



UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Matemática

Inferencia robusta en modelos no lineales con respuestas faltantes

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires en el
área Ciencias Matemáticas

Paula Mercedes Spano

Director de tesis: Dra. Ana M. Bianco

Consejero de estudios: Dra. Ana M. Bianco

Lugar de trabajo: Instituto de Cálculo

Buenos Aires, 11 de noviembre de 2015

Fecha de defensa: 28 de marzo de 2016

Inferencia Robusta en Modelos No Lineales con Respuestas Faltantes

El modelo lineal es uno de los más populares en Estadística. Sin embargo, en muchas situaciones la naturaleza del fenómeno en estudio es intrínsecamente no lineal y por lo tanto, no se pueden utilizar aproximaciones lineales teniéndose que ajustar a los datos un modelo no lineal. Es bien sabido que, en este contexto, el estimador clásico del parámetro de regresión basado en mínimos cuadrados es muy sensible a los datos atípicos y por esta razón, se necesitan métodos robustos de inferencia. Asimismo, en algunas ocasiones, los datos están incompletos por haberse perdido algunas respuestas al azar.

En esta tesis, proponemos una familia de estimadores robustos para estimar al parámetro de regresión en un modelo no lineal con posibles respuestas faltantes al azar. Bajo condiciones de regularidad, estudiamos la consistencia y la distribución asintótica de los estimadores propuestos, dando un tratamiento unificado a muestras completas y a muestras con respuestas faltantes. Para estudiar la robustez local, deducimos su función de influencia y para comparar su comportamiento con el del estimador clásico y con otras alternativas robustas, consideramos dos ejemplos numéricos. Llevamos a cabo simulaciones de Monte Carlo para estudiar el comportamiento en muestras finitas de los estimadores propuestos bajo distintos esquemas de pérdida de respuestas y en diferentes escenarios de contaminación.

Por otro lado, introducimos una familia de tests robustos utilizando un estadístico de tipo Wald basado en el estimador propuesto con el objetivo de testear hipótesis que involucren al parámetro de regresión. Deducimos su distribución asintótica bajo la hipótesis nula, como también bajo alternativas contiguas. Asimismo, deducimos su función de influencia. Realizamos un estudio de simulación para evaluar el comportamiento de los tests propuestos.

Finalmente, el procedimiento propuesto también se ilustra a través del análisis de un conjunto de datos reales.

Palabras Claves: Modelos de regresión no lineal; Estimación robusta; Tests de hipótesis robustos; Datos Faltantes.

Robust Inference in Nonlinear Models with Missing Responses

The linear model is one of the most popular models in Statistics. However, in many situations the nature of the phenomenon is intrinsically nonlinear and so, linear approximations are not valid and the data must be fitted using a nonlinear model. It is well known that, in this setting, the classical estimator of the regression parameter, based on least squares, is very sensitive to outliers and for this reason, robust methods are needed. Besides, in some occasions the data are incomplete since some responses are missing at random.

In this thesis, we propose a family of robust estimators to estimate the regression parameter in a nonlinear model with responses missing at random. Under mild conditions, we study the consistency and asymptotic distribution of the proposal, giving a unified treatment to incomplete and complete samples. To study local robustness, we derive their influence function and, to compare its behaviour with that of the classical estimator and other robust alternatives, we consider two numerical examples. We perform Monte Carlo simulations to study the finite sample behavior of the proposed estimators under different schemes of missing responses and different scenarios of contaminations.

On the other hand, we introduce a family of robust tests using a Wald-type statistic based on the proposed estimators in order to check hypotheses that involve the regression parameter. We derive their asymptotic distribution under the null hypothesis as well as under contiguous alternatives. Besides, we derive their influence function. We perform a simulation study to evaluate the behaviour of the proposed test.

Finally, the proposed procedure is also illustrated analysing a real data set.

Keywords: Nonlinear regression models; Robust estimation; Robust hypothesis testing; Missing data.

Agradecimientos

En primer lugar quería agradecer a mi directora Ana por haber aceptado acompañarme en este largo camino con toda su dedicación y atención. No solo por dirigirme para realizar este trabajo de tesis sino también por ser tan hermosa persona y amiga.

Por otra parte quería agradecer a los jurados Ricardo Maronna, Daniela Rodríguez y Matías Salibián–Barrera por tomarse el trabajo de leer mi trabajo y por sus valiosas observaciones y correcciones.

También estoy muy agradecida de haber recibido, a lo largo de toda la carrera, los conocimientos de Víctor, Graciela, Mariela y Daniela. Como así también compartir el día a día y los viajes con todos mis compañeros del Instituto, no los nombro porque son muchos!! también aportando en conocimiento pero lo más importante aportando con su amistad!

En especial quiero agradecer a mis dos amores Seba y Oli. Enamorada de mi compañero y amigo que la vida me regaló, siempre acompañándome en todo momento y aguantándome con todas mis locuras. Y a mi Olivita hermosa que llegó para llenarnos de amor y alegría y que se portó súper bien para que yo pueda preparar la presentación.

A mis papás y a mis dos hermanos. Y bueno, la familia no se elige, si hubiera podido... elegiría la mía una y otra vez!! Gracias por ser incondicionales y por ser mis ejemplos de vida.

También gracias al resto de mi familia y amig@s de afuera del Instituto que si bien entre tanto trabajo no los veo muy seguido, saber que los tengo me hacen sentir completa.

Gracias infinitas (∞) a todos y todas por ser parte de mi vida y por llenarme de cariño y alegría.

Los quiero con todo el corazón!!
Paulita

Índice general

1. Introducción	1
2. Modelos No Lineales con Posibles Respuestas Faltantes	7
2.1. Introducción	7
2.2. Estimación Clásica	8
2.3. Estimación Robusta	9
2.3.1. M - y MM -estimadores	9
2.3.2. Propuesta	11
3. Comportamiento Asintótico	15
3.1. Consistencia de Fisher	15
3.2. Resultados previos: Teoría de Procesos Empíricos	18
3.3. Consistencia Fuerte	22
3.4. Distribución Asintótica	23
4. Curva de Influencia	33
4.1. Definición	33
4.2. Función de Influencia y Función de Influencia Esperada	34
4.3. Ejemplos Numéricos	36
4.3.1. Modelo de Michaelis–Menten	37
4.3.2. Modelo Exponencial	38
5. Estudio de Simulación	41
5.1. Modelo de Crecimiento Exponencial	41
5.2. Modelo Lineal–Exponencial	47

6. Tests de Hipótesis	53
6.1. Comportamiento asintótico	54
6.1.1. Bajo Hipótesis Fijas	54
6.1.2. Bajo Hipótesis Alternativas Contiguas	55
6.2. Función de Influencia	60
6.2.1. Ejemplo Numérico	62
6.3. Estudio de Simulación	65
6.4. Modelo Lineal–exponencial	65
6.5. Modelo de Crecimiento Exponencial	69
7. Ejemplo: Datos de Calidad del Aire	77
8. Apéndice	83

Capítulo 1

Introducción

El modelo lineal es uno de los más populares en Estadística, de fácil interpretación, y bajo ciertas condiciones, los estimadores óptimos de los parámetros tienen una forma cerrada de cálculo sencillo. No obstante, en el campo de las ciencias aplicadas, en áreas tan variadas como la Biología, la Ingeniería, pasando por la Medicina y la Farmacología, son frecuentes los problemas en los que es necesario el uso de regresión no lineal. Esto ocurre porque en muchas situaciones la naturaleza del fenómeno en estudio es intrínsecamente no lineal y por lo tanto, las simplificaciones lineales no dan aproximaciones satisfactorias. En otras situaciones, la relación no lineal puede ser transformada con el fin de obtener un modelo lineal que conserve las principales características del problema original. Sin embargo, es frecuente que en el campo de las Ciencias Física o Químicas, encontremos situaciones en las que los parámetros del modelo transformado no resultan de tanto interés como los provistos por el ajuste no lineal, en el cual tienen una interpretación natural. En estos casos, es necesario modelar los datos mediante una función que es una combinación no lineal de los parámetros del modelo y que depende de una o más covariables. En general, estos modelos suponen un conocimiento previo que conduce a una estructura funcional específica. Suelen ser utilizados tanto en contextos empíricos como teóricos que sugieren una forma funcional adecuada para expresar una relación no lineal conocida entre una variable de respuesta y un conjunto de variables predictoras.

En este sentido, los modelos de regresión no lineal suponen que la variable de respuesta $y \in \mathbb{R}$ satisface la siguiente relación

$$y = f(\mathbf{x}, \boldsymbol{\theta}) + \sigma\epsilon, \quad (1.1)$$

donde $\mathbf{x} \in \mathbb{R}^k$ es el vector de covariables con distribución G_x , $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^p$ es el vector de parámetros desconocidos a estimar, $\sigma \in \mathbb{R}_+$ es el parámetro de escala y ϵ es el término de error independiente de las variables explicativas \mathbf{x} . Finalmente, suponemos que la función de regresión f es conocida. En el contexto clásico, se supone además que $\mathbb{E}(\epsilon) = 0$ y $\text{Var}(\epsilon) = 1$.

Entre los modelos no lineales más conocidos, encontramos el modelo de crecimiento

(o de decrecimiento) exponencial dado por

$$f(x, \boldsymbol{\theta}) = \theta_1 \exp(\theta_2 x) + \theta_3, \quad (1.2)$$

donde $\theta_2 > 0$ (o $\theta_2 < 0$), que, por ejemplo, se utiliza para el modelado del crecimiento de poblaciones. Asimismo, es muy usado el modelo de crecimiento logístico de la forma

$$f(x, \boldsymbol{\theta}) = \frac{\theta_1}{1 + \exp(-\theta_2 - \theta_3 x)},$$

con el que se modelan datos con un rápido crecimiento inicial y luego más lento a medida que se aproximan a su valor asintótico, como en el caso de la difusión de un virus o la saturación de un mercado. Si bien los dos ejemplos anteriores tienen una sola covariable, en otras ocasiones se dispone de más variables regresoras. Estos dos ejemplos son casos particulares de los modelos no lineales llamados separables en los que la función de regresión puede escribirse como

$$f(\mathbf{x}, \boldsymbol{\theta}) = \boldsymbol{\theta}_{(1)}^t \mathbf{h}(\mathbf{x}, \boldsymbol{\theta}_{(2)}) = \sum_{i=1}^{p_1} \theta_{(1)i} h_i(\mathbf{x}, \boldsymbol{\theta}_{(2)}),$$

para $\boldsymbol{\theta}^t = (\boldsymbol{\theta}_{(1)}^t, \boldsymbol{\theta}_{(2)}^t)$, $\boldsymbol{\theta}_{(1)} \in \mathbb{R}^{p_1}$, $\boldsymbol{\theta}_{(2)} \in \mathbb{R}^{p_2}$ y $\mathbf{h}(\mathbf{x}, \boldsymbol{\theta}_{(2)}) = (h_1(\mathbf{x}, \boldsymbol{\theta}_{(2)}), \dots, h_{p_1}(\mathbf{x}, \boldsymbol{\theta}_{(2)}))$.

Por otro lado, la mayoría de los métodos estadísticos clásicos suponen que el modelo que genera los datos es conocido y suelen ser muy sensibles a la presencia de datos atípicos, al incumplimiento de las hipótesis que subyacen debajo del modelo asumido como verdadero, aunque sean pequeños apartamientos del mismo. En el contexto de modelos no lineales, el estimador clásico está basado en el procedimiento de mínimos cuadrados y es bien sabido, que puede ser seriamente afectado por una pequeña cantidad de puntos atípicos.

A modo de ejemplo, consideremos el siguiente conjunto de datos en el que se estudia la relación entre el peso del ojo y la edad de un conejo. El conejo europeo *Oryctolagus cuniculus* es una plaga en Australia. Desde el punto de vista ecológico sería de interés disponer de un método fiable para determinar la edad de los conejos capturados en el medio silvestre. En un estudio realizado por Dudzinski y Mykytowycz (1961), se midió el peso seco de la lente del ojo para 71 conejos salvajes, con edades conocidas, viviendo a la intemperie. El peso de la lente del ojo tiende a variar mucho menos que el peso total del cuerpo con las condiciones ambientales y, por este motivo, podría ser un mejor indicador de la edad. El scatter-plot de los datos se muestra en el panel de la izquierda de la Figura 1.1.

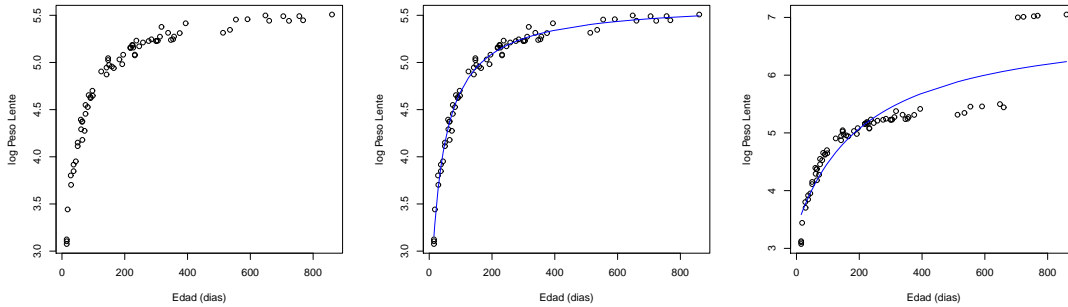


Figura 1.1: Datos originales y contaminados del estudio de conejos.

A partir del estudio de Dudzinski y Mykytowycz (1961) se sugiere una relación no lineal entre el logaritmo del peso de la lente (y) y la edad del conejo (x) a través de la función

$$f(x, \theta_1, \theta_2, \theta_3) = \theta_1 - \frac{\theta_2}{\theta_3 + x}.$$

En el panel central de la Figura 1.1, se muestra en azul la curva ajustada por mínimos cuadrados a los datos obtenidos, mientras que en el panel de la derecha se muestra el ajuste obtenido por el método de mínimos cuadrados después de haber modificado las respuestas de sólo 5 observaciones y se puede apreciar cómo el ajuste cambia en forma drástica y se empeora notablemente.

Este ejemplo da evidencia de que el estimador de mínimos cuadrados puede ser muy influenciado por la presencia de unos pocos datos anómalos y por esta razón es que se necesitan estimadores resistentes.

En contraste con esta situación, la Estadística Robusta propone procedimientos que son resistentes a desviaciones del modelo central, conduciendo a inferencias confiables aún cuando estas desviaciones se produzcan y que, al mismo tiempo, pierdan poca eficiencia cuando el modelo central es válido. En este sentido, volviendo al ejemplo presentado de los conejos, cuando ajustamos el modelo propuesto al conjunto de datos modificados con el procedimiento robusto que proponemos en este trabajo, obtenemos la curva en rojo de la Figura 1.2, que como vemos no es influenciada por los datos atípicos introducidos.

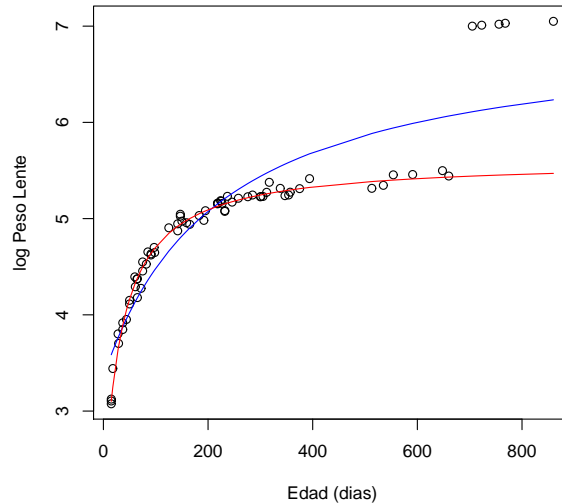


Figura 1.2: Datos contaminados del estudio de conejos.

Algunas propuestas robustas para regresión lineal han sido extendidas al caso no lineal. En uno de los primeros artículos que siguen estos lineamientos, Fraiman (1983) propone un M -estimador general de influencia acotada. Stromberg y Ruppert (1992) abordan el concepto de punto de ruptura bajo el modelo de regresión no lineal y muestran que, para la mayoría de las funciones de regresión no lineal, el punto de ruptura del estimador de mínimos cuadrados es $1/n$, donde n es el tamaño de la muestra. Stromberg (1993) introduce un algoritmo para computar estimadores de alto punto de ruptura en regresión no lineal que solo requiere una pequeña cantidad de ajustes por mínimos cuadrados a p puntos. Tabatabai y Argyros (1993) extienden los L -estimadores para los problemas de estimación y testeo de hipótesis. Mukherjee (1996) trata una clase de estimadores robustos basados en mínimas distancias, mientras que Markatou y Manos (1996) consideran el problema de tests de hipótesis robustos. Sakata y White (2001) extienden el uso de S -estimadores a modelos no lineales con observaciones dependientes. Más recientemente, Fasano (2009) estudia la teoría asintótica de los MM - y τ -estimadores bajo condiciones regulares, entre las que se incluye el segundo momento de las covariables \mathbf{x} . Luego, Fasano *et al.* (2012) prueban la continuidad débil, la consistencia de Fisher y la diferenciabilidad de funcionales de estimación correspondientes a estimadores de alto punto de ruptura, tales como S - y MM -estimadores, en regresión lineal y no lineal, en este último caso asumiendo segundo momento finito de las covariables y compacidad del espacio paramétrico.

Todos estos procedimientos robustos fueron aplicados en muestras completas, sin embargo, en la práctica algunas respuestas pueden perderse ya sea simplemente por azar o por diseño, como es el caso de estudios en dos etapas. De hecho, algunas personas pueden rehusarse a responder sobre temas privados o la variable de respuesta puede ser costosa de medir. En otros casos, factores incontrolables pueden introducir pérdida de la información

o pueden ocurrir problemas en el proceso de registro de la misma. Müller (2009) encara el problema de estimar las esperanzas de funciones de las covariables y de la respuesta en regresión no lineal por medio de un estimador completamente imputado bajo el supuesto de respuestas *perdidas al azar* (*missing at random*). Sued y Yohai (2013) introducen estimadores consistentes para funcionales de posición continuos de la variable dependiente en un modelo no lineal con datos faltantes e implementan para la estimación del parámetro de regresión los *MM*-estimadores. Koul, Müller y Schick (2012) dan un método general para deducir, bajo condiciones generales, el comportamiento asintótico de métodos desarrollados para datos completamente observados cuando se aplican a datos con observaciones faltantes.

En el presente trabajo, se unifican los aspectos arriba mencionados a fin de proponer estimadores y tests robustos. Uno de los aportes de esta tesis es el de dar un enfoque unificado al problema de inferencia robusta para modelos no lineales en un contexto en el que puede haber o no respuestas faltantes, pero la covariable \mathbf{x} es completamente observada. De manera que consideramos observaciones (y_i, \mathbf{x}_i^t) que satisfacen un modelo no lineal, donde algunas de las respuestas pueden no haber sido observadas, y en base a los datos disponibles estimaremos el vector de parámetros desconocido y realizaremos tests de hipótesis que involucren a dicho vector.

En primera instancia, en este marco se propondrá una familia de estimadores robustos del parámetro de regresión basados en *M*-estimadores pesados que serán computados a través de un procedimiento *MM*. En segundo término, a partir de estos estimadores, se introducirá un estadístico de tipo Wald, con el fin de obtener tests robustos para chequear hipótesis que involucren al parámetro de interés.

Cabe mencionar que además de incluir respuestas faltantes se obtienen mejoras respecto de los procedimientos ya existentes en dos sentidos. Por un lado, al introducir pesos que controlan el leverage, es decir, a aquellas observaciones con valores atípicos en las covariables que hacen palanca, se puede prescindir de supuestos sobre la existencia del segundo momento de estas últimas y a la vez estos pesos pueden controlar la influencia de los datos atípicos en el estadístico de tipo Wald, permitiendo obtener tests robustos, tanto desde el punto de vista del comportamiento del nivel como de la potencia. Por otro lado, dado que las propiedades asintóticas de los procedimientos propuestos se prueban usando técnicas de procesos empíricos, es posible obtener resultados bajo supuestos más débiles sobre la función de regresión f y sobre el espacio paramétrico Θ .

Todo el trabajo realizado se encuentra organizado de la siguiente manera. En el Capítulo 2 definimos el modelo no lineal con posibles respuestas faltantes. Asimismo describimos los procedimientos de estimación ya estudiados como así también el método robusto propuesto.

En el Capítulo 3 estudiamos las propiedades asintóticas del estimador presentado en esta tesis. Por otra parte, en el Capítulo 4 deducimos su función de influencia y estudiamos su comportamiento a través de dos ejemplos numéricos. Con el objetivo de estudiar el comportamiento de los estimadores propuestos realizamos un estudio de Monte Carlo el cual se encuentra en el Capítulo 5.

En el Capítulo 6 proponemos tests de hipótesis robustos a partir del estimador estudiado. Obtuvimos el comportamiento asintótico de los estadísticos introducidos bajo la hipótesis nula y bajo alternativas contiguas, como así también su función de influencia. Asimismo, realizamos un estudio de simulación con el fin de analizar el comportamiento, tanto en nivel como en potencia, de los tests propuestos y comparar su performance con otros tests existentes.

Finalmente, en el Capítulo 7 presentamos una aplicación de nuestra propuesta a un conjunto de datos reales.

Capítulo 2

Modelos No Lineales con Posibles Respuestas Faltantes

2.1. Introducción

Dada una muestra aleatoria de tamaño n , (y_i, \mathbf{x}_i^t) , $i = 1, \dots, n$, donde $y_i \in \mathbb{R}$, $\mathbf{x}_i \in \mathbb{R}^k$, decimos que sigue un modelo de regresión no lineal si

$$y_i = f(\mathbf{x}_i, \boldsymbol{\theta}) + \sigma \epsilon_i, \quad (2.1)$$

donde f es una función conocida que es no lineal en el vector de parámetros $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^p$, los cuales son desconocidos y son el objetivo de interés, $\sigma \in \mathbb{R}_+$ es el parámetro de escala y ϵ_i son los errores. En el contexto de este trabajo asumiremos que ϵ_i , con $i = 1, \dots, n$, son independientes, idénticamente distribuidos (i.i.d.), con distribución simétrica e independientes de $\mathbf{x}_i \sim G_x$.

Como ya hemos mencionado daremos un enfoque unificado al problema de inferencia robusta para modelos no lineales en un contexto en el que puede haber o no respuestas faltantes, pero la covariable \mathbf{x} es completamente observada. Bajo esta situación, se asume que se tienen n observaciones $(y_i, \mathbf{x}_i^t, \delta_i)$, $i = 1, \dots, n$, con función de distribución G , donde δ_i es una variable aleatoria dicotómica utilizada para indicar la pérdida de las respuestas, es decir, $\delta_i = 1$ si y_i es observada y $\delta_i = 0$ si y_i es faltante.

Existen diferentes mecanismos que conducen a la pérdida de datos. Nosotros consideraremos el caso en el que subyace un mecanismo ignorable que produce los datos faltantes, imponiendo que la variable binaria y la variable de respuesta son condicionalmente independientes dada la covariable, a esta condición se la denomina *perdidos al azar* (missing at random - MAR). En otras palabras, si el vector $(y, \mathbf{x}^t, \delta)$ tiene distribución G igual que $(y_i, \mathbf{x}_i^t, \delta_i)$, decimos que cumple la condición MAR si

$$P(\delta = 1 | (y, \mathbf{x})) = P(\delta = 1 | \mathbf{x}) = p(\mathbf{x}). \quad (2.2)$$

De aquí en más supondremos que la muestra aleatoria $(y_i, \mathbf{x}_i^t, \delta_i)$, $1 \leq i \leq n$, cumple la condición (2.2).

2.2. Estimación Clásica

Supongamos primero que tenemos una muestra aleatoria completa, es decir $(y_i, \mathbf{x}_i^t, \delta_i)$, con $\delta_i = 1$ para todo $i = 1, \dots, n$, que cumplen el modelo de regresión no lineal (2.1). En el caso clásico habitualmente el supuesto que se realiza respecto de los errores es que $\mathbb{E}(\epsilon_i) = 0$ y $Var(\epsilon_i) = E(\epsilon_i^2) = 1$ para todo i , mientras que en el contexto de robustez se evitan suposiciones acerca de la existencia de momentos de los errores.

Como mencionamos en la introducción, el estimador clásico está basado en el método de mínimos cuadrados (MC), es decir

$$\begin{aligned}\widehat{\boldsymbol{\theta}}_{MC} &= \operatorname{argmin}_{\mathbf{b} \in \Theta} \sum_{i=1}^n (y_i - f(\mathbf{x}_i, \mathbf{b}))^2, \\ &= \operatorname{argmin}_{\mathbf{b} \in \Theta} \sum_{i=1}^n \delta_i (y_i - f(\mathbf{x}_i, \mathbf{b}))^2.\end{aligned}$$

Bajo ciertos supuestos de regularidad (ver Seber y Wild, 1989), $\widehat{\boldsymbol{\theta}}_{MC}$ y

$s^2 = \frac{1}{n-p} \sum_{i=1}^n \delta_i (y_i - f(\mathbf{x}_i, \widehat{\boldsymbol{\theta}}_{MC}))^2$ son estimadores consistentes de $\boldsymbol{\theta}$ y σ^2 respectivamente y además, $\widehat{\boldsymbol{\theta}}_{MC}$ resulta asintóticamente normal.

De esta manera, volviendo al caso en el que hay respuestas perdidas al azar, es decir $\delta_i = 0$ para algunos $i \in \{1, \dots, n\}$, podemos considerar el estimador clásico simplificado definido por

$$\widehat{\boldsymbol{\theta}}_{MC} = \operatorname{argmin}_{\mathbf{b} \in \Theta} \sum_{i=1}^n \delta_i (y_i - f(\mathbf{x}_i, \mathbf{b}))^2, \quad (2.3)$$

que es el estimador de mínimos cuadrados calculado a partir de los datos completos disponibles. Notemos que este estimador es solución del sistema

$$\sum_{i=1}^n \delta_i r_i(\mathbf{b}) \dot{\mathbf{f}}(\mathbf{x}_i, \mathbf{b}) = \mathbf{0}, \quad (2.4)$$

donde $r_i(\mathbf{b}) = y_i - f(\mathbf{x}_i, \mathbf{b})$ son los residuos y el gradiente $\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b}) = (\dot{f}_1(\mathbf{x}, \mathbf{b}), \dots, \dot{f}_p(\mathbf{x}, \mathbf{b}))^t$ con $\dot{f}_j(\mathbf{x}, \mathbf{b}) = \frac{\partial f(\mathbf{x}, \mathbf{b})}{\partial b_j}$. Por lo tanto, una observación atípica puede influenciar la determinación del estimador ya sea a través del residuo o del leverage, medido en términos del gradiente $\dot{\mathbf{f}}$.

Cabe destacar que, a partir del principio de transferencia definido por Koul *et al.* (2012), las propiedades asintóticas del estimador de mínimos cuadrados obtenidas para muestras completas pueden ser transferidas, bajo el supuesto MAR, al estimador obtenido en (2.3).

2.3. Estimación Robusta

Si bien el procedimiento de mínimos cuadrados resulta óptimo bajo normalidad, es sensible al apartamiento de esta hipótesis, en tanto que pequeñas desviaciones podrían producir grandes fluctuaciones del estimador, invalidando las conclusiones basadas en este método. En la introducción hemos ilustrado mediante un ejemplo cuán sensible puede ser este estimador a la presencia de unos pocos datos atípicos y cómo estos pueden volver completamente inestable al estimador clásico.

El objetivo de los métodos robustos es desarrollar estimadores que tengan un “buen” comportamiento en un “entorno” del modelo. Es decir, que sean resistentes a estos apartamientos y que a la vez sean eficientes cuando el modelo central es válido.

2.3.1. M - y MM -estimadores

Una forma de controlar las observaciones atípicas es reemplazar, en el proceso de minimización (2.3) la función cuadrática por una ρ -función en el sentido de Maronna *et al.* (2006), esto es una función que satisfaga las siguientes condiciones

1. $\rho(x)$ es creciente como función del $|x|$,
2. $\rho(0) = 0$,
3. $\rho(x)$ es estrictamente creciente para $x > 0$ tal que $\rho(x) < \rho(\infty)$,
4. si ρ es acotada, se supone también que $\rho(\infty) = 1$.

De esta forma, el M -estimador simplificado basado en las observaciones completas disponibles quedaría definido por

$$\widehat{\boldsymbol{\theta}} = \underset{\mathbf{b} \in \Theta}{\operatorname{argmin}} \sum_{i=1}^n \delta_i \rho(y_i - f(\mathbf{x}_i, \mathbf{b})).$$

Sin embargo, el estimador resultante no sería invariante por cambios de escala, por este motivo es necesario dividir el argumento de la función ρ por $\widehat{\sigma} = \widehat{\sigma}_n$ un estimador robusto preliminar de la escala de los residuos. De esta manera, se define el M -estimador como

$$\widehat{\boldsymbol{\theta}}_M = \underset{\mathbf{b} \in \Theta}{\operatorname{argmin}} \sum_{i=1}^n \delta_i \rho\left(\frac{y_i - f(\mathbf{x}_i, \mathbf{b})}{\widehat{\sigma}}\right). \quad (2.5)$$

Una elección posible que satisfaga las condiciones de una ρ -función es la familia de funciones bicuadrada de Tukey dada por

$$\rho(t) = \begin{cases} 1 - [1 - (t/c)^2]^3 & \text{si } |t| \leq c \\ 1 & \text{si } |t| > c, \end{cases} \quad (2.6)$$

donde c es una constante de calibración que se elige con el fin de asegurar una eficiencia asintótica deseada.

El rol del estimador de σ en (2.5) es crítico ya que determinará qué residuos son grandes y por lo tanto, deben ser acotados por la función de pérdida ρ . Por este motivo, es necesario que $\widehat{\sigma}$ sea un estimador robusto de la escala. Huber (1981) introduce los M -estimadores de escala de la siguiente manera. Sea ρ una ρ -función en el sentido de Maronna *et al.* (2006), luego dada una muestra aleatoria de tamaño n , $\mathbf{u} = (u_1, \dots, u_n)$, su M -estimador de escala $\widehat{\sigma}(\mathbf{u})$ es el valor s que es solución de la siguiente ecuación

$$\frac{1}{n} \sum_{i=1}^n \rho\left(\frac{u_i}{s}\right) = a, \quad (2.7)$$

donde $a = \mathbb{E}_\phi(\rho(u))$, y ϕ es la función de distribución de una normal estándar.

Con el objetivo de obtener un estimador con alto punto de ruptura, que además sea altamente eficiente bajo distribución normal, Yohai (1987) introdujo los MM -estimadores para el parámetro de regresión lineal, que fue extendido para el modelo no lineal en Fasano (2009). El MM -estimador en un modelo no lineal con respuestas faltantes se puede definir a través de los siguientes tres pasos:

Paso 1. Calcular un estimador inicial $\widehat{\boldsymbol{\theta}}_1$ de $\boldsymbol{\theta}$ con alto punto de ruptura.

Paso 2. Calcular los residuos $r_i(\widehat{\boldsymbol{\theta}}_1) = y_i - f(\mathbf{x}_i, \widehat{\boldsymbol{\theta}}_1)$, para i tal que $\delta_i = 1$ y computar la M -escala basada en la muestra de los residuos observados usando una ρ -función ρ_o con $a = 0,5 \max_t \rho_o(t)$, es decir que si $m = \sum_{i=1}^n \delta_i$, $\widehat{\sigma}$ es tal que

$$\frac{1}{m} \sum_{i=1}^n \delta_i \rho_o\left(\frac{r_i(\widehat{\boldsymbol{\theta}}_1)}{\widehat{\sigma}}\right) = a.$$

Paso 3. Sea ρ_1 otra ρ -función tal que $\rho_1(t) \leq \rho_o(t)$, $\sup_t \rho_1(t) = \sup_t \rho_o(t)$, luego el MM -estimador se define como

$$\widehat{\boldsymbol{\theta}}_{MM} = \operatorname{argmin}_{\mathbf{b} \in \Theta} \sum_{i=1}^n \delta_i \rho_1\left(\frac{r_i(\mathbf{b})}{\widehat{\sigma}}\right). \quad (2.8)$$

Observación 2.3.1 a) Si derivamos con respecto a \mathbf{b} la función objetivo en (2.8), $\widehat{\boldsymbol{\theta}}_{MM}$ es solución del siguiente sistema de ecuaciones

$$\sum_{i=1}^n \delta_i \psi_1\left(\frac{r_i(\mathbf{b})}{\widehat{\sigma}}\right) \dot{\mathbf{f}}(\mathbf{x}_i, \mathbf{b}) = \mathbf{0},$$

donde $\psi_1 = \rho_1'$.

b) Las propiedades asintóticas de los M - y MM -estimadores del parámetro de regresión en modelos no lineales con datos completos fueron estudiadas por Fasano (2009)

y Fasano et. al. (2012). Sued y Yohai (2013) introducen estimadores consistentes de un funcional de posición continuo de la respuesta en un modelo no lineal con observaciones faltantes. El estimador marginal necesita un estimador previo del parámetro de regresión no lineal y con este fin estos autores implementan los MM-estimadores. Asimismo, estudian sus propiedades asintóticas cuando hay respuestas perdidas al azar (MAR) bajo ciertas condiciones de regularidad que incluyen la existencia del segundo momento de las covariables \mathbf{x} y que el espacio paramétrico Θ sea compacto.

- c) Respecto del estimador inicial necesario en el PASO 1 distintas alternativas son posibles, entre ellas, podemos mencionar los S-estimadores y el estimador LMS. Para el caso de datos completos, Fasano (2009) implementa en su estudio numérico los MM-estimadores tomando como estimador inicial un S-estimador, que se define mediante la minimización de una medida robusta de la escala de los residuos. Para el caso que tratamos aquí los llamados S-estimadores quedan definidos por

$$\widehat{\boldsymbol{\theta}}_S = \underset{\mathbf{b} \in \Theta}{\operatorname{argmin}} \widehat{\sigma}(\mathbf{r}(\mathbf{b})), \quad (2.9)$$

donde $\mathbf{r}(\mathbf{b})$ es el vector de los m residuos observados y $\widehat{\sigma}(\mathbf{r})$ es la M -escala definida por (2.7). Un S-estimador muy popular es el estimador LMS que se obtiene al minimizar la mediana de los residuos al cuadrado, es decir

$$\widehat{\boldsymbol{\theta}}_{\text{LMS}} = \underset{\mathbf{b} \in \Theta}{\operatorname{argmin}} \operatorname{MED}_{i:\delta_i=1} (r_i(\mathbf{b})^2). \quad (2.10)$$

El punto de ruptura del estimador LMS en modelos no lineales fue estudiado por Stromberg y Ruppert (1992) y en Stromberg (1993) se presenta un algoritmo para su cómputo. Por otra parte, en ocasiones, es posible desarrollar un estimador inicial robusto ad hoc. Este es el caso, por ejemplo, de aquellos modelos que se pueden aproximar mediante una linealización y en tal situación, pueden tomarse como estimadores iniciales los propuestos en su trabajo original por Yohai (1987), entre ellos, los estimadores basados en medianas repetidas de Siegel (1982).

2.3.2. Propuesta

En la ecuación (2.4) notamos que una observación atípica puede tener influencia en la determinación del estimador de mínimos cuadrados tanto a través del residuo como del leverage. Nuestra propuesta surge con el doble propósito de controlar estos dos efectos, ya que si bien la función ρ es acotada y el estimador resultante será robusto, cuando se presentan outliers de alto leverage se espera aumentar la robustez al acotar su influencia mediante la introducción de una función de peso. Por este motivo, incluimos en la función objetivo de (2.8) una función de peso w , que será una función de las covariables \mathbf{x} y por lo general, dependerá de otros parámetros, que llamaremos genéricamente $\boldsymbol{\Gamma}$. Distintas elecciones de la función de peso son posibles. Podría ser que el usuario elija acotar el efecto del gradiente en forma directa y en este caso, podría tomar $w(\mathbf{x}, \widehat{\boldsymbol{\theta}}_1) = w_o(|\dot{\mathbf{f}}(\mathbf{x}, \widehat{\boldsymbol{\theta}}_1)|)$, donde $\widehat{\boldsymbol{\theta}}_1$ es el estimador inicial del Paso 1 y $w_o(t) = \psi_o(t)/t$, siendo $\psi_o = \rho'_o$ y ρ_o una

ρ -función como las ya definidas. También es frecuente el uso de funciones w_o de tipo *hard rejection*, que son funciones indicadoras o características de un intervalo. En otros casos, el control del gradiente, puede lograrse penalizando directamente las covariables de alta palanca y en esta situación, el usuario podría elegir una función de peso que penalice la distancia de Mahalanobis robusta de las covariables, para lo cual, será necesario estimar los parámetros de posición y escala de \mathbf{x} , digamos $\boldsymbol{\mu}$ y matriz de escala $\boldsymbol{\Upsilon}$. Así, si $\widehat{\boldsymbol{\mu}} = \widehat{\boldsymbol{\mu}}_n$ y $\widehat{\boldsymbol{\Upsilon}} = \widehat{\boldsymbol{\Upsilon}}_n$ son estimadores robustos de la posición y escala de \mathbf{x} basados en $\mathbf{x}_1, \dots, \mathbf{x}_n$, entonces podrían considerarse pesos de la forma

$$w(\mathbf{x}, \widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Upsilon}}) = w_o \left(\sqrt{(\mathbf{x} - \widehat{\boldsymbol{\mu}})^t \widehat{\boldsymbol{\Upsilon}}^{-1} (\mathbf{x} - \widehat{\boldsymbol{\mu}})} \right). \quad (2.11)$$

En forma genérica, consideraremos pesos $w = w(\mathbf{x}, \widehat{\boldsymbol{\Gamma}})$ que dependen de estimadores $\widehat{\boldsymbol{\Gamma}} = \widehat{\boldsymbol{\Gamma}}_n$ de parámetros desconocidos $\boldsymbol{\Gamma} \in \Lambda$, de manera que $w : \mathbb{R}^k \times \Lambda \mapsto \mathbb{R}_{\geq 0}$. Asimismo, es de esperar que al introducir estos pesos también se obtenga un beneficio al realizar inferencia sobre el parámetro de interés $\boldsymbol{\theta}$, ya sea al computar intervalos de confianza o tests de hipótesis de tipo Wald basados en estos estimadores. Estas ventajas de los estimadores pesados como los aquí presentados han sido comprobadas en el contexto de modelos lineales generalizados (ver, por ejemplo, Bianco, Boente y Rodrigues, 2013a y 2013b) ya que los pesos también intervienen en el control del crecimiento de la matriz de covarianza asintótica estimada, necesaria para realizar inferencia.

Teniendo en cuenta todas estas consideraciones, dada una muestra aleatoria $(y_i, \mathbf{x}_i^t, \delta_i)$, $1 \leq i \leq n$, que sigue un modelo no lineal con respuestas perdidas al azar tal como fue descrito en la Sección 2.1, proponemos el estimador $\widehat{\boldsymbol{\theta}}$ definido como

$$\widehat{\boldsymbol{\theta}} = \underset{\mathbf{b} \in \Theta}{\operatorname{argmin}} \sum_{i=1}^n \delta_i \rho_1 \left(\frac{r_i(\mathbf{b})}{\widehat{\sigma}} \right) w(\mathbf{x}_i, \widehat{\boldsymbol{\Gamma}}). \quad (2.12)$$

Considerando los aspectos descriptos en la sección anterior, será de capital importancia la estimación robusta del desvío y del parámetro auxiliar $\boldsymbol{\Gamma}$. Por esta razón, proponemos el siguiente procedimiento de cuatro pasos:

Paso 1. Computar un estimador inicial $\widehat{\boldsymbol{\theta}}_1$ robusto de $\boldsymbol{\theta}$ con alto punto de ruptura.

Paso 2. Calcular los residuos $r_i(\widehat{\boldsymbol{\theta}}_1) = y_i - f(\mathbf{x}_i, \widehat{\boldsymbol{\theta}}_1)$, para i tal que $\delta_i = 1$ y computar la M -escala basada en la muestra de los residuos observados usando una ρ -función ρ_o con $a = 0,5 \max_t \rho_o(t)$, es decir que si $m = \sum_{i=1}^n \delta_i$, $\widehat{\sigma}$ es tal que

$$\frac{1}{m} \sum_{i=1}^n \delta_i \rho_o \left(\frac{r_i(\widehat{\boldsymbol{\theta}}_1)}{\widehat{\sigma}} \right) = a. \quad (2.13)$$

Paso 3. Calcular $\widehat{\boldsymbol{\Gamma}}$ un estimador robusto de $\boldsymbol{\Gamma}$ y computar $w(\mathbf{x}_i, \widehat{\boldsymbol{\Gamma}})$ para todo $i = 1, \dots, n$.

PASO 4. Considerar ρ_1 otra ρ -función tal que $\rho_1(t) \leq \rho_o(t)$, $\sup_t \rho_1(t) = \sup_t \rho_o(t)$ y computar el estimador como

$$\widehat{\boldsymbol{\theta}} = \operatorname{argmin}_{\mathbf{b} \in \Theta} \sum_{i=1}^n \delta_i \rho_1 \left(\frac{r_i(\mathbf{b})}{\widehat{\sigma}} \right) w(\mathbf{x}_i, \widehat{\Gamma}). \quad (2.14)$$

El estimador resultante de este procedimiento *MM* lo llamaremos *WMM*-estimador (*WMM*).

Observación 2.3.2 a) Si derivamos la función objetivo de (2.12) respecto a \mathbf{b} e igualamos a cero, el estimador $\widehat{\boldsymbol{\theta}}$ es solución del siguiente sistema de ecuaciones

$$\sum_{i=1}^n \delta_i \psi_1 \left(\frac{r_i(\mathbf{b})}{\widehat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \mathbf{b}) w(\mathbf{x}_i, \widehat{\Gamma}) = \mathbf{0},$$

donde $\psi_1 = \rho'_1$ y $\dot{\mathbf{f}}$ es el gradiente de la función de regresión f con respecto a \mathbf{b} .

- b) En el PASO 1 podría computarse como estimador inicial cualquiera de las opciones descritas para los *MM*-estimadores en la sección anterior. Los mismos *MM*-estimadores podrían usarse como estimadores iniciales en esta etapa del procedimiento.
- c) Tanto para el PASO 2 como para el PASO 4, se pueden elegir las funciones ρ_0 y ρ_1 dentro de la familia de funciones bicuadradas de Tukey, que fue definida en (2.6).
- d) La función w_o se puede tomar dentro de la familia de funciones de peso inducida por la función bicuadrada de Tukey. Es decir que, si definimos $\psi_o(t) = \rho'_o(t)$, siendo ρ_o como en (2.6), luego $w_o(t) = \psi_o(t)/t$.

En el próximo capítulo analizaremos el comportamiento asintótico de los estimadores propuestos.

Capítulo 3

Comportamiento Asintótico

Sean $(y_i, \mathbf{x}_i^t, \delta_i)$, $i = 1, \dots, n$, vectores aleatorios independientes con distribución G que siguen un modelo de regresión no lineal con respuestas perdidas al azar, tal como fue descrito en la Sección 2.1. A lo largo de todo este capítulo y de los siguientes asumiremos, con el objetivo de que los parámetros queden identificados, que se cumple la siguiente condición de identificabilidad

I. $P(f(\mathbf{x}, \boldsymbol{\theta}) = f(\mathbf{x}, \boldsymbol{\theta}^*)) < 1$, para todo $\boldsymbol{\theta}^* \neq \boldsymbol{\theta}$.

3.1. Consistencia de Fisher

En esta sección comenzaremos por estudiar la consistencia de Fisher del estimador robusto propuesto en la Sección 2.3.2.

Definiendo

$$\begin{aligned} H_n(\mathbf{b}, s, \mathbf{G}) &= \frac{1}{n} \sum_{i=1}^n \delta_i \rho_1 \left(\frac{y_i - f(\mathbf{x}_i, \mathbf{b})}{s} \right) w(\mathbf{x}_i, \mathbf{G}) \quad y \\ \sigma_n(\mathbf{b}) &: \sum_{i=1}^n \delta_i \rho_o \left(\frac{y_i - f(\mathbf{x}_i, \mathbf{b})}{\sigma_n(\mathbf{b})} \right) = a \sum_{i=1}^n \delta_i \quad , \end{aligned} \quad (3.1)$$

el estimador de $\boldsymbol{\theta}$ propuesto se define como

$$\widehat{\boldsymbol{\theta}} = \underset{\mathbf{b} \in \Theta}{\operatorname{argmin}} H_n(\mathbf{b}, \widehat{\sigma}, \widehat{\Gamma}),$$

donde, $\widehat{\sigma} = \sigma_n(\widehat{\boldsymbol{\theta}}_1)$ con $\widehat{\boldsymbol{\theta}}_1$ el estimador inicial del Paso 1 y $\widehat{\Gamma}$ es un estimador robusto del parámetro auxiliar Γ . Cuando la muestra está totalmente observada, $\delta_i = 1$ para todo i y en (3.1) tenemos que $\sum_{i=1}^n \delta_i = n$.

Recordemos que si ρ_1 es diferenciable, tal que $\psi_1 = \rho_1'$ y $\dot{\mathbf{f}}_j = \partial f / \partial b_j$, $\widehat{\boldsymbol{\theta}}$ es solución

del sistema de ecuaciones $H_n^{(1)}(\widehat{\boldsymbol{\theta}}, \widehat{\sigma}, \widehat{\boldsymbol{\Gamma}}) = \mathbf{0}$, donde

$$H_n^{(1)}(\mathbf{b}, s, \mathbf{G}) = \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{y_i - f(\mathbf{x}_i, \mathbf{b})}{s} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \mathbf{b}) w(\mathbf{x}_i, \mathbf{G}). \quad (3.2)$$

Consideremos ahora los funcionales relacionados con $H_n(\mathbf{b}, s, \mathbf{G})$ y $\sigma_n(\mathbf{b})$ dados por

$$H(\mathbf{b}, s, \mathbf{G}) = \mathbb{E}_G \left[\delta \rho_1 \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{s} \right) w(\mathbf{x}, \mathbf{G}) \right] \quad (3.3)$$

$$\sigma(\mathbf{b}) : \mathbb{E}_G \left[\delta \rho_o \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{\sigma(\mathbf{b})} \right) \right] = a \mathbb{E}(p(\mathbf{x})). \quad (3.4)$$

Probaremos que el funcional es Fisher-consistente, es decir que $H(\mathbf{b}, \sigma, \boldsymbol{\Gamma})$ tiene un único mínimo en $\boldsymbol{\theta}$ en el contexto de regresión no lineal con respuestas faltantes. Para ello necesitamos suponer que se satisfacen las siguientes condiciones,

A0. dado $\tau > 0$, $\mathbb{E}(\rho_1[(\epsilon - a)/\tau]) > \mathbb{E}(\rho_1(\epsilon/\tau))$, $\forall a \neq 0$.

A1. $P(\{\mathbf{x} \in S_w \cap S_p : f(\mathbf{x}, \boldsymbol{\theta}) \neq f(\mathbf{x}, \boldsymbol{\theta}^*)\}) > 0$, para todo $\boldsymbol{\theta}^* \neq \boldsymbol{\theta}$, donde S_w y S_p son los soportes de $w(\mathbf{x}, \boldsymbol{\Gamma})$ y $p(\mathbf{x})$ respectivamente.

Observación 3.1.1 La primera condición, **A0**, es necesaria ya que al utilizar funciones de pérdida redescendentes para controlar los residuos se necesitan supuestos más fuertes para obtener la unicidad deseada, tales como simetría y unimodalidad de la distribución subyacente. De hecho, denotemos por $\lambda(a, \tau) = \mathbb{E}(\rho_1[(\epsilon - a)/\tau])$ y definamos $v = \operatorname{argmin}_a \lambda(a, \tau_o)$, siendo τ_o el valor de la escala. Por el Teorema 10.2 de Maronna et al. (2006), si la función de densidad de ϵ es estrictamente decreciente como función de $|\epsilon - v|$ y ρ_1 es una ρ -función, entonces, $\lambda(a, \tau)$ tiene un único mínimo en $a = v$ para todo $\tau > 0$. Se pueden ver más detalles en Fasano et al. (2012).

La segunda hipótesis **A1**, es un requerimiento natural para preservar la información contenida en el conjunto de datos aún cuando haya datos faltantes y haya observaciones con peso nulo. Por otro lado, por la condición de identificabilidad **I, A1** se cumple directamente si $w(\mathbf{x}, \boldsymbol{\Gamma})$ y $p(\mathbf{x})$ son estrictamente positivas o si $P(S_w \cap S_p) = 1$. Asimismo, ésta es una condición más débil que pedir que la función de peso sea estrictamente positiva o que el $\inf_{\mathbf{x}} p(\mathbf{x}) > 0$, lo cual significa que la variable de respuesta puede ser observada para cualquier valor de la covariable.

Proposición 3.1.1 Supongamos que $(y, \mathbf{x}^t, \delta)$ es un vector aleatorio tal que $y = f(\mathbf{x}, \boldsymbol{\theta}) + \sigma\epsilon$, donde ϵ es independiente del vector de covariables \mathbf{x} y que cumple la condición **MAR** dada en (2.2). Si la función de pérdida ρ_1 verifica **A0** y la función de peso w cumple **A1**, entonces el funcional

$$\boldsymbol{\theta}(G) = \operatorname{argmin}_{\mathbf{b} \in \Theta} \mathbb{E}_G \left[\delta \rho_1 \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{\sigma} \right) w(\mathbf{x}, \boldsymbol{\Gamma}) \right]$$

es Fisher-consistente.

Demostración: Tenemos que, por la condición MAR dada en (2.2)

$$\begin{aligned} \mathbb{E}_G \left[\delta \rho_1 \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{\sigma} \right) w(\mathbf{x}, \Gamma) \right] &= \mathbb{E}_G \left[p(\mathbf{x}) \rho_1 \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{\sigma} \right) w(\mathbf{x}, \Gamma) \right] \\ &= \mathbb{E}_G \left[p(\mathbf{x}) \rho_1 \left(\frac{y - f(\mathbf{x}, \boldsymbol{\theta}) + f(\mathbf{x}, \boldsymbol{\theta}) - f(\mathbf{x}, \mathbf{b})}{\sigma} \right) w(\mathbf{x}, \Gamma) \right] \\ &= \mathbb{E}_G \left[p(\mathbf{x}) w(\mathbf{x}, \Gamma) \mathbb{E} \left\{ \rho_1 \left(\frac{\sigma \epsilon + f(\mathbf{x}, \boldsymbol{\theta}) - f(\mathbf{x}, \mathbf{b})}{\sigma} \right) \middle| \mathbf{x} \right\} \right]. \end{aligned}$$

Cuando $f(\mathbf{x}, \boldsymbol{\theta}) - f(\mathbf{x}, \mathbf{b}) \neq 0$, por **A0**, tenemos que

$$\begin{aligned} \mathbb{E} \left\{ \rho_1 \left(\frac{\sigma \epsilon + f(\mathbf{x}, \boldsymbol{\theta}) - f(\mathbf{x}, \mathbf{b})}{\sigma} \right) \middle| \mathbf{x} \right\} &= \mathbb{E} \left\{ \rho_1 \left(\epsilon + \frac{f(\mathbf{x}, \boldsymbol{\theta}) - f(\mathbf{x}, \mathbf{b})}{\sigma} \right) \middle| \mathbf{x} \right\} \\ &> \mathbb{E} \{ \rho_1(\epsilon) \}. \end{aligned}$$

Finalmente, como vale **A1**, $\boldsymbol{\theta}(G) = \boldsymbol{\theta}$, como queríamos demostrar. \square

El *WMM*-estimador fue definido como el resultado de un procedimiento de 4 pasos. El primero de todos ellos involucra el cálculo de un estimador inicial robusto y éste podría ser un *S*-estimador. En el siguiente lema probaremos que los *S*-estimadores, que resultan Fisher-consistentes en un modelo no lineal con datos completos, también lo son para el caso de respuestas faltantes de acuerdo con un mecanismo de pérdida MAR.

