



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DE CHILE

ANÁLISIS BAYESIANO SEMIPARAMÉTRICO DEL PROBLEMA DE CALIBRACIÓN EN MODELOS DE REGRESIÓN ELÍPTICO

Maria Paz Casanova Laudien

Profesores Guía: Pilar Iglesias
 Víctor Salinas

SOMETIDO A LOS REQUERIMIENTOS PARA OPTAR EL GRADO DE DOCTOR

EN EL

**PROGRAMA DE DOCTORADO EN ESTADÍSTICA
PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
SANTIAGO DE CHILE
JULIO DE 2005**

A mis hijos,

Paulina, Raquel, Gabriel y Priscila.

INDICE

<i>Dedicatoria</i>	<i>i</i>
<i>Tabla de contenidos</i>	<i>ii</i>
<i>Agradecimientos</i>	<i>iii</i>
<i>Resumen</i>	<i>v</i>
<i>Capitulo 1: Introducción</i>	<i>1</i>
<i>Capitulo 2: Modelo de Regresión Lineal bajo Distribuciones Elípticas Representables</i>	<i>12</i>
<i>Capitulo 3: Solución Semiparamétrica al Problema de Calibración Lineal Continua bajo distrib. Elípticas Representables</i>	<i>24</i>
<i>Capitulo 4: Solución Semiparamétrica al Problema de Calibración Lineal Continua bajo distrib. Elípticas Representables</i>	<i>62</i>
<i>Capitulo 5: Solución Semiparamétrica del Modelo con Errores en las Variables</i>	<i>90</i>
<i>Capitulo 6: Análisis Semiparamétrico del Modelo de Calibración Comparativa</i>	<i>123</i>
<i>Anexo : Glosario de Notaciones</i>	<i>150</i>
<i>Conclusiones</i>	<i>151</i>
<i>Proyecciones Futuras</i>	<i>153</i>
<i>Referencias</i>	<i>154</i>

AGRADECIMIENTOS

Quisiera agradecer, en primer lugar, a Dios por darme la oportunidad de crecer en un medio donde el estudio y la responsabilidad eran altamente valorados y por regalarme la conciencia de su presencia y amor incondicional y constante en mi vida, tanto en los malos como en los buenos momentos, tanto frente a los difíciles desafíos, como frente a los triunfos.

No puedo menos que agradecer a mis padres su ejemplo de tenacidad y su respaldo a lo largo de mi vida; en especial a mi madre su devoción y cuidados, mas allá de lo que esperaba y a mi padre su presencia silenciosa pero significativa de los últimos años.

Quisiera aprovechar esta oportunidad para agradecer a mis hijos por su comprensión y paciencia y por su apoyo en el logro de esta meta.

También quisiera agradecer a mis queridos amigos y amigas de todas partes: de la quinta región, amigos de mis años más difíciles; de Santiago, amigos y compañeros de mis años de estudio; de Concepción, amigos de mi nueva vida. Todos ellos me ayudaron a lo largo de estos años para poder llegar a este momento.

Esta ayuda me llegó muchas veces de forma práctica: con mis hijos, con mis estudios, económicamente, con alojamiento; otras veces su ayuda me llegó de forma espiritual: con palabras de ánimo, con consejos, con reprensiones, con oraciones, con un abrazo, con paciencia y sobre todo, con su cariño.

No puedo olvidar a toda la comunidad no docente de la Pontificia Universidad Católica, secretarias, auxiliares, portero, etc., quienes cada día realizan su valioso trabajo y junto a ello brindan una sonrisa que motiva a seguir adelante.

También agradezco a todos mis profesores, quienes creyeron en mí y transfirieron su saber y también su forma de vida. Gracias por hacerme sentir parte de ustedes.

En especial quiero agradecer a mis profesores guía, Pilar iglesias y Víctor Salinas.

A Pilar porque me levantó cuando ya no tenía fuerzas, luchó por mí, me dio ánimo y sembró en mí su maravilloso ejemplo de vida. Pilita, maestra y amiga, estás en mis oraciones y ojalá Dios nos conceda muchos años más de tu presencia en medio nuestro, formando a muchos jóvenes como profesionales, investigadores y hombres de bien.

A Víctor por su apoyo y preocupación constante y por su amistad.

También quiero agradecer a Alexis Peña por su valioso apoyo en el trabajo computacional y por su disposición y amistad gratuita.

Finalmente quiero expresar mi eterna gratitud al doctor Mario Mellado Sánchez, por su apoyo técnico, logístico y emocional en la etapa final de mi trabajo, y sin cuya amistad y desinteresada ayuda, más allá de sus fuerzas, no hubiese sido capaz de lograr esta meta.

RESUMEN

El presente trabajo reúne cuatro extensiones del modelo de regresión y los analiza en forma bayesiana y semiparamétrica.

Los modelos aquí abordados corresponden a: Calibración Lineal Continua, Calibración Lineal Discreta, Modelo con Errores de Medición y Calibración Comparativa. Estos modelos son abordados en los capítulos 3, 4, 5 y 6, respectivamente.

El modelo de Calibración Lineal Discreta considera una variable binaria o una variable discreta para la respuesta, a la cual subyace una variable elíptica. Los otros tres modelos consideran errores elípticos representables, con el fin de generalizar el modelo normal, abordándose los casos independiente y dependiente.

El artículo de Arellano – Valle, Galea, e Iglesias (1999) permite relacionar las distribuciones elípticas de los errores con los modelos normales, por medio de la inclusión de variables subyacentes que, en este caso, siguen un proceso de Dirichlet. Aquí se contemplan tanto una priori discreta, como una priori continua para el proceso. Estas herramientas se utilizan para los cuatro modelos.

Por otra parte, el artículo de Basu y Mukhopadhyay (2000) permite vincular una distribución binaria con un modelo normal, lo cual junto a las propiedades de la distribución elíptica y del proceso Dirichlet, dan forma al análisis utilizado en el capítulo 4. Finalmente, el artículo de Escobar y West (1998) permite la implementación computacional de los resultados teóricos.

Este trabajo incluye dos capítulos previos introductorios. El primero entrega el marco conceptual general y el segundo se aboca a describir cada uno de los modelos a analizar en capítulos posteriores, así como las principales proposiciones que se utilizarán para la obtención de los resultados.

Los capítulos 3 a 6 presentan el análisis de cada modelo en particular, incluyendo, para cada caso, ejemplos y simulaciones, cuyos datos son analizados con la metodología propuesta en forma computacional.

Finalmente, y antes de las conclusiones, se presentan un anexo con el glosario de notaciones.

CAPÍTULO 1:

INTRODUCCIÓN

El objetivo de este capítulo es presentar el marco general, definiciones y propiedades necesarias para el desarrollo del trabajo. En la primera sección se discuten las diferencias fundamentales entre inferencia paramétrica, semiparamétrica y no paramétrica, como también aspectos de implementación, en particular los modelos abordados por Escobar y West (1998). Se incluyen definiciones y teoremas fundamentales para la comprensión del trabajo. En la segunda sección se presenta el algoritmo propuesto por Escobar y West (1998), en el cual se basa la implementación computacional de los problemas tratados en el presente trabajo. Los modelos lineales elípticos son introducidos en la tercera sección, como una extensión del modelo lineal normal usual. Finalmente la cuarta sección es dedicada al plan general de esta tesis, incluyendo los objetivos principales.

1.1 INFERENCIA BAYESIANA SEMIPARAMÉTRICA

En esta sección se discuten algunos aspectos de la inferencia Bayesiana no paramétrica y semi paramétrica para modelos jerárquicos, con énfasis en los modelos abordados por Escobar y West (1998) y el proceso de Dirichlet introducido por Ferguson (1973). Muller y Quintana (2004) presentan una extensa revisión bibliográfica del tema.

Un modelo estadístico es un trío $(X, \mathcal{A}, \mathbb{P})$, donde X es el espacio de posibles observaciones, \mathcal{A} es la σ - álgebra para X y \mathbb{P} es una familia de medidas de probabilidad definidas sobre (X, \mathcal{A}) . El objetivo del análisis estadístico es la inferencia, que intuitivamente se puede definir como el proceso de adivinar o estimar qué $\mathbb{P} \in \mathcal{P}$ genera los datos $x \in X$ o bien, verificar o refutar alguna hipótesis acerca de la verdadera medida \mathbb{P} .

Dependiendo de la naturaleza de la familia de medidas de probabilidad \mathcal{P} , es posible distinguir tres enfoques en la inferencia estadística.

La inferencia paramétrica considera que cada medida de probabilidad de la familia está indexada por un parámetro real o vectorial, en tal caso escribimos $\mathcal{P} = \{ \mathbb{P}_\theta: \theta \in \Theta \}$, donde Θ es el llamado espacio de parámetros, el cual es finito-dimensional. Es usual suponer, en este caso, que a cada medida de probabilidad se le puede asociar una función de densidad (o cuantía), de esta forma surgen los modelos estadísticos más habituales en la literatura: modelos binomial, Poisson, exponencial, normal, etc.

La inferencia no-paramétrica se refiere al caso cuando no se supone forma alguna para las medidas de probabilidad de la familia \mathcal{P} , en tal caso puede considerarse como conjunto de índices el conjunto de todas las funciones de distribución F de interés para el problema. Luego anotamos $\mathcal{P} = \{ \mathbb{P}_F: F \in \mathcal{F} \}$ y es claro que en este caso el espacio de parámetros es infinito-dimensional.

Finalmente, se ha convenido en llamar modelo estadístico semi-paramétrico, a aquel modelo que toma en consideración tanto parámetros finito-dimensionales como infinito – dimensionales, habitualmente considerando una estructura jerárquica en su definición.

Por otra parte, el enfoque Clásico considera que θ (o F) es determinístico, pero desconocido; mientras que el enfoque Bayesiano considera que θ (o F) es aleatorio. En este último caso, además del modelo (X, \mathcal{A}, P) se debe considerar un trío – en rigor un espacio de probabilidad – que describa probabilísticamente información inicial (subjetiva) respecto de θ (o F), esto es, $(\Theta, \mathcal{B}, \pi)$ o (F, \mathcal{B}, π) , donde \mathcal{B} es la respectiva σ - álgebra y π es la distribución a priori.

Denotaremos por $p(\cdot | \theta)$ ó $p(\cdot | F)$ la correspondiente función de densidad o probabilidad definida por P .

De esta forma, las medidas de probabilidad del modelo estadístico inicial pueden ser consideradas condicionales en θ o F , es decir,

$$P = \{ p(\cdot | \theta), \theta \in \Theta \} \text{ o } P = \{ p(\cdot | F), F \in F \}.$$

La actualización de la información a priori se hace a través de la regla de Bayes, obteniéndose la llamada distribución a posteriori

$$\pi(\theta | \mathbf{x}) \propto p(\mathbf{x} | \theta) \cdot \pi(\theta), \quad \theta \in \Theta.$$

Es usual visualizar esta situación como un problema que incluye variables aleatorias observables (los datos) y variables aleatorias no-observables (parámetros).

Una forma útil de visualizar el modelo Bayesiano es en dos etapas o jerarquías, la primera corresponde al modelo observacional y la segunda a la distribución a priori, es decir,

$$(1.1) \quad \mathbf{x} | \theta \sim p(\mathbf{x} | \theta)$$

$$(1.2) \quad \theta \sim \pi(\theta).$$

Aquí se utiliza la notación $\mathbf{x} | \theta \sim f(\mathbf{x} | \theta)$ que quiere decir que la función de densidad condicional de \mathbf{x} dado θ es $f(\mathbf{x} | \theta)$.

Cuando la etapa (1.2) se especifica en sub-etapas, se habla de modelos jerárquicos.

Como ejemplo, consideremos el modelo lineal jerárquico, que fue introducido por Lindley y Smith (1972):

$$\begin{aligned} \mathbf{y} | \beta_1, \phi &\sim N(\mathbf{X}_1 \beta_1, \phi^{-1} \mathbf{I}_n) \\ \beta_1 | \beta_2 &\sim N(\mathbf{X}_2 \beta_2, \mathbf{C}) \\ \beta_2 &\sim N(\mathbf{b}, \mathbf{B}) \\ \phi &\sim \text{Gamma}(n_0 / 2, n_0 \sigma_0^2 / 2) \end{aligned}$$

y fue reformulado en la década de los 90 por Mac Eachern (1994), West y Turner (1994), Escobar y West (1995) y Müller, Erkanli y West (1996), considerando componentes no paramétricos en la jerarquización.

Escobar & West (1998) relajaron el supuesto de normalidad y propusieron la siguiente alternativa:

$$\begin{aligned}
 y_i | \theta_i, \zeta &\sim^{\text{cond. ind.}} F(y_i | \theta_i, \zeta) \\
 \theta_i | \lambda &\sim^{\text{cond. i.i.d.}} G(\theta_i | \lambda) \\
 \zeta &\sim H_1 \\
 \lambda &\sim H_2,
 \end{aligned}$$

donde H_1 y H_2 son las distribuciones a priori para ζ y λ , respectivamente.

Este modelo paramétrico se extiende al caso semiparamétrico, modelando la incerteza relativa a la forma funcional de G , la cual depende de una función de distribución conocida G_0 indexada o dependiente en λ (parámetro de posición) y de un parámetro de precisión α (que para el desarrollo de este trabajo hemos denominado \mathbf{c}):

$$\begin{aligned}
 (1.3) \quad y_i | \theta_i, \zeta &\sim^{\text{cond. ind.}} F(y_i | \theta_i, \zeta) \\
 \theta_i | G, \lambda &\sim^{\text{cond. i.i.d.}} G(\theta_i) \\
 G | \lambda, \alpha &\sim D(G | G_0(\cdot | \lambda), \alpha) \\
 \zeta &\sim H_1 \\
 \lambda &\sim H_2 \\
 \alpha &\sim H_3.
 \end{aligned}$$

La especificación de una distribución a priori para G requiere de la consideración de una medida de probabilidad sobre el espacio de funciones de distribución. En este contexto, el proceso de Dirichlet ha tenido un rol fundamental y fue presentado por primera vez por Ferguson (1973), quien prueba su existencia y lo caracteriza a partir de procesos con incrementos independientes Gama.

El objetivo de este autor fue, precisamente, hallar una distribución a *priori* sobre el espacio de todas las funciones de distribución que satisfaga dos condiciones:

- Que tenga un soporte suficientemente grande, de manera que represente a la mayor cantidad de posibles opiniones acerca del modelo.
- Que las distribuciones a *posteriori* resultantes sean analíticamente manejables.

