0$

N4) $E\|\Psi(\mathbf{z}, \boldsymbol{\theta})\|^2 < \infty$

Sólo nos falta verificar que N3) se cumple.

i) Como $\phi(\boldsymbol{\theta}_0) = 0$,

$$\begin{aligned} \|\phi(\boldsymbol{\theta})\| &= \|\phi(\boldsymbol{\theta}) - \phi(\boldsymbol{\theta}_0)\| = \|E(\Psi(\mathbf{z}, \boldsymbol{\theta}) - \Psi(\mathbf{z}, \boldsymbol{\theta}_0))\| \\ &\geq a.\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \end{aligned}$$

para $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \leq d_0$, para cierto d_0 y siendo a por ejemplo: $\frac{1}{2}E\left\|\frac{\partial}{\partial\boldsymbol{\theta}}\Psi(\mathbf{z}, \boldsymbol{\theta}_1)\right\|$ con $\boldsymbol{\theta}_1$ entre $\boldsymbol{\theta}$ y $\boldsymbol{\theta}_0$.

ii) Como $\dfrac{\partial}{\partial\boldsymbol{\theta}}\psi$ es acotada y g tiene derivadas F_0-integrables, entonces

$$\begin{aligned} E\sup_{\|\boldsymbol{\tau}-\boldsymbol{\theta}\|}\|\Psi(\mathbf{z}, \boldsymbol{\tau}) - \Psi(\mathbf{z}, \boldsymbol{\theta})\| &\leq E\left(\sup_{\|\boldsymbol{\tau}-\boldsymbol{\theta}\|\leq d}\left\|\frac{\partial}{\partial\boldsymbol{\theta}}\Psi\left(\mathbf{z}, \tilde{\boldsymbol{\theta}}\right)\right\|\|\boldsymbol{\tau} - \boldsymbol{\theta}\|\right) \\ &\leq E\left(\sup_{\|\boldsymbol{\tau}-\boldsymbol{\theta}\|\leq d}\left\|\frac{\partial}{\partial\boldsymbol{\theta}}\Psi\left(\mathbf{z}, \tilde{\boldsymbol{\theta}}\right)\right\|\right).d \leq b.d \end{aligned}$$

para cierto $b > 0$ y cierto $\tilde{\boldsymbol{\theta}}$ entre $\boldsymbol{\tau}$ y $\boldsymbol{\theta}$.

iii) Análogo a ii) pero con $\sup_{\|\boldsymbol{\tau}-\boldsymbol{\theta}\|\leq d}\left\|\dfrac{\partial}{\partial\boldsymbol{\theta}}\Psi\left(\mathbf{z}, \tilde{\boldsymbol{\theta}}\right)\right\|^2$.

Quedando así probado el teorema. ■

5. Demostraciones y definiciones para la Sección 3.2.

Para analizar el punto de ruptura de los MM-estimadores hemos utilizado la versión dada por Stromberg y Ruppert (1992).

Supongamos que

$$\sup_{\boldsymbol{\theta}} g(\mathbf{x},\boldsymbol{\theta}) = +\infty \text{ y que } \inf_{\boldsymbol{\theta}} g(\mathbf{x},\boldsymbol{\theta}) = -\infty$$

Sea $\mathbf{Z}_n = (\mathbf{z}_1, ..., \mathbf{z}_n)$ una muestra de tamaño n, $\hat{\boldsymbol{\theta}}(\mathbf{Z}_n)$ el estimador $\hat{\boldsymbol{\theta}}$ evaluado en $\mathbf{Z}_n$ y D^m el conjunto de todos los conjuntos de datos, $\mathbf{y}$, de tamaño n que tienen en común $n-m$ elementos con $\mathbf{Z}_n$, esto es:

$$D^m = \{\mathbf{y} : \#(\mathbf{y}) = n,\ \#(\mathbf{Z}_n \cap \mathbf{y}) = n-m\}$$

Entonces definimos las *funciones de ruptura superior e inferior* como:

$$\varepsilon^*_{n,+}\left(\mathbf{x}, g, \hat{\boldsymbol{\theta}}, \mathbf{Z}_n\right) = \min_{0\le m\le n}\left\{\frac{m}{n} : \sup_{\mathbf{y}\in D^m} g\left(\mathbf{x}, \hat{\boldsymbol{\theta}}(\mathbf{y})\right) = \infty\right\}$$

$$\varepsilon^*_{n,-}\left(\mathbf{x}, g, \hat{\boldsymbol{\theta}}, \mathbf{Z}_n\right) = \min_{0\le m\le n}\left\{\frac{m}{n} : \inf_{\mathbf{y}\in D^m} g\left(\mathbf{x}, \hat{\boldsymbol{\theta}}(\mathbf{y})\right) = -\infty\right\}.$$

Se define, además, la *función de ruptura* como

$$\varepsilon^*_{n}\left(\mathbf{x}, g, \hat{\boldsymbol{\theta}}, \mathbf{Z}_n\right) = \min\left\{\varepsilon^*_{n,+}\left(\mathbf{x}, g, \hat{\boldsymbol{\theta}}, \mathbf{Z}_n\right), \varepsilon^*_{n,-}\left(\mathbf{x}, g, \hat{\boldsymbol{\theta}}, \mathbf{Z}_n\right)\right\}.$$

Finalmente, el *punto de ruptura* se define como

$$\varepsilon^*_{n}\left(g, \hat{\boldsymbol{\theta}}, \mathbf{Z}_n\right) = \inf_{\mathbf{x}\in\mathbb{R}^p} \varepsilon^*_{n}\left(\mathbf{x}, g, \hat{\boldsymbol{\theta}}, \mathbf{Z}_n\right).$$

Sean $\mathbf{z}_1 = (y_1, \mathbf{x}_1), ..., \mathbf{z}_n = (y_n, \mathbf{x}_n), ...$ observaciones iid. Utilizaremos la siguiente versión asintótica de la robustez cualitativa.

Definición 5.1 $\{\mathbf{T}_n\}_{n\ge p}$ *es* asintóticamente cualitativamente robusta *en* H_0 *si dado* $\varepsilon > 0$, $\exists\delta > 0$ *y* n_0 *tal que* $\Pi_{p+1}(H^*, H_0) < \delta$ *y* $n > n_0$, *implican* $\Pi_p(L(T_n, H^*), L(T_n, H_0)) > \varepsilon$, *siendo* $L(\mathbf{T},H)$ *la distribución de* $\mathbf{T}$ *bajo* H *y* Π_j *la métrica de Prohorov entre distribuciones en* $\mathbb{R}^j$.

Papantoni-Kazakos y Gray (1979), demostraron que para la robustez cualitativamente asintótica en H_0 es suficiente con probar que:

- **T** es continua en H_0 como funcional definida en el espacio de distribuciones en $\mathbb{R}^{p+1}$ dotado de la métrica Π_{p+1} y con valores en $\mathbb{R}^p$.

Para las demostraciones de los resultados supondremos, sin pérdida de generalidad, que $P\left(g\left(\mathbf{x}, \boldsymbol{\theta}_0\right) = 0\right) = 1$.

Demostración del teorema 3.5. Sea $\varepsilon > 0$, como ρ_0 es una función-ρ, podemos hallar $\delta > 0$ tal que

$$E_{H_0}\left(\inf_{\|\theta - \boldsymbol{\theta_0}\| \leq \delta} \rho_0\left(\frac{y - g(\mathbf{x}, \boldsymbol{\theta})}{\sigma_0 - \varepsilon}\right)\right) \geq b + \delta$$

y

$$E_{H_0}\left(\sup_{|||\theta - \boldsymbol{\theta_0}||| \leq \delta} \rho_0\left(\frac{y - g(\mathbf{x}, \boldsymbol{\theta})}{\sigma_0 + \varepsilon}\right)\right) \leq b - \delta$$

por lo tanto, por la Ley de Grandes Números se tiene

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \inf_{\|\theta - \boldsymbol{\theta_0}\| \leq \delta} \rho_0\left(\frac{y_i - g(\mathbf{x}_i, \boldsymbol{\theta})}{\sigma_0 - \varepsilon}\right) \geq b + \delta \qquad c.s.$$

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \sup_{\|\theta - \boldsymbol{\theta_0}\| \leq \delta} \rho_0\left(\frac{y_i - g(\mathbf{x}_i, \boldsymbol{\theta})}{\sigma_0 + \varepsilon}\right) \leq b - \delta \qquad c.s.$$

Entonces, como $\lim_{n \to \infty} \mathbf{T}_{0,n} = \theta_0 \;\; c.s.$, se tiene

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \rho_0\left(\frac{y_i - g(\mathbf{x}_i, \mathbf{T}_{0,n})}{\sigma_0 - \varepsilon}\right) \geq b + \delta \qquad c.s.$$

y

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \rho_0\left(\frac{y_i - g(\mathbf{x}_i, \mathbf{T}_{0,n})}{\sigma_0 + \varepsilon}\right) \leq b - \delta \qquad c.s.$$

Por lo tanto como ρ_0 es monótona, con probabilidad uno,

$$\exists\; n_0 \text{ tal que } \forall n \geq n_0,\; \sigma_0 - \varepsilon \leq s_n \leq \sigma_0 + \varepsilon$$

es decir s_n es consistente para σ_0. ■

Lema 5.1 *Supongamos que F_0 satisface (B) y que G_0 satisface (C). Definimos para $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, $h^*(\boldsymbol{\theta}) = E_{H_0}(\rho(y - g(\mathbf{x}, \boldsymbol{\theta})))$. Entonces h^* tiene un único mínimo en $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.*

Demostración. Por el lema 3.1 de Yohai (1987), h^* tiene único mínimo en $g(\mathbf{x}, \boldsymbol{\theta}) = \mathbf{0}$ y como $\forall \varepsilon > 0, \exists \delta > 0$ tal que $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| > \delta$ implica

$$P\{\mathbf{x} : |g(\mathbf{x}, \boldsymbol{\theta}) - g(\mathbf{x}, \boldsymbol{\theta}_0)| > \varepsilon\} \geq 0{,}5,$$

entonces $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ y h^* tiene único mínimo en $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. ■

Demostración del teorema 3.6. Se quiere probar que $\mathbf{T}_{1,n} \to \boldsymbol{\theta}_0$ *c.s.*

Veamos que basta con probar que dado $\varepsilon > 0$, $\exists\ \gamma > 0$ y $\sigma_1 > \sigma_0$ tales que

$$\liminf_{n\to\infty} \inf_{\boldsymbol{\theta}\in C_\varepsilon} \frac{1}{n} \sum_{i=1}^{n} \rho_1\left(\frac{y_i - g(\mathbf{x}_i, \boldsymbol{\theta})}{\sigma_1}\right) \geq E_{F_0}\left(\rho_1\left(\frac{u}{\sigma_0}\right)\right) + \gamma \quad c.s. \tag{41}$$

y

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \rho_1\left(\frac{y_i - g(\mathbf{x}_i, \boldsymbol{\theta})}{s_n}\right) = E_{F_0}\left(\rho_1\left(\frac{u}{\sigma_0}\right)\right) \quad c.s. \tag{42}$$

siendo $C_\varepsilon = \{\boldsymbol{\theta} : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \geq \epsilon\}$.

