

ROBUSTEZ A DISTRIBUCIÓN PARA REGRESIÓN CUANTÍLICA EN DATOS LONGITUDINALES

Fernando Alonso Velez Reyes

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Estadística

2011

ROBUSTEZ A DISTRIBUCIÓN PARA REGRESIÓN CUANTÍLICA EN DATOS LONGITUDINALES

Fernando Alonso Velez Reyes

Tesis presentada como requisito parcial para optar al título de:
Magister en Estadística

Director

Msc. Luis Guillermo Díaz M.

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Estadística

2011

*A mis padres Juan y Francelina, mi hermano Carlos y mi novia
Monica ...*

Agradecimientos

Agradezco a Cesar Augusto Serna M. profesor de la universidad central por su colaboración revisando y corrigiendo el documento, a Yuan Liu y Matteo Bottai de la universidad de Carolina del Sur (USA) por su cordial colaboración en la comprensión del algoritmo, a Belcorp por su autorización para el uso de los datos empleados para en la aplicación.

Resumen

La regresión cuantílica es una técnica de modelamiento de gran flexibilidad que permite la descripción del cuantil condicional de una variable respuesta Y dado unas variables explicativas X , siendo la regresión mediana un caso particular. Este tipo de regresión se ha desarrollado durante los últimos 30 años, logrando extenderse a varias áreas. La extensión al análisis de datos longitudinales tiene unos pocos años de desarrollo, en los cuales dos metodologías de estimación se destacan, mínimos cuadrados y máxima verosimilitud. Este segundo método utiliza la distribución asimétrica de Laplace para obtener la estimación de los parámetros, sin embargo esto implica suponer que la variable respuesta sigue la distribución antes mencionada. En este trabajo, se evalúa el efecto que tiene la modificación de este supuesto sobre las estimaciones de los parámetros, las varianzas de los efectos aleatorios y el ajuste del modelo. Para ello, se emplea un estudio de simulación en el cual se determina el sesgo relativo de los parámetros, la bondad de ajuste y una medida de desempeño de las colas de los estimadores.

La estimación por máxima verosimilitud fue utilizada para describir un experimento sensorial en productos hidroalcohólicos, permitiendo evaluar el desempeño de dos productos en un tiempo de 8 horas. Se desarrolló la estimación para un modelo de medidas repetidas en 19 cuantiles diferentes, encontrando diferencia estadísticamente significativa entre los productos en la mayoría de casos. Igualmente, un modelo lineal mixto fue ajustado vía REML, cuyos resultados son muy próximos a los obtenidos para el caso de regresión mediana. En general, se observa un mejor ajuste en términos del criterio de Akaike para los cuantiles $\tau = 0.55$ y $\tau = 0.60$.

Abstract

Quantile regression is a technique for modeling flexibility that allows the description of the conditional quantile of a response variable Y given some explanatory variables X , with a median regression case. This type of regression has been developed over the past 30 years, allowing spread to multiple areas. The extension to the analysis of longitudinal data has a few years of development, in which two estimation methodologies are highlighted, least squares and maximum likelihood. This second method uses the asymmetric Laplace distribution for the estimation of the parameters, but this implies the assumption that the response variable is the distribution above. In this paper, we evaluate the effect of changing this assumption on parameter estimates, the variances of the random-effects model fit. To that end, uses a simulation study which determines the relative bias of the parameters, the goodness of fit and performance measurement of the tails of the estimates.

The maximum likelihood estimation was used to describe a sensory experiment hydroalcoholic products, allowing to evaluate the performance of two products in a time of 8 hours. The estimate was developed for a model of repeated measures in 19 different quantiles, finding statistically significant difference between the products in most cases. Similarly, a linear mixed model was fitted via REML, the results are very close to those obtained for the median regression. In general, there is a better fit in terms of the Akaike criterion for quantile $\tau = 0.55$ and $\tau = 0.60$.

Contenido

Agradecimientos	V
Resumen	VI
Abstract	VII
Lista de Figuras	XI
Lista de Tablas	XIII
1. Introducción	1
2. Elementos Generales	3
2.1. Datos Longitudinales	3
2.1.1. Modelo	4
2.2. Regresión cuantílica	5
2.2.1. Cuantiles	6
2.2.2. Modelo	6
2.2.3. Estimación	8
2.2.4. Propiedades de estimadores	8
2.2.5. Interpretación de los modelos de regresión cuantílica	9
2.2.6. Bondad de ajuste	9
2.3. Elementos de estadística robusta	11
2.4. Muestreo de aceptación y rechazo adaptativo	13
3. Regresión cuantílica para datos longitudinales	16
3.1. Modelo	16
3.2. Estimación	19
3.2.1. Paso E	19
3.2.2. Paso M	20
3.3. Distribución de Efectos Aleatorios	21
3.4. Robustez a distribución	21
3.4.1. Cálculo de $B(a, T_n)$	21

3.4.2. Probabilidad del sesgo	22
4. Simulación	23
4.1. Modelo	23
4.2. Selección de parámetros	24
4.3. Criterio de convergencia y tamaño de la simulación	25
5. Resultados	27
5.1. Estimación de efectos fijos	27
5.2. Estimación de varianzas y criterio de ajuste	30
5.3. Desempeño de colas	31
6. Aplicación	36
6.1. Análisis sensorial de productos hidroalcohólicos	36
7. Conclusiones y Recomendaciones	41
7.1. Conclusiones	41
7.2. Recomendaciones	42
A. Distribución Asimétrica de Laplace	43
A.1. Estimadores de los parámetros	45
B. Código R	47
C. Resultados aplicación	56
Bibliografía	59

Lista de Figuras

2.1. Función de Pérdida	7
5.1. Desempeño de colas $B(a, T_n)$ dist. simétricas	32
5.2. Probabilidad de sesgos para la distribución normal sesgada	33
5.3. Probabilidad de sesgos para las distribuciones t-Student y Ji-cuadrado	34
5.4. Probabilidad de sesgos para la distribución asimétrica de Laplace	35
6.1. Diagrama de perfiles	37
6.2. Variación de los parámetros de acuerdo al cuantil	39
6.3. Variación en el criterio de información de Akaike	40
A.1. Función de densidad	44

Lista de Tablas

4.1. Distribuciones usadas para el término de error	24
5.1. Estimaciones de β_0 y β_1 en la Dist. Normal sesgada $NS(\mu, \sigma^2, \psi)$	27
5.2. Estimaciones de β_0 y β_1 en la Dist. Asimétrica de Laplace $DAL(\mu, \sigma, \tau)$	28
5.3. Estimaciones de β_0 y β_1 en la Dist. t-Student t_ν	29
5.4. Estimaciones de β_0 y β_1 en la Dist. Ji-Cuadrado χ_ν^2	29
5.5. Estimaciones de σ	30
5.6. Estimaciones de φ	31
5.7. Estimaciones del criterio de información de Akaike	31
6.1. Estimación de parámetros y el errores estándar para los efectos fijos	38
C.1. Estimación de parámetros para los efectos fijos	56
C.2. Estimación de parámetros, errores estándar y valores P. $\tau < 0.5$	57
C.3. Estimación de parámetros, errores estándar y valores P. $\tau \geq 0.5$	58

Capítulo 1

Introducción

La regresión cuantílica es una generalización de la regresión mediana para cualquier cuantil τ entre 0 y 1. Esta metodología fue formalizada por [Koenker y Bassett \(1978\)](#), quienes establecieron las bases que permitirían la gran mayoría de los avances en el tema. Esta forma de modelamiento presenta ventajas importantes en relación con la regresión clásica, puesto que permite el modelamiento para cuantiles extremos como $\tau = 0.01, 0.05, 0.95$ y 0.99 o similares que describen los contornos de un conjunto de datos, lo cual no es posible en el modelamiento de $E[y | x]$. La formulación general se basa en el método de mínimos cuadrados, requiriendo el uso de procedimientos de optimización para determinar la solución. Algunos algoritmos son discutidos en [Koenker y D'Orey \(1987\)](#) y [Koenker y Park \(1996\)](#).

La aplicación de la regresión cuantílica a datos longitudinales fue propuesta por [Koenker \(2004\)](#), considerado para ello un modelo de medidas repetidas con intercepto aleatorio. La estimación la desarrolla por el método de mínimos cuadrados, sin embargo considera un término de penalización λ sobre los efectos aleatorios a fin de controlar la pérdida de grados de libertad tras la estimación de tales efectos. La principal dificultad con este método es que λ debe ser seleccionado sin un criterio claramente definido y eso según [Geraci y Bottai \(2007\)](#) afecta la estimación. El cálculo óptimo de λ es aún tema de investigación.

Algunos métodos de estimación alternativos han sido propuestos por [Geraci y Bottai \(2007\)](#) y [Canay \(2008\)](#). Este último autor propone una metodología diferente mediante una estimación en dos etapas. Según [Canay \(2008\)](#), los estimadores obtenidos de esta manera presentan un buen comportamiento en cuanto a sesgo, consistencia, normalidad asintótica y muestran ser mejor en términos de error cuadrático medio, con relación al

estimador propuesto por [Koenker \(2004\)](#). Sin embargo, esta propuesta aún no ha sido publicada y solo se presenta a nivel informativo. En contraste, el método propuesto por [Geraci y Bottai \(2007\)](#) se fundamenta en la distribución asimétrica de Laplace para desarrollar la estimación vía máxima verosimilitud, sin embargo esto supone la solución de integrales de gran complejidad que impide tener una forma funcional de la solución. En lugar de ello, los autores emplean herramientas computacionales como el algoritmo EM, el muestreador de Gibbs y el muestreo de aceptación y rechazo adaptativo. El eje central de esta metodología consiste en generar valores con la distribución condicional de los efectos aleatorios y así poderlos sustraer del modelo original, permitiendo reducir el modelo a un caso de regresión cuantílica normal obteniendo las estimaciones de los efectos fijos. Sin embargo, los resultados obtenidos con esta metodología tienen el supuesto que la variable respuesta sigue una distribución asimétrica de Laplace, lo que en la realidad podría resultar poco probable.

En el presente trabajo se estudia algunas características que permiten evaluar la robustez de la metodología propuesta por [Geraci y Bottai \(2007\)](#) cuando se tienen alteraciones del supuesto distribucional. Para ello, se consideran diferentes distribuciones para el término de error, considerando casos simétricos y no simétricos así como distribuciones con colas pesadas. Este estudio se desarrollo mediante simulación, permitiendo estimar el sesgo relativo y la varianza de la estimación. Por otra parte, [Jurečkova \(1981\)](#) desarrolla una medida de desempeño de colas para estimadores de localización que luego es extendida a estimadores de regresión por [He et al. \(1990\)](#). Esta medida es considerada en el trabajo para determinar el efecto que ejerce la distribución del término de error sobre la distribución de los estimadores. Sin embargo, el trabajo desarrollado en los trabajos antes citados parte del supuesto según el cual, la distribución de la respuesta es simétrica, lo que en la realidad no siempre es correcto. Para ello, se explora el comportamiento de las colas en presencia de distribuciones no simétricas para el termino de error para lo cual se calculó la probabilidad de sesgo.

Finalmente, la metodología propuesta por [Geraci y Bottai \(2007\)](#) es aplicada a un caso de análisis sensorial, en el cual se evalúa la intensidad de productos hidroalcohólicos, mediante jueces entrenados olfativamente. Un modelo lineal mixto es ajustado, considerando el efecto individual de los jueces como término aleatorio. Así mismo, diferentes modelos de regresión cuantílica son ajustados encontrando una clara concordancia en el caso de la mediana con el modelo clásico.

Capítulo 2

Elementos Generales

2.1. Datos Longitudinales

Con frecuencia, los estudios estadísticos considerados en muchas áreas del conocimiento permiten describir fenómenos o problemáticas de manera transversal, es decir tomando información en un momento en particular para el respectivo análisis. Esta forma de proceder, aunque correcta, no permite evaluar como cambian las principales variables involucradas a través del tiempo. Esto es importante en muchos casos por cuanto permite describir la evolución de las variables en estudio. Así, la medición de una o más características en un conjunto de individuos en varios tiempos se denomina estudio de medidas repetidas y constituye el caso más común de datos longitudinales, permitiendo separar el efecto del tiempo de los factores o covariables de interés. Este tipo de análisis permite tener en cuenta la asociación existente entre las mediciones de un mismo individuo, modelando así la estructura de correlación presente.

Cabe mencionar que la elección del modelo estadístico depende del tipo de respuesta y del tipo de inferencia que se desea hacer. Existen varias metodologías usadas para el modelado de datos longitudinales: los modelos marginales, dirigidos a modelar la media y varianza de los datos considerando ciertas estructuras de correlación; los modelos de efectos aleatorios, donde un componente aleatorio explica la correlación entre los individuos, así como la heterogeneidad no observada que está implícita en esta correlación; y los modelos de transición, enfocados en la distribución condicional de una respuesta, dada la historia pasada, tanto de las respuestas anteriores como de las variables explicativas.

En los casos donde la variable respuesta se distribuye aproximadamente normal, se hace uso de modelos lineales mixtos, mientras que las ecuaciones de estimación generalizadas resultan más apropiadas cuando la naturaleza de la respuesta es de tipo binomial, conteo u otra distribución de la familia exponencial.

2.1.1. Modelo

En general, el modelo para datos longitudinales puede expresarse como un modelo lineal mixto, para lo cual considérese N individuos sobre los cuales se realiza n_i mediciones (en un diseño balanceado $n_i = n$) entonces el modelo se puede expresar como:

$$\mathbf{Y}_i = \mathbf{X}_i\beta + \mathbf{Z}_i\alpha_i + \varepsilon_i \quad (2.1)$$

donde $\mathbf{Y}_i = \mathbf{Y}_{(n_i \times 1)}$ es un vector de respuestas, \mathbf{X}_i es la matriz $n_i \times p$ de covariables con efectos fijos β , \mathbf{Z}_i es una matriz $n_i \times q$ de covariables con efectos aleatorios α_i y ε_i son los vectores de errores.

Los supuestos usuales que se hacen en este modelo son:

1. $\alpha_i \sim N(0, D)$
2. $\varepsilon_i \sim N(0, \Sigma_i)$
3. ε_i y α_i son independientes

donde D y Σ_i son matrices de covarianzas de tamaño $q \times q$ y $n_i \times n_i$ respectivamente.

Estos supuestos implican que

$$\mathbf{Y}_i \sim N\left(\mathbf{X}_i\beta, \mathbf{Z}_iD\mathbf{Z}_i^\top + \Sigma_i\right) \quad (2.2)$$

Para la estimación de los parámetros del modelo se emplea el método de máxima verosimilitud para lo cual considérese γ el vector de todos los parámetros de varianzas y covarianzas halladas en $V_i = \mathbf{Z}_iD\mathbf{Z}_i^\top + \Sigma_i$ y además sea $\theta = (\beta^\top, \gamma^\top)$ un vector s -dimensional de todos los parámetros en el modelo marginal para \mathbf{Y}_i . El estimador de máxima verosimilitud para β condicionado sobre γ está dado por:

$$\hat{\beta}(\gamma) = \left(\sum_{i=1}^N X_i^\top W_i X_i \right)^{-1} \sum_{i=1}^N X_i^\top W_i y_i \quad (2.3)$$

donde $W_i = V_i^{-1}$.

Cuando γ no es conocido, se puede estimar β reemplazando W_i por \widehat{W}_i . Para la estimación de γ se emplea nuevamente el método de máxima verosimilitud (ML) o máxima verosimilitud restringida (REML).

En general, al emplear REML para modelos con una cantidad importante de parámetros en la estructura de la media, se induce una pérdida de grados de libertad que puede afectar considerablemente las estimaciones de las componentes de varianza de tal manera que este método permite obtener estimaciones de las componentes de varianza con un error cuadrático medio, uniformemente más pequeño [Verbeke y Molenberghs \(2000, Cap. 3\)](#).