Dos conceptos importantes para comprender los hallazgos de Ferguson (1973) son los siguientes:

Medida de Probabilidad Aleatoria

Sea $(\Omega, \mathcal{F}, \mu)$ espacio de probabilidad y (X, \mathcal{A}) espacio medible. Una *medida de probabilidad aleatoria sobre* $\Omega \times \mathcal{A}$, es una función \mathbf{P} de $\Omega \times \mathcal{A}$ en $[0, 1]$, tal que:

- Para cada $\omega \in \Omega$, $\mathbf{P}(\omega, \cdot)$ es una medida de probabilidad sobre (X, \mathcal{A}) y
- Para cada $A \in \mathcal{A}$, $\mathbf{P}(\cdot, A)$ es una variable aleatoria definida sobre $(\Omega, \mathcal{F}, \mu)$ asumiendo valores en $[0,1]$.

En este contexto es que Ferguson prueba que asociada a cada medida finita α sobre un conjunto X y asociada a cada $\mathbf{c} \in \mathbb{R}^+$, existe una medida de probabilidad aleatoria que resulta ser discreta con probabilidad 1. A la clase a la cual pertenecen estas medidas de probabilidad aleatoria, Ferguson las denomina *Procesos de Dirichlet*.

Proceso de Dirichlet

Sea α una medida finita no – nula sobre (X, \mathcal{A}) y sea $(\Omega, \mathcal{F}, \mu)$ espacio de probabilidad. Se llama proceso de Dirichlet de parámetro α a una medida de probabilidad aleatoria, \mathbf{P} , sobre $\Omega \times \mathcal{A}$ tal que para toda partición medible $\{A_1, \dots, A_k\}$ de X , el vector aleatorio $(\mathbf{P}(A_1), \dots, \mathbf{P}(A_k))$ tiene distribución Dirichlet Singular de parámetro $(\alpha(A_1), \dots, \alpha(A_k))$, donde la variable aleatoria $\mathbf{P}(A_i)$ denota a la variable aleatoria $\mathbf{P}(\cdot, A_i)$.

Lo anterior, se anota $\mathbf{P} \sim \mathcal{D}(\alpha)$.

Ferguson muestra que la clase, es cerrada bajo la operación Bayesiana, es decir, *priori de Dirichlet conduce a posteriori de Dirichlet*, en otras palabras,

$$\text{Si } X_1, \dots, X_n | \mathbf{P} \sim \mathbf{P},$$

$$\text{y } \mathbf{P} \sim \mathcal{D}(\alpha)$$

$$\text{entonces } \mathbf{P} | X_1, \dots, X_n \sim \mathcal{D}\left(\alpha + \sum_{i=1}^n \delta_{X_i}\right),$$

donde $\delta_x(A) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$, es la medida de Dirac en x .

Una caracterización alternativa del Proceso Dirichlet, a partir de secuencias de Polya, es debida a Blackwell y Mac Queen (1973), la cual facilita la deducción de las distribuciones condicionales necesarias para la implementación computacional en la inferencia Bayesiana semi-paramétrica.

Sea μ una medida finita sobre (X, \mathcal{A}) , se dice que $\{X_n\}$ es una secuencia de Polya con parámetro μ si:

$$(1.4) \quad \mathbf{P}(X_1 \in A) = \mu(A) / \mu(X)$$

$$(1.5) \quad \mathbf{P}(X_{n+1} \in A | X_1, \dots, X_n) = \mu_n(A) / \mu_n(X)$$

donde:

$$(1.6) \quad \mu_n = \mu + \sum_{i=1}^n \delta_{X_i} .$$

Note que esta representación permite expresar la medida μ_n como la suma de la información a priori μ y la información muestral acumulada hasta el momento.

El siguiente Teorema muestra que un proceso Dirichlet μ^* se puede aproximar mediante una medida m_n , que corresponde a un promedio ponderado de la medida original μ y la función de distribución empírica F_n asociada a la muestra aleatoria (m.a.) X_1, \dots, X_n .

Teorema de Blackwell & Mac Queen (1973)

Si $\{X_n\}$ es una secuencia de Polya de parámetro μ , entonces:

$$(1.7) \quad m_n = \mu_n / \mu(X) \xrightarrow{c.s.} \mu^*$$

$$(1.8) \quad \mu^* \sim D(\mu)$$

$$(1.9) \quad X_1, X_2, \dots | \mu^* \sim \text{ind } \mu^*$$

Note que:

$$(1.10) \quad m_n = \mu_n / \mu_n(X) = [\mu + \sum_{i=1}^n \delta_{X_i}] / [\mu(X) + \sum_{i=1}^n \delta_{X_i}(X)].$$

Así si $c = \mu(X)$ es la medida del espacio, y si se define una medida de probabilidad en función de μ como $P(A) = \mu(A) / \mu(X)$, se tiene que:

$$m_n(\cdot) = \frac{\mu + \sum_{i=1}^n \delta_{X_i}}{c + n} = \frac{c}{c + n} \cdot P(\cdot) + \frac{n}{c + n} \cdot F_n(\cdot).$$

Aunque hay muchas representaciones del Proceso Dirichlet, esta es la que permitirá la implementación computacional en nuestro caso, en particular el muestreo de Gibbs introducido por Gelfand y Smith (1990).

Estos resultados fueron utilizados por Escobar y West (1998) para obtener soluciones Bayesianas en el modelo jerárquico (1.3).

Basados en el teorema anterior, ellos construyen un algoritmo, el cual permite, usando Gibbs Sampling, simular valores de los parámetros desde las distribuciones condicionales completas, aprovechando el extenso desarrollo computacional de la década pasada.

A manera de ilustración, consideraremos uno de los modelos presentados por estos autores.

Consideremos el siguiente modelo jerárquico,

$$(1.11) \quad \begin{aligned} y_i | \omega_i &\sim f(y_i | \omega_i) \\ \omega_i | G &\sim \text{c.i.i.d. } G \\ G &\sim D(c G_0) \end{aligned}$$

Escobar & West (1998) prueban que la densidad a posteriori de $[(\omega_1, \dots, \omega_n) | y_1, \dots, y_n]$ viene dada por:

$$\pi[(\omega_1, \dots, \omega_n) | y_1, \dots, y_n] \propto \prod_{i=1}^n f(y_i / \omega_i) \cdot \frac{c \cdot G_0(d\omega_i) + \sum_{k < i} \delta(\omega_i / \omega_k)}{c + i - 1}$$

$$\text{donde } \delta(\omega_i | \omega_k) =: \delta_{\{\omega_k\}}(\omega_i) = \begin{cases} 1, & \text{si } \omega_i = \omega_k \\ 0, & \text{si } \omega_i \neq \omega_k \end{cases}.$$

Un esbozo de la demostración es presentada a continuación.

Considere el modelo jerárquico (1.11).

$$\begin{aligned}\pi(\omega_1, \dots, \omega_n | y_1, \dots, y_n) &= \int \pi(\omega_1, \dots, \omega_n | y_1, \dots, y_n; G) \cdot \pi(G | y_1, \dots, y_n) dG \\ &= \int \pi(\omega_1, \dots, \omega_n | y_1, \dots, y_n; G) \cdot d\pi^*(G),\end{aligned}$$

donde $\pi^* \sim D(c + \sum \delta_{y_i})$.

Luego,

$$\begin{aligned}\pi(\omega_1, \dots, \omega_n | y_1, \dots, y_n) &\propto \int \pi(y_1, \dots, y_n | \omega_1, \dots, \omega_n; G) \cdot \pi(\omega_1, \dots, \omega_n | G) d\pi^*(G). \\ &\propto \int \prod_{i=1}^n f(y_i | \omega_i) \cdot \pi(\omega_1, \dots, \omega_n | G) d\pi^*(G). \\ &\propto \prod_{i=1}^n f(y_i | \omega_i) \cdot \int \pi(\omega_1, \dots, \omega_n | G) d\pi^*(G) \text{ (Blackwell y Mac Queen, 1973)}.\end{aligned}$$

Pero $\{\omega_i\}$ es una secuencia de Polyà, de parámetro $c + \sum \delta_{y_i}$; de donde, desarrollando lo anterior se llega al resultado.

Denotemos por ω_{-i} a $\{\omega_k; k \neq i\}$. De lo anterior, se deduce la distribución condicional completa para cada uno de los ω_i , la cual resulta expresarse de forma más conveniente:

$$(1.12) \quad \omega_i | \omega_{-i}, y_1, \dots, y_n, G \propto q_0 \cdot G_b(d\omega_i) + \sum_{k \neq i} q_k \cdot \delta(d\omega_i | \omega_k),$$

donde la distribución base G_b y los pesos q_0, \dots, q_n se obtienen desde:

$$(1.13) \quad \begin{aligned}G_b(d\omega_i | y_i) &\propto f(y_i | \omega_i) \cdot G_0(d\omega_i), \\ q_{0i} &\propto \int f(y_i | \omega_i) \cdot dG_0(\omega_i), \\ q_{ki} &\propto f(y_i | \omega_k),\end{aligned}$$

con q_0 dependiendo sólo de y_i y con q_k dependiendo de y_i y ω_{-i} . En efecto,

$$\pi(\omega_n | \omega_{-n}, \mathbf{y}) \propto f(y_n | \omega_n) \cdot \frac{c \cdot G_0(d\omega_n) + \sum_{j \neq n} \delta(d\omega_n / \omega_j)}{c + n - 1}.$$

Ahora bien, como los $\{\omega_i\}$ son permutables, se tiene que la distribución proporcional hallada arriba es válida para cualquiera de ellos, de donde:

$$\pi(\omega_i | \omega_{-i}, \mathbf{y}) \propto \frac{c \cdot f(y_i | \omega_i) \cdot G_0(d\omega_i) + \sum_{j \neq i} f(y_i | \omega_j) \cdot \delta(d\omega_i / \omega_j)}{c + n - 1}.$$

Haciendo $q_{0i} = c \cdot f(y_i)$, $q_{ki} = f(y_i | \omega_k)$, $k \neq i$, y $G_b(d\omega_i) = \frac{f(y_i | \omega_i) \cdot G_0(d\omega_i)}{f(y_i)} \propto f(y_i | \omega_i) \cdot G_0(d\omega_i)$, se

tiene el resultado.

Por ejemplo, si f corresponde a la función de densidad Normal con media μ y varianza $\sigma^2 = \phi^{-1} \omega_i$, donde la priori para ϕ es Gama y G_0 es una distribución Gama Invertida, se obtiene el siguiente modelo jerárquico:

$$(1.14) \quad \begin{aligned} y_i | \phi, \omega_i &\sim \text{ind. } N(\mu; \phi^{-1} \omega_i) \\ \phi &\sim \text{Gamma}(r_1; \lambda_1) \\ \omega_1, \dots, \omega_n | G &\sim \text{i.i.d. } G \\ G &\sim D(c G_0); G_0 = IG(r_0, \lambda_0) \end{aligned}$$

Entonces la distribución base es $G_b(\cdot | y_i)$, la distribución $IG(r_0 + 1/2; \lambda_0 + \frac{\phi}{2}(y_i - \mu)^2)$ y los pesos vienen dados por q_0 y q_k , que son proporcionales, respectivamente, a la densidad $\int f(y_i | \omega_i) \cdot dG_0(\omega_i) = f(y_i)$ de la $T(\mu; 2\lambda_0\phi^{-1}; 2r_0)$ y a la densidad $f(y_i | \omega_k)$ de la $N(\mu; \omega_k\phi^{-1})$, evaluadas en y_i . Estos pesos no pueden ser negativos.

Así se obtienen las distribuciones condicionales siguientes (ver Glosario de Notaciones en Anexo):

$$(1.15) \quad \begin{aligned} y_i | \phi, \omega_i &\sim \text{ind. } N(\mu; \phi^{-1} \omega_i) \\ \omega_i | \omega_{-i}, y_1, \dots, y_n &\propto q_0 \cdot G_b(d\omega_i) + \sum_{k \neq i} q_k \cdot \delta(d\omega_i | \omega_k) \\ \phi | \omega_1, \dots, \omega_n, y_1, \dots, y_n &\sim \text{Gamma}\left(\frac{n}{2} + r_1, \lambda_1 + \sum_{i=1}^n \frac{(y_i - \mu)^2}{2\omega_i}\right). \end{aligned}$$

Recientemente, West (1992), Mengersen y Robert (1993), Diebolt y Robert (1994), Nobile (1994), Escobar y West (1995), Roeder y Wasserman (1995), Raftery (1995) y Carroll et al. (1999) mostraron que las mezclas de normales proveen una familia simple y flexible. Esta familia forma parte de la familia de modelos elípticos (Fang et al, 1990), que serán el foco de este trabajo.

1.2 ALGORITMO DE ESCOBAR & WEST

Escobar y West (1998) basan su algoritmo computacional en los resultados de Bush y MacEachern (1996), quienes resuelven las dificultades técnicas que se presentan al muestrear directamente de las distribuciones condicionales, utilizando la estructura de conglomerados implícita en el proceso de *Dirichlet*.

Se denota por ω_k^* a los valores distintos de ω_i , formando de esta manera conglomerados que los agrupan y se denota por n_k^* al número de elementos de cada conglomerado. Puesto que es importante hallar la constante de proporcionalidad que permita expresar de forma más exacta la distribución presentada en (1.12), se estandarizan los pesos q_k de la siguiente forma

$$q_0^* = q_0 / \sum_{j \neq i} q_j \quad \text{y} \quad q_k^* = q_k / \sum_{j \neq i} q_j.$$

Si se denota por I^* al conjunto de índices de los ω_k^* y se denota por K al número de conglomerados ($K = \# I^*$), la distribución condicional completa de los ω_i se puede re-expresar como:

$$\omega_i | \omega_{-i}, y_1, \dots, y_n, G, \alpha \sim q_0^* \cdot G_b(d\omega_i | y_i) + \sum_{k=1}^K n_k^* q_k^* \cdot \delta(d\omega_i | d\omega_k).$$

El algoritmo incluye una etapa adicional en el muestreo de Gibbs que se denomina “re-mezcla”, la cual tiene como objetivo prevenir que la cadena de Markov se estabilice en un grupo pequeño de cluster, lo cual puede ocurrir si la suma de los q_k^* es grande respecto de q_0^* .

Esta re-mezcla se debe realizar después de cada paso del algoritmo estándar, y el procedimiento es el siguiente:

Sea $S = (S_1, \dots, S_n)$, donde $S_i = j$ si $\omega_i = \omega_j^*$, es decir, si ω_i pertenece al conglomerado j .

Sea $J_j = \{i \in \{1, \dots, n\} : S_i = j\}$ el conjunto de los índices de objetos en el conglomerado j .