Supongamos probado esto, entonces

$$\liminf_{n\to\infty} \inf_{\boldsymbol{\theta}\in C_\varepsilon} \frac{1}{n} \sum_{i=1}^{n} \rho_1\left(\frac{y_i - g(\mathbf{x}_i, \boldsymbol{\theta})}{\sigma_1}\right) > \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \rho_1\left(\frac{y_i - g(\mathbf{x}_i, \mathbf{T}_{0,n})}{s_n}\right)$$

$$\geq \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \rho_1\left(\frac{y_i - g(\mathbf{x}_i, \mathbf{T}_{1,n})}{s_n}\right)$$

por (9); y como

$$\rho_1\left(\frac{y_i - g(\mathbf{x}_i, \boldsymbol{\theta})}{\sigma_1}\right) \leq \rho_1\left(\frac{y_i - g(\mathbf{x}_i, \boldsymbol{\theta})}{\sigma_0}\right)$$

entonces

$$\liminf_{n\to\infty} \inf_{\boldsymbol{\theta}\in C_\varepsilon} \frac{1}{n} \sum_{i=1}^{n} \rho_1\left(\frac{y_i - g(\mathbf{x}_i, \boldsymbol{\theta})}{\sigma_0}\right) \geq \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \rho_1\left(\frac{y_i - g(\mathbf{x}_i, \mathbf{T}_{1,n})}{s_n}\right)$$

y por el teorema 3.5, s_n es fuertemente consistente para σ_0 entonces se concluye que $\mathbf{T}_{1,n} \notin C_\varepsilon$ con lo cual $\|\mathbf{T}_{1,n} - \boldsymbol{\theta}_0\| < \varepsilon$, lo que implica la consistencia de $\mathbf{T}_{1,n}$.

Sólo resta probar (41) y (42).

Sea $\{C_{\boldsymbol{\theta}}\}_{\boldsymbol{\theta}}$ un cubrimiento por compactos de Θ tales que

$$\inf_{\boldsymbol{\theta} \in C_{\boldsymbol{\theta}}} \rho_1\left(\frac{y - g(\mathbf{x}, \boldsymbol{\theta})}{\sigma_1}\right) \underset{\boldsymbol{\theta} \to \boldsymbol{\theta}_0}{\longrightarrow} \rho_1\left(\frac{u}{\sigma_1}\right).$$

Además, como ρ_1 es acotada se tiene

$$\left|\inf_{\boldsymbol{\theta} \in C_{\boldsymbol{\theta}}} \rho_1\left(\frac{y - g(\mathbf{x}, \boldsymbol{\theta})}{\sigma_1}\right)\right| \leq 1, \forall \boldsymbol{\theta}$$

entonces por el Teorema de la Convergencia Dominada

$$E_{H_0}\left(\inf_{\boldsymbol{\theta} \in C_{\boldsymbol{\theta}}} \rho_1\left(\frac{y - g(\mathbf{x}, \boldsymbol{\theta})}{\sigma_1}\right)\right) \underset{\boldsymbol{\theta} \to \boldsymbol{\theta}_0}{\longrightarrow} E_{F_0}\rho_1\left(\frac{u}{\sigma_1}\right)$$

Por el lema 5.1, $\rho_1\left(\dfrac{y - g(\mathbf{x}, \boldsymbol{\theta})}{\sigma_1}\right)$ tiene único mínimo en $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Por lo tanto,

$$E_{H_0}\left(\inf_{\boldsymbol{\theta} \in C_{\boldsymbol{\theta}}} \rho_1\left(\frac{y - g(\mathbf{x}, \boldsymbol{\theta})}{\sigma_1}\right)\right) > E_{F_0}\rho_1\left(\frac{u}{\sigma_1}\right) \ , \forall \boldsymbol{\theta}$$

Como $C_\varepsilon \subset \Theta$ y es compacto, existe un subcubrimiento finito de él $\{C_i\}_{i=1}^{r}$,

$$E_{H_0}\left(\inf_{\boldsymbol{\theta} \in C_{\mathbf{i}}} \rho_1\left(\frac{y - g(\mathbf{x}, \boldsymbol{\theta})}{\sigma_1}\right)\right) > E_{F_0}\rho_1\left(\frac{u}{\sigma_1}\right)$$

entonces para cierto $\gamma_1 > 0$,

$$E_{H_0}\left(\inf_{\boldsymbol{\theta} \in C_{\mathbf{i}}} \rho_1\left(\frac{y - g(\mathbf{x}, \boldsymbol{\theta})}{\sigma_1}\right)\right) \geq E_{F_0}\rho_1\left(\frac{u}{\sigma_1}\right) + \gamma_1$$

Luego, elegimos $\sigma_1 > \sigma_0$ tal que

$$\left|E_{F_0}\rho_1\left(\frac{u}{\sigma_0}\right) - E_{F_0}\rho_1\left(\frac{u}{\sigma_1}\right)\right| < \frac{\gamma_1}{2}$$

y por lo tanto,

$$E_{H_0}\left(\inf_{\boldsymbol{\theta}\in C_{\mathbf{i}}}\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_1}\right)\right) \tag{43}$$

$$\geq E_{F_0}\rho_1\left(\frac{u}{\sigma_1}\right)+E_{F_0}\rho_1\left(\frac{u}{\sigma_0}\right)-E_{F_0}\rho_1\left(\frac{u}{\sigma_1}\right)+\frac{\gamma_1}{2} \tag{44}$$

$$= E_{F_0}\rho_1\left(\frac{u}{\sigma_0}\right)+\gamma$$

con $\gamma=\dfrac{\gamma_1}{2}$.

Luego, como $C_\varepsilon\subset\bigcup_{i=1}^{r}C_i$

$$\begin{aligned}&\lim\inf_{n\to\infty}\inf_{\boldsymbol{\theta}\in C_\varepsilon}\frac{1}{n}\sum_{i=1}^{n}\rho_1\left(\frac{y_i-g(\mathbf{x}_i,\boldsymbol{\theta})}{\sigma_1}\right)\\ &\geq \min_{1\leq i\leq r}\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\inf_{\boldsymbol{\theta}\in C_i}\rho_1\left(\frac{y_i-g(\mathbf{x}_i,\boldsymbol{\theta})}{\sigma_1}\right),\end{aligned}$$

por la Ley de Grandes Números, el miembro derecho es

$$\min_{1\leq i\leq r}E_{H_0}\left(\inf_{\boldsymbol{\theta}\in C_{\mathbf{i}}}\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_1}\right)\right)$$

y, por (43), se tiene

$$\lim\inf_{n\to\infty}\inf_{\boldsymbol{\theta}\in C_\varepsilon}\frac{1}{n}\sum_{i=1}^{n}\rho_1\left(\frac{y_i-g(\mathbf{x}_i,\boldsymbol{\theta})}{\sigma_1}\right)\geq E_{F_0}\rho_1\left(\frac{u}{\sigma_0}\right)+\gamma$$

quedando así probado (41).

Finalmente, para probar (42), consideramos

$$h(\mathbf{z},\mathbf{T}_{0,n})=\rho_1\left(\frac{y-g(\mathbf{x},\mathbf{T}_{0,n})}{s_n}\right)$$

Como $\mathbf{T}_{0,n}\to\boldsymbol{\theta}_0$ *c.s.*y $s_n\to\sigma_0$ *c.s.*, entonces por el lema 4.2 de Yohai (1987), se tiene (42). ■

Observación 5.1 *Si el espacio de parámetros Θ no es compacto, pero la función g cumple la condición (18), se procede como se hizo en la sección anterior para los M-estimadores, trabajando con la compactificación de Stone-Čech.*

Para demostrar el teorema 3.7 se necesita el siguiente lema.

Lema 5.2

$$p\lim \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\Psi\left(\mathbf{x}_i, y_i, \boldsymbol{\theta}_0, s_n\right) - \Psi\left(\left(\mathbf{x}_i, y_i, \boldsymbol{\theta}_0, \sigma_0\right)\right)\right] = 0$$

con

$$\Psi\left(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta}, \sigma\right) = \psi_1\left(\frac{\mathbf{y} - g\left(\mathbf{x}, \boldsymbol{\theta}\right)}{\sigma}\right) \dot{\mathbf{g}}\left(\mathbf{x}, \boldsymbol{\theta}\right)$$

siendo $\dot{\mathbf{g}}\left(\mathbf{x}, \boldsymbol{\theta}\right) = \dfrac{\partial \mathbf{g}\left(\mathbf{x}, \boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}}$ *y* $p\lim$ *el límite en probabilidad.*

Demostración. Sea para $0{,}5\sigma_0 \leq t \leq 1{,}5\sigma_0$

$$A_{n,j}(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Psi\left(x_{ij}, y_i, \boldsymbol{\theta}_0, t\right)$$

$A_{n,j} \in C_{[0{,}5\sigma_0, 1{,}5\sigma_0]}$. Como $\lim_{n\to\infty} s_n = \sigma_0$ *c.s.*, para probar el lema basta con mostrar que $A_{n,j}(t)$ es tight para $1 \leq j \leq p$.

Supongamos esto probado, queremos ver que $A_{n,j}(s_n) - A_{n,j}(\sigma_0) \xrightarrow{p} 0$, es decir $\forall \epsilon, \eta > 0$, $\exists N$ tal que

$$n \geq N \Rightarrow P\left(\|A_{n,j}(s_n) - A_{n,j}(\sigma_0)\| \geq \epsilon\right) \leq \eta \tag{45}$$

Consideremos $\epsilon, \eta > 0$ dados. Como $s_n \xrightarrow{p} \sigma_0$, $\exists n_0$ tal que

$$n \geq n_0 \Rightarrow P\left(|s_n - \sigma_0| > 0{,}5\sigma_0\right) < \frac{\eta}{2}$$

por lo tanto para $n \geq n_0$, $P\left(s_n \in [0{,}5\sigma_0, 1{,}5\sigma_0]\right) \geq 1 - \dfrac{\eta}{2}$.