2.2. Regresión cuantílica

El propósito del análisis de regresión es determinar la relación entre una variable respuesta y un conjunto de variables predictoras, sin embargo, dado que las variaciones de la variable respuesta no pueden ser explicadas en su totalidad por las predictoras, se prefiere describir la respuesta en términos de una medida de tendencia central como la media o la mediana. En el caso de la media, se tiene la regresión clásica a través del método de mínimos cuadrados, el cual permite modelar una amplia familia de relaciones funcionales y ha sido extensamente estudiada. Cuando la medida de centralización empleada no es la media sino la mediana, entonces el método de mínimos cuadrados se cambia por las mínimas distancias absolutas y se tiene la regresión mediana. Esta forma de modelamiento resulta igualmente acertada en la descripción de la relación entre las variable de estudio, sin embargo en situaciones de asimetría de la respuesta, la regresión mediana resulta más informativa que su contraparte.

La mediana es un caso particular de cuantil ($\tau = 0.5$), que describe el punto central de la distribución, de tal manera que la regresión mediana resulta ser un caso particular de regresión cuantílica. En términos generales, la relación entre una variable respuesta y un grupo de variables predictoras puede ser explorada para diversos cuantiles, lo que permite

conocer mejor los cambios que ocurren en la respuesta de acuerdo a los cambios de las variables predictoras.

2.2.1. Cuantiles

Sea X una variable aleatoria real caracterizada por la función de distribución $F_X(x)$, así:

$$F_X(x) = P(X \leq x)$$

de tal manera que se puede definir el τ -ésimo cuantil de X como:

$$F^{-1}(\tau) = \inf \{x : F(X) \geq \tau\} \quad \tau \in (0, 1) \quad (2.4)$$

En general los cuantiles surgen de un problema de optimización, es decir, se desea hallar un valor x de la variable aleatoria, tal que $F_X(x) = \tau$, para lo cual se considera la función de pérdida lineal asimétrica dada por:

$$\rho_\tau(u) = u[\tau - I(u < 0)] \quad \tau \in (0, 1) \quad (2.5)$$

donde $I(\cdot)$ denota la función indicadora. Esta función de pérdida se puede expresar como:

$$\rho_\tau(u) = \begin{cases} u\tau & \text{si } u \geq 0 \\ u(\tau - 1) & \text{si } u < 0 \end{cases}$$

esta función permite hallar el τ -ésimo cuantil muestral como la solución de:

$$\min_{\varepsilon \in \mathbb{R}} \sum_{i=1}^n \rho_\tau(x_i - \varepsilon) \quad (2.6)$$

2.2.2. Modelo

Mediante el método de mínimos cuadrados Koenker (2004) desarrolla una nueva clase de modelos para funciones cuantiles condicionales. En el caso de la media por ejemplo, se plantea solucionar el problema:

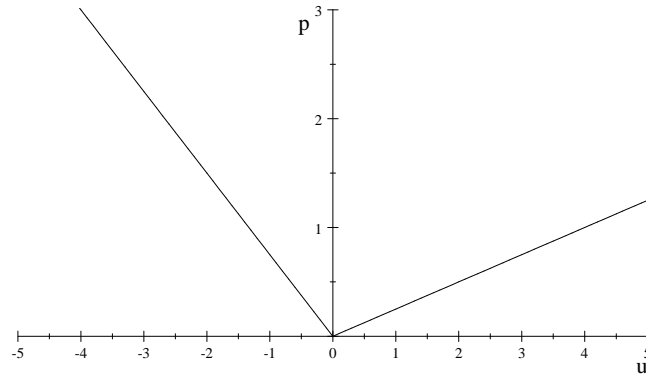


FIGURA 2.1: Función de Pérdida

$$\min_{\mu \in \mathbb{R}} \sum_{i=1}^n (y_i - \mu)^2$$

que bajo el supuesto de una relación lineal $y = x^\top \beta + \varepsilon$ se tiene que la media condicional está dada por $E[y | x] = x^\top \beta$ haciendo que el problema a solucionar sea:

$$\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n (y_i - x_i^\top \beta)^2 \quad (2.7)$$

De manera similar, dado que el τ -ésimo cuantil $\xi(\tau)$ soluciona (2.6) entonces la τ -ésima función cuantil condicional Q_y será:

$$Q_y(\tau | X) = X^\top \beta \quad (2.8)$$

con lo cual el problema de optimización se puede expresar como:

$$\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_\tau(y_i - x_i^\top \beta) \quad (2.9)$$

Ventajas de la regresión cuantílica

La regresión en la cual se modela el valor esperado condicional $E[y | x]$ constituye la manera más común de modelamiento, sin embargo [Koenker \(2004, Cap. 2\)](#) expone ejemplos en los cuales la regresión cuantílica provee ventajas como:

1. Permite modelar los extremos de la variable respuesta.
2. Permite identificar mejor el efecto de las covariables sobre la distribución condicional.
3. En datos con elevada presencia de ceros el modelamiento por regresión cuantílica no se ve afectada por la selección de δ en la transformación $\log(y + \delta)$.
4. Brinda mayor flexibilidad en el modelamiento de datos con altos niveles de variabilidad, describiendo el comportamiento para cada cuantil deseado.

2.2.3. Estimación

En la regresión clásica el método de mínimos cuadrados emplea las distancias verticales desde las observaciones hasta el modelo a ajustar $G(X)$ minimizando para ello la suma de sus cuadrados (2.7). De manera análoga, en el modelo paramétrico de regresión cuantílica se minimizan tales distancias verticales (2.9) pero son ponderadas con τ si están por encima de la función a ajustar, o con $1 - \tau$ si están por debajo de ella. Así el objetivo es hallar el mejor estimador de los parámetros para cada valor de τ . Hao y Naiman (2007, Pag. 33).

Dado que la función de pérdida (2.6) no es diferenciable en el origen, no se logra obtener una solución explícita para los coeficientes de regresión de este modelo, por lo cual es necesario emplear algoritmos numéricos para solucionar el problema de optimización planteado en (2.9). Yu et al. (2003). Particularmente, el método de punto interior brinda una vía de estimación discutida y desarrollada por Portnoy y Koenker (1997), Koenker y D'Orey (1987), Koenker y Park (1996) tal como lo cita Yu et al. (2003).

Otras vías de estimación del modelo paramétrico de regresión cuantílica se obtiene a través de la transformación de Box-Cox o el método en dos etapas Yu et al. (2003). Así mismo, versiones no paramétricas y semiparamétricas son descritas por Yu y Jones (1998) y Cole y Green (1992), respectivamente.

2.2.4. Propiedades de estimadores

En Koenker y Bassett (1978), se demuestran propiedades importantes relacionadas con los cambios que sufren los estimadores de una regresión cuantílica cuando los datos originales son transformados, ya sea la variable respuesta o las variables regresoras, estas se resumen a continuación:

Teorema 2.1. Sea $\widehat{\beta}(\tau; y, X)$ el estimador de regresión cuantílica basado en las observaciones (y, X) . Si A es cualquier matriz no singular de orden $(p \times p)$, $\gamma \in \mathbb{R}^p$ y $a > 0$, entonces para cualquier $\tau \in [0, 1]$ se satisface que:

1. $\widehat{\beta}(\tau; ay, X) = a\widehat{\beta}(\tau; y, X)$
2. $\widehat{\beta}(\tau; -ay, X) = -a\widehat{\beta}(1 - \tau; y, X)$
3. $\widehat{\beta}(\tau; y + X\gamma, X) = \widehat{\beta}(\tau; y, X) + \gamma$
4. $\widehat{\beta}(\tau; y, XA) = A^{-1}\widehat{\beta}(\tau; y, X)$

Los cuantiles poseen además, la propiedad de equivarianza ante transformaciones monótonas, esto significa que el cuantil de una variable aleatoria transformada $h(Y)$ es la transformada de los cuantiles originales de Y . Esto se resume como:

$$Q_{h(Y)}(\tau) = h(Q_Y(\tau))$$

2.2.5. Interpretación de los modelos de regresión cuantílica

La interpretación de las estimaciones de los modelos de regresión cuantílica es análoga a la que se realiza para las estimaciones de modelos de regresión lineal, al considerar el efecto del cambio en una unidad de una covariable sobre la respuesta, sin embargo, en el caso lineal se interpreta en términos de la media, en el caso de la regresión cuantílica se realiza en términos del τ -ésimo cuantil de la variable respuesta. [Hao y Naiman \(2007, Pag. 56\)](#). En general, el cambio realizado sobre las covariables incluye variables dummy al pasar de 0 a 1 o viceversa.

2.2.6. Bondad de ajuste

El ajuste de un modelo de regresión cuantílica es estudiado por [Koenker y Machado \(1999\)](#), donde presenta el coeficiente de ajuste $R^1(\tau)$ el cual se deduce de manera análoga a la regresión clásica por el método de mínimos cuadrados. Sea el modelo lineal de regresión cuantílica dado en (2.8) y considere la partición:

$$Q_y(\tau | x) = x_1^\top \beta_1(\tau) + x_2^\top \beta_2(\tau) \tag{2.10}$$

si $\hat{\beta}(\tau)$ denota el estimador del modelo completo entonces éste será el minimizador de:

$$\hat{V}(\tau) = \min_{\mathbf{b} \in \mathbb{R}^p} \sum \rho_\tau(y_i - x_i^\top \mathbf{b})$$

y sea $\tilde{\beta}(\tau) = (\tilde{\beta}_1(\tau), \mathbf{0})$ el estimador del modelo restringido y minimizador de:

$$\tilde{V}(\tau) = \min_{\mathbf{b}_1 \in \mathbb{R}^{p-q}} \sum \rho_\tau(y_i - x_{1i}^\top \mathbf{b}_1)$$

de tal manera que se define un criterio de bondad de ajuste como

$$R^1(\tau) = 1 - \frac{\hat{V}(\tau)}{\tilde{V}(\tau)} \quad (2.11)$$

Dado que $\hat{V}(\tau) \leq \tilde{V}(\tau)$, entonces los valores del coeficiente $R^1(\tau)$ estarán entre 0 y 1, de manera que valores próximos a cero sugieren un buen ajuste, ya que en este caso $\tilde{V}(\tau)$ será ligeramente mayor que $\hat{V}(\tau)$, indicando que el modelo restringido se ajusta casi como el modelo completo. Así, $R^1(\tau)$ mide el éxito relativo del modelo de regresión cuantílica a un cuantil específico en términos de la suma de residuos absolutos ponderados, constituyendo una medida de ajuste local, en lugar de una global, como lo es el coeficiente de determinación R^2 en la regresión clásica.

Otras medidas de bondad de ajuste para un modelo de regresión cuantílica son presentados por [Koenker \(2005\)](#). El criterio de información de Akaike dado por:

$$AIC = \log \hat{\sigma} + p \quad (2.12)$$

donde p denota la dimensión del modelo y $\hat{\sigma}$ está dado por:

$$\hat{\sigma} = \frac{1}{n} \sum_{i=1}^n \rho_\tau(y_i - x_i^\top \hat{\beta}_n(\tau)) \quad (2.13)$$

Dado que este criterio puede llegar a ser sobrestimado por la dimensión del modelo, entonces se puede considerar el criterio de información de Schwarz dado por:

$$SIC = \log \hat{\sigma} + \frac{1}{2}p \log n \quad (2.14)$$

2.3. Elementos de estadística robusta

El término robustez es entendido en el contexto estadístico como la insensibilidad a pequeñas variaciones de los supuestos. [Huber y Ronchetti \(2009\)](#)[Pag. 2]. Supuestos como el ajuste a un modelo probabilístico, la homogeneidad de varianzas y la simetría de la distribución son algunos de los más usuales, en la mayoría de los métodos estadísticos. Así, estudiar el efecto que las variaciones sobre tales supuestos en los resultados de un método estadístico, constituye un análisis de robustez del mismo. Las herramientas actuales para el estudio de la robustez emplean el concepto matemático de *funcional*, a partir del cual se definen varios nuevos elementos que permiten evaluar las características robustas de una estadística. Para ello, considere una variable aleatoria X con distribución de probabilidad P_θ y función de distribución F_θ , donde $P_\theta \in \mathcal{P} = \{P_\theta : \theta \in \Theta \subseteq \mathbb{R}^p\}$. De esta manera, en muchos casos θ es vista como una funcional definida sobre \mathcal{P} . Así, un estimador de θ basado en una muestra aleatoria (X_1, \dots, X_n) se denota como $\hat{\theta} = T(P_n)$ donde P_n es la distribución de probabilidad empírica, [Jurečkova y Picek \(2009\)](#)[Cap. 1]. Este concepto es usado en la definición de la función de influencia y el punto de ruptura de un estimador.

Un aspecto importante en el análisis de robustez de un estimador es el comportamiento de las colas cuando la distribución subyacente presenta colas pesadas o ligeras. [Jurečkova \(1981\)](#) analiza tal comportamiento, para estimadores de localización, caracterizando para ello las funciones de distribución con colas exponenciales o ligeras y distribuciones con colas algebraicas o pesadas. Las distribuciones con cola exponenciales están caracterizadas como aquellas que satisfacen:

$$\lim_{a \rightarrow \infty} \frac{-\log(1 - F(a))}{ca^r} = 1 \quad c > 0 \quad r > 0 \quad a \neq 0 \quad (2.15)$$

y las distribuciones con colas pesadas o algebraicas aquellas en las que se verifica:

$$\lim_{a \rightarrow \infty} \frac{-\log(1 - F(a))}{m \log a} = 1 \quad m > 0 \quad a \neq 1 \quad (2.16)$$

Estos dos casos son denominados posteriormente Tipo I y Tipo II respectivamente por [He et al. \(1990\)](#). En el caso de estimadores de localización con observaciones independientes X_1, \dots, X_n de una función de distribución simétrica $F(x - \theta)$ absolutamente continua, con densidad $f(z) = f(-z) > 0$, $z \in \mathbb{R}^1$, [Jurečková \(1981\)](#), considera la siguiente medida de desempeño:

$$B(a, T_n) = \frac{-\log P_\theta (|T_n - \theta| > a)}{-\log(1 - F(a))} \quad (2.17)$$

para n fijo y $a \rightarrow \infty$. De esta manera, se espera que la probabilidad del numerador sea pequeña para un a fijo y tienda a cero cuando a tiende a infinito. [Jurečková \(1981\)](#) muestra que esa probabilidad es controlada por el comportamiento de las colas de F y que para cualquier estimador equivariante a traslaciones, $T_n(X_1 + s, \dots, X_n + s) = T_n(X_1, \dots, X_n) + s$ se tiene que:

$$1 \leq \liminf_{a \rightarrow \infty} B(a, T_n) \leq \limsup_{a \rightarrow \infty} B(a, T_n) \leq n \quad (2.18)$$

Esta medida de desempeño de las colas de un estimador es extendida a los modelos lineales por [He et al. \(1990\)](#), donde se considera nuevamente un estimador equivariante a traslaciones ahora expresado como $T_n(Y_1 + X_1 b, \dots, Y_n + X_n b) = T_n(Y_1, \dots, Y_n) + b$ con $b \in \mathbb{R}^p$. Sea el modelo lineal

$$Y = X^\top \beta + \varepsilon \quad (2.19)$$

donde $Y_{(n \times 1)}$ es el vector respuesta, $X_{(n \times p)}$ la matriz diseño de rango p , β el vector de parámetros y ε es el vector de errores independientes con función de distribución absolutamente continua F , tal que

$$0 < F(z) < 1 \quad F(z) + F(-z) = 1 \quad z \in \mathbb{R}^1$$

Entonces la medida de desempeño de colas para los estimadores del modelo (2.19) está dada por:

$$B(a, T_n) = \frac{-\log \{P_\beta (\max_i |x_i(T_n - \beta)| > a)\}}{-\log[1 - F(a)]} \quad 0 < F(a) < 1 \quad (2.20)$$

Para este caso, [He et al. \(1990\)](#) logran establecer un límite superior: $\limsup_{a \rightarrow \infty} B(a, T_n) \leq n$, mientras que para la cota inferior es posible obtener valores de $B(a, T_n)$ menores a

uno para distribuciones con colas tipo I. En general, el análisis desarrollado por estos autores muestra que los estimadores de mínimos cuadrados exhiben un bajo desempeño para distribuciones con colas pesadas (Tipo II), pues una sola observación puede generar un cambio desmesurado en el estimador.

Por otra parte, esta medida de desempeño de colas es considerada para estimadores de regresión cuantílica, presentados en la sección 2.2 y dados por:

$$T_n = \hat{\beta} = \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_{\tau} \left(y_i - x_i^{\top} \beta \right) \quad (2.21)$$

Para los cuales [He et al. \(1990\)](#), obtienen $B(a, T_n) \leq [\text{mín} \{ \tau, 1 - \tau \} n] + 1$, como límite superior. En general, esta medida de desempeño es mejor cuanto más cercana sea al límite superior.