Sea $Y(j) = \{y_i : S_i = j\}$ el conjunto de observaciones correspondientes al conglomerado j .

Luego, los ω_j^* dada la estructura de conglomerado son condicionalmente independientes e idénticamente distribuidos con la siguiente distribución:

$$\omega_j^* | y_1, \dots, y_n, S, I^* = \omega_j^* | Y(j), S, I^* \propto \prod_{i \in J_j} f_{Y_i}(y_i | \omega_j^*) \cdot G_0(d\omega_j^*) = \prod_{i \in J_j} G_b(d\omega_j^* | y_i)$$

Distribución a priori sobre el parámetro de precisión $c = \alpha(\mathbf{x})$

Si en el modelo (1.11) se agrega una etapa que consiste en especificar una distribución inicial para c , entonces será necesario obtener la distribución condicional completa de c :

$$\begin{aligned} y_i | \omega_i &\sim \text{i.nd. } f(y_i | \omega_i) \\ \omega_1, \dots, \omega_n | G &\sim \text{c.i.i.d. } G \\ G &\sim D(c | G_0) \\ c &\sim F \end{aligned}$$

Este parámetro c incide en el peso asignado a la distribución Priori G_0 para G , es decir, para valores pequeños de c , G se concentra en un pequeño conjunto de valores de los ω_i (se habla de cluster) y para valores grandes de c la priori de G , G_0 , asume mayor importancia.

Escobar y West (1998) sugieren una priori Gamma para c , ya que esta distribución permite tanto valores grandes como pequeños.

Si la priori para c es Gamma(a, b), con a parámetro de forma y $b > 0$ parámetro de escala, la posteriori de c corresponde a la mezcla de Gammas (1.16), que se presenta más abajo.

Escobar y West (1998) incluyen en su algoritmo una variable latente η con priori Beta:

$$\eta / c, I^* \sim \text{Beta}(c + 1, n).$$

Esta variable latente depende de c y de los cluster que se forman con los valores distintos que asumen los ω_i .

La distribución condicional completa de c resulta no depender de otros parámetros aparte de η y de I^* (I^* contiene la estructura de los cluster, y la dependencia aquí se vincula únicamente al número de cluster K), y corresponde a una mezcla de Gammas. En efecto:

$$(1.16) \quad c / \eta, I^* \sim \pi_\eta \text{Gamma}(a+K, b - \log(\eta)) + (1 - \pi_\eta) \text{Gamma}(a+K - 1, b - \log(\eta)),$$

donde π_η se define en función de un cuociente de chances:

$$\frac{\pi_\eta}{1 - \pi_\eta} = \frac{a + K - 1}{n \cdot (b - \log(\eta))} \Rightarrow \pi_\eta = \left[1 + \frac{n \cdot (b - \log(\eta))}{a + K - 1} \right]^{-1}$$

1.3 MODELOS LINEALES JERÁRQUICOS ELÍPTICOS

Box y Tiao (1973) fueron los primeros en extender las inferencias desde el modelo con errores normales a un modelo no normal dentro de la clase elíptica (modelo con errores en la familia exponencial potencia) desde una perspectiva bayesiana. Por otra parte, Zellner (1976) resulta ser el pionero en el estudio del modelo t – student multivariado (elíptico dependiente). Mucho después, Geweke (1993) aborda el mismo problema, considerando errores independientes con distribución t-Student. Por otra parte, Jammalamadaka et al. (1987) y Chib et al. (1988) abordan este tema, modelando los errores como mezclas de normales.

Estos resultados son extendidos por Osiewalski y Steel (1993) a la clase de distribuciones elípticas, definidas a continuación.

Distribución Elíptica

Sea \mathbf{y} un v.a. $n \times 1$. Se dice que \mathbf{y} tiene una distribución elíptica con parámetro de posición $\boldsymbol{\mu}$ ($n \times 1$) y matriz de dispersión Σ ($n \times n$) definida positiva ($\Sigma > 0$), si existe h , conocida como función generadora de densidad y definida para reales no – negativos, tal que la densidad de \mathbf{y} está dada por

$$(1.17) \quad f_{\mathbf{Y}}(\mathbf{y}) = |\Sigma|^{-1/2} h [(\mathbf{y} - \boldsymbol{\mu})^t \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu})]$$

y h satisface:

$$(1.18) \quad \int_0^{\infty} \frac{\pi^{n/2}}{\Gamma(n/2)} u^{n/2} h(u) du = 1.$$

La notación, $\mathbf{y} \sim El_n(\boldsymbol{\mu}, \Sigma, h)$, es frecuentemente utilizada para denotar este hecho.

El Modelo Elíptico tiene como casos particulares al Modelo Normal, t – Student y otros. Para la distribución normal multivariada la función generadora de densidad corresponde a:

$$h(u) = (2\pi)^{-n/2} \exp(-u/2)$$

Por otra parte, la distribución elíptica, al igual que la distribución normal, admite estandarización, es decir, si $\mathbf{y} \sim El_n(\boldsymbol{\mu}, \Sigma, h)$, entonces

$$(1.19) \quad \mathbf{z} = \Sigma^{-1/2} (\mathbf{y} - \boldsymbol{\mu}) \sim El_n(\mathbf{0}, I_n, h).$$

Distribución Esférica

Se dice que \mathbf{z} tiene distribución Esférica si existe h satisfaciendo (1.18), y la f.d.p. de \mathbf{z} corresponde a:

$$(1.20) \quad f_{\mathbf{Z}}(\mathbf{z}) = h(\|\mathbf{z}\|^2).$$

Es claro que una distribución elíptica estándar, definida por (1.19), corresponde a una distribución esférica.

Existen dos alternativas para este modelo y son las siguientes:

- El modelo elíptico dependiente que preserva la esfericidad y relaja la independencia:

En este caso, $\boldsymbol{\varepsilon} \sim El_n(\mathbf{0}; \sigma^2 \mathbf{I}_n, h)$, la f.d.p. de \mathbf{y} corresponde a:

$$f_{\mathbf{y}}(\mathbf{y}) = \sigma^{-n} \cdot h \left[\sum_{i=1}^n (\mathbf{y}_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / \sigma^2 \right].$$

- El modelo elíptico independiente que relaja la esfericidad conjunta:

En este caso, $\varepsilon_1, \dots, \varepsilon_n \sim \text{ind. } El_1(\mathbf{0}; \sigma_i^2, h)$, la f.d.p. de \mathbf{y} corresponde a:

$$f_{\mathbf{y}}(\mathbf{y}) = \prod_{i=1}^n \sigma_i^{-n} \cdot h \left((\mathbf{y}_i - \boldsymbol{\mu}_i)^2 / \sigma_i^2 \right).$$

Ambos modelos coinciden sólo bajo normalidad (Kelker, 1970 y Arellano – Valle et. al., 1994).

Modelo Elíptico Representable

Los Modelos Elípticos también pueden clasificarse de acuerdo a la naturaleza de la función generadora de densidad h , la cual puede ser Representable o No Representable.

Se dice que h es representable (y por lo tanto el Modelo Elíptico es Representable) si y sólo si existe una función de distribución H con $H(0) = 0$, tal que:

$$(1.21) \quad h(\mathbf{u}) = \int_0^{\infty} \frac{e^{-\mathbf{u}/2v}}{(2\pi v)^{1/2}} dH(v).$$

La relación entre h y H es 1 a 1, y la función generador en el caso Representable es una mezcla en parámetro de escala de distribuciones normales.

De manera análoga, se dice que h (y por lo tanto el Modelo Elíptico) no es representable, si no existe H con $H(0) = 0$, que satisfaga (1.21) para la función generador h .

De esta manera se define la subclase de distribuciones Normales Compuestas dentro de la Familia Elíptica, la cual incluye a todos aquellos modelos Elípticos Representables.

La importancia de la subclase de Modelos Elípticos Representables se expresa en las *Proposiciones 2.1.a)* y *2.1. b)*.

1.4 PLAN DE TESIS

El objetivo del presente trabajo de tesis es abordar el problema de inferencia bayesiana relativa a los parámetros para tres modelos de interés, los cuales son el modelo de calibración, el modelo con errores de medición y el modelo de calibración comparativa.

En el capítulo 2 se clasifican las distribuciones elípticas, se presentan los modelos a abordar y se expone un conjunto de proposiciones y lemas que serán utilizados en los capítulos siguientes, entre los cuales destacan dos proposiciones que lo vinculan al modelo normal.

Los capítulos 3 y 4 abordan el problema de calibración para variables continuas y discretas, respectivamente, y el análisis para los modelos de errores de medición y calibración comparativa es presentado en los capítulos 5 y 6, respectivamente .

En los capítulos 3, 5 y 6 se presentan las distribuciones condicionales completas para la implementación computacional. Estas distribuciones se obtuvieron en primer lugar para el caso en que el modelo contempla errores normales, extendiéndose aquí los resultados para el caso elíptico, dependiente e independiente.

El capítulo 4 aborda el caso de calibración discreta con respuesta dicotómica o bernoulli (presencia / ausencia de cierta característica) y con respuesta binomial (número de ocurrencias del fenómeno de interés en una muestra de tamaño fijo), haciendo uso de variables subyacentes continuas elípticas para la respuesta, extendiendo así los resultados de Basu y Mukhopadhyay (2000).

Cada capítulo presenta también aplicaciones y simulaciones, que permiten ejemplificar en forma práctica la resolución de los problemas estudiados.

Todos estos modelos han sido tradicionalmente estudiados desde la perspectiva clásica y desde la perspectiva bayesiana paramétrica, generalmente limitándose a asumir a una distribución normal para los errores. Algunas contribuciones y extensiones en esta área, considerando distribuciones elípticas en el modelo paramétrico, son debidas a Fernández y Steel (1998), Vidal et. al. (2005) y Arellano et. al. (2003).

La contribución del presente trabajo consiste en extender los resultados de los modelos de regresión y afines mencionados al modelo elíptico semi-paramétrico, desde la perspectiva bayesiana. Para este fin se compara la resolución clásica (vinculada a la función de verosimilitud) con la alternativa bayesiana Semiparamétrica.

El abordaje de los modelos mencionados requirió, en todos los casos, la obtención de distribuciones condicionales completas que permitieran utilizar el Muestreo de Gibbs, debido a que las expresiones para las distribuciones a posteriori no son obtenibles analíticamente, así como tampoco lo son los estimadores clásicos tradicionales.

Los programas utilizados para implementar los algoritmos que configuran el Muestreo de Gibbs, fueron compilados en el programa THE SAS SYSTEM V8. Para cada caso se realizaron 70.000 iteraciones, y las cadenas fueron analizadas con el programa BOA (Bayesian Output Analysis v 1.0.0) en la plataforma R V 1.8.1 PROJECT. De acuerdo a los resultados entregados por el programa BOA, se decidió eliminar las primeras 10.000 iteraciones y realizar un muestreo sistemático de un 10%.

Finalmente se incluye un breve capítulo de conclusiones, el cual también menciona las extensiones futuras que tiene el presente trabajo.

CAPÍTULO 2:

MODELO DE REGRESIÓN LINEAL BAJO DISTRIBUCIONES ELÍPTICAS REPRESENTABLES

El presente capítulo expone los modelos y problemas abordados en este trabajo y que serán desarrollados extensamente en los capítulos 3 al 6. Los problemas de los que hablamos aquí son: el problema de Calibración (discreta y continua), el problema de errores de medición y el problema de calibración comparativa. Todos estos problemas se presentan en la primera sección y constituyen extensiones del problema de regresión. La segunda sección expone las relaciones existentes entre el modelo elíptico, asumido para los errores, y el modelo normal, lo cual permite extender los resultados para el caso normal al caso elíptico. La tercera sección expone las distribuciones a posteriori para diferentes componentes de los modelos jerárquicos básicos abordados en capítulos posteriores, lo cual permitirá derivar las distribuciones condicionales completas requeridas para la implementación computacional y obtener de este modo, sin necesidad de contar explícitamente con las distribuciones a posteriori para los parámetros de interés, los estimadores de Bayes.

2.1 MODELOS LINEALES ABORDADOS

El análisis Bayesiano de modelos de regresión elípticos ha atraído la atención de diversos autores en las últimas décadas, a partir de los trabajos pioneros de Box y Tiao (1973) y Zellner (1976), quienes trataron la familia exponencial potencia y t – student, respectivamente. Estos resultados fueron extendidos a mezclas en parámetro de escala de modelos de regresión normal por Jamalamadaka et. al (1987) y Chib et. al (1988) y la clase entera de las distribuciones elípticas por Osiewalski y Stell (1993). Todos estos autores describen una clase de distribuciones a priori no informativas, que producen distribuciones predictivas y posterioris para los coeficientes de regresión, las cuales coinciden con aquellas obtenidas bajo normalidad. Los alejamientos de la normalidad dentro de la clase de las distribuciones elípticas solamente afectan a las inferencias sobre el parámetro de escala. Extensiones sobre este mismo tema son presentadas en Arellano – Valle, Del Pino e Iglesias (2002a) y Arellano – Valle, Iglesias y Vidal (2002b). Este último contempla una revisión de los trabajos previos e incluye comparación de modelos.

El problema de calibración bajo modelos de regresión elíptico ha sido abordado en Branco et al. (2001) y el problema de diagnóstico en modelos de regresión elípticos por Arellano – Valle et al. (2000). Todos estos autores abordan los temas dentro de la perspectiva paramétrica. En la misma dirección, Viviani et. al (2003) y Arellano y Bolfarine (2004) presentan soluciones para inferencia en modelos de regresión elíptica con errores en las variables y modelos de calibración comparativa.

Como se mencionó en la introducción, el objetivo de este trabajo es entregar un tratamiento unificado para el análisis Bayesiano de modelos de regresión elípticos, restringiéndonos a la subfamilia de mezcla en parámetro de escala de modelos de regresión elíptica, bajo un enfoque semiparamétrico.

Adicionalmente consideramos también modelos de regresión discreta. Respecto de este tema, Basu y Mukhopadhyay (2000) realizan un gran aporte en el tratamiento del modelo de regresión binaria.

Todos estos problemas son descritos y abordados a continuación.

Formulación del Modelo de Calibración

Este problema fue planteado originalmente por Hoadley (1970) y Hunter y Lamboy (1981) dentro del contexto de modelos normales y bajo la perspectiva bayesiana. Brown (1993) resume los principales resultados concernientes a este problema, el cual es extendido por Branco et al. (2000) a la familia elíptica, bajo los modelos dependiente e independiente.