Llamemos $a = 0{,}5\sigma_0$ y $b = 1{,}5\sigma_0$. Definimos

$$w_n\left(\delta\right) = \sup_{\substack{|s-t|<\delta \\ s,t\in[a,b]}} \left\|A_{n,j}\left(s\right) - A_{n,j}\left(t\right)\right\|$$

Como suponemos probado $A_{n,j}$ es tight, el teorema 8.2 de Billingsley implica que $\exists \delta_1, n_1$ tal que

$$n \geq n_1 \Rightarrow P\left(w_n\left(\delta_1\right) \geq \varepsilon\right) \leq \frac{\eta}{2} \tag{46}$$

Sea $\delta = \text{mín}\left(\delta_1, a\right)$. Como $s_n \xrightarrow{p} \sigma_0$, $\exists n_2 \geq n_0$ tal que

$$n \geq n_2 \Rightarrow P\left(\left|s_n - \sigma_0\right| > \delta\right) \leq \frac{\eta}{2} \tag{47}$$

Además, $s_n \in [a, b] \Rightarrow$

$$\left\|A_{n,j}\left(s_n\right) - A_{n,j}\left(\sigma_0\right)\right\| \leq w_n\left(\left|s_n - \sigma_0\right|\right) \tag{48}$$

Finalmente, sea $N = \text{máx}\left(n_1, n_2\right)$, entonces uniendo (48), (46) y (47) para $n \geq N$

$$\begin{aligned}
&P\left(\left\|A_{n,j}(s_n) - A_{n,j}(\sigma_0)\right\| \geq \epsilon\right) \\
&= P\left(\left\|A_{n,j}(s_n) - A_{n,j}(\sigma_0)\right\| \geq \epsilon \wedge \left|s_n - \sigma_0\right| < \delta\right) \\
&+ P\left(\left\|A_{n,j}(s_n) - A_{n,j}(\sigma_0)\right\| \geq \epsilon \wedge \left|s_n - \sigma_0\right| \geq \delta\right) \\
&\leq P\left(w_n\left(\left|s_n - \sigma_0\right|\right) \geq \varepsilon\right) + P\left(\left|s_n - \sigma_0\right| \geq \delta\right) \leq \frac{\eta}{2} + \frac{\eta}{2} = \eta
\end{aligned}$$

Por lo tanto, $A_{n,j}\left(s_n\right) - A_{n,j}\left(\sigma_0\right) \xrightarrow{p} 0$ como se quería probar.

Sólo resta probar que $A_{n,j}(t)$ es "*tight*". Usando el teorema 12.3 de Billingsley (1968), será suficiente mostrar que se cumplen las siguientes dos condiciones:

1. $A_{n,j}(a)$ es tight
2. $\exists \gamma > 0, \alpha > 0$ y una función R monótona , continua sobre $[a, b]$ tal que $\forall a \leq t_1 \leq t_2 \leq b$ y $\forall \lambda > 0$, se tiene

$$P_{H_0}\left(\left|A_{n,j}(t_1) - A_{n,j}(t_2)\right| \geq \lambda\right) \leq \left(\frac{1}{\lambda}\right)^{\gamma}\left(R\left(t_2\right) - R\left(t_1\right)\right)^{\alpha}$$

La condición 1 es válida por el Teorema Central del Límite.

Para la condición 2, observamos que, por la desigualdad de Chebychev, es suficiente con probar que existe una constante K tal que

$$E_{H_0}\left(A_{n,j}(t_1) - A_{n,j}(t_2)\right)^2 \leq K\left(t_1 - t_2\right)^2 \tag{49}$$

Tenemos

$$
\begin{aligned}
&E_{H_0}\left(A_{n,j}(t_1) - A_{n,j}(t_2)\right)^2 \\
&\leq E_{G_0}\left(\dot{\mathbf{g}}\left(\mathbf{x}_{ij}, \boldsymbol{\theta}\right) \dot{\mathbf{g}}\left(\mathbf{x}_{ij}, \boldsymbol{\theta}\right)'\right) E_{F_0}\left[\psi_1\left(\frac{y}{t_2}\right) - \psi_1\left(\frac{y}{t_1}\right)\right]^2 \qquad (50) \\
&= V E_{F_0}\left[\psi_1\left(\frac{y}{t_2}\right) - \psi_1\left(\frac{y}{t_1}\right)\right]^2 \qquad (51)
\end{aligned}
$$

De acuerdo a la suposición (D), sea $K_0 = \sup|\psi_1'(u)|$. Entonces, como $|u| > m$ implica $\psi_1(u) = 0$, se tiene

$$
\begin{aligned}
\left[\psi_1\left(\frac{y}{t_2}\right) - \psi_1\left(\frac{y}{t_1}\right)\right]^2 &\leq \psi_1'\left(\frac{y}{t^*}\right)^2\left(\frac{y}{(t^*)^2}\right)^2 \sigma_0^2\left(t_1 - t_2\right)^2 \\
&\leq K_0^2\sigma_0^2\left(\frac{y^2}{\left(0{,}5\sigma_0\right)^4} I\left(|y| \leq 1{,}5m\sigma_0\right)\right)\left(t_1 - t_2\right)^2 \\
&= \frac{K_0^2 16\left(t_1 - t_2\right)^2}{\sigma_0^2} y^2 I\left(|y| \leq 1{,}5m\sigma_0\right)
\end{aligned}
$$

con $t^* \in (t_1, t_2)$. Por lo tanto,

$$
\begin{aligned}
E_{F_0}\left[\psi_1\left(\frac{y}{t_2}\right) - \psi_1\left(\frac{y}{t_1}\right)\right]^2 &\leq \frac{K_0^2 16\left(t_1 - t_2\right)^2}{\sigma_0^2} E_{F_0}\left(y^2 I\left(|y| \leq 1{,}5m\sigma_0\right)\right) \qquad (52) \\
&\leq \frac{K_0^2 16\left(t_1 - t_2\right)^2}{\sigma_0^2}\left(1{,}5m\sigma_0\right)^2 = 36K_0^2 m^2\left(t_1 - t_2\right)^2
\end{aligned}
$$

Luego, juntando (51) y (52), se tiene (49), quedando así probado el lema. ■

Ahora estamos en condiciones de demostrar el teorema 3.7.

Demostración del teorema 3.7. Por los teoremas 3.5 y 3.6, $\text{lím}\, s_n = \sigma_0$ *c.s.* y $\text{lím}\, \mathbf{T}_{1,n} = \boldsymbol{\theta}_0$ *c.s.*

Además, por (8)

$$
\sum_{i=1}^{n} \psi_1\left(\frac{y - g\left(\mathbf{x}, \mathbf{T}_{1,n}\right)}{s_n}\right) \dot{\mathbf{g}}\left(\mathbf{x}, \mathbf{T}_{1,n}\right) = 0.
$$

Por el Teorema del Valor Medio,

$$\psi_1\left(\frac{y-g\left(\mathbf{x},\mathbf{T}_{1,n}\right)}{s_n}\right)-\psi_1\left(\frac{y-g\left(\mathbf{x},\boldsymbol{\theta}_0\right)}{s_n}\right)$$
$$=\psi_1'\left(\frac{y-g\left(\mathbf{x},\boldsymbol{\theta}_n^*\right)}{s_n}\right)\frac{\dot{\mathbf{g}}\left(\mathbf{x},\boldsymbol{\theta}_n^*\right)}{s_n}\left(\mathbf{T}_{1,n}-\boldsymbol{\theta}_0\right)$$

con $\boldsymbol{\theta}_n^*$ entre $\mathbf{T}_{1,n}$ y $\boldsymbol{\theta}_0$. Luego,

$$\frac{1}{n}\sum_{i=1}^{n}\psi_1\left(\frac{u}{s_n}\right)\dot{\mathbf{g}}\left(\mathbf{x},\mathbf{T}_{1,n}\right)$$
$$=\frac{\mathbf{T}_{1,n}-\boldsymbol{\theta}_0}{s_n}\frac{1}{n}\sum_{i=1}^{n}\psi_1'\left(\frac{y-g\left(\mathbf{x},\boldsymbol{\theta}_n^*\right)}{s_n}\right)\dot{\mathbf{g}}\left(\mathbf{x},\boldsymbol{\theta}_n^*\right)\dot{\mathbf{g}}\left(\mathbf{x},\mathbf{T}_{1,n}\right)'$$

es decir,

$$\frac{1}{\sqrt{n}}\mathbf{a}_n=\frac{\mathbf{T}_{1,n}-\boldsymbol{\theta}_0}{s_n}\mathbf{w}_n$$

donde

$$\frac{1}{\sqrt{n}}\mathbf{a}_n=\frac{1}{n}\sum_{i=1}^{n}\psi_1\left(\frac{u}{s_n}\right)\dot{\mathbf{g}}\left(\mathbf{x},\mathbf{T}_{1,n}\right)$$
$$\text{y}\quad \mathbf{w}_n=\frac{1}{n}\sum_{i=1}^{n}\psi_1'\left(\frac{y-g\left(\mathbf{x},\boldsymbol{\theta}_n^*\right)}{s_n}\right)\dot{\mathbf{g}}\left(\mathbf{x},\boldsymbol{\theta}_n^*\right)\dot{\mathbf{g}}\left(\mathbf{x},\mathbf{T}_{1,n}\right)'$$

luego,

$$\mathbf{a}_n\ \mathbf{w}_n^{-1}s_n=\sqrt{n}\left(\mathbf{T}_{1,n}-\boldsymbol{\theta}_0\right)$$

y

$$\text{lím}\,\boldsymbol{\theta}_n^*=\boldsymbol{\theta}_0\ c.s.$$

por el lema 4.2 Yohai (1987),

$$\text{lím}\,\mathbf{w}_n=E_{H_0}\left[\psi_1'\left(\frac{u}{\sigma_0}\right)\dot{\mathbf{g}}\left(\mathbf{x},\boldsymbol{\theta}_0\right)\dot{\mathbf{g}}\left(\mathbf{x},\boldsymbol{\theta}_0\right)'\right]\ c.s.$$

$$E_{H_0}\left[\psi_1'\left(\frac{u}{\sigma_0}\right)\dot{\mathbf{g}}\left(\mathbf{x},\boldsymbol{\theta}_0\right)\dot{\mathbf{g}}\left(\mathbf{x},\boldsymbol{\theta}_0\right)'\right]=E_{F_0}\psi_1'\left(\frac{u}{\sigma_0}\right)E_{G_0}\left[\dot{\mathbf{g}}\left(\mathbf{x},\boldsymbol{\theta}_0\right)\dot{\mathbf{g}}\left(\mathbf{x},\boldsymbol{\theta}_0\right)'\right]$$
$$=E_{F_0}\psi_1'\left(\frac{u}{\sigma_0}\right)V=\left[B(\psi_1,F_0)\right]^{1/2}V$$

Luego, $\lim \mathbf{w}_n = [B(\psi_1, F_0)]^{1/2} V \;\; c.s.$

Por otro lado, como $E_H \left[\psi_1 \left(\dfrac{u}{\sigma_0} \right) \dot{\mathbf{g}}(\mathbf{x}, \boldsymbol{\theta}_0) \right] = 0$ entonces

$$
\begin{aligned}
\mathbf{Var}_{H_0} \left[\psi_1 \left(\frac{u}{\sigma_0} \right) \dot{\mathbf{g}}(\mathbf{x}, \boldsymbol{\theta}_0) \right] &= E_{H_0} \left[\psi_1^2 \left(\frac{u}{\sigma_0} \right) \dot{\mathbf{g}}(\mathbf{x}, \boldsymbol{\theta}_0) \, \dot{\mathbf{g}}(\mathbf{x}, \boldsymbol{\theta}_0)' \right] \\
&= E_{F_0} \psi_1^2 \left(\frac{u}{\sigma_0} \right) E_{G_0} \left[\dot{\mathbf{g}}(\mathbf{x}, \boldsymbol{\theta}_0) \, \dot{\mathbf{g}}(\mathbf{x}, \boldsymbol{\theta}_0)' \right] \\
&= E_{F_0} \psi_1^2 \left(\frac{u}{\sigma_0} \right) V = A(\psi_1, F_0) V
\end{aligned}
$$

Finalmente, por el lema de Slutsky

$$
\sqrt{n} \, (\mathbf{T}_n - \boldsymbol{\theta}_0) = \mathbf{a}_n \; \mathbf{w}_n^{-1} s_n \xrightarrow{d} N \left(0, \sigma_0^2 \frac{A(\psi_1, F_0)}{B(\psi_1, F_0)} V^{-1} \right)
$$

■

Para probar el teorema 3.8 es necesario un lema previo.