2.4. Muestreo de aceptación y rechazo adaptativo

Un método de obtener una muestra aleatoria de una función de distribución $f(x)$ es el muestreo de aceptación y rechazo. Para esto, se requiere una distribución de muestreo $g(x)$ proporcional a $f(x)$ desde la cual se puede generar fácilmente una muestra. Sin embargo, un requisito para su aplicación es que exista una constante finita c tal que $cg(x) = f(x)$, $\forall x \in D$, donde D denota el dominio de $f(x)$. En general el proceso implica los siguientes pasos:

1. Generar x de $g(x)$
2. Generar $u \sim U(0, 1)$
3. Si $u > f(x)/cg(x)$, entonces rechazar x ; de lo contrario hacer $X_m = x$
4. Devolver X_m

La constante c no solo debe ser calculable, sino que debe ser lo más pequeña posible, pues el número esperado de iteraciones (evaluaciones de $f(x)$) necesarias para aceptar un valor es c , de tal manera que para casos donde tal valor sea muy grande o no sea fácil hallar una función $g(x)$, se tendrá dificultad para el uso de este método.

Una versión mejorada es el Muestreo de Aceptación y Rechazo Adaptativo (ARS por sus siglas en inglés) propuesto por [Gilks y Wild \(1992\)](#), en el cual el número de evaluaciones

de $f(x)$ se logra reducir al incorporar a la distribución de muestreo $g(x)$, la información de cada punto previamente rechazado. Este método aplica a densidades univariadas log-cóncavas con respecto a la medida de Lebesgue y con dominio D en la recta real. La condición de log-concavidad para una función $f(x)$, establece que:

$$\ln f(a) - 2 \ln f(b) + \ln f(c) < 0 \quad \forall a, b, c \in D \quad \text{tal que} \quad a < b < c \quad (2.22)$$

Para la descripción del método considere $S_n = \{x_i; i = 0, \dots, n+1\}$ que denota un conjunto de abscisas en orden ascendente, donde x_0 y x_{n+1} son posiblemente los límites inferior y superior de D respectivamente. Sea además, $L_{ij}(x; S_n)$ la recta que pasa por los puntos $[x_i, \ln f(x_i)]$ y $[x_j, \ln f(x_j)]$ con $1 \leq i \leq j \leq n$ y para cualquier otra pareja (i, j) $L_{ij}(x; S_n)$ no está definida.

Considere ahora la función definida a trozos $h_n(x)$:

$$h_n(x) = \min(L_{i-1,i}(x; S_n), L_{i+1,i+2}(x; S_n)) \quad x_i \leq x \leq x_{i+1}$$

teniendo en cuenta que si b es indefinido entonces $\min(a, b) = \min(b, a) = a$ [Gilks y Wild \(1992\)](#).

Como consecuencia de asumir la log-concavidad de $f(x)$, $h_n(x)$ es una envolvente para $\ln f(x)$, es decir $h_n(x) \geq \ln f(x) \quad \forall x \in D$. La distribución de muestreo para el proceso de aceptación y rechazo esta dada por:

$$g_n = \frac{1}{m_n} \exp h_n(x)$$

donde

$$m_n = \int \exp h_n(x) dx$$

De esta manera, el algoritmo para el muestreo de aceptación y rechazo adaptativo comprende los siguientes pasos:

1. Inicializar n y S_n
2. Generar $X \sim g_n(x)$
3. Generar $u \sim U(0, 1)$
4. Si $u > f(X)/\exp h_n(X)$

entonces rechazar X

Hacer $S_{n+1} = S_n \cup \{X\}$

Actualizar los subíndices de S_{n+1} y ordenar

Hacer $n = n + 1$ y volver al paso 1

de lo contrario aceptar X

Hacer $X_m = X$

5. Devolver X_m

Este algoritmo se encuentra implementado en la función `arms` del paquete `HI` en el software `R`.

Capítulo 3

Regresión cuantílica para datos longitudinales

La extensión de la regresión cuantílica al análisis de datos longitudinales fue propuesta por [Koenker \(2004\)](#), para lo cual considera un modelo de medidas repetidas con un término aleatorio correspondiente al efecto de cada individuo sobre la respuesta. Para la estimación de los efectos fijos emplea el método de mínimos cuadrados penalizados, introduciendo así un parámetro de penalización que debe ser seleccionado. Otras formas de estimación han sido propuestas por [Canay \(2008\)](#), [Geraci y Bottai \(2007\)](#), [Liu y Bottai \(2009\)](#) donde se presenta la estimación en dos etapas y por máxima verosimilitud usando la distribución asimétrica de Laplace univariada y multivariada.

3.1. Modelo

El análisis de datos longitudinales se emplea generalmente sobre variables respuesta continuas donde el supuesto de normalidad para el término aleatorio y de error pueden resultar de gran ayuda en la simplificación del proceso de estimación, sin embargo aún en casos donde tales supuestos no sean correctos o donde la naturaleza de la respuesta no sea continua, es posible desarrollar un modelo apropiado empleando ecuaciones de estimación generalizadas, las cuales aplican a variables respuestas cuya distribución pertenezca a la familia exponencial. Otra clase de modelos ha sido explorada con la aplicación de la regresión cuantílica al análisis de este tipo de datos.

Para esto, considere el siguiente modelo para la función cuantil condicional de la j -ésima observación sobre el i -ésimo individuo

$$Q_{y_{ij}}(\tau | x_{ij}) = x_{ij}^\top \beta(\tau) + \alpha_i \quad i = 1, \dots, N \quad j = 1, \dots, n_i \quad (3.1)$$

en el cual el término α_i tiene solo un efecto de cambio de localización sobre los cuantiles condicionales de la respuesta. La covariable x_{ij} puede depender de τ más no así el término aleatorio α_i . Para estimar este modelo para varios cuantiles, [Koenker \(2004\)](#) propone solucionar el siguiente problema:

$$\min_{(\alpha, \beta)} \sum_{k=1}^q \sum_{j=1}^{n_i} \sum_{i=1}^N w_k \rho_{\tau_k} \left(y_{ij} - \alpha_i - x_{ij}^\top \beta(\tau_k) \right) \quad (3.2)$$

donde w_k son los pesos que controla la influencia relativa de los q cuantiles sobre la estimación de los parámetros α_i .

Generalmente, los α_i se reducen a un valor común a fin de lograr estimadores óptimos bajo el supuesto de normalidad en el modelo lineal clásico para datos longitudinales. Esta reducción también es considerada en el caso de la regresión cuantílica, de tal manera que si n es relativamente mayor que m_i la reducción puede presentar ventajas en el control de la variabilidad introducida por el número de parámetros α_i a ser estimados. Para esto [Koenker \(2004\)](#) considera en el modelo el siguiente término de penalidad:

$$P(\alpha) = \sum_{i=1}^N |\alpha_i|$$

en lugar del supuesto de normalidad, de tal manera que el problema a solucionar ahora es:

$$\min_{(\alpha, \beta)} \sum_{k=1}^q \sum_{j=1}^{n_i} \sum_{i=1}^N w_k \rho_{\tau_k} \left(y_{ij} - \alpha_i - x_{ij}^\top \beta(\tau_k) \right) + \lambda \sum_{i=1}^N |\alpha_i| \quad (3.3)$$

siendo λ un parámetro de penalización que debe ser seleccionado apropiadamente, ya que su influencia sobre la estimación del modelo puede ser importante según [Geraci y Bottai \(2007\)](#). Un estimador óptimo de λ es obtenido por [Lamarche \(2006\)](#), a partir del cual obtiene estimadores de mínima varianza para el modelo de regresión cuantílica en datos longitudinales.

Por otra parte, la distribución asimétrica de Laplace propuesta por [Koenker y Machado \(1999\)](#) es considerada por [Geraci y Bottai \(2007\)](#), para desarrollar la estimación mediante el método de máxima verosimilitud. Para ello parten de la función cuantil condicional mixta de la respuesta y_{ij} :

$$Q_{y_{ij}|\alpha_i}(\tau | x_{ij}, \alpha_i) = x_{ij}^\top \beta + \alpha_i \quad i = 1, \dots, N \quad j = 1, \dots, n_i \quad (3.4)$$

donde $Q_{y_{ij}|\alpha_i}(\cdot) \equiv F_{y_{ij}|\alpha_i}^{-1}$ es la inversa de la función de distribución acumulada de la respuesta condicionada sobre los efectos aleatorios α_i . En este punto, se considera el supuesto según el cual, la variable respuesta sigue una distribución asimétrica de Laplace ([Apéndice A.1](#)) es decir $y_{ij} \sim DAL(\mu, \sigma, \tau)$ de tal manera que si $\mu_{ij} = x_{ij}^\top \beta + \alpha_i$ entonces la respuesta seguirá una distribución de acuerdo a $y_{ij} \sim DAL(\beta, \alpha_i, \sigma, \tau)$. Adicionalmente, otros supuestos del modelo son:

1. Los efectos aleatorios α_i son i.i.d. de acuerdo con determinada función de densidad f_α caracterizada por algún parámetro $\varphi(\tau)$.
2. El término de error aleatorio ε_{ij} se asume independiente de los efectos aleatorios α_i .

Sea el vector $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{in_i})$ de observaciones correspondientes al i –ésimo individuo, de tal manera que la función de distribución de éste es:

$$f(\mathbf{y}_i | \beta, \alpha_i, \sigma) = \prod_{j=1}^{n_i} f(y_{ij} | \beta, \alpha_i, \sigma)$$

y por lo tanto la distribución conjunta de (\mathbf{y}_i, α_i) está dada por:

$$f(\mathbf{y}_i, \alpha_i | \beta, \sigma, \varphi) = f(\mathbf{y}_i | \beta, \alpha_i, \sigma) f_\alpha(\alpha_i | \varphi) \quad (3.5)$$

esto a su vez permite expresar la distribución conjunta de (\mathbf{y}, α) como:

$$f(\mathbf{y}, \alpha | \beta, \sigma, \varphi) = \prod_{i=1}^N f(\mathbf{y}_i | \beta, \alpha_i, \sigma) f_\alpha(\alpha_i | \varphi) \quad (3.6)$$

3.2. Estimación

La estimación del modelo (3.1) es desarrollada por Geraci y Bottai (2007) empleando máxima verosimilitud, para lo cual obtienen la función de densidad marginal de \mathbf{y} por integración a partir de (3.6):

$$f(\mathbf{y} | \eta) = \int_{\mathbb{R}^N} f(\mathbf{y}, \alpha | \eta) d\alpha \quad (3.7)$$

donde $\eta = (\beta, \sigma, \varphi)$. Esto permite plantear la inferencia para η basada en la verosimilitud marginal:

$$\mathbf{L}(\eta; \mathbf{y}) = \sum_{i=1}^N \mathbf{L}(\eta; \mathbf{y}_i)$$

Sin embargo, debido a que la integral en (3.7) puede no tener una solución explícita por vía analítica, Geraci y Bottai (2007) hacen uso del algoritmo EM para su estimación.

3.2.1. Paso E

El paso E para el i -ésimo individuo en la iteración $(t + 1)$ se expresa como:

$$Q_i(\eta | \eta^t) = E \{ \ell(\eta; \mathbf{y}_i, \alpha_i) | \mathbf{y}_i, \eta^t \}$$

que es equivalente a:

$$Q_i(\eta | \eta^t) = \int \{ \log f(\mathbf{y}_i | \beta, \alpha_i, \sigma) + \log f(\alpha_i | \varphi) \} \cdot f(\alpha_i | \mathbf{y}_i, \eta^t) d\alpha_i \quad (3.8)$$

donde $\eta^{(t)} = (\beta^{(t)}, \sigma^{(t)}, \varphi^{(t)})$, $\ell(\eta; \mathbf{y}_i, \alpha_i)$ es el logaritmo de la verosimilitud para el i -ésimo individuo y la esperanza es tomada con respecto a la distribución de los efectos aleatorios no observados α_i dadas las observaciones \mathbf{y}_i , cuya densidad es:

$$f(\alpha_i | \mathbf{y}_i, \eta) \propto f(\mathbf{y}_i | \beta, \alpha_i, \sigma) f_\alpha(\alpha_i | \varphi) \quad (3.9)$$

Para aproximar tal valor esperado Geraci y Bottai (2007) y Liu y Bottai (2009), proponen usar el método de Monte Carlo obteniendo para ello una muestra $\mathbf{v}_i = (\mathbf{v}_{i1}, \dots, \mathbf{v}_{im_i})$ a

partir de la distribución condicional de α_i dada en (3.9), de tal manera que la aproximación del valor esperado se obtiene mediante:

$$Q_i^*(\eta | \eta^t) \approx \frac{1}{m_i} \sum_{j=1}^{m_i} \{\log f(\mathbf{y}_i | \beta, \mathbf{v}_{ij}, \sigma) + \log f(\mathbf{v}_{ij} | \varphi)\} \quad (3.10)$$

Esto permite considerar la variable $\tilde{y}_{ikj} = y_{ij} - \mathbf{v}_{ik}$ con lo cual los \tilde{y}_{ij} pueden considerarse independientes y así emplear la regresión cuantílica usual para estimar los parámetros.

3.2.2. Paso M

El estimador de máxima verosimilitud del parámetro η para una función cuantil τ se obtiene mediante el siguiente proceso iterativo:

Paso 1. Inicializar el parámetro $\eta^{(t)} = (\beta^{(t)}, \sigma^{(t)}, \varphi^{(t)})$ para $t = 0$ y sustituir en (3.9).

Paso 2. Para cada individuo obtener una muestra $\mathbf{v}_i^{(t)} = (\mathbf{v}_{i1}^{(t)}, \dots, \mathbf{v}_{im_i}^{(t)})$ a partir de (3.9) de tamaño m_i mediante algún método de MCMC; por ejemplo mediante el muestreador de Gibbs y muestreo de aceptación y rechazo adaptativo propuesto por Gilks et al. (1995).

Paso 3. Obtener la estimación de $\beta^{(t+1)}$ usando regresión cuantílica:

$$\beta^{(t+1)} = \arg \min_{\beta \in B} \sum_{i=1}^N \frac{1}{m_i} \sum_{k=1}^{m_i} \sum_{j=1}^{n_i} \rho_{\tau}(\tilde{y}_{ikj} - \mathbf{x}_{ij}^{\top} \beta)$$

y calcular

$$\sigma^{(t+1)} = \frac{1}{\sum_{i=1}^N n_i} \frac{1}{\sum_{i=1}^N m_i} \sum_{i=1}^N \sum_{k=1}^{m_i} \sum_{j=1}^{n_i} \rho_{\tau}(\tilde{y}_{ikj} - \mathbf{x}_{ij}^{\top} \beta^{(t+1)})$$

$$\varphi^{(t+1)} = \widehat{\Phi}(\mathbf{v}_1, \dots, \mathbf{v}_N)$$

donde $\widehat{\Phi}$ es un estimador de máxima verosimilitud de φ , condicionado sobre \mathbf{y} .

Paso 4. Hacer $t = t + 1$ y repetir los pasos del 1 al 3 hasta lograr convergencia del parámetro η .

3.3. Distribución de Efectos Aleatorios

Para la distribución de los efectos aleatorios α_i es posible considerar una amplia gama de opciones, sin embargo una selección frecuente es el uso de efectos aleatorios normalmente distribuidos, más aun, [Liu y Bottai \(2009\)](#) afirman que la selección de la distribución del término aleatorio no parece ser crucial. En general, si los efectos aleatorios son independientes e idénticamente distribuidos de acuerdo con $\alpha_i \sim N(0, \varphi)$, entonces para el i -ésimo individuo, la distribución conjunta $f(\mathbf{y}_i, \alpha_i | \eta)$ es:

$$f(\mathbf{y}_i, \alpha_i | \eta) = f(\alpha_i | \varphi) \prod_{j=i}^{n_i} f(\mathbf{y}_{ij} | \beta, \alpha_i, \sigma)$$

$$f(\mathbf{y}_i, \alpha_i | \eta) = \left[\frac{\tau(1-\tau)}{\sigma} \right]^{n_i} \frac{1}{\sqrt{2\pi\varphi}} \exp \left[- \sum_{j=1}^{n_i} \left\{ \rho_\tau \left(\frac{y_{ij} - \mu_{ij}}{\sigma} \right) \right\} - \frac{1}{2\varphi} \alpha_i^2 \right] \quad (3.11)$$

Esta distribución conjunta es empleada para obtener las muestras de los efectos aleatorios, mediante el uso coordinado del muestreador de Gibbs y el método de aceptación y rechazo adaptativo. Esta función cumple con el requisito de log-concavidad en α_i mencionado en la sección 2.4.