De esta manera, si consideramos $\widehat{\boldsymbol{\theta}}_S = \operatorname{argmin}_{\mathbf{b} \in \Theta} \widehat{\sigma}(\mathbf{r}(\mathbf{b}))$ su funcional asociado se define por

$$\underline{\boldsymbol{\theta}}(G) = \operatorname{argmin}_{\mathbf{b} \in \Theta} \sigma(\mathbf{b}),$$

donde $\sigma(\mathbf{b})$ está definido en (3.4).

Proposición 3.1.2 *Supongamos que $(y, \mathbf{x}^t, \delta)$ es un vector aleatorio tal que $y = f(\mathbf{x}, \boldsymbol{\theta}) + \sigma \epsilon$, que cumple la condición (2.2), donde ϵ es independiente del vector de covariables \mathbf{x} y además tiene función de densidad positiva en algún entorno alrededor del cero. Si se verifican **A0** y **A1**, entonces el funcional*

$$\underline{\boldsymbol{\theta}}(G) = \operatorname{argmin}_{\mathbf{b} \in \Theta} \sigma(\mathbf{b})$$

es Fisher-consistente.

Demostración: Análogamente a la Proposición 3.1.1, tenemos que

$$\begin{aligned} \mathbb{E}_G \left[\delta \rho_o \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{\sigma(\mathbf{b})} \right) \right] &= \mathbb{E}_G \left[\delta \rho_o \left(\frac{y - f(\mathbf{x}, \boldsymbol{\theta}) + f(\mathbf{x}, \boldsymbol{\theta}) - f(\mathbf{x}, \mathbf{b})}{\sigma(\mathbf{b})} \right) \right] \\ &= \mathbb{E}_G \left[p(\mathbf{x}) \rho_o \left(\frac{\sigma \epsilon + f(\mathbf{x}, \boldsymbol{\theta}) - f(\mathbf{x}, \mathbf{b})}{\sigma(\mathbf{b})} \right) \right]. \end{aligned} \quad (3.5)$$

Por **A0**, cuando $f(\mathbf{x}, \boldsymbol{\theta}) - f(\mathbf{x}, \mathbf{b}) \neq 0$, tenemos que para todo $\tau > 0$, se cumple que

$$\begin{aligned} \mathbb{E} \left[p(\mathbf{x}) \rho_o \left(\frac{\sigma \epsilon + f(\mathbf{x}, \boldsymbol{\theta}) - f(\mathbf{x}, \mathbf{b})}{\tau} \right) \middle| \mathbf{x} \right] &= p(\mathbf{x}) \mathbb{E} \left[\rho_o \left(\frac{\sigma \epsilon + f(\mathbf{x}, \boldsymbol{\theta}) - f(\mathbf{x}, \mathbf{b})}{\tau} \right) \middle| \mathbf{x} \right] \\ &> p(\mathbf{x}) \mathbb{E} \left[\rho_o \left(\frac{\sigma \epsilon}{\tau} \right) \middle| \mathbf{x} \right], \end{aligned}$$

por lo tanto, por **A1** y la independencia entre ϵ y \mathbf{x} tenemos que

$$\mathbb{E}_G \left[\delta \rho_o \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{\tau} \right) \right] > \mathbb{E} \left[\delta \rho_o \left(\frac{\sigma \epsilon}{\tau} \right) \right]. \quad (3.6)$$

Asimismo, como $\sigma(\mathbf{b})$ es tal que

$$\begin{aligned} a\mathbb{E}(p(\mathbf{x})) &= \mathbb{E}_G \left[\delta \rho_o \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{\sigma(\mathbf{b})} \right) \right] \\ &= \mathbb{E}_G \left[p(\mathbf{x}) \rho_o \left(\frac{\sigma \epsilon + f(\mathbf{x}, \boldsymbol{\theta}) - f(\mathbf{x}, \mathbf{b})}{\sigma(\mathbf{b})} \right) \right], \end{aligned}$$

usando (3.5) y (3.6) resulta que

$$\mathbb{E} \left[\delta \rho_o \left(\frac{\sigma \epsilon}{\sigma(\mathbf{b})} \right) \right] < a\mathbb{E}(p(\mathbf{x})).$$

Con lo cual

$$\mathbb{E}_G \left[\delta \rho_o \left(\frac{y - f(\mathbf{x}, \boldsymbol{\theta})}{\sigma(\mathbf{b})} \right) \right] < a\mathbb{E}(p(\mathbf{x})).$$

Finalmente, dado que $\mathbb{E}(\delta \rho_o(\sigma \epsilon / \tau))$ es estrictamente decreciente en τ por tener ϵ densidad positiva en un entorno del 0, tenemos que

$$\sigma(\boldsymbol{\theta}) < \sigma(\mathbf{b}),$$

lo que completa la demostración. \square

3.2. Resultados previos: Teoría de Procesos Empíricos

Con el objetivo de demostrar que los estimadores propuestos en la Sección 2.3.2 son fuertemente consistentes y asintóticamente normales necesitamos algunas definiciones y resultados de procesos empíricos. A los fines de completitud de esta tesis, presentamos en esta Sección un resumen de estos resultados, en su mayoría, extraídos de Pollard (1984) y de van der Vaart y Wellner (1996).

Definición 3.2.1 Sea \mathcal{F} una clase de funciones, decimos que \mathcal{F} es permisible si puede ser indexada por un conjunto T , es decir $\mathcal{F} = \{f(\cdot, t) : t \in T\}$.

Definición 3.2.2 Sea \mathcal{F} una clase de funciones, decimos que $F : \mathcal{X} \rightarrow \mathbb{R}$ es una envolvente para \mathcal{F} si $|f(x)| \leq F(x)$ para todo $x \in \mathcal{X}$ y para toda $f \in \mathcal{F}$.

Sea $L_r(\mathbb{P})$ el conjunto de funciones $g : \mathcal{X} \rightarrow \mathbb{R}$ tales que $\|g\|_{r,\mathbb{P}} = [\mathbb{E}_{\mathbb{P}}(|g(x)|^r)]^{1/r} < \infty$.

Definición 3.2.3 Definimos ε -bracket en $L_r(\mathbb{P})$ a un par de funciones $l, u \in L_r(\mathbb{P})$ tales que $P[l(x) \leq u(x)] = 1$ y $\|l - u\|_{r,\mathbb{P}} \leq \varepsilon$. Decimos que dada $f \in \mathcal{F}$ pertenece al bracket l, u si $P[l(x) \leq f(x) \leq u(x)] = 1$.

Definición 3.2.4 Se define número de bracketing a $N_{[\cdot]}(\varepsilon, \mathcal{F}, L_r(\mathbb{P}))$, el mínimo número de ε -brackets en $L_r(\mathbb{P})$ necesario para asegurar que toda $f \in \mathcal{F}$ pertenece a al menos uno de ellos. Al logaritmo del número de bracketing se lo denomina entropía con bracketing.

Definición 3.2.5 En $L_r(\mathbb{P})$ definimos ε -bola alrededor de una función $g \in L_r(\mathbb{P})$ al conjunto $\{h \in L_r(\mathbb{P}) : \|h - g\|_{r,\mathbb{P}} < \varepsilon\}$.

Definición 3.2.6 Se define número de cubrimiento $N(\varepsilon, \mathcal{F}, L_r(\mathbb{P}))$ al mínimo número de ε -bolas en $L_r(\mathbb{P})$ necesarias para cubrir \mathcal{F} , es decir que todos los elementos de \mathcal{F} deben estar contenidos en al menos una de las bolas, pero no es necesario que el centro de las mismas pertenezca a la familia \mathcal{F} . La entropía es el logaritmo del número de cubrimiento.

El número de cubrimiento se puede interpretar como la riqueza que tiene la clase \mathcal{F} a escala ε .

Teorema 3.2.1 Sea \mathcal{F} una clase de funciones tal que para cada $\varepsilon > 0$, $N_{[\cdot]}(\varepsilon, \mathcal{F}, L_1(\mathbb{P})) < \infty$, entonces $\sup_{f \in \mathcal{F}} |\mathbb{E}_{\mathbb{P}_n}(f) - \mathbb{E}_{\mathbb{P}}(f)| \xrightarrow{c.s.} 0$, donde \mathbb{P}_n es la distribución empírica.

Teorema 3.2.2 Sea \mathcal{F} una clase de funciones permisible con función envolvente F tal que $\mathbb{E}_{\mathbb{P}}(F) < \infty$. Si \mathbb{P}_n es la distribución empírica y $N_{[\cdot]}(\varepsilon, \mathcal{F}, L^1(\mathbb{P})) < \infty$ o $\log N(\varepsilon, \mathcal{F}, L_1(\mathbb{P}_n)) = o_p(n)$ para cada $\varepsilon > 0$ fijo, entonces $\sup_{f \in \mathcal{F}} |\mathbb{E}_{\mathbb{P}_n}(f) - \mathbb{E}_{\mathbb{P}}(f)| \xrightarrow{c.s.} 0$.

Definición 3.2.7 Una clase \mathcal{F} de funciones medibles $f : \mathcal{X} \rightarrow \mathbb{R}$ sobre un espacio de probabilidad $(\mathcal{X}, \mathcal{A}, \mathbb{P})$ es una clase \mathbb{P} -medible si la aplicación $(x_1, \dots, x_n) \mapsto \sup_{f \in \mathcal{F}} \left| \sum_{i=1}^n e_i f(x_i) \right|$ es medible en la completación $(\mathcal{X}^n, \mathcal{A}^n, \mathbb{P}^n)$ para todo n y para todo vector $(e_1, \dots, e_n) \in \mathbb{R}^n$.

Proposición 3.2.1 Si \mathcal{F} es una clase de funciones medibles que contiene un subconjunto G contable tal que para toda $f \in \mathcal{F}$ existe una sucesión $\{g_m\} \subset G$ tal que $g_m(x) \rightarrow_{m \rightarrow \infty} f(x)$ para todo x , entonces \mathcal{F} es \mathbb{P} -medible para toda medida de probabilidad \mathbb{P} .

La noción de entropía fue introducida por Kolmogorov y mide la complejidad de un conjunto, en tanto poseer una alta entropía significa que es necesaria mucha información para describir un elemento con una exactitud ε .

Definición 3.2.8 Sea \mathcal{F} una familia de funciones con envolvente F . Se llama integral de bracketing a $J_{[\cdot]}(\delta, \mathcal{F}) = \int_0^\delta \sqrt{1 + \log(N_{[\cdot]}(\varepsilon\|F\|_{2,\mathbb{P}}, \mathcal{F}, L_2(\mathbb{P})))} d\varepsilon$.

Proposición 3.2.2 Si \mathcal{F} , con envolvente F , satisface la condición de entropía de bracketing, es decir $\int_0^\infty \sqrt{\log(N_{[\cdot]}(\varepsilon\|F\|_{2,\mathbb{P}}, \mathcal{F}, L_2(\mathbb{P})))} d\varepsilon < \infty$, entonces

$$J_{[\cdot]}(\varepsilon, \mathcal{F}) \rightarrow_{\varepsilon \rightarrow 0} 0.$$

Proposición 3.2.3 Desigualdad Maximal para Números de Bracketing: Sean $\mathbf{x}_1, \dots, \mathbf{x}_n$ vectores aleatorios i.i.d. con distribución común \mathbb{P} . Sea \mathcal{F} una clase de funciones \mathbb{P} -medible con envolvente F tal que $\|F\|_{2,\mathbb{P}} < \infty$ y dado $\varepsilon > 0$, sea

$$a(\varepsilon) = \frac{\varepsilon\|F\|_{2,\mathbb{P}}}{\sqrt{1 + \log(N_{[\cdot]}(\varepsilon\|F\|_{2,\mathbb{P}}, \mathcal{F}, L_2(\mathbb{P})))}}.$$

Si $\|f\|_{2,\mathbb{P}} < \varepsilon\|F\|_{2,\mathbb{P}}$ para toda $f \in \mathcal{F}$, entonces existe una constante D , que no depende de n , tal que

$$\begin{aligned} \left\| \sup_{f \in \mathcal{F}} |T_n f| \right\|_{1,\mathbb{P}} &\leq DJ_{[\cdot]}(\varepsilon, \mathcal{F})\|F\|_{2,\mathbb{P}} + \sqrt{n}\mathbb{E}_{\mathbb{P}}\left(F I_{\{F > \sqrt{na}(\varepsilon)\}}\right) \\ &\leq DJ_{[\cdot]}(1, \mathcal{F})\|F\|_{2,\mathbb{P}}. \end{aligned}$$

Teorema 3.2.3 Sea $\mathcal{F} = \{f_t : t \in T\}$ una clase de funciones tal que existe F fija tal que $|f_s(\mathbf{x}) - f_t(\mathbf{x})| \leq d(s, t)F(\mathbf{x})$, es decir que es Lipschitz en el parámetro, entonces para cualquier $\|\cdot\|$,

$$N_{[\cdot]}(2\varepsilon\|F\|, \mathcal{F}, \|\cdot\|) \leq N(\varepsilon, T, \|\cdot\|).$$

Por otra parte, Nolan y Pollard (1987), introducen una noción de clase de funciones que facilita el trabajo, esta es la llamada *Clase Euclídea*.

Definición 3.2.9 Diremos que una clase de funciones \mathcal{F} es Euclídea para la función envolvente F si existen constantes A y V tales que

$$N(\varepsilon\|F\|_{1,\mathbb{P}}, \mathcal{F}, L_1(\mathbb{P})) \leq A\varepsilon^{-V}, \quad 0 < \varepsilon \leq 1,$$

donde $0 < \|F\|_{1,\mathbb{P}} < \infty$.

Luego, a partir de la definición de número de cubrimiento, se tiene que si \mathcal{F} es Euclídea, y para cada $r > 1$, $0 < \|F\|_{r,\mathbb{P}} < \infty$, entonces

$$N(\varepsilon\|F\|_{r,\mathbb{P}}, \mathcal{F}, L_r(\mathbb{P})) \leq A2^{rV}\varepsilon^{-rV}, \quad 0 < \varepsilon \leq 1.$$

Asimismo, dedujeron una serie de propiedades de permanencia para estas clases enunciadas en los siguientes resultados, que pueden hallarse en Pakes y Pollard (1989).

Proposición 3.2.4 Si \mathcal{F} es Euclídea para la función envolvente F y \mathcal{G} es Euclídea para G , entonces

- a) la clase $\{f + g : f \in \mathcal{F}, g \in \mathcal{G}\}$ es Euclídea para la envolvente $F + G$.
- b) la clase $\{f \cdot g : f \in \mathcal{F}, g \in \mathcal{G}\}$ es Euclídea para la envolvente $F \cdot G$.
- c) para cada M , la clase $\{\alpha f : f \in \mathcal{F}, \alpha \in \mathbb{R}, |\alpha| \leq M\}$ es euclídea con envolvente MF .

Proposición 3.2.5 Si T es medible y \mathcal{F} es Euclídea para la función envolvente F , la clase de funciones $\{T^\circ f : f \in \mathcal{F}\}$ es Euclídea con envolvente $F^\circ T$.

Proposición 3.2.6 i) Sea $\rho : \mathbb{R}_+ \rightarrow \mathbb{R}$, una función de variación acotada, entonces la clase de funciones $\mathcal{F} = \{f_{\mathbf{A}, \mathbf{b}}(\mathbf{x}) = \rho(|\mathbf{A}\mathbf{x} + \mathbf{b}|) : \mathbf{A} \text{ es una matriz de } m \times k, \mathbf{b} \in \mathbb{R}^m\}$ es Euclídea para la función envolvente $F = \|\rho\|_\infty$.

ii) Sea $\rho : \mathbb{R} \rightarrow \mathbb{R}$, una función de variación acotada, entonces la clase de funciones $\mathcal{F} = \{f_{\mathbf{a}, b}(\mathbf{x}) = \rho(|\mathbf{a}^t \mathbf{x} + b|) : \mathbf{a} \in \mathbb{R}^k, b \in \mathbb{R}\}$ es Euclídea para la función envolvente $F = \|\rho\|_\infty$.

Asimismo, en Wellner (2012) puede encontrarse la siguiente proposición sobre la preservación de la entropía con bracketing.

Proposición 3.2.7 Supongamos que para cada $t \in T$ un conjunto compacto de \mathbb{R}^p , la clase de funciones $\mathcal{F}_t = \{f_{t,u} : u \in U\}$ cumple que

$$\log(N_{[\cdot]}(\varepsilon, \mathcal{F}_t, L_2(\mathbb{P}))) \leq K \left(\frac{1}{\varepsilon}\right)^W,$$

para $W > 2$ y K que no dependa de t . Si suponemos además que

$$|f_{t_1, u} - f_{t_2, u}| \leq F|t_1 - t_2|,$$

con F una función tal que $\|F\|_{2, \mathbb{P}} < \infty$, entonces la clase $\mathcal{F} = \cup_{t \in T} \mathcal{F}_t$ cumple que

$$\log(N_{[\cdot]}(\varepsilon, \mathcal{F}, L_2(\mathbb{P}))) \lesssim p \log\left(\frac{1}{\varepsilon}\right) + K \left(\frac{1}{\varepsilon}\right)^W.$$

En las próximas dos secciones supondremos que w y ρ_1 son tales que $H(\boldsymbol{\theta}, \sigma, \boldsymbol{\Gamma}) = \min_{\mathbf{b} \in \Theta} H(\mathbf{b}, \sigma, \boldsymbol{\Gamma})$ y que el estimador propuesto es Fisher-consistente.

Por otra parte, para resumir notación, para cualquier vector o matriz \mathbf{D} , indicaremos con $|\mathbf{D}|$ a la norma euclídea de \mathbf{D} y definimos $\mathcal{B} = \{\mathbf{b} \in \Theta : |\mathbf{b} - \boldsymbol{\theta}| \leq \varepsilon_o\}$, $\mathcal{S} = \{s \in \mathbb{R}^+ : |s - \sigma| \leq \varepsilon_o\}$, $\mathcal{G} = \{\mathbf{G} \in \Lambda : |\mathbf{G} - \boldsymbol{\Gamma}| \leq \varepsilon_o\}$ y $\mathcal{X}_K = \{\mathbf{x} \in \mathbb{R}^k : |\mathbf{x}| \leq K\}$.

3.3. Consistencia Fuerte

Consideremos el siguiente conjunto de hipótesis para demostrar la consistencia fuerte del WMM -estimador.

A2. La función de peso w es tal que $\mathbb{E} \left[\sup_{\mathbf{G} \in \mathcal{G}} p(\mathbf{x})w(\mathbf{x}, \mathbf{G}) \right] < \infty$.

A3. La función ρ_1 es continua y la función $H(\mathbf{b}, s, \mathbf{G})$ cumple la siguiente condición de equicontinuidad: para todo $\varepsilon > 0$, existe $\tau > 0$ tal que para todo $s_1, s_2 \in \mathcal{K}$, un conjunto compacto en \mathcal{S} , y para todo $\mathbf{G}_1, \mathbf{G}_2 \in \mathcal{L}$, un conjunto compacto en \mathcal{G} ,

$$|s_1 - s_2| < \tau \text{ y } |\mathbf{G}_1 - \mathbf{G}_2| < \tau \Rightarrow \sup_{\mathbf{b} \in \Theta} |H(\mathbf{b}, s_1, \mathbf{G}_1) - H(\mathbf{b}, s_2, \mathbf{G}_2)| < \varepsilon.$$

A4. La clase de funciones $\mathcal{F} = \{g_{\mathbf{b},s,\mathbf{G}}(y, \mathbf{x}, \delta) = \delta \rho_1 \left(\frac{y-f(\mathbf{x},\mathbf{b})}{s} \right) w(\mathbf{x}, \mathbf{G}), s \in \mathcal{S}, \mathbf{b} \in \Theta, \mathbf{G} \in \mathcal{G}\}$ tiene número de bracketing finito, es decir, $N_{[\cdot]}(\varepsilon, \mathcal{F}, L^1(\mathbb{P})) < \infty$, para todo $0 < \varepsilon < 1$, donde \mathbb{P} es la distribución de (y, \mathbf{x}^\top) o $\log N(\varepsilon, \mathcal{F}, L^1(\mathbb{P}_n)) = o_P(n)$ con \mathbb{P}_n la distribución empírica correspondiente.

Observación 3.3.1 *El primer supuesto, A2, se verifica inmediatamente dado que la función de peso w habitualmente se elige acotada.*

La hipótesis A3 se puede verificar con una adecuada elección de las funciones de pérdida ρ y de peso w , como por ejemplo que la función $\eta(t) = \psi(t)t$, con $\psi = \rho'$, sea acotada y que w sea una función de Lipschitz de soporte compacto.

El supuesto A4 puede verificarse en el contexto paramétrico en distintas situaciones. Cabe observar que si tenemos dos familias de funciones \mathcal{F}_1 y \mathcal{F}_2 con número de bracketing finito, $N_{\mathcal{F}_1} = N_{[\cdot]}(\varepsilon, \mathcal{F}_1, L_1(\mathbb{P})) < \infty$ y $N_{\mathcal{F}_2} = N_{[\cdot]}(\varepsilon, \mathcal{F}_2, L_1(\mathbb{P})) < \infty$, y consideramos la familia de funciones $\mathcal{F} = \{f(\mathbf{z}) = f_1(\mathbf{z})f_2(\mathbf{z}) : f_1 \in \mathcal{F}_1 \text{ y } f_2 \in \mathcal{F}_2\}$, entonces $N_{[\cdot]}(\varepsilon, \mathcal{F}, L_1(\mathbb{P})) < \infty$, con lo cual bastará con pedir que cada uno de los dos factores de $g_{\mathbf{b},s,\mathbf{G}}(y, \mathbf{x}, \delta)$ pertenezcan a familias de funciones de bracketing finito.

Por ejemplo, cuando $w(\mathbf{x}, \mathbf{\Gamma}) = w_o(d(\mathbf{x}, \mathbf{\Gamma}))$ siendo $d(\mathbf{x}, \mathbf{\Gamma})$ la distancia de Mahalanobis de \mathbf{x} y $\mathbf{\Gamma}$ representa posición y escala como en (2.11), si w_o es una función continua y acotada la clase $\mathcal{F}^ = \{g_{\mathbf{G}}(\mathbf{x}) = w(\mathbf{x}, \mathbf{G}), \mathbf{G} \in \mathcal{G}\}$ tiene número de bracketing finito (ver Lema 3.10 de van de Geer, 2000).*

Además, si la función de pérdida ρ es de variación acotada o Lipschitz, dado que s varía en un intervalo real acotado, en condiciones bastante generales, el número de bracketing de $\mathcal{F}^+ = \{g_{\mathbf{b},s}(y, \mathbf{x}, \delta) = \delta \rho_1 \left(\frac{y-f(\mathbf{x},\mathbf{b})}{s} \right), s \in \mathcal{S}, \mathbf{b} \in \Theta\}$ esencialmente dependerá de la familia de funciones de regresión f y de características del espacio paramétrico Θ .

Las propiedades de permanencia para el número de bracketing y las clases euclídeas enunciadas en la Sección 3.2 permiten verificar que el número de bracketing de la familia \mathcal{F}^+ es finito en una amplia variedad de funciones de regresión f . Como por ejemplo,

si combinamos las Proposiciones 3.2.4 y 3.2.6, podemos tomar g_1 y g_2 dentro de las funciones de densidad, cuya variación total es 1, o que sean estrictamente monótonas o Lipschitz, resultando entonces de variación acotada y dados $\boldsymbol{\theta}^t = (\theta_0, \theta_1, \dots, \theta_p) = (\theta_0, \boldsymbol{\theta}_p^t) \in \mathbb{R}^{p+1}$ y $\mathbf{x} \in \mathbb{R}^p$, podemos considerar $f(\mathbf{x}, \boldsymbol{\theta}) = g_1(\boldsymbol{\theta}_p^t \mathbf{x} + \theta_0) + g_2(\boldsymbol{\theta}_p^t \mathbf{x} + \theta_0)$ o bien $f(\mathbf{x}, \boldsymbol{\theta}) = g_1(\boldsymbol{\theta}_p^t \mathbf{x} + \theta_0)g_2(\boldsymbol{\theta}_p^t \mathbf{x} + \theta_0)$.

Asimismo, para el caso de funciones de pérdida ρ acotadas y Lipschitz, Anthony y Bartlett (1999) obtienen cotas para el número de cubrimiento de la familia \mathcal{F}^+ en base al número de cubrimientos de la familia de funciones definida por la función de regresión f , es decir $\{f(\mathbf{x}, \mathbf{b}), \mathbf{b} \in \Theta\}$.

El siguiente teorema enuncia la consistencia fuerte de los estimadores propuestos que resuelven (2.12).

Teorema 3.3.2 *Supongamos que valen A2-A4 y que se satisface la condición MAR dada en (2.2). Si además, $\widehat{\sigma} \xrightarrow{c.s.} \sigma$ y $\widehat{\Gamma} \xrightarrow{c.s.} \Gamma$, entonces $\widehat{\boldsymbol{\theta}} \xrightarrow{c.s.} \boldsymbol{\theta}$.*

Demostración: En primer lugar probaremos que $H_n(\widehat{\boldsymbol{\theta}}, \widehat{\sigma}, \widehat{\Gamma}) - H_n(\widehat{\boldsymbol{\theta}}, \sigma, \Gamma) \xrightarrow{c.s.} 0$. Notemos que $\mathbb{E}(H_n(\mathbf{b}, s, \mathbf{G})) = H(\mathbf{b}, s, \mathbf{G})$ definida en (3.3). Además, tenemos que

$$\begin{aligned} \sup_{\mathbf{b} \in \Theta} \left| H_n(\mathbf{b}, \widehat{\sigma}, \widehat{\Gamma}) - H_n(\mathbf{b}, \sigma, \Gamma) \right| &\leq \sup_{\mathbf{b} \in \Theta} \left| H_n(\mathbf{b}, \widehat{\sigma}, \widehat{\Gamma}) - H(\mathbf{b}, \widehat{\sigma}, \widehat{\Gamma}) \right| + \sup_{\mathbf{b} \in \Theta} |H_n(\mathbf{b}, \sigma, \Gamma) - H(\mathbf{b}, \sigma, \Gamma)| \\ &\quad + \sup_{\mathbf{b} \in \Theta} \left| H(\mathbf{b}, \widehat{\sigma}, \widehat{\Gamma}) - H(\mathbf{b}, \sigma, \Gamma) \right| \\ &\leq 2W_n + \sup_{\mathbf{b} \in \Theta} \left| H(\mathbf{b}, \widehat{\sigma}, \widehat{\Gamma}) - H(\mathbf{b}, \sigma, \Gamma) \right|, \end{aligned}$$

donde

$$W_n = \sup_{\mathbf{b} \in \Theta, s \in \mathcal{S}, \mathbf{G} \in \mathcal{G}} \left| \frac{1}{n} \sum_{i=1}^n \delta_i \rho \left(\frac{y_i - f(\mathbf{x}_i, \mathbf{b})}{s} \right) w(\mathbf{x}_i, \mathbf{G}) - \mathbb{E} \left(\delta \rho \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{s} \right) w(\mathbf{x}, \mathbf{G}) \right) \right|.$$

Como valen A2 y A4, se puede aplicar el Teorema 3.2.2 resultando que $W_n \xrightarrow{c.s.} 0$.

De esta manera, como, por A3, $H(\mathbf{b}, \sigma, \Gamma)$ es equicontinua y $\widehat{\sigma}$ y $\widehat{\Gamma}$ son consistentes, $\sup_{\mathbf{b} \in \Theta} \left| H(\mathbf{b}, \widehat{\sigma}, \widehat{\Gamma}) - H(\mathbf{b}, \sigma, \Gamma) \right| \xrightarrow{c.s.} 0$.

Finalmente, $H_n(\widehat{\boldsymbol{\theta}}, \widehat{\sigma}, \widehat{\Gamma}) - H_n(\widehat{\boldsymbol{\theta}}, \sigma, \Gamma) \xrightarrow{c.s.} 0$, entonces la sucesión de estimadores $\widehat{\boldsymbol{\theta}}$ cumple que $\inf_{\mathbf{b} \in \Theta} H_n(\mathbf{b}, \widehat{\sigma}, \widehat{\Gamma}) - H_n(\widehat{\boldsymbol{\theta}}, \sigma, \Gamma) \xrightarrow{c.s.} 0$. Por lo tanto, se pueden aplicar los resultados de Huber (1967) obteniendo la conclusión deseada. \square

3.4. Distribución Asintótica

En esta sección probaremos, bajo el supuesto de que Θ y Λ son espacios separables, que $\widehat{\boldsymbol{\theta}}$, el estimador pesado propuesto para estimar el parámetro de regresión en un modelo

no lineal con respuestas faltantes MAR que resulta de (2.12), tiene distribución asintóticamente normal.

Consideremos el siguiente conjunto de hipótesis.

- N1.** $\psi_1 = \rho'_1$ es una función impar, acotada y continua.
- N2.** ψ'_1 es una función continua y acotada.
- N3.** La función $\eta_1(t) = \psi'_1(t)t$ es acotada.
- N4.** La función de peso $w(\mathbf{x}, \mathbf{G})$ es acotada y continua como función de \mathbf{G} .
- N5.** Existen funciones $\alpha(\mathbf{x})$ con $\mathbb{E}(\alpha^2(\mathbf{x})) < \infty$ y $\tilde{\alpha}(\mathbf{x}, \mathbf{G})$ con $\mathbb{E} \left[(\sup_{\mathcal{G}} |\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})| \tilde{\alpha}(\mathbf{x}, \mathbf{G}))^2 \right] < \infty$, tales que $|\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) w(\mathbf{x}, \boldsymbol{\Gamma})| < \alpha(\mathbf{x})$ y $|(w(\mathbf{x}, \mathbf{G}) - w(\mathbf{x}, \boldsymbol{\Gamma}))| < \tilde{\alpha}(\mathbf{x}, \mathbf{G})|\mathbf{G} - \boldsymbol{\Gamma}|$ si $\mathbf{G} \in \mathcal{G}$.
- N6.** La función de regresión $f(\mathbf{x}, \mathbf{b})$ es dos veces derivable como función de \mathbf{b} y existe ε_o tal que

$$\mathbb{E} \left[\sup_{\mathcal{B}, \mathcal{G}} |\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b}) \dot{\mathbf{f}}(\mathbf{x}, \mathbf{b})^t w(\mathbf{x}, \mathbf{G})| \right] < \infty$$

$$\mathbb{E} \left[\sup_{\mathcal{B}, \mathcal{G}} |\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G})| \right] < \infty,$$

donde $\ddot{\mathbf{f}}$ es la matriz hessiana de $f(\mathbf{x}, \mathbf{b})$ con respecto a \mathbf{b} .

- N7.** La función de regresión $f(\mathbf{x}, \mathbf{b})$ es equicontinua en el sentido en que dado $\varepsilon > 0$, existe $\tau > 0$ tal que si $|\mathbf{b}_1 - \mathbf{b}_2| \leq \tau$, $|\mathbf{b}_i - \boldsymbol{\theta}| \leq \tau$, $i = 1, 2$ y $\mathbf{x} \in \mathcal{X}_K$, entonces $|f(\mathbf{x}, \mathbf{b}_1) - f(\mathbf{x}, \mathbf{b}_2)| < \varepsilon$, y además $\sup_{\mathcal{X}_K} |f(\mathbf{x}, \boldsymbol{\theta})| < \infty$, para todo $K > 0$.
- N8.** La función $\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b}) \sqrt{w(\mathbf{x}, \mathbf{G})}$ es equicontinua en el sentido en que dado $\varepsilon > 0$, existe $\tau > 0$ tal que si $|\mathbf{b}_1 - \mathbf{b}_2| \leq \tau$, $|\mathbf{G}_1 - \mathbf{G}_2| < \tau$, $|\mathbf{b}_i - \boldsymbol{\theta}| \leq \tau$, $|\mathbf{G}_i - \boldsymbol{\Gamma}| < \tau$, $i = 1, 2$ y $\mathbf{x} \in \mathcal{X}_K$, entonces $|\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b}_1) \sqrt{w(\mathbf{x}, \mathbf{G}_1)} - \dot{\mathbf{f}}(\mathbf{x}, \mathbf{b}_2) \sqrt{w(\mathbf{x}, \mathbf{G}_2)}| < \varepsilon$, y además $\sup_{\mathcal{X}_K} |\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) \sqrt{w(\mathbf{x}, \boldsymbol{\Gamma})}| < \infty$, para todo $K > 0$.
- N9.** La función $\ddot{\mathbf{f}}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G})$ es equicontinua en el sentido en que dado $\varepsilon > 0$, existe $\tau > 0$ tal que si $|\mathbf{b}_1 - \mathbf{b}_2| \leq \tau$, $|\mathbf{G}_1 - \mathbf{G}_2| < \tau$, $|\mathbf{b}_i - \boldsymbol{\theta}| \leq \tau$, $|\mathbf{G}_i - \boldsymbol{\Gamma}| < \tau$, $i = 1, 2$ y $\mathbf{x} \in \mathcal{X}_K$, entonces $|\ddot{\mathbf{f}}(\mathbf{x}, \mathbf{b}_1) w(\mathbf{x}, \mathbf{G}_1) - \ddot{\mathbf{f}}(\mathbf{x}, \mathbf{b}_2) w(\mathbf{x}, \mathbf{G}_2)| < \varepsilon$, y además $\sup_{\mathcal{X}_K} |\ddot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) w(\mathbf{x}, \boldsymbol{\Gamma})| < \infty$, para todo $K > 0$.

N10. La matriz

$$\mathbf{A} = \mathbb{E} \psi'_1(\varepsilon) \mathbb{E} \left[p(\mathbf{x}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t w(\mathbf{x}, \boldsymbol{\Gamma}) \right]$$

es no singular.

Observación 3.4.1 *Los supuestos N1–N4 son condiciones estándar sobre la función de pérdida y la función de peso en modelos de regresión. El conjunto de hipótesis N5 a N9 es más débil que pedir que para algún ε_o $\mathbb{E} \left[\sup_{\mathcal{B}} |\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b})|^2 \right] < \infty$ y $\mathbb{E} \left[\sup_{\mathcal{B}} |\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})| \right] < \infty$, que es la condición requerida para derivar la distribución asintótica para MM–estimadores (ver Fasano et al., 2012). La condición N5 establece que el peso a elegir debe controlar el crecimiento del gradiente y se requiere cierta suavidad del mismo respecto de \mathbf{G} . Por ejemplo, tal como se describió en la Sección 2.3.2, si el peso se elige de la forma $w(\mathbf{x}, \mathbf{b}) = w_o(|\dot{\mathbf{f}}(\mathbf{x}_i, \mathbf{b})|)$, donde $w_o(t) = \psi_o(t)/t$, siendo $\psi_o = \rho'_o$ y ρ_o una ρ –función, entonces $|\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b}) w_o(\mathbf{x}, \mathbf{b})| \leq |\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b})| \frac{\psi_o(|\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b})|)}{|\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b})|} \leq \|\psi_o\|_\infty$, por lo tanto $\alpha \equiv \|\psi_o\|_\infty$. Cuando w_o tiene derivada w'_o , entonces $|w_o(|\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b}_1)|) - w_o(|\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b}_2)|)| \leq |w'_o(|\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b})|)| |\dot{\mathbf{f}}(\mathbf{x}, \mathbf{b})| |\mathbf{b}_1 - \mathbf{b}_2|$ siendo \mathbf{b} un punto intermedio y obtenemos $\tilde{\alpha}$, si w'_o controla el crecimiento de $\dot{\mathbf{f}}$ y $\ddot{\mathbf{f}}$ en el entorno \mathcal{B} .*

N6 establece que el peso debe ser elegido de manera que el producto del gradiente de la función de regresión por el peso y su derivada por el peso, sean integrables, debilitando el supuesto de existencia de momentos de $\dot{\mathbf{f}}$ y $\ddot{\mathbf{f}}$ y por lo tanto, amplía la variedad de funciones de regresión f a tratar. N7 se cumple, por ejemplo si $f(\mathbf{x}, \mathbf{b})$ es continua como función de las covariables y N8 y N9 se verifican si $\dot{\mathbf{f}}$ y $\ddot{\mathbf{f}}$ son continuas como función de las covariables \mathbf{x} y con una elección adecuada del peso w . Finalmente, N10 nos dice que la matriz $\mathbb{E} \left[p(\mathbf{x}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t w(\mathbf{x}, \boldsymbol{\Gamma}) \right]$ es no singular y este supuesto está relacionado con la no colinealidad de $\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})$ sobre $\mathcal{S}_w \cap \mathcal{S}_p \cap \mathcal{S}_x$.

En primer lugar, enunciaremos un lema auxiliar general sobre la convergencia en probabilidad para ciertos procesos empíricos.

Consideremos las funciones $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ y $\mathbf{M}(\mathbf{x}, \mathbf{b}) : \mathbb{R}^k \times \Theta \rightarrow \mathbb{R}^{p \times p}$ tales que cumplan las siguientes hipótesis, las cuales son versiones generales de algunos supuestos arriba descriptos:

L1. φ es una función continua y acotada.

L2. Existe ε_o tal que

$$\mathbb{E} \left[\sup_{\mathcal{B}, \mathcal{G}} |\mathbf{M}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G})| \right] < \infty.$$

L3. La función $\mathbf{M}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G}) : \mathbb{R}^k \times \Theta \times \Lambda \rightarrow \mathbb{R}^{p \times p}$ es equicontinua en el sentido en que dado $\varepsilon > 0$, existe $\tau > 0$ tal que si $|\mathbf{b}_1 - \mathbf{b}_2| \leq \tau$, $|\mathbf{G}_1 - \mathbf{G}_2| < \tau$, $|\mathbf{b}_i - \boldsymbol{\theta}| \leq \tau$, $|\mathbf{G}_i - \boldsymbol{\Gamma}| < \tau$, $i = 1, 2$, entonces $|\mathbf{M}(\mathbf{x}, \mathbf{b}_1) w(\mathbf{x}, \mathbf{G}_1) - \mathbf{M}(\mathbf{x}, \mathbf{b}_2) w(\mathbf{x}, \mathbf{G}_2)| < \varepsilon$ si $\mathbf{x} \in \mathcal{X}_K$ y $\sup_{\mathcal{X}_K} |\mathbf{M}(\mathbf{x}, \boldsymbol{\theta}) w(\mathbf{x}, \boldsymbol{\Gamma})| < \infty$ para todo $K > 0$.

Lema 3.4.1 *Sean $(y_i, \mathbf{x}_i^t, \delta_i) \sim G$, $1 \leq i \leq n$, vectores aleatorios independientes que cumplan el modelo (2.1) y la condición (2.2). Supongamos que valen N4, N7 y L1-L3 y que los estimadores $\widehat{\sigma}$, $\widehat{\boldsymbol{\Gamma}}$ y $\widetilde{\boldsymbol{\theta}}$ son tales que $\widehat{\sigma} \xrightarrow{p} \sigma$, $\widehat{\boldsymbol{\Gamma}} \xrightarrow{p} \boldsymbol{\Gamma}$ y $\widetilde{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}$. Entonces, $\mathbf{D}_n \xrightarrow{p} \mathbf{D}$,*

donde

$$\mathbf{D}_n = \frac{1}{n} \sum_{i=1}^n \delta_i \varphi \left(\frac{y_i - f(\mathbf{x}_i, \tilde{\boldsymbol{\theta}})}{\hat{\sigma}} \right) \mathbf{M}(\mathbf{x}_i, \tilde{\boldsymbol{\theta}}) w(\mathbf{x}_i, \hat{\boldsymbol{\Gamma}}) \quad y \quad \mathbf{D} = \mathbb{E} \varphi(\epsilon) \mathbb{E} \left[p(\mathbf{x}) \mathbf{M}(\mathbf{x}, \boldsymbol{\theta}) w(\mathbf{x}, \boldsymbol{\Gamma}) \right].$$

Demostración: Consideremos el vector aleatorio $(y, \mathbf{x}^t, \delta)$ con la misma distribución que $(y_i, \mathbf{x}_i^t, \delta_i)$.

Es suficiente probar que

- i) $\lim_{(\mathbf{b}, s, \mathbf{G}) \rightarrow (\boldsymbol{\theta}, \sigma, \boldsymbol{\Gamma})} \mathbb{E} \left[\delta \varphi \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{s} \right) \mathbf{M}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G}) \right] = \mathbf{D},$
- ii) $\sup_{\mathcal{B}, \mathcal{S}, \mathcal{G}} \left| \frac{1}{n} \sum_{i=1}^n \mathbb{I}(y_i, \mathbf{x}_i, \delta_i, \mathbf{b}, s, \mathbf{G}) - \mathbb{E} [\mathbb{I}(y, \mathbf{x}, \delta, \boldsymbol{\theta}, \sigma, \boldsymbol{\Gamma})] \right| \xrightarrow{P} \mathbf{0},$ con $\mathbb{I}(y, \mathbf{x}, \delta, \mathbf{b}, s, \mathbf{G}) = \delta \varphi \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{s} \right) \mathbf{M}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G}).$

El ítem i) se deduce de **L1–L3** y del Teorema de Convergencia Dominada ya que \mathbb{I} converge puntualmente por **N4** y **N7**, la función φ es continua y se tiene que $\|\mathbb{I}\| \leq \|\varphi\|_\infty \sup_{\mathcal{B}, \mathcal{G}} |\mathbf{M}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G})|$, que es integrable.

Para probar ii), como \mathbb{I} es una matriz, probaremos la convergencia de cada elemento.

Por el Teorema 3.2.1, basta con probar que dado $\eta > 0$ existe una clase finita de funciones \mathcal{I}_η tal que para todo $\mathbf{b} \in \mathcal{B}$, $s \in \mathcal{S}$ y $\mathbf{G} \in \mathcal{G}$ existe $\mathbb{I}_{\eta, U}^{kl} \in \mathcal{I}_\eta$ y $\mathbb{I}_{\eta, L}^{kl} \in \mathcal{I}_\eta$ tales que

$$\mathbb{I}_{\eta, L}^{kl}(y, \mathbf{x}) \leq \mathbb{I}^{kl}(y, \mathbf{x}, \delta, \mathbf{b}, s, \mathbf{G}) \leq \mathbb{I}_{\eta, U}^{kl}(y, \mathbf{x}) \quad (3.7)$$

$$\mathbb{E} \left(\mathbb{I}_{\eta, U}^{kl}(y, \mathbf{x}) - \mathbb{I}_{\eta, L}^{kl}(y, \mathbf{x}) \right) < \eta. \quad (3.8)$$

Dado $K > 0$, por **N7**, consideremos $C_1 = \sup_{\mathcal{X}_K, \mathcal{B}} |f(\mathbf{x}, \mathbf{b})|$.

Definamos $A_K = \left\{ \left| \frac{y - f(\mathbf{x}, \boldsymbol{\theta})}{\sigma} \right| \leq K, |\mathbf{x}| \leq K \right\}$ y tomemos $K \in \mathbb{N}$ de forma tal que dado $\eta_1 = \eta/5 \|\varphi\|_\infty$, $P(A_K) > 1 - \eta_1$, y, como vale **L2**, tal que

$$\mathbb{E} \left(\sup_{\mathcal{B}, \mathcal{G}} |\mathbf{M}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G})| I_{A_K^c}(y, \mathbf{x}) \right) < \eta_1. \quad (3.9)$$

Si $(y, \mathbf{x}^t) \in A_K$, entonces $|y| \leq K_1$, con $K_1 = K\sigma + C_1$. Luego, llamando $B_K = \{|\mathbf{x}| \leq K, |y| \leq K_1\}$, $A_K \subset B_K$. Por otro lado, sea c tal que para todo $\mathbf{b} \in \mathcal{B}$ y $s \in \mathcal{S}$, $|\mathbf{b}| < c$ y $c^{-1} < s < c$. Luego, si $(y, \mathbf{x}^t) \in B_K$, $\mathbf{b} \in \mathcal{B}$ y $s \in \mathcal{S}$, entonces $|y - f(\mathbf{x}, \mathbf{b})|/s \leq (K_1 + C_1)c = b_1$.

De **L3** existe $b_2 \geq 0$ tal que $|\mathbf{M}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G})| \leq b_2$ para cada $\mathbf{x} \in B_K$, $\mathbf{b} \in \mathcal{B}$ y $\mathbf{G} \in \mathcal{G}$. Por **L1** tenemos que $\varphi(t)\mathbf{z}$ es uniformemente continua en $C_K = \{|t| \leq b_1, |\mathbf{z}| \leq b_2\}$ y por lo tanto, existe $\tau > 0$ tal que $|\varphi(t)\mathbf{z} - \varphi(u)\mathbf{v}| < \eta/10$, para $|t - u| < \tau$, $|\mathbf{z} - \mathbf{v}| < \tau$, $(t, \mathbf{z}^t) \in C_K$ y $(u, \mathbf{v}^t) \in C_K$.

Como valen **N7** y **L3** resulta que si $(y, \mathbf{x}^t) \in B_K$, entonces $(y-f(\mathbf{x}, \mathbf{b}))/s$ y $\mathbf{M}(\mathbf{x}, \mathbf{b})w(\mathbf{x}, \mathbf{G})$ son funciones equicontinuas de $(\mathbf{b}, s, \mathbf{G}) \in \mathcal{B} \times \mathcal{S} \times \mathcal{G}$, es decir que existe $\varepsilon_1 > 0$ tal que si $|\mathbf{b} - \tilde{\mathbf{b}}| < \varepsilon_1$, $|s - \tilde{s}| < \varepsilon_1$, $|\mathbf{G} - \tilde{\mathbf{G}}| < \varepsilon_1$, $(\mathbf{b}, s, \mathbf{G}), (\tilde{\mathbf{b}}, \tilde{s}, \tilde{\mathbf{G}}) \in \mathcal{B} \times \mathcal{S} \times \mathcal{G}$ y $(y, \mathbf{x}^t) \in B_K$,

$$\left| \frac{y - f(\mathbf{x}, \mathbf{b})}{s} - \frac{y - f(\mathbf{x}, \tilde{\mathbf{b}})}{\tilde{s}} \right| < \tau$$

$$\left| \mathbf{M}(\mathbf{x}, \mathbf{b})w(\mathbf{x}, \mathbf{G}) - \mathbf{M}(\mathbf{x}, \tilde{\mathbf{b}})w(\mathbf{x}, \tilde{\mathbf{G}}) \right| < \tau.$$

Sean ahora $(\mathcal{B}_i)_{1 \leq i \leq N_1}$ una colección finita de bolas centradas en $\theta_i \in \mathcal{B}$ de radio menor que ε_1 tal que $\mathcal{B} = \bigcup_{i=1}^{N_1} \mathcal{B}_i$, $(\mathcal{S}_m)_{1 \leq m \leq N_2}$ otra colección finita de bolas centradas en $\sigma_m \in \mathcal{S}$ de radio menor que ε_1 tal que $\mathcal{S} = \bigcup_{m=1}^{N_2} \mathcal{S}_m$ y $(\mathcal{G}_r)_{1 \leq r \leq N_3}$ otra colección finita de bolas centradas en $\Gamma_r \in \mathcal{G}$ también de radio menor que ε_1 tal que $\mathcal{G} = \bigcup_{r=1}^{N_3} \mathcal{G}_r$. Por lo que dados $\mathbf{b} \in \mathcal{B}$, $s \in \mathcal{S}$ y $\mathbf{G} \in \mathcal{G}$ existe una tira de índices (i, m, r) tal que $\mathbf{b} \in \mathcal{B}_i$, $s \in \mathcal{S}_m$ y $\mathbf{G} \in \mathcal{G}_r$ y

$$\left| \mathbb{I}^{kl}(y, \mathbf{x}, \delta, \mathbf{b}, s, \mathbf{G}) - \mathbb{I}^{kl}(y, \mathbf{x}, \delta, \theta_i, \sigma_m, \Gamma_r) \right| \leq \frac{\eta}{10} + I_{A_K^c}(y, \mathbf{x}) \|\varphi\|_{\infty} 2 \sup_{\mathcal{B}, \mathcal{G}} |\mathbf{M}(\mathbf{x}, \mathbf{b})w(\mathbf{x}, \mathbf{G})|.$$

Así, si definimos $e(y, \mathbf{x}) = \left(\frac{\eta}{10} + 2\|\varphi\|_{\infty} I_{A_K^c}(y, \mathbf{x}) \sup_{\mathcal{B}, \mathcal{G}} |\mathbf{M}(\mathbf{x}, \mathbf{b})w(\mathbf{x}, \mathbf{G})| \right)$ y consideramos

$$\mathbb{I}_{\eta, L}^{kl}(y, \mathbf{x}) = \mathbb{I}^{kl}(y, \mathbf{x}, \delta, \theta_i, \sigma_m, \Gamma_r) - e(y, \mathbf{x}) \quad \text{y} \quad \mathbb{I}_{\eta, U}^{kl}(y, \mathbf{x}) = \mathbb{I}^{kl}(y, \mathbf{x}, \delta, \theta_i, \sigma_m, \Gamma_r) + e(y, \mathbf{x})$$

obtenemos (3.7). Por otro lado,

$$\mathbb{E} \left(\mathbb{I}_{\eta, U}^{kl}(y, \mathbf{x}) - \mathbb{I}_{\eta, L}^{kl}(y, \mathbf{x}) \right) = \frac{\eta}{5} + 4\|\varphi\|_{\infty} \mathbb{E} \left(I_{A_K^c}(y, \mathbf{x}) \sup_{\mathcal{B}, \mathcal{G}} |\mathbf{M}(\mathbf{x}, \mathbf{b})w(\mathbf{x}, \mathbf{G})| \right),$$

entonces, usando (3.9), probamos (3.8) y podemos concluir que $\mathbf{D}_n \xrightarrow{p} \mathbf{D}$. \square

Como consecuencia directa del Lema 3.4.1 se puede probar la convergencia de los procesos que involucran matrices referentes a nuestro problema y que nos servirá para deducir la normalidad asintótica deseada.

Lema 3.4.2 *Supongamos que $(y_i, \mathbf{x}_i^t, \delta_i)$, $1 \leq i \leq n$ cumplen el modelo (2.1) y la condición MAR dada en (2.2). Supongamos que valen **N1**, **N2**, **N4**, **N6** y **N7-N10** y que $\widehat{\sigma} \xrightarrow{p} \sigma$, $\widehat{\Gamma} \xrightarrow{p} \Gamma$ y $\widehat{\theta} \xrightarrow{p} \theta$. Entonces,*

$$a) \mathbf{A}_n^1 = \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1' \left(\frac{y_i - f(\mathbf{x}_i, \widehat{\theta})}{\widehat{\sigma}} \right) \mathbf{f}(\mathbf{x}_i, \widehat{\theta}) \mathbf{f}(\mathbf{x}_i, \widehat{\theta})^t w(\mathbf{x}_i, \widehat{\Gamma}) \xrightarrow{p} \mathbf{A}, \text{ donde } \mathbf{A} \text{ está dada en } \mathbf{N10}.$$

$$b) \mathbf{A}_n^2 = \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{y_i - f(\mathbf{x}_i, \tilde{\boldsymbol{\theta}})}{\hat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \tilde{\boldsymbol{\theta}}) w(\mathbf{x}_i, \hat{\boldsymbol{\Gamma}}) \xrightarrow{p} \mathbf{0}.$$

Ahora sí ya estamos en condiciones de probar la normalidad asintótica de los estimadores propuestos, cuya demostración está realizada utilizando argumentos similares a los usados por Bianco y Boente (2002) y (2004).