Lema 5.3 *Supongamos* $\mathbf{T}_0(H)$ *es continua en* H_0 *y* $\mathbf{T}_0(H_0) = \boldsymbol{\theta}_0$*, donde* H_0 *está dada por (14). Supongamos también que* ρ_0 *es una función-*ρ *acotada y* F_0 *satisface (B). Entonces* $s(H)$*, definida por*

$$
E_H \left(\rho_0 \left(\frac{y - \mathbf{g}(\mathbf{x}, \mathbf{T}_0(H))}{s(H)} \right) \right) = b \tag{53}
$$

es continua en H_0*.*

Demostración. Queremos ver que dado $\varepsilon > 0$, $\exists \delta > 0$ tal que

$$
\Pi_{p+1}(H, H_0) < \delta \Rightarrow |s(H) - s(H_0)| < \varepsilon
$$

Sea $\varepsilon > 0$ dado y sea $\sigma_0 = s(H_0)$, es decir

$$
E_{H_0} \left(\rho_0 \left(\frac{y - \mathbf{g}(\mathbf{x}, \mathbf{T}_0(H_0))}{\sigma_0} \right) \right) = b.
$$

Como ρ_0 es creciente, $\mathbf{T}_0(H_0)$ continua y como

$$
\frac{y - \mathbf{g}(\mathbf{x}, \mathbf{T}_0(H_0))}{\sigma_0 - \varepsilon} > \frac{y - \mathbf{g}(\mathbf{x}, \mathbf{T}_0(H_0))}{\sigma_0} > \frac{y - \mathbf{g}(\mathbf{x}, \mathbf{T}_0(H_0))}{\sigma_0 + \varepsilon}
$$

se tiene

$$E_{H_0}\left(\rho_0\left(\frac{y-\mathbf{g}(\mathbf{x},\mathbf{T}_0(H_0))}{\sigma_0+\varepsilon}\right)\right)<b$$

y

$$E_{H_0}\left(\rho_0\left(\frac{y-\mathbf{g}(\mathbf{x},\mathbf{T}_0(H_0))}{\sigma_0-\varepsilon}\right)\right)>b$$

Definimos

$$q_1\left(y,\mathbf{x},\boldsymbol{\gamma}\right)=\sup\left\{\rho_0\left(\frac{y-\mathbf{g}(\mathbf{x},\boldsymbol{\theta})}{\sigma_0+\varepsilon}\right):\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\leq\gamma\right\}$$

$$q_2\left(y,\mathbf{x},\boldsymbol{\gamma}\right)=\text{ínf}\left\{\rho_0\left(\frac{y-\mathbf{g}(\mathbf{x},\boldsymbol{\theta})}{\sigma_0-\varepsilon}\right):\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\leq\gamma\right\}$$

Entonces, por el Teorema de la Convergencia Dominada, $\exists\gamma_0>0$ y $\eta>0$ tal que

$$E_{H_0}q_1\left(y,\mathbf{x},\boldsymbol{\gamma}_0\right)\leq b-\eta$$

$$E_{H_0}q_2\left(y,\mathbf{x},\boldsymbol{\gamma}_0\right)\geq b+\eta$$

Como las funciones q_i son uniformemente continuas y acotadas, podemos hallar δ_1 tal que $\Pi_{p+1}(H,H_0)<\delta_1$ implique

$$E_Hq_1\left(y,\mathbf{x},\boldsymbol{\gamma}_0\right)\leq b-\frac{\eta}{2}$$

$$E_Hq_2\left(y,\mathbf{x},\boldsymbol{\gamma}_0\right)\geq b+\frac{\eta}{2}$$

Sea δ_2 tal que $\Pi_{p+1}(H,H_0)<\delta_2\Rightarrow\|\mathbf{T}_0(H)-\theta_0\|<\gamma_0$. Tomemos $\delta=\text{mín}\left(\delta_1,\delta_2\right)$, entonces si $\Pi_{p+1}(H,H_0)<\delta$ se tiene

$$E_H\left(\rho_0\left(\frac{y-\mathbf{g}(\mathbf{x},\mathbf{T}_0(H))}{\sigma_0+\varepsilon}\right)\right)\leq b-\frac{\eta}{2}$$

y

$$E_H\left(\rho_0\left(\frac{y-\mathbf{g}(\mathbf{x},\mathbf{T}_0(H))}{\sigma_0-\varepsilon}\right)\right)\geq b+\frac{\eta}{2}$$

y por lo tanto $\sigma_0-\varepsilon\leq s(H)\leq\sigma_0+\varepsilon$. ■

Demostración del teorema 3.8. Sea $B_\varepsilon=\{\boldsymbol{\theta}\in\boldsymbol{\Theta}:\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\geq\varepsilon\}$. Como Θ es compacto, B_ε también lo es.

Queremos ver que dado $\varepsilon > 0$, $\exists\delta > 0$ tal que

$$\Pi_{p+1}(H, H_0) < \delta \Rightarrow \|\mathbf{T}_1(H) - \mathbf{T}_1(H_0)\| < \varepsilon.$$

$\mathbf{T}_1(H)$ es el que minimiza $E_H\left(\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{s(H)}\right)\right)$ y por el lema 3.2 de Yohai (1987), $\mathbf{T}_1(H_0) = \boldsymbol{\theta}_0$. Es decir que debemos probar $\|\mathbf{T}_1(H) - \boldsymbol{\theta}_0\| < \varepsilon$.

Si logramos probar que $\forall\varepsilon > 0$, $\exists\sigma_1 > \sigma_0$, $\delta > 0$ y $\eta > 0$ tales que $\Pi_{p+1}(H, H_0) < \delta$ implique

$$\inf_{\boldsymbol{\theta}\in\mathbf{B}_\varepsilon} E_H\left(\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_1}\right)\right) \geq E_{F_0}\rho_1\left(\frac{u}{\sigma_0}\right) + \eta \tag{54}$$

y

$$E_H\left(\rho_1\left(\frac{y-g(\mathbf{x},\mathbf{T}_0\,(H)\,)}{s(H)}\right)\right) \leq E_{F_0}\rho_1\left(\frac{u}{\sigma_0}\right) + \frac{\eta}{2} \tag{55}$$

entonces por (9)

$$E_H\left(\rho_1\left(\frac{y-g(\mathbf{x},\mathbf{T}_1\,(H)\,)}{s(H)}\right)\right) \leq E_H\left(\rho_1\left(\frac{y-g(\mathbf{x},\mathbf{T}_0\,(H)\,)}{s(H)}\right)\right)$$

y por (54) y (55)

$$\begin{aligned}\inf_{\boldsymbol{\theta}\in B_\varepsilon} E_H\left(\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_1}\right)\right) &\geq E_H\left(\rho_1\left(\frac{y-g(\mathbf{x},\mathbf{T}_0\,(H)\,)}{s(H)}\right)\right) + \frac{\eta}{2}\\ &> E_H\left(\rho_1\left(\frac{y-g(\mathbf{x},\mathbf{T}_1\,(H)\,)}{s(H)}\right)\right)\end{aligned}$$

entonces $\mathbf{T}_1(H) \notin B_\varepsilon$ y por lo tanto $\|\mathbf{T}_1(H) - \boldsymbol{\theta}_0\| < \varepsilon$.

Sólo resta probar (54) y (55).

Por ser B_ε compacto, existe un cubrimiento finito por compactos $\{C_i\}_{i=1}^r$ tales que

$$\inf_{\boldsymbol{\theta}\in C_i} \rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_1}\right) \underset{\boldsymbol{\theta}\to\boldsymbol{\theta}_0}{\longrightarrow} \rho_1\left(\frac{u}{\sigma_1}\right)$$

además, por ser ρ acotada

$$\left|\inf_{\boldsymbol{\theta}\in C_i} \rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_1}\right)\right| \leq 1 \quad \forall\boldsymbol{\theta}.$$

Luego, por el Teorema de la Convergencia Dominada

$$E_{H_0}\left(\inf_{\boldsymbol{\theta}\in C_i}\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_1}\right)\right)\underset{\boldsymbol{\theta}\to\boldsymbol{\theta}_0}{\longrightarrow}E_{F_0}\rho_1\left(\frac{u}{\sigma_1}\right)$$

Por el lema 5.1, $\rho_1\left(\dfrac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_1}\right)$ tiene único mínimo en $\boldsymbol{\theta}=\boldsymbol{\theta}_0$ y por lo tanto

$$E_{H_0}\left(\inf_{\boldsymbol{\theta}\in C_i}\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_1}\right)\right)>E_{F_0}\rho_1\left(\frac{u}{\sigma_1}\right)\quad\forall\boldsymbol{\theta}.$$

Entonces elegimos $\sigma_1>\sigma_0$ tal que

$$\left|E_{F_0}\left(\rho_1\left(\frac{u}{\sigma_1}\right)-\rho_1\left(\frac{u}{\sigma_0}\right)\right)\right|<\frac{\gamma_1}{2}$$

entonces

$$E_{H_0}\left(\inf_{\boldsymbol{\theta}\in C_i}\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_1}\right)\right)\geq E_{F_0}\rho_1\left(\frac{u}{\sigma_0}\right)+\gamma$$

con $\gamma=\dfrac{\gamma_1}{2}$.