3.4. Robustez a distribución

3.4.1. Cálculo de $B(a, T_n)$

La medida de desempeño de colas bajo simetría $B(a, T_n)$ resulta considerablemente compleja para ser obtenida matemáticamente, más aún si se considera que los estimadores de regresión cuantílica no tienen una solución explícita como en el caso de los mínimos cuadrados.

Por tal motivo se propone calcular esta medida al término del proceso de estimación descrito en la sección 3.2, para lo cual se debe tener en cuenta los siguientes aspectos:

1. El proceso de estimación de los parámetros debe ser iterado \mathcal{N} veces.

2. Para el caso del modelo de medidas repetidas (3.1) se tiene que $x_{ij} = x_{i'j}$ para todo $i \neq i'$ y $j = 1, \dots, n_i$, donde n_i es el número total de mediciones realizadas sobre cada individuo.
3. $T_n = \hat{\beta}$ donde $\hat{\beta}$ es el vector de parámetros ($p \times 1$) de efectos fijos del modelo.

Sea $d_{lj} = \max_j |x_j(T_n - \beta)|$ donde x_j es el vector ($1 \times p$) de covariable correspondiente a la j -ésima medición sobre cada uno de los individuos y l denota la iteración, de tal manera que la medida de desempeño de colas $B(a, T_n)$ se puede obtener como:

$$\hat{B}(a, T_n) = \frac{-\log \left\{ \frac{1}{N} \sum_{l=1}^N I(d_{lj} > a) \right\}}{-\log \{1 - F(a)\}} \quad (3.12)$$

donde $I(\cdot)$ es la función indicadora, F es la función de distribución del término de error y a es un valor fijo.

3.4.2. Probabilidad del sesgo

La medida de desempeño de colas $B(a, T_n)$ propuesta por He et al. (1990) tiene como principal supuesto, la simetría de la función de distribución F expresada como $F(z) + F(-z) = 1$, sin embargo, un amplio grupo de distribuciones no satisfacen esta condición y su efecto sobre los estimadores podría resultar importante si el término de error se distribuye de manera asimétrica.

Considere el modelo lineal $Y = X\beta + \varepsilon$ donde Y es el vector de respuestas ($n \times 1$), X es la matriz diseño de orden ($n \times p$) y rango p , β es el vector de parámetros ($p \times 1$) y ε es el vector de errores independientes con función de distribución absolutamente continua F , tal que $0 < F(z) < 1$ con $z \in \mathbb{R}^1$. Sea además a un valor en el dominio de F entonces el comportamiento de T_n puede ser descrito por la probabilidad:

$$P_c(a, T_n) = P_\beta \left[\min_i (x_i(T_n - \beta) < -a) \right] + P_\beta \left[\max_i (x_i(T_n - \beta) > a) \right] \quad (3.13)$$

Un comportamiento ideal supone una rápida convergencia de P_c hacia cero, lo que se traduce en que los estimadores se aproximan rápidamente a los parámetros.

Capítulo 4

Simulación

4.1. Modelo

Para el estudio se utilizó un modelo lineal mixto con intercepto aleatorio dado por:

$$y_{ij} = 1 + 2x_{ij} + \alpha_i + \varepsilon_{ij} \quad i = 1, \dots, 50 \quad j = 1, \dots, 15 \quad (4.1)$$

donde x_{ij} es la covariable cuyos valores son fijos através de toda la simulación y están dados por $x_{ij} = j$, α_i es el efecto aleatorio correspondiente a cada individuo y ε_{ij} es el término de error del i – *ésimo* individuo en la j – *ésima* medición. Así mismo, para la estructura de correlación entre las medidas se utilizó el modelo de correlación exponencial descrito por [Diggle et al. \(2002, Pag. 56\)](#) según el cual el elemento ρ_{st} está dado por:

$$\rho_{st} = \sigma^2 \exp(-\phi|x_s - x_t|) \quad (4.2)$$

para lo que se tomó a $\phi = 0.8$. Este modelo, considera que la correlación decae exponencialmente en proporción a la distancia en el tiempo. Entre las diversas estructuras para la matriz de correlación, esta es una opción general y lógica para el modelo propuesto. El valor de ϕ se seleccionó de tal manera que se obtuviera una correlación clara y fuerte entre las observaciones.

Para el término de error se empleó la distribución asimétrica de Laplace, normal estándar sesgada, Ji-cuadrado y t-Student. Por su parte, los efectos aleatorios fueron considerados normales con media cero y varianza uno para todos los escenarios simulados, en razón a

los resultados presentados por [Liu y Bottai \(2009\)](#).

La estimación de parámetros fue realizada en cada escenario para tres cuantiles diferentes $\tau = (0.05, 0.5, 0.8)$ iterando 1000 veces en cada uno. Así mismo, se estimó el modelo lineal mixto clásico a fin de contrastar los resultados para $\tau = 0.5$ con los del procedimiento regular. En cada iteración, el tamaño de la simulación fue determinado de manera dinámica de acuerdo con [Eickhoff et al. \(2004\)](#) y un máximo de 3000 ciclos en cada una, como se describe en la sección 4.3.

Finalmente, se calculó la medida de desempeño de colas $B(a, T_n)$ descrita en la sección 3.9 y el sesgo relativo SR como sigue:

$$SR(\hat{\beta}_k) = \frac{1}{\mathcal{N}} \left(\sum_{l=1}^{\mathcal{N}} \frac{\hat{\beta}_k^l}{\beta_k} - 1 \right) \quad (4.3)$$

donde β_k $k = 0, 1$ es el valor real del parámetro, \mathcal{N} es el número de iteraciones. El valor real de β_0 se determina adicionando el τ -ésimo cuantil estandarizado de la distribución asignada al término de error al valor real en cada escenario.

4.2. Selección de parámetros

Las distribuciones usadas para el término de error en el modelo (4.1) son tales que permiten explorar diferentes niveles de simetría con los cuales evaluar la estimación de los parámetros por regresión cuantílica. En los casos donde los parámetros seleccionados para la distribución modifican el valor de la media, se centró apropiadamente. Así mismo, se consideró casos simétricos con los cuales evaluar la influencia de colas pesadas.

Distribución	Notación	Parámetros	C.Simetría
Normal sesgada	$NS(\mu, \sigma^2, \gamma)$	$\gamma = -20$	-0.99
		$\gamma = 0$	0.00
		$\gamma = 20$	0.99
Asimétrica de Laplace	$DAL(\mu, \sigma, \tau)$	$\tau = 0.05$	-1.99
		$\tau = 0.50$	0.00
		$\tau = 0.80$	1.80
t-Student	t_ν	$\nu = 3$	0.00
		$\nu = 5$	0.00
Ji-Cuadrado	χ_ν^2	$\nu = 2$	2.00
		$\nu = 4$	1.41

TABLA 4.1: Distribuciones usadas para el término de error

Para la distribución normal sesgada y asimétrica de Laplace se tomó el caso estándar al considerar $\mu = 0$ y $\sigma^2 = 1$.

4.3. Criterio de convergencia y tamaño de la simulación

El criterio de convergencia para la estimación de los parámetros está basado en la metodología propuesta por [Eickhoff et al. \(2004\)](#), según la cual, para el algoritmo MCEM resulta más apropiado un criterio de convergencia basado en distancias de verosimilitudes en lugar de distancias entre estimaciones sucesivas de los parámetros. Para ello se considera $\hat{f}_i^{(t)}$ la aproximación de $f(\mathbf{y} | \eta)$ dada en la expresión (3.7), de tal manera que una estimación de la verosimilitud es:

$$\hat{\ell}(\eta^{(t)}) = \sum_{i=1}^N \log \hat{f}_i^{(t)}$$

donde $\hat{f}_i^{(t)} = \frac{1}{N} \sum_{m=1}^N f(\mathbf{y}_i | \mathbf{v}_{im}, \eta^{(t)})$. De esta manera, se establece como criterio de convergencia que:

$$|\hat{\ell}(\eta^{(t)}) - \hat{\ell}(\eta^{(t-1)})| < \frac{\sqrt{n}\delta}{3} \quad (4.4)$$

donde δ es un valor seleccionado para la distancia en verosimilitud.

Por otra parte, el tamaño de la simulación se determina para cada paso del algoritmo MCEM, para lo que se estima la varianza de los parámetros en el paso t como sigue:

$$\hat{V}(\eta^{(t-1)}) = \frac{1}{N-1} \sum_{i=1}^N \left[\frac{\sum_{i=1}^N f^2(\mathbf{y}_i | \mathbf{v}_{im}, \eta^{(t)})}{(\hat{f}_i^{(t-1)})^2} - N \right] \quad (4.5)$$

De esta forma, [Eickhoff et al. \(2004\)](#) establecen que el tamaño de la simulación en el paso t esta dado por:

$$\mathcal{N}_0^{(t)} \geq \frac{\hat{V}(\eta^{(t)})}{\varepsilon \delta_1^2} \quad (4.6)$$

donde $\delta_1 = \frac{\sqrt{n}\delta}{3}$ y ε es la probabilidad que la distancia entre la verosimilitud de dos iteraciones sucesivas sea mayor a δ_1 . Finalmente, si $\mathcal{N}_0^{(t)} > \mathcal{N}_0^{(t-1)}$ entonces el tamaño de la simulación para el paso $(t+1)$ es la diferencia entre los anteriores valores, de lo contrario el tamaño de la simulación continua siendo el correspondiente al paso $(t-1)$.

En general la selección de δ y ε debe ser cuidadosa para evitar tamaños que puedan exceder la capacidad computacional disponible. En este estudio se tomo $\delta = 10$ y $\varepsilon = 0.05$.

Capítulo 5

Resultados

5.1. Estimación de efectos fijos

La estimación puntual, el sesgo relativo y la varianza para β_1 y β_0 se presentan en la siguientes tablas para cada uno de los escenarios considerados:

	β_0				β_1			
	Cuantil							
	MLM	0.05	0.50	0.80	MLM	0.05	0.50	0.80
$\psi = -20$								
Estimación	0.9933	-1.3081	1.0497	2.1410	2.0002	2.0039	2.0016	2.0013
Sesgo	-0.0067	0.4139	-0.1272	0.1270	0.0001	0.0078	0.0049	0.0055
Var	0.0240	0.1036	0.0333	0.0527	0.0000	0.0004	0.0002	0.0002
$\psi = 0$								
Estimación	0.9998	-1.6318	0.9943	2.3747	1.9998	2.0001	2.0008	1.9976
Sesgo	-0.0002	0.0504	-0.0057	0.2894	-0.0001	0.0103	0.0055	0.0071
Var	0.0279	0.0357	0.0422	0.0709	0.0001	0.0007	0.0002	0.0003
$\psi = 20$								
Estimación	1.0110	-1.4594	0.9515	2.3150	1.9999	1.9958	1.9981	2.0015
Sesgo	0.0110	0.6353	0.1932	0.2845	-0.0001	0.0098	0.0057	0.0066
Var	0.0248	0.1086	0.0393	0.0497	0.0001	0.0006	0.0002	0.0003

TABLA 5.1: Estimaciones de β_0 y β_1 en la Normal sesgada $NS(\mu, \sigma^2, \psi)$

En general se encuentra que la estimación de la pendiente presenta varianzas bajas y sesgos relativos bajos en todos los escenarios, sin embargo el intercepto presenta niveles de

variabilidad mucho mayores, así como elevados sesgos que superan en ocasiones el 40%. Los sesgos más bajos se encuentran cuando la distribución de los errores es simétrica y se estima para el cuantil $\tau = 0.5$, que corresponde a la regresión mediana.

	β_0				β_1			
	Cuantil							
	MLM	0.05	0.50	0.80	MLM	0.05	0.50	0.80
$\tau = 0.05$								
Estimación	1.0044	-1.3503	1.0097	2.2709	1.9995	2.0003	1.9992	2.0009
Sesgo	0.0044	-0.0320	0.0530	0.4106	-0.0002	0.0082	0.0059	0.0070
Var	0.0236	0.1077	0.0391	0.0620	0.0001	0.0004	0.0002	0.0003
$\tau = 0.5$								
Estimación	1.0007	-1.6711	1.0021	2.3052	1.9995	2.0011	1.9994	2.3052
Sesgo	0.0007	0.6602	0.0021	0.3989	-0.0002	0.0093	0.0049	0.3989
Var	0.0262	0.1304	0.0367	0.0602	0.0000	0.0005	0.0002	0.0602
$\tau = 0.8$								
Estimación	0.9868	-1.3534	1.0689	2.3026	2.0007	1.9998	2.0007	2.0011
Sesgo	-0.0132	0.4066	-0.1094	0.3329	0.0003	0.0086	0.0051	0.0053
Var	0.0262	0.1056	0.0395	0.0578	0.0000	0.0005	0.0002	0.0002

TABLA 5.2: Estimaciones de β_0 y β_1 en la distribución Asimétrica de Laplace $DAL(\mu, \sigma, \tau)$

Aunque la distribución asimétrica de Laplace es la base del proceso de estimación, los sesgos resultan ser altos en el caso del intercepto, particularmente en el caso de $\tau = 0.8$. Adicionalmente, se observa un incremento en el valor del mismo parámetro cuando el cuantil de estimación es 0.8 para cada uno de los tres valores de τ , esto sugiere un efecto del cuantil de estimación sobre los parámetros, lo cual no resulta deseable.

En la distribución t-Student, el aumento en grados de libertad no parece tener un efecto importante sobre las estimaciones, en tanto que para la distribución Ji-cuadrado se encuentran cambios apreciables en las estimaciones de β_0 para τ igual a 0.05 y 0.5.

Las estimaciones puntuales y el sesgo muestran una clara variación en los valores reportados para el intercepto, sin embargo estos cambios resultan, en algunos casos mayores de los esperados, sugiriendo un efecto del cuantil de estimación, lo que no es considerado en el método de máxima verosimilitud. Este efecto si es controlado en el método propuesto por [Koenker \(2004\)](#). Otro aspecto que presenta variación respecto al cuantil, es la varianza de la estimación la que tiende a ser menor cuando $\tau = 0.5$ a la vez que se incrementa para

	β_0				β_1			
	Cuantil				Cuantil			
	MLM	0.05	0.50	0.80	MLM	0.05	0.50	0.80
$\nu = 3$								
Estimación	1.0281	-1.5483	1.0339	2.2475	1.9997	1.9976	1.9989	2.0011
Sesgo	0.0281	0.1163	0.0339	0.4362	-0.0001	0.0096	0.0048	0.0058
Var	0.0188	0.1606	0.0297	0.0462	0.0001	0.0006	0.0002	0.0002
$\nu = 5$								
Estimación	1.0266	-1.6361	1.0226	2.2770	1.9979	2.0016	1.9980	2.0010
Sesgo	0.0266	0.3172	0.0226	0.3298	-0.0011	0.0096	0.0058	0.0068
Var	0.0225	0.1308	0.0309	0.0608	0.0001	0.0006	0.0002	0.0003

TABLA 5.3: Estimaciones de β_0 y β_1 en la distribución t-Student t_ν

	β_0				β_1			
	Cuantil				Cuantil			
	MLM	0.05	0.50	0.80	MLM	0.05	0.50	0.80
$\nu = 2$								
Estimación	1.0540	-2.1427	0.7640	4.1960	1.9995	1.9959	1.9961	1.9952
Sesgo	0.0540	0.0995	0.1022	0.6071	-0.0003	0.0109	0.0083	0.0216
Var	0.0358	0.1178	0.0634	0.3981	0.0002	0.0008	0.0004	0.0032
$\nu = 4$								
Estimación	1.0657	-4.6503	0.3896	4.1193	2.0018	1.9998	2.0021	2.0064
Sesgo	0.0657	0.4139	-0.4957	0.4187	0.0009	0.0177	0.0149	0.0221
Var	0.1381	0.2525	0.1694	0.2804	0.0010	0.0020	0.0015	0.0029

TABLA 5.4: Estimaciones de β_0 y β_1 en la distribución Ji-Cuadrado χ_ν^2

otros cuantiles. Ésto no ocurre en el caso de la pendiente, pues sus estimaciones presentan bajo sesgo y varianza.