Teorema 3.4.2 *Supongamos que valen N1–N10 y que $\hat{\sigma} \xrightarrow{p} \sigma$ y $\hat{\boldsymbol{\Gamma}} \xrightarrow{p} \boldsymbol{\Gamma}$. Entonces, si $\hat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}$, tenemos que*

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} N(\mathbf{0}, \sigma^2 \mathbf{A}^{-1} \mathbf{B} \mathbf{A}^{-1}).$$

donde la matriz simétrica \mathbf{A} está definida en N10 y $\mathbf{B} = \mathbb{E} \psi_1^2(\epsilon) \mathbb{E} [p(\mathbf{x}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t w^2(\mathbf{x}, \boldsymbol{\Gamma})]$.

Demostración: Como $\hat{\boldsymbol{\theta}}$ está definido por (2.14), entonces tenemos que $H_n^{(1)}(\hat{\boldsymbol{\theta}}, \hat{\sigma}, \hat{\boldsymbol{\Gamma}}) = \mathbf{0}$, donde $H_n^{(1)}$ está definida en (3.2). Si aplicamos un desarrollo de Taylor de primer orden alrededor de $\boldsymbol{\theta}$ a esta ecuación, tenemos que

$$\begin{aligned} \mathbf{0} &= \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{y_i - f(\mathbf{x}_i, \boldsymbol{\theta})}{\hat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \boldsymbol{\theta}) w(\mathbf{x}_i, \hat{\boldsymbol{\Gamma}}) \\ &\quad - \frac{1}{\hat{\sigma}} \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1' \left(\frac{y_i - f(\mathbf{x}_i, \tilde{\boldsymbol{\theta}})}{\hat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \tilde{\boldsymbol{\theta}}) \dot{\mathbf{f}}(\mathbf{x}_i, \tilde{\boldsymbol{\theta}})^t w(\mathbf{x}_i, \hat{\boldsymbol{\Gamma}}) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \\ &\quad + \frac{1}{\hat{\sigma}} \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{y_i - f(\mathbf{x}_i, \tilde{\boldsymbol{\theta}})}{\hat{\sigma}} \right) \ddot{\mathbf{f}}(\mathbf{x}_i, \tilde{\boldsymbol{\theta}}) w(\mathbf{x}_i, \hat{\boldsymbol{\Gamma}}) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \end{aligned}$$

siendo $\tilde{\boldsymbol{\theta}}$ un punto intermedio entre $\hat{\boldsymbol{\theta}}$ y $\boldsymbol{\theta}$. Si denotamos $\mathbf{A}_n = \mathbf{A}_n^1 - \hat{\sigma} \mathbf{A}_n^2$ con \mathbf{A}_n^1 y \mathbf{A}_n^2 definidas en el Lema 3.4.2, resulta

$$\mathbf{0} = \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{y_i - f(\mathbf{x}_i, \boldsymbol{\theta})}{\hat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \boldsymbol{\theta}) w(\mathbf{x}_i, \hat{\boldsymbol{\Gamma}}) - \frac{1}{\hat{\sigma}} \mathbf{A}_n (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$$

y por lo tanto,

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = \hat{\sigma} \mathbf{A}_n^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{y_i - f(\mathbf{x}_i, \boldsymbol{\theta})}{\hat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \boldsymbol{\theta}) w(\mathbf{x}_i, \hat{\boldsymbol{\Gamma}}).$$

Del Lema 3.4.2 tenemos que $\mathbf{A}_n \xrightarrow{p} \mathbf{A}$, por lo tanto, como $\hat{\sigma} \xrightarrow{p} \sigma$, es suficiente probar que

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{y_i - f(\mathbf{x}_i, \boldsymbol{\theta})}{\hat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \boldsymbol{\theta}) w(\mathbf{x}_i, \hat{\boldsymbol{\Gamma}}) \xrightarrow{D} N(\mathbf{0}, \mathbf{B}). \quad (3.10)$$

Definamos $L(y, \mathbf{x}, \delta, s, \mathbf{G}) = \delta \psi_1 \left(\frac{y - f(\mathbf{x}, \boldsymbol{\theta})}{s} \right) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) w(\mathbf{x}, \mathbf{G})$, donde $\boldsymbol{\theta}$ es el valor verdadero del parámetro y está fijo. Notemos que $L(y_i, \mathbf{x}_i, \delta_i, \sigma, \boldsymbol{\Gamma})$ tiene esperanza nula en virtud de la

simetría de los errores, imparidad de ψ_1 y la condición (2.2). Por el Teorema Central del Límite sabemos que

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n L(y_i, \mathbf{x}_i, \delta_i, \sigma, \mathbf{\Gamma}) \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{B}),$$

ergo, basta ver que

$$T_n(\widehat{\sigma}, \widehat{\mathbf{\Gamma}}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n [L(y_i, \mathbf{x}_i, \delta_i, \widehat{\sigma}, \widehat{\mathbf{\Gamma}}) - L(y_i, \mathbf{x}_i, \delta_i, \sigma, \mathbf{\Gamma})] \xrightarrow{p} \mathbf{0}, \quad (3.11)$$

lo cual probaremos utilizando los argumentos de procesos empíricos mencionados en la sección anterior.

Para demostrar la convergencia (3.11) tenemos que probar que para cualquier $\zeta > 0$, $P(|T_n(\widehat{\sigma}, \widehat{\mathbf{\Gamma}})| > \zeta) \rightarrow 0$. Sean \mathcal{S} y \mathcal{G} los entornos de σ y $\mathbf{\Gamma}$, respectivamente, mencionados anteriormente y sea c tal que $c^{-1} < s < c$ para todo $s \in \mathcal{S}$.

Sumando y restando $L(y, \mathbf{x}, \delta, s, \mathbf{G})$ y haciendo un desarrollo de Taylor de primer orden para ψ_1 , después de un poco de álgebra, obtenemos que

$$\begin{aligned} |L(y, \mathbf{x}, \delta, s, \mathbf{G}) - L(y, \mathbf{x}, \delta, \sigma, \mathbf{\Gamma})| &\leq \left| \delta \psi_1' \left(\frac{y - f(\mathbf{x}, \boldsymbol{\theta})}{\sigma^*} \right) \frac{f(\mathbf{x}, \boldsymbol{\theta}) - y}{\sigma^{*2}} \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) w(\mathbf{x}, \mathbf{\Gamma}) \right| |s - \sigma| \\ &\quad + \left| \delta \psi_1 \left(\frac{y - f(\mathbf{x}, \boldsymbol{\theta})}{s} \right) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) (w(\mathbf{x}, \mathbf{G}) - w(\mathbf{x}, \mathbf{\Gamma})) \right| \\ &\leq \left(\|\eta_1\|_\infty \frac{1}{\sigma^*} \alpha(\mathbf{x}) + \|\psi_1\|_\infty |\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})| \tilde{\alpha}(\mathbf{x}, \mathbf{G}) \right) |(s, \mathbf{G}) - (\sigma, \mathbf{\Gamma})|, \end{aligned}$$

donde η_1 fue definida en **N3**, α y $\tilde{\alpha}$ están definidas en **N5** y σ^* es un punto intermedio entre s y σ .

Tomando $q(\mathbf{x}) = \sup_{\mathcal{G}} \left(k_1 \alpha(\mathbf{x}) + k_2 |\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})| \tilde{\alpha}(\mathbf{x}, \mathbf{G}) \right)$ y usando **N1** y **N3** resulta que para cada $s \in \mathcal{S}$ y $\mathbf{G} \in \mathcal{G}$

$$|L(y, \mathbf{x}, \delta, s, \mathbf{G}) - L(y, \mathbf{x}, \delta, \sigma, \mathbf{\Gamma})| \leq q(\mathbf{x}) |(s, \mathbf{G}) - (\sigma, \mathbf{\Gamma})|. \quad (3.12)$$

Consideremos, entonces, la siguiente familia de funciones

$$\mathcal{F} = \{g_{s, \mathbf{G}}(y, \mathbf{x}, \delta) = L(y, \mathbf{x}, \delta, s, \mathbf{G}) - L(y, \mathbf{x}, \delta, \sigma, \mathbf{\Gamma}) : s \in \mathcal{S}, \mathbf{G} \in \mathcal{G}\}.$$

Notemos que como los errores son simétricos y ψ_1 es impar, $\mathbb{E}(g_{s, \mathbf{G}}(y, \mathbf{x}, \delta)) = \mathbf{0}$. Además, dado que valen (3.12) y **N5**, tenemos que $r(\mathbf{x}) = q(\mathbf{x}) \sqrt{2}\varepsilon_o$ es una envolvente para \mathcal{F} .

Por otra parte, valen **N1** y **N4**, por lo tanto $L(y, \mathbf{x}, \delta, s, \mathbf{G})$ es continua en (s, \mathbf{G}) , entonces como $\mathcal{S} \times \mathcal{G}$ es separable, por la Proposición 3.2.1, \mathcal{F} es \mathbb{P} -medible.

Asimismo, debido a que vale (3.12), por el Teorema 3.2.3, tenemos que \mathcal{F} cumple la siguiente condición de bracketing

$$N_{[\cdot]}(2\zeta \|q\|_{\mathbb{P}, 2}, \mathcal{F}, L^2(\mathbb{P})) \leq N(\zeta, \mathcal{S} \times \mathcal{G}, |\cdot|).$$

Luego, por la Proposición 3.2.2

$$J_{[\]}(\nu, \mathcal{F}) \xrightarrow{\nu \rightarrow 0} 0.$$

Por consiguiente, como $\widehat{\sigma} \xrightarrow{p} \sigma$ y $\widehat{\Gamma} \xrightarrow{p} \Gamma$, tenemos que dados $\zeta > 0$ y $\nu > 0$, existe n suficientemente grande tal que

$$\begin{aligned} P\left(\left|T_n(\widehat{\sigma}, \widehat{\Gamma})\right| > \zeta\right) &\leq P\left(|\widehat{\sigma} - \sigma| > \nu, |\widehat{\Gamma} - \Gamma| > \nu\right) + P\left(\left|T_n(\widehat{\sigma}, \widehat{\Gamma})\right| > \zeta, |\widehat{\sigma} - \sigma| < \nu, |\widehat{\Gamma} - \Gamma| < \nu\right) \\ &\leq \nu + P\left(\left|T_n(\widehat{\sigma}, \widehat{\Gamma})\right| > \zeta, |\widehat{\sigma} - \sigma| < \nu, |\widehat{\Gamma} - \Gamma| < \nu\right) \\ &\leq \nu + P\left(\sup_{|s-\sigma|<\nu, |\mathbf{G}-\Gamma|<\nu} |T_n(s, \mathbf{G})| > \zeta\right) \\ &\leq \nu + \frac{1}{\zeta} \mathbb{E}\left(\sup_{|s-\sigma|<\nu, |\mathbf{G}-\Gamma|<\nu} |T_n(s, \mathbf{G})| > \zeta\right). \end{aligned}$$

Luego, basta ver que

$$\lim_{\nu \rightarrow 0} \lim_{n \rightarrow \infty} \mathbb{E}\left(\sup_{|s-\sigma|<\nu, |\mathbf{G}-\Gamma|<\nu} |T_n(s, \mathbf{G})| > \zeta\right) = 0.$$

Dado $\nu > 0$, consideremos la subclase

$$\mathcal{F}_\nu = \{g_{s, \mathbf{G}}(y, \mathbf{x}, \delta) : s \in \mathcal{S}, \mathbf{G} \in \mathcal{G}, |s - \sigma| < \nu, |\mathbf{G} - \Gamma| < \nu\}.$$

\mathcal{F}_ν cumple con las condiciones de la desigualdad maximal enunciada en la Proposición 3.2.3 con función envolvente $r(\mathbf{x})$, definida más arriba, de modo que

$$\begin{aligned} \mathbb{E}\left(\sup_{|s-\sigma|<\nu, |\mathbf{G}-\Gamma|<\nu} |T_n(s, \mathbf{G})| > \zeta\right) &\leq D_2 J_{[\]}(\nu, \mathcal{F}_\nu) \left[\mathbb{E}\left(r^2(\mathbf{x})\right)\right]^{1/2} + \sqrt{n} \mathbb{E}\left(r(\mathbf{x}) I_{\{r(\mathbf{x}) > \sqrt{na}(\nu)\}}\right) \\ &\leq D_2 J_{[\]}(\nu, \mathcal{F}) \left[\mathbb{E}\left(r^2(\mathbf{x})\right)\right]^{1/2} + \sqrt{n} \left[\mathbb{E}\left(r^2(\mathbf{x})\right)\right]^{1/2} \\ &\quad + \sqrt{n} \mathbb{E}\left(\frac{r^2(\mathbf{x})}{r(\mathbf{x})} I_{\{r(\mathbf{x}) > \sqrt{na}(\nu)\}}\right) \\ &\leq D_2 J_{[\]}(\nu, \mathcal{F}) \left[\mathbb{E}\left(r^2(\mathbf{x})\right)\right]^{1/2} + \sqrt{n} \frac{1}{\sqrt{na}(\nu)} \mathbb{E}\left(r^2(\mathbf{x}) I_{\{r(\mathbf{x}) > \sqrt{na}(\nu)\}}\right) \\ &= D_2 J_{[\]}(\nu, \mathcal{F}) \left[\mathbb{E}\left(r^2(\mathbf{x})\right)\right]^{1/2} + \frac{1}{a(\nu)} \mathbb{E}\left(r^2(\mathbf{x}) I_{\{r(\mathbf{x}) > \sqrt{na}(\nu)\}}\right) \quad (3.13) \end{aligned}$$

Dado que la función envolvente r tiene segundo momento finito, el segundo término de (3.13) tiende a 0 y por lo tanto,

$$\lim_{n \rightarrow \infty} \mathbb{E}\left(\sup_{|s-\sigma|<\nu, |\mathbf{G}-\Gamma|<\nu} |T_n(s, \mathbf{G})| > \zeta\right) \leq D_2 J_{[\]}(\nu, \mathcal{F}) \mathbb{E}\left(r^2(\mathbf{x})\right)^{1/2}.$$

Finalmente, los resultados deseados se deducen del hecho que $J_{[\cdot]}(\nu, \mathcal{F}) \rightarrow 0$ cuando ν decrece a 0. \square

La condición **N5** podría reemplazarse por la siguiente hipótesis en el caso en que se elija una función de peso w derivable respecto del parámetro \mathbf{G} , y el resultado del Teorema anterior seguiría siendo válido, utilizando para su prueba argumentos semejantes a los empleados en su demostración.

N5'. La función de peso $w(\mathbf{x}, \mathbf{G})$ es derivable como función de \mathbf{G} , tal que si $\dot{w}(\mathbf{x}, \mathbf{G})$ es el gradiente de $w(\mathbf{x}, \mathbf{G})$ con respecto a \mathbf{G} ,

- a) para cada componente del gradiente \dot{w} , \dot{w}_l , se tiene que $\|\dot{w}_l\|_\infty < \infty$,
- b) existe ε_o tal que para cada componente \dot{f}_j , $j = 1, \dots, p$, del gradiente de la función de regresión se verifica que

$$\mathbb{E} \left[\sup_{\mathcal{G}} |\dot{f}_j(\mathbf{x}, \boldsymbol{\theta}) \dot{w}(\mathbf{x}, \mathbf{G})|^2 \right] < \infty.$$

Cabe notar que cuando $p(\mathbf{x}) = 1$, que corresponde al caso en que no hay respuestas faltantes, obtenemos que $\mathbf{A} = \mathbb{E} \psi_1'(\epsilon) \mathbb{E} [\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t w(\mathbf{x}, \boldsymbol{\Gamma})]$ y $\mathbf{B} = \mathbb{E} \psi_1^2(\epsilon) \mathbb{E} [\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t w^2(\mathbf{x}, \boldsymbol{\Gamma})]$. Cuando además $w \equiv 1$, $\mathbf{A} = \mathbb{E} \psi_1'(\epsilon) \mathbb{E} [\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t]$ y $\mathbf{B} = \mathbb{E} \psi_1^2(\epsilon) \mathbb{E} [\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t]$ y en consecuencia, la matriz de covarianza asintótica del estimador $\widehat{\boldsymbol{\theta}}$ corresponde a la de un *MM*-estimador (ver Fasano, 2009).

Por otra parte, para el estimador de mínimos cuadrados $\psi_1(t) = 2t$ y, por lo tanto, $\boldsymbol{\Sigma} = \mathbb{E} [\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t]$

Capítulo 4

Curva de Influencia

4.1. Definición

La curva o función de influencia es una medida de la robustez respecto a un outlier que permite estudiar la estabilidad local de un estimador. Esta importante herramienta constituye la piedra angular de lo que se conoce como el *enfoque infinitesimal*. Fue introducida por Hampel (1974) a fin de evaluar el comportamiento infinitesimal de un estimador dado. Esta medida se utiliza con el fin de cuantificar el “buen” comportamiento del estimador en un “entorno” del modelo propuesto para el cual fue definido. También habilita a estudiar la eficiencia asintótica de un estimador dando a la vez una justificación teórica para la selección de la función de peso y las constantes de calibración. De hecho, se la puede pensar como la primera derivada del funcional asociado al estimador y, bajo condiciones de regularidad, permite la deducción de la matriz de covarianza asintótica del estimador en cuestión.

La curva de influencia es una versión asintótica de lo que conocemos como la *curva de sensibilidad*. Comenzaremos por abordar este primer concepto en una situación genérica en la que $\widehat{\mu}_n$ es un estimador basado en una muestra $\mathbf{z}_1, \dots, \mathbf{z}_n$ de tamaño n de vectores aleatorios i.i.d. con distribución G .

La *curva de sensibilidad* del estimador $\widehat{\mu}_n$ es la diferencia

$$\begin{aligned} SC(\mathbf{z}_o, \mathbf{z}_1, \dots, \mathbf{z}_n, \widehat{\mu}_n) &= \frac{\widehat{\mu}_{n+1}(\mathbf{z}_1, \dots, \mathbf{z}_n, \mathbf{z}_o) - \widehat{\mu}_n(\mathbf{z}_1, \dots, \mathbf{z}_n)}{1/(n+1)} \\ &= (n+1) (\widehat{\mu}_{n+1}(\mathbf{z}_1, \dots, \mathbf{z}_n, \mathbf{z}_o) - \widehat{\mu}_n(\mathbf{z}_1, \dots, \mathbf{z}_n)) \end{aligned}$$

como función del valor atípico \mathbf{z}_o , la cual sirve para medir el efecto de diferentes posiciones de un outlier introducido a la muestra. De hecho, la curva de sensibilidad es n veces el cambio en el estimador al agregar una observación adicional \mathbf{z}_o .

Asumamos que $\widehat{\mu}_n$ es un estimador que puede escribirse como un funcional, es decir $\widehat{\mu}_n = T(G_n)$, donde G_n es la distribución empírica que asigna masa $1/n$ a cada $\mathbf{z}_1, \dots, \mathbf{z}_n$, y que existe $\mu = T(G)$, donde $T(G)$ es el funcional basado en la distribución G . Con esta

notación, la curva de sensibilidad puede escribirse como

$$SC = \frac{T\left(\left(1 - \frac{1}{n+1}\right)G_n + \frac{1}{n+1}\Delta_{\mathbf{z}_o}\right) - T(G_n)}{1/(n+1)},$$

donde $\Delta_{\mathbf{z}_o}$ es la medida de Dirac que da masa 1 al punto \mathbf{z}_o . De esta manera, resulta natural extender esta noción al concepto de curva de influencia, para la cual será necesario estudiar el comportamiento de los estimadores cuando el tamaño de la muestra tiende a infinito. Como dijimos anteriormente nos interesa saber cómo se comporta el estimador en un entorno de una distribución G . Hay varias maneras de caracterizar estos entornos, nos enfocaremos en la familia conocida como *entornos de contaminación* definidos como

$$\mathcal{F}(\varepsilon) = \{G_{\mathbf{z}_o, \varepsilon} : G_{\mathbf{z}_o, \varepsilon} = (1 - \varepsilon)G + \varepsilon\Delta_{\mathbf{z}_o}\}.$$

Introducimos entonces la *función de influencia* (IF) de un estimador como la versión asintótica de su curva de sensibilidad. Esta función es una aproximación del comportamiento del estimador cuando la muestra contiene una pequeña proporción ε de idénticos valores atípicos. Demos ahora la definición formal.

Definición 4.1.1 *Supongamos que el vector aleatorio \mathbf{z} tiene distribución G . La función de influencia del funcional T en G se define como*

$$IF(\mathbf{z}_o, T(G)) = \lim_{\varepsilon \rightarrow 0} \frac{T((1 - \varepsilon)G + \varepsilon\Delta_{\mathbf{z}_o}) - T(G)}{\varepsilon}.$$

Podemos decir entonces que la función de influencia mide aproximadamente el efecto de una contaminación infinitesimal ubicada en un valor fijo. Es claro que si tomamos $\varepsilon = 1/n$ y $G = G_{n-1}$, se reduce a SC , es decir mide aproximadamente n veces el cambio de T debido a una observación adicional en \mathbf{z}_o para una muestra de tamaño grande $n - 1$.

En la siguiente sección se deduce la función de influencia para el estimador solución de (2.14).

4.2. Función de Influencia y Función de Influencia Esperada

Sea G_1 una distribución definida en $\mathbb{R}^{k+1} \times \{0, 1\}$ y notemos por $\theta(G_1)$, $\sigma(G_1)$ y $\Gamma(G_1)$ a los funcionales relacionados a los estimadores $\hat{\theta}$, $\hat{\sigma}$ y $\hat{\Gamma}$, respectivamente. Es decir, $\theta(G_1)$ minimiza $H(\mathbf{b}, \sigma(G_1), \Gamma(G_1), G_1)$, donde

$$H(\mathbf{b}, \sigma(G_1), \Gamma(G_1), G_1) = \mathbb{E}_{G_1} \left[\delta\rho_1 \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{\sigma(G_1)} \right) w(\mathbf{x}, \Gamma(G_1)) \right]$$

y por lo tanto, resuelve el siguiente sistema de ecuaciones

$$H^{(1)}(\mathbf{b}, \sigma(G_1), \Gamma(G_1), G_1) = \mathbb{E}_{G_1} \left[p(\mathbf{x})\psi_1 \left(\frac{y - f(\mathbf{x}, \mathbf{b})}{\sigma(G_1)} \right) \dot{\mathbf{f}}(\mathbf{x}, \mathbf{b})w(\mathbf{x}, \Gamma(G_1)) \right] = \mathbf{0}.$$

Bajo el supuesto de que $\theta(G_1)$ es un funcional Fisher-consistente en G , es decir $\theta(G) = \theta$, demostraremos la siguiente proposición.

Proposición 4.2.1 *Supongamos que se cumplen N1, N2, N6 y N10 y que la función de peso $w(\mathbf{x}, \mathbf{G})$ es derivable con respecto a \mathbf{G} tal que existe ε_o de manera que*

$\mathbb{E} \left[\sup_{\mathcal{B}, \mathcal{G}} |\dot{f}_j(\mathbf{x}, \theta) \dot{w}(\mathbf{x}, \mathbf{G})|^2 \right] < \infty$ para cada componente \dot{f}_j , $j = 1, \dots, p$, del gradiente de la función de regresión. Asumamos, además que $\mathbb{E}_G \left[\psi_1 \left(\frac{y-f(\mathbf{x}, \theta)}{\sigma} \right) \middle| \mathbf{x} \right] = \mathbb{E}_G \left[\psi'_1 \left(\frac{y-f(\mathbf{x}, \theta)}{\sigma} \right) \frac{y-f(\mathbf{x}, \theta)}{\sigma} \middle| \mathbf{x} \right] = 0$, que se satisface la condición MAR dada en (2.2) y que las funciones de influencia $IF(\mathbf{z}_o, \sigma(G))$ y $IF(\mathbf{z}_o, \Gamma(G))$ existen y $\sigma(G) = \sigma$ y $\Gamma(G) = \Gamma$. Entonces $IF(\mathbf{z}_o, \theta(G))$ existe y tenemos que

$$IF(\mathbf{z}_o, \theta(G)) = \delta_o \sigma \psi_1 \left(\frac{y_o - f(\mathbf{x}_o, \theta)}{\sigma} \right) w(\mathbf{x}_o, \Gamma) \mathbf{A}^{-1} \dot{\mathbf{f}}(\mathbf{x}_o, \theta), \quad (4.1)$$

donde $\mathbf{z}_o = (y_o, \mathbf{x}_o^t, \delta_o)$ y \mathbf{A} es la matriz que está definida en N10.

Demostración: Sea $G_{\mathbf{z}_o, \varepsilon} \in \mathcal{F}(\varepsilon)$, como

$$\mathbb{E}_{G_{\mathbf{z}_o, \varepsilon}} \left[\psi_1 \left(\frac{y - f(\mathbf{x}, \theta(G_{\mathbf{z}_o, \varepsilon}))}{\sigma(G_{\mathbf{z}_o, \varepsilon})} \right) \dot{\mathbf{f}}(\mathbf{x}, \theta(G_{\mathbf{z}_o, \varepsilon})) w(\mathbf{x}, \Gamma(G_{\mathbf{z}_o, \varepsilon})) \delta \right] = \mathbf{0}$$

y $G_{\mathbf{z}_o, \varepsilon} = (1 - \varepsilon)G + \varepsilon \Delta_{\mathbf{z}_o}$, entonces

$$\begin{aligned} \mathbf{0} &= (1 - \varepsilon) \mathbb{E}_G \left[\psi_1 \left(\frac{y - f(\mathbf{x}, \theta(G_{\mathbf{z}_o, \varepsilon}))}{\sigma(G_{\mathbf{z}_o, \varepsilon})} \right) \dot{\mathbf{f}}(\mathbf{x}, \theta(G_{\mathbf{z}_o, \varepsilon})) w(\mathbf{x}, \Gamma(G_{\mathbf{z}_o, \varepsilon})) p(\mathbf{x}) \right] \\ &+ \varepsilon \psi_1 \left(\frac{y_o - f(\mathbf{x}_o, \theta(G_{\mathbf{z}_o, \varepsilon}))}{\sigma(G_{\mathbf{z}_o, \varepsilon})} \right) \dot{\mathbf{f}}(\mathbf{x}_o, \theta(G_{\mathbf{z}_o, \varepsilon})) w(\mathbf{x}_o, \Gamma(G_{\mathbf{z}_o, \varepsilon})) \delta_o. \end{aligned} \quad (4.2)$$

Por lo tanto, si derivamos (4.2) con respecto a ε y evaluamos en $\varepsilon = 0$, nos queda que

$$\begin{aligned} \mathbf{0} &= -\mathbb{E}_G \left[p(\mathbf{x}) \psi_1 \left(\frac{y - f(\mathbf{x}, \theta)}{\sigma} \right) \dot{\mathbf{f}}(\mathbf{x}, \theta) w(\mathbf{x}, \Gamma) \right] \\ &- \frac{1}{\sigma} \mathbb{E}_G \left[p(\mathbf{x}) \psi'_1 \left(\frac{y - f(\mathbf{x}, \theta)}{\sigma} \right) \dot{\mathbf{f}}(\mathbf{x}, \theta) \dot{\mathbf{f}}(\mathbf{x}, \theta)^t w(\mathbf{x}, \Gamma) \right] IF(\mathbf{z}_o, \theta(G)) \\ &- \mathbb{E}_G \left[p(\mathbf{x}) \psi'_1 \left(\frac{y - f(\mathbf{x}, \theta)}{\sigma} \right) \frac{y - f(\mathbf{x}, \theta)}{\sigma^2} \dot{\mathbf{f}}(\mathbf{x}, \theta) w(\mathbf{x}, \Gamma) \right] IF(\mathbf{z}_o, \sigma(G)) \\ &+ \mathbb{E}_G \left[p(\mathbf{x}) \psi_1 \left(\frac{y - f(\mathbf{x}, \theta)}{\sigma} \right) \dot{\mathbf{f}}(\mathbf{x}, \theta) w(\mathbf{x}, \Gamma) \right] IF(\mathbf{z}_o, \theta(G)) \\ &+ \mathbb{E}_G \left[p(\mathbf{x}) \psi_1 \left(\frac{y - f(\mathbf{x}, \theta)}{\sigma} \right) \dot{\mathbf{f}}(\mathbf{x}, \theta) \dot{w}(\mathbf{x}, \Gamma)^t IF(\mathbf{z}_o, \Gamma(G)) \right] \\ &+ \delta_o \psi_1 \left(\frac{y_o - f(\mathbf{x}_o, \theta)}{\sigma} \right) \dot{\mathbf{f}}(\mathbf{x}_o, \theta) w(\mathbf{x}_o, \Gamma). \end{aligned}$$

Dado que $\mathbb{E}_G \left(\psi_1 \left(\frac{y-f(\mathbf{x}, \boldsymbol{\theta})}{\sigma} \right) \middle| \mathbf{x} \right) = \mathbb{E}_G \left(\psi_1' \left(\frac{y-f(\mathbf{x}, \boldsymbol{\theta})}{\sigma} \right) \frac{y-f(\mathbf{x}, \boldsymbol{\theta})}{\sigma} \middle| \mathbf{x} \right) = 0$, obtenemos que

$$\text{IF}(\mathbf{z}_o, \boldsymbol{\theta}(G)) = \mathbf{A}^{-1} \delta_o \psi_1 \left(\frac{y_o - f(\mathbf{x}_o, \boldsymbol{\theta})}{\sigma} \right) \dot{\mathbf{f}}(\mathbf{x}_o, \boldsymbol{\theta}) w(\mathbf{x}_o, \boldsymbol{\Gamma}) \sigma,$$

como queríamos demostrar. \square

Notemos que como bajo el modelo (1.1), $y = f(\mathbf{x}, \boldsymbol{\theta}) + \sigma\epsilon$, luego si ϵ tiene distribución simétrica alrededor de 0, es independiente de \mathbf{x} y valen **N1** y **N2**, $\psi_1(t)$ y $t\psi_1(t)$ son funciones impares y tenemos que $\mathbb{E}_G \left[\psi_1 \left(\frac{y-f(\mathbf{x}, \boldsymbol{\theta})}{\sigma} \right) \middle| \mathbf{x} \right] = \mathbb{E}_G \left[\psi_1' \left(\frac{y-f(\mathbf{x}, \boldsymbol{\theta})}{\sigma} \right) \frac{y-f(\mathbf{x}, \boldsymbol{\theta})}{\sigma} \middle| \mathbf{x} \right] = 0$.

Cabe destacar que la $\text{IF}(\mathbf{z}_o, \boldsymbol{\theta}(G))$ depende de δ_o que es una variable binaria, con lo cual la función de influencia valdrá cero si y_o es una respuesta faltante. Por este motivo, Bianco *et al.* (2013) consideraron la función de influencia esperada, dada por

$$\text{EIF}(\bar{\mathbf{z}}_o, T(G)) = \mathbb{E}_G[\text{IF}(\mathbf{z}_o, T(G)) | \bar{\mathbf{z}}_o],$$

siendo $\bar{\mathbf{z}}_o = (y_o, \mathbf{x}_o^t)$, como una herramienta alternativa para medir el efecto de outliers sobre un funcional dado cuando hay que pérdida de datos.

La EIF mide la influencia esperada condicional de una observación, sin importar si la respuesta fue observada o no. Nos da una idea del efecto de la probabilidad de pérdida sobre el funcional y sobre el estimador relacionado.

Corolario 4.2.1 *En las condiciones de la Proposición 4.2.1, tenemos que*

$$\text{EIF}(\bar{\mathbf{z}}_o, \boldsymbol{\theta}(G)) = p(\mathbf{x}_o) \sigma \psi_1 \left(\frac{y_o - f(\mathbf{x}_o, \boldsymbol{\theta})}{\sigma} \right) w(\mathbf{x}_o, \boldsymbol{\Gamma}) \mathbf{A}^{-1} \dot{\mathbf{f}}(\mathbf{x}_o, \boldsymbol{\theta}).$$

Por lo tanto, la EIF de nuestro estimador resultará acotada de acuerdo a una correcta elección de la función de score ψ_1 y la función de peso w , de manera de poder controlar a los residuos y al gradiente.

Observemos que para el caso del estimador clásico $\psi_1(t) = 2t$ y $w \equiv 1$, por lo tanto su función de influencia esperada resulta ser

$$\text{EIF}_{\text{MC}} = p(\mathbf{x}_o)(y_o - f(\mathbf{x}_o, \boldsymbol{\theta})) \mathbb{E}_G \left[p(\mathbf{x}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t \right]^{-1} \dot{\mathbf{f}}(\mathbf{x}_o, \boldsymbol{\theta}), \text{ la cual no será acotada.}$$

El caso del *MM*-estimador se obtiene tomando $w \equiv 1$ (ver Fasano *et al.*, 2012) y queda que $\text{EIF}_{\text{MM}} = p(\mathbf{x}_o) \sigma \psi_1 \left(\frac{y_o - f(\mathbf{x}_o, \boldsymbol{\theta})}{\sigma} \right) \mathbf{A}^{-1} \dot{\mathbf{f}}(\mathbf{x}_o, \boldsymbol{\theta})$.

4.3. Ejemplos Numéricos

Con el objetivo de comparar el comportamiento de la función de influencia esperada de los estimadores pesados (*wmm*) con la del estimador clásico (*mc*) y la de los

MM-estimadores, los cuales fueron presentados en el Capítulo 2, consideramos dos ejemplos numéricos usando tres esquemas para la función de probabilidad de pérdida de las respuestas:

$$\begin{aligned} p_{\text{COMP}}(x) &= 1, \\ p_{\text{LOG}}(x) &= (1 + \exp(-ax - 0,2))^{-1}, \\ p_{\text{COS}}(x) &= 0,8 + 0,2 \cos^2(bx + 0,4), \end{aligned}$$

donde las constantes a y b serán elegidas para obtener distintos porcentajes de respuestas faltantes.

4.3.1. Modelo de Michaelis–Menten

En el primer ejemplo, consideramos el modelo de Michaelis–Menten en la parametrización dada por Ratkowsky (1983), es decir que la función de regresión está dada por

$$f(x, \theta) = \frac{\theta_1 x}{\exp(\theta_2) + x}, \quad x > 0.$$

Este modelo se utiliza frecuentemente para ajustar la cinética de enzimas.

Para nuestro estudio consideramos

$$\theta_1 = 10, \quad \theta_2 = 0, \quad \sigma = 1, \quad x \sim U(0, 10), \quad y \quad \epsilon \sim N(0, 1).$$

Las constantes utilizadas para las probabilidades de pérdida fueron $a = 0,2$ y $b = 0,2$, las cuales nos llevan a tener aproximadamente un 25 % y un 15 % de respuestas faltantes con p_{LOG} y p_{COS} , respectivamente. La ρ -función que utilizamos para calcular las EIF de los estimadores robustos, es decir de los *WMM*- y de los *MM*-estimadores, la elegimos dentro de la familia de la bicuadrada definida en (2.6) con $c = 4$. Asimismo, para el cómputo del estimador pesado utilizamos una función de peso w_o dentro de la familia inducida por la bicuadrada con argumento igual a la distancia de Mahalanobis robusta basada en la covariable x . Esto es

$$w_o(t) = \left(1 - \frac{t^2}{c_w^2}\right)^2 \frac{6}{c_w^2} \mathbb{I}\{|t| \leq c_w\}, \quad (4.3)$$

$$w(x, \mathbf{\Gamma}) = w_o(d(x, \mu_x, \sigma_x)), \quad \text{con } \mathbf{\Gamma} = (\mu_x, \sigma_x), \quad (4.4)$$

donde d es la distancia de Mahalanobis y μ_x y σ_x fueron tomadas como la mediana y la MAD de x , respectivamente, y tomamos $c_w = \chi_{1,0,95}^2$.

Para ilustrar el comportamiento de las EIF, calculamos su norma euclídea $|EIF|$ sobre una grilla de valores (x_0, y_0) , equidistantes, de tamaño 134×51 dentro de los intervalos $[0; 39,9] \times [0; 15]$. En la Figura 4.1 graficamos los valores obtenidos de $|EIF|$ en función de x_0 e y_0 .

Observando estos gráficos podemos ver que la $|EIF|$ del funcional relacionado al estimador de mínimos cuadrados, aun cuando está acotada, puede alcanzar valores muy altos

debido tanto a grandes residuos como a grandes valores de la variable x . Este comportamiento empeora bajo el esquema de pérdida logístico, mientras que cuando la probabilidad de respuestas faltantes sigue un esquema dado por el coseno aparece un efecto de fluctuación. El comportamiento de la norma en estudio de los WMM - y MM -estimadores se asemeja al del estimador de mínimos cuadrados para valores moderados de $\bar{\mathbf{z}}_o = (y_o, x_o)^t$ bajo los tres esquemas de pérdida considerados. Si bien la norma de la función de influencia esperada del MM -estimador no crece indefinidamente, debido a que el gradiente de la función de regresión es acotado, ésta resulta ser mucho mayor que la del WMM -estimador, ya que éste penaliza aquellas observaciones con valores altos en la covariable por medio de la función de peso. Esto sucede para los tres esquemas de falta de respuestas considerados.

4.3.2. Modelo Exponencial

Como un segundo estudio, consideramos el modelo de crecimiento exponencial dado por la relación

$$y = \theta_1 \exp(\theta_2 x) + \sigma \epsilon, \quad (4.5)$$

con $\theta_1 = 5$, $\theta_2 = 2$, $\sigma = 1$ y $x \sim U(0, 1)$ y $\epsilon \sim N(0, 1)$, valores que fueron considerados por Fasano (2009).

Para las probabilidades de pérdida de las respuestas elegimos los valores $a = 2$ y $b = 10$ de manera de obtener aproximadamente un 25 % y un 10 % de respuestas faltantes para $p_{\text{LOG}}(x)$ y $p_{\text{COS}}(x)$, respectivamente. Para calcular la EIF del WMM - y del MM -estimador utilizamos la misma ρ -función y función de peso que en el ejemplo anterior con constantes de calibración, $c = 4$ y $c_w = \chi_{1,0,95}^2$, respectivamente. Asimismo, en la Figura 4.2, graficamos la norma de la función de influencia esperada, $|EIF|$ en función de x_o e y_o a lo largo de una grilla de 141×101 valores equidistantes dentro de los intervalos $[0; 1,4] \times [0; 50]$.

Para este modelo podemos ver que la $|EIF|$ del estimador clásico no es acotada. Aunque el MM -estimador logra controlar la influencia en una gran parte de la región en estudio, vemos que a medida que x_o se aproxima a 1 e y_o crece la norma de la esperanza de la influencia crece, indefinidamente. Luego, cuando utilizamos la función de peso para acotar las variables explicativas cuya distancia de Mahalanobis sea grande, la EIF resulta siempre acotada y vale cero tanto para valores extremos de la covariable como de la variable de respuesta.

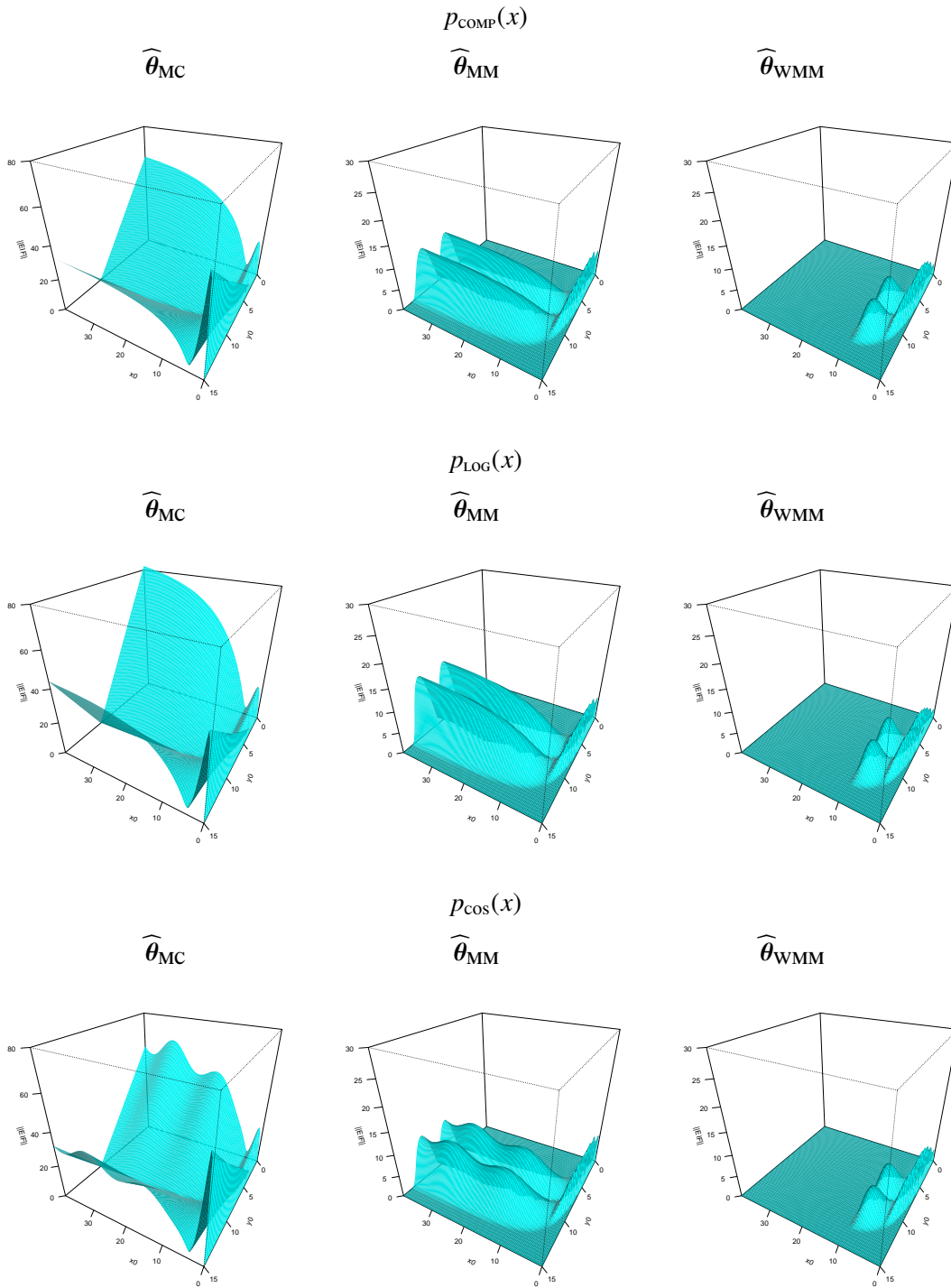


Figura 4.1: Gráficos de la |EIF|, en función de x_0 e y_0 para el Modelo de Michaelis-Menten.

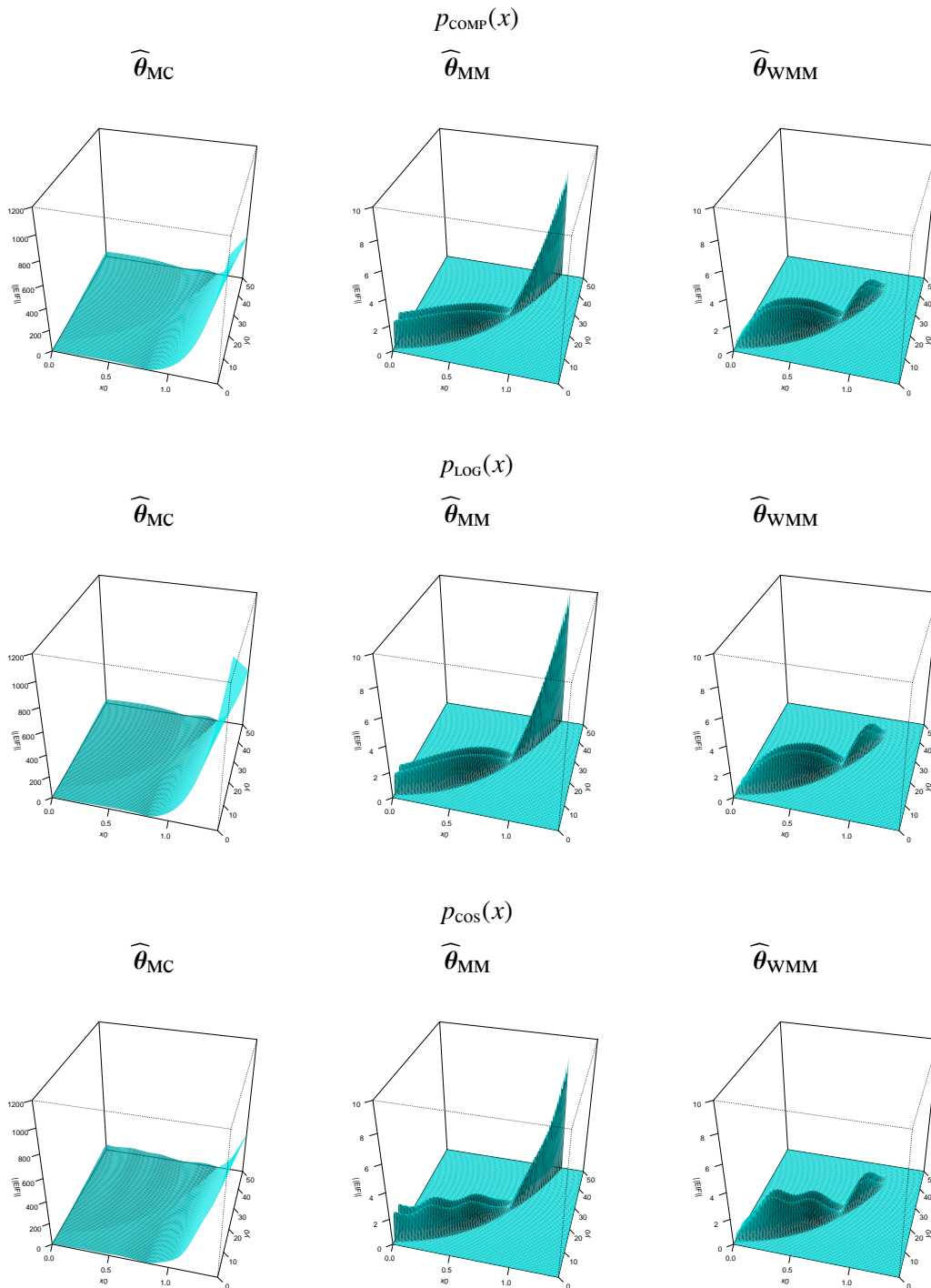


Figura 4.2: Gráficos de la $|\text{EIF}|$, en función de x_0 e y_0 para el Modelo Exponencial.

Capítulo 5

Estudio de Simulación

Por medio de un estudio de Monte Carlo comparamos el desempeño del procedimiento propuesto con el del estimador clásico y con el de otros dos estimadores robustos. Para este estudio consideramos dos modelos para la función de regresión con diferentes mecanismos de pérdida de las respuestas y distintas configuraciones de contaminación.

5.1. Modelo de Crecimiento Exponencial

Para el primer experimento numérico consideramos el modelo exponencial dado por (4.5). Siguiendo este modelo, generamos $NR = 1000$ réplicas de muestras (y_i, x_i) de tamaño $n = 50$ con

$$\theta_1 = 5, \theta_2 = 2, \sigma = 1, x_i \sim U(0, 1) \text{ y } \epsilon_i \sim N(0, 1).$$

Implementamos tres esquemas para la función de pérdida de las respuestas:

$$\begin{aligned} p_{\text{COMP}}(x) &= 1, \\ p_{\text{LOG}}(x) &= (1 + \exp(-ax - 0,2))^{-1}, \text{ con } a = 2 \\ p_{\text{COS}}(x) &= 0,8 + 0,2 \cos^2(bx + 0,4), \text{ con } b = 2 \end{aligned}$$

las cuales tienen una probabilidad esperada de pérdida de las respuestas aproximadamente igual a 25 % y 15 %, respectivamente.

Los estimadores se evaluaron tanto en muestras sin contaminar, las que notaremos por C_0 , como bajo muestras contaminadas según distintas configuraciones variando la proporción de outliers presentes en la muestra. Si bien se simuló variadas configuraciones para las contaminaciones, en esta tesis solo mostraremos aquellas que resultaron más relevantes.

A los modelos contaminados los llamaremos $C_{1,m}$, $C_{2,m}$, $C_{3,m}$ y $C_{4,m}$ donde m es la cantidad de observaciones reemplazadas por valores atípicos, y los definimos en la siguiente tabla

	x_o	y_o
$C_{1,m}$	1,2	$0,9[\theta_1 \exp(\theta_2 x_o)]$
$C_{2,m}$	1,4	$0,9[\theta_1 \exp(\theta_2 x_o)]$
$C_{3,m}$	1,8	$0,9[\theta_1 \exp(\theta_2 x_o)]$
$C_{4,m}$	2	$0,9[\theta_1 \exp(\theta_2 x_o)]$

Se reemplazaron $m = 5, 7, 10$ y 12 observaciones con estos valores, que representan aproximadamente el 10 %, el 15 %, el 20 % y el 25 % de los datos originales, respectivamente.

En todos los casos calculamos el estimador de mínimos cuadrados (MC) y los estimadores robustos previamente mencionados, el M –, el MM – y el WMM –estimador. Al igual que para la curva de influencia, elegimos la ρ –función dentro de la familia bicuadrada con diferentes constantes de calibración c . Además para el cómputo del estimador propuesto se utilizaron las funciones de peso definidas en (4.3) y (4.4) utilizando las versiones muestrales de la mediana y la MAD y con dos constantes distintas, $c_{w,1} = \chi_{0,95}^2$ y $c_{w,2} = \chi_{0,9}^2$, con el objetivo de obtener dos versiones del WMM –estimador de manera que la segunda sea más robusta que la primera y que notaremos por wmm_1 y wmm_2 , respectivamente.

Con el fin de determinar las constantes c utilizadas para las ρ –funciones para que los estimadores robustos tengan eficiencias asintóticas iguales a 0,85 y 0,90, recordemos que si $\widehat{\boldsymbol{\theta}}$ es un estimador de $\boldsymbol{\theta}$ que resulta asintóticamente normal con matriz de covarianza \mathbf{V} , y $\widehat{\boldsymbol{\theta}}_{MC}$ es el estimador de máxima verosimilitud, que bajo el modelo no lineal coincide con el estimador de mínimos cuadrados, y tiene matriz de covarianza \mathbf{V}_o una manera de medir la eficiencia asintótica de $\widehat{\boldsymbol{\theta}}$ (ver Maronna *et al.*, 2006) es a través de

$$ef(\widehat{\boldsymbol{\theta}}) = \lambda_1(\mathbf{V}^{-1}\mathbf{V}_o),$$

donde $\lambda_1(\mathbf{M})$ es el mayor autovalor de la matriz \mathbf{M} .

Resulta entonces que las eficiencias asintóticas para los estimadores utilizados se calculan de la siguiente manera

$$\begin{aligned} ef(\widehat{\boldsymbol{\theta}}_M) &= \frac{\mathbb{E}[\psi'_1(\epsilon)]^2}{\mathbb{E}[\psi_1^2(\epsilon)]}, \\ ef(\widehat{\boldsymbol{\theta}}_{MM}) &= ef(\widehat{\boldsymbol{\theta}}_M), \\ ef(\widehat{\boldsymbol{\theta}}_{WMM}) &= \lambda_1(\mathbf{A}\mathbf{B}^{-1}\mathbf{A}\mathbb{E}[\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t]), \end{aligned}$$

donde \mathbf{A} y \mathbf{B} están definidas en N10 y en el Teorema 3.4.2, respectivamente, y $\mathbb{E}[\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t]$ es la matriz de covarianza del estimador de mínimos cuadrados.