Sean $q_i\left(y,\mathbf{x}\right)=\inf_{\boldsymbol{\theta}\in C_i}\rho_1\left(\dfrac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_1}\right)$ para $i=1,...,r$. Como las q_i son uniformemente continuas y acotadas, $\exists\delta_0>0$ y $\eta>0$ tales que $\forall i$

$$\Pi_{p+1}(H,H_0)<\delta_0\Rightarrow E_Hq_i\left(y,\mathbf{x}\right)\geq E_{F_0}\rho_1\left(\frac{u}{\sigma_0}\right)+\eta$$

$$\Rightarrow E_H\inf_{i=1,\ldots,r}q_i\left(y,\mathbf{x}\right)\geq E_{F_0}\rho_1\left(\frac{u}{\sigma_0}\right)+\eta$$

$$\Rightarrow E_H\inf_{B_\varepsilon}\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_1}\right)\geq E_{F_0}\rho_1\left(\frac{u}{\sigma_0}\right)+\eta$$

que es (54). Ahora se probará (55).

Como

$$\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma}\right)\underset{\boldsymbol{\theta}\to\boldsymbol{\theta}_0}{\longrightarrow}\rho_1\left(\frac{u}{\sigma}\right)$$

y

$$\left|\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma}\right)\right|\leq 1$$

entonces, por el Teorema de la Convergencia Dominada

$$E_{H_0}\left(\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma}\right)\right)\underset{\boldsymbol{\theta}\to\boldsymbol{\theta}_0}{\longrightarrow} E_{F_0}\rho_1\left(\frac{u}{\sigma}\right).$$

Por lo tanto, $\exists\varphi_1>0$ tal que $\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\leq\varphi_1$ implica

$$\left|E_{H_0}\left(\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma}\right)\right)-E_{F_0}\rho_1\left(\frac{u}{\sigma}\right)\right|<\frac{\eta}{8}$$

y $\exists\varphi_2>0$ tal que $\|\sigma-\sigma_0\|\leq\varphi_2$ implica

$$\left|E_{H_0}\left(\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma}\right)\right)-E_{H_0}\left(\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma_0}\right)\right)\right|<\frac{\eta}{8}.$$

Por lo tanto, $\exists\varphi_1>0$, $\varphi_2>0$ tales que

$$E_{H_0}\left(\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\leq\varphi_1,\ \|\sigma-\sigma_0\|\leq\varphi_2}\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma}\right)\right)\leq E_{F_0}\rho_1\left(\frac{u}{\sigma_0}\right)+\frac{\eta}{4}.$$

Sea $\varphi=\min(\varphi_1,\varphi_2)$ y sea

$$q(y,\mathbf{x})=\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\leq\varphi,\ \|\sigma-\sigma_0\|\leq\varphi}\rho_1\left(\frac{y-g(\mathbf{x},\boldsymbol{\theta})}{\sigma}\right).$$

Como $q(y,\mathbf{x})$ es uniformemente continua y acotada, $\exists\delta_1>0$ tal que $\Pi_{p+1}(H,H_0)<\delta_1$ implica

$$E_H q(y,\mathbf{x})\leq E_{F_0}\rho_1\left(\frac{u}{\sigma_0}\right)+\frac{\eta}{2}.$$

Además, $\mathbf{T}_0(H_0)=\boldsymbol{\theta}_0$, $\mathbf{T}_0$ es continua en H_0 y por el lema 5.3, $s(H)$ es continua en H_0. Entonces $\exists\delta_2>0$ tal que

$$\Pi_{p+1}(H,H_0)<\delta_2\Rightarrow|\mathbf{T}_0(H)-\mathbf{T}_0(H_0)|\leq\varphi \text{ y } |s(H)-s(H_0)|\leq\varphi.$$

Entonces tomando $\delta=\min(\delta_0,\delta_1,\delta_2)$ se tiene (55), quedando así probado el teorema. ■

Demostración del teorema 3.9. Dado $\varepsilon<\min\left(\varepsilon^*_{n,+}(g,\mathbf{T}_0,\chi),\dfrac{1-2c_n}{2-2c_n}\right)$, queremos ver que

$$\forall\chi^m \text{ con } \frac{m}{n}<\varepsilon \text{ se tenga } g(\mathbf{x},\mathbf{T}_1(\chi^m))<\infty.$$

Sea

$$\varepsilon_{+}(g,\mathbf{T}_0)=\operatorname*{\acute{i}nf}_{i}\varepsilon_{+}(\mathbf{x}_i,g,\mathbf{T}_0)=\frac{m}{n}$$

entonces

$$\forall i=1,...,n\ \ \varepsilon_{+}(\mathbf{x}_i,g,\mathbf{T}_0)\geq\frac{m}{n}$$

es decir que

$$\exists i_0,\ \sup_{\chi^m} g(x_{i_0},\mathbf{T}_0(\chi^m))=\infty,$$

Luego, existe una sucesión $\boldsymbol{\theta}_N=\mathbf{T}_0\left(\chi_N^m\right)$ tal que $g(x_{i_0},\boldsymbol{\theta}_N)\underset{N\to\infty}{\rightarrow}\infty$

Por (22), $\exists M>\frac{n}{2}$ tal que $\forall A,\exists B$ tal que

$$\min_{\boldsymbol{\theta}\in U_B}\frac{\#\left\{i:g(\mathbf{x}_i,\boldsymbol{\theta})>A\right\}}{n}\geq\frac{M}{n}=1-c_n>\frac{1}{2}$$

con $U_B=\left\{\boldsymbol{\theta}:g(\mathbf{x}_{i_0},\boldsymbol{\theta})>B\right\}$.

Sea A suficientemente grande y k_0 tal que

$$\operatorname*{\acute{i}nf}_{\boldsymbol{\theta}:g(\mathbf{x},\boldsymbol{\theta})>A}\rho_1\left(\frac{r_i\left(\boldsymbol{\theta}\right)}{k_0}\right)=a$$

Para ese A, existe B tal que

$$\min_{\boldsymbol{\theta}\in U_B}\frac{\#\left\{i:g(\mathbf{x}_i,\boldsymbol{\theta})>A\right\}}{n}\geq 1-c_n$$

es decir

$$\frac{\#\left\{i:g(\mathbf{x}_i,\boldsymbol{\theta})>A\right\}}{n}\geq 1-c_n\quad\forall\boldsymbol{\theta}\in U_B$$

Luego,

$$\operatorname*{\acute{i}nf}_{\substack{\boldsymbol{\theta}:g(\mathbf{x},\boldsymbol{\theta})>A\\ S_n\leq k_0}}\frac{1}{n}\sum_{i=1}^{n}\rho_1\left(\frac{r_i(\boldsymbol{\theta})}{S_n}\right)\geq\frac{M}{n}.a=\left(1-c_n\ \right).a>\frac{a}{2}$$

Además,

$$\frac{1}{n}\sum_{i=1}^{n}\rho_1\left(\frac{r_i(\mathbf{T}_0)}{S_n}\right)\leq\frac{1}{n}\sum_{i=1}^{n}\rho_0\left(\frac{r_i(\mathbf{T}_0)}{S_n}\right)=b=\frac{a}{2}$$

y resulta

$$\inf_{\substack{\boldsymbol{\theta}\,:\,g(\mathbf{x},\boldsymbol{\theta})>A \\ S_n\leq k_0}} \frac{1}{n}\sum_{i=1}^{n}\rho_1\left(\frac{r_i(\boldsymbol{\theta}\left(\chi^m\right))}{S_n}\right) > \frac{1}{n}\sum_{i=1}^{n}\rho_1\left(\frac{r_i(\mathbf{T}_0\left(\chi^m\right))}{S_n}\right)$$

Por lo tanto, por (9), $g\left(\mathbf{x},\mathbf{T}_1\left(\chi^m\right)\right) \leq A \quad \forall\chi^m$, con lo cual m puntos no podrían causar la ruptura superior.

Sólo resta probar que $\exists k_0$ tal que $S_n \leq k_0$.

Como

$$\varepsilon < \varepsilon^*_{n,+}(g,\mathbf{T}_0,\chi) \qquad \text{y} \qquad \frac{m}{n} > \varepsilon$$

existe un k_1 tal que

$$g(\mathbf{x},\mathbf{T}_0(\chi^m)) \leq k_1,\, \forall\chi^m \in D^m$$

Sea k_2 tal que

$$\sup_{i=1,\dots,n} r_i(\mathbf{T}_0(\chi^m)) \leq k_2$$

Como

$$\varepsilon < \frac{1-2c_n}{2-2c_n} < 1/2,$$

existe $\gamma > 0$ tal que $\varepsilon a + \gamma < b$

Sea δ tal que $\rho_0(\delta) = \gamma$ y $k_0 = k_2/\delta$.

Entonces

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\rho_0\left(\frac{r_i(\mathbf{T}_0(\chi^m))}{k_0}\right) &\leq \frac{n-m}{n}\rho_0\left(\frac{k_2}{k_0}\right) + \frac{m}{n}a \\ &= \frac{n-m}{n}\gamma + \frac{m}{n}a < \varepsilon a + \gamma < b\end{aligned}$$

y por lo tanto, $S_n \leq k_0$ como se quería probar. ■

6. Demostraciones correspondientes a la Sección 3.3.

Para demostrar el teorema 3.10 se necesitan los siguientes lemas.

Lema 6.1 *Supongamos que F_0 satisface (B) y que G_0 satisface (C). Sea $s(\boldsymbol{\theta})$ definido por*

$$E_{H_0}\left(\rho\left(\frac{y-\mathbf{g}(\mathbf{x},\boldsymbol{\theta})}{s(\boldsymbol{\theta})}\right)\right)=b \tag{56}$$

Entonces $s(\boldsymbol{\theta})$ tiene único mínimo en $\boldsymbol{\theta}=\boldsymbol{\theta}_0$.

Demostración. Sea $\boldsymbol{\theta}\neq\boldsymbol{\theta}_0$

$$b=E_{F_0}\left(\rho\left(\frac{\mathbf{u}}{s(\boldsymbol{\theta}_0)}\right)\right)<E_{H_0}\left(\rho\left(\frac{y-\mathbf{g}(\mathbf{x},\boldsymbol{\theta})}{s(\boldsymbol{\theta}_0)}\right)\right)$$

por lo tanto,

$$E_{H_0}\left(\rho\left(\frac{y-\mathbf{g}(\mathbf{x},\boldsymbol{\theta})}{s(\boldsymbol{\theta}_0)}\right)\right)>b=E_{H_0}\left(\rho\left(\frac{y-\mathbf{g}(\mathbf{x},\boldsymbol{\theta})}{s(\boldsymbol{\theta})}\right)\right)$$

entonces por ser ρ una función-ρ, se tiene $s(\boldsymbol{\theta}_0)<s(\boldsymbol{\theta})$. ■

Lema 6.2 *Sea H_0 una distribución sobre $\mathbb{R}^{p+1}$ que satisface (14) donde F_0 satisface (B) y G_0 satisface (C). Supongamos que ρ_2 satisface (F) y (G). Definimos*

$$\tau^2(\boldsymbol{\theta})=s^2(\boldsymbol{\theta})\,E_{H_0}\rho_2\left(\frac{y-\mathbf{g}(\mathbf{x},\boldsymbol{\theta})}{s(\boldsymbol{\theta})}\right) \tag{57}$$

donde $s(\boldsymbol{\theta})$ está definida por (56) usando como ρ la función ρ_1, entonces $\tau(\boldsymbol{\theta})$ tiene un único mínimo en $\boldsymbol{\theta}=\boldsymbol{\theta}_0$.