El uso específico de la distribución asimétrica de Laplace para el término de error, no muestra un resultado importante en comparación con las otras distribuciones consideradas, lo cual indica que alteraciones en este supuesto pueden no ser un factor significativo en las estimaciones de los parámetros.

5.2. Estimación de varianzas y criterio de ajuste

La desviación estándar σ del término de error, la varianza φ del efecto aleatorio y el criterio de información de Akaike, también fueron estimados. Las tablas 5.5, 5.6 y 5.7 presentan los resultados para todos los escenarios simulados. Para el caso de σ se encuentra que los valores mas cercanos al correcto ($\sigma = 1$) se obtienen para $\tau = 0.5$ en todas las distribuciones empleadas, sin embargo el modelo lineal mixto logra estimar con mayor exactitud dicho valor. La estimación para cuantiles diferentes a $\tau = 0.5$ genera valores notablemente sesgados del parámetro mostrando ser menores en casi todos los casos.

Distribución	Parámetro	MLM	Cuantil		
			0.05	0.50	0.80
Normal Sesgada	-20	0.995	0.191	0.819	0.557
	0	0.988	0.212	0.823	0.571
	20	1.028	0.209	0.837	0.578
t-student	3	0.999	0.222	0.827	0.546
	5	1.002	0.217	0.843	0.566
Ji-cuadrado	2	0.988	0.234	1.044	1.361
	4	0.979	0.356	1.694	1.364
Asimétrica de Laplace	0.05	0.990	0.197	0.825	0.590
	0.5	0.988	0.216	0.816	0.572
	0.8	0.991	0.201	0.816	0.560

TABLA 5.5: Estimaciones de σ

En general, la estimación de φ resulta más consistente que σ . Los valores estimados son menores al valor correcto ($\varphi = 1$) pero no presentan variaciones importantes en relación al cuantil de estimación. Se destaca el valor obtenido para este parámetro en el modelo lineal mixto cuando la distribución del error sigue una Ji-cuadrado, pues los valores de $varphi$ son cercanos a los grados de libertad empleados.

Los valores del criterio de información de Akaike son mayores en el caso de la regresión cuantílica en comparación con los obtenidos en el modelo lineal mixto. La tendencia general es a disminuir cuando el cuantil de estimación se aproxima a $\tau = 0.5$. Así mismo, los valores en la distribución Ji-cuadrado son más altos que en las otras distribuciones, conservando la tendencia antes mencionada.

Distribución	Parámetro	MLM	Cuantil		
			0.05	0.50	0.80
Normal sesgada	-20	1,001	0,808	0,792	0,908
	0	1,000	0,868	0,806	0,789
	20	0,999	0,832	0,841	0,784
t-student	3	0,969	1,112	0,868	0,718
	5	0,994	0,960	0,869	0,787
Ji-cuadrado	2	2,011	0,722	0,805	1,278
	4	4,005	0,750	0,869	1,306
Asimétrica de Laplace	0,05	1,000	0,726	0,821	0,839
	0,5	1,001	0,965	0,801	0,800
	0,8	0,994	0,760	0,801	0,791

TABLA 5.6: Estimaciones de φ

Distribución	Parámetro	MLM	Cuantil		
			0.05	0.50	0.80
Normal sesgada	-20	2281,98	3539,54	3247,44	3347,86
	0	2280,09	3701,94	3252,78	3373,13
	20	2283,18	3686,70	3280,39	3385,36
t-student	3	2231,77	3750,03	3265,96	3296,60
	5	2273,09	3725,90	3296,21	3355,73
Ji-cuadrado	2	3264,32	3857,64	3617,80	4696,87
	4	4253,20	4502,27	4354,53	4699,19
Asimétrica de Laplace	0,05	2280,01	3594,88	3259,12	3418,95
	0,5	2280,69	3722,14	3240,83	3373,25
	0,8	2270,84	3622,47	3241,49	3343,94

TABLA 5.7: Estimaciones del criterio de información de Akaike

5.3. Desempeño de colas

Con las diferentes distribuciones consideradas para el término de error en este estudio, se determinó el valor de $B(a, T_n)$ para valores de a entre 0 y 3 en los casos simétricos y la probabilidad de sesgos para todos los escenarios. La figura 5.1 presenta las curvas obtenidas para las distribuciones simétricas. De esta manera, se obtiene que en todos los casos el mejor desempeño se obtiene para el cuantil 0.5 en la distribución asimétrica de Laplace y normal. En general, en la evaluación de los estimadores mediante $B(a, T_n)$ se espera un rápido crecimiento de la medida, sin embargo, la figura 5.1 muestra que para el cuantil 0.8 tal comportamiento es más lento.

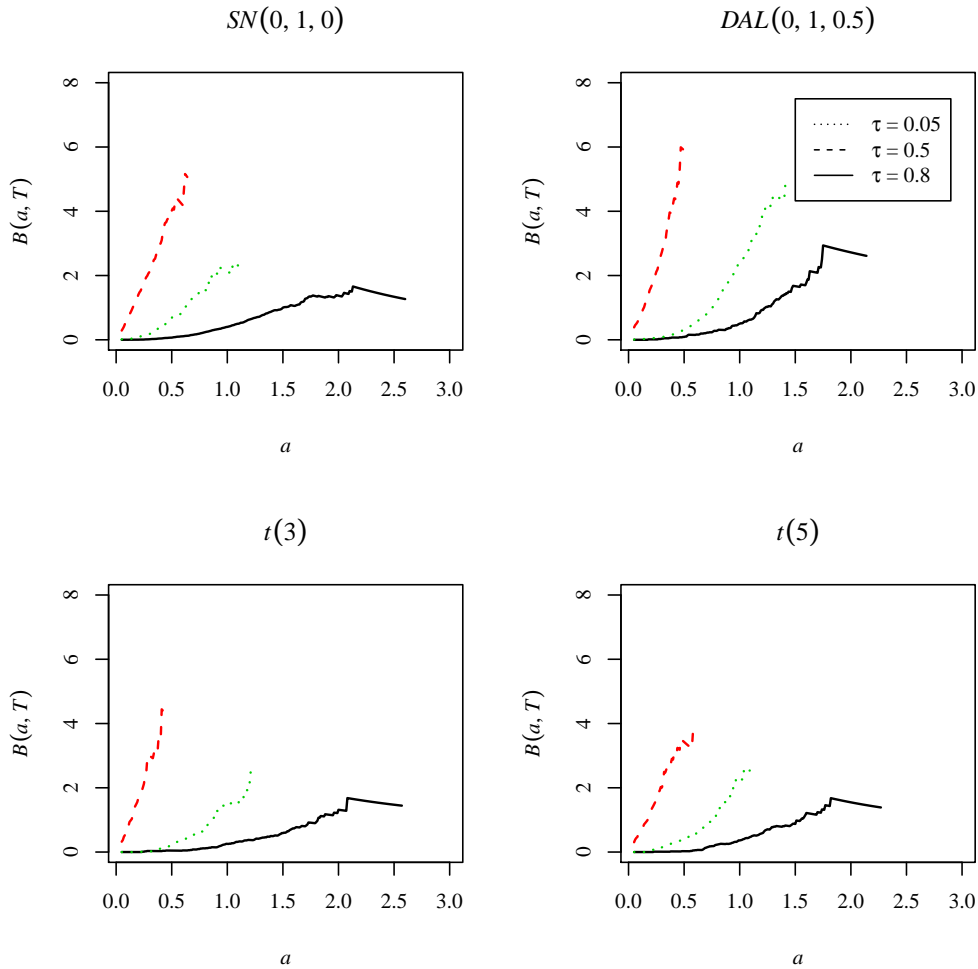


FIGURA 5.1: Medida de desempeño de colas $B(a, T_n)$ para distribuciones simétricas

En general, para la distribución normal sesgada con parámetro de forma -20 y la distribución asimétrica de Laplace con $\tau = 0.8$ se encuentra que la probabilidad P_c se reduce a cero mucho más rápido que en otros escenarios.

Por su parte, la probabilidad de sesgo de los estimadores tarda en decaer para el cuantil $\tau = 0.8$ en la distribución Ji-cuadrado, sin embargo tiende rápidamente a cero cuando se considera el caso particular de la regresión mediana. La distribución asimétrica de Laplace muestra un comportamiento más consistente al obtenerse probabilidades de sesgo menores para todos los cuantiles a partir de $a > 2$. De esta manera, se encuentra que las probabilidades de sesgo obtenidas resultan importantes y pueden afectar las estimaciones para valores de τ cercanos a cero o uno.

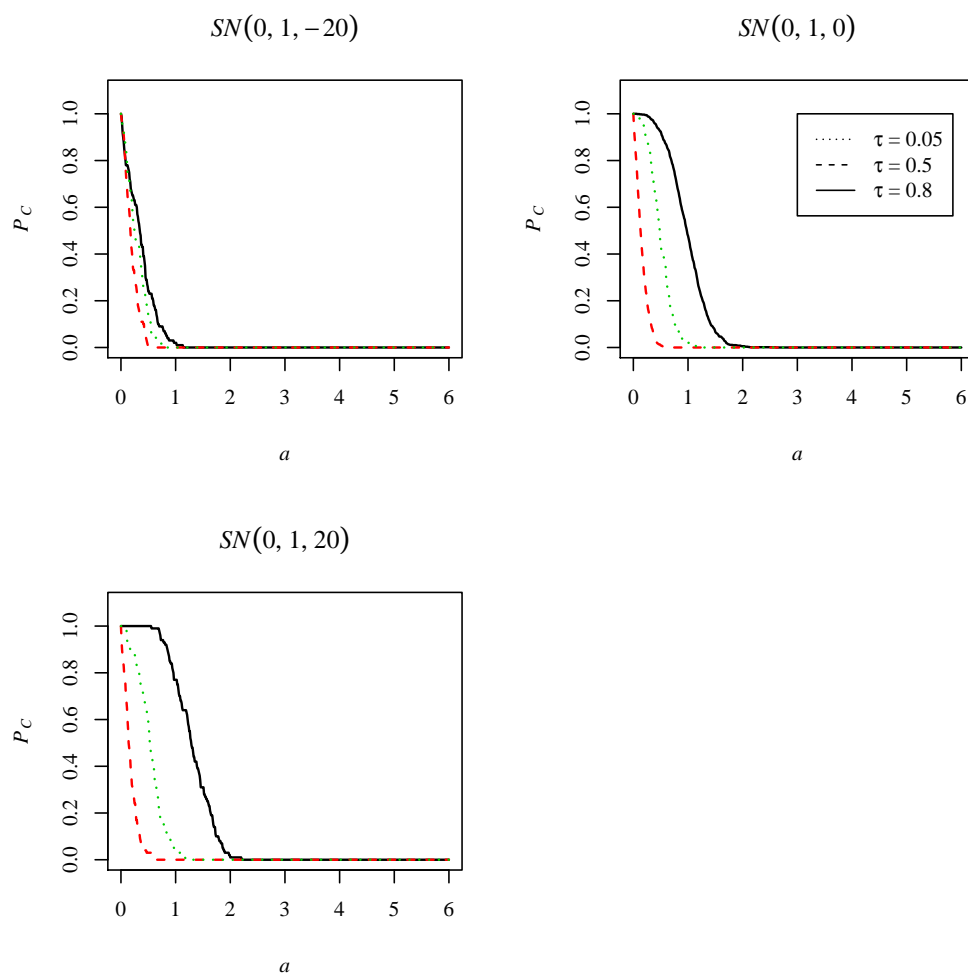


FIGURA 5.2: Probabilidad de sesgos para la distribución normal sesgada

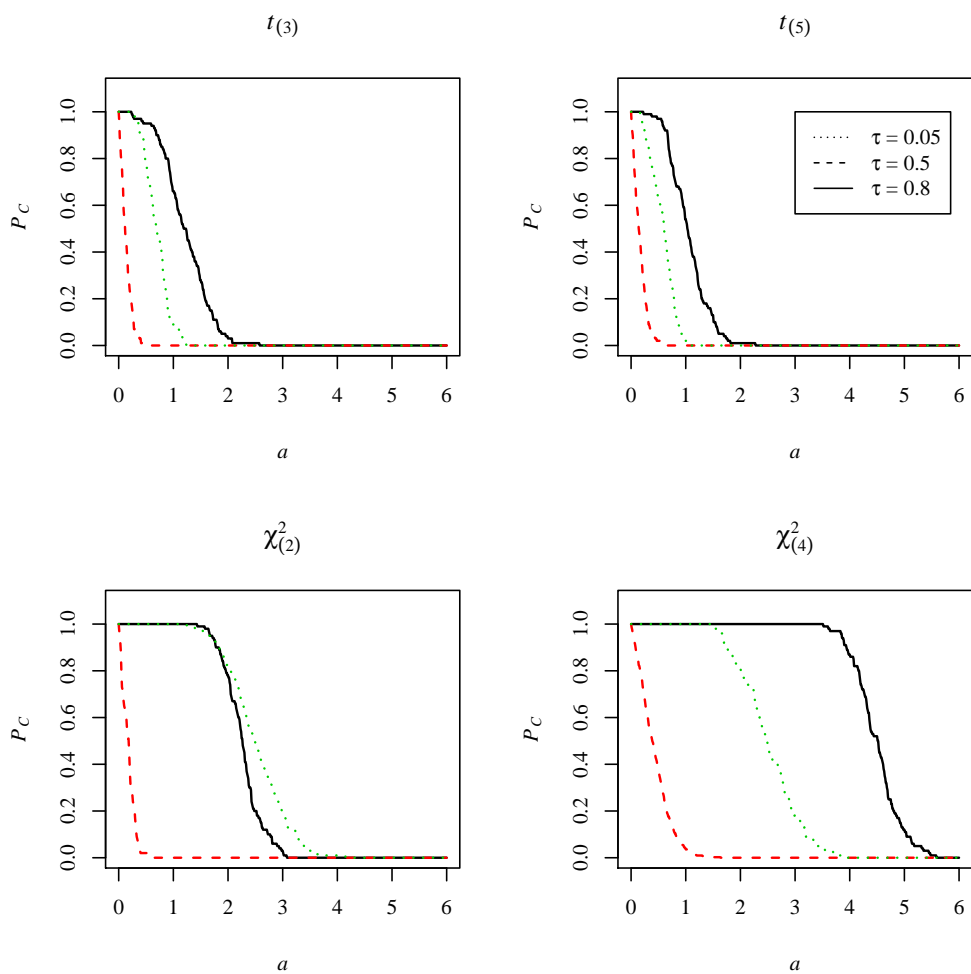


FIGURA 5.3: Probabilidad de sesgos para las distribuciones t-Student y Ji-cuadrado

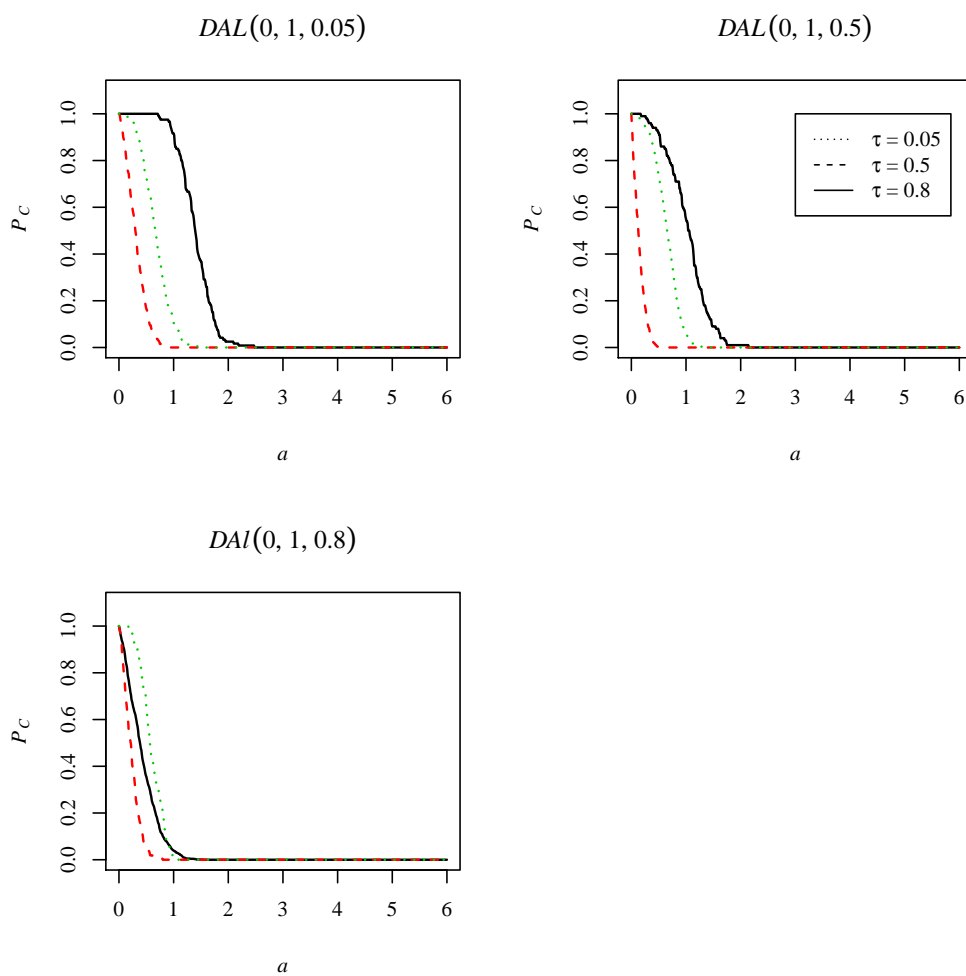


FIGURA 5.4: Probabilidad de sesgos para la distribución asimétrica de Laplace

Capítulo 6

Aplicación

6.1. Análisis sensorial de productos hidroalcohólicos

En el desarrollo de productos hidroalcohólicos, el análisis sensorial permite evaluar diferentes características del producto cuya cuantificación es realizada por seres humanos en lugar de instrumentos. Un aspecto de principal importancia en este tipo de productos es la duración, es decir el tiempo hasta el cual el usuario percibe el aroma del producto en su piel. Para ello, se evalúa la intensidad del producto hidroalcohólico durante un tiempo determinado.