Se entiende por modelo de calibración a la situación en la cual se cuenta con n pares de mediciones observadas (x_i, y_i) , para las cuales se tiene que y_i depende funcionalmente de x_i más un error aditivo. Las variables x_i son fijas y las y_i son variables aleatorias observables. En este contexto se supone que las mediciones x_i se han realizado con un instrumento mucho más preciso que las mediciones y_i , pero cuya operación también es mucho más costosa. Por otra parte, se cuenta con una medición observada y_0 , pero no se ha observado x_0 , que es el parámetro de interés y respecto del cual se desea concluir. Precisamente, el problema de inferencia acerca de x_0 es lo que se conoce como problema de calibración.

El modelo de calibración es el siguiente:

$$(2.1) \quad y_i = f(x_i) + \varepsilon_i, \quad i = 0, \dots, n; \quad \text{con } \varepsilon_i \sim \text{ind} (E(\varepsilon_i) = 0, V(\varepsilon_i) = \sigma^2), \quad i = 0, \dots, n, \quad \text{y con } x_0 \text{ desconocido.}$$

Es usual suponer que las variables aleatorias ε_i son i.i.d. con $E(\varepsilon_i) = 0$ y $V(\varepsilon_i) = \sigma^2$.

Los modelos de calibración más frecuentes en la literatura son el modelo lineal continuo y modelo lineal discreto, dependiendo de la naturaleza de las variables y_i .

Modelo de Calibración Lineal Continua

Si las variables y_i son continuas, se puede definir el siguiente modelo:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 0, \dots, n;$$

donde β_0 , β_1 y σ^2 son los parámetros del modelo, x_0 es desconocido y los x_1, \dots, x_n se consideran constantes conocidas. Las variantes del modelo surgen en función de la distribución para los términos de error, respecto de la cual asumiremos, dentro del contexto de este trabajo, que pertenece a la clase de las distribuciones elípticas representables.

Modelo de Calibración Lineal Discreta

Si, por otra parte, las variables y_i son discretas, se puede definir el siguiente Modelo:

$$P(Y_i = y) = H(\beta_0 + \beta_1 x_i); \quad y \in X.$$

Donde X es un conjunto discreto.

Un caso importante es el que se deriva del Modelo de Regresión Binaria, el cual se puede expresar del modo siguiente:

$$P(Y_i = y) = [H(\beta_0 + \beta_1 x_i)]^y [1 - H(\beta_0 + \beta_1 x_i)]^{1-y}; \quad y \in \{0,1\}; \quad i = 1, \dots, n;$$

modelo para el cual,

$$y_i \sim \text{ind Bernoulli}(\theta), \quad \text{con } \theta = H(\beta_0 + \beta_1 x_i), \quad \theta \in [0, 1].$$

En forma análoga, se deriva el Modelo de Regresión Binomial, el cual se puede expresar en la forma siguiente:

$$P(y_i = y) = \binom{n_i}{y} [H(\beta_0 + \beta_1 x_i)]^y [1 - H(\beta_0 + \beta_1 x_i)]^{n_i - y}; y \in \{0, \dots, n_i\}; i = 1, \dots, k;$$

de donde,

$$y_i \sim \text{binomial}(n_i, p_i), \text{ donde } p_i = H(\beta_0 + \beta_1 x_i); i = 1, \dots, k, p_i \in [0, 1].$$

En ambos casos H se denomina la Función de Enlace y se escoge como una función de distribución acumulada, para que θ asuma valores en el intervalo unitario.

Dos funciones de Enlace conocidas son:

$H = \Phi$, la cual da origen al modelo probito, y

$H = (1 - e^{-x})^{-1}$ (la f.d.a. logística estándar), la cual lleva al modelo logito.

En este trabajo H será función de distribución elíptica representable con distribución de mezcla desconocida.

Modelo con Errores en las Variables (MEM)

El Modelo con Errores en las Variables o Modelo con Errores de Medición (MEM) es una generalización del Modelo de Regresión Lineal, para el cual las variables aleatorias x_i no son observables directamente, sino que se observan con Errores de Medición Aditivos, es decir, este Modelo incluye ciertas variables z_i , tales que:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, i = 1, \dots, n$$

$$z_i = x_i + v_i, i = 1, \dots, n.$$

Desde el punto de vista clásico se admiten dos variantes para este modelo: si los x_i son parámetros *incidentales*, es decir constantes desconocidas, se habla de un *Modelo Funcional*; por otra parte, si los x_i son variables aleatorias, recibe el nombre de *Modelo Estructural*. La perspectiva Bayesiana no hace diferencia en este sentido.

Modelo de Calibración Comparativa

El modelo que se presenta a continuación, puede ser visto como una generalización del modelo aditivo con errores en las variables, para el cual las predictoras observadas z_i se denotan por y_{i0} , y se cuenta con p diferentes variables observadas y_{ij} para cada x_i .

Por otra parte, también puede ser considerado como una generalización del Modelo de Calibración Continua, para el cual, se cuenta con p instrumentos inexactos, en lugar de uno:

$$y_{ij} = \beta_{0j} + \beta_{1j} x_i + \varepsilon_{ij}, i = 1, \dots, n; j = 1, \dots, p.$$

$$y_{i0} = x_i + \varepsilon_{i0}, i = 1, \dots, n.$$

En este contexto, x_i es el verdadero valor de la medición, la cual no se puede observar, pero se cuenta con una medición exacta y_{i0} con error aditivo y p mediciones inexactas y_{ij} , las cuales se relacionan linealmente con x_i incluyendo, naturalmente, un error aditivo.

El objetivo del análisis se centra en la calibración de los p instrumentos inexactos en función del instrumento exacto, es decir en el problema de inferencia sobre los parámetros de regresión $\beta_j = (\beta_{0j}, \beta_{1j})^t$.

2.2 RELACIONES ENTRE MODELOS NORMALES Y ELÍPTICOS

En esta sección se presentarán dos casos en los cuales la estrategia de inclusión de variables latentes permite la resolución del problema de forma más sencilla. Ver por ejemplo Arellano – Valle et al. (1999) y Basu y Mukhopadhyay (2000). La extensión considerada en este trabajo radica en el hecho de que las variables latentes siguen una distribución F desconocida y F tiene una distribución a priori de *Dirichlet*, lo que le imprime el carácter de Semiparamétrico al análisis.

La subclase de modelos elípticos representables será el foco en este trabajo

Relación entre el Modelo Elíptico Representable y el Normal

Proposición 2.1 a): Caso Dependiente

$$\mathbf{y} | \mathbf{X}, \boldsymbol{\beta}, \phi \sim El_n(\mathbf{X} \boldsymbol{\beta}, \phi^{-1} \mathbf{I}_n, h),$$

con h representable, puede ser expresado equivalentemente si existe una variable aleatoria ω , independiente de $(\boldsymbol{\beta}, \sigma^2)$, tal que

$$\mathbf{y} | \mathbf{X}, \boldsymbol{\beta}, \phi, \omega \sim^{ind} N_n(\mathbf{X} \boldsymbol{\beta}, \omega \phi^{-1} \mathbf{I}_n),$$

donde $\omega \sim G$; con $G(0) = 0$.

Demostración:

Por (1.21) y la definición de modelo elíptico representable en la sección 1.3 del capítulo primero, si h es representable entonces existe G tal que:

$$f_Y(\mathbf{y} | \boldsymbol{\beta}, \phi) = \int_0^{\infty} \left(\frac{\phi}{2\pi\omega} \right)^{n/2} \exp\{-\phi (\mathbf{y} - \mathbf{X} \boldsymbol{\beta})^t (\mathbf{y} - \mathbf{X} \boldsymbol{\beta}) / 2\omega\} dG(\omega).$$

Notando que el integrando corresponde a una $N(\mathbf{X} \boldsymbol{\beta}; \phi^{-1} \omega \mathbf{I})$ y que G es una función de distribución acumulada con $G(0) = 0$, se obtiene la primera parte de la prueba.

En la otra dirección basta notar que:

$$f_Y(\mathbf{y} | \boldsymbol{\beta}, \phi) = \int f_Y(\mathbf{y} | \boldsymbol{\beta}, \phi, \omega) dG(\omega). \quad \square$$

Observación 2.1:

En la práctica esta proposición permite que los desarrollos ya conocidos para modelos jerárquicos con errores normales se extiendan a la clase de modelos elípticos representables.

A diferencia de lo existente en la literatura, la función de distribución G que determina el modelo elíptico será considerada desconocida con priori *Dirichlet*.

Proposición 2.1b): Caso Independiente

$$y_i | \mathbf{x}_i, \boldsymbol{\beta}, \phi \sim^{ind} El_1(x_i^t \boldsymbol{\beta}, \phi^{-1}, h); i = 1, \dots, n$$

si y sólo si existen variables aleatorias $\omega_1, \dots, \omega_n$ mutuamente independientes e independientes de $(\boldsymbol{\beta}, \sigma^2)$ tales que

$$y_i | \mathbf{x}_i, \omega_i, \boldsymbol{\beta}, \phi \sim^{ind} N_1(\mathbf{x}_i^t \boldsymbol{\beta}, \phi^{-1} \omega_i); i = 1, \dots, n;$$

donde las variables latentes $\omega_1, \dots, \omega_n$ satisfacen

$$\omega_i \sim^{iid} G; i = 1, \dots, n.$$

Demostración:

Análoga a la demostración de la *Proposición 2.1. (a)*. □

Variables Subyacentes Elípticas en el Modelo Lineal Discreto

Basu & Mukhopadhyay (2000) usan enlaces especiales para enfrentar el modelo de regresión binaria, seleccionando funciones de distribución acumuladas (f.d.a.) en la clase \mathfrak{S} de las distribuciones elípticas representables:

$$\mathfrak{S} = \{F \text{ en el espacio de las f.d.a: } F(\cdot) = \int_0^\infty \Phi(\cdot / \sigma) dG(\sigma), G \text{ f.d.a. sobre } \mathbb{R}_0^+\}.$$

Las siguientes dos proposiciones permitirán relacionar el análisis del modelo de calibración lineal discreto (binario) con el desarrollo de los modelos con errores normales y elípticos, mediante la inclusión de variables subyacentes a la respuesta binaria. La primera (*Proposición 2.2 a*) se debe a Basu & Mukhopadhyay (2000) y la segunda (*Proposición 2.2 b*) es una extensión de la primera.

Proposición 2.2 a): Caso Normal

Se cumple que

$$(2.2) \quad y_i | \beta, G \sim^{ind} \text{Bernoulli} \left(\theta_i = \int_0^\infty \Phi(\mathbf{x}_i^t \beta / \sigma_i) dG(\sigma_i) \right), \quad i = 1, \dots, n.$$

donde Φ es la función de distribución acumulada normal estándar,

si existen variables aleatorias Z_i, σ_i^2 tales que

$$(2.3) \quad \begin{aligned} y_i | Z_i, \beta, \sigma_i^2, G &\sim^{ind} \text{Bernoulli}[\theta_i = P(Z_i > 0)] \\ Z_i | \beta, \sigma_i^2, G &\sim^{ind} N(\mathbf{x}_i^t \beta, \sigma_i^2) & i = 1, \dots, n. \\ y \quad \sigma_i^2 | G &\sim^{i.i.d.} G. \end{aligned}$$

Observación:

La dicotomía de las variables y_i se vinculan a un punto de corte para una variable subyacente continua, estandarizando se vincula al signo de las variables subyacentes Z_i , es decir: $y_i = 1$ ssi $Z_i \geq 0$, $y_i = 0$ ssi $Z_i < 0$.

Demostración:

Usando la especificación (2.3), se tiene que

$$P(y_i = 1 | \beta, G) = P(Z_i > 0 | \beta, G) = \int_0^\infty P(Z_i > 0 | \beta, \sigma_i, G) dG(\sigma_i) = \int_0^\infty P\left(\frac{Z_i - \mathbf{x}_i^t \beta}{\sigma_i} > \frac{-\mathbf{x}_i^t \beta}{\sigma_i} | \beta, \sigma_i, G\right) dG(\sigma_i).$$

Usando ahora la simetría de la distribución normal, se tiene que

$$P(y_i = 1 | \beta, G) = \int_0^\infty P\left(\frac{Z_i - \mathbf{x}_i^t \beta}{\sigma_i} < \frac{\mathbf{x}_i^t \beta}{\sigma_i} | \beta, \sigma_i, G\right) dG(\sigma_i) = \int_0^\infty \Phi(\mathbf{x}_i^t \beta / \sigma_i) dG(\sigma_i) = \theta_i. \quad \square$$

Dado que en la demostración anterior lo esencial es la simetría, el resultado puede ser extendido al caso elíptico.

Proposición 2.2 b): Caso Elíptico

Se cumple que

$$(2.4) \quad y_i | \beta, G \sim^{ind} \text{Bernoulli}(\theta_i = \Phi_E(\mathbf{x}_i^t \beta / \sigma_i)), \quad i = 1, \dots, n$$

donde Φ_E es la función de distribución acumulada elíptica,

si existen variables aleatorias Z_i , tales que

$$(2.5) \quad y_i | Z_i, \beta, G \sim^{ind} \text{Bernoulli}[\theta_i = P(Z_i > 0)],$$

$$y \quad Z_i | \beta, \omega \sim^{ind} \text{EL}_1(\mathbf{x}_i^t \beta, 1, h); i = 1, \dots, n.$$

Demostración:

La nueva función de enlace corresponde a mezclas escala de f.d.a. normales. Note que por la *Proposición 2.1. (b)* se tiene que (2.3) es equivalente a (2.5). Este resultado será utilizado en el capítulo 4.

2.3 ALGUNOS RESULTADOS BÁSICOS PARA INFERENCIA EN MODELOS DE REGRESIÓN ELÍPTICOS REPRESENTABLES

Los siguientes lemas son resultados que serán frecuentemente utilizados para el cálculo de las distribuciones condicionales de los coeficientes de regresión y dispersión requeridos en la implementación computacional de Gibbs Sampling, para los modelos abordados en los subsecuentes capítulos.

2.3.1 MODELO DE REGRESIÓN NORMAL HOMOSCEDÁSTICO

Los siguientes dos lemas presentan, respectivamente, las distribuciones a posteriori para los parámetros de dispersión σ^2 y de calibración \mathbf{x}_0 , en el contexto del modelo de regresión con errores normales de varianza común.