Demostración. Por el lema 5.1, se tiene

$$\tau^2(\boldsymbol{\theta})>s^2(\boldsymbol{\theta})\,E_{H_0}\rho_2\left(\frac{y}{s(\boldsymbol{\theta})}\right)\quad\text{para todo }\boldsymbol{\theta}\neq\boldsymbol{\theta}_0$$

entonces por el lema 6.1 para probar este lema es suficiente con probar que

$$h(s)=s^2E_{F_0}\rho_2\left(\frac{y}{s}\right)$$

es no decreciente en s. Como (F) y (G) implican que

$$h'(s) = sE_{F_0}\left\{2\rho_2\left(\frac{y}{s}\right) - \psi_2\left(\frac{y}{s}\right)\frac{y}{s}\right\} \geq 0$$

entonces $h(s)$ es no decreciente y $h(s(\boldsymbol{\theta}))$ tiene único mínimo en $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Por lo tanto,

$$\tau^2(\boldsymbol{\theta}) > s^2(\boldsymbol{\theta})\, E_{H_0}\rho_2\left(\frac{y}{s(\boldsymbol{\theta})}\right) > s^2(\boldsymbol{\theta}_0)\, E_{H_0}\rho_2\left(\frac{y}{s(\boldsymbol{\theta}_0)}\right) = \tau^2(\boldsymbol{\theta}_0)$$

$$\Rightarrow \tau^2(\boldsymbol{\theta}) > \tau^2(\boldsymbol{\theta}_0) \quad \forall \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$$

quedando así probado el lema. ■

Lema 6.3 *Supongamos que F_0 satisface (B) y G_0 satisface (C). Entonces*

$$\lim_{n\to\infty} \sup_{\boldsymbol{\theta}} |s_n(\mathbf{r}(\boldsymbol{\theta})) - s(\boldsymbol{\theta})| = 0 \quad c.s. \tag{58}$$

$$\lim_{n\to\infty} \sup_{\boldsymbol{\theta}} |\tau_n(\mathbf{r}(\boldsymbol{\theta})) - \tau(\boldsymbol{\theta})| = 0 \quad c.s. \tag{59}$$

donde $s(\boldsymbol{\theta})$ y $\tau(\boldsymbol{\theta})$ están definidos por (56) y (57), respectivamente, con $\rho = \rho_1$.

Demostración. $s(\boldsymbol{\theta})$ es continua y positiva. Sea $h_1 = \inf_{\boldsymbol{\theta}} s(\boldsymbol{\theta})$ y $h_2 = \sup_{\boldsymbol{\theta}} s(\boldsymbol{\theta})$. Entonces $h_1 > 0$ y $h_2 < \infty$.

Por el lema 4.4 de Yohai (1987) se tiene para $j = 1, 2$

$$\lim_{n\to\infty} \sup_{\substack{\boldsymbol{\theta} \\ s\in[h_1/2,2h_2]}} \left| \frac{1}{n}\sum_{i=1}^{n} \rho_j\left(\frac{r_i(\boldsymbol{\theta})}{s}\right) - E_{H_0}\rho_j\left(\frac{\mathbf{u}(\boldsymbol{\theta})}{s}\right) \right| = 0 \quad c.s. \tag{60}$$

Primero veremos que (60) con $j = 1$ implica (58).

Sea $0 \leq \varepsilon \leq \frac{h_1}{2}$ y sean

$$g_1(\boldsymbol{\theta}) = E_{H_0}\rho_1\left(\frac{\mathbf{u}(\boldsymbol{\theta})}{s(\boldsymbol{\theta}) + \varepsilon}\right) \quad \text{ y } \quad g_2(\boldsymbol{\theta}) = E_{H_0}\rho_1\left(\frac{\mathbf{u}(\boldsymbol{\theta})}{s(\boldsymbol{\theta}) - \varepsilon}\right)$$

Claramente, $g_1(\boldsymbol{\theta}) < b$ y $g_2(\boldsymbol{\theta}) > b$ y, además, son continuas. Entonces

$$\gamma_1 = \sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} g_1(\boldsymbol{\theta}) < b \ \text{ y } \ \gamma_2 = \inf_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} g_2(\boldsymbol{\theta}) > b$$

Sea $\delta = \text{mín}\,(b-\gamma_1, \gamma_2 - b)$. Por (60), $\exists n_0$ tal que $\forall n \geq n_0$

$$\sup_{\substack{\boldsymbol{\theta} \\ s\in[h_1/2, 2h_2]}} \left| \frac{1}{n}\sum_{i=1}^{n} \rho_1\left(\frac{r_i(\boldsymbol{\theta})}{s}\right) - E_{H_0}\rho_1\left(\frac{\mathbf{u}(\boldsymbol{\theta})}{s}\right) \right| \leq \frac{\delta}{2}$$

Luego, $\forall n \geq n_0$ se tiene

$$\begin{aligned} \inf_{\boldsymbol{\theta}} \frac{1}{n}\sum_{i=1}^{n} \rho_1\left(\frac{r_i(\boldsymbol{\theta})}{s(\boldsymbol{\theta})-\varepsilon}\right) &\geq \inf_{\boldsymbol{\theta}} E_{H_0}\rho_1\left(\frac{\mathbf{u}(\boldsymbol{\theta})}{s(\boldsymbol{\theta})-\varepsilon}\right) + \frac{\delta}{2} \\ &\geq b + \delta - \frac{\delta}{2} = b + \frac{\delta}{2} \end{aligned}$$

Análogamente,

$$\sup_{\boldsymbol{\theta}} \frac{1}{n}\sum_{i=1}^{n} \rho_1\left(\frac{r_i(\boldsymbol{\theta})}{s(\boldsymbol{\theta})+\varepsilon}\right) \leq b - \frac{\delta}{2}$$

entonces

$$\sup_{\boldsymbol{\theta}} |s_n(\mathbf{r}(\boldsymbol{\theta})) - s(\boldsymbol{\theta})| \leq \varepsilon$$

Finalmente, (59) se sigue de (58) y (60). ■

Demostración del teorema 3.10. Por el lema 6.3, la continuidad de $\tau(\boldsymbol{\theta})$ y $\mathbf{T}_{0,n} \to \boldsymbol{\theta}_0$ *c.s.* se tiene que

$$\tau_n(\mathbf{r}(\mathbf{T}_{0,n})) \to \tau(\boldsymbol{\theta}_0) \quad c.s.$$

Si probamos que $\forall \varepsilon > 0$

$$\limsup_{n\to\infty} \inf_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\geq\varepsilon} \tau_n(\mathbf{r}(\boldsymbol{\theta})) > \tau(\boldsymbol{\theta}_0) \tag{61}$$

entonces por (19), $\|\mathbf{T}_{1,n} - \boldsymbol{\theta}_0\| < \varepsilon$ lo cual mostraría la consistencia de $\mathbf{T}_{1,n}$.

Finalmente, por el lema 6.2 y la continuidad de $\tau(\boldsymbol{\theta})$

$$\inf_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\geq\varepsilon} \tau(\boldsymbol{\theta}) > \tau(\boldsymbol{\theta}_0)$$

y por el lema 6.3 se tiene (61). ■

Observación 6.1 *Si el espacio de parámetros Θ no es compacto, pero la función g cumple la condición (18), se trabaja con la compactificación de Stone-Čech.*

Demostración del teorema 3.11. Por la definición de $\mathbf{T}_n$ y el Teorema del Valor Medio, se tiene

$$\frac{1}{n}\sum_{i=1}^{n}\left[W_n\psi_1\left(\frac{r_i\left(\boldsymbol{\theta}\right)}{s_n}\right)+\psi_2\left(\frac{r_i\left(\boldsymbol{\theta}\right)}{s_n}\right)\right]\dot{\mathbf{g}}\left(\mathbf{x}_i,\boldsymbol{\theta}\right)-V_n\left(\boldsymbol{\theta}_n^*\right)\frac{\left(\mathbf{T}_n-\boldsymbol{\theta}_0\right)}{s_n}=0$$

con $\|\boldsymbol{\theta}_n^*-\boldsymbol{\theta}_0\|\leq\|\mathbf{T}_n-\boldsymbol{\theta}_0\|$ y

$$V_n\left(\boldsymbol{\theta}\right)=\frac{1}{n}\sum_{i=1}^{n}\left[W_n\psi_1'\left(\frac{r_i\left(\boldsymbol{\theta}\right)}{s_n}\right)+\psi_2'\left(\frac{r_i\left(\boldsymbol{\theta}\right)}{s_n}\right)\right]\dot{\mathbf{g}}\left(\mathbf{x}_i,\boldsymbol{\theta}\right)\dot{\mathbf{g}}\left(\mathbf{x}_i,\boldsymbol{\theta}\right)'$$

Luego,

$$\sqrt{n}\left(\mathbf{T}_n-\boldsymbol{\theta}_0\right)=s_nV_n^{-1}\left(\boldsymbol{\theta}_n^*\right)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[W_n\psi_1\left(\frac{r_i\left(\boldsymbol{\theta}\right)}{s_n}\right)+\psi_2\left(\frac{r_i\left(\boldsymbol{\theta}\right)}{s_n}\right)\right]\dot{\mathbf{g}}\left(\mathbf{x}_i,\boldsymbol{\theta}\right)$$

por el lema 4.2 de Yohai (1987) y el lema 6.3

$$W_n\rightarrow W_0\quad c.s.\quad \text{y}\quad V_n\left(\boldsymbol{\theta}_n^*\right)\rightarrow V_0\quad c.s.$$

siendo $V_0=B\left(\psi_0,F_0\right)^{1/2}V$.

Además por el lema 5.1 Yohai (1987) y el lema 6.3

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_j\left(\frac{r_i\left(\boldsymbol{\theta}_0\right)}{s_n}\right)$$

es asintóticamente equivalente a

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_j\left(\frac{r_i\left(\boldsymbol{\theta}_0\right)}{\sigma_0}\right)\text{ para } j=1,2$$

Luego, por el Teorema Central del Límite y el lema de Slutsky

$$\sqrt{n}\left(\mathbf{T}_n-\boldsymbol{\theta}_0\right)\rightarrow^d N\left(\mathbf{0},\sigma_0^2\frac{A\left(\psi_0,F_0\right)}{B\left(\psi_0,F_0\right)}V^{-1}\right)$$

■

Para la demostración del teorema 3.12 es necesario el siguiente lema.