Este tipo de evaluaciones requiere un panel de jueces extensamente entrenados en el uso de una escala de medida continua de la intensidad olfativa, de tal manera que tengan puntos de referencia de intensidad respecto a los cuales evaluar cada nuevo producto. Esto también implica el reentrenamiento permanente del panel de jueces, pues todo el procedimiento de evaluación se basa en la memoria olfativa del juez.

Se realiza una prueba sensorial para comparar el desempeño en intensidad de dos productos. Para ello, se emplearon dos personas como portadores, a las cuales se les aplicó cada producto en una zona delimitada de sus antebrazos, considerando usar el mismo volumen de producto en cada uno de ellos. Luego, los 26 jueces que participaron en el estudio fueron asignados aleatoriamente a cada portador para realizar la evaluación de la intensidad de ambos productos durante toda la prueba, de tal manera que cada portador fue evaluado por 13 jueces. Los tiempos de evaluación fueron 0, 1.5, 3, 4.5, 6 y 8 horas. La figura 6.1 muestra el comportamiento del logaritmo de la intensidad respecto a los tiempos de control

para los dos productos evaluados.

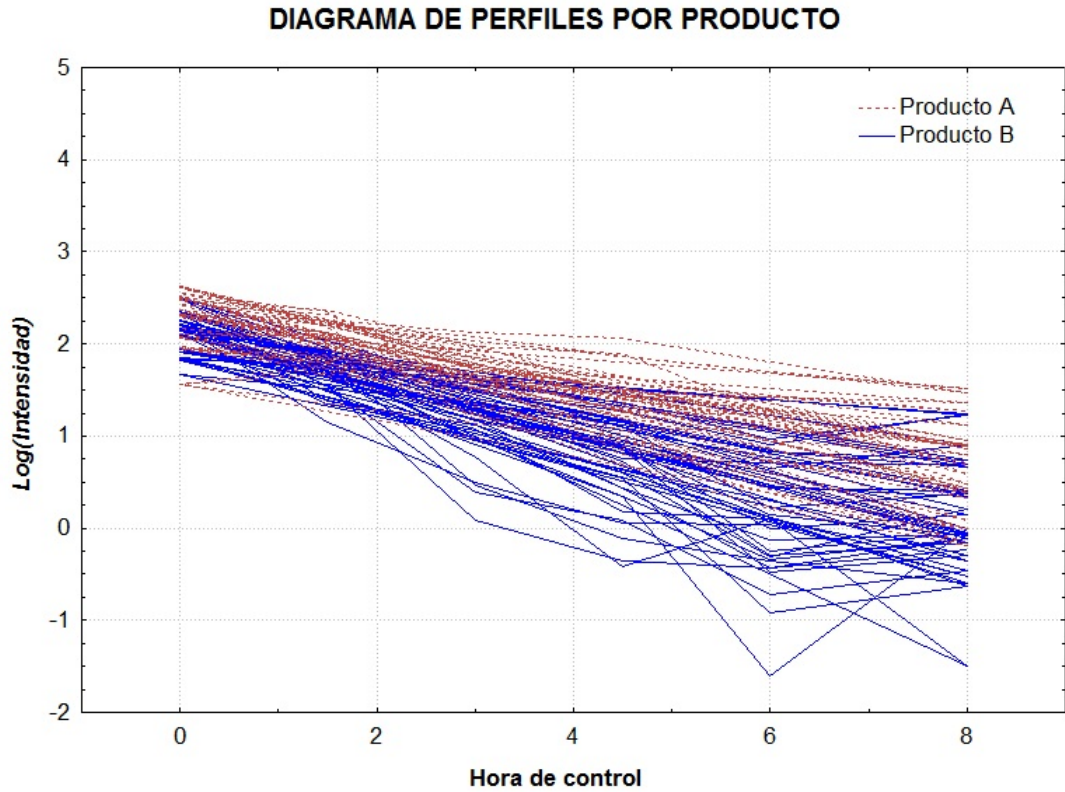


FIGURA 6.1: Diagrama de perfiles

Para la descripción de esta información se emplearon dos opciones: modelos lineales mixtos y regresión cuantílica para datos longitudinales. Para el primer caso se propone el siguiente modelo:

$$y_{ij} = \beta_0 + \beta_1 t_j + \beta_2 X_i + \beta_3 X_i t_{ij} + \alpha_i + \varepsilon_{ij} \quad i = 1, \dots, 26 \quad j = 1, \dots, 6 \quad (6.1)$$

donde y_{ij} denota el logaritmo de la intensidad dada por el i -ésimo juez en el j -ésimo tiempo de control, t_j es el tiempo de control en horas, X_i es una variable indicadora que denota el producto evaluado por el juez i siendo 1 si califica el producto A y 0 en otro caso, α_i es el efecto aleatorio correspondiente al i -ésimo juez, mientras que ε_{ij} es el error aleatorio.

La distribución de la intensidad presenta una ligera asimetría a la izquierda (Coef. de simetría -0.44 para el producto A y -0.32 para el producto B), además la prueba de Shapiro-Wilks permite rechazar el supuesto de normalidad en la variable respuesta para

ambos productos. De igual forma ocurre con los residuos obtenidos luego de ajustar el modelo lineal mixto (6.1) vía máxima verosimilitud restringida (REML).

La segunda clase de modelos a ajustar, nos permite modelar la mediana, que es menos sensible a la asimetría de la distribución y a datos extremos, para lo cual se propone el siguiente modelo de regresión cuantílica:

$$Q_{y_{ij}|\alpha_i}(\tau | t_j, X_i, \alpha_i) = \beta_0 + \beta_1 t_j + \beta_2 X_i + \beta_3 X_i t_{ij} + \alpha_i + \varepsilon_{ij} \quad (6.2)$$

Para el efecto aleatorio α_i se consideró la distribución normal estándar, teniendo en cuenta el poco efecto que su selección puede tener en el modelo según Liu y Bottai (2009), lo cual implica que la distribución conjunta para (y_i, u_i) será normal-Laplace de acuerdo con (3.11). El algoritmo MCEM fue empleado considerando un tamaño de muestra inicialmente constante de 1000 así como 5000 iteraciones de calentamiento (burn-in size) sin embargo, resultados igualmente precisos se lograron obtener tomando un tamaño inicial de 500 y 2000 iteraciones. Como criterio de convergencia se empleó la distancia entre verosimilitudes descrito por Eickhoff et al. (2004).

La tabla 6.1 presenta las estimaciones del modelo lineal mixto y la regresión cuantílica para el caso de la mediana y cuatro cuantiles más. En general la estimación se realizó para 19 valores diferentes (Ver apéndice C.3) del cuantil, desde $\tau = 0.05$ hasta $\tau = 0.95$.

Método	β_0	β_1	β_2	β_3
MLM	2,000(0,0636)	-0,281(0,0114)	0,311(0,0754)	0,069(0,0161)
$\tau = 0.5$	2,127(0,0445)	-0,283(0,0171)	0,213(0,0592)	0,088(0,0215)
$\tau = 0.1$	1,811(0,1166)	-0,404(0,0376)	0,240(0,1460)*	0,136(0,0421)
$\tau = 0.25$	1,957(0,0670)	-0,327(0,0239)	0,221(0,0940)	0,101(0,0321)
$\tau = 0.75$	2,233(0,0482)	-0,252(0,0179)	0,312(0,0811)	0,053(0,0225)
$\tau = 0.9$	2,309(0,0895)	-0,212(0,0256)	0,331(0,1105)	0,063(0,0400)*

TABLA 6.1: Estimación de parámetros y el errores estándar para los efectos fijos

De acuerdo con los modelos ajustados los parámetros β_0 y β_1 resultan significativos en todos los casos indicando una clara tendencia decreciente en el tiempo, así mismo el modelo lineal mixto permite obtener valores muy próximos al caso de la mediana para estos dos parámetros. Por su parte, el producto hidroalcohólico no resulta ser significativamente diferente para todos los cuantiles, pues para $\tau = 0.05, 0.10, 0.95$ (valores con asterisco (*)) en la tabla 6.1) el parámetro β_2 no resulta significativo. De igual manera ocurre para β_3 en $\tau = 0.05, 0.80, 0.85, 0.90, 0.95$ indicando que la interacción no es significativa. El

comportamiento de los coeficientes del modelo se presenta en la figura 6.2. Allí, las líneas en gris denotan las estimaciones del MLM.

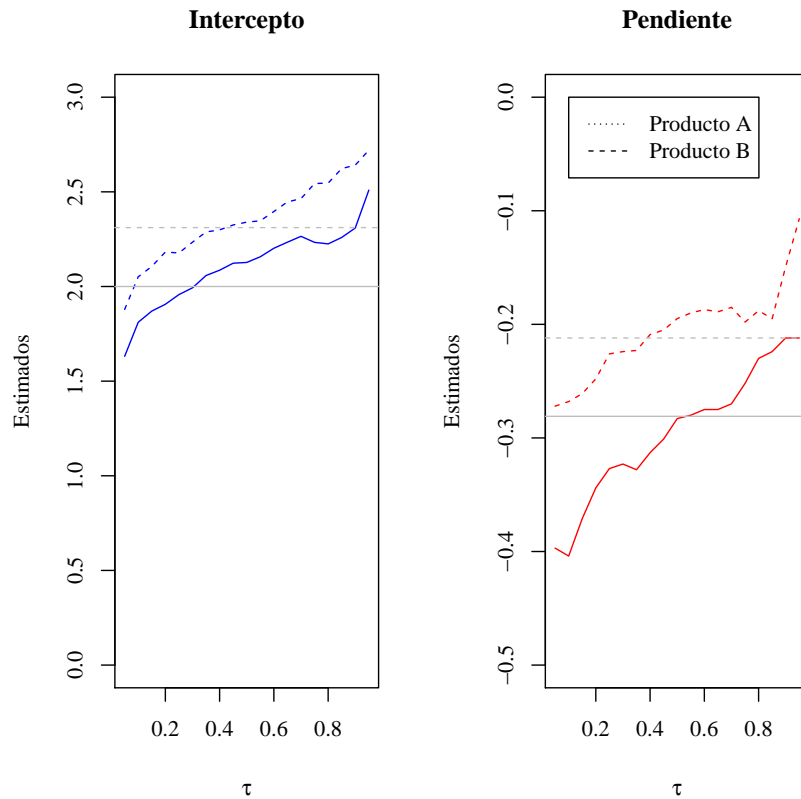


FIGURA 6.2: Variación de los parámetros de acuerdo al cuantil

Las estimaciones del parámetro β_0 y β_1 muestran una tendencia creciente indicando en el caso de β_1 que el cuantil de la respuesta decrece menos en tanto τ va creciendo, como se espera de acuerdo a la figura 6.1. En cuanto a β_2 y β_3 se observa poca variación para todos los valores del cuantil. Adicionalmente, entre el cuantil $\tau = 0.15$ y $\tau = 0.90$ la diferencia entre los productos en evaluación es significativa siendo mayor la intensidad del producto A en 0,247 en la escala logarítmica o 1.282 en la escala original.

En cuanto al ajuste de los diferentes modelos se encuentra que el mejor ajuste se obtiene para $\tau = 0.55$ y $\tau = 0.60$. El modelo lineal mixto obtuvo un valor de 349.5 para el mismo criterio. La figura 6.3 muestra el comportamiento general, el cual aumenta para los cuantiles extremos.

La posibilidad de describir más generalmente la distribución condicional de la respuesta a través de la estimación de sus cuantiles, mientras se explica la dependencia entre las

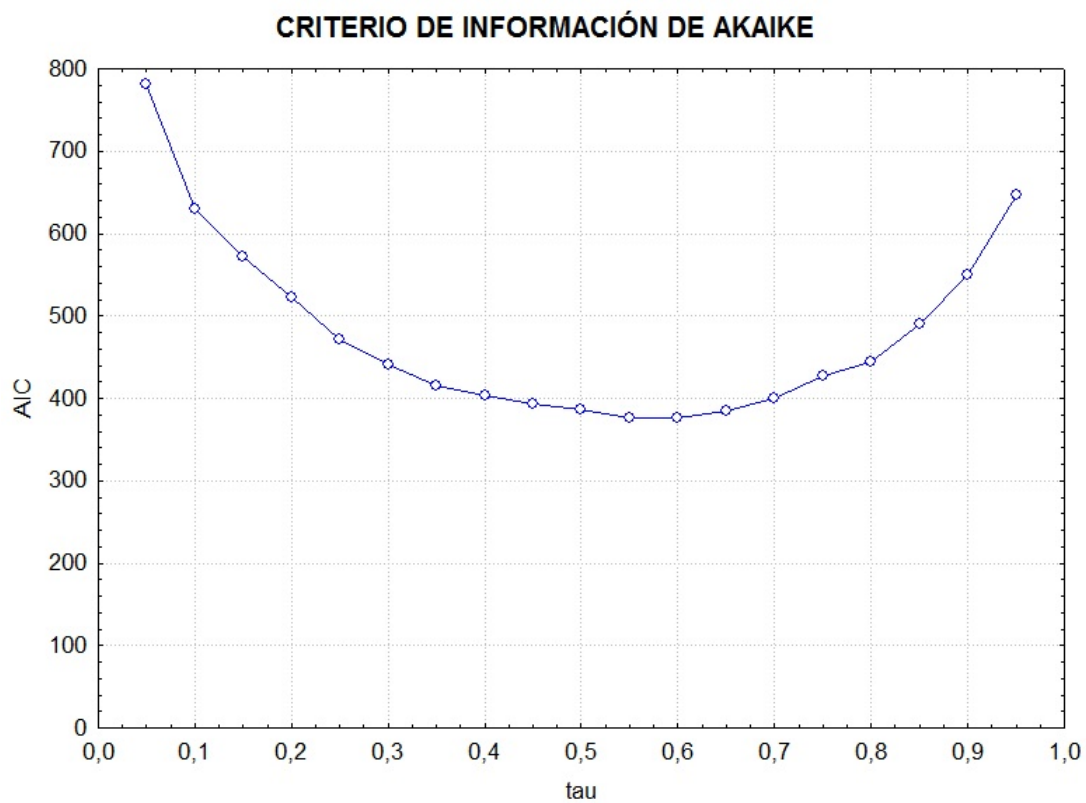


FIGURA 6.3: Variación en el criterio de información de Akaike

observaciones, representa una gran ventaja de los modelos de regresión cuantílica frente a los modelos lineales mixtos.