De esta manera obtuvimos la siguiente tabla de constantes de calibración, en base a las distribuciones mencionadas previamente, para el cálculo de los estimadores robustos bajo las eficiencias deseadas.

Estimadores	Eficiencia	c_w	c
M-MM	85 %	-	3,44
M-MM	90 %	-	4
WMM ₁	85 %	$\chi^2_{1,0,95}$	3,75
WMM ₁	90 %	$\chi^2_{1,0,95}$	4,5
WMM ₂	85 %	$\chi^2_{1,0,9}$	4,75

Utilizamos dos tipos de estimadores como iniciales para las versiones robustas. El primero se dedujo de la linealización del modelo que resulta de aplicar logaritmo a ambos lados de la igualdad (4.5), obteniendo

$$\widehat{\theta}_{1,\text{MED}} = \exp\left(\text{Med}(\log(y_i)) - \widehat{\theta}_{2,\text{MED}}\text{Med}(x_i)\right), \quad (5.1)$$

$$\widehat{\theta}_{2,\text{MED}} = \text{Med}\left(\frac{\log(y_i) - \text{Med}(\log(y_i))}{x_i - \text{Med}(x_i)}\right), \quad (5.2)$$

donde las medianas se calculan sobre las observaciones con $\delta_i = 1$.

El segundo estimador inicial que utilizamos fue el S -estimador definido en la Sección 2.3.

Cabe observar que establecer un punto inicial es un problema crucial para los modelos no lineales, por este motivo también realizamos simulaciones utilizando como punto de partida el LMS -estimador definido por (2.10) pero estos resultados no serán mostrados ya que este estimador si bien es altamente robusto, resulta muy inestable computacionalmente para este modelo en las condiciones que consideramos en este trabajo.

En base a los resultados obtenidos calculamos como medidas de resumen a la Media, la Mediana, el Desvío Estándar (SD), la MAD y el Error Cuadrático Medio (ECM). Estos valores se encuentran en el Apéndice en los Cuadros 8.1 al 8.12 para los estimadores con punto inicial basado en las medianas, mientras que los valores correspondientes a los computados a partir de un S -estimador se encuentran en los Cuadros 8.13 al 8.24.

Finalmente, en los Cuadros 5.1 y 5.2 se encuentran los errores cuadráticos medios conjuntos calculados de la siguiente forma. Dado un estimador $\widehat{\theta}$ el error cuadrático medio lo hemos computado como

$$ECM = \frac{1}{NR} \sum_{j=1}^{NR} \left((\widehat{\theta}_{1,j} - \theta_1)^2 + (\widehat{\theta}_{2,j} - \theta_2)^2 \right),$$

donde $(\widehat{\theta}_{1,j}, \widehat{\theta}_{2,j})$ es el valor del estimador obtenido en la replicación j , $j = 1, \dots, NR$. En el Cuadro 5.1 los errores cuadráticos medios de los estimadores robustos corresponden al caso en que se utilizaron como iniciales los valores obtenidos de las ecuaciones (5.1) y (5.2), y los que figuran en el Cuadro 5.2 corresponden a tomar como inicial un S -estimador.

Aunque los errores cuadráticos medios mostrados en estos dos cuadros son muy similares, los obtenidos utilizando en los robustos como inicial un S -estimador, en general,

son más pequeños. Por estas razones, las conclusiones del análisis las haremos a partir de los resultados exhibidos en el Cuadro 5.2. En él podemos observar que, si bien, como es de esperar, hay cierta pérdida de eficiencia, el comportamiento de los estimadores robustos considerados cuando la muestra está sin contaminar (C_0), es similar al del estimador de mínimos cuadrados. Esta conclusión es válida tanto para el caso completo, como para los casos en que hay pérdida de respuestas, ya sea por un esquema logístico como por un esquema de coseno. Al introducir datos atípicos en la muestra se puede ver como el estimador de mínimos cuadrados es altamente afectado aumentando su error cuadrático medio desde cinco hasta veinte veces dependiendo de la contaminación considerada. Asimismo podemos ver que a medida que aumenta el valor de contaminación en la covariable esta suma de los errores va disminuyendo, comportamiento que también ocurre en algunos sectores de la curva de influencia.

Si bien los M - y MM - estimadores tienen errores menores a los del estimador clásico ante la presencia de las contaminaciones consideradas, esta diferencia es casi mínima siendo del mismo orden que los del estimador de mc y su aumento con respecto al caso sin contaminar va desde el doble hasta a más de diez veces. Esto sucede tanto para aquellos que alcanzan un 85 % de eficiencia como los que tienen un 90 % de eficiencia.

Al introducir la función de peso en el cálculo del estimador obteniendo los WMM -estimadores podemos ver que al haber un 10 % de contaminación sus errores cuadráticos medios no se ven afectados siendo muy similares al caso C_0 . Esto también ocurre cuando hay 7 observaciones atípicas (15 % de $n = 50$). Cuando el porcentaje aumenta al 20 % vemos este mismo comportamiento para los casos $C_{2,10}$, $C_{3,10}$ y $C_{4,10}$ aunque para el caso $C_{1,10}$ los WMM -estimadores se ven afectados al aumentar su error en dos veces o más, pero siempre permaneciendo menor que las otras versiones robustas consideradas. Finalmente, al introducir 12 outliers, es decir aproximadamente un 25 % de contaminación, vemos un aumento considerable en los errores para los casos $C_{1,12}$, en el que los errores crecen al mismo nivel que los de los M - y MM -estimadores, y $C_{2,12}$ en que los errores crecen pero permanecen menores a los de sus contrapartidas robustas. Mientras que para los casos $C_{3,12}$ y $C_{4,12}$ los WMM -estimadores considerados logran controlar la presencia de estos tipos de datos atípicos de manera más eficiente. Esto ocurre tanto para el caso de datos completos como para los casos en los que hay pérdida de las respuestas.

Estos resultados permiten concluir que el estimador propuesto resiste hasta un 20 % de contaminación en el sentido en que fue elegida en este estudio.

Por otro lado, vale la pena destacar que si bien los valores de las varianzas de los estimadores WMM que se encuentran en los Cuadros 8.1 al 8.24 aumentan a medida que se incrementa la cantidad de datos atípicos presente en la muestra, la ganancia en el sesgo es mucho mayor ya que su disminución es considerable en relación a sus contrapartidas robustas.

	Eficiencia	85 %				90 %		
p_{COMP}	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁
C_0	0,0277	0,0305	0,0294	0,0352	0,0428	0,0286	0,0279	0,0336
$C_{1,5}$	0,3369	0,2608	0,2821	0,0424	0,0400	0,2977	0,3083	0,0484
$C_{1,7}$	0,3721	0,3412	0,3522	0,0882	0,0494	0,3598	0,3656	0,1061
$C_{1,10}$	0,4085	0,3863	0,3899	0,2893	0,1528	0,3932	0,3953	0,2948
$C_{1,12}$	0,4349	0,3932	0,3999	0,4048	0,3184	0,3998	0,4049	0,4110
$C_{2,5}$	0,2931	0,2154	0,2253	0,0359	0,0399	0,2332	0,2469	0,0348
$C_{2,7}$	0,2925	0,2540	0,2642	0,0383	0,0409	0,2688	0,2783	0,0371
$C_{2,10}$	0,3012	0,2829	0,2869	0,0650	0,0515	0,2874	0,2896	0,0652
$C_{2,12}$	0,3032	0,2849	0,2868	0,1398	0,0921	0,2882	0,2898	0,1405
$C_{3,5}$	0,1431	0,1019	0,1074	0,0366	0,0402	0,1099	0,1138	0,0351
$C_{3,7}$	0,1606	0,1167	0,1205	0,0389	0,0413	0,1224	0,1275	0,0372
$C_{3,10}$	0,2036	0,1370	0,1386	0,0439	0,0436	0,1392	0,1404	0,0406
$C_{3,12}$	0,1442	0,1412	0,1410	0,0509	0,0467	0,1416	0,1417	0,0457
$C_{4,5}$	0,1050	0,0761	0,0780	0,0366	0,0401	0,0803	0,0824	0,0350
$C_{4,7}$	0,1063	0,0858	0,0890	0,0390	0,0412	0,0896	0,0900	0,0372
$C_{4,10}$	0,1316	0,1002	0,1013	0,0444	0,0440	0,1016	0,1021	0,0409
$C_{4,12}$	0,1062	0,1043	0,1041	0,0503	0,0474	0,1045	0,1041	0,0450
p_{LOG}	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁
C_0	0,0358	0,0441	0,0407	0,0487	0,0600	0,0408	0,0387	0,0463
$C_{1,5}$	0,4095	0,3469	0,3638	0,0646	0,0548	0,3741	0,3870	0,0688
$C_{1,7}$	0,4567	0,4300	0,4415	0,1305	0,0674	0,4415	0,4489	0,1434
$C_{1,10}$	0,5112	0,4800	0,4836	0,3588	0,1881	0,4825	0,4856	0,3634
$C_{1,12}$	0,5146	0,4937	0,5009	0,4984	0,3775	0,4935	0,5020	0,4936
$C_{2,5}$	0,3270	0,2626	0,2784	0,0493	0,0544	0,2835	0,2976	0,0476
$C_{2,7}$	0,3382	0,3116	0,3195	0,0544	0,0552	0,3228	0,3262	0,0518
$C_{2,10}$	0,3479	0,3315	0,3362	0,0895	0,0711	0,3336	0,3370	0,0867
$C_{2,12}$	0,3483	0,3451	0,3471	0,1918	0,1272	0,3448	0,3457	0,1840
$C_{3,5}$	0,1583	0,1261	0,1272	0,0506	0,0549	0,1290	0,1341	0,0484
$C_{3,7}$	0,1701	0,1441	0,1485	0,0562	0,0570	0,1489	0,1498	0,0523
$C_{3,10}$	0,1943	0,1594	0,1608	0,0709	0,0641	0,1597	0,1602	0,0611
$C_{3,12}$	0,1588	0,1644	0,1648	0,0888	0,0770	0,1623	0,1640	0,0763
$C_{4,5}$	0,1149	0,0943	0,0964	0,0504	0,0546	0,0981	0,1001	0,0481
$C_{4,7}$	0,1165	0,1072	0,1080	0,0565	0,0570	0,1093	0,1099	0,0523
$C_{4,10}$	0,1203	0,1188	0,1178	0,0715	0,0643	0,1175	0,1172	0,0616
$C_{4,12}$	0,1156	0,1261	0,1242	0,0881	0,0761	0,1220	0,1220	0,0745
p_{COS}	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁
C_0	0,0286	0,0335	0,0322	0,0400	0,0498	0,0315	0,0307	0,0380
$C_{1,5}$	0,3412	0,2833	0,3062	0,0515	0,0464	0,3148	0,3279	0,0570
$C_{1,7}$	0,3821	0,3555	0,3646	0,1076	0,0566	0,3673	0,3723	0,1226
$C_{1,10}$	0,4167	0,3858	0,3914	0,3079	0,1681	0,3929	0,3968	0,3151
$C_{1,12}$	0,4323	0,3950	0,4030	0,4271	0,3392	0,3992	0,4065	0,4276
$C_{2,5}$	0,2862	0,2283	0,2423	0,0408	0,0464	0,2467	0,2621	0,0392
$C_{2,7}$	0,2971	0,2681	0,2791	0,0431	0,0468	0,2800	0,2863	0,0424
$C_{2,10}$	0,3050	0,2859	0,2893	0,0707	0,0580	0,2899	0,2926	0,0688
$C_{2,12}$	0,3078	0,2896	0,2940	0,1608	0,1024	0,2911	0,2946	0,1586
$C_{3,5}$	0,1439	0,1037	0,1070	0,0412	0,0466	0,1100	0,1149	0,0395
$C_{3,7}$	0,1454	0,1163	0,1203	0,0438	0,0473	0,1240	0,1275	0,0416
$C_{3,10}$	0,1933	0,1384	0,1405	0,0494	0,0493	0,1410	0,1425	0,0453
$C_{3,12}$	0,1805	0,1404	0,1407	0,0573	0,0545	0,1411	0,1415	0,0523
$C_{4,5}$	0,1061	0,0744	0,0777	0,0409	0,0464	0,0783	0,0799	0,0393
$C_{4,7}$	0,1086	0,0855	0,0856	0,0432	0,0471	0,0874	0,0897	0,0414
$C_{4,10}$	0,1083	0,0979	0,1001	0,0488	0,0491	0,1016	0,1014	0,0449
$C_{4,12}$	0,1077	0,1055	0,1057	0,0561	0,0541	0,1056	0,1060	0,0506

Cuadro 5.1: Errores cuadráticos medios utilizando en los estimadores robustos como iniciales los descriptos en (5.1) y (5.2), $n = 50$.

	Eficiencia	85 %				90 %		
p_{COMP}	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁
C_0	0,0277	0,0316	0,0316	0,0371	0,0436	0,0293	0,0293	0,0345
$C_{1,5}$	0,3369	0,2366	0,2353	0,0423	0,0415	0,2696	0,2686	0,0435
$C_{1,7}$	0,3721	0,3331	0,3280	0,0826	0,0495	0,3467	0,3457	0,0897
$C_{1,10}$	0,4085	0,3725	0,3694	0,2665	0,1412	0,3781	0,3761	0,2791
$C_{1,12}$	0,4349	0,3725	0,3718	0,3861	0,2999	0,3776	0,3770	0,3928
$C_{2,5}$	0,2931	0,2248	0,2253	0,0401	0,0421	0,2346	0,2330	0,0370
$C_{2,7}$	0,2925	0,2699	0,2694	0,0447	0,0441	0,2743	0,2730	0,0406
$C_{2,10}$	0,3012	0,2804	0,2801	0,0727	0,0565	0,2831	0,2830	0,0655
$C_{2,12}$	0,3032	0,2820	0,2819	0,1484	0,0957	0,2832	0,2827	0,1421
$C_{3,5}$	0,1431	0,1191	0,1182	0,0398	0,0418	0,1203	0,1208	0,0368
$C_{3,7}$	0,1606	0,1341	0,1349	0,0426	0,0432	0,1337	0,1337	0,0391
$C_{3,10}$	0,2036	0,1373	0,1383	0,0495	0,0467	0,1383	0,1380	0,0436
$C_{3,12}$	0,1442	0,1424	0,1422	0,0587	0,0515	0,1414	0,1415	0,0498
$C_{4,5}$	0,1050	0,0755	0,0760	0,0374	0,0407	0,0790	0,0784	0,0354
$C_{4,7}$	0,1063	0,0850	0,0852	0,0392	0,0415	0,0879	0,0872	0,0373
$C_{4,10}$	0,1316	0,0902	0,0896	0,0411	0,0425	0,0919	0,0924	0,0391
$C_{4,12}$	0,1062	0,0941	0,0928	0,0435	0,0439	0,0962	0,0952	0,0412
p_{LOG}	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁
C_0	0,0358	0,0467	0,0466	0,0540	0,0630	0,0428	0,0425	0,0492
$C_{1,5}$	0,4095	0,3235	0,3188	0,0632	0,0571	0,3466	0,3469	0,0638
$C_{1,7}$	0,4567	0,4157	0,4127	0,1224	0,0664	0,4247	0,4230	0,1238
$C_{1,10}$	0,5112	0,4639	0,4605	0,3437	0,1814	0,4647	0,4619	0,3445
$C_{1,12}$	0,5146	0,4853	0,4804	0,4830	0,3687	0,4792	0,4739	0,4771
$C_{2,5}$	0,3270	0,2795	0,2795	0,0579	0,0588	0,2834	0,2817	0,0525
$C_{2,7}$	0,3382	0,3194	0,3186	0,0673	0,0622	0,3209	0,3213	0,0588
$C_{2,10}$	0,3479	0,3327	0,3333	0,1227	0,0868	0,3317	0,3329	0,1015
$C_{2,12}$	0,3483	0,3418	0,3425	0,2227	0,1522	0,3356	0,3346	0,2022
$C_{3,5}$	0,1583	0,1428	0,1434	0,0575	0,0587	0,1419	0,1420	0,0521
$C_{3,7}$	0,1701	0,1543	0,1542	0,0660	0,0630	0,1541	0,1541	0,0580
$C_{3,10}$	0,1943	0,1622	0,1611	0,0898	0,0743	0,1580	0,1586	0,0723
$C_{3,12}$	0,1588	0,1679	0,1692	0,1194	0,0955	0,1647	0,1652	0,0944
$C_{4,5}$	0,1149	0,0976	0,0964	0,0523	0,0555	0,0967	0,0972	0,0491
$C_{4,7}$	0,1165	0,1010	0,1002	0,0564	0,0567	0,1020	0,1016	0,0522
$C_{4,10}$	0,1203	0,1120	0,1113	0,0651	0,0611	0,1120	0,1118	0,0584
$C_{4,12}$	0,1156	0,1148	0,1135	0,0764	0,0700	0,1160	0,1142	0,0682
p_{COS}	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁
C_0	0,0286	0,0351	0,0353	0,0431	0,0512	0,0324	0,0325	0,0398
$C_{1,5}$	0,3412	0,2563	0,2535	0,0512	0,0484	0,2854	0,2832	0,0506
$C_{1,7}$	0,3821	0,3449	0,3418	0,0961	0,0583	0,3552	0,3556	0,1064
$C_{1,10}$	0,4167	0,3767	0,3721	0,2928	0,1576	0,3782	0,3758	0,3015
$C_{1,12}$	0,4323	0,3833	0,3822	0,4114	0,3248	0,3804	0,3785	0,4128
$C_{2,5}$	0,2862	0,2429	0,2419	0,0461	0,0491	0,2506	0,2495	0,0420
$C_{2,7}$	0,2971	0,2757	0,2757	0,0529	0,0519	0,2788	0,2788	0,0471
$C_{2,10}$	0,3050	0,2864	0,2866	0,0898	0,0685	0,2865	0,2867	0,0789
$C_{2,12}$	0,3078	0,2908	0,2919	0,1787	0,1147	0,2897	0,2887	0,1611
$C_{3,5}$	0,1439	0,1226	0,1207	0,0455	0,0486	0,1228	0,1229	0,0414
$C_{3,7}$	0,1454	0,1334	0,1331	0,0498	0,0499	0,1337	0,1344	0,0450
$C_{3,10}$	0,1933	0,1391	0,1400	0,0597	0,0552	0,1394	0,1394	0,0514
$C_{3,12}$	0,1805	0,1461	0,1457	0,0714	0,0602	0,1462	0,1449	0,0590
$C_{4,5}$	0,1061	0,0809	0,0810	0,0426	0,0473	0,0824	0,0820	0,0400
$C_{4,7}$	0,1086	0,0872	0,0871	0,0444	0,0477	0,0884	0,0888	0,0419
$C_{4,10}$	0,1083	0,0951	0,0946	0,0463	0,0472	0,0963	0,0973	0,0429
$C_{4,12}$	0,1077	0,0978	0,0975	0,0525	0,0506	0,0977	0,0985	0,0472

Cuadro 5.2: Errores cuadráticos medios utilizando en los estimadores robustos como iniciales un S -estimador, $n = 50$.

5.2. Modelo Lineal–Exponencial

Como una segunda evaluación de nuestra propuesta elegimos un modelo separable discutido en Gallant (1975). De esta manera, generamos $NR = 1000$ muestras $(y_i, \mathbf{x}_i^\dagger)$, con $i = 1, \dots, 100$, tales que

$$y_i = \theta_1 + \theta_2 x_{i2} + \theta_3 \exp(\theta_0 x_{i1}) + \sigma \epsilon_i,$$

donde $\theta_0 = -1, \theta_1 = 1, \theta_2 = 0, \theta_3 = -0,5, x_{i1} \sim U(0, 10), x_{i2} \sim N(0,5; 0,1^2), \sigma = 0,03$ y $\epsilon_i \sim N(0, 1)$ para todo i .

En primera instancia, consideramos el conjunto de datos completos, es decir, en el cual la probabilidad de la variable dicotómica es $p_{\text{COMP}}(\mathbf{x}) = 1$. Estudiamos el caso sin contaminar, denotado por C_0^* y tres esquemas de contaminación con distintos porcentajes de outliers cada uno, dados por

$C_{1,m}^*$: en los cuales m observaciones fueron reemplazadas por los valores (y_0, x_{01}, x_{02}) , con $x_{01} = 11, x_{02} = 0$ e $y_0 = 0,9(\theta_1 + \theta_2 x_{02} + \theta_3 \exp(\theta_0 x_{01}))$, para $m = 10, 15, 20$ y 25 .

$C_{2,m}^*$: en los cuales m observaciones fueron reemplazadas por los valores (y_0, x_{01}, x_{02}) , con $x_{01} = 11, x_{02} = 0$ e $y_0 = 0,8(\theta_1 + \theta_2 x_{02} + \theta_3 \exp(\theta_0 x_{01}))$, para $m = 10, 15, 20$ y 25 .

$C_{3,m}^*$: en los cuales m observaciones fueron reemplazadas por los valores (y_0, x_{01}, x_{02}) , con $x_{01} = 11, x_{02} = 0$ e $y_0 = 1,1(\theta_1 + \theta_2 x_{02} + \theta_3 \exp(\theta_0 x_{01}))$, para $m = 10, 15, 20$ y 25 .

En cada esquema de contaminación, cuando $m = 10, 15, 20$ y 25 , la proporción de datos atípicos es de $0,10, 0,15, 0,20$ y $0,25$, respectivamente.

En la Figura 5.1 podemos ver una de las muestras que generamos contaminada con cada una de las configuraciones recién mencionadas para el caso en que $m = 10$. En el panel de la izquierda marcamos en color rojo a las observaciones atípicas dadas por $C_{1,m}^*$ y se puede ver como afectan a la curvatura del gráfico correspondiente. Mientras que en los gráficos del centro y de la derecha representamos en verde y en azul a los puntos atípicos definidos por $C_{2,m}^*$ y $C_{3,m}^*$, respectivamente. Como veremos, estos datos anómalos tienen un severo impacto en la estimación por el método de mínimos cuadrados y resultan ser las contaminaciones más relevantes de las que hemos probado.

Con el objetivo de comparar el desempeño estadístico de los estimadores propuestos en términos de robustez y eficiencia, computamos, al igual que para el modelo exponencial, el estimador de mínimos cuadrados, mc , y las alternativas robustas m, mm y wmm . Para todos los estimadores robustos considerados para el modelo lineal–exponencial utilizamos las mismas funciones de pérdida y de peso que las usadas para el primer modelo. Las constantes elegidas a partir del cálculo de las eficiencias asintóticas fueron las siguientes

Estimadores	Eficiencia	c_w	c
M–MM	85 %	-	3,4
M–MM	90 %	-	3,9
WMM	85 %	$\chi_{2,0,975}^2$	4
WMM	90 %	$\chi_{2,0,975}^2$	4,75

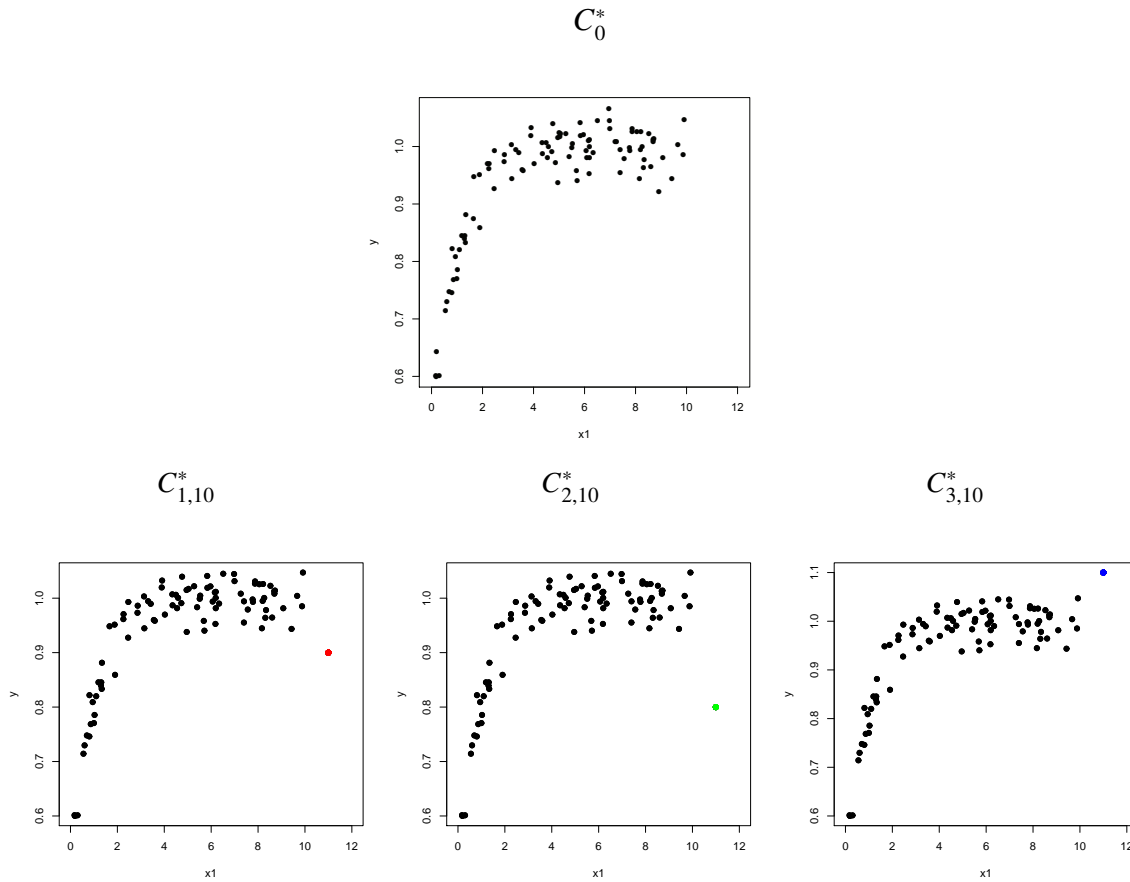


Figura 5.1: Una muestra de los datos simulados: Los puntos coloreados en cada uno de los gráficos inferiores corresponden a los distintos tipos de observaciones atípicas.

y como estimador inicial para el cómputo de los estimadores robustos se utilizaron los S -estimadores definidos en (2.9).

Debido a que la cantidad de parámetros es más alta que en ejemplo anterior, para comparar el comportamiento de los distintos estimadores, solo mostraremos como medida resumen la suma de los errores cuadráticos medios (ECM), es decir

$$ECM = \frac{1}{NR} \sum_{j=1}^{NR} \left((\widehat{\theta}_{0,j} - \theta_0)^2 + (\widehat{\theta}_{1,j} - \theta_1)^2 + (\widehat{\theta}_{2,j} - \theta_2)^2 + (\widehat{\theta}_{3,j} - \theta_3)^2 \right).$$

En el Cuadro 5.3 mostramos los errores cuadráticos medios del estimador de mínimos cuadrados bajo C_0^* y las distintas contaminaciones $C_{i,m}^*$, $i = 1, 2$ y 3 , para $m = 10, 15, 20$ y 25 . Como vemos el ECM llega a crecer hasta seis u ocho veces bajo los esquemas de contaminación $C_{1,m}^*$ y $C_{3,m}^*$ y más aún bajo $C_{2,m}^*$, mostrando la severidad de los outliers introducidos.

Respecto de los estimadores robustos, como es bien sabido, estos pueden tener colas pesadas cuando hay datos atípicos en la muestra y este parece ser el caso del MM -estimador.

<i>ECM</i>					
C_0^*	0,0062	C_0^*	0,0062	C_0^*	0,0062
$C_{1,10}^*$	0,0346	$C_{2,10}^*$	0,1207	$C_{3,10}^*$	0,0337
$C_{1,15}^*$	0,0414	$C_{2,15}^*$	0,1464	$C_{3,10}^*$	0,0402
$C_{1,20}^*$	0,0458	$C_{2,20}^*$	0,1628	$C_{3,10}^*$	0,0441
$C_{1,25}^*$	0,0493	$C_{2,25}^*$	0,1748	$C_{3,10}^*$	0,0474

Cuadro 5.3: Errores cuadráticos medios del estimador de mínimos cuadrados bajo C_0^* , $C_{1,m}^*$, $C_{2,m}^*$ y $C_{3,m}^*$ para $m = 10, 15, 20$ y 25 para muestras completas.

De hecho este fenómeno fue reportado por Fasano (2009) en su estudio de simulación en un escenario semejante. Por esta razón, calculamos para evaluar los casos con contaminación la suma de los errores cuadráticos medios podada al 10 % superior (ECM_p), esto es, si dado un estimador $\widehat{\theta}$ y $\widehat{\theta}_{ecm}$ es el vector de coordenadas $\widehat{\theta}_{ecm,j} = (\widehat{\theta}_{0,j} - \theta_0)^2 + (\widehat{\theta}_{1,j} - \theta_1)^2 + (\widehat{\theta}_{2,j} - \theta_2)^2 + (\widehat{\theta}_{3,j} - \theta_3)^2$, con $j = 1, \dots, NR$,

$$ECM_p = \sum_{j=1}^{0,9NR} \widehat{\theta}_{ecm,(j)},$$

donde $\widehat{\theta}_{ecm,(j)}$ es el j -ésimo estadístico de orden. Estos valores se encuentran en el Cuadro 5.4. A partir de este cuadro podemos concluir que el efecto de las contaminaciones $C_{1,m}^*$ y $C_{3,m}^*$ es semejante sobre todos los estimadores robustos considerados. En él podemos ver que el error cuadrático medio podado del *WMM*-estimador es muy estable en todos los escenarios considerados. En general, bajo las contaminaciones $C_{2,m}^*$ el estimador propuesto tiene errores cuadráticos más bajos que bajo los otros dos esquemas de contaminación. Para los otros estimadores robustos computados los errores cuadráticos sufren un incremento, de hecho con tan solo el 10 % de valores atípicos presentes en la muestra hay un incremento de más de 4 veces para el caso en que no hay pérdida de las respuestas. Esto revela que tanto el *M*- como el *MM*-estimador se vuelven muy inestables con las contaminaciones introducidas, mientras que el *WMM*-estimador logra enfrentar los datos anómalos con los que se contaminó. Cabe observar que en el Cuadro 5.4 se pueden ver valores que resultan inferiores a los reportados para el estimador clásico, cuando la muestra está sin contaminar (C_0^*), en el Cuadro 5.3, esto se debe a que los del primer cuadro están podados y los del segundo no.

En segunda término, hemos introducido un mecanismo de pérdida de respuestas siguiendo un esquema logístico dado por $p_{\text{LOG}}(\mathbf{x}) = (1 + \exp(-0,15x_1 - 0,2))^{-1}$, obteniendo aproximadamente una media de 30 % de respuestas faltantes. Dado que, de acuerdo al Cuadro 5.4, el efecto que las contaminaciones de tipo $C_{1,m}^*$ y $C_{3,m}^*$ tienen sobre nuestro estimador es similar y, a la vez, son de las que producen un mayor incremento en el error cuadrático medio del mismo, hemos considerado el tipo $C_{1,m}^*$ como representativo de lo que ocurre bajo contaminación en ausencia de algunas respuestas. En el Cuadro 5.5, donde se encuentran los errores cuadráticos medios del estimador de mínimos cuadrados bajo $C_{1,m}^*$ para $m = 10, 15, 20$ y 25 , podemos observar que el efecto de los outliers introducidos sobre el estimador clásico es severo, ya que el *ECM* aumenta entre cuatro y seis veces.

		85 %			90 %		
p_{COMP}	Eficiencia	M	MM	WMM	M	MM	WMM
ECM	C_0^*	0,0100	0,0098	0,0095	0,0079	0,0076	0,0084
ECM_p	$C_{1,10}^*$	0,0385	0,0384	0,0063	0,0372	0,0372	0,0058
	$C_{1,15}^*$	0,0498	0,0499	0,0068	0,0468	0,0468	0,0061
	$C_{1,20}^*$	0,0575	0,0573	0,0080	0,0532	0,0531	0,0069
	$C_{1,25}^*$	0,0680	0,0677	0,0118	0,0619	0,0618	0,0091
		M	MM	WMM	M	MM	WMM
ECM_p	$C_{2,10}^*$	0,0069	0,0069	0,0059	0,0064	0,0064	0,0056
	$C_{2,15}^*$	0,0769	0,0770	0,0058	0,0737	0,0737	0,0055
	$C_{2,20}^*$	0,1822	0,1827	0,0064	0,1747	0,1748	0,0060
	$C_{2,25}^*$	0,2205	0,2204	0,0076	0,2079	0,2078	0,0069
		M	MM	WMM	M	MM	WMM
ECM_p	$C_{3,10}^*$	0,0388	0,0388	0,0063	0,0372	0,0373	0,0058
	$C_{3,15}^*$	0,0493	0,0494	0,0068	0,0462	0,0463	0,0061
	$C_{3,20}^*$	0,0557	0,0555	0,0082	0,0521	0,0520	0,0069
	$C_{3,25}^*$	0,0661	0,0661	0,0123	0,0602	0,0602	0,0096

Cuadro 5.4: Errores cuadráticos medios para C_0^* y Errores cuadráticos medios podados para $C_{i,m}^*$, $i = 1, 2$ y 3 , para $m = 10, 15, 20$ y 25 para muestras completas.

p_{LOG}	ECM
C_0^*	0,0094
$C_{1,10}^*$	0,0381
$C_{1,15}^*$	0,0460
$C_{1,20}^*$	0,0511
$C_{1,25}^*$	0,0555

Cuadro 5.5: Errores cuadráticos medios del estimador de mínimos cuadrados bajo C_0^* y $C_{1,m}^*$ para $m = 10, 15, 20$ y 25 bajo p_{LOG} .

Finalmente, en el Cuadro 5.6 exhibimos los errores cuadráticos medios podados para el caso en que las respuestas están perdidas según el esquema logístico. En el mismo podemos apreciar, por un lado, la pérdida de eficiencia de los estimadores robustos bajo el modelo central, que es muy similar entre todos ellos y por el otro, la mayor estabilidad del estimador propuesto respecto de los otros estimadores robustos.

	Eficiencia	85 %			90 %		
p_{LOG}		M	MM	WMM	M	MM	WMM
ECM	C_0^*	0,0222	0,0221	0,0207	0,0192	0,0191	0,0164
ECM_p	$C_{1,10}^*$	0,0430	0,0429	0,0101	0,0405	0,0404	0,0090
	$C_{1,15}^*$	0,0578	0,0578	0,0116	0,0526	0,0525	0,0101
	$C_{1,20}^*$	0,0703	0,0699	0,0146	0,0641	0,0641	0,0117
	$C_{1,25}^*$	0,0891	0,0889	0,0241	0,0813	0,0810	0,0177

Cuadro 5.6: Errores cuadráticos medios para C_0^* y Errores cuadráticos medios podados para $C_{1,m}^*$ para $m = 10, 15, 20$ y 25 .

Capítulo 6

Tests de Hipótesis

La estimación puntual suele ser una de las primeras etapas de un análisis estadístico, generalmente es seguida por una etapa de inferencia en la que se calculan intervalos de confianza y tests de hipótesis. Markatou y Manos (1996) y Liu *et al.* (2005) consideran tests de hipótesis robustos en modelos no lineales basados en M -estimadores o M -estimadores generales. Sin embargo, el problema de testear hipótesis mediante procedimientos robustos, en general, ha recibido menos atención que el de estimación robusta y el caso de los modelos no lineales no es una excepción. Por este motivo, en esta segunda parte nos concentramos en el problema de testear robustamente en un modelo no lineal, en las condiciones descritas en la Sección 2.1, hipótesis que involucren al parámetro de regresión. Nos enfocaremos en testear hipótesis de la forma

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_o \quad \text{vs.} \quad H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_o \quad (6.1)$$

o subhipótesis que involucren un subconjunto de coordenadas del vector de parámetros desconocido, mediante un estadístico de tipo Wald basado en $\widehat{\boldsymbol{\theta}}$, el estimador pesado propuesto. En el Capítulo 3 probamos que si se cumple el modelo (2.1), bajo ciertas condiciones de regularidad, el estimador propuesto es asintóticamente normal, más precisamente

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} N(\mathbf{0}, \sigma^2 \boldsymbol{\Sigma}(\boldsymbol{\theta})),$$

donde

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \mathbf{A}^{-1}(\boldsymbol{\theta})\mathbf{B}(\boldsymbol{\theta})\mathbf{A}^{-1}(\boldsymbol{\theta}), \quad (6.2)$$

$$\mathbf{A}(\boldsymbol{\theta}) = \mathbb{E}\psi'_1(\epsilon)\mathbb{E}\left[p(\mathbf{x})\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t w(\mathbf{x}, \boldsymbol{\Gamma})\right], \quad (6.3)$$

$$\mathbf{B}(\boldsymbol{\theta}) = \mathbb{E}\psi_1^2(\epsilon)\mathbb{E}\left[p(\mathbf{x})\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})^t w^2(\mathbf{x}, \boldsymbol{\Gamma})\right]. \quad (6.4)$$

Para poder construir el estadístico de un test de tipo Wald necesitamos estimar la matriz de covarianza de la distribución asintótica. A la escala σ la estimaremos con $\widehat{\sigma}$ una M -escala computada a partir de la ecuación (2.13) y basada en el estimador $\widehat{\boldsymbol{\theta}}$ propuesto. A las matrices $\mathbf{A}(\boldsymbol{\theta})$ y $\mathbf{B}(\boldsymbol{\theta})$ las estimaremos usando sus versiones muestrales $\widehat{\mathbf{A}} = \widehat{\mathbf{A}}(\widehat{\boldsymbol{\theta}})$ y

$\widehat{\mathbf{B}} = \widehat{\mathbf{B}}(\widehat{\boldsymbol{\theta}})$, donde

$$\widehat{\mathbf{A}}(\widehat{\boldsymbol{\theta}}) = \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1' \left(\frac{y_i - f(\mathbf{x}_i, \widehat{\boldsymbol{\theta}})}{\widehat{\sigma}} \right) \mathbf{f}(\mathbf{x}_i, \widehat{\boldsymbol{\theta}}) \mathbf{f}(\mathbf{x}_i, \widehat{\boldsymbol{\theta}})^t w(\mathbf{x}_i, \widehat{\boldsymbol{\Gamma}}), \quad (6.5)$$

$$\widehat{\mathbf{B}}(\widehat{\boldsymbol{\theta}}) = \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1^2 \left(\frac{y_i - f(\mathbf{x}_i, \widehat{\boldsymbol{\theta}})}{\widehat{\sigma}} \right) \mathbf{f}(\mathbf{x}_i, \widehat{\boldsymbol{\theta}}) \mathbf{f}(\mathbf{x}_i, \widehat{\boldsymbol{\theta}})^t w^2(\mathbf{x}_i, \widehat{\boldsymbol{\Gamma}}), \quad (6.6)$$

siendo $\widehat{\boldsymbol{\Gamma}}$ un estimador previo del parámetro auxiliar $\boldsymbol{\Gamma}$. El Lema 3.4.1 establece que para cada $\boldsymbol{\theta}$ fijo, las matrices $\widehat{\mathbf{A}}$ y $\widehat{\mathbf{B}}$ convergen en probabilidad a $\mathbf{A}(\boldsymbol{\theta})$ y $\mathbf{B}(\boldsymbol{\theta})$, respectivamente. Luego, para testear las hipótesis dadas en (6.1) proponemos el siguiente estadístico de tipo Wald basado en la forma cuadrática

$$\widehat{\mathcal{W}}_n = \frac{n}{\widehat{\sigma}^2} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o)^t \widehat{\boldsymbol{\Sigma}}^{-1} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o), \quad (6.7)$$

donde $\widehat{\boldsymbol{\Sigma}} = \widehat{\mathbf{A}}^{-1} \widehat{\mathbf{B}} \widehat{\mathbf{A}}^{-1}$.

6.1. Comportamiento asintótico

6.1.1. Bajo Hipótesis Fijas

Como la distribución del estadístico propuesto está basada en la del *WMM*-estimador del parámetro de regresión, utilizaremos las hipótesis enumeradas en la Sección 3.4.

Teorema 6.1.1 *Supongamos que valen N1–N10 y que $\widehat{\sigma} \xrightarrow{p} \sigma$ y $\widehat{\boldsymbol{\Gamma}} \xrightarrow{p} \boldsymbol{\Gamma}$. Si $\widehat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}_o$, entonces tenemos que*

i) bajo $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_o$

$$\widehat{\mathcal{W}}_n \xrightarrow{\mathcal{D}} \chi_p^2,$$

ii) bajo $H_{1,\boldsymbol{\theta}} : \boldsymbol{\theta} \neq \boldsymbol{\theta}_o$

$$\widehat{\mathcal{W}}_n \xrightarrow{p} \infty,$$

donde el estadístico $\widehat{\mathcal{W}}_n$ fue definido en (6.7).

Demostración: i) La distribución asintótica del estadístico basado en el estimador propuesto es una consecuencia directa del Teorema 3.4.2. Dado que, como valen N2, N4 y N6–N8 podemos aplicar el Lema 3.4.2 para decir que $\widehat{\mathbf{A}} \xrightarrow{p} \mathbf{A}(\boldsymbol{\theta}_o)$. Luego, como además vale N1 podemos aplicar el Lema 3.4.1 para concluir que $\widehat{\mathbf{B}} \xrightarrow{p} \mathbf{B}(\boldsymbol{\theta}_o)$. Por otro lado, tenemos que

$$\begin{aligned} \widehat{\mathcal{W}}_n &= \frac{n}{\widehat{\sigma}^2} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o)^t \widehat{\boldsymbol{\Sigma}}^{-1} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o) \\ &= \frac{n}{\sigma^2} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o)^t \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_o) (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o) + \sqrt{n} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o)^t \left(\frac{\widehat{\boldsymbol{\Sigma}}^{-1}}{\widehat{\sigma}^2} - \frac{\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_o)}{\sigma^2} \right) \sqrt{n} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o). \end{aligned}$$

Finalmente, como $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o) \xrightarrow{D} N(\mathbf{0}, \sigma^2 \boldsymbol{\Sigma}(\boldsymbol{\theta}_o))$, $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o) = O_p(1)$ y como además $\widehat{\boldsymbol{\Sigma}}$ y $\widehat{\sigma}$ son estimadores consistentes de $\boldsymbol{\Sigma}(\boldsymbol{\theta}_o)$ y σ , aplicando el Lema de Slutsky obtenemos el resultado deseado.

ii) Dado que

$$\widehat{\mathcal{W}}_n = \frac{n}{\widehat{\sigma}^2} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^t \widehat{\boldsymbol{\Sigma}}^{-1} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + \frac{n}{\widehat{\sigma}^2} (\boldsymbol{\theta} - \boldsymbol{\theta}_o)^t \widehat{\boldsymbol{\Sigma}}^{-1} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o + \widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}),$$

por el Teorema 3.4.2, la consistencia de $\widehat{\boldsymbol{\Sigma}}$ y $\widehat{\sigma}$ y el hecho de que $\widehat{\boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\theta}$, tenemos que $(\boldsymbol{\theta} - \boldsymbol{\theta}_o)^t \widehat{\boldsymbol{\Sigma}}^{-1} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o - \widehat{\boldsymbol{\theta}} + \boldsymbol{\theta}) \xrightarrow{p} (\boldsymbol{\theta} - \boldsymbol{\theta}_o)^t \boldsymbol{\Sigma}^{-1} (\boldsymbol{\theta} - \boldsymbol{\theta}_o) > 0$ y dado que $(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^t \widehat{\boldsymbol{\Sigma}}^{-1} (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) > 0$ con probabilidad tendiendo a 1, se obtiene el resultado. \square

Por el resultado anterior, tenemos que para chequear las hipótesis

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_o \quad \text{vs.} \quad H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_o, \quad (6.8)$$

el test de tipo Wald que rechaza si

$$\widehat{\mathcal{W}}_n > \chi_{p,\alpha}^2$$

es un test consistente de nivel asintótico α .

6.1.2. Bajo Hipótesis Alternativas Contiguas

Otro comportamiento de relevancia sobre los tests de hipótesis es su distribución bajo hipótesis alternativas contiguas. Estas alternativas son hipótesis que dependen de n y que tienden a la hipótesis nula cuando n tiende a infinito. Lo que interesa estudiar es la distribución límite del estadístico propuesto.

Formalmente, las hipótesis alternativas contiguas que estudiaremos se definen como $H_{1,n} : \boldsymbol{\theta} = \boldsymbol{\theta}_o + \mathbf{c}n^{-1/2}$, donde $\mathbf{c} \neq \mathbf{0}$.

Antes de deducir la distribución asintótica del estadístico del test probaremos algunos resultados y como consecuencia de ellos obtendremos la distribución deseada.

En primer lugar deduciremos la distribución asintótica del estimador robusto propuesto bajo hipótesis alternativas contiguas como así también probaremos la convergencia de las matrices estimadas que forman parte del estadístico del test. Para ello enunciaremos un lema auxiliar, que es una generalización del Lema 3.4.1 y para la cual necesitamos considerar la siguiente hipótesis.

L4. La función $u(\mathbf{x}, \mathbf{b})$ es equicontinua en el sentido en que dado $\varepsilon > 0$, existe $\tau > 0$ tal que si $|\mathbf{b}_1 - \mathbf{b}_2| \leq \tau$, $|\mathbf{b}_i - \boldsymbol{\theta}| \leq \tau$, $i = 1, 2$ y $\mathbf{x} \in \mathcal{X}_K$, entonces $|u(\mathbf{x}, \mathbf{b}_1) - u(\mathbf{x}, \mathbf{b}_2)| < \varepsilon$, y además $\sup_{\mathbf{x} \in \mathcal{X}_K} |u(\mathbf{x}, \boldsymbol{\theta})| < \infty$, para todo $K > 0$.

Lema 6.1.1 Sean $(\epsilon_i, \mathbf{x}_i^t, \delta_i)$, $1 \leq i \leq n$ vectores aleatorios con distribución G_* tales que δ_i son variables aleatorias dicotómicas. Supongamos que valen **N4** y **L1-L4** y que tenemos estimadores $\widehat{\sigma}$, $\widehat{\Gamma}$ y $\widetilde{\theta}$ tales que $\widehat{\sigma} \xrightarrow{p} \sigma$, $\widehat{\Gamma} \xrightarrow{p} \Gamma$ y $\widetilde{\theta} \xrightarrow{p} \theta$. Entonces, $\mathbf{D}_n \xrightarrow{p} \mathbf{D}$, donde, dado $a \in \mathbb{R}$,

$$\mathbf{D}_n = \frac{1}{n} \sum_{i=1}^n \delta_i \varphi \left(\frac{a\epsilon_i + u(\mathbf{x}_i, \widetilde{\theta})}{\widehat{\sigma}} \right) \mathbf{M}(\mathbf{x}_i, \widetilde{\theta}) w(\mathbf{x}_i, \widehat{\Gamma}) \quad \text{y} \quad \mathbf{D} = \mathbb{E} \left[\delta \varphi \left(\frac{a\epsilon + u(\mathbf{x}, \theta)}{\sigma} \right) \mathbf{M}(\mathbf{x}, \theta) w(\mathbf{x}, \Gamma) \right].$$

Demostración: Consideremos el vector aleatorio $(\epsilon, \mathbf{x}^t, \delta)$ con distribución G_* .

Es suficiente probar que

$$\text{i) } \lim_{(\mathbf{b}, s, \mathbf{G}) \rightarrow (\theta, \sigma, \Gamma)} \mathbb{E} \left[\delta \varphi \left(\frac{a\epsilon + u(\mathbf{x}, \mathbf{b})}{s} \right) \mathbf{M}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G}) \right] = \mathbf{D},$$

ii)

$$\sup_{\mathcal{B}, s, \mathcal{G}} \left| \frac{1}{n} \sum_{i=1}^n \mathbb{I}(\epsilon_i, \mathbf{x}_i, \delta_i, \mathbf{b}, s, \mathbf{G}) - \mathbb{E} [\mathbb{I}(\epsilon, \mathbf{x}, \delta, \mathbf{b}, s, \mathbf{G})] \right| \xrightarrow{p} \mathbf{0},$$

$$\text{con } \mathbb{I}(\epsilon, \mathbf{x}, \delta, \mathbf{b}, s, \mathbf{G}) = \delta \varphi \left(\frac{a\epsilon + u(\mathbf{x}, \mathbf{b})}{s} \right) \mathbf{M}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G}).$$

El ítem i) Se deduce de **L1-L3** y del Teorema de Convergencia Dominada ya que \mathbb{I} converge puntualmente por **N4** y **L4**, la función φ es continua y se tiene que $|\mathbb{I}| \leq \|\varphi\|_\infty \sup_{\mathcal{B}, \mathcal{G}} |\mathbf{M}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G})|$, que es integrable.

Para probar ii), como \mathbb{I} es una matriz, probaremos la convergencia de cada elemento.

Por el Teorema 3.2.1, basta con probar que dado $\eta > 0$ existe una clase finita de funciones \mathcal{I}_η tal que para todo $\mathbf{b} \in \mathcal{B}$, $s \in \mathcal{S}$ y $\mathbf{G} \in \mathcal{G}$ existe $\mathbb{I}_{\eta, U}^{kl} \in \mathcal{I}_\eta$ y $\mathbb{I}_{\eta, L}^{kl} \in \mathcal{I}_\eta$ tales que

$$\mathbb{I}_{\eta, L}^{kl}(\epsilon, \mathbf{x}) \leq \mathbb{I}^{kl}(\epsilon, \mathbf{x}, \delta, \mathbf{b}, s, \mathbf{G}) \leq \mathbb{I}_{\eta, U}^{kl}(\epsilon, \mathbf{x}) \quad (6.9)$$

$$\mathbb{E} \left(\mathbb{I}_{\eta, U}^{kl}(\epsilon, \mathbf{x}) - \mathbb{I}_{\eta, L}^{kl}(\epsilon, \mathbf{x}) \right) < \eta. \quad (6.10)$$

Dado $K > 0$, por **L4**, consideremos $C_1 = \sup_{\mathcal{X}_K, \mathcal{B}} |u(\mathbf{x}, \mathbf{b})|$.

Definamos $A_K = \{|\epsilon| \leq K, |\mathbf{x}| \leq K\}$ y tomemos $K \in \mathbb{N}$ de forma tal que dado $\eta_1 = \eta/5\|\psi'\|_\infty$, $P(A_K) > 1 - \eta_1$, y, como vale **L2**, tal que

$$\mathbb{E} \left(\sup_{\mathcal{B}, \mathcal{G}} |\mathbf{M}(\mathbf{x}, \mathbf{b}) w(\mathbf{x}, \mathbf{G})| I_{A_K^c}(\epsilon, \mathbf{x}) \right) < \eta_1. \quad (6.11)$$

Por otro lado, sea c tal que para todo $\mathbf{b} \in \mathcal{B}$ y $s \in \mathcal{S}$, $|\mathbf{b}| < c$ y $c^{-1} < s < c$. Luego, si $(\epsilon, \mathbf{x}^t) \in A_K$, $\mathbf{b} \in \mathcal{B}$ y $s \in \mathcal{S}$, entonces

$$\left| \frac{a\epsilon + u(\mathbf{x}, \mathbf{b})}{s} \right| \leq \frac{a|\epsilon| + |u(\mathbf{x}, \mathbf{b})|}{s} \leq (aK + C_1)c = b_1.$$

De **L3** existe $b_2 \geq 0$ tal que $|\mathbf{M}(\mathbf{x}, \mathbf{b})w(\mathbf{x}, \mathbf{G})| \leq b_2$ para cada $\mathbf{x} \in A_K$, $\mathbf{b} \in \mathcal{B}$ y $\mathbf{G} \in \mathcal{G}$. Por **L1** tenemos que $\varphi(t)\mathbf{z}$ es uniformemente continua en $B_K = \{|t| \leq b_1, |\mathbf{z}| \leq b_2\}$ y por lo tanto, existe $\tau > 0$ tal que $|\varphi(t)\mathbf{z} - \varphi(u)\mathbf{v}| < \eta/10$, para $|t - u| < \tau$, $|\mathbf{z} - \mathbf{v}| < \tau$, $(t, \mathbf{z}^t) \in B_K$ y $(u, \mathbf{v}^t) \in B_K$.