Lema 2.1: Distribución Condicional del Parámetro de Calibración

$$\text{Si} \quad y_i | \mathbf{x}_0 \sim^{ind} N(\mathbf{x}_i^{*t} \beta, \sigma_0^2) \text{ con } \mathbf{x}_i^* = \begin{bmatrix} 1 \\ \mathbf{x}_i \end{bmatrix}_{p \times 1} \quad \text{y con } \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}_{p \times 1} \quad \text{para } i = 0, \dots, n.$$

$$\text{y} \quad \mathbf{x}_0 \sim N_{p-1}(\mathbf{m}_0, v_0)$$

Entonces

$$\mathbf{x}_0 | \mathbf{y} \sim N_{p-1}(\mathbf{m}_1, v_1) \quad \text{con } \mathbf{m}_1 = v_1 \left(\frac{(y_0 - \beta_0)}{\sigma_0^2} \cdot \beta_1 + v_0^{-1} \mathbf{m}_0 \right), \quad v_1 = \left(\frac{\beta_1^t \cdot \beta_1}{\sigma_0^2} + v_0^{-1} \right)^{-1}$$

Demostración:

$$\begin{aligned} \pi(\mathbf{x}_0 | \mathbf{y}) &\propto \pi(\mathbf{y} | \mathbf{x}_0) \cdot \pi(\mathbf{x}_0) \propto \exp\{ -\frac{1}{2} [(y_0 - \mathbf{x}_0^{*t} \beta) / \sigma_0^2] \} \cdot \pi(\mathbf{x}_0) \\ &\propto \exp\{ -\frac{1}{2} [(y_0 - \beta_0 - \mathbf{x}_0^t \beta_1) / \sigma_0^2 + (\mathbf{x}_0 - \mathbf{m}_0)^t v_0^{-1} + (\mathbf{x}_0 - \mathbf{m}_0)^t] \} \\ &\propto \exp\{ -\frac{1}{2} [-2(y_0 - \beta_0) \mathbf{x}_0^t \beta_1 / \sigma_0^2 + \mathbf{x}_0^t (\beta_1 \beta_1^t) \mathbf{x}_0 / \sigma_0^2 + \mathbf{x}_0^t v_0^{-1} \mathbf{x}_0 - 2\mathbf{x}_0^t v_0^{-1} \mathbf{m}_0] \} \\ &\propto \exp\{ -\frac{1}{2} [-2 \mathbf{x}_0^t ((y_0 - \beta_0) \beta_1 / \sigma_0^2 + v_0^{-1} \mathbf{m}_0) + \mathbf{x}_0^t ((\beta_1 \beta_1^t) / \sigma_0^2 + v_0^{-1}) \mathbf{x}_0] \} \\ &\propto \exp\{ -\frac{1}{2} (\mathbf{x}_0 - \mathbf{m}_1)^t v_1^{-1} (\mathbf{x}_0 - \mathbf{m}_1) \}, \text{ lo que completa la demostración. } \square \end{aligned}$$

Ejemplo 2.1: En particular, si el modelo de regresión abordado es el lineal simple ($p = 2$), se obtiene el siguiente resultado:

$$\text{Si } y_i | x_0 \sim^{\text{ind.}} N(\beta_0 + \beta_1 \cdot x_i, \sigma_0^2), \text{ para } i = 0, \dots, n.$$

$$y \quad x_0 \sim N(m_0, v_0)$$

Entonces

$$x_0 | y \sim N(m_1, v_1) \text{ con } m_1 = v_1 \left(\frac{(y_0 - \beta_0)}{\sigma_0^2} \cdot \beta_1 + v_0^{-1} m_0 \right), v_1 = \left(\frac{\beta_1^2}{\sigma_0^2} + v_0^{-1} \right)^{-1}$$

Lema 2.2: Distribución condicional de la Componente de Dispersión para el caso dependiente discreto

$$\text{Si } y_i | \omega \sim^{\text{cond. ind.}} N(\mu_i, \phi^{-1} \omega)$$

$$\omega \sim f, \text{ con } f(x) = \sum_{j=1}^{k+1} p_j \cdot \delta_{a_j}(x) \text{ y con } \sum_{j=1}^{k+1} p_j = 1$$

Entonces

$$\omega | y \sim g, \text{ con } g(x) = \frac{\exp\left\{-\frac{\phi}{2x} \sum_{i=1}^n (y_i - \mu_i)^2\right\} \sum_{j=1}^{k+1} \delta_{a_j}(x) \cdot p_j}{x^{\frac{n}{2}} \cdot \sum_{j=1}^{k+1} \frac{\exp\left\{-\frac{\phi}{2 \cdot a_j} \sum_{i=1}^n (y_i - \mu_i)^2\right\}}{(a_j)^{\frac{n}{2}}} \cdot p_j} \text{ y con } x \in \{a_1, \dots, a_{k+1}\}$$

Ejemplo 2.2:

En este caso se aplicará el *Lema 2.2* al Modelo de Regresión Lineal homoscedástico, donde el parámetro de dispersión ω tiene una distribución discreta binaria, es decir los errores son mezcla escala de dos Normales.

$$\text{Si } y_i | \omega \sim^{\text{cond. ind.}} N(\mathbf{x}_i^t \boldsymbol{\beta}, \omega)$$

$$p(\omega = a_i) = p_i, \quad i = 1, 2, \text{ donde } p_1 + p_2 = 1,$$

entonces

$$\omega | y \sim g, \text{ donde } g(a_j) \propto \omega^{-\frac{n}{2}} \cdot p_j \cdot \exp\left\{-\frac{1}{2\omega} \sum_{i=1}^n (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2\right\}, \quad j \in \{1, 2\}.$$

2.3.2 MODELO DE REGRESIÓN NORMAL HETEROSCEDÁSTICO

Los siguientes cuatro lemas presentan, respectivamente, las distribuciones condicionales completas para los coeficientes de regresión β , parámetro de dispersión común σ^2 , parámetros de dispersión no comunes ω_i , parámetro de calibración \mathbf{x}_0 y probabilidades de la distribución discreta en el contexto del modelo de regresión con errores normales con varianzas diferentes.

Lema 2.3: Distribución Condicional de los Coeficientes de Regresión

Si $\mathbf{y} | \beta \sim N(\mathbf{X}\beta, \Sigma_0)$, Σ_0 conocida y $\beta \sim N(\mathbf{b}_0, \mathbf{B}_0)$,

entonces

$$\beta | \mathbf{y} \sim N(\mathbf{b}_1, \mathbf{B}_1),$$

donde $\mathbf{b}_1 = \mathbf{B}_1 (\mathbf{X}^t \Sigma_0^{-1} \mathbf{y} + \mathbf{B}_0^{-1} \mathbf{b}_0)$ y $\mathbf{B}_1 = (\mathbf{X}^t \Sigma_0^{-1} \mathbf{X} + \mathbf{B}_0^{-1})^{-1}$

Ejemplo 2.3. A continuación se aplica el *Lema 2.1* al caso de un modelo de regresión lineal simple heteroscedástico. Los errores se han asumido normales, así como los coeficientes del modelo, y las distribuciones a posteriori de dichos coeficientes resultan ser normales.

Si $y_i | \beta_0, \beta_1 \sim \text{ind. } N(\beta_0 + \beta_1 x_i, \phi^{-1} \omega_i)$, $\beta_0 \sim N(b_{00}, v_{00})$ y $\beta_1 \sim N(b_{10}, v_{10})$,

entonces

$$\beta_0 | \beta_1, \mathbf{y} \sim N(b_{01}, v_{01}),$$

$$\beta_1 | \beta_0, \mathbf{y} \sim N(b_{11}, v_{11}),$$

$$\text{con } v_{01} = \left(\phi \sum_{i=1}^n \frac{1}{\omega_i} + v_{00}^{-1} \right)^{-1}, v_{11} = \left(\phi \sum_{i=1}^n \frac{x_i^2}{\omega_i} + v_{10}^{-1} \right)^{-1}, b_{01} = v_{01} \cdot \left(\phi \sum_{i=1}^n \frac{y_i - \beta_1 \cdot x_i}{\omega_i} + b_{00} v_{00}^{-1} \right) \text{ y}$$

$$b_{11} = v_{11} \cdot \left(\phi \sum_{i=1}^n \frac{(y_i - \beta_0) \cdot x_i}{\omega_i} + b_{10} v_{10}^{-1} \right)$$

Lema 2.4: Distribución Condicional del Parámetro de Dispersión

Si $\mathbf{y} | \sigma^2 \sim N_n(\boldsymbol{\mu}_0, \sigma^2 \Sigma_0)$ y $\sigma^2 \sim IG(r_0, \lambda_0)$,

entonces

$$\sigma^2 | \mathbf{y} \sim IG(r_1, \lambda_1), \text{ con } r_1 = \frac{n}{2} + r_0 \text{ y } \lambda_1 = \frac{1}{2} (\mathbf{y} - \boldsymbol{\mu}_0)^t \Sigma_0^{-1} (\mathbf{y} - \boldsymbol{\mu}_0) + \lambda_0.$$

Ejemplo 2.4:

A continuación se aplica el *Lema 2.2* al modelo de regresión normal con errores independientes homoscedásticos.

Si $\mathbf{y} | \sigma^2 \sim N_n(\mathbf{X}^t \beta, \sigma^2 \mathbf{I})$ y $\sigma^2 \sim IG(r_0, \lambda_0)$,

entonces

$$\sigma^2 | \mathbf{y} \sim IG(r_1, \lambda_1), \text{ donde } r_1 = \frac{n}{2} + r_0 \text{ y } \lambda_1 = \frac{1}{2} (\mathbf{y} - \mathbf{X}^t \beta)^t (\mathbf{y} - \mathbf{X}^t \beta) + \lambda_0.$$

Lema 2.5: Distribución condicional de la Componente de Dispersión para el caso independiente discreto

Si $y_i | \omega_i \sim \text{cond. ind. } N(\mu_i, \phi^{-1} \omega_i)$

y $\omega_i \sim \text{i.i.d. } f$, con $f(x) = \sum_{j=1}^{k+1} p_j \cdot \delta_{a_j}(x)$ y con $\sum_{j=1}^{k+1} p_j = 1$,

entonces

$$\omega_i | y_i \sim \text{ind. } g, \text{ con } g(\omega_i) = \frac{p_j}{\sqrt{\omega_i}} \cdot \exp\left\{-\frac{\phi}{2x}(y_i - \beta_0 - \beta_1 \cdot x_i)^2\right\} \sum_{j=1}^{k+1} \delta_{a_j}(\omega_i)$$

y con $\omega_i \in \{a_1, \dots, a_{k+1}\}$

Lema 2.6: Distribución Condicional para el Parámetro de calibración x_0

Si $y_i | \mathbf{x}_0 \sim \text{ind. } N(\mathbf{x}_i^{*t} \boldsymbol{\beta}, \sigma^2_i)$ y $\mathbf{x}_0 \sim N_{p-1}(\mathbf{m}_0, v_0)$, con \mathbf{x}_i^* y $\boldsymbol{\beta}$ definidas como en el Lema 2.1.,

entonces

$$\mathbf{x}_0 | \mathbf{y} \sim N_{p-1}(\mathbf{m}_1, v_1), \text{ con } \mathbf{m}_1 \text{ y } v_1 \text{ definidas como en Lema 2.1.}$$

Demostración: Idéntica a Lema 2.1.

Lema 2.7: Distribución Condicional para las Probabilidades de la Distribución Discreta

Si $\omega_i | \mathbf{p} \sim \text{i.i.d. } f(x) = \sum_{j=1}^{k+1} p_j \cdot \delta_{a_j}(x)$ con $\sum_{j=1}^{k+1} p_j = 1$

y $\mathbf{p} \sim D(\boldsymbol{\alpha}_0)$,

entonces

$$\mathbf{p} | \boldsymbol{\omega} \sim D\left(\boldsymbol{\alpha}_0 + \sum_{i=1}^n \delta_{\omega_i}\right),$$

donde $\boldsymbol{\alpha}_0 + \sum_{i=1}^n \delta_{\omega_i}$ es una notación que alude al vector de componentes $\alpha_{0j} + \sum_{i=1}^n \delta_{\omega_i}$, donde $\boldsymbol{\alpha}_0$ es el

vector de componentes α_{0j} .

Observación:

Este vector $\boldsymbol{\alpha}_0$ puede, a su vez, tener una distribución a priori beta, por ejemplo.

El ejemplo que se presenta a continuación hace uso de todos los lemas anteriores y corresponde a un problema de calibración para una regresión lineal simple.

Ejemplo 2.5:

Si $y_i | \phi, \omega_i, \beta, \mathbf{p}, \mathbf{x}_0 \sim N(\beta_0 + \beta_1 \cdot x_i; \phi^{-1} \omega_i)$

$$\omega_i | \mathbf{p} \sim f(\omega_i) = \sum_{j=1}^{k+1} p_j \cdot \delta_{a_j}(\omega_i) \text{ con } \sum_{j=1}^{k+1} p_j = 1$$

$$\phi \sim \text{Gamma}(r_0, \lambda_0),$$

$$\beta_r \sim N(b_r, v_r), r = 0, 1,$$

$$\mathbf{p} \sim D(\boldsymbol{\alpha}_0)$$

y $\mathbf{x}_0 \sim N(\mathbf{m}_0, d_0),$

entonces

$$(2.6) \quad \omega_i | y_i, \phi, \beta, \mathbf{p}, \mathbf{x}_0 \sim^{\text{cond. ind.}} \mathbf{g}(\omega_i) = \frac{p_j}{\sqrt{\omega_i}} \cdot \exp\left\{-\frac{\phi}{2x} (y_i - \beta_0 - \beta_1 \cdot x_i)^2\right\} \sum_{j=1}^{k+1} \delta_{a_j}(\omega_i),$$

donde $\omega_i \in \{a_1, \dots, a_{k+1}\}$, desde el Lema 2.2.

$$(2.7) \quad \phi | \mathbf{y}, \omega, \beta, \mathbf{p}, \mathbf{x}_0 \sim \text{Gamma}(r_1, \lambda_1),$$

donde $r_1 = \frac{n}{2} + r_0$ y $\lambda_1 = \lambda_0 + \frac{1}{2} \sum_{i=0}^n (y_i - \beta_0 - \beta_1 \cdot x_i)^2$, desde Lema 2.4.

$$(2.8) \quad \beta_0 | \mathbf{y}, \omega, \phi, \beta_1, \mathbf{p}, \mathbf{x}_0 \sim N(b_0^*, v_0^*),$$

con $v_0^* = \left(\phi \sum_{i=1}^n \frac{1}{\omega_i} + v_0^{-1}\right)^{-1}$ y $b_0^* = v_0^* \cdot \left(\phi \sum_{i=1}^n \frac{y_i - \beta_1 \cdot x_i}{\omega_i} + b_0 v_0^{-1}\right)$, desde Ejemplo 2.3.

$$(2.9) \quad \beta_1 | \mathbf{y}, \omega, \phi, \beta_0, \mathbf{p}, \mathbf{x}_0 \sim N(b_1^*, v_1^*),$$

con $v_1^* = \left(\phi \sum_{i=1}^n \frac{x_i^2}{\omega_i} + v_{10}^{-1}\right)^{-1}$ y $b_1^* = v_1^* \cdot \left(\phi \sum_{i=1}^n \frac{(y_i - \beta_0) \cdot x_i}{\omega_i} + b_1 v_1^{-1}\right)$, desde Ejemplo 2.3.

$$(2.10) \quad \mathbf{p} | \mathbf{y}, \omega, \phi, \beta, \mathbf{x}_0 \sim D\left(\boldsymbol{\alpha}_0 + \sum_{i=1}^n \delta_{\omega_i}\right), \text{ desde el Lema 2.6.}$$

$$(2.11) \quad \mathbf{x}_0 | \mathbf{y}, \omega, \phi, \beta, \mathbf{p} \sim N(\mathbf{m}_1, d_1)$$

con $\mathbf{m}_1 = d_1 \cdot \left(\frac{(y_0 - \beta_0) \cdot \beta_1}{\sigma_0^2} + d_0^{-1} \mathbf{m}_0\right)$, $d_1 = \left(\frac{\beta_1 \cdot \beta_1}{\sigma_0^2} + d_0^{-1}\right)^{-1}$, desde el Lema 2.6.