Lema 6.4 *Sea $Z_n = \{z_1, ..., z_n\}$ una muestra de tamaño n y sea c_n dado por (22). Consideremos las mismas suposiciones que en el teorema 3.12. Entonces dado $\varepsilon < \dfrac{1-2c_n}{2-2c_n}$ y $k_2 > 0, \exists k_1$ tal que $\dfrac{m}{n} \leq \varepsilon$ implica*

$$\inf_{\boldsymbol{\theta}:|g(\mathbf{x},\boldsymbol{\theta})|\geq k_1} \tau_n\left(\mathbf{r}\left(\boldsymbol{\theta}\right)\right) \geq k_2$$

$\forall$ $\mathbf{W}_m$, $\#\left(\mathbf{W}_m \cap Z_n\right) = m$

Demostración. Como $c_n < 0{,}5$ entonces $\dfrac{M}{n} > 0{,}5$. Recordemos que

$$M = \min_{i_0} \min_{A} \max_{B} \min_{\theta \in U(B,i_0)} \#\left\{i : g\left(x_i, \theta\right) > A\right\}$$

siendo $U(B, i_0) = \{\theta : g\left(x_{i_0}, \theta\right) > B\}$.

Luego, $\dfrac{m}{n} < \varepsilon$ implica que $\varepsilon_1 = \dfrac{n-m}{n}(1-c_n) > 0{,}5$.

Por el lema 3.2 (ii) de Yohai y Zamar (1988), dado k_2 y $\varepsilon_1 > 0{,}5$, existe k tal que

$$\frac{\#\left\{i : |u_i| > k\right\}}{n} \geq \varepsilon_1 \text{ implica } \tau_n\left(r\left(\mathbf{u}\right)\right) \geq k_2 \tag{62}$$

para cualquier muestra $\mathbf{u} = (u_1, ..., u_n)$. Sea $A = \max\limits_{i=1,\dots,n} |y_i| + k$.

Entonces, para ese A y $\forall i_0$, $\exists B_{i_0}$ tal que

$$\frac{\#\left\{i : g\left(x_i, \theta\right) > A\right\}}{n} \geq 1 - c_n > 0{,}5$$

$\forall \theta \in U(B_{i_0}, i_0)$.

Sea $k_1 = \inf\limits_{i_0} B_{i_0}$. Supongamos, entonces, $\dfrac{m}{n} < \varepsilon$ y $\theta : g\left(x_i, \theta\right) > k_1$.

Sean $C = \{i : w_i = z_i\}$ y $D = \{i \in C : g\left(x_i, \theta\right) > A\}$.

Para $i \in D$ se tiene

$$\left|y_i - g\left(x_i, \theta\right)\right| \geq \left|\max|y_i| - A\right| \geq k$$

por lo tanto, $\forall i \in D$, $\left|r_i\left(\theta\right)\right| \geq k$.

Luego,

$$\#\left\{i \in C : \left|r_i\left(\theta\right)\right| \geq k\right\} \geq \#D \geq \left(1 - c_n\right)\left(n - m\right)$$

entonces,

$$\frac{\#\left\{i:\left|r_i\left(\theta\right)\right|\geq k\right\}}{n}\geq\frac{\#D}{n}\geq\left(1-c_n\right).\frac{n-m}{n}=\varepsilon_1>0{,}5$$

y por (62) $\tau_n\left(r\left(\boldsymbol{\theta}\right)\right)\geq k_2$ para $\theta:g\left(x_i,\theta\right)>k_1$. ■

Ahora estamos en condiciones de demostrar el teorema 3.12.

Demostración del teorema 3.12. Dado $\varepsilon<\min\left\{\varepsilon_{n,+}^{*}\left(g,\mathbf{T}_0,\mathbf{Z}_n\right),\dfrac{1-2c_n}{2-2c_n}\right\}$, se quiere ver que $\forall\chi^m$, $\frac{m}{n}<\varepsilon$ se tenga $g\left(\mathbf{x},\mathbf{T}_1\left(\chi^m\right)\right)<\infty$.

Como $\varepsilon<\varepsilon_{n,+}^{*}\left(g,\mathbf{T}_0,\mathbf{Z}_n\right)$ entonces para $\frac{m}{n}<\varepsilon$, $\exists\ k$ tal que

$$g\left(\mathbf{x},\mathbf{T}_0\left(\chi^m\right)\right)<k\forall\chi^m\in D^m$$

y por lo tanto,

$$\exists k_1 \text{ tal que } \sup_{i=1,..,n} r_i\left(\mathbf{T}_0\left(\chi^m\right)\right)\leq k_1$$

entonces,

$$\frac{\#\left\{i:\left|r_i\left(\mathbf{T}_0\left(\chi^m\right)\right)\right|\leq k_1\right\}}{n}\geq 1-\varepsilon,\ \forall\varepsilon<0{,}5.$$

Por el lema 3.2 (i) de Yohai y Zamar (1988), $\exists k_2$ tal que $\tau_n\left(\mathbf{r}\left(\mathbf{T}_0\left(\chi^m\right)\right)\right)\leq k_2$ y por el lema 6.4 existe K, $\frac{m}{n}\leq\varepsilon$ implica

$$\inf_{\boldsymbol{\theta}:|g(\mathbf{x},\boldsymbol{\theta})|\geq K}\tau_n\left(\mathbf{r}(\theta)\right)\geq 2k_2>\tau_n\left(\mathbf{r}\left(\mathbf{T}_0\left(\chi^m\right)\right)\right)\geq\tau_n\left(\mathbf{r}\left(\mathbf{T}_1\left(\chi^m\right)\right)\right)$$

con lo cual se tiene que $g\left(\mathbf{x},\mathbf{T}_1\left(\chi^m\right)\right)<K$. Por lo tanto, m puntos no pueden causar la ruptura superior y

$$\varepsilon_{n,+}^{*}\left(g,\mathbf{T}_1,\mathbf{Z}_n\right)\geq\min\left\{\varepsilon_{n,+}^{*}\left(g,\mathbf{T}_0,\mathbf{Z}_n\right),\frac{1-2c_n}{2-2c_n}\right\}$$

■

7. Simulación

Se ha simulado el modelo de crecimiento exponencial

$$y_i = \beta \exp(\alpha x_i) + e_i$$

considerando $x_i \sim U(0,1)$, $i = 1, ..., 20$; $\beta = 5$; $\alpha = 2$; $\sigma = 1$ y errores $e_i \sim N(0, \sigma^2)$.

Cada simulación consitió en 200 conjuntos de datos como los mencionados a cada uno de los cuales se lo contaminó con un 10 % de datos atípicos de la forma:

$$y_0 = \beta \exp(\alpha x_0) K$$

donde a x_0 se le dio valores entre 0,5 y 2 y la constante K varió entre 0,1 y 4.

Para cada simulación se calcularon el MM-estimador y el τ-estimador para (α, β) tomando como función-ρ, $\rho(r/c)$ siendo ρ la función bicuadrática: $\rho(r) = \min\left\{1, 1 - (1 - r^2)^3\right\}$ y eligiendo c de manera de obtener una eficiencia normal de 0.85. Como estimador inicial elegimos el S-estimador con igual función-ρ.

Como medida del comportamiento de cada estimador se utilizó el error medio cuadrático (EMC). Dado que bajo distribuciones contaminadas, los estimadores robustos suelen tener distribuciones con colas pesadas, se usó además un EMC "podado" obtenido omitiendo el 10 % *superior* de los errores cuadráticos.

En las tablas 1 y 2 se muestran algunos de los errores cuadráticos podados para $x_0 = 1{,}09$, que es el valor de x_0 donde se obtuvieron los EMC más altos. Las tablas 1 y 2 corresponden respectivamente al MM -estimador y al $\tau-$estimador. En ambas tablas, la primera columna contiene el valor de la constante K, en la segunda está el EMC podado para el parámetro β y en la última columna el correspondiente para α. Para el MM-estimador observamos que el valor máximo se obtiene para $K = 1{,}09$ y para el τ-estimador el máximo se encuentra en $K = 1{,}105$.

En las figuras 1 y 2 se han graficado los EMC podados para los distintos valores de K, para el $x_0 = 1{,}09$. Las figuras 1 y 2 corresponden respectivamente a β y α. Podemos observar que ambos estimadores se comportan de manera similar, aunque el MM-estimador da errores levemente inferiores a los del τ-estimador.

K	$\boldsymbol{\beta}$	$\boldsymbol{\alpha}$
0,7	0,050332	0,0033321
1	0,033137	0,0011764
1,08	0,16421	0,016192
1,09	**0,1802**	**0,01812**
1,1	0,16713	0,017025
1,15	0,11812	0,011771
1,2	0,061575	0,0038851
1,5	0,051611	0,003331

Tabla 1: EMC podados correspondientes al MM-estimador, para $x_0 = 1{,}09$.

K	$\boldsymbol{\beta}$	$\boldsymbol{\alpha}$
0,5	0,065867	0,004121
1	0,034564	0,0011971
1,08	0,17936	0,01745
1,105	**0,19969**	**0,020168**
1,12	0,17398	0,017446
1,15	0,14004	0,013862
1,2	0,087361	0,0054806
1,7	0,067418	0,0042487

Tabla 2: EMC podados correspondientes al τ-estimador, para $x_0 = 1{,}09$.

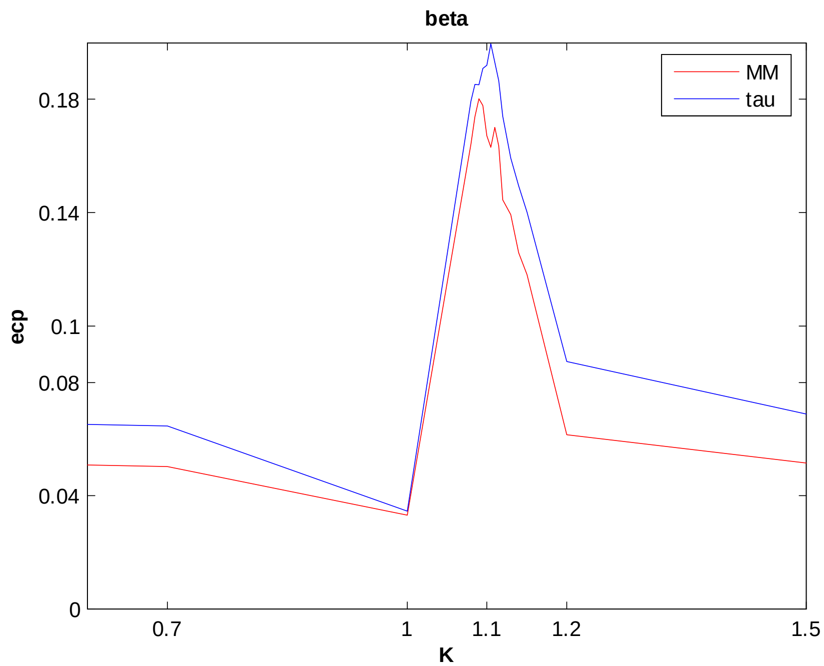

Figura 1: EMC podados vs K, para el parámetro β.