Como valen **L4** y **L3** resulta que si $(\epsilon, \mathbf{x}^t) \in A_K$, entonces $(a\epsilon + u(\mathbf{x}, \mathbf{b}))/s$ y $\mathbf{M}(\mathbf{x}, \mathbf{b})w(\mathbf{x}, \mathbf{G})$ son funciones equicontinuas de $(\mathbf{b}, s, \mathbf{G}) \in \mathcal{B} \times \mathcal{S} \times \mathcal{G}$, es decir que existe $\varepsilon_1 > 0$ tal que si $|\mathbf{b} - \tilde{\mathbf{b}}| < \varepsilon_1$, $|s - \tilde{s}| < \varepsilon_1$, $|\mathbf{G} - \tilde{\mathbf{G}}| < \varepsilon_1$, $(\mathbf{b}, s, \mathbf{G}), (\tilde{\mathbf{b}}, \tilde{s}, \tilde{\mathbf{G}}) \in \mathcal{B} \times \mathcal{S} \times \mathcal{G}$ y $(\epsilon, \mathbf{x}^t) \in A_K$,

$$\left| \frac{a\epsilon + u(\mathbf{x}, \mathbf{b})}{s} - \frac{a\epsilon + u(\mathbf{x}, \tilde{\mathbf{b}})}{\tilde{s}} \right| < \tau$$

$$\left| \mathbf{M}(\mathbf{x}, \mathbf{b})w(\mathbf{x}, \mathbf{G}) - \mathbf{M}(\mathbf{x}, \tilde{\mathbf{b}})w(\mathbf{x}, \tilde{\mathbf{G}}) \right| < \tau.$$

Sean ahora $(\mathcal{B}_i)_{1 \leq i \leq N_1}$ una colección finita de bolas centradas en $\theta_i \in \mathcal{B}$ de radio menor que ε_1 tal que $\mathcal{B} = \bigcup_{i=1}^{N_1} \mathcal{B}_i$, $(\mathcal{S}_m)_{1 \leq m \leq N_2}$ otra colección finita de bolas centradas en $\sigma_m \in \mathcal{S}$ de radio menor que ε_1 tal que $\mathcal{S} = \bigcup_{m=1}^{N_2} \mathcal{S}_m$ y $(\mathcal{G}_r)_{1 \leq r \leq N_3}$ otra colección finita de bolas centradas en $\Gamma_r \in \mathcal{G}$ también de radio menor que ε_1 tal que $\mathcal{G} = \bigcup_{r=1}^{N_3} \mathcal{G}_r$. Por lo que dados $\mathbf{b} \in \mathcal{B}$, $s \in \mathcal{S}$ y $\mathbf{G} \in \mathcal{G}$ existe una tira de índices (i, m, r) tal que $\mathbf{b} \in \mathcal{B}_i$, $s \in \mathcal{S}_m$ y $\mathbf{G} \in \mathcal{G}_r$ y

$$\left| \mathbb{I}^{kl}(\epsilon, \mathbf{x}, \delta, \mathbf{b}, s, \mathbf{G}) - \mathbb{I}^{kl}(\epsilon, \mathbf{x}, \delta, \theta_i, \sigma_m, \Gamma_r) \right| \leq \frac{\eta}{10} + I_{A_K^c}(\epsilon, \mathbf{x}) \|\varphi\|_{\infty} 2 \sup_{\mathcal{B}, \mathcal{G}} |\mathbf{M}(\mathbf{x}, \mathbf{b})w(\mathbf{x}, \mathbf{G})|.$$

Así, si definimos $e(\epsilon, \mathbf{x}) = \left(\frac{\eta}{10} + 2\|\varphi\|_{\infty} I_{A_K^c}(\epsilon, \mathbf{x}) \sup_{\mathcal{B}, \mathcal{G}} |\mathbf{M}(\mathbf{x}, \mathbf{b})w(\mathbf{x}, \mathbf{G})| \right)$ y consideramos

$$\mathbb{I}_{\eta, L}^{kl}(\epsilon, \mathbf{x}) = \mathbb{I}^{kl}(\epsilon, \mathbf{x}, \delta, \theta_i, \sigma_m, \Gamma_r) - e(\epsilon, \mathbf{x}) \quad \text{y} \quad \mathbb{I}_{\eta, U}^{kl}(\epsilon, \mathbf{x}) = \mathbb{I}^{kl}(\epsilon, \mathbf{x}, \delta, \theta_i, \sigma_m, \Gamma_r) + e(\epsilon, \mathbf{x})$$

obtenemos (6.9). Por otro lado,

$$\mathbb{E} \left(\mathbb{I}_{\eta, U}^{kl}(\epsilon, \mathbf{x}) - \mathbb{I}_{\eta, L}^{kl}(\epsilon, \mathbf{x}) \right) = \frac{\eta}{5} + 4\|\varphi\|_{\infty} \mathbb{E} \left(I_{A_K^c}(\epsilon, \mathbf{x}) \sup_{\mathcal{B}, \mathcal{G}} |\mathbf{M}(\mathbf{x}, \mathbf{b})w(\mathbf{x}, \mathbf{G})| \right),$$

entonces, usando (6.11), probamos (6.10) y podemos concluir que $\mathbf{D}_n \xrightarrow{p} \mathbf{D}$. \square

El siguiente teorema establece la distribución asintótica del estimador robusto propuesto bajo hipótesis alternativas contiguas.

Teorema 6.1.2 Sean $(y_i, \mathbf{x}_i^t, \delta_i)$ con $i = 1, \dots, n$ vectores aleatorios que siguen el modelo dado por $H_{1,n}$, es decir,

$$y_i = f(\mathbf{x}_i, \boldsymbol{\theta}_n) + \sigma\epsilon, \quad \text{con} \quad \boldsymbol{\theta}_n = \boldsymbol{\theta}_o + \mathbf{c} \frac{1}{\sqrt{n}}, \quad (6.12)$$

y que la probabilidad de la función de distribución de δ_i cumple la condición MAR dada en (2.2). Si además valen $N1-N10$ y $\widehat{\sigma} \xrightarrow{p} \sigma$ y $\widehat{\Gamma} \xrightarrow{p} \Gamma$, entonces bajo $H_{1,n}$

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o) \xrightarrow{D} N(\mathbf{c}, \sigma^2 \boldsymbol{\Sigma}(\boldsymbol{\theta}_o)),$$

donde $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ fue definida en (6.2).

Demostración: Sumando y restando $\boldsymbol{\theta}_n$ obtenemos que

$$\begin{aligned} \sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_o) &= \sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_n + \boldsymbol{\theta}_n - \boldsymbol{\theta}_o) \\ &= \sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_n) + \sqrt{n}(\boldsymbol{\theta}_n - \boldsymbol{\theta}_o) \\ &= \sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_n) + \sqrt{n}\left(\boldsymbol{\theta}_o + \mathbf{c} \frac{1}{\sqrt{n}} - \boldsymbol{\theta}_o\right) \\ &= \sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_n) + \mathbf{c}, \end{aligned}$$

por lo tanto, por el Lema de Slutsky, alcanza con probar que $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_n) \xrightarrow{D} N(\mathbf{0}, \sigma^2 \boldsymbol{\Sigma}(\boldsymbol{\theta}_o))$ bajo $H_{1,n}$.

Tenemos que $\widehat{\boldsymbol{\theta}}$ es solución del siguiente sistema de ecuaciones,

$$\frac{1}{n} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{y_i - f(\mathbf{x}_i, \widehat{\boldsymbol{\theta}})}{\widehat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \widehat{\boldsymbol{\theta}}) w(\mathbf{x}_i, \widehat{\Gamma}) = \mathbf{0},$$

luego de un desarrollo de Taylor alrededor de $\boldsymbol{\theta}_n$ se obtiene que

$$\begin{aligned} \mathbf{0} &= \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{y_i - f(\mathbf{x}_i, \boldsymbol{\theta}_n)}{\widehat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \boldsymbol{\theta}_n) w(\mathbf{x}_i, \widehat{\Gamma}) + \left[\frac{-1}{n} \sum_{i=1}^n \delta_i \psi_1' \left(\frac{y_i - f(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n)}{\widehat{\sigma}} \right) \frac{\dot{\mathbf{f}}(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n)}{\widehat{\sigma}} \dot{\mathbf{f}}(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n)^t w(\mathbf{x}_i, \widehat{\Gamma}) + \right. \\ &\quad \left. + \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{y_i - f(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n)}{\widehat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n) w(\mathbf{x}_i, \widehat{\Gamma}) \right] (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_n), \end{aligned}$$

donde $\widetilde{\boldsymbol{\theta}}_n$ es un punto intermedio entre $\widehat{\boldsymbol{\theta}}$ y $\boldsymbol{\theta}_n$. A partir de esta ecuación podemos despejar $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_n)$ de la siguiente manera

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_n) = \widehat{\sigma} \mathbf{C}_n^{-1} \mathbf{Z}_n,$$

donde

$$\begin{aligned} \mathbf{C}_n &= \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1' \left(\frac{y_i - f(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n)}{\widehat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n) \dot{\mathbf{f}}(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n)^t w(\mathbf{x}_i, \widehat{\Gamma}) - \widehat{\sigma} \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{y_i - f(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n)}{\widehat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n) w(\mathbf{x}_i, \widehat{\Gamma}), \\ \mathbf{Z}_n &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{y_i - f(\mathbf{x}_i, \boldsymbol{\theta}_n)}{\widehat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \boldsymbol{\theta}_n) w(\mathbf{x}_i, \widehat{\Gamma}). \end{aligned}$$

En primer lugar veamos que $\mathbf{C}_n = \mathbf{C}_{n,1} - \widehat{\sigma}\mathbf{C}_{n,2} \xrightarrow{p} \mathbf{A}(\boldsymbol{\theta}_o)$, donde $\mathbf{A}(\boldsymbol{\theta})$ está definida en (6.3). Sumando y restando $f(\mathbf{x}_i, \boldsymbol{\theta}_n)$ en el numerador de los argumentos de ψ_1 y ψ'_1 , tenemos que como la muestra aleatoria satisface el modelo dado en (6.12)

$$\begin{aligned}\mathbf{C}_{n,1} &= \frac{1}{n} \sum_{i=1}^n \delta_i \psi'_1 \left(\frac{\sigma \epsilon_i + f(\mathbf{x}_i, \boldsymbol{\theta}_n) - f(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n)}{\widehat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n) \dot{\mathbf{f}}(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n)^t w(\mathbf{x}_i, \widehat{\boldsymbol{\Gamma}}), \\ \mathbf{C}_{n,2} &= \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{\sigma \epsilon_i + f(\mathbf{x}_i, \boldsymbol{\theta}_n) - f(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n)}{\widehat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \widetilde{\boldsymbol{\theta}}_n) w(\mathbf{x}_i, \widehat{\boldsymbol{\Gamma}}).\end{aligned}$$

Como valen **N2**, **N4** y **N6–N8** tenemos que por el Lema 6.1.1

$$\mathbf{C}_{n,1} \xrightarrow{p} \mathbb{E} \left[\delta \psi'_1(\epsilon) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}_o) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}_o)^t w(\mathbf{x}, \boldsymbol{\Gamma}) \right].$$

Además $(y, \mathbf{x}^t, \delta)$ satisface la condición MAR y el término de error ϵ es independiente de \mathbf{x} , por lo tanto

$$\mathbf{C}_{n,1} \xrightarrow{p} \mathbb{E} \psi'_1(\epsilon) \mathbb{E} \left[p(\mathbf{x}) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}_o) \dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}_o)^t w(\mathbf{x}, \boldsymbol{\Gamma}) \right] = \mathbf{A}(\boldsymbol{\theta}_o).$$

De la misma manera, como además se cumplen los supuestos **N1** y **N9**, y el error ϵ tiene distribución simétrica respecto del 0, se deduce que

$$\mathbf{C}_{n,2} \xrightarrow{p} \mathbf{0}.$$

Para terminar con la demostración del resultado del teorema resta ver que

$$\mathbf{Z}_n \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{B}(\boldsymbol{\theta}_o)),$$

donde $\mathbf{B}(\boldsymbol{\theta})$ está definida en (6.4).

Dado que la muestra sigue el modelo (6.12), $\mathbf{Z}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{\sigma \epsilon_i}{\widehat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \boldsymbol{\theta}_n) w(\mathbf{x}_i, \widehat{\boldsymbol{\Gamma}})$,

haciendo un desarrollo de Taylor alrededor de $\boldsymbol{\theta}_o$ de primer orden podemos escribir a \mathbf{Z}_n como

$$\mathbf{Z}_n = \mathbf{Z}_{n,1} + \mathbf{Z}_{n,2} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{\sigma \epsilon_i}{\widehat{\sigma}} \right) \dot{\mathbf{f}}(\mathbf{x}_i, \boldsymbol{\theta}_o) w(\mathbf{x}_i, \widehat{\boldsymbol{\Gamma}}) + \frac{1}{\sqrt{n}} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{\sigma \epsilon_i}{\widehat{\sigma}} \right) \ddot{\mathbf{f}}(\mathbf{x}_i, \boldsymbol{\theta}_n^*) w(\mathbf{x}_i, \widehat{\boldsymbol{\Gamma}}) (\boldsymbol{\theta}_n - \boldsymbol{\theta}_o),$$

donde $\boldsymbol{\theta}_n^*$ es un punto intermedio entre $\boldsymbol{\theta}_n$ y $\boldsymbol{\theta}_o$.

Como $\boldsymbol{\theta}_n = \boldsymbol{\theta}_o + c n^{-1/2}$, $\mathbf{Z}_{n,2} = \frac{1}{n} \sum_{i=1}^n \delta_i \psi_1 \left(\frac{\sigma \epsilon_i}{\widehat{\sigma}} \right) \ddot{\mathbf{f}}(\mathbf{x}_i, \boldsymbol{\theta}_n^*) w(\mathbf{x}_i, \widehat{\boldsymbol{\Gamma}}) \mathbf{c}$, dado que se satisfacen **N1**, **N4**, **N6** y **N9**, podemos aplicar el Lema 6.1.1 para concluir que

$$\mathbf{Z}_{n,2} \xrightarrow{p} \mathbf{0}.$$

Siguiendo el mismo esquema que se realizó en el Teorema 3.4.2 para probar la expresión (3.10), se prueba que $\mathbf{Z}_{n,1} \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{B}(\boldsymbol{\theta}_o))$. Concluyendo, de esta manera, con la demostración de la convergencia asintótica planteada. \square

Vamos a enunciar el último resultado que necesitamos para obtener la distribución del estadístico construido como una forma cuadrática a partir del estimador propuesto.

Lema 6.1.2 Sea $(y_i, \mathbf{x}_i^t, \delta_i)$ una muestra aleatoria de tamaño n del modelo (6.12), tal que se cumple la condición de respuestas perdidas al azar. Bajo los supuestos **N1**, **N2**, **N4**, **N6–N8** tenemos que bajo $H_{1,n}$

$$\begin{aligned}\widehat{\mathbf{A}} &\xrightarrow{p} \mathbf{A}(\boldsymbol{\theta}_o), \\ \widehat{\mathbf{B}} &\xrightarrow{p} \mathbf{B}(\boldsymbol{\theta}_o),\end{aligned}$$

donde los estimadores $\widehat{\mathbf{A}}$ y $\widehat{\mathbf{B}}$ están definidos en (6.5) y (6.6), respectivamente y $\mathbf{A}(\boldsymbol{\theta})$ y $\mathbf{B}(\boldsymbol{\theta})$ están definidas en (6.3) y (6.4), respectivamente.

Demostración: La conclusión del enunciado del lema se deduce de la aplicación directa del Lema 6.1.1. \square

El siguiente resultado se obtiene como corolario del Teorema 6.1.2 junto con el Lema anterior y la aplicación del Lema de Slutsky obteniendo la distribución asintótica del estadístico propuesto bajo hipótesis alternativas contiguas.

Corolario 6.1.3 Sean $(y_i, \mathbf{x}_i^t, \delta_i)$ con $i = 1, \dots, n$ vectores aleatorios que siguen el modelo dado por $H_{1,n}$, y que la probabilidad de la función de distribución de δ_i cumple la condición **MAR** dada en (2.2). Si además valen **N1–N10**, entonces bajo $H_{1,n}$

$$\widehat{\mathcal{W}}_n \xrightarrow{D} \chi_p^2(\lambda),$$

donde $\lambda = \mathbf{c}^t \boldsymbol{\Sigma}^{-1} \mathbf{c}$.

6.2. Función de Influencia

Como ya hemos mencionado en el Capítulo 4, la función de influencia es una medida de la robustez respecto de un outlier la cual nos permite estudiar la estabilidad local y la eficiencia asintótica de un estimador.

Recordemos que la curva de influencia puede pensarse como la derivada primera del funcional asociado al estimador, es decir que dado un funcional $T(G)$ esta se define como

$$\text{IF}(\mathbf{z}_*, T(G)) = \lim_{\varepsilon \rightarrow 0} \frac{T(G_{\mathbf{z}_*, \varepsilon}) - T(G)}{\varepsilon},$$

donde $G_{\mathbf{z}_*, \varepsilon} = (1 - \varepsilon)G + \varepsilon\Delta_{\mathbf{z}_*}$ y $\Delta_{\mathbf{z}_*}$ denota la masa puntual en el punto $\mathbf{z}_* = (y_*, \mathbf{x}_*^t, \delta_*)$ y representa al modelo contaminado.

Dada cualquier distribución G_1 , sea $\mathbf{V}(G_1)$ un funcional de escala Fisher-consistente en G , es decir $\mathbf{V}(G) = \boldsymbol{\Sigma}$. Llamemos $\boldsymbol{\theta}(G_1)$, $\sigma(G_1)$ y $\boldsymbol{\Gamma}(G_1)$ a los funcionales relacionados a $\widehat{\boldsymbol{\theta}}$, $\widehat{\sigma}$ y $\widehat{\boldsymbol{\Gamma}}$, respectivamente, y asumamos que $\boldsymbol{\theta}(G_1)$, solución de

$$\mathbb{E}_{G_1} \left[p(\mathbf{x}) \psi_1 \left(\frac{y - f(\mathbf{x}, \boldsymbol{\theta}(G_1))}{\sigma(G_1)} \right) \hat{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta}(G_1)) w(\mathbf{x}, \boldsymbol{\Gamma}(G_1)) \right] = \mathbf{0},$$

es un funcional Fisher-consistente en G , o sea, $\theta(G) = \theta$.

Definimos los funcionales asociados a las matrices $\widehat{\mathbf{A}}$ y $\widehat{\mathbf{B}}$ definidas en (6.5) y (6.6), respectivamente, como

$$\mathbf{A}(G_1) = \mathbb{E}_{G_1} \left[p(\mathbf{x}) \psi_1' \left(\frac{y - f(\mathbf{x}, \theta(G_1))}{\sigma(G_1)} \right) \dot{\mathbf{f}}(\mathbf{x}, \theta(G_1)) \dot{\mathbf{f}}(\mathbf{x}, \theta(G_1))^t w(\mathbf{x}, \Gamma(G_1)) \right], \quad (6.13)$$

$$\mathbf{B}(G_1) = \mathbb{E}_{G_1} \left[p(\mathbf{x}) \psi_1^2 \left(\frac{y - f(\mathbf{x}, \theta(G_1))}{\sigma(G_1)} \right) \dot{\mathbf{f}}(\mathbf{x}, \theta(G_1)) \dot{\mathbf{f}}(\mathbf{x}, \theta(G_1))^t w^2(\mathbf{x}, \Gamma(G_1)) \right] \quad (6.14)$$

El funcional del test de tipo Wald relacionado al estadístico usado para testear

$$H_0 : \theta = \theta_0 \quad \text{vs.} \quad H_1 : \theta \neq \theta_0$$

está dado por

$$\mathcal{W}(G_1) = \sigma(G_1)^{-2} (\theta(G_1) - \theta_0)^t \mathbf{V}(G_1)^{-1} (\theta(G_1) - \theta_0).$$

Es sencillo comprobar que bajo H_0 , $\text{IF}(\mathbf{z}_*, \mathcal{W}(G)) = 0$. Luego, con el propósito de obtener una influencia no nula, consideramos, como en Hampel *et al.* (1986, p. 348), la raíz cuadrada del estadístico del test, es decir, $\mathcal{T}(G_1) = \mathcal{W}(G_1)^{1/2}$. Entonces, usando que bajo H_0 , $\mathcal{W}(G) = 0$, tenemos que

$$\text{IF}(\mathbf{z}_*, \mathcal{T}(G)) = \left[\frac{1}{2} \frac{\partial^2 \mathcal{W}(G_{\mathbf{z}_*, \varepsilon})}{\partial \varepsilon^2} \Big|_{\varepsilon=0} \right]^{1/2}.$$

De hecho, aplicando la regla de L'Hôpital dos veces resulta que

$$\begin{aligned} \text{IF}(\mathbf{z}_*, \mathcal{T}(G)) &= \lim_{\varepsilon \rightarrow 0} \frac{\mathcal{T}(G_{\mathbf{z}_*, \varepsilon}) - \mathcal{T}(G)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{\mathcal{T}(G_{\mathbf{z}_*, \varepsilon})}{\varepsilon} \\ &= \left(\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{W}(G_{\mathbf{z}_*, \varepsilon})}{\varepsilon^2} \right)^{1/2} \\ &= \left[\frac{1}{2} \frac{\partial^2 \mathcal{W}(G_{\mathbf{z}_*, \varepsilon})}{\partial \varepsilon^2} \Big|_{\varepsilon=0} \right]^{1/2}. \end{aligned}$$

El siguiente Teorema da la expresión de la función de influencia del funcional del test $\mathcal{T}(G)$ bajo condiciones de regularidad.

Teorema 6.2.1 *Asumamos que son válidas las condiciones de la Proposición 4.2.1 y que existe $\partial^2 \theta(G_{\mathbf{z}_*, \varepsilon}) / \partial \varepsilon^2 \Big|_{\varepsilon=0}$. Sean $\mathbf{V}(G_1)$, $\sigma(G_1)$ y $\Gamma(G_1)$ funcionales tales que $\mathbf{V}(G) = \Sigma$, donde $\Sigma = \mathbf{A}^{-1} \mathbf{B} \mathbf{A}^{-1}$, $\mathbf{A} = \mathbf{A}(G)$ y $\mathbf{B} = \mathbf{B}(G)$ son matrices simétricas definidas en (6.13) y (6.14), respectivamente, $\sigma(G) = \sigma$ y $\Gamma(G) = \Gamma$. Luego, la función de influencia en G del funcional $\mathcal{T}(G_1)$ para testear $H_0 : \theta = \theta_0$ vs. $H_1 : \theta \neq \theta_0$ está dada por*

$$\text{IF}(\mathbf{z}_*, \mathcal{T}(G))^2 = \delta_* \psi_1^2 \left(\frac{y_* - f(\mathbf{x}_*, \theta(G))}{\sigma(G)} \right) w^2(\mathbf{x}_*, \Gamma(G)) \dot{\mathbf{f}}(\mathbf{x}_*, \theta(G))^t \mathbf{B}(G)^{-1} \dot{\mathbf{f}}(\mathbf{x}_*, \theta(G)). \quad (6.15)$$

Demostración: Debemos calcular $\left. \frac{\partial^2 \mathcal{W}(G_{\mathbf{z}_*, \varepsilon})}{\partial \varepsilon^2} \right|_{\varepsilon=0}$.

Usando que bajo la hipótesis nula, $\boldsymbol{\theta}(G) = \boldsymbol{\theta}_0$, se puede comprobar de manera directa que

$$\left. \frac{\partial^2 \mathcal{W}(G_{\mathbf{z}_*, \varepsilon})}{\partial \varepsilon^2} \right|_{\varepsilon=0} = 2\sigma(G_{\mathbf{z}_*, \varepsilon})^{-2} \left(\frac{\partial}{\partial \varepsilon} (\boldsymbol{\theta}(G_{\mathbf{z}_*, \varepsilon}) - \boldsymbol{\theta}_0) \right)^t \mathbf{V}(G_{\mathbf{z}_*, \varepsilon})^{-1} \left(\frac{\partial}{\partial \varepsilon} (\boldsymbol{\theta}(G_{\mathbf{z}_*, \varepsilon}) - \boldsymbol{\theta}_0) \right) \Big|_{\varepsilon=0}$$

y por lo tanto,

$$\text{IF}(\mathbf{z}_*, \mathcal{T}(G))^2 = \sigma^{-2} \text{IF}(\mathbf{z}_*, \boldsymbol{\theta}(G))^t \boldsymbol{\Sigma}^{-1} \text{IF}(\mathbf{z}_*, \boldsymbol{\theta}(G)).$$

Luego, reemplazando la $\text{IF}(\mathbf{z}_*, \boldsymbol{\theta}(G))$ por (4.1), se obtiene la expresión (6.15). \square

Al igual que la $\text{IF}(\mathbf{z}_*, \boldsymbol{\theta}(G))$, la curva de influencia del funcional del test depende de δ_* , variable que puede valer cero si y_* es una respuesta perdida. Por esta razón también consideraremos la esperanza de la IF^2 que mide la influencia esperada condicional de una observación sin importar si la respuesta fue observada o no.

Corolario 6.2.1 *Tomando esperanza a la expresión dada en (6.15) condicional a $\bar{\mathbf{z}}_* = (y_*, \mathbf{x}_*^t)$, nos queda que*

$$E(\text{IF}(\mathbf{z}_*, \mathcal{T}(G))^2) = p(\mathbf{x}_*) \psi_1^2 \left(\frac{y_* - f(\mathbf{x}_*, \boldsymbol{\theta}(G))}{\sigma(G)} \right) w^2(\mathbf{x}_*, \boldsymbol{\Gamma}(G)) \dot{\mathbf{f}}(\mathbf{x}_*, \boldsymbol{\theta}(G))^t \mathbf{B}(G)^{-1} \dot{\mathbf{f}}(\mathbf{x}_*, \boldsymbol{\theta}(G)).$$

Cabe destacar que con una elección adecuada de las funciones ψ_1 y w tanto la IF^2 definida en (6.15) como su esperanza resultarán acotadas.

6.2.1. Ejemplo Numérico

Ilustraremos el comportamiento de la función de influencia esperada para los tests de tipo Wald cuando \widehat{W} se basa en los estimadores pesados (wmm) y los estimadores clásico (mc) y MM (mm) en el caso del modelo Lineal–exponencial tratado en la Sección 5.2.

Consideramos tres esquemas para la función de probabilidad de pérdida de las respuestas:

$$\begin{aligned} p_{\text{COMP}}(t) &= 1, \\ p_{\text{LOG}}(t) &= (1 + \exp(-0,2 t - 0,2))^{-1}, \\ p_{\text{COS}}(t) &= 0,8 + 0,2 \cos^2(0,8 t + 0,4), \end{aligned}$$

aplicadas a x_{i1} , obteniéndose un porcentaje de pérdida aproximadamente igual a 0.25 y 0.10 con p_{LOG} y p_{COS} , respectivamente.

La ρ –función que utilizamos para calcular las $E(\text{IF}(\mathbf{z}_*, \mathcal{T})^2)$ en el caso en que el estadístico del test se basa en los estimadores robustos WMM – y MM –estimadores, pertenece a la familia de la función bicuadrada definida en (2.6) con $c = 4$ y $3,9$, respectivamente.

Asimismo, para el cómputo del estimador pesado utilizamos una función de peso w_o tipo Hard Rejection aplicada al gradiente $|\dot{\mathbf{f}}(\mathbf{x}, \boldsymbol{\theta})|$. Las constantes de poda de w_o fueron elegidas como los percentiles 0,025 y el 0,975 de la norma de gradiente (1,049112 y 1,373228, respectivamente). Los estimadores robustos fueron calibrados de manera de tener una eficiencia asintótica del 90 %.

Para ilustrar el comportamiento de la esperanza de la función de influencia, calculamos $E(\text{IF}(\mathbf{z}_*, \mathcal{T})^2)$ sobre puntos arbitrarios $\mathbf{z}_* = (x_0, 1/2, y_0)$, donde (x_0, y_0) varían en el rectángulo $[0, 15] \times [0, 1,5]$. En la Figura 6.1 graficamos los valores obtenidos de $E(\text{IF}(\mathbf{z}_*, \mathcal{T})^2)$ en función de x_0 e y_0 para los casos en que el estadístico \widehat{W} se basa en $\widehat{\boldsymbol{\theta}}_{\text{MC}}$, $\widehat{\boldsymbol{\theta}}_{\text{MM}}$ y $\widehat{\boldsymbol{\theta}}_{\text{WMM}}$.

Como podemos observar en estos gráficos la esperanza de la función de influencia del test basado en el estimador de mínimos cuadrados, no es acotada y esto ocurre tanto ante valores grandes de y_0 como cercanos a 0 de x_0 . Bajo el esquema de respuestas missing $p_{\cos}(t)$ la $E(\text{IF}(\mathbf{z}_*, \mathcal{T})^2)$ del test clásico muestra una leve oscilación debida a que el patrón de pérdida depende de un coseno y esto se repite en los gráficos correspondientes a los test robustos. El comportamiento de la esperanza de la función de influencia cuando el test se basa en los estimadores robustos $WMM-$ y $MM-$ se asemeja al del estimador de mínimos cuadrados para valores moderados de x_0 y de y_0 bajo los tres esquemas de pérdida considerados. Para aquellos puntos arbitrarios con los mayores valores de y_0 y cercanos a 0 de x_0 , la $E(\text{IF}(\mathbf{z}_*, \mathcal{T})^2)$ de los dos tests robustos es controlada, mostrando que su influencia se mantiene acotada, pero puede observarse que cuando se utiliza el estimador pesado también se penalizan aquellas observaciones con x_0 cercanos a 0 y valores de y_0 cercanos a 0,5. Esto sucede para los tres esquemas de falta de respuestas considerados.

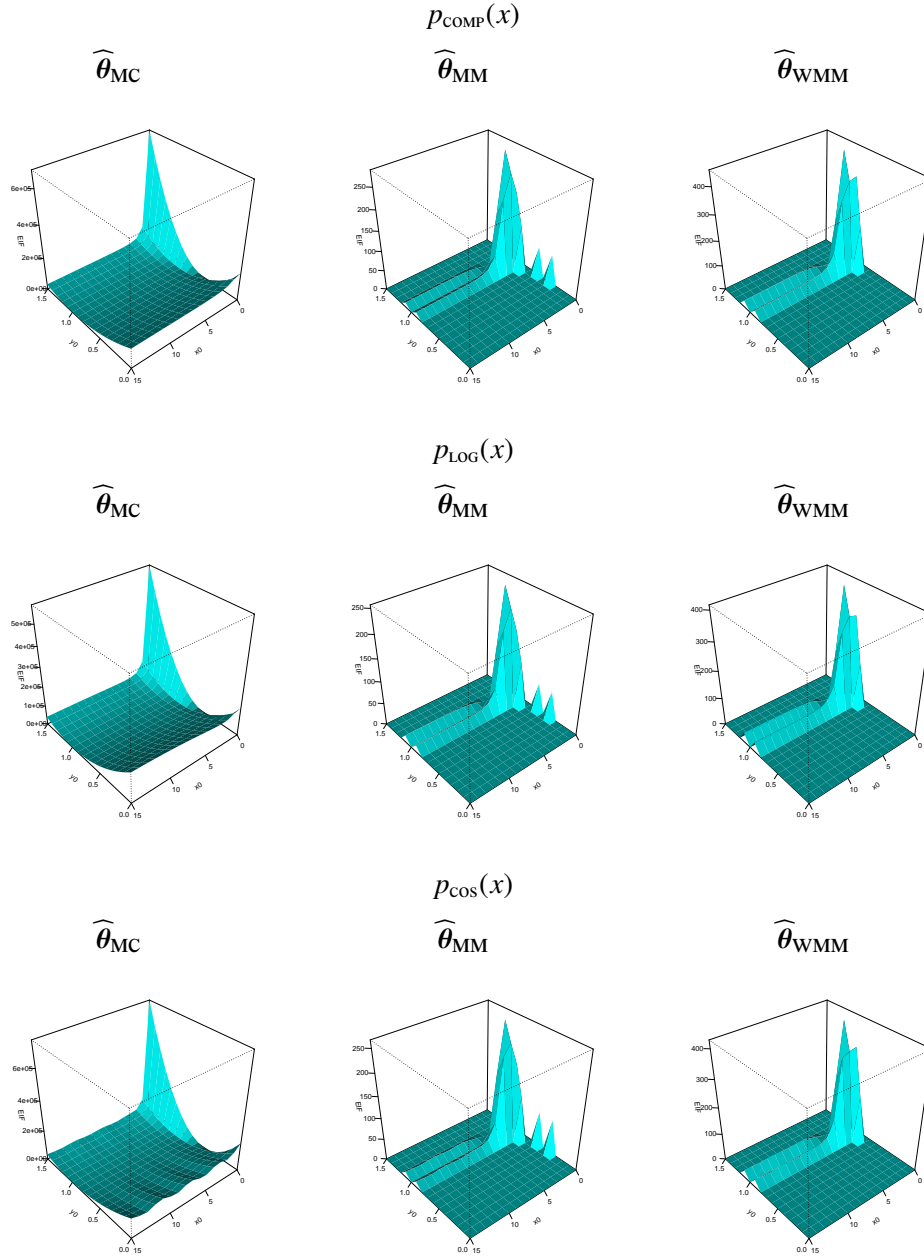


Figura 6.1: Gráficos de $E(\text{IF}(\mathbf{z}_*, \mathcal{T})^2)$ en función de x_0 e y_0 para el Modelo Lineal-exponencial cuando el estadístico $\widehat{\mathcal{W}}$ se basa en $\widehat{\theta}_{\text{MC}}$, $\widehat{\theta}_{\text{MM}}$ y $\widehat{\theta}_{\text{WMM}}$.

6.3. Estudio de Simulación

Hemos llevado a cabo dos experimentos de Monte Carlo con el objetivo de evaluar el desempeño de los tests propuestos comparándolos con el test clásico, como así también con los tests basados en otras propuestas robustas existentes.

6.4. Modelo Lineal–exponencial

En este primer experimento consideramos nuevamente el modelo Lineal–exponencial tratado en las Secciones 5.2 y 6.2.1. En este estudio realizamos pruebas de hipótesis sobre los parámetros del modelo lineal–exponencial definido en la Sección 5.2 para estudiar tanto el nivel como la potencia de los tests. Este análisis se realizó en base a $NR = 1000$ muestras aleatorias de tamaño $n = 300$. Denominaremos K_0 al diseño sin contaminación. Por otro lado, consideramos muestras con un 10 % de contaminación siguiendo los siguientes esquemas:

$K_{kx,ky}$: 30 observaciones son reemplazadas por puntos de la forma $\mathbf{x}_0 = (kx, 0)$ e $y_0 = 0,10 * ky (\theta_1 + \theta_2 x_{0,2} + \theta_3 e^{\theta_0 x_{0,1}})$ siendo $kx = 9, 10, 11, 12, 13$ y $ky = 8, 9, 11,$

con $\boldsymbol{\theta} = (-1, 1, 0, -0,5)^t$. En la Figura 6.2 se representan los distintos tipos de contaminación considerados. En el panel de la izquierda se grafica una muestra de los datos simulados originales, mientras que en los paneles central y derecho se indican en rojo los puntos atípicos introducidos según el esquema $K_{11,9}$ y los distintos esquemas de contaminación $K_{kx,ky}$, respectivamente.

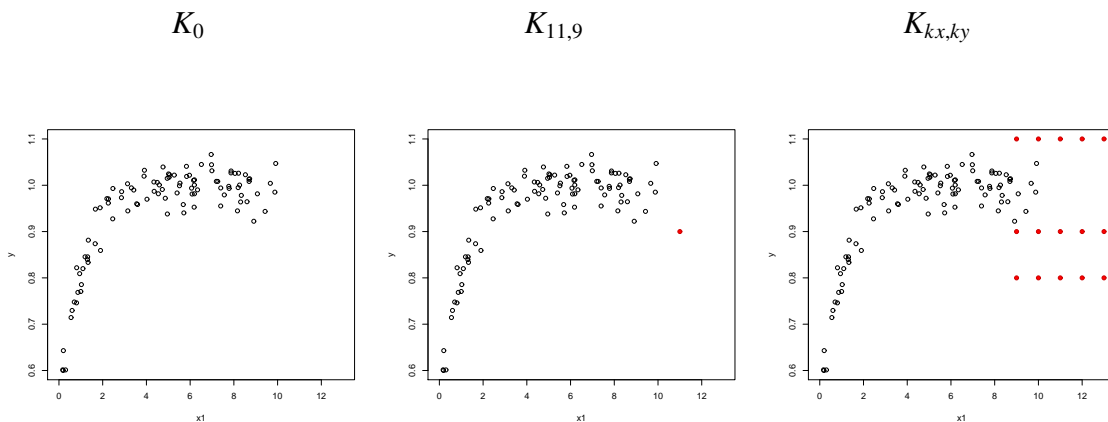


Figura 6.2: Una muestra de los datos simulados: en el panel de la izquierda los datos originales. En los paneles central y derecho los puntos rojos corresponden a las observaciones atípicas introducidas según los distintos esquemas de contaminación $K_{kx,ky}$.

Respecto de las probabilidades de pérdida, tomamos $p_{\text{COMP}}(t)$ y $p_{\text{LOG}}(t)$ descritas en la sección anterior. Asimismo, hemos computado los estadísticos basados en el estimador clásico y en los estimadores robustos tal como fueron considerados en la Sección 6.2.1. Se calculó para los estimadores robustos como estimador inicial, $\widehat{\theta}_{ini}$, un S -estimador y los pesos del estimador pesado fueron elegidos como en la sección anterior y se aplicaron al gradiente $|\dot{\mathbf{f}}(\mathbf{x}, \widehat{\theta}_{ini})|$.

La hipótesis a testear es

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \quad \text{vs.} \quad H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$$

con $\boldsymbol{\theta}_0 = (-1, 1, 0, -0,5)^t$. Para comparar el comportamiento del nivel de los tests calculamos las frecuencias observadas de rechazo en cada escenario considerado tomando un nivel nominal $\alpha = 0,05$. Estos valores están reportados en el Cuadro 6.1. Como se puede observar en dicho cuadro el efecto de las contaminaciones consideradas sobre el test clásico es enorme, ya que la frecuencia de rechazo es 1, es decir que siempre se rechaza la hipótesis nula aún cuando es cierta. En el caso de los tests robustos, las contaminaciones con $ky = 8$ son bien toleradas por el test basado en el MM -estimador en muestras completas y un poco menos toleradas en muestras con respuestas missing ya que el nivel aumenta con kx . Sin embargo, las frecuencias de rechazo observadas para este test son muy cercanas o iguales a 1 cuando ky crece. Por el contrario, el nivel observado del test basado en WMM es muy estable en todos los escenarios considerados.

			$p_{\text{COMP}}(t)$			$p_{\text{LOG}}(t)$		
			MC	MM	WMM			
K_0			0,065	0,083	0,083	K_0		
$K_{kx,ky}$	ky	kx				$K_{kx,ky}$	ky	kx
	8	9	1	0,082	0,077		8	9
	8	10	1	0,082	0,077		8	10
	8	11	1	0,082	0,077		8	11
	8	12	1	0,081	0,077		8	12
	8	13	1	0,082	0,077		8	13
	9	9	1	0,998	0,082		9	9
	9	10	1	0,998	0,082		9	10
	9	11	1	0,998	0,082		9	11
	9	12	1	0,998	0,082		9	12
	9	13	1	0,998	0,082		9	13
	11	9	1	1	0,089		11	9
	11	10	1	0,999	0,088		11	10
	11	11	1	0,999	0,088		11	11
	11	12	1	1	0,089		11	12
	11	13	1	1	0,088		11	13

Cuadro 6.1: Frecuencias observadas bajo H_0 correspondientes a $\widehat{\mathcal{W}}$ basado en MC , MM y WMM para K_0 y $K_{kx,ky}$.

Con el objetivo de comparar el comportamiento de la potencia de los distintos tests computamos sus frecuencias de rechazo bajo distintas hipótesis alternativas dadas por

$$H_{1,\Delta} : \boldsymbol{\theta}_n = \boldsymbol{\theta}_0 + \Delta n^{-1/2}(1, 1, 1, 1)^t$$

para $\Delta = \pm 0,100, \pm 0,075, \pm 0,050, \pm 0,025, 0$ nuevamente con nivel nominal $\alpha = 0,05$. A modo de ejemplo, nos enfocaremos en el tipo de contaminación $K_{11,9}$, que por lo visto hasta el momento, en términos de nivel y su efecto en la estimación (coincide con $C_{1,10}^*$ de la Sección 5.2), figura entre los esquemas de contaminación más graves.

A continuación, en las Figuras 6.3 y 6.4, graficamos las frecuencias observadas de rechazo en función de Δ bajo cada uno de los esquemas de pérdida de respuestas considerados. En dichos plots se representa con negro la potencia del test basado en mínimos cuadrados, en azul la correspondiente al MM -estimador y en verde a nuestro estimador pesado. Además, en los Cuadros 6.2 y 6.3 reportamos los valores que fueron graficados.

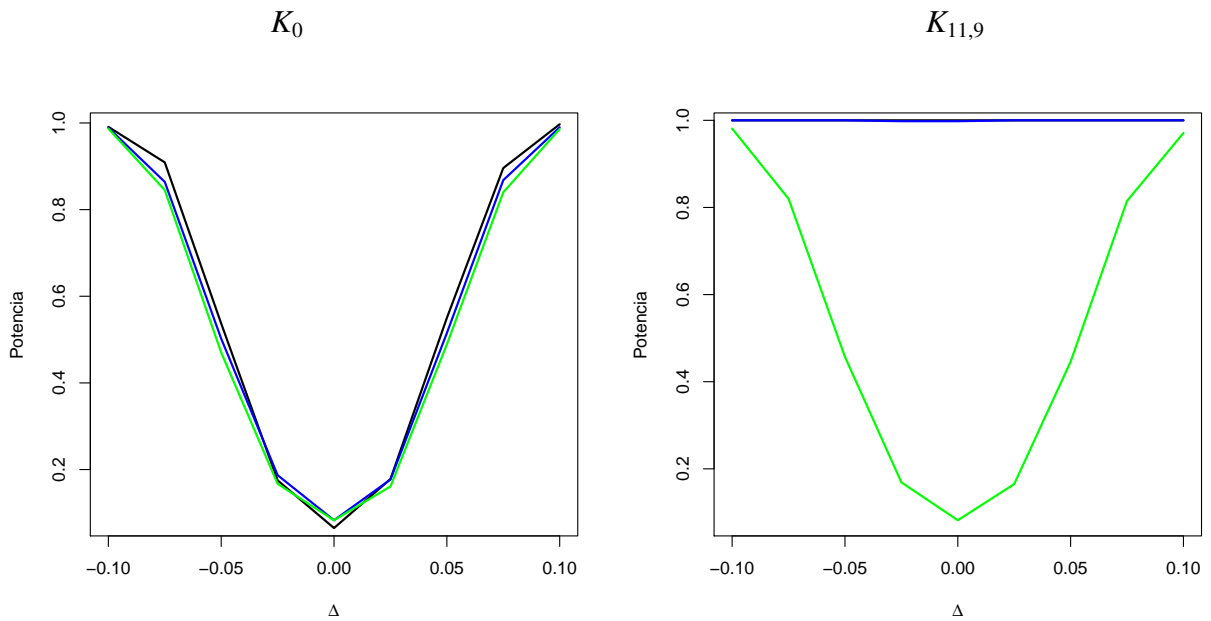


Figura 6.3: Frecuencias de rechazo en el caso $p_{\text{COMP}}(t)$.

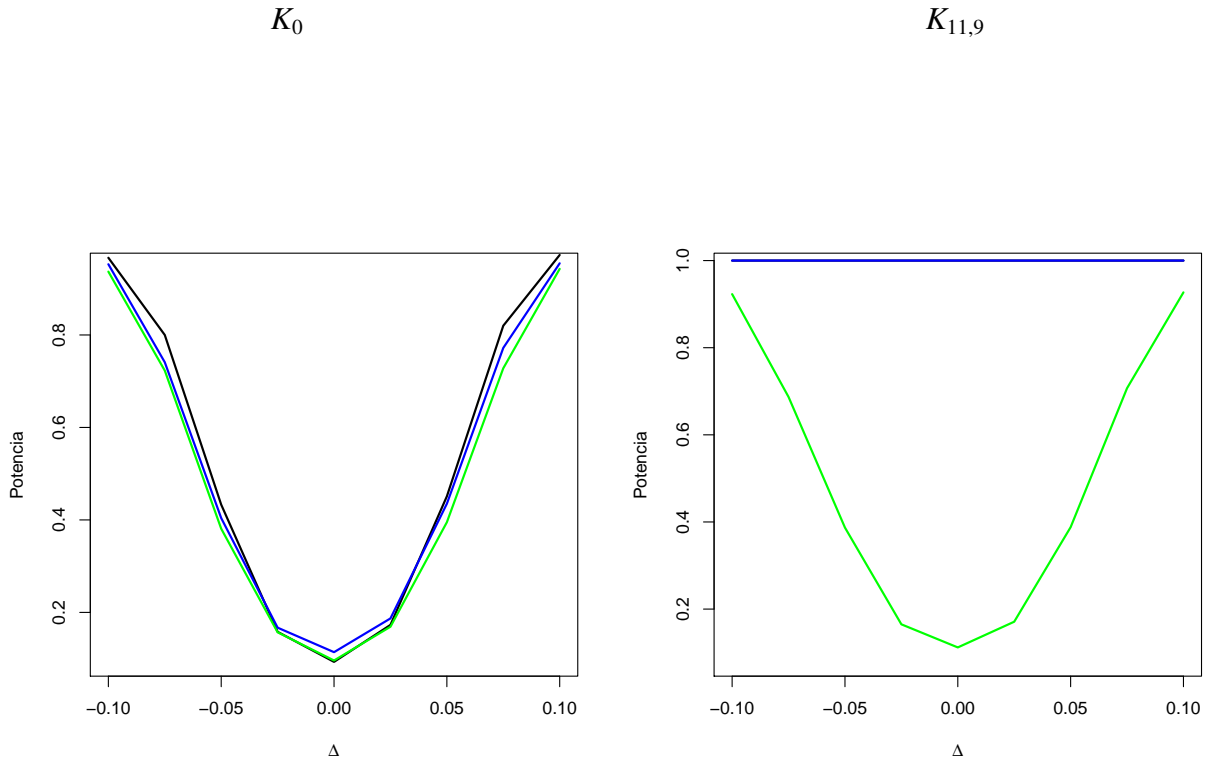


Figura 6.4: Frecuencias de rechazo en el caso $p_{\text{Log}}(t)$.

	K_0								
Delta	-0.100	-0.075	-0.050	-0.025	0.000	0.025	0.050	0.075	0.100
MC	0.991	0.909	0.537	0.175	0.065	0.178	0.550	0.896	0.997
MM	0.989	0.864	0.502	0.187	0.083	0.177	0.515	0.868	0.990
WMM	0.987	0.845	0.470	0.167	0.083	0.161	0.489	0.840	0.986
	$K_{11,9}$								
MC	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MM	1.000	1.000	1.000	0.998	0.998	1.000	1.000	1.000	1.000
WMM	0.981	0.820	0.457	0.169	0.082	0.165	0.446	0.815	0.971

Cuadro 6.2: Frecuencias de rechazo de $\widehat{\mathcal{W}}$ basado en MC, MM y WMM para $p_{\text{COMP}}(t)$.

	K_0								
Delta	-0.100	-0.075	-0.050	-0.025	0.000	0.025	0.050	0.075	0.100
MC	0.967	0.800	0.433	0.158	0.093	0.173	0.451	0.820	0.973
MM	0.953	0.741	0.404	0.167	0.114	0.187	0.435	0.772	0.955
WMM	0.937	0.723	0.381	0.157	0.096	0.169	0.395	0.728	0.943
	$K_{11,9}$								
MC	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MM	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
WMM	0.923	0.687	0.387	0.165	0.112	0.171	0.388	0.707	0.927

Cuadro 6.3: Frecuencias de rechazo de $\widehat{\mathcal{W}}$ basado en MC, MM y WMM para $p_{\text{LOG}}(t)$.

Tanto en estos gráficos como tablas podemos observar que en las muestras no contaminadas, el comportamiento de la potencia de todos los tests analizados es similar. Sin embargo, al introducir un 10 % de contaminación, la potencia del test basado en el estadístico clásico y la del estadístico basado en el *MM*-estimadores se ven severamente afectadas. Para el esquema de contaminación considerado, estos dos tests se vuelven prácticamente no informativos, ya que su función de potencia observada es prácticamente 1. Esta conclusión es válida tanto para muestras completas como para muestras con respuestas perdidas. Por el contrario, si bien el test basado en el estimador pesado tiene un leve aumento del nivel y una muy leve pérdida de potencia en muestras contaminadas, su función de potencia es muy estable y el test sigue siendo informativo, tanto bajo el esquema de pérdida $p_{\text{COMP}}(t)$ como $p_{\text{LOG}}(t)$.

6.5. Modelo de Crecimiento Exponencial

En este estudio realizamos pruebas de hipótesis sobre los parámetros del modelo exponencial definido en la Sección 5.1 para estudiar tanto el nivel como la potencia de los tests. Este análisis se realizó en base a $NR = 1000$ muestras aleatorias de tamaño $n = 100$ y consideramos los mismos casos de contaminación y de probabilidades de pérdida que los utilizados en la sección mencionada. Hemos considerado los estadísticos robustos basados en los estimadores que se computan con inicial un *S*-estimador ya que como hemos visto, en general, se obtienen errores cuadráticos menores para este modelo.

Para comparar el comportamiento del nivel de los tests calculamos las frecuencias observadas de rechazo para cada escenario considerado bajo un nivel nominal $\alpha = 0,05$. Estos valores están reportados en el Cuadro 6.4. Asimismo, en dicho Cuadro podemos observar que en el caso sin contaminación los niveles muestrales de los tests robustos se comportan de manera similar al nivel muestral del test basado en el estimador de mínimos cuadrados para los tres esquemas de funciones de pérdida considerados.

Con el objetivo de comparar el comportamiento de la potencia de los distintos tests robustos computamos sus frecuencias de rechazo bajo distintas hipótesis alternativas dadas

$\alpha = 0,05$	Eficiencia	Tests	p_{COMP}	p_{LOG}	p_{COS}
C_0		MC	0,064	0,071	0,063
	85 %	M	0,090	0,098	0,097
		MM	0,094	0,097	0,094
		WMM ₁	0,078	0,088	0,082
		WMM ₂	0,074	0,086	0,075
	90 %	M	0,084	0,087	0,092
		MM	0,087	0,090	0,094
		WMM ₁	0,078	0,083	0,080

Cuadro 6.4: Frecuencias observadas de rechazo bajo H_0 para C_0 .

por

$$H_{1,\Delta} : \theta_1 = 5 + \Delta n^{-1/2}, \theta_2 = 2 + \Delta n^{-1/2},$$

para $\Delta = \pm 0,05, \pm 0,1, \pm 0,15, \pm 0,2, \pm 0,25, \pm 0,3, \pm 0,35, \pm 0,4$ con nivel nominal $\alpha = 0,05$.