2.3.3 MODELO DE REGRESIÓN CON ERRORES EN LAS VARIABLES

El resultado que se presenta a continuación será utilizado para abordar el problema de errores de medición, ya que permitirá obtener el estimador de Bayes de x_i (la medición exacta no observable) para dicho modelo.

Lema 2.8: Distribución Condicional de la Medición Exacta

$$\begin{aligned} \text{Si } y_i | x_i &\sim^{\text{cond. ind.}} N(\beta_0 + \beta_1 x_i; \phi_\varepsilon^{-1} \omega_{\varepsilon i}) \\ z_i | x_i &\sim^{\text{cond. ind.}} N(x_i; \phi_U^{-1} \omega_{U i}) \\ x_i &\sim^{\text{i.i.d.}} N(\mu_0; \sigma_0^2), \end{aligned}$$

entonces

$$\begin{aligned} x_i | \mathbf{y}, \mathbf{z} &\sim N(\mu_{1i}; \sigma_{1i}^2), \\ \text{con } \sigma_{1i}^2 &= \left[\frac{\phi_\varepsilon \cdot \beta_1^2}{\omega_{\varepsilon i}} + \frac{\phi_u}{\omega_{u i}} + \frac{1}{\sigma_0^2} \right]^{-1} \text{ y } \mu_{1i} = \sigma_{1i}^2 \cdot \left[\frac{\phi_\varepsilon \cdot \beta_1 (y_i - \beta_0)}{\omega_{\varepsilon i}} + \frac{\phi_u \cdot z_i}{\omega_{u i}} + \frac{\mu_0}{\sigma_0^2} \right]. \end{aligned}$$

2.3.4 MODELO DE REGRESIÓN DISCRETA

El siguiente lema presenta todas las distribuciones a posteriori para el modelo de calibración discreto, y debe a Basu y Mukhopadhyay (2000).

Lema 2.8: Distribuciones Condicionales Completas para Modelo paramétrico de Calibración Binaria con variable subyacente mezcla escala discreta de normales

$$\begin{aligned} \text{Si } y_i | z_i, \boldsymbol{\beta}, \omega_i, \mathbf{p}, \mathbf{G} &\sim^{\text{ind.}} \text{Bernoulli}(\theta_i = P(z_i > 0 | \omega_i, \boldsymbol{\beta}, \mathbf{p})), \\ z_i | \boldsymbol{\beta}, \omega_i, \mathbf{p}, \mathbf{G} &\sim^{\text{ind.}} N(\beta_0 + \beta_1; \omega_i), \\ \boldsymbol{\beta} &\sim \pi_1(\boldsymbol{\beta}), \\ \omega_i | \mathbf{p} &\sim^{\text{ind.}} G(\omega_i) = \sum_{j=1}^{k+1} p_j \cdot \delta_{a_j}(\omega_i) \end{aligned}$$

$$\text{y } \mathbf{p} = (p_1, \dots, p_k) \sim D(\alpha_1, \dots, \alpha_{k+1}),$$

$$\text{con } \boldsymbol{\beta} \perp \omega | \mathbf{p}; \boldsymbol{\beta} \perp \mathbf{p},$$

entonces las distribuciones condicionales completas resultantes son las siguientes:

$$(2.12) \quad z_i | y_i, \omega_i, \mathbf{p}, \boldsymbol{\beta} \sim \text{NT}(\mathbb{R}^+; \beta_0 + \beta_1 \cdot x_i; \omega_i) \cdot I(y_i = 1) + \text{NT}(\mathbb{R}^-; \beta_0 + \beta_1 \cdot x_i; \omega_i) \cdot I(y_i = 0)$$

donde $\text{NT}(A; \mu; \sigma^2)$ corresponde a la distribución normal de parámetros μ y σ^2 , truncada sobre el conjunto A .

$$(2.13) \quad \omega_i | y_i, z_i, \mathbf{p}, \boldsymbol{\beta} \sim \sum_{j=1}^{k+1} q_{ij} \cdot \delta_{a_j}, \text{ con } q_{ij} = \frac{\phi\left(\frac{z_i - \beta_0 - \beta_1 \cdot x_i}{\tau_j}\right) \cdot \frac{p_j}{\tau_j}}{\sum_{k=1}^s \phi\left(\frac{z_i - \beta_0 - \beta_1 \cdot x_i}{\tau_k}\right) \cdot \frac{p_k}{\tau_k}}$$

$$(2.14) \quad \mathbf{p} | y_i, z_i, \omega_i, \boldsymbol{\beta} \sim D(\boldsymbol{\alpha}^*), \text{ con } \boldsymbol{\alpha}^* = (\alpha_1^*, \dots, \alpha_{k+1}^*) \text{ y } \alpha_j^* = \alpha_j + \sum_{i=0}^n \delta_{\omega_i}(a_j)$$

$$(2.15) \quad \beta_0 | y_i, z_i, \omega_i, \beta_1, \mathbf{p} \sim N(m_0; S_0^2),$$

$$\text{con } m_0 = S_0^2 \left[\phi \sum_{i=0}^n \frac{z_i - \beta_1 \cdot x_i}{\omega_i} + \frac{b_0}{v_0} \right] \text{ y } S_0^2 = \left[\phi \sum_{i=0}^n \frac{1}{\omega_i} + \frac{1}{v_0} \right]^{-1}$$

$$(2.16) \quad \beta_1 | y_i, z_i, \omega_i, \beta_0, \mathbf{p} \sim N(m_1; S_1^2),$$

$$\text{con } m_1 = S_1^2 \left[\phi \sum_{i=0}^n \frac{(z_i - \beta_0) \cdot x_i}{\omega_i} + \frac{b_1}{v_1} \right] \text{ y } S_1^2 = \left[\phi \sum_{i=0}^n \frac{x_i^2}{\omega_i} + \frac{1}{v_1} \right]^{-1}$$

Lema 2.9: Distribuciones Condicionales Completas para Modelo semi paramétrico de Calibración Binario con variable subyacente mezcla escala continua de normales

$$\text{Si } y_i | z_i, \boldsymbol{\beta}, \omega_i, \mathbf{G} \sim \text{ind. Bernoulli}(\theta_i = P(z_i > 0 | \boldsymbol{\beta}, \omega_i, \mathbf{G})),$$

$$z_i | \boldsymbol{\beta}, \omega_i \sim \text{ind. } N(\beta_0 + \beta_1; \omega_i),$$

$$\boldsymbol{\beta} \sim \pi_1(\boldsymbol{\beta}),$$

$$\omega_i | \mathbf{G} \sim \text{i.i.d. } \mathbf{G}$$

$$\text{y } \mathbf{G} \sim D(\text{c } \mathbf{G}_0),$$

$$\text{con } \boldsymbol{\beta} \perp \omega | \mathbf{G}; \boldsymbol{\beta} \perp \mathbf{G},$$

entonces las distribuciones condicionales completas resultantes coinciden con las del Lema 2.8. para \mathbf{z} y $\boldsymbol{\beta}$, y se agrega la siguiente distribución condicional completa para ω :

$$(2.17) \quad \omega_i | z_i, \boldsymbol{\beta}, y_i \sim \omega_i | \mathbf{z}, \boldsymbol{\beta}, x_0, \{\omega_k, k \neq i\}, \mathbf{G}, \mathbf{y} \sim q_{0i} \cdot dG_b^{(i)}(\omega_i) + \sum_{k \neq i} q_{ki} \cdot \delta_{\omega_k}(\omega_i)$$

$$\text{con } q_{0i} = f_{T_i}(z_i), \text{ donde } T_i \sim T(\mu = E(z_i) = \beta_0 + \beta_1 x_i, \sigma^2 = 2\lambda_1, \gamma = 2r_1),$$

$$\text{con } q_{ki} = \Phi_Z\left(\frac{z_i - \beta_0 - \beta_1 \cdot x_i}{\phi^{-1} \cdot \omega_k}\right)$$

y con $G_b^{(i)}$ la función de distribución acumulada Gamma Inversa de parámetros de posición $r = r_1 + 1/2$ y de escala $\lambda = \lambda_1 + (z_i - \beta_0 - \beta_1 x_i)^2 / 2$.

CAPÍTULO 3:

SOLUCIÓN SEMIPARAMÉTRICA AL PROBLEMA DE CALIBRACIÓN LINEAL CONTINUA BAJO DISTRIBUCIONES ELÍPTICAS REPRESENTABLES

Originalmente, el problema de calibración lineal desde una perspectiva bayesiana fue planteado por Hoadley (1970) y Hunter y Lamboy (1981) en el contexto de modelos normales. Extensiones al modelo de regresión con errores *t* – student y a los modelos de regresión elípticos, han sido consideradas por Branco, Bolfarine e Iglesias (1998). El presente trabajo, aborda el caso semiparamétrico y considera parámetros de dispersión con priori *Dirichlet*.

El hecho de considerar errores elípticos le da una mayor flexibilidad al modelo, permitiendo agrupar una gran familia de distribuciones.

En este capítulo abordamos el problema de calibración en el contexto de modelos de regresión elíptico, bajo una óptica bayesiana semiparamétrica. Específicamente, nuestro enfoque considera la incorporación de una priori para la distribución del parámetro de dispersión de acuerdo con un proceso de *Dirichlet*.

El modelo de calibración lineal continua, que se aborda en este capítulo, corresponde a:

$$y_i = \mathbf{x}_i^t \boldsymbol{\beta} + \varepsilon_i, i = 0, \dots, n;$$

donde $\boldsymbol{\beta}$ es el parámetro de los coeficientes del modelo de regresión lineal que relaciona a la variable respuesta y_i y las predictoras en el vector \mathbf{x}_i . Aquí y_i representa a las observaciones obtenidas del instrumento económico e inexacto, mientras que \mathbf{x}_i representa a las observaciones en el instrumento exacto pero costoso.

Se cuenta con n conjuntos de observaciones (y_i, \mathbf{x}_i) , las cuales permiten calibrar el instrumento. Luego se toma una nueva observación y_0 , pero \mathbf{x}_0 no es observable, y corresponde al parámetro de interés.

El problema de calibración es citado muchas veces en la literatura como el problema de regresión o predicción inversa, ya que se trata precisamente de recuperar la información que en una primera etapa era considerada conocida a partir de los datos observados y que son función de ella. Es por esta razón, que precisamente resulta más lógico utilizar el enfoque bayesiano para su análisis.

Existen dos tipos de calibración: aquella denominada calibración natural y que considera los \mathbf{x}_i como realizaciones de una variable aleatoria, y la denominada calibración controlada, que se vincula a nuestro problema y considera a los \mathbf{x}_i como no aleatorios y elegidos por el experimentador.

Desde el punto de vista clásico no ha habido una respuesta satisfactoria a este problema. En este sentido podemos mencionar los métodos tradicionales existentes en la literatura y que se relacionan con el enfoque clásico:

Estimador clásico (Eisenhart, 1939)

$\hat{\mathbf{x}}_0 = \frac{y_0 - \hat{\beta}_0}{\hat{\beta}_1}$, donde $\hat{\beta}_0$ y $\hat{\beta}_1$ son, respectivamente, los estimadores de mínimos cuadrados de los

coeficientes del modelo de regresión lineal simple $y_i = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i$, basados en los pares de observaciones (x_i, y_i) , y donde $\varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2)$, con $i = 1, \dots, n$.

Estimador inverso (Krutchkoff, 1967)

$\hat{\mathbf{x}}_0 = \hat{\delta}_0 + \hat{\delta}_1 \cdot y_0$, donde $\hat{\delta}_0$ y $\hat{\delta}_1$ son, respectivamente, los estimadores de mínimos cuadrados de los coeficientes del modelo de regresión lineal simple $x_i = \delta_0 + \delta_1 \cdot y_i + \xi_i$, basados en los pares de observaciones (y_i, x_i) , donde $y_i = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i$ y $\varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2)$, con $i = 1, \dots, n$.

Es claro que la utilización de estos dos estimadores clásicos se limita al caso del modelo de regresión simple, lo cual es una gran deficiencia de índole práctica.

Por otra parte, aunque el estimador clásico es un estimador máximo verosímil, no es insesgado ni tiene varianza mínima, al contrario, la varianza puede ser infinita por tratarse del cociente de dos variables aleatorias normales independientes, lo cual lleva, en ocasiones (por ejemplo para β_1 no significativamente distinto de cero) a intervalos confidenciales ilimitados y/o inconexos (Brown, 1993).

La situación del estimador inverso, sin embargo, no es mejor, ya que, aunque tiene varianza finita, asume incorrectamente que los ξ_i y los y_i son independientes, lo cual claramente no es efectivo, ya que la relación entre el modelo de regresión simple usual y el modelo inverso requieren que $\delta_0 = -\beta_0/\beta_1$, $\delta_1 = \beta_1^{-1}$ y $\xi_i = -\varepsilon_i/\beta_1$, de donde $\text{Cov}(\xi_i, y_i) = -\sigma^2/\beta_1$.

En el artículo de Osborne (1991) se encuentra una comparación más detallada de las propiedades de estos dos estimadores.