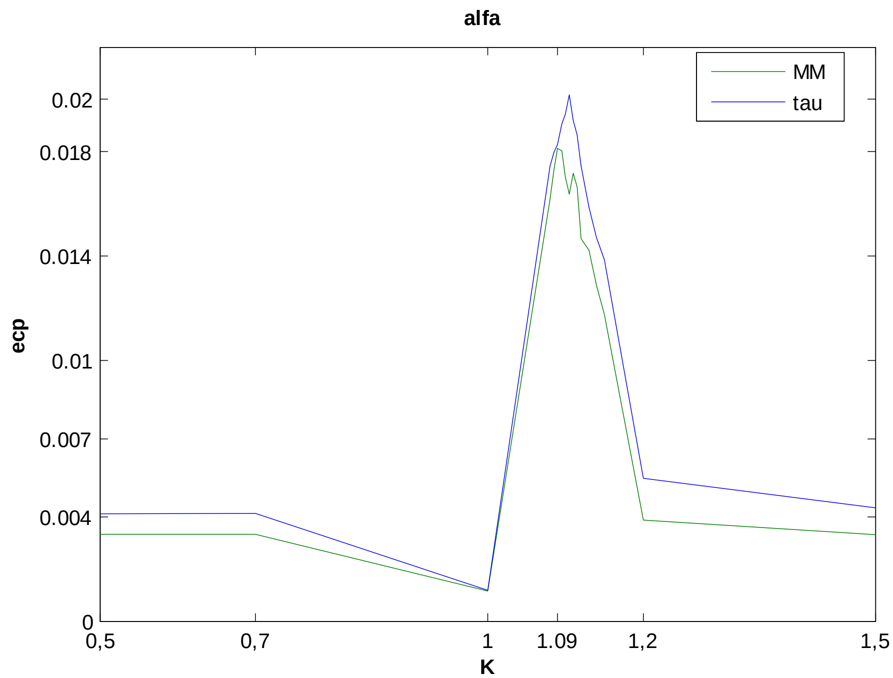

Figura 2: EMC podados vs K, para el parámetro α.

Referencias

[1] Amemiya, T. (1983). Nonlinear regression models, en *Handbook of Econometrics*, Vol **I,** 333-389, editores Z. Griliches y M. D. Intriligator. North-Holland: Amsterdan.

[2] Bianco, A. y Yohai, V.J. (1996), Robust estimation in the logistic regression model, en *Robust Statistics, Data Analysis and Computer Intensive Methods, Proceedings of the workshop in honor of Peter J. Huber*, editor H. Rieder, Lecture Notes in Statistics **109**, 17-34 Springer-Verlag, New York.

[3] Bunke, H. y Schmidt, W. (1980). Asymptotic results on nonlinear approximation of regression functions and weighted least squares. *Mathematische Operationsforschung und Statistik Series Statistics*, **11**, 3-22.

[4] Carroll, R. J. y Ruppert, D. (1987). Diagnostics and robust estimation when transforming the regression model and the response .*Technometrics*, **29**, 287-299.

[5] Čížek, P. (2006). Least trimmed squares in nonlinear regression under dependence. *Journal of Statistical Planning and Inference*, **136**, 3967-3988.

[6] Dutter, R. y Huber, P.J. (1981). Numerical methods for the nonlinear robust regression problem. *Journal of Statistical Computation and Simulation*, **13**, 79-113.

[7] Edlund, Ove, Ekblom, Hakan, Ekblom, Haakan, Ekblom, Håkan y Madsen, Kaj (1997). Algorithms for nonlinear M-estimation. *Computational Statistics*, **12**, 373-383.

[8] Fraiman, R. (1983) General M-estimators and applications to bounded influence estimation for non-linear regression. *Communications in Statistics. Theory and Methods*, Vol **A12**, **22**, 2617-2631.

[9] Gallant, A. R. (1975). The power of the likelihood ratio test of location in nonlinear regression models. *Journal of the American Statistical Association*, **70**, 198-203.

[10] Gallant, A.R. (1987). *Nonlinear Statistical Models*, Wiley.

[11] Genton, M. G. y Lucas, A. (2003). Comprehensive definitions of breakdown points for independent and dependent observations. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, **65**, 81-94.

[12] Hampel, F. (1971). A General Qualitative Definition of Robustness. *The Annals of Mathematical Statistics*, **42**, 1887-1896.

[13] Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions, en *Proceedings of the Fifth Berkeley Symposium in Mathematical Statistics and Probability*, Berkeley: University of California Press, Vol. **1**, 221-233.

[14] Huber, P. J. (1973). Robust Regression: Asymptotics, Conjectures and Monte Carlo. *The Annals of Statistics,* **1**, 799-821.

[15] Huber, P. J. y Donoho, D. L. (1983). The notion of breakdown point, en *A festschrift for Erich L. Lehmann in honor of his 65th birthday*, editores P. J: Bickel, K. A. Doksum, y J. L. Hodges, Jr., Belmont, CA: Wadsworth International Group, 157-184.

[16] Jennrich, R. I. (1969). Asymptotic propierties of nonlinear least squares estimators. *The Annals of Mathematical Statistics*, **40,** 633-643.

[17] Johansen, S. (1984). Functional Relations, Random Coefficients, and Nonlinear Regression with Application to Kinetic Data. *Lecture Notes in Statistics*, **22**, 126ss. Springer-Verlag: Berlin.

[18] Liese, F. y Vajda, I. (2003). A general asymptotic theory of M-estimators I. *Mathematical Methods of Statistics*, **12**, 454-477.

[19] Liese, F. y Vajda, I. (2004). A general asymptotic theory of M-estimators II *Mathematical Methods of Statistics*, **13**, 82-95.

[20] Malinvaud, E. (1970a). The consistency of nonlinear regressions. *The Annals of Mathematical Statistics*, **41,** 956-969.

[21] Malinvaud, E. (1970b). *Statistical Methods of Econometrics*, Amsterdam.

[22] Markatou, M.y Manos, G. (1996). Robust tests in nonlinear regression models. *Journal of Statistical Planning and Inference*, **55**, 205-217.

[23] Maronna, R., Yohai V. y Zamar, M. (2006). *Robust Statistics, Theory and Methods*, Wiley.

[24] Mukherjee, K. (1996). Robust estimation in nonlinear regression via minimum distance method . *Mathematical Methods of Statistics*, **5**, 99-112.

[25] Munkres, James R. (1975). *TOPOLOGY. A first course.* Prentice−Hall, Inc.

[26] Richardson, G.D: y Bhattacharyya, B.B. (1986). Consistent estimators in nonlinear regression for a noncompact parameter space. *The Annals of Statistics,* **14,** 1591-1596.

[27] Rousseeuw, P. (1984). Least median of squares regression. *Journal of the American Statistical Association* **79**, 871-880.

[28] Rousseeuw, P. (1985). Multivaritate estimation with high breakdown point. *Mathematical Statistics and its Applications (Vol. B)*, editores: W. Grossmann, G. Pflug, I. Vincze y W. Wertz, 283-297. Dordrecht: Riedel.

[29] Rousseeuw, P y Yohai, V. (1984). Robust regression by means of S-estimators, en *Robust and nonlinear time series analysis,* editor Heidelberg. Lecture Notes in Statististics, **26**, 256–272. Springer, New York.

[30] Sakata, S. y White, H.(1995). An alternative definition of finite-sample breakdown point with applications to regression model estimators. *Journal of the American Statistical Association*, **90**, 1099-1106.

[31] Sakata, S. y White, H. (2001). S-estimation of nonlinear regression models with dependent and heterogeneous observations. *Journal of Econometrics*, **103**, 5-72.

[32] Seber, G.A.F. y Wild, C.J. (1989). *Nonlinear Regression.* Wiley.

[33] Shao, J. (1992) Consistency of Least-Squares Estimator and Its Jackknife Variance Estimator in Nonlinear Models. *Canadian Journal of Statistics*, **20**, 415-428.

[34] Stromberg, A. (1992). High breakdown estimators in nonlinear regression, en L1-*Statistical Analysis and Related Methods*, 103-112 . Dodge, Yadolah (ed.). Elsevier/North-Holland.

[35] Stromberg, A.J. (1993). Computation of High Breakdown Nonlinear Regression Parameters. *Journal of the American Statistical Association* **88**, 237-244.

[36] Stromberg, A. J. (1994). Software for computing a robust efficient linear or nonlinear regression estimator. *American Statistical Association.* Proceedings of the Section on Statistical Education, 6-11.

[37] Stromberg, A. J. (1995). Consistency of the least median of squares estimator in nonlinear regression *Communications in Statistics: Theory and Methods*, **24**, 1971-1984.

[38] Stromberg, A. J. (1997). Some software for computing robust linear or nonlinear regression estimators. *Communications in Statistics: Simulation and Computation*, **26**, 947-959.

[39] Stromberg, A. J. y Ruppert, D (1992) . Breakdown in nonlinear regression. *Journal of the American Statistical Association*, **87**, 991-997.

[40] Tabatabai M. A. y Argyros I. K. (1993). Robust estimation and testing for general nonlinear regression models. *Applied mathematics and computation*, **58**, 85-101.

[41] Tiede, J. J. y Pagano, M. (1979). The application of robust calibration to radioimmunoassay. *Biometrics*, **35**, 567-574.

[42] Vainer, B. P. y Kukush, A. G. (1998). The consistency of M-estimators constructed from a concave weight function. *Theory of Probability and Mathematical Statistics*, **57**, 11-18.

[43] Vankeerberghen, P., Smeyers-Verbeke, J., Leardi, R., Karr, C. L. y Massart, D. L. (1995). Robust regression and outlier detection for non-linear models using genetic algorithms. *Chemometrics and Intelligent Laboratory Systems*, **28**, 73-87.

[44] Wu, C. F. (1981). Asymptotic theory of nonlinear least squares estimation. *The Annals of Statistics,* **9**, 501-513.

[45] Yohai, V. J. (1987). High breakdown-point and high efficiency estimates for regression. *The Annals of Statistics,* **15**, 642-656.

[46] Yohai, V. J. y Zamar, R. H. (1988). High breakdown estimates of regression by means of the minimization of an efficient scale, *Journal of the American Association,* **83**, 406-413.

[47] Zwanzig, S. (1980). The choice of approximatative models in nonlinear regression. *Mathematische Operationsforschung und Statistik Series Statistics*, **11**, 23-47.