Capítulo 7

Conclusiones y Recomendaciones

7.1. Conclusiones

La aplicación de la regresión cuantílica al análisis de datos longitudinales es todavía, un tema de investigación en el cual dos métodos de estimación se destacan, mínimos cuadrados (Koenker, 2004) y máxima verosimilitud (Geraci y Bottai, 2007). Sin embargo, este segundo método implica realizar el supuesto que la distribución condicional de la variable respuesta sigue una distribución asimétrica de Laplace, lo cual no resulta ser una característica común en la realidad. Teniendo en cuenta esto y los resultados obtenidos se tienen las siguientes conclusiones:

1. Las estimaciones puntuales y sus varianzas muestran que el cambio en la distribución del término de error no tiene un efecto significativo sobre los efectos fijos del modelo, logrando estimaciones con sesgos del orden de 10^{-2} o menos.
2. La estimación de la varianza de los efectos aleatorios, resulta sesgada en todos los escenarios, sin que el cambio en distribución del término de error pueda ser una causa asignable, pues incluso para la distribución asimétrica de Laplace presenta un valor menor al esperado. En casi todos los casos se identifica subestimación del valor correcto.
3. La medida $B(a, T_n)$ para el desempeño de colas muestra un ligero cambio cuando la distribución del error es la t-student, mostrando valores levemente más bajos. Lo mismo ocurre para la distribución normal pero en menor grado. Esto sugiere que

para distribuciones simétricas si se logra detectar un cambio en la distribución de los estimadores.

4. En general la probabilidad de sesgo tiende rápidamente a cero sin importar la distribución del término de error, sin embargo, resulta importante para otros valores diferentes de $\tau = 0.5$ y cuando la distribución del error es Ji-cuadrado.
5. El ajuste del modelo fue evaluado según el criterio de información de Akaike, siendo mayor que el modelo lineal mixto (MLM) calculado como referencia. Particularmente, para la distribución Ji-cuadrado se obtiene los valores más altos de AIC, sin embargo ocurre lo mismo en el MLM. En general, se encuentra relación no lineal del AIC con el cuantil de estimación, lo cual se hace más evidente en la aplicación desarrollada en el capítulo 6.

7.2. Recomendaciones

El presente estudio se realizó bajo un escenario de homocedasticidad, por lo cual una ampliación consiste en considerar una situación con varianza no constante. Así mismo, la generalización al caso multivariado presentada por [Liu y Bottai \(2009\)](#) es un reciente campo para continuar la evaluación de las propiedades robustas de la estimación por máxima verosimilitud en regresión cuantílica.

Anexo A

Distribución Asimétrica de Laplace

La distribución asimétrica de Laplace (DAL) mencionada por [Koenker y Machado \(1999\)](#) es presentada y analizada por [Yu y Zhang \(2005\)](#), quien cita otras tres formas de la distribución mencionada, debidas a [Johnson et al. \(1995\)](#), [Kotz et al. \(2002\)](#), [Fernández y Steel \(1998\)](#). La versión de la DAL empleada para la estimación en regresión cuantílica tiene la siguiente función de densidad:

$$f(x | \mu, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp \left\{ -\rho_{\tau} \left(\frac{x-\mu}{\sigma} \right) \right\} \quad (\text{A.1})$$

Donde, τ es el parámetro de asimetría ($0 < \tau < 1$), σ es el parámetro de escala ($\sigma > 0$), μ el de localización, ($-\infty < \mu < \infty$) y $\rho_{\tau}(u)$ es la función de pérdida dada en (4.6). Es decir, la variable aleatoria x sigue la distribución asimétrica de Laplace con parámetros (μ, σ, τ) , que en lo siguiente se denotará como $X \sim DAL(\mu, \sigma, \tau)$.

Se obtiene directamente de [A.1](#):

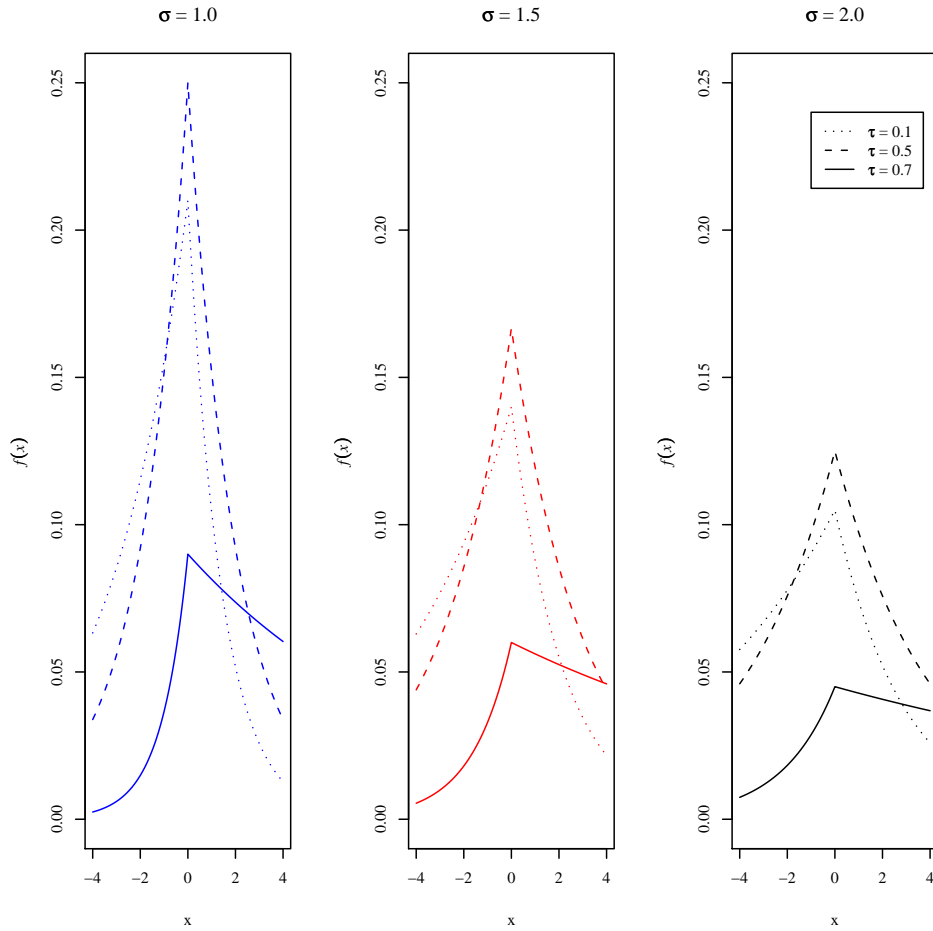


FIGURA A.1: Función de densidad

Función de distribución

$$F(x; \mu, \sigma, \tau) = \begin{cases} \tau \exp \left\{ \frac{1-\tau}{\sigma} (x - \mu) \right\} & \text{si } u \leq \mu \\ 1 - (1 - \tau) \exp \left\{ \frac{\tau}{\sigma} (x - \mu) \right\} & \text{si } u \geq \mu \end{cases} \quad (\text{A.2})$$

Función Inversa

$$F^{-1}(x; \mu, \sigma, \tau) = \begin{cases} \mu + \frac{\sigma}{1 - \tau} \log \left(\frac{x}{\tau} \right) & \text{si } 0 \leq x \leq \tau \\ \mu - \frac{\sigma}{\tau} \log \left(\frac{1 - x}{1 - \tau} \right) & \text{si } \tau < x \leq 1 \end{cases} \quad (\text{A.3})$$

Función Generatriz de momentos

$$\phi(t) = \tau(1 - \tau) \frac{\exp(\mu t)}{(\tau - \sigma t)(\sigma t + 1 - \tau)} \quad (\text{A.4})$$

donde $\frac{\tau-1}{\sigma} < t < \frac{\tau}{\sigma}$

k -ésimo momento central

$$E(X - \mu)^k = k! \sigma^k \tau(1 - \tau) \left(\frac{1}{\tau^{k+1}} + \frac{(-1)^k}{(1 - \tau)^{k+1}} \right) \quad (\text{A.5})$$

Media

$$E(X) = \mu + \frac{\sigma(1 - 2\tau)}{\tau(1 - \tau)} \quad (\text{A.6})$$

Varianza

$$\text{Var}(X) = \frac{\sigma^2(1 - 2\tau + 2\tau^2)}{\tau^2(1 - \tau)^2} \quad (\text{A.7})$$

Coficiente de Asimetría

$$C.\text{Asimetría} = \frac{2(\tau^3 - (1 - \tau)^3)}{((1 - \tau)^2 + \tau^2)^{3/2}} \quad (\text{A.8})$$

Coficiente de Curtosis

$$C.\text{Curtosis} = \frac{9\tau^4 + 6\tau^2(1 - \tau)^2 + 9(1 - \tau)^4}{(1 - 2\tau + 2\tau^2)^2} \quad (\text{A.9})$$

Nótese que estos dos últimos valores no dependen del parámetro de escala σ .

A.1. Estimadores de los parámetros

Los estimadores de máxima verosimilitud de los parámetros $(\hat{\mu}, \hat{\sigma}, \hat{\tau})$ están dados por la solución de las siguientes ecuaciones:

$$\hat{\mu} = \min_{\mu} \sum_{i=1}^n \rho_{\hat{\tau}}(x_i - \mu) \quad (\text{A.10})$$

$$\hat{\sigma} = \frac{1}{n} \sum_{i=1}^n \rho_{\hat{\tau}}(x_i - \mu) \quad (\text{A.11})$$

$$\hat{\tau} = \frac{a + \sqrt{a_2 - (\bar{x} - \hat{\mu})a}}{\bar{x} - \hat{\mu}} \quad (\text{A.12})$$

donde $a = \frac{1}{n} \sum_{i=1}^n (x_i - \mu) I(x_i \leq \hat{\mu})$. Un algoritmo iterativo es requerido para la solución de estos estimadores.

Anexo B

Código R

Programa para la generación de números aleatorios con distribución asimétrica de Laplace

```
rdal<-function(n,mu,sigma,tau){
  U<-runif(n)
  X<-numeric(n)
  for(i in 1:n){
    if(U[i]>tau){X[i]<-mu-(sigma/tau)*log((1-U[i])/(1-tau))}
    else{X[i]<-mu+(sigma/(1-tau))*log(U[i]/tau)}
  }
  return(X)
}
```

Programa para la generación de los efectos aleatorios y la corrección de la variable respuesta

```
GEA<-function(DT,M,tau,tsim,burn_in){
  DT<-as.matrix(DT)
  sigma<-DT[1]
  phi<-DT[2]
  b1<-DT[3]
  b0<-DT[4]
  U0<-rnorm(1,0,1)
  XX<-cbind(1,matrix(seq(1:M),ncol=1))
  YY<-DT[5:(M+4)]
```

```

# Simulación de efectos aleatorios
  Lista_Pr<-list(XX=XX,YY=YY,sigma=sigma,phi=phi,b1=b1,b0=b0,M=M,tau=tau)
  U0limite<-function(U0,Lista) {(min(U0)>-20 )*(max(U0)<20)}
  U<-arms(U0,U0sim,U0limite,burn_in+tsim,Lista=Lista_Pr)
# Guarda los efectos aleatorios posteriores al calentamiento
  u2<-as.matrix(U[(burn_in+1):(burn_in+tsim)])
# Se corrige la variable original con los efectos aleatorios simulados
  y2 <- YY%*%matrix(rep(1,tsim),nrow=1) - matrix(rep(1,M),ncol=1)%*%t(u2)
resultado<-list(t(u2),y2)
return(resultado)
}

```

Función conjunta condicional $f(y_i, \alpha_i | \eta)$

```

# Función de perdida para la distribución de Laplace #
laplace<-function(w,tau){
  if (tau <= 1 & tau >= 0) {
    w*ifelse(w<0,(tau-1),tau)}
  else {return("El valor del cuantil debe estar entre 0 y 1")}
# Distribucion condicional del efecto aleatorio
  U0sim<-function(U0,Lista) {
  b0 <- Lista$b0
  b1 <- Lista$b1
  YY <- Lista$YY
  XX <- Lista$XX
  sigma <- Lista$sigma
  phi <-Lista$phi
  tau <-Lista$tau
  M <-Lista$M
  beta0<-matrix(c(b0,b1),ncol=1)
  MUij<-(XX%*%beta0-matrix(rep(1,M),nrow=M)*U0)
  -sum(laplace((YY - MUij)/sigma, tau))-(1/(2*phi))*U0^2
}

```

Programa para la generación del modelo


```

Modelo<-function(N=50,M=15,dis.err = 'sn',var.err=1,
                 var.rand=1,n=1,tau=0.05,prm=20){

# N: Número de individuos
# M: Número de observaciones por individuo
# n: Número de bases a generar en la simulación
# p: Proporción de contaminación del termino de error
# dis.err : Distribución contaminante para el termino de error,
#           seleccionar una de normal, t3, chi2
# var.err: Varianza para el termino de error
# var.rand: Varianza par los efectos aleatorios

datos<-array(0,c(N*M,2,n))

for (i in 1:n){
  x<-matrix(rep(seq(1:M),N),ncol=1)
  Z <- kronecker(diag(N),rep(1,M))

  #if (dis.err=="normal") Eps <-rnorm(N*M,mean=0,sd=sqrt(var.err))
  if (dis.err=="sn") {
    require(sn)
    # Media de la normal sesgada
    lambda<-prm
    delta<-lambda/sqrt(1+lambda^2)
    b<-sqrt(2/pi)
    md<-var.err*b*delta
    vr<-var.err*(1-(2*delta^2/pi))
    Eps <-(rsn(N*M,0,sqrt(var.err),lambda)-md)*sqrt(var.err/vr)
  }
  if (dis.err=="chisq2") {
    gl<-prm
    Eps <-(rchisq(N*M,gl)-gl)*sqrt(var.err/2*prm)}

  if (dis.err=="dal") { md<-sqrt(var.err)*(1-2*tau)/(tau*(1-tau))
    vr<-(var.err^2)*(1-2*tau+2*tau^2)/((tau^2)*(1-tau)^2)
  }
}
}

```

```

        Eps <- (rdal(N*M,0,var.err,tau)-md)*sqrt(var.err/vr)}
if (dis.err=="ts") {
    gl<-prm
    Eps <-rt(N*M,gl)*sqrt(var.err/(prm/(prm-2)))}

eps<-t(VEC_INV(N,M,Eps))
MCOR<-matrix(0,ncol=M,nrow=M)
t<-seq(1:M)
phi<-0.8
for(f in 1:M){
  for(c in 1:M){
    MCOR[f,c]<- var.err*exp(-phi*abs(t[f]-t[c]))
  }
}
ERR<-eps%%t(chol(MCOR))
EPS<-VECF(t(ERR))
U <-matrix(rnorm(N,0,sqrt(var.rand)))

y<-cbind(1,x) %% matrix(c(1, 2),ncol=1) + Z %% U + EPS
datos[, ,i]<-cbind(y=y,x=x)

}
return(datos)
}

```

Programa para la aplicación del MCEM

```

MCEM<-function(data=data1,M=10,N=20,tau=0.5,e=0.1,delta=10){

# data:   Datos del modelo a estimar
# N:     Número de individuos
# M:     Número de observaciones por individuo
# delta: Distancia máxima para convergencia
# tsim:  Tamaño de la simulación
# burn_in:Tamaño de calentamiento para el arms