A continuación, en las Figuras 6.5, 6.6 y 6.7, graficamos las frecuencias observadas de rechazo en función de $\Omega = \Delta n^{-1/2}$, con $\Delta = 0, \pm 0,05, \pm 0,1, \pm 0,15, \pm 0,2, \pm 0,25, \pm 0,3, \pm 0,35, \pm 0,4$, bajo cada uno de los esquemas de pérdida de datos considerados.

En la Figura 6.5 se encuentran los gráficos correspondientes al caso de datos completos. En ella encontraremos graficada con línea negra (—) a la potencia bajo el procedimiento clásico, con líneas verde (—), azul (—), roja (—) y naranja (—) las correspondientes al M -estimador, al MM -estimador, al WMM_1 -estimador y al WMM_2 -estimador con 85 % de eficiencia, respectivamente. Mientras que las potencias correspondientes a los M -, MM - y WMM_1 - estimadores con 90 % de eficiencia están representadas con líneas punteadas de colores verde, azul y rojo, respectivamente. En estos gráficos podemos observar que cuando no se introduce contaminación a las muestras, el comportamiento de la potencia es similar entre todos los estadísticos analizados. Luego, al reemplazar algunos datos por observaciones atípicas según cada esquema de contaminación, vemos que tanto la potencia del test clásico como las de los estadísticos basados en los M - y MM -estimadores se ven seriamente afectadas con cualquier porcentaje de contaminación. El test clásico se vuelve prácticamente no informativo bajo cualquiera de las contaminaciones ya que su función de potencia observada es 1. Si bien no ocurre lo mismo para los tests basados en los M - y MM - estimadores vemos que su frecuencia observada de rechazo bajo H_0 es mucho mayor al nivel nominal resultando significativamente distinta del valor 5 %. Esto ocurre tanto para todos los tipos de contaminación, como en todos los esquemas de pérdida de las respuestas utilizados. En cambio, al considerar los tests basados en los WMM -estimadores se ve que para la mayoría de ellos el nivel muestral permanece estable y cercano al valor nominal en la mayoría de las situaciones. Como era de esperar de acuerdo a lo ya visto en el estudio numérico del Capítulo 5, el tipo de contaminación C_1 con porcentajes por encima del 20 % arroja resultados pobres para todos los tests considerados. Todas estas consideraciones son válidas también para la Figura 6.6 en el caso en que hay pérdida de datos por un esquema logístico como cuando la pérdida sigue un esquema por coseno, el cual se encuentra en la Figura 6.7.

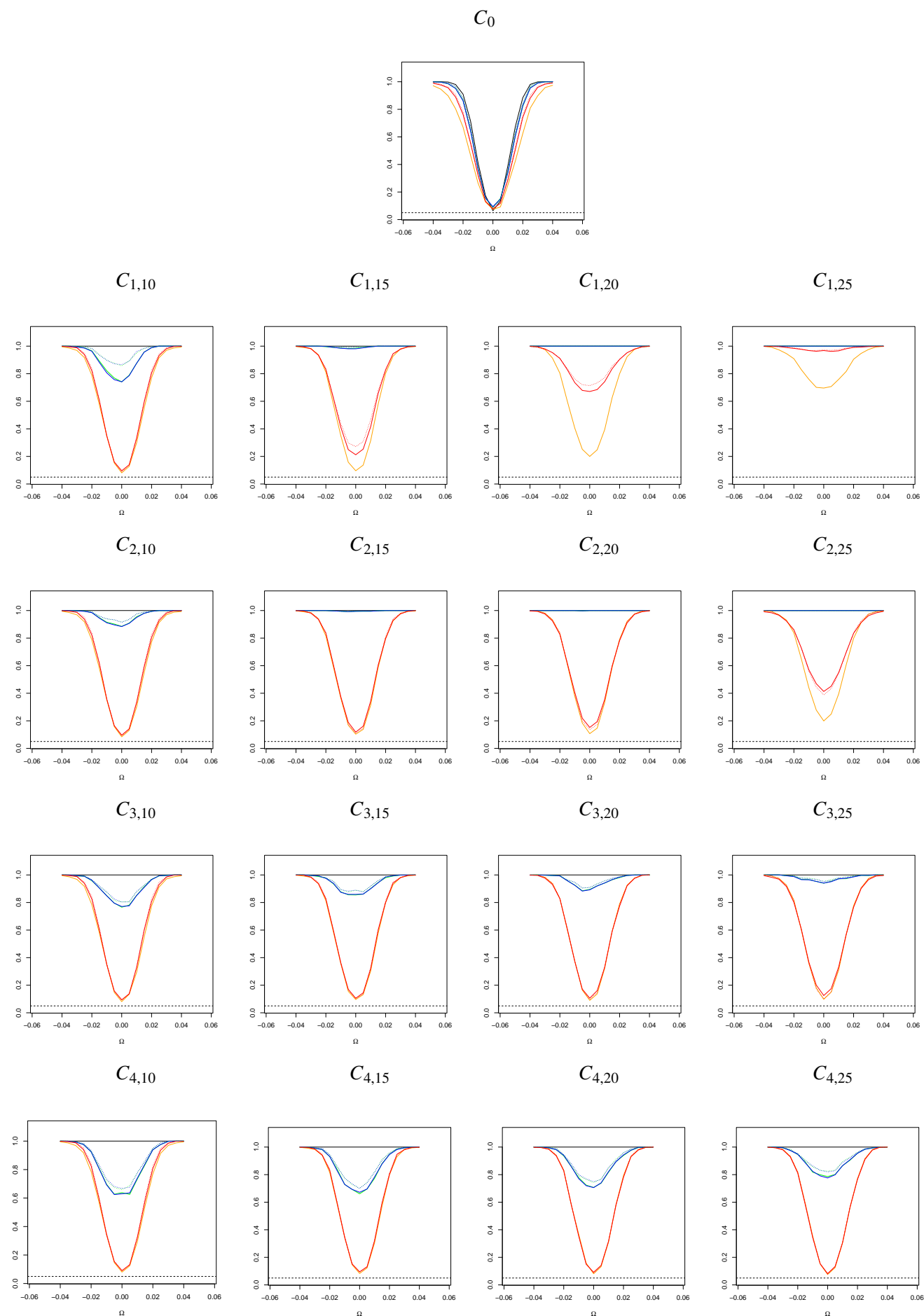


Figura 6.5: Frecuencias observadas de rechazo cuando la probabilidad de pérdida de las respuestas es p_{COMP}

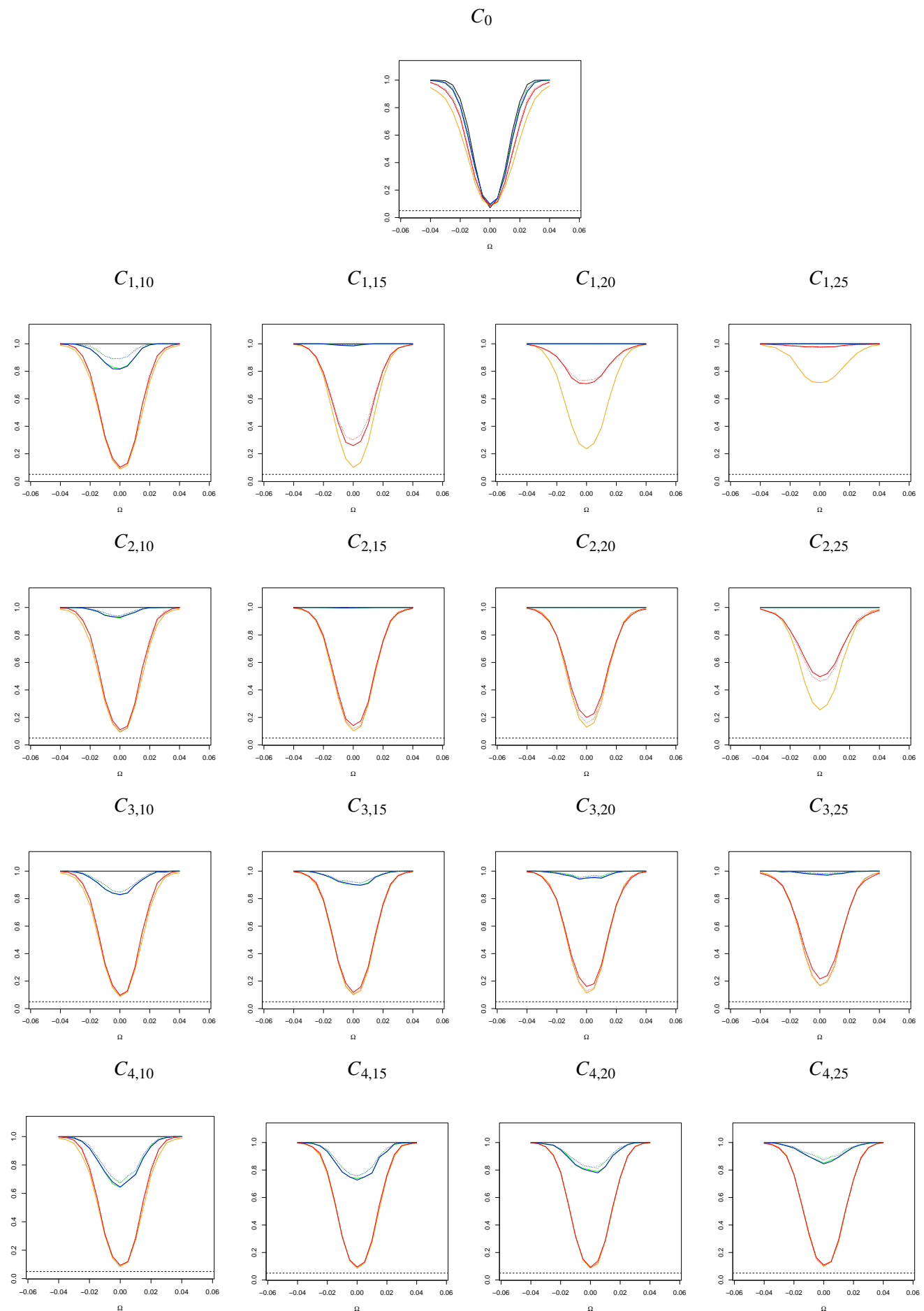


Figura 6.6: Frecuencias observadas de rechazo cuando la probabilidad de pérdida de las respuestas es p_{LOG}

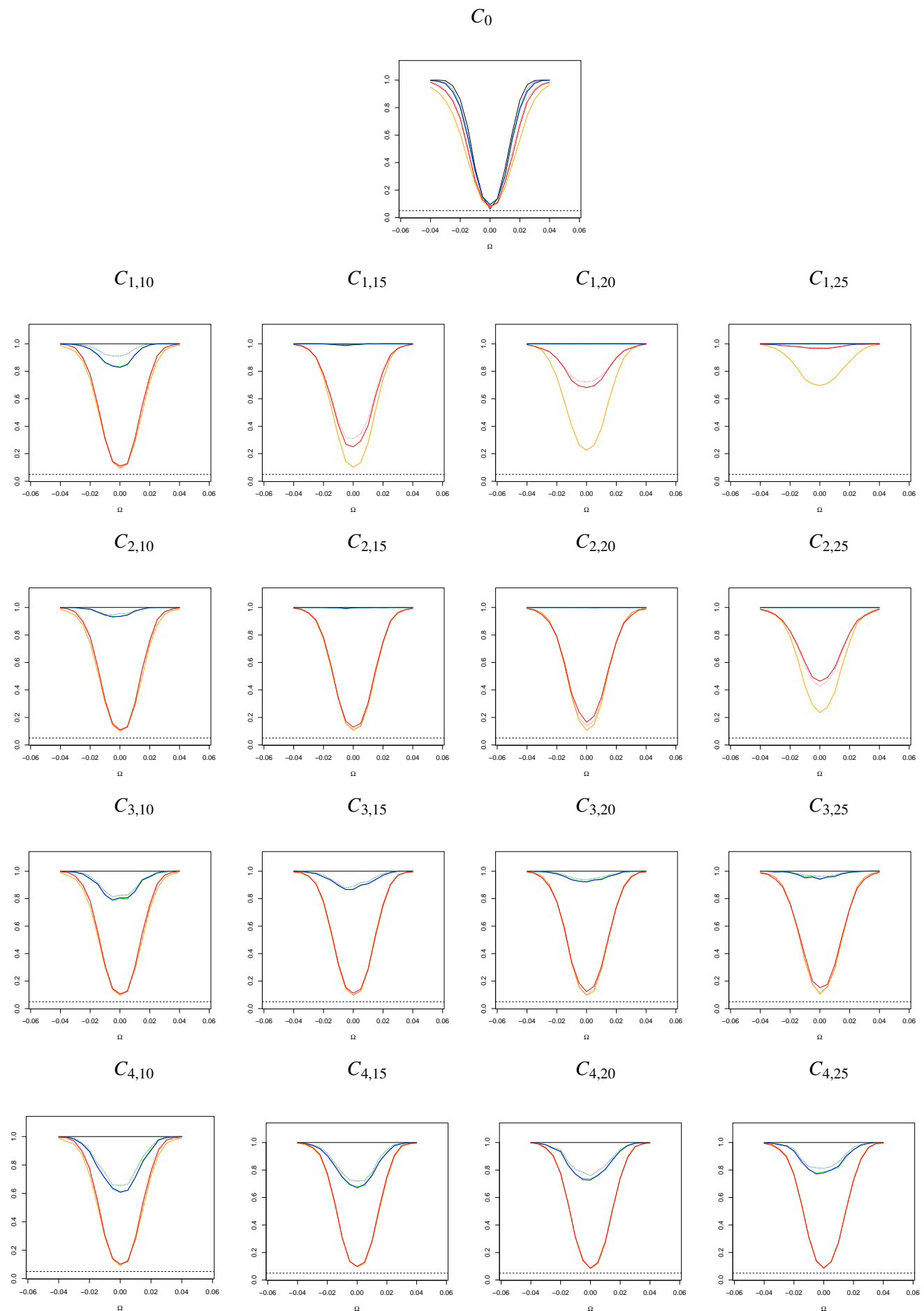


Figura 6.7: Frecuencias observadas de rechazo cuando la probabilidad de pérdida de las respuestas es p_{\cos}

Finalmente, para dar una medida del desempeño de la potencia de los tests robustos y poder compararlos, utilizamos la “magnitud corregida de la potencia exacta relativa” que fue definida por Morales *et al.* (2004). En su trabajo definen esta magnitud para comparar el comportamiento de la potencia entre dos tests cuyos estadísticos estén dados por $\widehat{W}_{1,n}$ y $\widehat{W}_{2,n}$ y se calcula en base a las frecuencias observadas de rechazo bajo la hipótesis nula, $\widehat{\alpha}_{1,H_0}$ y $\widehat{\alpha}_{2,H_0}$, respectivamente, y bajo la hipótesis alternativa, $\widehat{\alpha}_{1,H_1}$ y $\widehat{\alpha}_{2,H_1}$, respectivamente, como

$$\rho_{H_1}(\widehat{W}_{1,n}, \widehat{W}_{2,n}) = \left(\frac{\widehat{\alpha}_{1,H_1} - \widehat{\alpha}_{1,H_0}}{\widehat{\alpha}_{2,H_1} - \widehat{\alpha}_{2,H_0}} - 1 \right) 100. \quad (6.16)$$

Lo que hicimos fue calcularla para cada uno de los estadísticos robustos versus el estadístico clásico. Cuando no hay contaminación, el comportamiento es mejor cuanto más se parezca al del test clásico, es decir cuanto más cercano a cero, en valor absoluto, sea la magnitud definida en (6.16).

Como mencionamos previamente, ante la presencia de datos atípicos, tenemos que bajo H_0 el estadístico basado en mínimos cuadrados rechaza la hipótesis nula en todas las muestras consideradas, esto sucede también bajo cada una de las hipótesis alternativas $H_{1,\Delta}$, por lo que el denominador de la magnitud $\rho_{H_{1,\Delta}}$ da cero, siendo infinito el valor de esta medida. Por este motivo es que solo la calculamos para el caso C_0 . Por otra parte, vimos que el comportamiento de las frecuencias observadas de rechazo $\widehat{\alpha}$ para las distintas hipótesis alternativas $H_{1,\Delta}$ resultó simétrico respecto del cero, por lo cual informaremos los valores con Δ positivo. Estos resultados se encuentran en el Cuadro 6.5 para cada una de las probabilidades de pérdida. Como puede observarse, el test basado en el estimador pesado da valores de $\rho_{H_{1,\Delta}}$ menores que los basados en los $M-$ y $MM-$ estimadores, esta diferencia es más pronunciada a medida que Δ se acerca a 0. Esto indica que bajo H_0 estos tests son más eficientes para detectar este tipo de alternativas, que sería el precio a pagar por una mayor robustez, sin embargo, como es de esperar, a medida que Δ crece esta diferencia disminuye marcadamente.

P_{COMP}	Δ	$H_{1,\Delta}$							
		0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40
85 %	M	-11,27	-23,31	-15,08	-10,73	-5,45	-3,74	-2,78	-2,78
	MM	-22,54	-24,54	-16,23	-10,73	-6,21	-4,17	-3,21	-3,20
	WMM ₁	-42,25	-35,89	-29,67	-19,15	-12,43	-5,77	-3,31	-2,46
	WMM ₂	-74,65	-46,01	-42,62	-32,93	-19,85	-11,87	-5,55	-3,85
90 %	M	-12,68	-15,64	-9,34	-7,44	-3,27	-2,35	-2,14	-2,14
	MM	-19,72	-16,26	-10,00	-7,80	-3,82	-2,67	-2,46	-2,46
	WMM ₁	-46,48	-34,36	-25,08	-17,19	-10,36	-4,92	-2,99	-2,24
P_{LOG}	Δ	$H_{1,\Delta}$							
		0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4
85 %	M	-35,82	-24,91	-15,41	-9,70	-8,47	-4,64	-3,23	-2,91
	MM	-34,33	-24,54	-14,87	-10,35	-8,47	-4,53	-3,12	-2,80
	WMM ₁	-56,72	-38,83	-32,08	-24,32	-16,16	-9,17	-5,71	-3,44
	WMM ₂	-64,18	-49,82	-47,49	-37,39	-27,42	-16,50	-9,90	-6,14
90 %	M	-23,88	-9,89	-8,78	-5,43	-5,02	-2,48	-1,72	-1,72
	MM	-23,88	-10,26	-10,04	-6,08	-5,35	-2,80	-2,05	-2,05
	WMM ₁	-50,75	-34,80	-28,67	-21,35	-13,49	-7,34	-4,52	-2,48
P_{COS}	Δ	$H_{1,\Delta}$							
		0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40
85 %	M	-50,00	-29,76	-15,26	-12,28	-9,07	-5,78	-4,06	-3,63
	MM	-44,74	-28,72	-14,18	-12,03	-8,74	-5,35	-3,63	-3,31
	WMM ₁	-64,47	-39,45	-34,83	-25,57	-15,93	-9,53	-5,44	-3,74
	WMM ₂	-60,53	-50,52	-43,81	-37,97	-26,22	-15,74	-8,96	-5,12
90 %	M	-40,79	-24,57	-11,31	-9,37	-6,42	-3,75	-3,20	-3,09
	MM	-43,42	-23,88	-11,13	-9,62	-6,75	-4,07	-3,42	-3,31
	WMM ₁	-60,53	-37,02	-27,65	-22,41	-13,16	-8,14	-4,59	-3,09

Cuadro 6.5: Magnitud de Morales para el caso C_0 .

Capítulo 7

Ejemplo: Datos de Calidad del Aire

Para ilustrar el uso de los estimadores propuestos en el caso de respuestas faltantes, consideramos un conjunto de datos registrado en el área metropolitana de Nueva York. Con el objetivo de estudiar la calidad del aire, se registraron 153 observaciones que consisten en lecturas diarias tomadas entre el 1° de Mayo de 1973 y el 30 de septiembre de 1973 de cuatro variables: Ozono, Radiación Solar, Velocidad del Viento y Temperatura Máxima. Cleveland (1985) estudió la relación entre el Ozono (ppb) y la Velocidad del Viento (mph). Detectó una relación no lineal en la cual a medida que la Velocidad aumenta, el Ozono decrece. Esto se debe al incremento de ventilación que se produce cuando la velocidad del viento es mayor. En 37 observaciones el valor del Ozono está perdido. Al igual que Cleveland (1985) tomamos al Ozono (O) como variable de respuesta y a la Velocidad del Viento (V) como la variable explicativa. De esta manera, para nuestro análisis consideramos los 153 datos, de los cuales hay registrados 116 valores para la variable de respuesta Ozono, por lo que hay 37 respuestas perdidas.

Para estudiar esta relación con los distintos tipos de estimadores que estuvimos trabajando en los capítulos anteriores, realizamos varios ajustes con diferentes funciones de regresión no lineal que podrían responder al comportamiento observado. Entre ellas vimos que la que mejor ajustaba correspondía al modelo de decrecimiento exponencial dado por

$$O = \theta_1 \exp(\theta_2 V) + \theta_3, \quad \theta_2 < 0.$$

Bajo este modelo, computamos el estimador de mínimos cuadrados, MM - y WMM -estimadores. Para ambas clases de estimadores robustos utilizamos la función bicuadrada para controlar a las observaciones con residuos altos y para el WMM -estimador elegimos como función de peso a

$$w_o(t) = \begin{cases} 1 & \text{si } t^2 \leq (c_w - 1/2c_w)^2 \\ \frac{c_w^2 - t^2}{c_w^2 - (c_w - 1/2c_w)^2} & \text{si } (c_w - 1/2c_w)^2 < t^2 \leq c_w^2 \\ 0 & \text{si } t^2 > c_w^2, \end{cases}$$

aplicado a la distancia de Mahalanobis robusta como definimos en la Sección 4.3.1, ya que al utilizar este tipo de función Hard Rejection obtuvimos un comportamiento más

estable de los estimadores pesados que al tomar w_o dentro de la familia inducida por la bicuadrada.

Asimismo, para las versiones robustas utilizamos distintos estimadores iniciales. El primero, Ini_1 , se basó en el estimador utilizado para el modelo de crecimiento exponencial (4.5), que resulta de una aproximación lineal al modelo, definido por (5.1) y (5.2) y tomando $\hat{\theta}_{3,MED} = 0,9 \text{ mín}(O)$. Por otro lado, consideramos dos S -estimadores, uno de ellos con punto inicial basado en Ini_1 y el otro con punto inicial basado en el LMS -estimador, tal como fuera sugerido por Stromberg (1993). Llamaremos a estos dos últimos estimadores iniciales Ini_2 e Ini_3 , respectivamente. En primer lugar calculamos los estimadores de los parámetros obteniendo los valores que figuran en el Cuadro 7.1.

	Estimadores							
		Ini_1		Ini_2		Ini_3		
	MC	MM	WMM	MM	WMM	MM	WMM	MC ⁻⁵
θ_1	216,93	213,12	685,43	212,26	712,36	897,10	826,88	554,82
θ_2	-0,23	-0,26	-0,47	-0,26	-0,48	-0,54	-0,52	-0,42
θ_3	12,86	13,25	21,52	13,38	21,41	20,16	21,20	21,15

Cuadro 7.1: Datos del Aire.

Como podemos observar, los resultados del MM -estimador varían notablemente según los distintos valores iniciales usados. Cuando es computado a partir de Ini_1 e Ini_2 los resultados son muy similares a los obtenidos con el procedimiento de mínimos cuadrados. Por otro lado, al utilizar a Ini_3 como punto de partida, los valores del MM -estimador se acercan a las estimaciones obtenidas a partir del estimador propuesto, que se muestra muy estable respecto del punto inicial utilizado. Esta propiedad, que también fue observada en el marco de los estudios de simulación realizados, resulta relevante ya que la elección del punto inicial en el contexto de estimación en modelos no lineales siempre merece atención, más allá del problema de robustez.

En segunda instancia, con el objetivo de analizar el comportamiento de los estimadores considerados construimos intervalos de confianza (IC) de nivel asintótico 99 % deducidos a partir de la distribución asintótica obtenida en el Teorema 3.4.2, considerando las versiones muestrales definidas en (6.5) y (6.6) evaluadas en los estimadores correspondientes.

Al observar los intervalos de confianza que figuran en el Cuadro 7.2, podemos ver que los construidos para el tercer parámetro a partir del estimador de mínimos cuadrados y del MM -estimador basado en Ini_1 e Ini_2 , nos llevan a la conclusión de que θ_3 no es significativamente diferente de cero. Sin embargo, esta conclusión no es del todo satisfactoria ya que en el estudio inicial de estos datos, el modelo de decrecimiento exponencial con $\theta_3 = 0$ no dio un buen ajuste. Lo mismo ocurre para el parámetro θ_1 . Por el contrario, al mirar los IC calculados a partir de cualquiera de los WMM -estimadores y del MM -estimador con inicial Ini_3 , vemos que todos los parámetros del modelo resultan significativamente diferentes de cero. Cabe observar también que al comparar los intervalos

	Intervalos de Confianza					
	Ini_1		Ini_2		Ini_3	
	MC	MM	WMM	MM	WMM	WMM
θ_1	[69,87 ; 363,99]	[-57,44 ; 483,69]	[137,96 ; 1232,90]	[-11,58 ; 436,10]	[141,32 ; 1283,39]	[138,59 ; 1655,60]
θ_2	[-0,39 ; -0,08]	[-0,49 ; -0,02]	[-0,69 ; -0,26]	[-0,46 ; -0,06]	[-0,70 ; -0,27]	[-0,77 ; -0,30]
θ_3	[-4,16 ; 29,89]	[-2,51 ; 29,02]	[13,83 ; 29,20]	[-0,87 ; 27,63]	[13,91 ; 28,91]	[14,07 ; 26,26]

Cuadro 7.2: Intervalos de Confianza

obtenidos con el Ini_3 como punto inicial, la longitud de la mayoría de los computados con el WMM -estimador es menor que la de los computados con el MM -estimador.

Por otro lado, con el objetivo de detectar datos atípicos en la muestra calculamos el boxplot de los residuos estandarizados obtenidos a partir de los WMM -estimadores. En los tres casos se identificaron los mismos 5 outliers, que corresponden a las observaciones número 86, 100, 101, 121 y 126 de la muestra original. Estos datos atípicos también son detectados a partir de los residuos estandarizados correspondientes al ajuste dado por el MM -estimador basado en Ini_3 . Asimismo, las observaciones número 86, 100, 101 también son detectadas como anómalas a partir de los residuos computados con el estimador clásico y también con los MM -estimadores basados en Ini_1 e Ini_2 . Cabe observar que las observaciones 121 y 126, no detectadas como atípicas, corresponden a dos días en los que se registró la menor velocidad de viento y con valores relativamente bajos para la variable Ozono. En la Figura 7.1 graficamos en negro la curva ajustada por el método de mínimos cuadrados, en azul punteado la ajustada por el MM -estimador usando Ini_1 como inicial, en azul la correspondientes al MM -estimador con inicial Ini_3 y en rojo la obtenida a partir del WMM -estimador. En verde se representan los outliers identificados por todos los métodos y en rojo los identificados sólo por el WMM -estimador

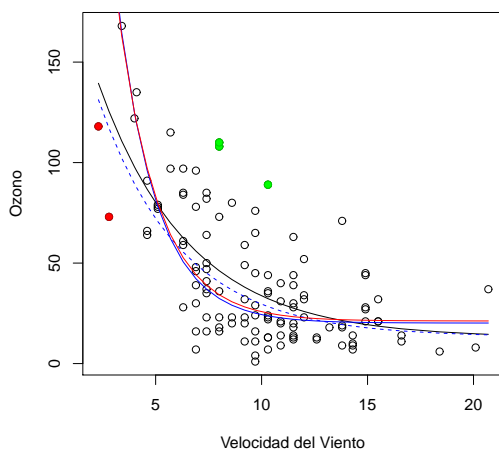


Figura 7.1: Gráfico de la presencia de Ozono en el aire en función de la Velocidad del viento. En negro la curva ajustada por mínimos cuadrados, en azul punteado por el MM -estimador con Ini_1 , en azul con el MM -estimador con Ini_3 , en rojo por el WMM -estimador. En verde se representan los outliers identificado por todos los métodos y en rojo los identificados sólo por el WMM -estimador.

Por último, eliminamos estas 5 observaciones de la muestra y recalculamos el estimador de mínimos cuadrados (mc^{-5}), cuyos resultados se encuentran en el Cuadro 7.1, en el cual podemos ver como los valores estimados se acercan a los calculados por los WMM -estimadores. Cabe destacar que el intervalo de confianza de nivel asintótico 99 % para θ_3 obtenido a partir del estimador clásico basado en esta muestra sin los cinco outliers detectados, quedó lejos del cero, siendo $[13,38; 28,93]$. Esto nos lleva a la misma conclu-

sión que la obtenida con los *WMM*-estimadores calculados sobre toda la muestra.

Capítulo 8

Apéndice

		$\theta_1 = 5$								$\theta_2 = 2$									
		85 %				90 %				85 %				90 %					
	Eficiencia	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2		
C_0	Media	5,0005	5,0001	4,9995	4,9975	4,9949	4,9993	4,9972	4,9998	4,9993	4,9993	4,9972	4,9998	4,9993	4,9993	4,9972	4,9998	4,9993	
	Mediana	4,9983	5,0042	5,0048	4,9995	5,0032	5,0052	5,0061	4,9985	5,0052	5,0061	4,9985	5,0052	5,0061	4,9985	5,0052	5,0061	4,9985	
	SD	0,1614	0,1693	0,1663	0,1807	0,1977	0,1640	0,1620	0,1764	0,1640	0,1620	0,1764	0,1640	0,1620	0,1764	0,1640	0,1620	0,1764	0,1640
	MAD	0,1548	0,1597	0,1584	0,1779	0,1816	0,1605	0,1561	0,1727	0,1605	0,1561	0,1727	0,1605	0,1561	0,1727	0,1605	0,1561	0,1727	0,1605
ECM	0,0260	0,0286	0,0276	0,0326	0,0391	0,0269	0,0262	0,0311	0,0269	0,0262	0,0311	0,0269	0,0262	0,0311	0,0269	0,0262	0,0311	0,0269	
	MC	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2		
$C_{1,5}$	Media	5,5342	5,3777	5,4185	5,0138	5,0027	5,4502	5,4722	5,0269	5,4502	5,4722	5,0269	5,4502	5,4722	5,0269	5,4502	5,4722	5,0269	
	Mediana	5,5402	5,4746	5,5034	5,0044	5,0016	5,5137	5,5216	5,0098	5,5137	5,5216	5,0098	5,5137	5,5216	5,0098	5,5137	5,5216	5,0098	
	SD	0,1618	0,3121	0,2910	0,1986	0,1928	0,2675	0,2473	0,2111	0,2675	0,2473	0,2111	0,2675	0,2473	0,2111	0,2675	0,2473	0,2111	
	MAD	0,1212	0,2276	0,1817	0,1819	0,1871	0,1610	0,1524	0,1850	0,1610	0,1524	0,1850	0,1610	0,1524	0,1850	0,1610	0,1524	0,1850	0,1610
ECM	0,3115	0,2400	0,2597	0,0396	0,0372	0,2742	0,2841	0,0453	0,2742	0,2841	0,0453	0,2742	0,2841	0,0453	0,2742	0,2841	0,0453	0,2742	
	MC	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2		
$C_{1,7}$	Media	5,5717	5,5081	5,5279	5,0777	5,0180	5,5414	5,5542	5,1224	5,5414	5,5542	5,1224	5,5414	5,5542	5,1224	5,5414	5,5542	5,1224	
	Mediana	5,5733	5,5462	5,5549	5,0305	5,0061	5,5600	5,5661	5,0745	5,5600	5,5661	5,0745	5,5600	5,5661	5,0745	5,5600	5,5661	5,0745	
	SD	0,1310	0,2363	0,2136	0,2767	0,2144	0,1959	0,1729	0,2903	0,1959	0,1729	0,2903	0,1959	0,1729	0,2903	0,1959	0,1729	0,2903	0,1959
	MAD	0,1257	0,1652	0,1568	0,2276	0,1962	0,1485	0,1423	0,2651	0,1485	0,1423	0,2651	0,1485	0,1423	0,2651	0,1485	0,1423	0,2651	0,1485
ECM	0,3440	0,3140	0,3243	0,0825	0,0462	0,3315	0,3370	0,0992	0,3315	0,3370	0,0992	0,3315	0,3370	0,0992	0,3315	0,3370	0,0992	0,3315	
	MC	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2		
$C_{1,10}$	Media	5,5987	5,5694	5,5753	5,3710	5,1629	5,5799	5,5825	5,3982	5,5799	5,5825	5,3982	5,5799	5,5825	5,3982	5,5799	5,5825	5,3982	
	Mediana	5,6011	5,5725	5,5764	5,4552	5,0806	5,5817	5,5871	5,4662	5,5817	5,5871	5,4662	5,5817	5,5871	5,4662	5,5817	5,5871	5,4662	
	SD	0,1390	0,1774	0,1679	0,3636	0,3421	0,1617	0,1589	0,3413	0,1617	0,1589	0,3413	0,1617	0,1589	0,3413	0,1617	0,1589	0,3413	0,1617
	MAD	0,1370	0,1603	0,1582	0,3969	0,2997	0,1523	0,1484	0,3499	0,1523	0,1484	0,3499	0,1523	0,1484	0,3499	0,1523	0,1484	0,3499	0,1523
ECM	0,3778	0,3557	0,3592	0,2697	0,1435	0,3624	0,3645	0,2749	0,3624	0,3645	0,2749	0,3624	0,3645	0,2749	0,3624	0,3645	0,2749	0,3624	
	MC	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2		
$C_{1,12}$	Media	5,6053	5,5753	5,5834	5,5405	5,3879	5,5830	5,5893	5,5553	5,5830	5,5893	5,5553	5,5830	5,5893	5,5553	5,5830	5,5893	5,5553	
	Mediana	5,6118	5,5778	5,5849	5,5878	5,4573	5,5832	5,5938	5,5949	5,5832	5,5938	5,5949	5,5832	5,5938	5,5949	5,5832	5,5938	5,5949	5,5832
	SD	0,1900	0,1764	0,1680	0,2904	0,3846	0,1692	0,1617	0,2721	0,1692	0,1617	0,2721	0,1692	0,1617	0,2721	0,1692	0,1617	0,2721	0,1692
	MAD	0,1489	0,1735	0,1696	0,2256	0,4559	0,1638	0,1669	0,2166	0,1638	0,1669	0,2166	0,1638	0,1669	0,2166	0,1638	0,1669	0,2166	0,1638
ECM	0,4025	0,3621	0,3685	0,3764	0,2982	0,3685	0,3734	0,3824	0,3685	0,3734	0,3824	0,3685	0,3734	0,3824	0,3685	0,3734	0,3824	0,3685	
	MC	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2		

Cuadro 8.1: Medidas resumen del estimador de mínimos cuadrados y de los M -, MM - y WMM -estimadores calculados a partir del estimador inicial basado en las mediana bajo C_0 y bajo las contaminaciones $C_{1,m}$ para $m = 5, 7, 10$ y 12 cuando todas las respuestas son observadas, es decir cuando la probabilidad de pérdida es $P_{comp} = 0$.

	$\theta_1 = 5$										$\theta_2 = 2$																			
	Eficiencia					85 %					90 %					Eficiencia					85 %					90 %				
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{2,5}$	Media	5,4958	5,3405	5,3613	5,0029	5,0024	5,3809	5,4092	5,0024	5,3816	1,8616	1,9031	1,8976	1,9995	2,0000	1,8921	1,8845	1,9997	1,9997											
	Mediana	5,4983	5,4408	5,4556	5,0027	5,0016	5,4616	5,4762	5,0016	1,8603	1,8667	1,8646	2,0002	1,9994	2,0000	1,8640	1,8627	1,9997	1,9997											
	SD	0,1648	0,2901	0,2808	0,1835	0,1927	0,2675	0,2489	0,1803	0,0320	0,0775	0,0744	0,0483	0,0530	0,0530	0,0705	0,0647	0,0475	0,0475											
	MAD	0,1105	0,1827	0,1616	0,1786	0,1867	0,1492	0,1366	0,1704	0,0138	0,0236	0,0205	0,0459	0,0499	0,0499	0,0192	0,0178	0,0442	0,0442											
	ECM	0,2729	0,2000	0,2093	0,0336	0,0371	0,2166	0,2294	0,0325	0,0202	0,0154	0,0160	0,0023	0,0028	0,0028	0,0166	0,0175	0,0023	0,0023											
$C_{2,7}$	Media	5,5095	5,4259	5,4442	5,0049	5,0041	5,4545	5,4732	5,0040	1,8582	1,8789	1,8742	1,9992	1,9996	1,9996	1,8715	1,8667	1,9994	1,9994											
	Mediana	5,5114	5,4758	5,4834	5,0036	5,0014	5,4902	5,4983	4,9998	1,8576	1,8614	1,8602	1,9993	1,9993	1,9993	1,8595	1,8587	1,9993	1,9993											
	SD	0,1123	0,2335	0,2194	0,1896	0,1954	0,2077	0,1862	0,1867	0,0144	0,0590	0,0544	0,0491	0,0520	0,0520	0,0506	0,0434	0,0484	0,0484											
	MAD	0,1101	0,1480	0,1402	0,1802	0,1901	0,1330	0,1256	0,1793	0,0142	0,0191	0,0181	0,0475	0,0508	0,0508	0,0172	0,0161	0,0461	0,0461											
	ECM	0,2722	0,2358	0,2454	0,0359	0,0382	0,2497	0,2586	0,0348	0,0203	0,0182	0,0188	0,0024	0,0027	0,0027	0,0191	0,0197	0,0023	0,0023											
$C_{2,10}$	Media	5,5163	5,4850	5,4919	5,0398	5,0186	5,4948	5,4985	5,0448	1,8564	1,8617	1,8603	1,9903	1,9959	1,9959	1,8597	1,8590	1,9890	1,9890											
	Mediana	5,5177	5,4939	5,4991	5,0078	5,0076	5,4981	5,5028	5,0129	1,8560	1,8585	1,8578	1,9961	1,9980	1,9980	1,8581	1,8575	1,9953	1,9953											
	SD	0,1176	0,1659	0,1568	0,2441	0,2193	0,1492	0,1435	0,2437	0,0151	0,0326	0,0291	0,0618	0,0558	0,0558	0,0272	0,0250	0,0616	0,0616											
	MAD	0,1160	0,1377	0,1357	0,2015	0,1947	0,1310	0,1307	0,2010	0,0150	0,0178	0,0174	0,0509	0,0514	0,0514	0,0167	0,0165	0,0507	0,0507											
	ECM	0,2803	0,2627	0,2665	0,0611	0,0484	0,2670	0,2691	0,0613	0,0209	0,0202	0,0204	0,0039	0,0031	0,0031	0,0204	0,0205	0,0039	0,0039											
$C_{2,12}$	Media	5,5170	5,4897	5,4942	5,1601	5,0807	5,4961	5,4991	5,1680	1,8561	1,8597	1,8589	1,9582	1,9801	1,9801	1,8587	1,8582	1,9563	1,9563											
	Mediana	5,5198	5,4940	5,5001	5,0904	5,0387	5,5022	5,5041	5,1018	1,8554	1,8584	1,8576	1,9772	1,9906	1,9906	1,8574	1,8571	1,9720	1,9720											
	SD	0,1219	0,1576	0,1489	0,3251	0,2834	0,1469	0,1420	0,3221	0,0156	0,0252	0,0222	0,0829	0,0711	0,0711	0,0222	0,0208	0,0820	0,0820											
	MAD	0,1222	0,1452	0,1413	0,3383	0,2360	0,1411	0,1376	0,3453	0,0159	0,0188	0,0181	0,0872	0,0602	0,0602	0,0180	0,0178	0,0927	0,0927											
	ECM	0,2822	0,2646	0,2664	0,1312	0,0867	0,2677	0,2693	0,1319	0,0210	0,0203	0,0204	0,0086	0,0054	0,0054	0,0205	0,0205	0,0086	0,0086											

Cuadro 8.2: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir del estimador inicial basado en las medianas bajo las contaminaciones $C_{2,m}$ para $m = 5, 7, 10$ y 12 cuando todas las respuestas son observadas, es decir cuando la probabilidad de pérdida es p_{comp} .

		$\theta_1 = 5$										$\theta_2 = 2$											
		85 %					90 %					85 %					90 %						
	Eficiencia	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2		
$C_{3,5}$	Media	5.3541	5.1997	5.2170	5.0041	5.0030	5.2257	5.2392	5.0032	5.3070	5.3127	5.0018	5.0015	5.3070	5.3127	5.0018	5.0018	5.3070	5.3127	5.0018	5.0018	5.3070	5.3127
	Mediana	5.3547	5.2855	5.2983	5.0018	5.0015	5.3070	5.3127	5.0018	5.3070	5.3127	5.0018	5.0015	5.3070	5.3127	5.0018	5.0018	5.3070	5.3127	5.0018	5.0018	5.3070	5.3127
	SD	0.0914	0.2354	0.2307	0.1851	0.1934	0.2276	0.2216	0.1813	0.2276	0.2216	0.1813	0.1934	0.2276	0.2216	0.1813	0.1813	0.2276	0.2216	0.1813	0.1934	0.2276	0.2216
	MAD	0.0795	0.1860	0.1643	0.1812	0.1865	0.1514	0.1372	0.1708	0.1514	0.1372	0.1708	0.1865	0.1514	0.1372	0.1708	0.1708	0.1514	0.1372	0.1708	0.1865	0.1514	0.1372
ECM	0.1338	0.0952	0.1003	0.0342	0.0374	0.1027	0.1063	0.0328	0.1027	0.1063	0.0328	0.0374	0.1027	0.1063	0.0328	0.0328	0.1027	0.1063	0.0328	0.0374	0.1027	0.1063	
$C_{3,7}$	Media	5.3504	5.2396	5.2551	5.0049	5.0044	5.2606	5.2747	5.0039	5.3266	5.3357	5.0007	5.0007	5.3266	5.3357	5.0007	5.0007	5.3266	5.3357	5.0007	5.0007	5.3266	5.3357
	Mediana	5.3582	5.3109	5.3220	5.0003	5.0010	5.3266	5.3357	5.0007	5.3266	5.3357	5.0007	5.0010	5.3266	5.3357	5.0007	5.0007	5.3266	5.3357	5.0007	5.0007	5.3266	5.3357
	SD	0.1677	0.2272	0.2182	0.1912	0.1965	0.2159	0.2091	0.1870	0.2159	0.2091	0.1870	0.1965	0.2159	0.2091	0.1870	0.1870	0.2159	0.2091	0.1870	0.1965	0.2159	0.2091
	MAD	0.0804	0.1426	0.1295	0.1851	0.1931	0.1229	0.1145	0.1801	0.1229	0.1145	0.1801	0.1931	0.1229	0.1145	0.1801	0.1801	0.1229	0.1145	0.1801	0.1931	0.1229	0.1145
ECM	0.1509	0.1090	0.1126	0.0365	0.0386	0.1144	0.1192	0.0349	0.1144	0.1192	0.0349	0.0386	0.1144	0.1192	0.0349	0.0349	0.1144	0.1192	0.0349	0.0386	0.1144	0.1192	
$C_{3,10}$	Media	5.3379	5.3022	5.3107	5.0082	5.0068	5.3151	5.3212	5.0061	5.3455	5.3464	4.9961	4.9961	5.3455	5.3464	4.9961	4.9961	5.3455	5.3464	4.9961	4.9961	5.3455	5.3464
	Mediana	5.3568	5.3420	5.3456	4.9993	5.0005	5.3455	5.3464	4.9961	5.3455	5.3464	4.9961	4.9961	5.3455	5.3464	4.9961	4.9961	5.3455	5.3464	4.9961	4.9961	5.3455	5.3464
	SD	0.2801	0.1919	0.1819	0.2028	0.2023	0.1755	0.1675	0.1951	0.1755	0.1675	0.1951	0.2023	0.1755	0.1675	0.1951	0.1951	0.1755	0.1675	0.1951	0.2023	0.1755	0.1675
	MAD	0.0866	0.1137	0.1089	0.1911	0.1912	0.1070	0.1034	0.1874	0.1070	0.1034	0.1874	0.1912	0.1070	0.1034	0.1874	0.1874	0.1070	0.1034	0.1874	0.1912	0.1070	0.1034
ECM	0.1926	0.1281	0.1296	0.0412	0.0409	0.1301	0.1312	0.0381	0.1301	0.1312	0.0381	0.0409	0.1301	0.1312	0.0381	0.0381	0.1301	0.1312	0.0381	0.0409	0.1301	0.1312	
$C_{3,12}$	Media	5.3557	5.3291	5.3308	5.0141	5.0105	5.3373	5.3401	5.0125	5.3461	5.3485	5.0021	5.0021	5.3461	5.3485	5.0021	5.0021	5.3461	5.3485	5.0021	5.0021	5.3461	5.3485
	Mediana	5.3566	5.3438	5.3473	5.0055	5.0116	5.3461	5.3485	5.0021	5.3461	5.3485	5.0021	5.0021	5.3461	5.3485	5.0021	5.0021	5.3461	5.3485	5.0021	5.0021	5.3461	5.3485
	SD	0.0913	0.1538	0.1498	0.2183	0.2091	0.1368	0.1293	0.2068	0.1368	0.1293	0.2068	0.2091	0.1368	0.1293	0.2068	0.2068	0.1368	0.1293	0.2068	0.2091	0.1368	0.1293
	MAD	0.0921	0.1155	0.1137	0.2056	0.2015	0.1076	0.1064	0.1952	0.1076	0.1064	0.1952	0.2015	0.1076	0.1064	0.1952	0.1952	0.1076	0.1064	0.1952	0.2015	0.1076	0.1064
ECM	0.1348	0.1320	0.1318	0.0478	0.0438	0.1324	0.1324	0.0429	0.1324	0.1324	0.0429	0.0438	0.1324	0.1324	0.0429	0.0429	0.1324	0.1324	0.0429	0.0438	0.1324	0.1324	
$C_{3,12}$	Media	1.9035	1.9096	1.9093	1.9969	1.9977	1.9078	1.9070	1.9973	1.9046	1.9043	1.9987	1.9972	1.9046	1.9043	1.9987	1.9972	1.9046	1.9043	1.9987	1.9972	1.9046	1.9043
	Mediana	1.9033	1.9109	1.9096	1.9999	1.9993	1.9087	1.9079	1.9996	1.9046	1.9043	1.9987	1.9993	1.9046	1.9043	1.9987	1.9993	1.9046	1.9043	1.9987	1.9993	1.9046	1.9043
	SD	0.0095	0.0615	0.0599	0.0485	0.0530	0.0590	0.0574	0.0476	0.0316	0.0307	0.0559	0.0535	0.0316	0.0307	0.0559	0.0535	0.0316	0.0307	0.0559	0.0535	0.0316	0.0307
	MAD	0.0094	0.0202	0.0174	0.0465	0.0496	0.0159	0.0142	0.0447	0.0202	0.0174	0.0465	0.0496	0.0159	0.0142	0.0447	0.0447	0.0202	0.0174	0.0465	0.0496	0.0159	0.0142
ECM	0.0094	0.0092	0.0092	0.0031	0.0029	0.0092	0.0093	0.0028	0.0092	0.0092	0.0031	0.0029	0.0092	0.0092	0.0031	0.0029	0.0028	0.0092	0.0092	0.0031	0.0029	0.0092	

Cuadro 8.3: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir del estimador inicial basado en las medianas bajo las contaminaciones $C_{3,m}$ para $m = 5, 7, 10$ y 12 cuando todas las respuestas son observadas, es decir cuando la probabilidad de pérdida es $P_{\text{comp}} = 0$.

	$\theta_1 = 5$											$\theta_2 = 2$																				
	Eficiencia					85 %						90 %					Eficiencia					85 %						90 %				
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	MC	M	MM	WMM ₁	WMM ₂	
$C_{4,5}$	Media	5.3039	5.1486	5.1592	5.0043	5.0030	5.1691	5.1794	5.0032									1.9180	1.9598	1.9570	1.9992	1.9999	1.9544	1.9517	1.9995							
	Mediana	5.3057	5.2228	5.2344	5.0027	5.0014	5.2487	5.2545	5.0018									1.9178	1.9258	1.9246	1.9997	1.9991	1.9232	1.9227	1.9995							
	SD	0.0765	0.2217	0.2184	0.1849	0.1931	0.2158	0.2122	0.1810									0.0072	0.0569	0.0561	0.0485	0.0530	0.0554	0.0545	0.0475							
	MAD	0.0753	0.1982	0.1837	0.1811	0.1844	0.1604	0.1458	0.1716									0.0070	0.0220	0.0191	0.0471	0.0496	0.0156	0.0143	0.0449							
	ECM	0.0982	0.0712	0.0730	0.0342	0.0373	0.0751	0.0771	0.0327									0.0068	0.0049	0.0050	0.0024	0.0028	0.0052	0.0053	0.0023							
$C_{4,7}$	Media	5.3048	5.1825	5.1950	5.0049	5.0044	5.2025	5.2051	5.0038									1.9178	1.9506	1.9474	1.9992	1.9996	1.9454	1.9448	1.9995							
	Mediana	5.3067	5.2535	5.2654	5.0001	5.0010	5.2712	5.2729	5.0013									1.9176	1.9227	1.9215	1.9999	1.9990	1.9210	1.9208	1.9993							
	SD	0.0815	0.2168	0.2128	0.1915	0.1963	0.2071	0.2052	0.1869									0.0077	0.0554	0.0542	0.0495	0.0520	0.0528	0.0523	0.0484							
	MAD	0.0717	0.1551	0.1349	0.1864	0.1939	0.1224	0.1190	0.1807									0.0068	0.0149	0.0130	0.0479	0.0511	0.0117	0.0114	0.0466							
	ECM	0.0995	0.0803	0.0833	0.0366	0.0385	0.0838	0.0842	0.0349									0.0058	0.0055	0.0057	0.0024	0.0027	0.0058	0.0058	0.0023							
$C_{4,10}$	Media	5.2975	5.2360	5.2452	5.0097	5.0083	5.2507	5.2512	5.0076									1.9190	1.9359	1.9336	1.9979	1.9983	1.9323	1.9321	1.9984							
	Mediana	5.3051	5.2878	5.2925	4.9994	5.0007	5.2924	5.2943	4.9978									1.9177	1.9193	1.9188	1.9992	1.9986	1.9189	1.9187	1.9985							
	SD	0.1892	0.1952	0.1864	0.2040	0.2031	0.1798	0.1804	0.1958									0.0295	0.0476	0.0451	0.0521	0.0522	0.0438	0.0437	0.0501							
	MAD	0.0790	0.1185	0.1097	0.1925	0.1935	0.1046	0.1037	0.1886									0.0074	0.0112	0.0104	0.0492	0.0497	0.0099	0.0098	0.0475							
	ECM	0.1242	0.0938	0.0949	0.0417	0.0413	0.0951	0.0956	0.0384									0.0074	0.0064	0.0064	0.0027	0.0027	0.0065	0.0065	0.0025							
$C_{4,12}$	Media	5.3041	5.2647	5.2674	5.0101	5.0105	5.2719	5.2768	5.0089									1.9180	1.9280	1.9273	1.9979	1.9978	1.9263	1.9249	1.9981							
	Mediana	5.3066	5.2907	5.2918	5.0035	5.0122	5.2953	5.2972	4.9998									1.9176	1.9191	1.9190	1.9996	1.9975	1.9187	1.9185	1.9996							
	SD	0.0834	0.1664	0.1613	0.2171	0.2108	0.1545	0.1443	0.2053									0.0079	0.0378	0.0366	0.0557	0.0539	0.0351	0.0319	0.0525							
	MAD	0.0821	0.1118	0.1118	0.2125	0.2023	0.1029	0.1001	0.1997									0.0077	0.0106	0.0106	0.0534	0.0516	0.0098	0.0095	0.0499							
	ECM	0.0994	0.0977	0.0975	0.0472	0.0445	0.0978	0.0974	0.0422									0.0068	0.0066	0.0066	0.0031	0.0029	0.0067	0.0067	0.0028							

Cuadro 8.4: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir del estimador inicial basado en las medianas bajo las contaminaciones $C_{4,m}$ para $m = 5, 7, 10$ y 12 cuando todas las respuestas son observadas, es decir cuando la probabilidad de pérdida es p_{comp} .