Naturalmente el estimador inverso no se justifica desde el punto de vista clásico para la calibración controlada, ya que en ese caso los x_i son fijos, y esto hace que resulte muy interesante descubrir que, para cierta bien escogida priori, el estimador de Bayes para \mathbf{x}_0 es, efectivamente, el estimador inverso.

Desde el punto de vista bayesiano, un primer trabajo en esta área se debe a Hoadley (1970), quien aborda el problema de calibración normal paramétrico para el modelo de regresión lineal simple. Este autor utiliza

prioris impropias para β y ϕ , y una priori t – student de parámetros $\mu = \bar{x}$, $\sigma^2 = \frac{(n^2 - 1)}{n(n - 3)} S_x^2$ y $\gamma = n - 3$,

para \mathbf{x}_0 , recuperando así el estimador inverso como un estimador de Bayes.

Algunas extensiones a los resultados de Hoadley (1970) halladas en la literatura corresponden a calibración multivariada y calibración polinomial (Brown, 1982).

La aparición de los procedimientos MCMC en la década de los 90, sumado al desarrollo tecnológico – computacional, hacen que el desarrollo de los resultados en esta área se expandan.

Es así como, dentro de la línea univariada, pero extendiéndose al caso de calibración en regresión múltiple, Branco, Bolfarine e Iglesias (1998) estudian el modelo de calibración lineal, considerando errores t – student, lo cual permite un mejor análisis en presencia de outliers. Estos autores consideran una priori normal para β y \mathbf{x}_0 , una priori Chi Cuadrado para ϕ y prioris Chi Cuadrado Inversa independientes para los ω_i que vinculan el modelo t – student y el modelo normal. Posteriormente, Branco, Bolfarine, Iglesias y Arellano – Valle (2000) extienden estos resultados al modelo con errores elípticos, siempre desde el enfoque paramétrico. Nuestro trabajo generaliza este análisis desde el punto de vista semi paramétrico.

En otra dirección, Eno (1999) en su tesis doctoral presenta el análisis paramétrico del modelo con errores normales utilizando prioris de referencia para resolver el problema de calibración polinomial y multivariado. Este autor se concentra en la derivación de prioris de referencia para el parámetro \mathbf{x}_0 y muestra la imposibilidad de determinar explícitamente la distribución a posteriori de cada uno de los parámetros, por lo cual sugiere la utilización de muestreo de Gibbs para la determinación de los estimadores de Bayes. El caso de calibración multivariado también es abordado por Sundberg (1999).

Otros aportes recientes se deben a Salini, Tiano y Zirilli (2002), y Salini, Zirilli y Tiano (2002), quienes abordan el problema de calibración mediante la teoría de filtros de Kalman desde el punto de vista bayesiano, para el modelo de regresión simple y múltiple, respectivamente.

El presente capítulo se divide en tres secciones. En las primeras dos secciones se desarrolla una metodología bayesiana semiparamétrica para el problema de calibración lineal continua para los casos en que los errores tienen una distribución elíptica representable independiente y dependiente respectivamente. Como subproducto se obtiene una metodología bayesiana semiparamétrica para el análisis de regresión bajo los supuestos antes mencionados. Para ambos casos se requiere la implementación de métodos de simulación MCMC. En este trabajo se utiliza específicamente el muestreo de Gibbs, explotando la relación entre el modelo elíptico representable y el modelo normal heteroscedástico. Para ambas secciones, con el fin de ilustrar, se comienza abordando un caso paramétrico que extiende los resultados hallados en la literatura, y que consiste en considerar que los errores son mezcla escala finita de normales o lo que se conoce como normal contaminada. El vector de los pesos en esta mezcla sigue una distribución Dirichlet, como una forma de robustificar el modelo normal. La extensión natural de este caso, corresponde al caso semiparamétrico donde no nos amarramos a una mezcla determinada, sino que los parámetros que definen la mezcla siguen un proceso Dirichlet. Todos los casos abordados utilizan prioris usuales en la literatura, las cuales permiten comparación de resultados.

Finalmente, en la tercera sección, los resultados son ilustrados con datos utilizados previamente en la literatura y con simulación.

3.1 SOLUCIÓN BAYESIANA SEMIPARAMÉTRICA BAJO EL MODELO DE REGRESIÓN ELÍPTICO REPRESENTABLE INDEPENDIENTE

Para empezar, en esta sección abordamos el problema de calibración bajo el modelo de regresión, considerando que los términos de error son independientes e idénticamente distribuidos elípticos, con función generadora de densidad h representable; el modelo viene, entonces, dado por:

$$(3.1) \quad y_i | \boldsymbol{\beta}, \mathbf{x}_0, \phi \sim \text{ind. } El_1(\mathbf{x}_i^t \boldsymbol{\beta}, \phi^{-1}, h) \text{ y } h(u) = \int_0^{\infty} (2\pi\omega)^{-1/2} e^{-u/2\omega} dG(\omega),$$

con $h \geq 0$ y donde G es función de distribución acumulada tal que $G(0) = 0$.

Ahora bien, la función G es denominada función mezcla o mezcladora. Los parámetros $(\phi, \boldsymbol{\beta}, \mathbf{x}_0)$ y G son desconocidos y el hecho de que este último sea infinito dimensional es lo que imprime el carácter semiparamétrico a este enfoque, y es, también, lo que distingue los resultados que se presentan a continuación de los presentados en la literatura.

Por la *Proposición 2.1b* (sección 2.2 del Capítulo 2), el problema puede ser abordado considerando un modelo jerárquico en un número mayor de etapas:

$$(3.2) \quad \begin{aligned} y_i | \boldsymbol{\beta}, \mathbf{x}_0, \phi, \omega_i &\sim \text{i.i.d. } N(\mathbf{x}_i^t \boldsymbol{\beta}, \phi^{-1} \omega_i) \\ \omega_i | G &\sim \text{c.i.i.d. } G \\ (G, \boldsymbol{\omega}) &\perp (\boldsymbol{\beta}, \mathbf{x}_0, \phi). \end{aligned}$$

El modelo que nos interesa une (3.1) o equivalentemente (3.2) a las siguientes etapas:

$$(3.3) \quad \begin{aligned} \boldsymbol{\beta} &\sim N_p(\mathbf{b}, \mathbf{V}) \\ \mathbf{x}_0 &\sim N_{p-1}(\mathbf{c}_0, \mathbf{D}_0) \\ \phi &\sim \text{Gama}(\tau_0, \lambda_0). \end{aligned}$$

Es importante destacar que la especificación (3.1) no permite encontrar en forma explícita las distribuciones condicionales completas, mientras que la especificación (3.2) si lo permite, ya que la verosimilitud es una productoria de normales.

Esto facilita las expresiones y admite, utilizando para ello el lema de Blackwell y McQueen (1973) y los resultados de Escobar y West (1998), la utilización de muestreo de Gibbs para hallar los estimadores de Bayes de los parámetros.

La distribución a posteriori para el parámetro \mathbf{x}_0 por supuesto es la misma, pero la expresión respectiva es analíticamente imposible de manejar.

En efecto, la función de verosimilitud asociada al modelo (3.1), es la siguiente:

$$L(\boldsymbol{\beta}, \mathbf{x}_0, \phi | \mathbf{y}, \mathbf{G}) = (\phi / 2\pi)^{\frac{1}{2}(n+1)} \left[\prod_{i=0}^n \int_0^{\infty} \omega_i^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi \cdot (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / \omega_i) dG(\omega_i) \right],$$

mientras que para el modelo (3.2) se obtiene una función de verosimilitud aumentada que coincide con la verosimilitud de un modelo de regresión normal heteroscedástico:

$$L(\boldsymbol{\beta}, \mathbf{x}_0, \phi, \boldsymbol{\omega} | \mathbf{y}, \mathbf{G}) = (\phi / 2\pi)^{\frac{1}{2}(n+1)} \left[\prod_{i=0}^n \omega_i^{-\frac{1}{2}} \right] \cdot \exp[-\frac{1}{2} \phi \cdot \sum_{i=0}^n \omega_i^{-1} (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2].$$

Desde el punto de vista clásico resulta infactible realizar la estimación máximo verosímil con esta última expresión sin tener información acerca de \mathbf{G} , y para la primera especificación, incluso para alguna \mathbf{G} particular, la expresión es muy complicada de resolver, pues requiere que las derivadas respecto de $\boldsymbol{\beta}$, \mathbf{x}_0 y ϕ de la verosimilitud o de la log verosimilitud sean nulos. Para la expresión $L(\boldsymbol{\beta}, \mathbf{x}_0, \phi | \mathbf{y}, \mathbf{G})$ las derivadas son complicadísimas y en el segundo caso, para $L(\boldsymbol{\beta}, \mathbf{x}_0, \phi, \boldsymbol{\omega} | \mathbf{y}, \mathbf{G})$, se requiere que \mathbf{G} sea conocida y que las variables instrumentales ω_i sean estimadas.

En efecto, para la primera expresión, suponiendo que \mathbf{G} tiene asociada una función de densidad g y que la naturaleza de esa función permite intercambiar la integral respecto de ω_i con las derivadas respecto de $\boldsymbol{\beta}$, \mathbf{x}_0 y ϕ , y definiendo la función:

$$k(\mathbf{e}_i, \phi) = \left[\int_0^{\infty} \omega_i^{-\frac{3}{2}} \exp(-\frac{1}{2} \phi \cdot \mathbf{e}_i^2 / \omega_i) \cdot g(\omega_i) \cdot d\omega_i \right] / \left[\int_0^{\infty} \omega_i^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi \cdot \mathbf{e}_i^2 / \omega_i) \cdot g(\omega_i) \cdot d\omega_i \right],$$

con $\mathbf{e}_i = (y_i - \mathbf{x}_i^t \boldsymbol{\beta})$, los estimadores máximo verosímiles deben satisfacer que:

$$\frac{\phi}{n+1} \cdot \sum_{i=0}^n \mathbf{e}_i^2 \cdot k(\mathbf{e}_i, \phi) = 1; \quad \sum_{i=0}^n \mathbf{e}_i \cdot k(\mathbf{e}_i, \phi) = 0; \quad \sum_{i=0}^n \mathbf{e}_i \cdot \mathbf{x}_i \cdot k(\mathbf{e}_i, \phi) = 0, \text{ y } \mathbf{e}_0 = 0 \text{ ó } k(\mathbf{e}_0, \phi) = 0.$$

Ahora bien, para la segunda expresión $L(\boldsymbol{\beta}, \mathbf{x}_0, \phi, \boldsymbol{\omega} | \mathbf{y}, \mathbf{G})$, función de verosimilitud aumentada, bajo los supuestos mencionados anteriormente, los estimadores máximo verosímiles deben satisfacer que:

$$\frac{\phi}{n+1} \cdot \sum_{i=0}^n \frac{\mathbf{e}_i^2}{\omega_i} = 1; \quad \sum_{i=0}^n \frac{\mathbf{e}_i}{\omega_i} = 0; \quad \sum_{i=0}^n \frac{\mathbf{x}_i \cdot \mathbf{e}_i}{\omega_i} = 0; \quad \mathbf{e}_0 = 0; \quad \omega_i = \phi \cdot \mathbf{e}_i^2 \text{ y } \omega_i \neq 0 \quad \forall i = 1, \dots, n.$$

Esto lleva a:

$$\hat{e}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 \cdot x_i \text{ y } \hat{\phi}^{-1} \cdot \hat{\omega}_i = \hat{e}_i^2 \quad \forall i = 1, \dots, n;$$

de donde se obtiene que los estimadores máximo verosímiles de β_0 , β_1 y x_0 son solución del sistema:

$$\sum_{i=0}^n \hat{\omega}_i^{-1} \cdot (y_i - \hat{\beta}_0 - \hat{\beta}_1 \cdot x_i) = \sum_{i=0}^n \hat{\omega}_i^{-1} \cdot x_i \cdot (y_i - \hat{\beta}_0 - \hat{\beta}_1 \cdot x_i) = 0; \hat{\omega}_i = \phi \cdot (y_i - \hat{\beta}_0 - \hat{\beta}_1 \cdot x_i)^2; y_0 = \hat{\beta}_0 + \hat{\beta}_1 \cdot x_0$$

Esto deja en evidencia que no existe solución única para ϕ , puesto que la base para el conjunto de funciones estimables está dado por $\{\phi^{-1} \cdot \omega_i, i = 1, \dots, n\}$, y de ahí se desprenden las siguientes ecuaciones, que revelan claramente que la solución requiere métodos numéricos:

$$\sum_{i=1}^n \frac{1}{y_i - \hat{\beta}_0 - \hat{\beta}_1 \cdot x_i} = 0; \sum_{i=1}^n \frac{x_i}{y_i - \hat{\beta}_0 - \hat{\beta}_1 \cdot x_i} = 0, \text{ e } y_0 = \hat{\beta}_0 - \hat{\beta}_1 \cdot \hat{x}_0.$$

Por otra parte, para una distribución G particular bien escogida como Gama Inversa, la estimación máximo verosímil requiere hallar los valores de los parámetros que satisfagan las siguientes ecuaciones:

$$\frac{2r+1}{n+1} \sum_{i=0}^n \left(1 + \frac{2\lambda}{\phi \cdot e_i^2}\right)^{-1} = 1; \sum_{i=0}^n \frac{1}{e_i} \left(1 + \frac{2\lambda}{\phi \cdot e_i^2}\right)^{-1} = 0; \sum_{i=0}^n \frac{x_i}{e_i} \left(1 + \frac{2\lambda}{\phi \cdot e_i^2}\right)^{-1} = 0, \text{ y } e_0 = 0.$$

La última expresión surge de $\frac{\beta_1}{e_0} \left(1 + \frac{2\lambda}{\phi \cdot e_0^2}\right)^{-1} = 0$, donde en general $\beta_1 \neq 0$ y donde $e_i = (y_i - \mathbf{x}_i^t \boldsymbol{\beta})$, que en

el caso múltiple corresponde a $e_i = (y_i - \beta_0 - \beta_1 \cdot x_{i1} - \dots - \beta_{p-1} \cdot x_{i,p-1})$ y en el simple a $e_i = (y_i - \beta_0 - \beta_1 \cdot x_i)$.