```

```

# Lectura de datos
Y <- matrix(data[,1])
X <- matrix(data[,2])
IB<-25
c<-0 # Llave del while
k<-0 # Contador de iteraciones
MY<-VEC_INV(N,M,Y)
Log_VR<-rep(0,50)
while(c==0){
  k<-k+1

  if(k<IB){tsim<-100; burn_in<-500}
### ----- PASO E ----- ###

# Simulación del efecto aleatorio y corrección de la variable
  respuesta para cada individuo

  if(k==1){
    #b1<-10
    #b0<-10
    b1<-rnorm(1,2,1)
    b0<-rnorm(1,1,1)
    sigma <-rnorm(1,0,1)
    phi <- rnorm(1,0,1)}
  else{
    b1<-mubeta[2]
    b0<-mubeta[1]
  }

  # Simulación de efectos aleatorios
MT<-rbind((matrix(c(sigma,phi,b1,b0),nrow=4)%*%matrix(1,ncol=N)),MY)
y4 <- apply(MT,2,GEA,M,tau,tsim,burn_in)
for(i in 1:N){
  if(i==1){

```

```

        y2<-y4[[i]][[2]]
        u3<-y4[[i]][[1]]}
    else{
        y2<-rbind(y2,y4[[i]][[2]])
        u3<-rbind(u3,y4[[i]][[1]])}
    }

### ----- PASO M ----- ####

# Valores medios de la respuesta corregida por el efecto
# aleatorio simulado
y3<-matrix(rowMeans(y2),ncol=1)
XV<-VECF(X)

# Maximización de los efectos fijos a través de regresión cuantílica
mrq<-rq(y3~XV,tau=tau)
srq<-summary(mrq)
aic<-extractAIC(mrq)[2]
# Vector de coeficientes estimados b0, b1
mubeta<-c(srq[[3]][1],srq[[3]][2],srq[[3]][3],srq[[3]][4],
          srq[[3]][5],srq[[3]][6])

# Estimación de Sigma
S<-apply(laplace((y2-(cbind(1,XV)%*%mubeta[1:2]))%*%
               matrix(1,ncol=tsim)),tau),2,sum)
sigma<-sum(S)/(dim(y2)[1]*dim(y2)[2])

phi<-mean(apply(u3,2,var)) # Estimación de phi
if(k==1){parametros<-matrix(c(k,mubeta[1],mubeta[2],mubeta[3],
                             mubeta[4],mubeta[5],mubeta[6],sigma,phi,aic),nrow=1)}
else{parametros<-rbind(parametros,matrix(c(k,mubeta[1],mubeta[2],
                             mubeta[3],mubeta[4],mubeta[5],mubeta[6],sigma,phi,aic),nrow=1))}

#---- CALCULO DEL CRITERIO DE CONVERGENCIA ----#

if(k>=IB){

```

```

#cat("Iteración: ",k, "Parametros", parametros[k,], sep=' ', '\n' )
#cat("Inicio convergencia ", sep=' ', '\n' )
vhat<-rep(0,N)
log_VR<-rep(0,N)
A<-kronecker(diag(1, N),matrix(rep(1,M),ncol=M))
B<-(A%*%laplace(y2,tau))/sigma
C<-((tau*(1-tau)/sigma)^M)*exp(-B)
Vhat<-(1/(tsim-1))*sum((apply(C^2,1,sum)/(apply(C,1,mean))^2)-rep(tsim,N))
Log_VR[k]<-sum(log(apply(C,1,mean)))

Tsim<- ceiling(Vhat/(e*(sqrt(N)*delta/3)^2))
if(is.nan(Tsim)){Tsim<-200}
else{if(Tsim>3000){Tsim<-3000}
else{Tsim<- ceiling(Vhat/(e*(sqrt(N)*delta/3)^2))}}

cat("Tamaño calculado =",Tsim,"Distancia Inicial = ",
    abs(Log_VR[k-1]-Log_VR[k]),"Criterio = ",sqrt(N)*delta/3, '\n')
if (Tsim <= tsim )
  {if (abs(Log_VR[k-1]-Log_VR[k]) < (sqrt(N)*delta/3)) {c<-1} else {c<-0;}}
  else {

    DIF<-Tsim-tsim
### ----- PASO E ----- ###

# Simulación del efecto aleatorio y corrección de la variable respuesta
para cada individuo

    MT<-rbind((matrix(c(sigma,phi,b1,b0),nrow=4)%*%matrix(1,ncol=N)),MY)
    b1<-mubeta[2]
    b0<-mubeta[1]
    U0<-rnorm(1,0,1)
    #U0<-u3[i*tsim,1]
    Y4 <- apply(MT,2,GEA,M,tau,DIF,burn_in)
    for(i in 1:N){
      if(i==1){
        y5<-Y4[[i]][[2]]

```

```

        u5<-Y4[[i]][[1]]}
    else{
        y5<-rbind(y5,Y4[[i]][[2]])
        u5<-rbind(u5,Y4[[i]][[1]])}
    }

    # Se corrige la variable original con los efectos aleatorios simulados

    Y2 <- cbind(y2,y5)
    U3 <- cbind(u3,u5)
### ----- PASO M ----- ###

# Valores medios de la respuesta corregida por el efecto aleatorio simulado
Y3<-matrix(rowMeans(Y2),ncol=1)

# Maximización de los efectos fijos a través de regresión cuantílica
mrq<-rq(Y3~XV,tau=tau)
srq<-summary(mrq)
aic<-extractAIC(mrq)[2]
# Vector de coeficientes estimados b0, b1
mubeta<-c(srq[[3]][1],srq[[3]][2],srq[[3]][3],srq[[3]][4],
srq[[3]][5],srq[[3]][6])

# Estimación de Sigma
S<-apply(laplace((Y2-(cbind(1,XV)%*%mubeta[1:2])%*%matrix(1,ncol=Tsim)),
tau),2,sum)
sigma<-sum(S)/(dim(Y2)[1]*dim(Y2)[2])

phi<-mean(apply(U3,2,var)) # Estimación de phi
parametros[k,]<-c(k,mubeta[1],mubeta[2],mubeta[3],mubeta[4],
mubeta[5],mubeta[6],sigma,phi,aic)

A<-kronecker(diag(1, N),matrix(rep(1,M),ncol=M))
B<-(A%*%laplace(Y2,tau))/sigma
C<-((tau*(1-tau)/sigma)^M)*exp(-B)
Log_VR[k]<-sum(log(apply(C,1,mean)))

```

```
if (abs(Log_VR[k-1]-Log_VR[k]) < (sqrt(N)*delta/3)) {c<-1}
else {c<-0; tsim<-Tsim}
  }

      } # Fin del condicional k>=20
} # Fin del While

return(t(parametros[k,]))
}
```

Anexo C

Resultados aplicación

τ	β_0	β_1	β_2	β_3	σ	φ	AIC
0,05	1,63	-0,3967	0,2481	0,1247	0,0644	0,0915	781,88
0,10	1,81	-0,4036	0,2397	0,1361	0,0929	0,0438	629,47
0,15	1,87	-0,3709	0,2361	0,1098	0,1192	0,0351	572,70
0,20	1,91	-0,3438	0,2758	0,0955	0,1385	0,0276	523,21
0,25	1,96	-0,3270	0,2209	0,1014	0,1489	0,0189	471,79
0,30	1,99	-0,3227	0,2412	0,0985	0,1581	0,0145	441,73
0,35	2,06	-0,3279	0,2300	0,1051	0,1639	0,0117	415,58
0,40	2,09	-0,3128	0,2132	0,1033	0,1692	0,0114	403,64
0,45	2,12	-0,3006	0,2027	0,0953	0,1721	0,0120	393,45
0,50	2,13	-0,2835	0,2129	0,0882	0,1725	0,0134	386,17
0,55	2,16	-0,2796	0,1895	0,0895	0,1686	0,0135	375,61
0,60	2,20	-0,2745	0,1939	0,0871	0,1647	0,0144	377,02
0,65	2,23	-0,2748	0,2136	0,0862	0,1580	0,0156	384,86
0,70	2,26	-0,2702	0,1986	0,0853	0,1488	0,0166	399,49
0,75	2,23	-0,2519	0,3116	0,0535	0,1378	0,0190	426,98
0,80	2,23	-0,2295	0,3208	0,0412	0,1212	0,0197	445,31
0,85	2,26	-0,2238	0,3643	0,0286	0,1045	0,0251	491,09
0,90	2,31	-0,2118	0,3315	0,0626	0,0807	0,0282	550,40
0,95	2,51	-0,2125	0,2108	0,1058	0,0508	0,0520	647,24

TABLA C.1: Estimación de parámetros para los efectos fijos

τ	Parámetro	Estimado	Error Std	Valor P
0,05	β_0	1,631	0,1396	0,000
0,05	β_1	-0,397	0,0518	0,000
0,05	β_2	0,248	0,2753	0,368
0,05	β_3	0,125	0,0681	0,068
0,10	β_0	1,811	0,1166	0,000
0,10	β_1	-0,404	0,0376	0,000
0,10	β_2	0,240	0,1460	0,102
0,10	β_3	0,136	0,0421	0,001
0,15	β_0	1,870	0,0695	0,000
0,15	β_1	-0,371	0,0339	0,000
0,15	β_2	0,236	0,0894	0,009
0,15	β_3	0,110	0,0395	0,006
0,20	β_0	1,906	0,0486	0,000
0,20	β_1	-0,344	0,0234	0,000
0,20	β_2	0,276	0,0891	0,002
0,20	β_3	0,095	0,0327	0,004
0,25	β_0	1,957	0,0670	0,000
0,25	β_1	-0,327	0,0239	0,000
0,25	β_2	0,221	0,0940	0,019
0,25	β_3	0,101	0,0321	0,002
0,30	β_0	1,993	0,0740	0,000
0,30	β_1	-0,323	0,0252	0,000
0,30	β_2	0,241	0,0934	0,010
0,30	β_3	0,098	0,0316	0,002
0,35	β_0	2,058	0,0692	0,000
0,35	β_1	-0,328	0,0208	0,000
0,35	β_2	0,230	0,0860	0,008
0,35	β_3	0,105	0,0265	0,000
0,40	β_0	2,086	0,0706	0,000
0,40	β_1	-0,313	0,0219	0,000
0,40	β_2	0,213	0,0835	0,011
0,40	β_3	0,103	0,0261	0,000
0,45	β_0	2,123	0,0517	0,000
0,45	β_1	-0,301	0,0196	0,000
0,45	β_2	0,203	0,0675	0,003
0,45	β_3	0,095	0,0240	0,000

TABLA C.2: Estimación de parámetros, errores estándar y valores P. $\tau < 0.5$

τ	Parámetro	Estimado	Error Std	Valor P
0,50	β_0	2,127	0,0445	0,000
0,50	β_1	-0,283	0,0171	0,000
0,50	β_2	0,213	0,0592	0,000
0,50	β_3	0,088	0,0215	0,000
0,55	β_0	2,157	0,0480	0,000
0,55	β_1	-0,280	0,0155	0,000
0,55	β_2	0,190	0,0605	0,002
0,55	β_3	0,090	0,0198	0,000
0,60	β_0	2,202	0,0348	0,000
0,60	β_1	-0,275	0,0160	0,000
0,60	β_2	0,194	0,0677	0,004
0,60	β_3	0,087	0,0200	0,000
0,65	β_0	2,234	0,0324	0,000
0,65	β_1	-0,275	0,0154	0,000
0,65	β_2	0,214	0,0725	0,003
0,65	β_3	0,086	0,0207	0,000
0,70	β_0	2,265	0,0387	0,000
0,70	β_1	-0,270	0,0172	0,000
0,70	β_2	0,199	0,0749	0,008
0,70	β_3	0,085	0,0215	0,000
0,75	β_0	2,233	0,0482	0,000
0,75	β_1	-0,252	0,0179	0,000
0,75	β_2	0,312	0,0811	0,000
0,75	β_3	0,053	0,0225	0,018
0,80	β_0	2,225	0,0369	0,000
0,80	β_1	-0,230	0,0167	0,000
0,80	β_2	0,321	0,0767	0,000
0,80	β_3	0,041	0,0250	0,101
0,85	β_0	2,259	0,0454	0,000
0,85	β_1	-0,224	0,0190	0,000
0,85	β_2	0,364	0,0734	0,000
0,85	β_3	0,029	0,0341	0,403
0,90	β_0	2,309	0,0895	0,000
0,90	β_1	-0,212	0,0256	0,000
0,90	β_2	0,331	0,1105	0,003
0,90	β_3	0,063	0,0400	0,118
0,95	β_0	2,509	0,1561	0,000
0,95	β_1	-0,212	0,0488	0,000
0,95	β_2	0,211	0,1768	0,234
0,95	β_3	0,106	0,0558	0,059

TABLA C.3: Estimación de parámetros, errores estándar y valores P. $\tau \geq 0.5$

Bibliografía

- Canay, I. A. (2008). A note on quantile regression for panel data models. Northwestern University.
- Cole, T. y Green, P. (1992). Smoothing reference centile curves: The lms method and penalized likelihood. *Statistics in Medicine*, 11:1305–1319.
- Diggle, P. J., Heagerty, P. J., Liang, K., y Zeger, S. (2002). *Analysis of longitudinal data*. Oxford University Press, Londres, 2nd edition.
- Eickhoff, J. C., Zhu, J., y Amemiya, Y. (2004). On the simulation size and the convergence of the monte carlo em algorithm via likelihood-based distances. *Statistics & Probability Letters*, 67(2):161–171.
- Fernández, C. y Steel, M. F. J. (1998). On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93(441):359–371.
- Geraci, M. y Bottai, M. (2007). Quantile regression for longitudinal data using the asymmetric laplace distribution. *Biostatistics*, 8:140–154.
- Gilks, W. R., Best, N. G., y Tan, K. K. C. (1995). Adaptive rejection metropolis sampling within gibbs sampling. *Applied Statistics*, 44(4):455–472.
- Gilks, W. R. y Wild, P. (1992). Adaptive rejection sampling for gibbs sampling. *Applied Statistics*, 41(2):337–348.
- Hao, L. y Naiman, D. Q. (2007). *Quantile regression*, volume 7 of *Quantitative Applications in the Social Sciences*. Sage Publications, Inc.
- He, X., Jurečková, J., Koenker, R., y Portnoy, S. (1990). Tail behavior of regression estimators and their breakdown points. *Econometrica*, 58(5):1195–1214.
- Huber, P. J. y Ronchetti, E. (2009). *Robust Statistics*. John Wiley & Sons, segunda edición.

- Johnson, N., Kotz, S., y Balakrishnan, N. (1995). *Distributions in Statistics, Volumen 2: Continuous Univariate Distributions*, volume 2. John Wiley, New York, 2nd edition.
- Jurečková, J. (1981). Tail-behavior of location estimators. *The Annals of Statistics*, 9(3):578–585.
- Jurečková, J. y Picek, J. (2009). *Robust statistical methods with R*. Chapman Hall/CRC.
- Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91:74–89.
- Koenker, R. (2005). *Quantile Regression*. University Press, New York.
- Koenker, R. y Bassett, G. (1978). Regression quantiles. *Econometrica*, 46:33–50.
- Koenker, R. y D'Orey, V. (1987). Algorithm as 229: Computing regression quantiles algorithm as 229: Computing regression quantiles. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 36(3):383–393.
- Koenker, R. y Machado, J. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, 99(448):1296–1310.
- Koenker, R. y Park, B. (1996). An interior point algorithm for nonlinear quantile regression. *Journal of Econometrics*, 71:265–283.
- Kotz, S., Kozabowski, T., y Podgórski, K. (2002). Maximum likelihood estimation of asymmetric laplace parameters. *Annals of the Institute of Statistical Mathematics*, 54(2):816–826.
- Lamarche, C. (2006). Robust penalized quantile regression estimation for panel data. Manuscrito.
- Liu, Y. y Bottai, M. (2009). Mixed-effects models for conditional quantiles with longitudinal data. *The International Journal of Biostatistics*, 5(1):1–22.
- Portnoy, S. y Koenker, R. (1997). The gaussian hare and the laplacian tortoise: computability of squared-error versus absolute-error estimators (with discussion). *Statistical Science*, 12(4):279–296.
- Verbeke, G. y Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer-Verlag, New York.

-
- Yu, K. y Jones, M. C. (1998). Local linear quantile regression. *Journal of the American Statistical Association*, 93(441):228–237.
- Yu, K., Lu, Z., y Stander, J. (2003). Quantile regression: applications and current research areas. *The Statistician*, 52(3):331–350.
- Yu, K. y Zhang, J. (2005). A three-parameter asymmetrical laplace distribution and its extension. *Communications in Statistics-Theory and Methods*, 34:1867–1879.