$\theta_1 = 5$										$\theta_2 = 2$									
85 %										90 %									
Eficiencia										Eficiencia									
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂		MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂
C_0	Media	4,9994	4,9991	4,9989	4,9992	4,9976	4,9988	4,9988	4,9989	Media	2,0005	2,0006	2,0005	2,0015	2,0028	2,0006	2,0006	2,0015	2,0015
	Mediana	4,9992	5,0022	5,0040	5,0009	5,0064	5,0019	5,0049	5,0005	Mediana	1,9993	1,9992	1,9991	1,9988	1,9991	1,9994	1,9992	2,0000	2,0000
	SD	0,1833	0,2038	0,1959	0,2127	0,2348	0,1960	0,1908	0,2074	SD	0,0465	0,0512	0,0493	0,0588	0,0703	0,0481	0,0572	0,0572	0,0572
$C_{1,5}$	MAD	0,1816	0,1982	0,1961	0,2116	0,2309	0,1925	0,1920	0,2060	MAD	0,0447	0,0494	0,0482	0,0575	0,0674	0,0481	0,0468	0,0542	0,0542
	ECM	0,0336	0,0415	0,0383	0,0452	0,0551	0,0384	0,0364	0,0430	ECM	0,0022	0,0026	0,0024	0,0035	0,0049	0,0024	0,0023	0,0033	0,0033
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{1,5}$	Media	5,6001	5,4569	5,4887	5,0291	5,0058	5,5068	5,5344	5,0426	Media	1,8321	1,8683	1,8600	1,9929	1,9993	1,8552	1,8479	1,9894	1,9894
	Mediana	5,5969	5,5454	5,5640	5,0173	5,0040	5,5734	5,5814	5,0265	Mediana	1,8324	1,8337	1,8319	1,9938	1,9982	1,8317	1,8312	1,9923	1,9923
	SD	0,1439	0,3361	0,3144	0,2443	0,2259	0,3007	0,2714	0,2506	SD	0,0218	0,0890	0,0820	0,0634	0,0613	0,0773	0,0680	0,0649	0,0649
$C_{1,7}$	MAD	0,1450	0,2301	0,2034	0,2255	0,2266	0,1843	0,1753	0,2329	MAD	0,0220	0,0335	0,0307	0,0561	0,0586	0,0279	0,0265	0,0588	0,0588
	ECM	0,3808	0,3216	0,3375	0,0605	0,0510	0,3472	0,3592	0,0645	ECM	0,0287	0,0253	0,0263	0,0041	0,0038	0,0269	0,0278	0,0043	0,0043
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{1,10}$	Media	5,6345	5,5779	5,5970	5,1202	5,0270	5,5994	5,6121	5,1540	Media	1,8230	1,8335	1,8290	1,9688	1,9938	1,8287	1,8259	1,9600	1,9600
	Mediana	5,6322	5,6059	5,6154	5,0678	5,0124	5,6123	5,6243	5,1001	Mediana	1,8229	1,8233	1,8221	1,9823	1,9944	1,8226	1,8219	1,9733	1,9733
	SD	0,1497	0,2548	0,2310	0,3288	0,2499	0,2246	0,2055	0,3330	SD	0,0220	0,0589	0,0509	0,0846	0,0649	0,0491	0,0423	0,0855	0,0855
$C_{1,12}$	MAD	0,1536	0,1860	0,1793	0,2984	0,2316	0,1727	0,1713	0,3250	MAD	0,0229	0,0271	0,0263	0,0770	0,0595	0,0256	0,0254	0,0892	0,0892
	ECM	0,4249	0,3988	0,4097	0,1224	0,0631	0,4097	0,4168	0,1345	ECM	0,0318	0,0312	0,0318	0,0081	0,0043	0,0318	0,0321	0,0089	0,0089
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{1,10}$	Media	5,6595	5,6328	5,6407	5,4310	5,1971	5,6396	5,6458	5,4508	Media	1,8165	1,8180	1,8167	1,8841	1,9498	1,8172	1,8163	1,8792	1,8792
	Mediana	5,6614	5,6325	5,6362	5,5055	5,1364	5,6362	5,6415	5,5228	Mediana	1,8152	1,8171	1,8167	1,8457	1,9652	1,8169	1,8163	1,8449	1,8449
	SD	0,2028	0,2126	0,1962	0,3876	0,3718	0,1974	0,1841	0,3704	SD	0,0394	0,0361	0,0312	0,0978	0,0928	0,0324	0,0282	0,0924	0,0924
$C_{1,12}$	MAD	0,1693	0,2057	0,1867	0,3886	0,3468	0,1949	0,1817	0,3569	MAD	0,0243	0,0289	0,0270	0,0742	0,0917	0,0278	0,0264	0,0660	0,0660
	ECM	0,4760	0,4456	0,4490	0,3358	0,1770	0,4480	0,4510	0,3403	ECM	0,0352	0,0344	0,0346	0,0230	0,0111	0,0345	0,0346	0,0231	0,0231
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{1,12}$	Media	5,6715	5,6389	5,6498	5,6061	5,4326	5,6426	5,6540	5,6073	Media	1,8131	1,8158	1,8144	1,8347	1,8871	1,8157	1,8140	1,8350	1,8350
	Mediana	5,6689	5,6287	5,6449	5,6390	5,4857	5,6313	5,6516	5,6392	Mediana	1,8136	1,8161	1,8154	1,8190	1,8521	1,8164	1,8149	1,8192	1,8192
	SD	0,1680	0,2245	0,2081	0,3144	0,4096	0,2134	0,1972	0,3048	SD	0,0242	0,0334	0,0302	0,0715	0,1005	0,0329	0,0285	0,0695	0,0695
$C_{1,12}$	MAD	0,1745	0,2219	0,2099	0,2680	0,4820	0,2135	0,2045	0,2589	MAD	0,0253	0,0313	0,0298	0,0408	0,1020	0,0304	0,0288	0,0397	0,0397
	ECM	0,4791	0,4586	0,4655	0,4660	0,3547	0,4584	0,4666	0,4616	ECM	0,0355	0,0351	0,0354	0,0324	0,0228	0,0351	0,0354	0,0230	0,0230
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	

Cuadro 8.5: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir del estimador inicial basado en las medianas bajo C_0 y bajo las contaminaciónes $C_{1,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es P_{log} .

	$\theta_1 = 5$					$\theta_2 = 2$							
	Eficiencia		85 %		90 %		Eficiencia		85 %		90 %		
	MC	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{2,5}$	Media	5.5396	5.4161	5.0058	5.0051	5.4272	5.4596	5.0047	1.8550	1.8952	1.8867	1.9989	1.9995
	Mediana	5.5359	5.4826	5.4994	5.0116	5.5047	5.5157	5.0067	1.8550	1.8612	1.8587	1.9959	1.9984
	SD	0.1206	0.3134	0.2950	0.2150	0.2871	0.2584	0.2113	0.0154	0.0794	0.0740	0.0561	0.0612
	MAD	0.1203	0.2027	0.1747	0.2141	0.1682	0.1501	0.2106	0.0155	0.0264	0.0225	0.0536	0.0590
	ECM	0.3057	0.2453	0.2601	0.0462	0.2649	0.2780	0.0446	0.0213	0.0173	0.0183	0.0031	0.0037
$C_{2,7}$	Media	5.5486	5.4864	5.0115	5.0076	5.5109	5.5212	5.0090	1.8527	1.8667	1.8625	1.9973	1.9987
	Mediana	5.5437	5.5202	5.5291	5.0094	5.5291	5.5320	5.0061	1.8531	1.8553	1.8543	1.9974	1.9981
	SD	0.1241	0.2334	0.2107	0.2256	0.2013	0.1819	0.2203	0.0158	0.0535	0.0465	0.0583	0.0598
	MAD	0.1270	0.1604	0.1509	0.2191	0.1489	0.1401	0.2188	0.0163	0.0207	0.0194	0.0543	0.0567
	ECM	0.3163	0.2910	0.2984	0.0510	0.3015	0.3047	0.0486	0.0219	0.0206	0.0211	0.0034	0.0036
$C_{2,10}$	Media	5.5550	5.5247	5.0575	5.0318	5.5326	5.5385	5.0583	1.8511	1.8558	1.8545	1.9858	1.9925
	Mediana	5.5548	5.5307	5.5367	5.0303	5.5356	5.5411	5.0286	1.8509	1.8531	1.8527	1.9916	1.9958
	SD	0.1319	0.1853	0.1754	0.2845	0.1671	0.1578	0.2798	0.0168	0.0328	0.0299	0.0715	0.0649
	MAD	0.1373	0.1627	0.1582	0.2477	0.1577	0.1524	0.2447	0.0173	0.0208	0.0204	0.0606	0.0583
	ECM	0.3254	0.3096	0.3141	0.0842	0.3116	0.3149	0.0816	0.0225	0.0219	0.0221	0.0053	0.0043
$C_{2,12}$	Media	5.5544	5.5346	5.5396	5.2067	5.5409	5.5433	5.2063	1.8511	1.8537	1.8529	1.9479	1.9737
	Mediana	5.5545	5.5345	5.5430	5.1532	5.5332	5.5403	5.1572	1.8509	1.8530	1.8518	1.9581	1.9822
	SD	0.1360	0.1924	0.1834	0.3716	0.1733	0.1679	0.3617	0.0174	0.0291	0.0266	0.0916	0.0809
	MAD	0.1436	0.1815	0.1812	0.4197	0.1713	0.1661	0.4050	0.0185	0.0233	0.0230	0.1164	0.0761
	ECM	0.3258	0.3228	0.3248	0.1807	0.3225	0.3234	0.1733	0.0225	0.0223	0.0223	0.0111	0.0072

Cuadro 8.6: Medidas resumen del estimador de mínimos cuadrados y de los M -, MM - y WMM -estimadores calculados a partir del estimador inicial basado en las medianas bajo las contaminaciones $C_{2,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es p_{Log} .

		$\theta_1 = 5$										$\theta_2 = 2$										
		85 %					90 %					85 %					90 %					
	Eficiencia	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	
$C_{3,5}$	Media	5.3709	5.2253	5.2374	5.0076	5.0058	5.2419	5.2595	5.0058	5.2419	5.2595	5.0058	5.0058	1.9021	1.9403	1.9372	1.9984	1.9993	1.9360	1.9316	1.9989	1.9989
	Mediana	5.3727	5.3158	5.3226	5.0081	5.0067	5.3262	5.3338	5.0081	5.3262	5.3338	5.0081	5.0081	1.9017	1.9078	1.9069	1.9951	1.9978	1.9066	1.9058	1.9961	1.9961
	SD	0.1051	0.2600	0.2509	0.2177	0.2262	0.2501	0.2418	0.2128	0.2501	0.2418	0.2128	0.2128	0.0111	0.0651	0.0629	0.0567	0.0614	0.0625	0.0604	0.0554	0.0554
	MAD	0.0900	0.1759	0.1639	0.2132	0.2271	0.1538	0.1382	0.2105	0.1538	0.1382	0.2105	0.2105	0.0094	0.0187	0.0174	0.0535	0.0586	0.0161	0.0145	0.0533	0.0533
ECM	0.1486	0.1183	0.1193	0.0474	0.0511	0.1210	0.1258	0.0453	0.1210	0.1258	0.0453	0.0453	0.0097	0.0078	0.0079	0.0032	0.0038	0.0080	0.0083	0.0031	0.0031	
$C_{3,7}$	Media	5.3709	5.2864	5.3010	5.0132	5.0105	5.3091	5.3176	5.0102	5.3091	5.3176	5.0102	5.0102	1.9022	1.9242	1.9206	1.9971	1.9980	1.9183	1.9163	1.9979	1.9979
	Mediana	5.3730	5.3475	5.3545	5.0118	5.0056	5.3565	5.3600	5.0059	5.3565	5.3600	5.0059	5.0059	1.9016	1.9042	1.9035	1.9976	1.9974	1.9033	1.9029	1.9975	1.9975
	SD	0.1499	0.2309	0.2209	0.2295	0.2309	0.2103	0.1993	0.2215	0.2103	0.1993	0.2215	0.2215	0.0229	0.0558	0.0532	0.0585	0.0604	0.0503	0.0472	0.0566	0.0566
	MAD	0.0930	0.1330	0.1238	0.2242	0.2279	0.1217	0.1162	0.2225	0.1217	0.1162	0.2225	0.2225	0.0096	0.0138	0.0131	0.0549	0.0576	0.0126	0.0119	0.0533	0.0533
ECM	0.1600	0.1352	0.1394	0.0528	0.0534	0.1397	0.1406	0.0491	0.1397	0.1406	0.0491	0.0491	0.0101	0.0089	0.0091	0.0034	0.0036	0.0092	0.0092	0.0032	0.0032	
$C_{3,10}$	Media	5.3642	5.3407	5.3505	5.0244	5.0191	5.3523	5.3570	5.0189	5.3523	5.3570	5.0189	5.0189	1.9033	1.9096	1.9073	1.9944	1.9959	1.9068	1.9057	1.9959	1.9959
	Mediana	5.3805	5.3674	5.3731	5.0098	5.0121	5.3721	5.3717	5.0119	5.3721	5.3717	5.0119	5.0119	1.9007	1.9021	1.9015	1.9969	1.9961	1.9016	1.9016	1.9960	1.9960
	SD	0.2239	0.1837	0.1683	0.2573	0.2449	0.1610	0.1515	0.2391	0.1610	0.1515	0.2391	0.2391	0.0373	0.0383	0.0336	0.0649	0.0617	0.0318	0.0295	0.0600	0.0600
	MAD	0.0973	0.1240	0.1187	0.2425	0.2404	0.1127	0.1113	0.2301	0.1127	0.1113	0.2301	0.2301	0.0100	0.0128	0.0123	0.0612	0.0595	0.0116	0.0115	0.0575	0.0575
ECM	0.1836	0.1498	0.1511	0.0667	0.0603	0.1500	0.1504	0.0575	0.1500	0.1504	0.0575	0.0575	0.0107	0.0096	0.0097	0.0042	0.0038	0.0097	0.0098	0.0036	0.0036	
$C_{3,12}$	Media	5.3731	5.3573	5.3598	5.0384	5.0284	5.3625	5.3645	5.0323	5.3625	5.3645	5.0323	5.0323	1.9017	1.9056	1.9049	1.9908	1.9935	1.9041	1.9037	1.9925	1.9925
	Mediana	5.3757	5.3692	5.3716	5.0194	5.0113	5.3687	5.3718	5.0152	5.3687	5.3718	5.0152	5.0152	1.9014	1.9020	1.9017	1.9958	1.9960	1.9021	1.9017	1.9967	1.9967
	SD	0.0992	0.1646	0.1599	0.2865	0.2679	0.1455	0.1460	0.2661	0.1455	0.1460	0.2661	0.2661	0.0103	0.0281	0.0269	0.0724	0.0671	0.0227	0.0235	0.0666	0.0666
	MAD	0.1023	0.1355	0.1323	0.2794	0.2574	0.1268	0.1261	0.2556	0.1268	0.1261	0.2556	0.2556	0.0106	0.0140	0.0136	0.0715	0.0660	0.0131	0.0130	0.0645	0.0645
ECM	0.1490	0.1547	0.1550	0.0835	0.0725	0.1526	0.1542	0.0718	0.1526	0.1542	0.0718	0.0718	0.0098	0.0097	0.0098	0.0053	0.0045	0.0097	0.0098	0.0045	0.0045	

Cuadro 8.7: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir del estimador inicial basado en las medianas bajo las contaminaciones $C_{3,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es P_{Log} .

	Eficiencia	$\theta_1 = 5$										$\theta_2 = 2$												
		85 %					90 %					85 %					90 %							
		M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂			
$C_{4,5}$	MC	5.3160	5.1850	5.0076	5.0057	5.1938	5.2034	5.0058	5.0058	5.1938	5.2034	5.0058	5.0058	1.9169	1.9555	1.9915	1.9994	1.9427	1.9395	1.9970	1.9979	1.9366	1.9357	1.9978
	Media	5.3187	5.2550	5.0086	5.0064	5.2740	5.2795	5.0082	5.0082	5.2740	5.2795	5.0082	5.0082	1.9165	1.9227	1.9214	1.9978	1.9227	1.9214	1.9947	1.9978	1.9208	1.9202	1.9963
	SD	0.0901	0.2448	0.2373	0.2258	0.2337	0.2295	0.2125	0.2125	0.2337	0.2295	0.2125	0.2125	0.0085	0.0613	0.0594	0.0612	0.0613	0.0594	0.0564	0.0612	0.0586	0.0573	0.0552
	MAD	0.0830	0.1824	0.1619	0.2165	0.1440	0.1355	0.2105	0.2105	0.1440	0.1355	0.2105	0.2105	0.0077	0.0181	0.0155	0.0538	0.0181	0.0155	0.0538	0.0589	0.0137	0.0128	0.0532
	ECM	0.1079	0.0886	0.0905	0.0472	0.0921	0.0940	0.0451	0.0451	0.0921	0.0940	0.0451	0.0451	0.0070	0.0057	0.0059	0.0032	0.0057	0.0059	0.0032	0.0037	0.0060	0.0061	0.0030
$C_{4,7}$	MC	5.3174	5.2183	5.0134	5.0108	5.2422	5.2461	5.0104	5.0104	5.2422	5.2461	5.0104	5.0104	1.9167	1.9427	1.9395	1.9970	1.9427	1.9395	1.9970	1.9979	1.9366	1.9357	1.9978
	Media	5.3208	5.2893	5.2956	5.0114	5.0059	5.3001	5.3023	5.0067	5.3001	5.3023	5.0067	5.0067	1.9163	1.9192	1.9186	1.9976	1.9192	1.9186	1.9976	1.9973	1.9182	1.9180	1.9972
	SD	0.0937	0.2307	0.2192	0.2303	0.2100	0.2069	0.2214	0.2214	0.2100	0.2069	0.2214	0.2214	0.0091	0.0561	0.0533	0.0585	0.0561	0.0533	0.0585	0.0604	0.0508	0.0498	0.0565
	MAD	0.0822	0.1387	0.1262	0.2271	0.1170	0.1153	0.2267	0.2267	0.1170	0.1153	0.2267	0.2267	0.0078	0.0131	0.0119	0.0559	0.0131	0.0119	0.0559	0.0576	0.0110	0.0108	0.0539
	ECM	0.1095	0.1008	0.1015	0.0531	0.1027	0.1033	0.0491	0.0491	0.1027	0.1033	0.0491	0.0491	0.0070	0.0064	0.0065	0.0034	0.0064	0.0065	0.0034	0.0036	0.0066	0.0066	0.0032
$C_{4,10}$	MC	5.3161	5.2750	5.0235	5.0181	5.2825	5.2880	5.0181	5.0181	5.2825	5.2880	5.0181	5.0181	1.9168	1.9278	1.9269	1.9946	1.9278	1.9269	1.9946	1.9962	1.9258	1.9244	1.9961
	Media	5.3241	5.3104	5.3125	5.0143	5.3118	5.3131	5.0127	5.0127	5.3118	5.3131	5.0127	5.0127	1.9159	1.9172	1.9170	1.9963	1.9172	1.9170	1.9963	1.9962	1.9170	1.9169	1.9963
	SD	0.1154	0.1902	0.1825	0.2583	0.1752	0.1653	0.2402	0.2402	0.1752	0.1653	0.2402	0.2402	0.0124	0.0423	0.0402	0.0652	0.0423	0.0402	0.0652	0.0619	0.0385	0.0359	0.0603
	MAD	0.0879	0.1169	0.1140	0.2464	0.1095	0.1052	0.2314	0.2314	0.1095	0.1052	0.2314	0.2314	0.0083	0.0110	0.0107	0.0611	0.0110	0.0107	0.0611	0.0598	0.0103	0.0098	0.0579
	ECM	0.1132	0.1118	0.1108	0.0672	0.1105	0.1102	0.0580	0.0580	0.1105	0.1102	0.0580	0.0580	0.0071	0.0070	0.0070	0.0043	0.0070	0.0070	0.0043	0.0038	0.0070	0.0070	0.0036
$C_{4,12}$	MC	5.3162	5.2952	5.3001	5.0331	5.3023	5.3040	5.0271	5.0271	5.3023	5.3040	5.0271	5.0271	1.9168	1.9226	1.9215	1.9923	1.9226	1.9215	1.9923	1.9941	1.9205	1.9204	1.9939
	Media	5.3196	5.3134	5.3162	5.0138	5.3120	5.3159	5.0113	5.0113	5.3120	5.3159	5.0113	5.0113	1.9164	1.9170	1.9167	1.9957	1.9170	1.9167	1.9957	1.9961	1.9171	1.9168	1.9968
	SD	0.0925	0.1782	0.1645	0.2861	0.1535	0.1501	0.2634	0.2634	0.1535	0.1501	0.2634	0.2634	0.0087	0.0349	0.0311	0.0721	0.0349	0.0311	0.0721	0.0665	0.0283	0.0274	0.0661
	MAD	0.0908	0.1260	0.1233	0.2759	0.1168	0.1162	0.2538	0.2538	0.1168	0.1162	0.2538	0.2538	0.0085	0.0119	0.0116	0.0694	0.0119	0.0116	0.0694	0.0661	0.0110	0.0110	0.0650
	ECM	0.1086	0.1189	0.1171	0.0828	0.1149	0.1149	0.0701	0.0701	0.1149	0.1149	0.0701	0.0701	0.0070	0.0072	0.0071	0.0053	0.0072	0.0071	0.0053	0.0045	0.0071	0.0071	0.0044

Cuadro 8.8: Medidas resumen del estimador de mínimos cuadrados y de los M -, MM - y WMM -estimadores calculados a partir del estimador inicial basado en las medianas bajo las contaminaciones $C_{4,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es p_{Log} .

		$\theta_1 = 5$										$\theta_2 = 2$											
		85 %					90 %					85 %					90 %						
	Eficiencia	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2		
C_0	Media	5,0004	4,9995	4,9995	4,9980	4,9969	4,9998	4,9998	4,9982	4,9998	4,9998	4,9982	4,9982	4,9998	4,9998	4,9982	4,9982	4,9998	4,9998	4,9982	4,9982		
	Mediana	4,9909	4,9956	4,9951	4,9949	4,9986	4,9960	4,9963	4,9916	4,9978	4,9978	4,9916	4,9916	4,9978	4,9978	4,9916	4,9916	4,9978	4,9978	4,9916	4,9916		
	SD	0,1642	0,1775	0,1741	0,1925	0,2134	0,1721	0,1700	0,1878	0,1721	0,1700	0,1878	0,1721	0,1700	0,1878	0,1721	0,1700	0,1878	0,1721	0,1700	0,1878	0,1721	
$C_{1,5}$	MAD	0,1578	0,1637	0,1606	0,1890	0,1997	0,1622	0,1600	0,1807	0,1622	0,1600	0,1807	0,1622	0,1600	0,1807	0,1622	0,1600	0,1807	0,1622	0,1600	0,1807	0,1622	
	ECM	0,0269	0,0315	0,0303	0,0370	0,0455	0,0296	0,0289	0,0352	0,0296	0,0289	0,0352	0,0296	0,0289	0,0352	0,0296	0,0289	0,0352	0,0296	0,0289	0,0352	0,0296	
	MC		M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	
$C_{1,7}$	Media	5,5764	5,5219	5,5394	5,1033	5,0227	5,5469	5,5583	5,1412	5,5469	5,5583	5,1412	5,5469	5,5583	5,1412	5,5469	5,5583	5,1412	5,5469	5,5583	5,1412	5,5469	
	Mediana	5,5788	5,5482	5,5577	5,0466	5,0152	5,5603	5,5645	5,0874	5,5603	5,5645	5,0874	5,5603	5,5645	5,0874	5,5603	5,5645	5,0874	5,5603	5,5645	5,0874	5,5603	
	SD	0,1453	0,2347	0,2121	0,3000	0,2290	0,1986	0,1780	0,3079	0,1986	0,1780	0,3079	0,1986	0,1780	0,3079	0,1986	0,1780	0,3079	0,1986	0,1780	0,3079	0,1986	0,1780
$C_{1,10}$	MAD	0,1416	0,1788	0,1686	0,2542	0,2091	0,1630	0,1555	0,2867	0,1630	0,1555	0,2867	0,1630	0,1555	0,2867	0,1630	0,1555	0,2867	0,1630	0,1555	0,2867	0,1630	0,1555
	ECM	0,3534	0,3274	0,3359	0,1006	0,0529	0,3385	0,3433	0,1146	0,3385	0,3433	0,1146	0,3385	0,3433	0,1146	0,3385	0,3433	0,1146	0,3385	0,3433	0,1146	0,3385	0,3433
	MC		M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	
$C_{1,12}$	Media	5,6114	5,5682	5,5777	5,5535	5,3981	5,5755	5,5832	5,5621	5,5755	5,5832	5,5621	5,5755	5,5832	5,5621	5,5755	5,5832	5,5621	5,5755	5,5832	5,5621	5,5755	
	Mediana	5,6090	5,5749	5,5807	5,5885	5,4520	5,5760	5,5844	5,5982	5,5760	5,5844	5,5982	5,5760	5,5844	5,5982	5,5760	5,5844	5,5982	5,5760	5,5844	5,5982	5,5760	
	SD	0,1623	0,2027	0,1947	0,3020	0,3995	0,1916	0,1868	0,2868	0,1916	0,1868	0,2868	0,1916	0,1868	0,2868	0,1916	0,1868	0,2868	0,1916	0,1868	0,2868	0,1916	0,1868
$C_{1,12}$	MAD	0,1638	0,1901	0,1838	0,2563	0,4818	0,1830	0,1818	0,2486	0,1830	0,1818	0,2486	0,1830	0,1818	0,2486	0,1830	0,1818	0,2486	0,1830	0,1818	0,2486	0,1830	0,1818
	ECM	0,4001	0,3639	0,3716	0,3975	0,3179	0,3679	0,3750	0,3981	0,3679	0,3750	0,3981	0,3679	0,3750	0,3981	0,3679	0,3750	0,3981	0,3679	0,3750	0,3981	0,3679	0,3750
	MC		M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	M	MM	WMN1	WMN2	

Cuadro 8.9: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WM-$ estimadores calculados a partir del estimador inicial basado en las medianas bajo C_0 y bajo las contaminaciónes $C_{1,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es P_{cos} .

	$\theta_1 = 5$						$\theta_2 = 2$						
	Eficiencia		85 %		90 %		Eficiencia		85 %		90 %		
	MC	MM	MM	WMM ₁	WMM ₂	WMM ₁	MC	MM	MM	WMM ₁	WMM ₂	MM	WMM ₁
$C_{2,5}$	Media	5.5019	5.3630	5.3956	5.0044	5.0037	1.8600	1.8964	1.8877	1.9988	1.9996	1.8847	1.8777
	Mediana	5.5037	5.4468	5.4620	5.0016	5.0019	1.8596	1.8659	1.8639	1.9993	1.9987	1.8634	1.8616
	SD	0.1205	0.2836	0.2620	0.1952	0.2077	0.0154	0.0737	0.0673	0.0523	0.0576	0.0646	0.0588
	MAD	0.1175	0.1942	0.1689	0.1948	0.2052	0.0148	0.0258	0.0218	0.0500	0.0525	0.0210	0.0185
	ECM	0.2664	0.2121	0.2251	0.0381	0.0431	0.0198	0.0162	0.0172	0.0027	0.0033	0.0175	0.0184
$C_{2,7}$	Media	5.5111	5.4421	5.4622	5.0059	5.0058	1.8577	1.8742	1.8691	1.9986	1.9990	1.8665	1.8632
	Mediana	5.5122	5.4827	5.4911	5.0031	5.0055	1.8574	1.8602	1.8593	1.9984	1.9997	1.8590	1.8585
	SD	0.1241	0.2318	0.2144	0.2006	0.2088	0.0159	0.0557	0.0497	0.0534	0.0562	0.0432	0.0368
	MAD	0.1202	0.1501	0.1455	0.2000	0.2054	0.0156	0.0195	0.0187	0.0500	0.0524	0.0175	0.0172
	ECM	0.2766	0.2492	0.2595	0.0402	0.0436	0.0205	0.0189	0.0196	0.0029	0.0032	0.0197	0.0201
$C_{2,10}$	Media	5.5162	5.4848	5.4920	5.0457	5.0236	1.8562	1.8610	1.8596	1.9885	1.9944	1.8595	1.8586
	Mediana	5.5150	5.4877	5.4950	5.0135	5.0043	1.8561	1.8591	1.8581	1.9958	1.9988	1.8577	1.8569
	SD	0.1324	0.1748	0.1640	0.2538	0.2323	0.0171	0.0305	0.0272	0.0645	0.0592	0.0269	0.0245
	MAD	0.1287	0.1596	0.1550	0.2140	0.2049	0.0167	0.0205	0.0202	0.0535	0.0521	0.0191	0.0187
	ECM	0.2840	0.2656	0.2689	0.0664	0.0545	0.0210	0.0203	0.0204	0.0043	0.0035	0.0205	0.0206
$C_{2,12}$	Media	5.5182	5.4860	5.4915	5.1745	5.0872	1.8558	1.8598	1.8592	1.9547	1.9784	1.8587	1.8581
	Mediana	5.5194	5.4837	5.4914	5.1178	5.0410	1.8553	1.8597	1.8583	1.9688	1.9886	1.8590	1.8576
	SD	0.1349	0.1818	0.1785	0.3474	0.2979	0.0173	0.0258	0.0262	0.0877	0.0749	0.0217	0.0209
	MAD	0.1341	0.1685	0.1663	0.3897	0.2615	0.0173	0.0217	0.0214	0.1056	0.0648	0.0210	0.0206
	ECM	0.2867	0.2693	0.2735	0.1511	0.0963	0.0211	0.0203	0.0205	0.0097	0.0061	0.0204	0.0206

Cuadro 8.10: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir del estimador inicial basado en las medianas bajo las contaminaciones $C_{2,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es p_{\cos} .

	Eficiencia	$\theta_1 = 5$						$\theta_2 = 2$					
		85 %			90 %			85 %			90 %		
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂
$C_{3,5}$	Media	5.3552	5.1879	5.2042	5.0050	5.0041	5.2145	5.2310	5.0038				
	Mediana	5.3566	5.2741	5.2915	4.9991	5.0012	5.3036	5.3129	5.0027				
	SD	0.0916	0.2485	0.2415	0.1963	0.2081	0.2384	0.2325	0.1922				
	MAD	0.0871	0.2222	0.1841	0.1954	0.2063	0.1710	0.1507	0.1947				
ECM	0.1345	0.0970	0.1000	0.0385	0.0433	0.1028	0.1074	0.0369					
$C_{3,7}$	Media	5.3567	5.2286	5.2448	5.0073	5.0071	5.2561	5.2682	5.0057				
	Mediana	5.3587	5.3061	5.3155	5.0070	5.0055	5.3229	5.3295	5.0078				
	SD	0.0939	0.2378	0.2294	0.2024	0.2100	0.2245	0.2176	0.1976				
	MAD	0.0891	0.1606	0.1460	0.2009	0.2062	0.1365	0.1284	0.1982				
ECM	0.1360	0.1087	0.1125	0.0410	0.0441	0.1159	0.1192	0.0390					
$C_{3,10}$	Media	5.3408	5.3060	5.3161	5.0089	5.0076	5.3184	5.3238	5.0070				
	Mediana	5.3583	5.3389	5.3433	5.0057	5.0013	5.3448	5.3476	5.0011				
	SD	0.2579	0.1894	0.1780	0.2153	0.2151	0.1747	0.1689	0.2060				
	MAD	0.0921	0.1252	0.1210	0.2027	0.2016	0.1154	0.1117	0.1975				
ECM	0.1826	0.1295	0.1315	0.0464	0.0463	0.1319	0.1333	0.0425					
$C_{3,12}$	Media	5.3446	5.3194	5.3211	5.0143	5.0113	5.3287	5.3326	5.0133				
	Mediana	5.3569	5.3425	5.3438	5.0053	5.0021	5.3462	5.3490	5.0094				
	SD	0.2270	0.1717	0.1688	0.2316	0.2261	0.1549	0.1476	0.2213				
	MAD	0.0974	0.1276	0.1247	0.2212	0.2171	0.1175	0.1157	0.2078				
ECM	0.1702	0.1314	0.1316	0.0538	0.0512	0.1320	0.1324	0.0491					
$C_{3,5}$	Media	1.9033	1.9490	1.9446	1.9987	1.9995	1.9418	1.9374	1.9991				
	Mediana	1.9035	1.9122	1.9102	1.9994	1.9980	1.9090	1.9081	2.0000				
	SD	0.0095	0.0640	0.0624	0.0524	0.0576	0.0614	0.0596	0.0510				
	MAD	0.0090	0.0254	0.0208	0.0501	0.0538	0.0184	0.0158	0.0493				
ECM	0.0094	0.0067	0.0070	0.0027	0.0033	0.0072	0.0075	0.0026					
$C_{3,7}$	Media	1.9034	1.9374	1.9334	1.9984	1.9988	1.9302	1.9272	1.9989				
	Mediana	1.9031	1.9086	1.9076	1.9986	1.9982	1.9068	1.9061	1.9989				
	SD	0.0097	0.0605	0.0583	0.0529	0.0562	0.0566	0.0545	0.0514				
	MAD	0.0092	0.0171	0.0154	0.0513	0.0527	0.0143	0.0133	0.0503				
ECM	0.0094	0.0076	0.0078	0.0028	0.0032	0.0081	0.0083	0.0026					
$C_{3,10}$	Media	1.9061	1.9160	1.9135	1.9980	1.9985	1.9130	1.9118	1.9985				
	Mediana	1.9031	1.9050	1.9046	1.9996	1.9996	1.9045	1.9042	1.9988				
	SD	0.0430	0.0429	0.0395	0.0552	0.0550	0.0390	0.0375	0.0525				
	MAD	0.0096	0.0131	0.0125	0.0520	0.0511	0.0120	0.0116	0.0496				
ECM	0.0107	0.0089	0.0090	0.0030	0.0030	0.0091	0.0092	0.0028					
$C_{3,12}$	Media	1.9055	1.9116	1.9113	1.9966	1.9975	1.9095	1.9086	1.9969				
	Mediana	1.9033	1.9048	1.9046	1.9991	1.9983	1.9044	1.9041	1.9995				
	SD	0.0374	0.0350	0.0346	0.0593	0.0576	0.0306	0.0278	0.0564				
	MAD	0.0100	0.0132	0.0129	0.0554	0.0556	0.0122	0.0120	0.0537				
ECM	0.0103	0.0090	0.0091	0.0035	0.0033	0.0091	0.0091	0.0032					

Cuadro 8.11: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir del estimador inicial basado en las medianas bajo las contaminaciones $C_{3,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es P_{cos} .

	Eficiencia	$\theta_1 = 5$						$\theta_2 = 2$														
		M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂									
$C_{4,5}$	Media	5.3038	5.1295	5.1429	5.1754	5.1708	5.1848	5.1505	5.1561	5.0033	1.9181	1.9647	1.9536	1.9180	1.9180	1.9536	1.9524	1.9985	1.9989	1.9499	1.9482	1.9990
	Mediana	5.3056	5.2001	5.2213	5.2505	5.2473	5.2583	5.2651	5.0088	5.0037	1.9178	1.9289	1.9234	1.9176	1.9176	1.9234	1.9230	1.9977	1.9978	1.9222	1.9215	1.9985
	SD	0.0836	0.2302	0.2290	0.1955	0.2077	0.2183	0.2175	0.1969	0.2077	0.0079	0.0591	0.0585	0.0095	0.0095	0.0572	0.0564	0.0526	0.0561	0.0555	0.0551	0.0512
	MAD	0.0800	0.2372	0.2080	0.1953	0.2049	0.1458	0.1371	0.1978	0.2077	0.0075	0.0292	0.0242	0.0076	0.0076	0.0171	0.0159	0.0505	0.0528	0.0144	0.0133	0.0500
	ECM	0.0993	0.0697	0.0728	0.0382	0.0431	0.0818	0.0840	0.0388	0.0431	0.0068	0.0047	0.0049	0.0068	0.0068	0.0054	0.0028	0.0031	0.0031	0.0056	0.0057	0.0026
$C_{4,7}$	Media	5.3038	5.1708	5.1754	5.0064	5.0067	5.1848	5.1918	5.0051	5.0067	1.9180	1.9536	1.9536	1.9180	1.9180	1.9536	1.9524	1.9985	1.9989	1.9499	1.9482	1.9990
	Mediana	5.3075	5.2473	5.2505	5.0070	5.0066	5.2583	5.2651	5.0088	5.0066	1.9176	1.9234	1.9234	1.9176	1.9176	1.9234	1.9230	1.9977	1.9978	1.9222	1.9215	1.9985
	SD	0.0979	0.2259	0.2224	0.2011	0.2097	0.2183	0.2175	0.1969	0.2097	0.0095	0.0572	0.0564	0.0095	0.0095	0.0572	0.0564	0.0526	0.0561	0.0555	0.0551	0.0512
	MAD	0.0806	0.1656	0.1579	0.2008	0.2038	0.1458	0.1371	0.1978	0.2038	0.0076	0.0171	0.0159	0.0076	0.0076	0.0171	0.0159	0.0505	0.0528	0.0144	0.0133	0.0500
	ECM	0.1018	0.0801	0.0802	0.0404	0.0440	0.0818	0.0840	0.0388	0.0440	0.0068	0.0054	0.0028	0.0068	0.0068	0.0054	0.0028	0.0031	0.0031	0.0056	0.0057	0.0026
$C_{4,10}$	Media	5.3041	5.2249	5.2334	5.0094	5.0083	5.2414	5.2442	5.0071	5.0083	1.9179	1.9389	1.9368	1.9179	1.9179	1.9389	1.9368	1.9980	1.9984	1.9347	1.9341	1.9986
	Mediana	5.3062	5.2822	5.2847	5.0022	5.0024	5.2893	5.2885	4.9986	5.0024	1.9176	1.9198	1.9196	1.9176	1.9176	1.9198	1.9196	1.9995	1.9999	1.9192	1.9192	1.9994
	SD	0.0953	0.2028	0.1984	0.2139	0.2146	0.1921	0.1882	0.2055	0.2146	0.0091	0.0495	0.0484	0.0091	0.0091	0.0495	0.0484	0.0545	0.0548	0.0465	0.0456	0.0521
	MAD	0.0806	0.1344	0.1247	0.2036	0.2011	0.1153	0.1117	0.1976	0.2011	0.0076	0.0127	0.0118	0.0076	0.0076	0.0127	0.0118	0.0520	0.0507	0.0109	0.0105	0.0495
	ECM	0.1015	0.0917	0.0938	0.0458	0.0461	0.0952	0.0950	0.0422	0.0461	0.0068	0.0062	0.0063	0.0068	0.0068	0.0062	0.0063	0.0030	0.0030	0.0064	0.0064	0.0027
$C_{4,12}$	Media	5.3038	5.2487	5.2542	5.0106	5.0110	5.2594	5.2656	5.0090	5.0110	1.9180	1.9322	1.9308	1.9180	1.9180	1.9322	1.9308	1.9978	1.9977	1.9296	1.9280	1.9982
	Mediana	5.3065	5.2882	5.2903	5.0041	5.0040	5.2953	5.2966	5.0065	5.0040	1.9177	1.9194	1.9192	1.9177	1.9177	1.9194	1.9192	2.0003	1.9992	1.9187	1.9186	2.0001
	SD	0.0928	0.1926	0.1857	0.2295	0.2252	0.1780	0.1697	0.2179	0.2252	0.0088	0.0442	0.0426	0.0088	0.0088	0.0442	0.0426	0.0586	0.0572	0.0408	0.0384	0.0554
	MAD	0.0865	0.1228	0.1174	0.2171	0.2176	0.1106	0.1073	0.2092	0.2176	0.0082	0.0116	0.0111	0.0082	0.0082	0.0116	0.0111	0.0550	0.0560	0.0105	0.0102	0.0540
	ECM	0.1009	0.0989	0.0991	0.0527	0.0508	0.0990	0.0993	0.0475	0.0508	0.0068	0.0066	0.0066	0.0068	0.0068	0.0066	0.0066	0.0034	0.0033	0.0066	0.0067	0.0031

Cuadro 8.12: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir del estimador inicial basado en las medianas bajo las contaminaciones $C_{4,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es p_{\cos} .

		$\theta_1 = 5$										$\theta_2 = 2$										
		85 %					90 %					85 %					90 %					
	Eficiencia	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	
C_0	Media	5.0005	4.9995	4.9993	4.9987	4.9994	4.9993	4.9992	4.9992	4.9994	4.9993	4.9992	4.9992	4.9994	4.9993	4.9992	4.9992	4.9992	4.9994	4.9993	4.9992	4.9992
	Mediana	4.9983	5.0086	5.0094	5.0008	5.0048	5.0055	4.9992	5.0048	5.0055	5.0055	4.9992	5.0048	5.0055	5.0055	4.9992	5.0048	5.0055	5.0055	4.9992	5.0048	5.0055
	SD	0.1614	0.1725	0.1723	0.1856	0.166	0.1658	0.1789	0.166	0.1658	0.1658	0.1789	0.166	0.1658	0.1658	0.1789	0.166	0.1658	0.1658	0.1789	0.166	0.1658
$C_{1,5}$	MAD	0.1548	0.1639	0.1635	0.1855	0.1603	0.1587	0.1719	0.1603	0.1587	0.1587	0.1719	0.1603	0.1587	0.1587	0.1719	0.1603	0.1587	0.1587	0.1719	0.1603	0.1587
	ECM	0.026	0.0297	0.0297	0.0344	0.0275	0.0275	0.032	0.0275	0.0275	0.0275	0.032	0.0275	0.0275	0.032	0.0275	0.0275	0.032	0.0275	0.0275	0.032	0.0275
	MC	0.3115	0.2174	0.2162	0.0396	0.2481	0.2472	0.0407	0.2481	0.2472	0.2472	0.0407	0.2481	0.2472	0.2472	0.0407	0.2481	0.2472	0.2472	0.0407	0.2481	0.2472
$C_{1,7}$	Media	5.5717	5.4841	5.4782	5.0574	5.5081	5.5101	5.0774	5.5081	5.5101	5.5101	5.0774	5.5081	5.5101	5.5101	5.0774	5.5081	5.5101	5.5101	5.0774	5.5081	5.5101
	Mediana	5.5733	5.5267	5.5236	5.0179	5.5439	5.5423	5.0309	5.5439	5.5423	5.5423	5.0309	5.5439	5.5423	5.5423	5.0309	5.5439	5.5423	5.5423	5.0309	5.5439	5.5423
	SD	0.131	0.2679	0.2695	0.2722	0.2468	0.2404	0.2794	0.2468	0.2404	0.2404	0.2794	0.2468	0.2404	0.2404	0.2794	0.2468	0.2404	0.2404	0.2794	0.2468	0.2404
$C_{1,10}$	MAD	0.1257	0.189	0.1919	0.2162	0.169	0.169	0.2219	0.169	0.169	0.169	0.2219	0.169	0.169	0.169	0.2219	0.169	0.169	0.169	0.2219	0.169	0.169
	ECM	0.344	0.3061	0.3013	0.0773	0.319	0.318	0.084	0.319	0.318	0.318	0.084	0.319	0.318	0.318	0.084	0.319	0.318	0.318	0.084	0.319	0.318
	MC	5.5987	5.5484	5.5465	5.3116	5.5586	5.5575	5.3417	5.5586	5.5575	5.5575	5.3417	5.5586	5.5575	5.5575	5.3417	5.5586	5.5575	5.5575	5.3417	5.5586	5.5575
$C_{1,12}$	Media	5.6011	5.5428	5.5435	5.3343	5.5565	5.5558	5.4173	5.5565	5.5558	5.5558	5.4173	5.5565	5.5558	5.5558	5.4173	5.5565	5.5558	5.5558	5.4173	5.5565	5.5558
	Mediana	5.6118	5.5422	5.5394	5.5453	5.5551	5.5555	5.5593	5.5551	5.5555	5.5555	5.5593	5.5551	5.5555	5.5555	5.5593	5.5551	5.5555	5.5555	5.5593	5.5551	5.5555
	SD	0.19	0.214	0.2145	0.3367	0.1962	0.1967	0.3267	0.1962	0.1967	0.1967	0.3267	0.1962	0.1967	0.1967	0.3267	0.1962	0.1967	0.1967	0.3267	0.1962	0.1967
$C_{1,12}$	MAD	0.1489	0.2085	0.2052	0.2815	0.1968	0.1993	0.2543	0.1968	0.1993	0.1993	0.2543	0.1968	0.1993	0.1993	0.2543	0.1968	0.1993	0.1993	0.2543	0.1968	0.1993
	ECM	0.4025	0.3421	0.3414	0.3581	0.3469	0.3463	0.3647	0.3469	0.3463	0.3463	0.3647	0.3469	0.3463	0.3463	0.3647	0.3469	0.3463	0.3463	0.3647	0.3469	0.3463
	MC	1.8240	1.8284	1.8285	1.8551	1.8272	1.8273	1.834	1.8272	1.8273	1.8273	1.834	1.8272	1.8273	1.8273	1.834	1.8272	1.8273	1.8273	1.834	1.8272	1.8273

Cuadro 8.13: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir de un $S-$ estimador inicial bajo C_0 y bajo las contaminaciones $C_{1,m}$ para $m = 5, 7, 10$ y 12 cuando todas las respuestas son observadas, es decir cuando la probabilidad de pérdida es P_{comp} .

	$\theta_1 = 5$					$\theta_2 = 2$												
	Eficiencia		85 %			90 %			Eficiencia		85 %			90 %				
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{2,5}$	Media	5,4958	5,3541	5,3549	5,008	5,0057	5,3802	5,3771	5,0058	1,8616	1,8981	1,8980	1,9982	1,9991	1,8914	1,8923	1,9988	1,9988
	Mediana	5,4983	5,4415	5,4391	5,0076	5,0018	5,4597	5,4574	5,0049	1,8603	1,8662	1,8662	1,9979	1,9987	1,8639	1,8643	1,9996	1,9996
	SD	0,1648	0,2886	0,2885	0,1936	0,1977	0,2708	0,2724	0,1861	0,0320	0,0761	0,0761	0,0510	0,0543	0,0711	0,0717	0,0490	0,0490
	MAD	0,1105	0,1876	0,1911	0,1927	0,1934	0,1563	0,1557	0,1809	0,0138	0,0245	0,0245	0,0486	0,0500	0,0199	0,0200	0,0478	0,0478
	ECM	0,2729	0,2086	0,2091	0,0375	0,0391	0,2178	0,2163	0,0346	0,0202	0,0162	0,0162	0,0026	0,0030	0,0168	0,0167	0,0024	0,0024
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{2,7}$	Media	5,5095	5,4538	5,4535	5,0111	5,0082	5,465	5,4634	5,0082	1,8582	1,8694	1,8695	1,9975	1,9985	1,8673	1,8677	1,9983	1,9983
	Mediana	5,5114	5,4798	5,4813	5,0098	5,0046	5,4878	5,4891	5,0054	1,8576	1,8604	1,8602	1,9982	1,9987	1,8594	1,8596	1,9987	1,9987
	SD	0,1123	0,2116	0,2108	0,2045	0,2029	0,1963	0,1969	0,195	0,0144	0,0479	0,0478	0,0528	0,0538	0,0441	0,0446	0,0503	0,0503
	MAD	0,1101	0,1488	0,1501	0,2031	0,1987	0,139	0,1391	0,1895	0,0142	0,0193	0,0196	0,0514	0,0522	0,0178	0,0179	0,0493	0,0493
	ECM	0,2722	0,2506	0,2501	0,0419	0,0412	0,2547	0,2535	0,0381	0,0203	0,0193	0,0193	0,0028	0,0029	0,0196	0,0195	0,0025	0,0025
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{2,10}$	Media	5,5163	5,4792	5,4782	5,0423	5,0222	5,487	5,487	5,0388	1,8564	1,8610	1,8612	1,9896	1,9950	1,8599	1,8599	1,9906	1,9906
	Mediana	5,5177	5,4784	5,4804	5,0195	5,0074	5,4852	5,4861	5,0116	1,8560	1,8600	1,8599	1,9945	1,9981	1,8593	1,8592	1,9960	1,9960
	SD	0,1176	0,1751	0,177	0,258	0,2295	0,1603	0,16	0,2453	0,0151	0,0278	0,0284	0,0653	0,0581	0,0254	0,0254	0,0618	0,0618
	MAD	0,116	0,1617	0,1666	0,2187	0,2039	0,1466	0,1453	0,2059	0,0150	0,0213	0,0217	0,0574	0,0526	0,0188	0,0188	0,0533	0,0533
	ECM	0,2803	0,2603	0,26	0,0683	0,0531	0,2628	0,2627	0,0616	0,0209	0,0201	0,0201	0,0044	0,0034	0,0203	0,0203	0,0039	0,0039
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{2,12}$	Media	5,517	5,4768	5,4765	5,1491	5,0716	5,4838	5,4834	5,1499	1,8561	1,8606	1,8606	1,9612	1,9825	1,8597	1,8597	1,9611	1,9611
	Mediana	5,5198	5,4728	5,4714	5,0922	5,033	5,4808	5,48	5,086	1,8554	1,8607	1,8609	1,9758	1,9925	1,8597	1,8598	1,9770	1,9770
	SD	0,1219	0,1863	0,1866	0,3425	0,2917	0,1701	0,17	0,3334	0,0156	0,0243	0,0243	0,0868	0,0730	0,0221	0,0221	0,0843	0,0843
	MAD	0,1222	0,1785	0,1818	0,3497	0,2419	0,1645	0,1687	0,3258	0,0159	0,0231	0,0236	0,0927	0,0627	0,0216	0,0218	0,0867	0,0867
	ECM	0,2822	0,262	0,2619	0,1394	0,0901	0,263	0,2625	0,1335	0,0210	0,0200	0,0200	0,0090	0,0056	0,0202	0,0202	0,0086	0,0086

Cuadro 8.14: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir de un $S-$ estimador inicial bajo las contaminaciones $C_{2,m}$ para $m = 5, 7, 10$ y 12 cuando todas las respuestas son observadas, es decir cuando la probabilidad de pérdida es p_{COMP} .

		$\theta_1 = 5$										$\theta_2 = 2$										
		85 %					90 %					85 %					90 %					
	Eficiencia	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	
$C_{3,5}$	Media	5.3541	5.2376	5.2366	5.0085	5.0058	5.2488	5.2508	5.0058	5.2488	5.2508	5.0058	5.0058	1.9039	1.9351	1.9353	1.9980	1.9991	1.9322	1.9317	1.9988	1.9988
	Mediana	5.3547	5.3103	5.3104	5.0059	5.0042	5.3168	5.3179	5.0048	5.3168	5.3179	5.0048	5.0048	1.9038	1.9082	1.9082	1.9979	1.9980	1.9075	1.9075	1.9999	1.9999
	SD	0.0914	0.2343	0.2337	0.1929	0.1972	0.2248	0.2238	0.1855	0.2248	0.2238	0.1855	0.1855	0.0096	0.0599	0.0597	0.0506	0.0541	0.0575	0.0572	0.0487	0.0487
	MAD	0.0795	0.1469	0.1495	0.1878	0.1936	0.1302	0.13	0.1792	0.1302	0.13	0.1792	0.1792	0.0082	0.0155	0.0156	0.0489	0.0502	0.0135	0.0135	0.0473	0.0473
ECM	0.1338	0.1113	0.1105	0.0372	0.0389	0.1124	0.1129	0.0344	0.1124	0.1129	0.0344	0.0344	0.0093	0.0078	0.0077	0.0026	0.0029	0.0079	0.0079	0.0024	0.0024	
$C_{3,7}$	Media	5.3504	5.2845	5.2873	5.0104	5.0084	5.2889	5.2898	5.0075	5.2889	5.2898	5.0075	5.0075	1.9044	1.9226	1.9220	1.9978	1.9985	1.9216	1.9214	1.9986	1.9986
	Mediana	5.3582	5.3376	5.3391	5.0066	5.0057	5.339	5.3397	5.0042	5.339	5.3397	5.0042	5.0042	1.9032	1.9053	1.9050	1.9979	1.9978	1.9051	1.9050	1.9989	1.9989
	SD	0.1677	0.2109	0.209	0.1999	0.2008	0.2039	0.2026	0.1915	0.2039	0.2026	0.1915	0.1915	0.0247	0.0518	0.0511	0.0514	0.0532	0.0503	0.0501	0.0495	0.0495
	MAD	0.0804	0.1217	0.1205	0.1997	0.1947	0.1162	0.1146	0.1875	0.1162	0.1146	0.1875	0.1875	0.0083	0.0126	0.0125	0.0506	0.0521	0.0121	0.0119	0.0483	0.0483
ECM	0.1509	0.1254	0.1262	0.04	0.0404	0.125	0.125	0.0367	0.125	0.125	0.0367	0.0367	0.0097	0.0087	0.0087	0.0026	0.0028	0.0087	0.0087	0.0024	0.0024	
$C_{3,10}$	Media	5.3379	5.3113	5.3136	5.014	5.0107	5.3187	5.3191	5.0105	5.3187	5.3191	5.0105	5.0105	1.9066	1.9146	1.9142	1.9969	1.9979	1.9129	1.9129	1.9978	1.9978
	Mediana	5.3568	5.3423	5.3442	5.0085	5.0042	5.3453	5.3449	5.0055	5.3453	5.3449	5.0055	5.0055	1.9032	1.9047	1.9045	1.9972	1.9983	1.9044	1.9044	1.9980	1.9980
	SD	0.2801	0.1779	0.1763	0.2152	0.2091	0.1667	0.1653	0.2022	0.1667	0.1653	0.2022	0.2022	0.0475	0.0392	0.0389	0.0547	0.0536	0.0371	0.0368	0.0515	0.0515
	MAD	0.0866	0.1258	0.1241	0.1972	0.1982	0.1129	0.1126	0.1931	0.1129	0.1126	0.1931	0.1931	0.0090	0.0131	0.0129	0.0507	0.0511	0.0118	0.0117	0.0488	0.0488
ECM	0.1926	0.1285	0.1294	0.0465	0.0438	0.1294	0.1291	0.041	0.1294	0.1291	0.041	0.041	0.0110	0.0088	0.0089	0.0030	0.0029	0.0089	0.0089	0.0026	0.0026	
$C_{3,12}$	Media	5.3557	5.3259	5.3251	5.016	5.0109	5.3306	5.3319	5.0129	5.3306	5.3319	5.0129	5.0129	1.9035	1.9099	1.9101	1.9966	1.9978	1.9090	1.9087	1.9973	1.9973
	Mediana	5.3566	5.3441	5.3449	5.01	5.0016	5.3454	5.3448	5.0053	5.3454	5.3448	5.0053	5.0053	1.9033	1.9046	1.9045	1.9974	1.9990	1.9045	1.9045	1.9986	1.9986
	SD	0.0913	0.1647	0.1656	0.2345	0.2199	0.1516	0.1491	0.216	0.1516	0.1491	0.216	0.216	0.0095	0.0314	0.0318	0.0594	0.0558	0.0289	0.0284	0.0546	0.0546
	MAD	0.0921	0.1338	0.1338	0.2266	0.2108	0.1219	0.119	0.2124	0.1219	0.119	0.2124	0.2124	0.0094	0.0139	0.0138	0.0594	0.0557	0.0127	0.0124	0.0531	0.0531
ECM	0.1348	0.1333	0.1331	0.0552	0.0484	0.1323	0.1324	0.0468	0.1323	0.1324	0.0468	0.0468	0.0094	0.0091	0.0091	0.0035	0.0031	0.0091	0.0091	0.0030	0.0030	

Cuadro 8.15: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir de un $S-$ estimador inicial bajo las contaminaciones $C_{3,m}$ para $m = 5, 7, 10$ y 12 cuando todas las respuestas son observadas, es decir cuando la probabilidad de pérdida es p_{comp} .

	$\theta_1 = 5$										$\theta_2 = 2$																			
	Eficiencia					85 %					90 %					Eficiencia					85 %					90 %				
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	MC	M	MM	WMM ₁	WMM ₂			
$C_{4,5}$	Media	5.3039	5.143	5.1458	5.004	5.0029	5.162	5.1607	5.0029	5.162	5.1607	5.0029	5.0029	1.9180	1.9615	1.9609	1.9993	1.9999	1.9502	1.9507	1.9997	1.9996	1.9178	1.9178	1.9178	1.9178	1.9178	1.9178		
	Mediana	5.3057	5.2198	5.2239	5.0048	5.0008	5.2438	5.2424	4.9989	5.2438	5.2424	4.9989	5.0008	1.9176	1.9229	1.9228	2.0001	1.9989	1.9216	1.9215	2.0000	1.9993	1.9176	1.9176	1.9176	1.9176	1.9176	1.9176		
	SD	0.0765	0.2244	0.2235	0.1872	0.1944	0.2184	0.2183	0.1821	0.2184	0.2183	0.1821	0.1944	0.0077	0.0578	0.0577	0.0495	0.0522	0.0560	0.0559	0.0484	0.0495	0.0077	0.0077	0.0077	0.0077	0.0077	0.0077		
	MAD	0.0753	0.2041	0.1994	0.1824	0.1878	0.171	0.1736	0.1724	0.171	0.1736	0.1724	0.1878	0.0070	0.0237	0.0227	0.0479	0.0489	0.0564	0.0564	0.0479	0.0489	0.0070	0.0070	0.0070	0.0070	0.0070	0.0070		
	ECM	0.0982	0.0707	0.0711	0.035	0.0378	0.0739	0.0734	0.0331	0.0739	0.0734	0.0331	0.0378	0.0068	0.0048	0.0049	0.0024	0.0029	0.0051	0.0050	0.0023	0.0023	0.0068	0.0068	0.0068	0.0068	0.0068	0.0068		
$C_{4,7}$	Media	5.3048	5.169	5.1701	5.0037	5.0037	5.1855	5.1841	5.003	5.1855	5.1841	5.003	5.0037	1.9178	1.9546	1.9545	1.9996	1.9998	1.9502	1.9507	1.9997	1.9996	1.9178	1.9178	1.9178	1.9178	1.9178	1.9178		
	Mediana	5.3067	5.252	5.2528	4.9993	5.0033	5.2646	5.2652	5	5.2646	5.2652	5	5.0033	1.9176	1.9229	1.9228	2.0001	1.9989	1.9216	1.9215	2.0000	1.9993	1.9176	1.9176	1.9176	1.9176	1.9176	1.9176		
	SD	0.0815	0.2261	0.2256	0.192	0.197	0.219	0.2187	0.1873	0.219	0.2187	0.1873	0.197	0.0077	0.0578	0.0577	0.0495	0.0522	0.0560	0.0559	0.0484	0.0495	0.0077	0.0077	0.0077	0.0077	0.0077	0.0077		
	MAD	0.0717	0.1618	0.1572	0.1861	0.1916	0.1353	0.1327	0.1812	0.1353	0.1327	0.1812	0.1916	0.0068	0.0164	0.0160	0.0481	0.0503	0.0564	0.0564	0.0479	0.0489	0.0068	0.0068	0.0068	0.0068	0.0068	0.0068		
	ECM	0.0995	0.0796	0.0798	0.0368	0.0388	0.0823	0.0817	0.035	0.0823	0.0817	0.035	0.0388	0.0068	0.0054	0.0054	0.0024	0.0027	0.0056	0.0055	0.0023	0.0023	0.0068	0.0068	0.0068	0.0068	0.0068	0.0068		
$C_{4,10}$	Media	5.2975	5.198	5.1984	5.0046	5.0045	5.2087	5.2099	5.0037	5.2087	5.2099	5.0037	5.0045	1.9190	1.9466	1.9467	1.9994	1.9995	1.9439	1.9437	1.9996	1.9996	1.9190	1.9190	1.9190	1.9190	1.9190	1.9190		
	Mediana	5.3051	5.2647	5.2662	4.9977	4.9971	5.2761	5.279	4.9963	5.2761	5.279	4.9963	4.9971	1.9177	1.9215	1.9214	1.9998	1.9988	1.9204	1.9202	1.9993	1.9993	1.9177	1.9177	1.9177	1.9177	1.9177	1.9177		
	SD	0.1892	0.213	0.2112	0.1964	0.1999	0.206	0.206	0.1917	0.206	0.206	0.1917	0.1999	0.0295	0.0535	0.0534	0.0500	0.0513	0.0521	0.0520	0.0489	0.0489	0.0295	0.0295	0.0295	0.0295	0.0295	0.0295		
	MAD	0.079	0.1413	0.1366	0.1854	0.1946	0.1225	0.1194	0.1806	0.1225	0.1194	0.1806	0.1946	0.0074	0.0137	0.0134	0.0476	0.0497	0.0520	0.0520	0.0473	0.0473	0.0074	0.0074	0.0074	0.0074	0.0074	0.0074		
	ECM	0.1242	0.0845	0.0839	0.0386	0.0399	0.086	0.0865	0.0367	0.086	0.0865	0.0367	0.0399	0.0074	0.0057	0.0057	0.0025	0.0026	0.0059	0.0059	0.0024	0.0024	0.0074	0.0074	0.0074	0.0074	0.0074	0.0074		
$C_{4,12}$	Media	5.3041	5.2186	5.2162	5.0062	5.0062	5.2309	5.23	5.0052	5.2309	5.23	5.0052	5.0062	1.9180	1.9411	1.9418	1.9989	1.9989	1.9378	1.9381	1.9991	1.9991	1.9180	1.9180	1.9180	1.9180	1.9180	1.9180		
	Mediana	5.3066	5.2711	5.2706	5.0048	5.0029	5.2812	5.2814	5.0012	5.2812	5.2814	5.0012	5.0029	1.9176	1.9210	1.9210	2.0000	1.9982	1.9200	1.9200	1.9990	1.9990	1.9176	1.9176	1.9176	1.9176	1.9176	1.9176		
	SD	0.0834	0.2011	0.2005	0.202	0.2029	0.192	0.1903	0.1967	0.192	0.1903	0.1967	0.2029	0.0079	0.0494	0.0498	0.0515	0.0519	0.0473	0.0472	0.0503	0.0503	0.0079	0.0079	0.0079	0.0079	0.0079	0.0079		
	MAD	0.0821	0.1325	0.1321	0.1903	0.1948	0.12	0.1167	0.1906	0.12	0.1167	0.1906	0.1948	0.0077	0.0127	0.0127	0.0494	0.0507	0.0507	0.0507	0.0484	0.0484	0.0077	0.0077	0.0077	0.0077	0.0077	0.0077		
	ECM	0.0994	0.0882	0.0869	0.0408	0.0412	0.0901	0.0891	0.0387	0.0901	0.0891	0.0387	0.0412	0.0068	0.0059	0.0059	0.0027	0.0027	0.0061	0.0061	0.0025	0.0025	0.0068	0.0068	0.0068	0.0068	0.0068	0.0068		

Cuadro 8.16: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir de un $S-$ estimador inicial bajo las contaminaciones $C_{4,m}$ para $m = 5, 7, 10$ y 12 cuando todas las respuestas son observadas, es decir cuando la probabilidad de pérdida es p_{COMP} .

		$\theta_1 = 5$										$\theta_2 = 2$										
		85 %					90 %					85 %					90 %					
	Eficiencia	M	MM	WMN ₁	WMN ₂	M	MM	WMN ₁	WMN ₂	M	MM	WMN ₁	WMN ₂	M	MM	WMN ₁	WMN ₂	M	MM	WMN ₁	WMN ₂	
C_0	Media	4,9987	4,9984	4,9988	4,9975	4,9986	4,9984	4,9987	4,9987	4,9986	4,9984	4,9987	4,9987	4,9986	4,9984	4,9987	4,9987	4,9986	4,9984	4,9987	4,9987	4,9987
	Mediana	5,0031	5,0029	5,0001	5,0068	5,0042	5,0032	5,0012	5,0012	5,0042	5,0032	5,0012	5,0012	5,0042	5,0032	5,0012	5,0012	5,0042	5,0032	5,0012	5,0012	5,0012
	SD	0,2096	0,2094	0,2241	0,2241	0,2007	0,2	0,2139	0,2139	0,2007	0,2	0,2139	0,2139	0,2007	0,2	0,2139	0,2139	0,2007	0,2	0,2139	0,2139	0,2139
$C_{1,5}$	MAD	0,1816	0,2057	0,2066	0,2172	0,2309	0,1978	0,207	0,207	0,1978	0,1962	0,207	0,207	0,1978	0,1962	0,207	0,207	0,1978	0,1962	0,207	0,207	0,207
	ECM	0,0336	0,0439	0,0438	0,0502	0,0403	0,04	0,0457	0,0457	0,0403	0,04	0,0457	0,0457	0,0403	0,04	0,0457	0,0457	0,0403	0,04	0,0457	0,0457	0,0457
	MC	5,6001	5,3966	5,392	5,0181	5,4434	5,4412	5,0232	5,0232	5,4434	5,4412	5,0232	5,0232	5,4434	5,4412	5,0232	5,0232	5,4434	5,4412	5,0232	5,0232	5,0232
$C_{1,7}$	Media	5,5969	5,5042	5,4893	5,0093	5,5393	5,5332	5,0122	5,0122	5,5393	5,5332	5,0122	5,0122	5,5393	5,5332	5,0122	5,0122	5,5393	5,5332	5,0122	5,0122	5,0122
	Mediana	0,1439	0,3777	0,3765	0,2428	0,3534	0,3566	0,2435	0,2435	0,3534	0,3566	0,2435	0,2435	0,3534	0,3566	0,2435	0,2435	0,3534	0,3566	0,2435	0,2435	0,2435
	SD	0,145	0,3074	0,327	0,2275	0,2426	0,2549	0,2247	0,2247	0,2426	0,2549	0,2247	0,2247	0,2426	0,2549	0,2247	0,2247	0,2426	0,2549	0,2247	0,2247	0,2247
$C_{1,10}$	MAD	0,3808	0,2998	0,2953	0,0592	0,3214	0,3217	0,0598	0,0598	0,3214	0,3217	0,0598	0,0598	0,3214	0,3217	0,0598	0,0598	0,3214	0,3217	0,0598	0,0598	0,0598
	ECM	0,4249	0,3852	0,3823	0,1147	0,3937	0,392	0,1161	0,1161	0,3937	0,392	0,1161	0,1161	0,3937	0,392	0,1161	0,1161	0,3937	0,392	0,1161	0,1161	0,1161
	MC	5,6345	5,5389	5,5358	5,0929	5,5631	5,5615	5,1016	5,1016	5,5631	5,5615	5,1016	5,1016	5,5631	5,5615	5,1016	5,1016	5,5631	5,5615	5,1016	5,1016	5,1016
$C_{1,12}$	Media	5,6322	5,577	5,5755	5,0475	5,5918	5,5906	5,0442	5,0442	5,5918	5,5906	5,0442	5,0442	5,5918	5,5906	5,0442	5,0442	5,5918	5,5906	5,0442	5,0442	5,0442
	Mediana	0,1497	0,3079	0,3088	0,3258	0,2769	0,2772	0,3254	0,3254	0,2769	0,2772	0,3254	0,3254	0,2769	0,2772	0,3254	0,3254	0,2769	0,2772	0,3254	0,3254	0,3254
	SD	0,1536	0,2235	0,2249	0,2774	0,2014	0,1967	0,2748	0,2748	0,2014	0,1967	0,2748	0,2748	0,2014	0,1967	0,2748	0,2748	0,2014	0,1967	0,2748	0,2748	0,2748
$C_{1,12}$	MAD	0,4249	0,3852	0,3823	0,1147	0,3937	0,392	0,1161	0,1161	0,3937	0,392	0,1161	0,1161	0,3937	0,392	0,1161	0,1161	0,3937	0,392	0,1161	0,1161	0,1161
	ECM	0,4249	0,3852	0,3823	0,1147	0,3937	0,392	0,1161	0,1161	0,3937	0,392	0,1161	0,1161	0,3937	0,392	0,1161	0,1161	0,3937	0,392	0,1161	0,1161	0,1161
	MC	5,6345	5,5389	5,5358	5,0929	5,5631	5,5615	5,1016	5,1016	5,5631	5,5615	5,1016	5,1016	5,5631	5,5615	5,1016	5,1016	5,5631	5,5615	5,1016	5,1016	5,1016

Cuadro 8.17: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WM-$ estimadores calculados a partir de un $S-$ estimador inicial bajo C_0 y bajo las contaminaciones $C_{1,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es P_{LOG} .

	$\theta_1 = 5$						$\theta_2 = 2$					
	Eficiencia		85 %		90 %		Eficiencia		85 %		90 %	
	MC	MM	WMM ₁	WMM ₂	M	WMM ₁	MC	MM	WMM ₁	WMM ₂	M	WMM ₁
$C_{2,5}$	Media	5,5396	5,4018	5,4024	5,0146	5,0108	1,8550	1,8890	1,9965	1,9979	1,8855	1,8861
	Mediana	5,5359	5,483	5,4867	5,0114	5,0091	1,8550	1,8600	1,9943	1,9950	1,8588	1,8592
	SD	0,1206	0,3156	0,3149	0,2325	0,234	0,0154	0,0787	0,0605	0,0634	0,0752	0,0755
	MAD	0,1203	0,2047	0,209	0,2262	0,233	0,0155	0,0265	0,0271	0,0275	0,0230	0,0236
	ECM	0,3057	0,261	0,261	0,0542	0,0548	0,0213	0,0185	0,0037	0,0040	0,0188	0,0187
$C_{2,7}$	Media	5,5486	5,4931	5,4909	5,0232	5,0153	1,8527	1,8632	1,9941	1,9966	1,8621	1,8619
	Mediana	5,5437	5,5154	5,5133	5,0144	5,0112	1,8531	1,8555	1,9956	1,9960	1,8549	1,8547
	SD	0,1241	0,2349	0,2381	0,2504	0,2409	0,0158	0,0487	0,0500	0,0642	0,0464	0,0460
	MAD	0,127	0,1763	0,1813	0,2383	0,2333	0,0163	0,0225	0,0232	0,0261	0,0208	0,0208
	ECM	0,3163	0,2983	0,2976	0,0632	0,0582	0,0219	0,0211	0,0210	0,0041	0,0212	0,0212
$C_{2,10}$	Media	5,555	5,5127	5,5128	5,0746	5,0389	1,8511	1,8564	1,9815	1,9909	1,8555	1,8553
	Mediana	5,5548	5,5116	5,5115	5,0427	5,0224	1,8509	1,8555	1,9888	1,9938	1,8546	1,8547
	SD	0,1319	0,2199	0,2207	0,3315	0,2832	0,0168	0,0320	0,0321	0,0833	0,0293	0,0291
	MAD	0,1373	0,207	0,2083	0,3141	0,2593	0,0173	0,0268	0,0264	0,0764	0,0244	0,0247
	ECM	0,3254	0,3111	0,3116	0,1154	0,0817	0,0225	0,0216	0,0217	0,0073	0,0217	0,0218
$C_{2,12}$	Media	5,5544	5,5061	5,5068	5,1978	5,1036	1,8511	1,8569	1,9493	1,9743	1,8565	1,8563
	Mediana	5,5545	5,5049	5,5087	5,1658	5,0632	1,8509	1,8564	1,8558	1,9511	1,8559	1,8560
	SD	0,136	0,2532	0,2532	0,4133	0,3643	0,0174	0,0330	0,0336	0,1016	0,0308	0,0298
	MAD	0,1436	0,2444	0,2455	0,4594	0,3424	0,0185	0,0315	0,0317	0,1267	0,0287	0,0290
	ECM	0,3258	0,3202	0,3209	0,2098	0,1433	0,0225	0,0216	0,0216	0,0129	0,0216	0,0215

Cuadro 8.18: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir de un $S-$ estimador inicial bajo las contaminaciones $C_{2,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es P_{Log} .

		$\theta_1 = 5$										$\theta_2 = 2$										
		85 %					90 %					85 %					90 %					
	Eficiencia	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{3,5}$	Media	5.3709	5.2669	5.2696	5.0155	5.0108	5.2757	5.2772	5.0112	5.2757	5.2772	5.0112	5.0112	1.9021	1.9295	1.9288	1.9964	1.9980	1.9272	1.9270	1.9975	1.9975
	Mediana	5.3727	5.3399	5.3409	5.016	5.007	5.3428	5.3443	5.0111	5.3428	5.3443	5.0111	5.0111	1.9017	1.9051	1.9051	1.9931	1.9966	1.9048	1.9048	1.9951	1.9951
	SD	0.1051	0.2509	0.2489	0.2317	0.2338	0.2391	0.2375	0.2208	0.2391	0.2375	0.2208	0.2208	0.0111	0.0614	0.0610	0.0598	0.0632	0.0588	0.0585	0.0571	0.0571
	MAD	0.09	0.1487	0.1467	0.2272	0.23	0.1359	0.1358	0.2156	0.1359	0.1358	0.2156	0.2156	0.0094	0.0154	0.0152	0.0568	0.0596	0.0141	0.0141	0.0547	0.0547
ECM	0.1486	0.1341	0.1346	0.0539	0.0547	0.1331	0.1332	0.0488	0.1331	0.1332	0.0488	0.0488	0.0097	0.0087	0.0088	0.0036	0.0040	0.0088	0.0088	0.0033	0.0033	
	MC	M	M	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{3,7}$	Media	5.3709	5.3121	5.312	5.023	5.0166	5.3215	5.3218	5.0172	5.3215	5.3218	5.0172	5.0172	1.9022	1.9175	1.9176	1.9948	1.9965	1.9151	1.9151	1.9962	1.9962
	Mediana	5.373	5.3585	5.3555	5.014	5.0115	5.3592	5.3579	5.0121	5.3592	5.3579	5.0121	5.0121	1.9016	1.9030	1.9034	1.9955	1.9962	1.9030	1.9031	1.9957	1.9957
	SD	0.1499	0.2182	0.2182	0.2482	0.2425	0.2034	0.203	0.233	0.2034	0.203	0.233	0.233	0.0229	0.0504	0.0503	0.0626	0.0630	0.0467	0.0467	0.0590	0.0590
	MAD	0.093	0.1354	0.1383	0.237	0.2335	0.1244	0.1241	0.2259	0.1244	0.1241	0.2259	0.2259	0.0096	0.0140	0.0144	0.0616	0.0601	0.0129	0.0129	0.0569	0.0569
ECM	0.16	0.145	0.1449	0.0621	0.059	0.1447	0.1447	0.0545	0.1447	0.1447	0.0545	0.0545	0.0101	0.0093	0.0093	0.0039	0.0040	0.0094	0.0094	0.0035	0.0035	
	MC	M	M	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{3,10}$	Media	5.3642	5.3358	5.3375	5.0312	5.0229	5.3394	5.3416	5.024	5.3394	5.3416	5.024	5.024	1.9033	1.9109	1.9105	1.9924	1.9949	1.9098	1.9095	1.9946	1.9946
	Mediana	5.3805	5.3636	5.3598	5.0125	5.0125	5.3619	5.3641	5.0127	5.3619	5.3641	5.0127	5.0127	1.9007	1.9025	1.9028	1.9939	1.9949	1.9026	1.9024	1.9949	1.9949
	SD	0.2259	0.1999	0.1942	0.2889	0.2635	0.1827	0.1802	0.2597	0.1827	0.1802	0.2597	0.2597	0.0373	0.0401	0.0388	0.0735	0.0664	0.0369	0.0367	0.0651	0.0651
	MAD	0.0973	0.1549	0.1494	0.2718	0.2543	0.1313	0.1319	0.2516	0.1313	0.1319	0.2516	0.2516	0.0100	0.0160	0.0155	0.0701	0.0636	0.0136	0.0137	0.0615	0.0615
ECM	0.1836	0.1527	0.1516	0.0843	0.0699	0.1485	0.1491	0.068	0.1485	0.1491	0.068	0.068	0.0107	0.0095	0.0095	0.0055	0.0044	0.0095	0.0095	0.0043	0.0043	
	MC	M	M	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{3,12}$	Media	5.3731	5.3463	5.3488	5.0513	5.0293	5.3504	5.3546	5.036	5.3504	5.3546	5.036	5.036	1.9017	1.9064	1.9061	1.9874	1.9929	1.9058	1.9050	1.9912	1.9912
	Mediana	5.3757	5.3587	5.3608	5.0272	5.022	5.3595	5.3586	5.0253	5.3595	5.3586	5.0253	5.0253	1.9014	1.9031	1.9029	1.9938	1.9951	1.9030	1.9031	1.9920	1.9920
	SD	0.0992	0.1961	0.1945	0.3313	0.2984	0.1801	0.1724	0.2959	0.1801	0.1724	0.2959	0.2959	0.0103	0.0288	0.0289	0.0834	0.0749	0.0274	0.0253	0.0741	0.0741
	MAD	0.1023	0.1675	0.1663	0.322	0.2871	0.1525	0.1491	0.2729	0.1525	0.1491	0.2729	0.2729	0.0106	0.0173	0.0173	0.0800	0.0714	0.0157	0.0154	0.0700	0.0700
ECM	0.149	0.1583	0.1595	0.1123	0.0898	0.1551	0.1555	0.0888	0.1551	0.1555	0.0888	0.0888	0.0098	0.0096	0.0097	0.0071	0.0057	0.0096	0.0097	0.0056	0.0056	
	MC	M	M	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	

Cuadro 8.19: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir de un $S-$ estimador inicial bajo las contaminaciones $C_{3,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es P_{\log} .

	Eficiencia	$\theta_1 = 5$						$\theta_2 = 2$									
		85 %						85 %									
		M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂				
$C_{4,5}$	MC	5,316	5,1656	5,1666	5,0077	5,0058	5,1746	5,1768	5,0056	1,9169	1,9569	1,9567	1,9988	1,9994	1,9545	1,9541	1,9991
	Media	5,3187	5,2531	5,2514	5,0098	5,0052	5,2616	5,2665	5,0111	1,9165	1,9227	1,9229	1,9983	1,9975	1,9219	1,9215	1,9962
	SD	0,0901	0,2536	0,2509	0,2213	0,2274	0,2457	0,2453	0,2145	0,0085	0,0633	0,0628	0,0573	0,0615	0,0615	0,0615	0,0556
	MAD	0,083	0,1915	0,1944	0,22	0,2289	0,1716	0,1649	0,2127	0,0077	0,0192	0,0192	0,0543	0,0579	0,0169	0,0163	0,0532
	ECM	0,1079	0,0917	0,0906	0,049	0,0517	0,0908	0,0913	0,046	0,0070	0,0059	0,0058	0,0033	0,0038	0,0059	0,0059	0,0031
	MC	5,3174	5,1975	5,1975	5,0114	5,0089	5,2075	5,2064	5,0087	1,9167	1,9480	1,9481	1,9976	1,9985	1,9455	1,9460	1,9983
$C_{4,7}$	Media	5,3208	5,2726	5,2728	5,0062	5,0035	5,2789	5,2813	5,0034	1,9163	1,9208	1,9208	1,9983	1,9984	1,9202	1,9200	1,9974
	SD	0,0937	0,2364	0,2349	0,2299	0,2304	0,2299	0,23	0,2213	0,0091	0,0589	0,0586	0,0583	0,0601	0,0573	0,0574	0,0564
	MAD	0,0822	0,1588	0,1586	0,2187	0,2269	0,1439	0,1446	0,2199	0,0078	0,0151	0,0152	0,0550	0,0575	0,0138	0,0137	0,0527
	ECM	0,1095	0,0948	0,0941	0,053	0,0531	0,0958	0,0954	0,049	0,0070	0,0062	0,0061	0,0034	0,0036	0,0062	0,0062	0,0032
	MC	5,3161	5,2369	5,2322	5,0137	5,0111	5,2439	5,2492	5,0109	1,9168	1,9377	1,9393	1,9970	1,9978	1,9361	1,9349	1,9979
$C_{4,10}$	Media	5,3241	5,2928	5,2923	5,0124	5,0084	5,2977	5,2968	5,012	1,9159	1,9189	1,9189	1,9952	1,9975	1,9185	1,9184	1,9967
	SD	0,1154	0,222	0,2256	0,2472	0,2396	0,2143	0,2074	0,2344	0,0124	0,0520	0,0537	0,0622	0,0603	0,0506	0,0492	0,0587
	MAD	0,0879	0,1434	0,1464	0,2321	0,227	0,1301	0,1264	0,2293	0,0083	0,0136	0,0138	0,0589	0,0575	0,0124	0,0121	0,0563
	ECM	0,1132	0,1054	0,1047	0,0612	0,0575	0,1054	0,1051	0,055	0,0071	0,0066	0,0066	0,0039	0,0036	0,0066	0,0067	0,0034
	MC	5,3162	5,2513	5,2523	5,0187	5,0148	5,2642	5,2603	5,0143	1,9168	1,9333	1,9334	1,9959	1,9968	1,9302	1,9317	1,9970
$C_{4,12}$	Media	5,3196	5,2923	5,2934	5,0049	5,0072	5,3	5,2988	5,0053	1,9164	1,9190	1,9189	1,9983	1,9983	1,9183	1,9184	1,9986
	SD	0,0925	0,2124	0,208	0,2677	0,2562	0,1986	0,1993	0,2528	0,0087	0,0467	0,0465	0,0671	0,0648	0,0440	0,0453	0,0640
	MAD	0,0908	0,1507	0,1459	0,2466	0,2371	0,1296	0,1305	0,2277	0,0085	0,0142	0,0139	0,0628	0,0612	0,0121	0,0123	0,0581
	ECM	0,1086	0,1082	0,1069	0,0719	0,0658	0,1092	0,1075	0,0641	0,0070	0,0066	0,0066	0,0045	0,0042	0,0068	0,0067	0,0041

Cuadro 8.20: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir de un $S-$ estimador inicial bajo las contaminaciones $C_{4,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es P_{Log} .

		$\theta_1 = 5$										$\theta_2 = 2$										
		85 %					90 %					85 %					90 %					
	Eficiencia	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	
C_0	Media	5,0004	5	4,9955	4,9984	4,9972	5,0002	4,9999	4,9982	4,9982	4,9999	4,9982	4,9982	2,0000	2,0002	2,0015	2,0029	2,0000	2,0001	2,0016	2,0001	2,0016
	Mediana	4,9909	4,9974	4,9971	4,9982	4,9971	4,9962	4,9951	4,9927	4,9927	1,9998	1,9991	1,9990	1,9990	1,9998	1,9999	1,9997	1,9990	1,9991	1,9991	2,0001	2,0001
	SD	0,1642	0,1816	0,1822	0,1995	0,2162	0,1747	0,175	0,192	0,192	0,0414	0,0458	0,0572	0,0668	0,0437	0,0435	0,0508	0,0542	0,0440	0,0440	0,0548	0,0548
$C_{1,5}$	MAD	0,1578	0,166	0,1678	0,1897	0,2057	0,1611	0,1617	0,1851	0,1851	0,0438	0,0435	0,0508	0,0566	0,0424	0,0426	0,0499	0,0424	0,0426	0,0499	0,0437	0,0438
	ECM	0,0269	0,033	0,0332	0,0398	0,0467	0,0305	0,0306	0,0368	0,0368	0,0017	0,0021	0,0033	0,0045	0,0019	0,0019	0,0030	0,0019	0,0019	0,0030	0,0030	0,0030
	MC		M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2
$C_{1,7}$	Media	5,5764	5,5008	5,497	5,0742	5,0182	5,5255	5,5225	5,0964	5,0964	1,8318	1,8441	1,8450	1,9799	1,9959	1,8387	1,8397	1,8387	1,8397	1,9742	1,9742	
	Mediana	5,5788	5,5292	5,5261	5,0397	5,0078	5,5446	5,5454	5,0443	5,0443	1,8312	1,8325	1,8329	1,9900	1,9970	1,8314	1,8313	1,8314	1,8313	1,9889	1,9889	
	SD	0,1453	0,2571	0,259	0,2903	0,2328	0,2251	0,2326	0,3005	0,3005	0,0213	0,0604	0,0614	0,0773	0,0618	0,0501	0,0529	0,0501	0,0529	0,0789	0,0789	
$C_{1,10}$	MAD	0,1416	0,2025	0,2024	0,2464	0,2129	0,1808	0,1817	0,2414	0,2414	0,0209	0,0294	0,0307	0,0620	0,0550	0,0269	0,0267	0,0269	0,0267	0,0629	0,0629	
	ECM	0,3534	0,3169	0,314	0,0897	0,0545	0,3267	0,3271	0,0995	0,0995	0,0287	0,0280	0,0278	0,0064	0,0038	0,0285	0,0285	0,0285	0,0285	0,0069	0,0069	
	MC		M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2
$C_{1,12}$	Media	5,6114	5,5357	5,534	5,5022	5,334	5,5438	5,5423	5,5144	5,5144	1,8222	1,8296	1,8298	1,8515	1,9094	1,8286	1,8288	1,8286	1,8288	1,8501	1,8501	
	Mediana	5,609	5,5277	5,5296	5,545	5,3219	5,5393	5,5384	5,5588	5,5588	1,8219	1,8304	1,8295	1,8286	1,8982	1,8285	1,8288	1,8285	1,8288	1,8280	1,8280	
	SD	0,1623	0,2566	0,2582	0,3605	0,4391	0,2326	0,2324	0,345	0,345	0,0235	0,0383	0,0386	0,0856	0,1115	0,0346	0,0346	0,0346	0,0346	0,0826	0,0826	
$C_{1,12}$	MAD	0,1638	0,2461	0,2455	0,3103	0,5296	0,2231	0,2218	0,2869	0,2869	0,0234	0,0368	0,0365	0,0483	0,1443	0,0337	0,0337	0,0337	0,0337	0,0438	0,0438	
	ECM	0,4001	0,3528	0,3518	0,382	0,3042	0,3498	0,348	0,3835	0,3835	0,0322	0,0305	0,0304	0,0294	0,0206	0,0306	0,0305	0,0306	0,0305	0,0293	0,0293	
	MC		M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2	M	MM	WMM1	WMM2

Cuadro 8.21: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir de un $S-$ estimador inicial bajo C_0 y bajo las contaminaciones $C_{1,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es P_{coos} .

	$\theta_1 = 5$						$\theta_2 = 2$						
	Eficiencia		85 %		90 %		Eficiencia		85 %		90 %		
	MC	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	
$C_{2,5}$	Media	5.5019	5.3873	5.3867	5.0104	5.0072	5.4062	5.4085	5.0074	1.8600	1.8881	1.8882	1.9986
	Mediana	5.5037	5.448	5.4466	5.0044	5.0034	5.4628	5.4609	5.002	1.8596	1.8649	1.8652	1.9971
	SD	0.1205	0.2751	0.2741	0.2072	0.2134	0.2604	0.2546	0.1981	0.0154	0.0693	0.0692	0.0555
	MAD	0.1175	0.1874	0.1912	0.2022	0.21	0.1697	0.1714	0.1968	0.0148	0.0241	0.0249	0.0524
	ECM	0.2664	0.2256	0.2246	0.043	0.0456	0.2328	0.2317	0.0392	0.0198	0.0173	0.0173	0.0035
$C_{2,7}$	Media	5.5111	5.457	5.4583	5.0138	5.0104	5.4673	5.4675	5.0107	1.8577	1.8676	1.8673	1.9965
	Mediana	5.5122	5.4738	5.4757	5.0066	5.0074	5.4849	5.4832	5.0088	1.8574	1.8610	1.8607	1.9966
	SD	0.1241	0.2175	0.2146	0.2222	0.2199	0.2017	0.2011	0.2098	0.0159	0.0452	0.0441	0.0585
	MAD	0.1202	0.1712	0.168	0.218	0.2117	0.1585	0.1572	0.2055	0.0156	0.0221	0.0217	0.0556
	ECM	0.2766	0.2561	0.2561	0.0495	0.0484	0.259	0.259	0.0441	0.0205	0.0196	0.0196	0.0034
$C_{2,10}$	Media	5.5162	5.4763	5.4758	5.0537	5.0285	5.4828	5.4832	5.0479	1.8562	1.8610	1.8611	1.9862
	Mediana	5.515	5.4739	5.4779	5.0252	5.0083	5.4793	5.4802	5.0193	1.8561	1.8603	1.8598	1.9944
	SD	0.1324	0.1986	0.2004	0.2855	0.2522	0.1822	0.1816	0.2681	0.0171	0.0286	0.0294	0.0731
	MAD	0.1287	0.1886	0.1876	0.2508	0.221	0.1722	0.173	0.2254	0.0167	0.0245	0.0244	0.0644
	ECM	0.284	0.2663	0.2665	0.0843	0.0643	0.2662	0.2664	0.0741	0.0210	0.0201	0.0201	0.0055
$C_{2,12}$	Media	5.5182	5.4698	5.4706	5.1667	5.0784	5.4775	5.4767	5.1603	1.8558	1.8615	1.8614	1.9561
	Mediana	5.5194	5.4712	5.4721	5.1108	5.0352	5.4774	5.4761	5.1045	1.8553	1.8607	1.8607	1.9689
	SD	0.1349	0.2241	0.2246	0.3743	0.3191	0.2037	0.2033	0.3546	0.0173	0.0293	0.0294	0.0945
	MAD	0.1341	0.2126	0.2127	0.4042	0.2912	0.195	0.1947	0.3752	0.0173	0.0277	0.0275	0.1142
	ECM	0.2867	0.2708	0.2718	0.1678	0.1078	0.2695	0.2686	0.1513	0.0211	0.0200	0.0201	0.0109

Cuadro 8.22: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir de un $S-$ estimador inicial bajo las contaminaciones $C_{2,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es P_{cos} .

		$\theta_1 = 5$										$\theta_2 = 2$											
		85 %					90 %					85 %					90 %						
	Eficiencia	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂		
$C_{3,5}$	Media	5.3552	5.2374	5.2349	5.0106	5.007	5.2485	5.2491	5.0073														
	Mediana	5.3566	5.3147	5.3115	5.0018	5.0047	5.3216	5.3225	5.0009														
	SD	0.0916	0.2416	0.2405	0.206	0.2124	0.2304	0.23	0.1966														
	MAD	0.0871	0.1597	0.1611	0.2055	0.2156	0.1422	0.143	0.1964														
ECM	0.1345	0.1147	0.1129	0.0425	0.0451	0.1148	0.1149	0.0387															
$C_{3,7}$	Media	5.3567	5.2799	5.2803	5.0129	5.0106	5.2894	5.2887	5.0099														
	Mediana	5.3587	5.3332	5.3326	5.0068	5.0113	5.3376	5.3385	5.0079														
	SD	0.0939	0.2156	0.2144	0.2156	0.2157	0.2035	0.2059	0.205														
	MAD	0.0891	0.1371	0.135	0.212	0.2113	0.1262	0.1243	0.2031														
ECM	0.136	0.1248	0.1245	0.0466	0.0466	0.1251	0.1257	0.0421															
$C_{3,10}$	Media	5.3408	5.3058	5.3093	5.0195	5.014	5.3136	5.3128	5.0136														
	Mediana	5.3583	5.3341	5.3352	5.0068	5.003	5.3387	5.3389	5.0043														
	SD	0.2579	0.1915	0.1883	0.2359	0.2272	0.1791	0.1805	0.2192														
	MAD	0.0921	0.1405	0.1395	0.2151	0.2139	0.1287	0.1275	0.2022														
ECM	0.1826	0.1302	0.1311	0.056	0.0518	0.1304	0.1304	0.0482															
$C_{3,12}$	Media	5.3446	5.3206	5.3215	5.0231	5.0172	5.3278	5.3282	5.0182														
	Mediana	5.3569	5.3421	5.344	5.0156	5.0096	5.347	5.3475	5.0113														
	SD	0.227	0.1852	0.1824	0.258	0.2375	0.1718	0.1674	0.2348														
	MAD	0.0974	0.1492	0.1498	0.2371	0.2297	0.1304	0.1329	0.2254														
ECM	0.1702	0.137	0.1366	0.067	0.0566	0.137	0.1357	0.0554															
$C_{3,12}$	Media	1.9055	1.9115	1.9113	1.9945	1.9960	1.9099	1.9098	1.9957														
	Mediana	1.9033	1.9048	1.9046	1.9966	1.9977	1.9043	1.9043	1.9971														
	SD	0.0374	0.0357	0.0354	0.0664	0.0602	0.0332	0.0325	0.0595														
	MAD	0.0100	0.0155	0.0156	0.0631	0.0584	0.0136	0.0139	0.0568														
ECM	0.0103	0.0091	0.0091	0.0044	0.0036	0.0092	0.0092	0.0036															

Cuadro 8.23: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir de un $S-$ estimador inicial bajo las contaminaciones $C_{3,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es P_{cos} .

	Eficiencia	$\theta_1 = 5$						$\theta_2 = 2$					
		85 %			90 %			85 %			90 %		
	MC	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂	M	MM	WMM ₁	WMM ₂
$C_{4,5}$	Media	5.3038	5.1429	5.1458	5.0041	5.003	5.1545	5.1558	5.003	1.9181	1.9615	1.9607	1.9990
	Mediana	5.3056	5.2221	5.2217	4.9981	5.0038	5.2345	5.2374	5.0004	1.9178	1.9261	1.9258	1.9990
	SD	0.0836	0.2355	0.2338	0.1995	0.2096	0.231	0.2294	0.1935	0.0079	0.0600	0.0595	0.0528
	MAD	0.08	0.2165	0.214	0.1989	0.2067	0.1918	0.1848	0.1934	0.0075	0.0231	0.0227	0.0507
	ECM	0.0993	0.0758	0.0759	0.0398	0.0439	0.0772	0.0768	0.0374	0.0068	0.0051	0.0051	0.0028
$C_{4,7}$	Media	5.3038	5.1683	5.17	5.0042	5.0049	5.1821	5.1847	5.0034	1.9180	1.9547	1.9544	1.9994
	Mediana	5.3075	5.245	5.245	5.0021	5.0074	5.26	5.2612	5.0042	1.9176	1.9239	1.9237	1.9984
	SD	0.0979	0.2314	0.2298	0.2041	0.2109	0.2231	0.2219	0.1984	0.0095	0.0579	0.0576	0.0529
	MAD	0.0806	0.1868	0.1871	0.2025	0.208	0.1615	0.1552	0.199	0.0076	0.0193	0.0192	0.0516
	ECM	0.1018	0.0818	0.0817	0.0416	0.0445	0.0829	0.0833	0.0393	0.0068	0.0054	0.0054	0.0028
$C_{4,10}$	Media	5.3041	5.2046	5.2055	5.0056	5.0049	5.2144	5.2175	5.0037	1.9179	1.9448	1.9447	1.9987
	Mediana	5.3062	5.2702	5.2712	4.9994	5.0018	5.2754	5.2799	4.9994	1.9176	1.9210	1.9209	1.9997
	SD	0.0953	0.2175	0.2156	0.2085	0.2106	0.2103	0.2095	0.2008	0.0091	0.0539	0.0536	0.0537
	MAD	0.0806	0.1494	0.1458	0.1993	0.2013	0.1326	0.1297	0.1933	0.0076	0.0144	0.0140	0.0499
	ECM	0.1015	0.0892	0.0887	0.0434	0.0443	0.0902	0.0912	0.0403	0.0068	0.0059	0.0059	0.0029
$C_{4,12}$	Media	5.3038	5.2162	5.2194	5.0077	5.0081	5.2258	5.2271	5.0061	1.9180	1.9418	1.9411	1.9987
	Mediana	5.3065	5.2732	5.2719	5.0066	5.0043	5.2794	5.2812	5.0032	1.9177	1.9208	1.9209	1.9998
	SD	0.0928	0.2124	0.2083	0.222	0.2179	0.2019	0.2022	0.2107	0.0088	0.0509	0.0504	0.0563
	MAD	0.0865	0.1435	0.1389	0.2075	0.2144	0.1264	0.1284	0.2033	0.0082	0.0140	0.0132	0.0516
	ECM	0.1009	0.0918	0.0915	0.0493	0.0475	0.0917	0.0924	0.0444	0.0068	0.0060	0.0060	0.0032

Cuadro 8.24: Medidas resumen del estimador de mínimos cuadrados y de los $M-$, $MM-$ y $WMM-$ estimadores calculados a partir de un $S-$ estimador inicial bajo las contaminaciones $C_{4,m}$ para $m = 5, 7, 10$ y 12 cuando la probabilidad de pérdida es P_{cos} .

Bibliografía

- [1] Anthony, M. y Bartlett, P. (1999). *Neural Network Learning: Theoretical Foundations*, Cambridge University Press.
- [2] Bianco, A. y Boente, G. (2002). On the asymptotic behavior of one-step estimates in heteroscedastic regression models. *Statistics & Probability Letters*, **60**, 33–47.
- [3] Bianco, A. y Boente, G. (2004). Robust estimators in semiparametric partly linear regression models. *Journal of Statistical Planning and Inference*, **122**, 229–252.
- [4] Bianco, A., Boente, G. y Rodrigues, I. (2013). Resistant estimators in Poisson and Gamma models with missing responses and an application to outlier detection. *Journal of Multivariate Analysis*, **114**, 209–226.
- [5] Cleveland, W. (1985). *The elements of graphing data*, Bell Telephone Laboratories Inc., New Jersey.
- [6] Dudzinski, M. y Mykytowycz, R. (1961). The eye lens as an indicator of age in the wild rabbit in Australia. *CSIRO Wildlife Research*, **6**, 156–159.
- [7] Fasano, V. (2009). *Teoría asintótica de estimadores robustos en regresión lineal*. Tesis Doctoral, Universidad Nacional de la Plata. Disponible en http://www.mate.unlp.edu.ar/tesis/tesis_fasano_v.pdf.
- [8] Fasano, V., Maronna, R., Sued, M. y Yohai, V. (2012). Continuity and differentiability of regression M -functionals. *Bernoulli*, **18**, 1284–1309.
- [9] Fraiman, R. (1983). General M -estimators and applications to bounded influence estimation for non-linear regression. *Communication in Statistics - Theory and Methods*, **22**, 2617–2631.
- [10] Gallant, A. (1975). Nonlinear regression. *The American Statistician*, **29**, 73–81
- [11] Hampel, F. (1974). The influence curve and its role in robust estimation. *The Annals of Statistics*, **69**, 383–393.
- [12] Hampel, F., Ronchetti, E., Rousseeuw, P. y Stahel, W. (1986). *Robust Statistics: The Approach Based on Influence Functions*, John Wiley & Sons, Inc.

- [13] Huber, P. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. *5th Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 221–233.
- [14] Huber, P. (1981). *Robust Statistics*, John Wiley & Sons, Inc.
- [15] Koul, H., Müller, U. y Schick, A. (2012). The transfer principle: a tool for complete case analysis. *The Annals of Statistics*, **40**, 3031–3049.
- [16] Liu, R., Markatou, M., y Tsai, C. (2005). Robust estimation and testing in nonlinear regression models. *International Journal of Pure and Applied Mathematics*, **21**, 525–552.
- [17] Markatou, M. y Manos, G. (1996). Robust tests in nonlinear regression models. *Journal of Statistical Planning and Inference*, **55**, 205–217.
- [18] Maronna, R., Martin, D. y Yohai, V. (2006). *Robust statistics: Theory and methods*, Wiley, New York.
- [19] Morales, D., Pardo, L., Pardo, M. y Vajda, I. (2004) Rényi statistics for testing composite hypothesis in general exponential models. *Statistics: A Journal of Theoretical and Applied Statistics*, **38(2)**, 133–147.
- [20] Müller, U. (2009). Estimating linear functionals in nonlinear regression with responses missing at random. *The Annals of Statistics*, **37**, 2245–2277.
- [21] Mukherjee, K. (1996). Robust estimation in nonlinear regression via minimum distance method. *Mathematical Methods of Statistics*, **5**, 99–112.
- [22] Nolan, D. y Pollard, D. (1987). U–Processes: Rates of Convergence. *The Annals of Statistics*, **15(2)**, 780–799.
- [23] Pakes, A. y Pollard, D. (1989). Simulation and the Asymptotics of Optimization Estimators. *Econometrica*, **57**, 1027–1057.
- [24] Pollard, D. (1984). *Convergence of Stochastic Processes*, Springer–Verlag New York Inc.
- [25] Ratkowsky, D. (1983). *Nonlinear regression modeling*. Marcel Dekker, New York.
- [26] Sakata, S. y White, H. (2001). S–estimation of nonlinear regression models with dependent and heterogeneous observations. *Journal of Econometrics*, **103**, 5–72.
- [27] Seber, G. y Wild, C. (1989). *Nonlinear regression*, John Wiley & Sons, Inc.
- [28] Siegel, A. (1982). Robust regression using repeated medians. *Biometrika*, **69(1)**, 242–244.
- [29] Stromberg, A. (1993). Computation of high breakdown nonlinear regression parameters. *Journal of the American Statistical Association*, **88**, 237–244.

- [30] Stromberg, A. y Ruppert, D. (1992). Breakdown in nonlinear regression. *Journal of the American Statistical Association*, **87**, 991–997.
- [31] Sued, M. y Yohai, V. (2013). Robust location estimation with missing data. *The Canadian Journal of Statistics*, **41**, 111–132
- [32] Tabatabai, M. y Argyros, I. (1993). Robust estimation and testing for general nonlinear regression models. *Applied Mathematics and Computation*, **58**, 85–101.
- [33] van der Vaart, A. y Wellner, J. (1996). *Weak Convergence and Empirical Processes*, Springer–Verlag New York Inc.
- [34] van de Geer, S. (2000). *Empirical processes in M–estimation*, Cambridge University Press.
- [35] Wellner, J. (2012). *Empirical process theory for statistics*. Curso corto cuyas notas se encuentran disponibles en <http://www.stat.washington.edu/jaw/RESEARCH/TALKS/talks.html>.
- [36] Yohai, V. (1987). High Breakdown point and high efficiency estimates for regression. *Journal of the American Statistical Association*, **15**, 642–656.