En efecto, la distribución a posteriori asociada al modelo de interés bajo la especificación (3.1), aún considerando G conocida (en caso contrario las expresiones son mucho más complicadas), es de la forma siguiente:

$$\pi(\boldsymbol{\beta}, \mathbf{x}_0, \phi | \mathbf{y}, G) \propto L(\boldsymbol{\beta}, \mathbf{x}_0, \phi | \mathbf{y}, G) \cdot \pi(\boldsymbol{\beta}) \cdot \pi(\mathbf{x}_0) \cdot \pi(\phi),$$

es decir, la distribución a posteriori para los parámetros es:

$$\pi(\boldsymbol{\beta}, \mathbf{x}_0, \phi | \mathbf{y}, G) \propto (\phi)^{r_0 - \frac{1}{2}(n+3)} \cdot \exp\left(-\frac{1}{2} [(\boldsymbol{\beta}-\mathbf{b})^t \mathbf{V}^{-1}(\boldsymbol{\beta}-\mathbf{b}) + (\mathbf{x}_0-\mathbf{c}_0)^t \mathbf{D}_0^{-1}(\mathbf{x}_0-\mathbf{c}_0)] - \lambda_0 \phi\right) \cdot \prod_{i=0}^n \left[\int_0^{\infty} (\omega_i)^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi \cdot (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / \omega_i) dG(\omega_i) \right],$$

donde la constante de proporcionalidad es prácticamente imposible de evaluar explícitamente (esto se puede ver en el artículo de Branco, Bolfarine e Iglesias, 1998).

Como se mencionó anteriormente, la distribución a priori para x_0 es normal $(p-1)$ variada y de parámetros \mathbf{c}_0 y \mathbf{D}_0 . Marginalizando en la distribución a posteriori para los parámetros, la distribución a posteriori del parámetro de interés x_0 tiene la siguiente forma:

$$\pi(\mathbf{x}_0 | \mathbf{y}) \propto \exp\left[-\frac{1}{2} (\mathbf{x}_0 - \mathbf{c}_0)^t \mathbf{D}_0^{-1} (\mathbf{x}_0 - \mathbf{c}_0)\right] \cdot \iiint_{\mathbb{R}^+ \times \mathbb{R}^2} (\phi)^{r_0 - \frac{1}{2}(n+3)} \cdot \exp\left(-\frac{1}{2} [(\boldsymbol{\beta}-\mathbf{b})^t \mathbf{V}^{-1}(\boldsymbol{\beta}-\mathbf{b}) + 2 \lambda_0 \phi]\right) \cdot \prod_{i=0}^n \left[\int_0^{\infty} (\omega_i)^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi \cdot (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / \omega_i) dG(\omega_i) \right] d\phi d\boldsymbol{\beta}$$

la cual depende del proceso G y de su distribución a priori.

Incluso si G fuera la distribución Gama Inversa de parámetros r y λ , esta expresión resultaría igualmente complicada, ya que la expresión para la productoria de integrales respecto de ω es proporcional a:

$$\prod_{i=0}^n \left[\int_0^{\infty} (\omega_i)^{-1/2} \exp(-1/2 \phi \cdot (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / \omega_i) dG(\omega_i) \right] \propto \prod_{i=0}^n [1 + \phi \cdot (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / 2 \lambda]^{-1/2(2r+1)},$$

de donde la expresión resultante resulta inmanejable y el estimador de Bayes (la media de dicha distribución) no es obtenible analíticamente.

Por otro lado, para la especificación (3.2), la posteriori presenta una expresión semejante a la anterior y que conduce a los mismos hallazgos:

$$\begin{aligned} \pi(\boldsymbol{\beta}, \mathbf{x}_0, \phi, \boldsymbol{\omega} | \mathbf{y}) &\propto L(\boldsymbol{\beta}, \mathbf{x}_0, \phi, \boldsymbol{\omega} | \mathbf{y}, G) \cdot \pi(\boldsymbol{\beta}) \cdot \pi(\mathbf{x}_0) \cdot \pi(\phi) \cdot \pi(\boldsymbol{\omega} | G), \\ &\propto (\phi)^{r_0 - 1/2(n+3)} \cdot \exp(-1/2 [(\boldsymbol{\beta}-\mathbf{b})^t \mathbf{V}^{-1}(\boldsymbol{\beta}-\mathbf{b}) + (\mathbf{x}_0-\mathbf{c}_0)^t \mathbf{D}_0^{-1}(\mathbf{x}_0-\mathbf{c}_0)] - \lambda_0 \phi) \\ &\quad \cdot \left[\prod_{i=0}^n (\omega_i)^{-1/2} \cdot dG(\omega_i) \right] \cdot \exp[-1/2 \phi \cdot \sum_{i=0}^n \omega_i^{-1} (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2] \end{aligned}$$

Debido a lo expuesto anteriormente, hemos optado por la alternativa de utilizar MCMC para implementar el muestreo de Gibbs, ya que las distribuciones condicionales completas requeridas para dicha simulación se pueden obtener analíticamente, como se verá a continuación.

Ahora bien, las dos proposiciones siguientes abordan, el modelo que une (3.1) ó (3.2) a (3.3), considerando dos elecciones de la distribución a priori para la función de mezcla G . En ambos casos la distribución a posteriori para \mathbf{x}_0 es evidentemente muy difícil de tratar analíticamente, mientras que las distribuciones condicionales completas para todos los parámetros son simples, lo cual justifica considerar la implementación de muestreo de Gibbs.

Proposición 3.1

Bajo las especificaciones (3.1), (3.3), se escoge para los parámetros de dispersión una distribución a priori discreta $p(\omega_i | G) = \sum_{j=1}^{k+1} p_j \cdot \delta_{a_j}(\omega_i)$, cuyos pesos, a su vez, tienen una distribución a priori Dirichlet

$$(p_1, \dots, p_k) \sim \mathcal{D}(\alpha_1, \dots, \alpha_k).$$

En este caso la distribución a posteriori de \mathbf{x}_0 para \mathbf{p} conocido tiene la forma siguiente:

$$\begin{aligned} \pi(\mathbf{x}_0 | \mathbf{y}) &\propto \exp[-1/2 (\mathbf{x}_0 - \mathbf{c}_0)^t \mathbf{D}_0^{-1} (\mathbf{x}_0 - \mathbf{c}_0)] \cdot \\ &\quad \iiint_{\mathbb{R}^+ \times \mathbb{R}^2} (\phi)^{r_0 - 1/2(n+3)} \cdot \exp(-1/2 [(\boldsymbol{\beta}-\mathbf{b})^t \mathbf{V}^{-1}(\boldsymbol{\beta}-\mathbf{b}) + 2 \lambda_0 \phi]) \\ &\quad \cdot \prod_{i=0}^n \left[\sum_{j=1}^k p_j (a_j)^{-1/2} \exp(-1/2 \phi \cdot (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / \omega_i) \cdot \delta_{a_j}(\omega_i) \right] d\boldsymbol{\beta} d\phi, \end{aligned}$$

expresión analíticamente intratable; pero las distribuciones condicionales completas para implementar el muestreo de Gibbs son las siguientes:

$$(3.4) \quad \boldsymbol{\beta} \mid \mathbf{x}_0, \phi, \mathbf{p}, \boldsymbol{\omega}, \mathbf{y} \sim N_p(\mathbf{m}_\beta, S_\beta)$$

donde W y $\boldsymbol{\omega}$ son, respectivamente, la matriz diagonal y el vector de elementos ω_i ($i = 1, \dots, n$),

$$\mathbf{m}_\beta = S_\beta [\phi X^t W^{-1} \mathbf{y} + V^{-1} \mathbf{b}] \text{ y } S_\beta = [\phi X^t W^{-1} X + V^{-1}]^{-1},$$

$$X^t W^{-1} \mathbf{y} \text{ es un vector } p \times 1 \text{ con } j\text{-ésima componente: } \sum_{i=0}^n \frac{x_{ij} \cdot y_i}{\omega_i} \text{ y}$$

$$X^t W^{-1} X \text{ es matriz } p \times p \text{ de componente } (i,j): \sum_{m=0}^n \frac{x_{mi} \cdot x_{mj}}{\omega_m}.$$

$$(3.5) \quad \mathbf{x}_0 \mid \boldsymbol{\beta}, \phi, \mathbf{p}, \boldsymbol{\omega}, \mathbf{y} \sim N(\mathbf{m}_{x_0}, S_{x_0}^2),$$

$$\text{donde } \mathbf{m}_{x_0} = S_{x_0}^2 (D_0^{-1} \mathbf{c}_0 + \frac{\phi}{\omega_0} (y_0 - \beta_0) \cdot \boldsymbol{\beta}_1) \text{ y } S_{x_0}^2 = [\frac{\phi}{\omega_0} \boldsymbol{\beta}_1 \boldsymbol{\beta}_1^t + D_0^{-1}]^{-1}$$

$$(3.6) \quad \phi \mid \boldsymbol{\beta}, \mathbf{x}_0, \mathbf{p}, \boldsymbol{\omega}, \mathbf{y} \sim \text{Gama} \left(r_0 + \frac{n+1}{2}, \lambda_0 + \sum_{i=0}^n (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / 2\omega_i \right)$$

donde \mathbf{x}_i^t es la fila i -ésima de X .

$$(3.7) \quad \mathbf{p} \mid \boldsymbol{\beta}, \mathbf{x}_0, \phi, \boldsymbol{\omega}, \mathbf{y} \sim D \left(\boldsymbol{\alpha} + \sum_{i=0}^n \delta_{\omega_i} \right)$$

$$(3.8) \quad \omega_i \mid \boldsymbol{\beta}, \mathbf{x}_0, \phi, \mathbf{p}, y_i \sim \text{ind } g_i(\omega_i) = \sum_{j=1}^K \frac{\exp\{-\phi(y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / 2\omega_i\} \cdot p_j \cdot \delta_{a_j}(\omega_i)}{C_i \cdot \sqrt{\omega_i}},$$

$$\text{donde } C_i = \sum_{j=1}^K \frac{\exp\{-\phi(y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / 2a_j\} \cdot p_j}{C_i \cdot \sqrt{a_j}}.$$

Demostración

Las distribuciones condicionales completas se obtienen de forma directa a partir de los lemas de la sección 2.3 del capítulo 2. \square

La implementación se puede realizar confeccionando un programa para muestrear desde las condicionales o, en este caso, utilizando Winbugs.

Observación 3.1

El caso abordado arriba considera una mezcla discreta en parámetro escala de normales. Ahora bien, si $p_i = 1$ y $p_j = 0 \forall j \neq i$, es decir, si ω_i tiene una distribución degenerada en un punto, ya no se produce una mezcla de normales. Las distribuciones condicionales completas de los parámetros $\boldsymbol{\beta}$, \mathbf{x}_0 y ϕ se mantienen intactas, pero \mathbf{p} y ω_i pasan a ser constantes, de modo que $h(u) = (2\pi a)^{-1/2} e^{-u/2a}$, $u > 0$, con lo que se recupera el modelo de calibración lineal normal.

Observación 3.2

Considerando una muestra de pares (\mathbf{x}_i, y_i) con $i = 1, \dots, n$, es decir, omitiendo \mathbf{x}_0 , se recupera el modelo paramétrico de regresión con errores mezcla finita de normales heteroscedásticas, abordado por Arellano – Valle, Galea – Rojas e Iglesias (2000), para el cual las distribuciones condicionales completas que permiten la implementación de muestreo de Gibbs o la utilización de Winbuggs, difieren de las halladas únicamente en que las sumatorias omiten el término $i = 0$, en que la constante $(n + 1)$ se reemplaza por n y en que la expresión para \mathbf{x}_0 no existe.

El principal resultado de esta sección es presentado a continuación, y corresponde a la solución bayesiana semiparamétrica del problema de calibración bajo el modelo elíptico representable independiente.

Proposición 3.2

Bajo las especificaciones (3.1), (3.3), se considera que la priori para los parámetros de dispersión es un proceso Dirichlet G centrado en una distribución Gama Inversa G_0 , es decir, $G \sim D(c, G_0)$, con $G_0 \sim IG(r_1, \lambda_1)$.

La distribución a posteriori para \mathbf{x}_0 no permite obtener analíticamente el estimador de Bayes para \mathbf{x}_0 , como se ha mostrado anteriormente, sin embargo, las distribuciones condicionales completas para implementar el muestreo de Gibbs son las siguientes:

$$(3.9) \quad \boldsymbol{\beta} \mid \mathbf{x}_0, \phi, G, \boldsymbol{\omega}, \mathbf{y} \sim N_p(\mathbf{m}_\beta = S_\beta [\phi \mathbf{X}^t \mathbf{W}^{-1} \mathbf{y} + \mathbf{V}^{-1} \mathbf{b}], S_\beta = [\phi \mathbf{X}^t \mathbf{W}^{-1} \mathbf{X} + \mathbf{V}^{-1}]^{-1}),$$

donde \mathbf{W} se define igual que en la proposición anterior,

$$(3.10) \quad \mathbf{x}_0 \mid \boldsymbol{\beta}, \phi, G, \boldsymbol{\omega}, \mathbf{y} \sim N_{p-1}(\mathbf{m}_{\mathbf{x}_0} = S_{\mathbf{x}_0}^2 \left(\frac{\mathbf{c}_0}{D_0} + \frac{\phi}{\omega_0} (y_0 - \beta_0) \cdot \boldsymbol{\beta}_1 \right); S_{\mathbf{x}_0}^2 = \left[\frac{\phi}{\omega_0} \boldsymbol{\beta}_1 \boldsymbol{\beta}_1^t + \frac{1}{D_0} \right]^{-1}),$$

$$(3.11) \quad \phi \mid \boldsymbol{\beta}, \mathbf{x}_0, G, \boldsymbol{\omega}, \mathbf{y} \sim \text{Gama} \left(r_0 + \frac{n+1}{2}, \lambda_0 + \sum_{i=0}^n (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / 2\omega_i \right),$$

$$(3.12) \quad \omega_i \mid \boldsymbol{\beta}, \mathbf{x}_0, \phi, G, \{\omega_k\}_{k \neq i}, y_i \sim q_{0i} dG_b^{(i)}(\omega_i) + \sum_{k \neq i} q_{ki} \cdot \delta_{\omega_k}(\omega_i),$$

con q_{0i} es la función de densidad de la distribución $T(\mathbf{x}_i^t \boldsymbol{\beta}, 2\lambda_1/\phi, 2r_1)$ evaluada en y_i

q_{ki} es la función de densidad $N(\mathbf{x}_i^t \boldsymbol{\beta}, \phi^{-1}\omega_k)$ evaluada en y_i .

$$(3.13) \quad G_b^{(i)} \text{ es la función de distribución acumulada Gama Inversa } IG(r_1 + 1/2; \lambda_1 + \phi(y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / 2).$$

Demostración

Es claro que las primeras tres distribuciones condicionales completas no se ven influenciadas por la especificación de G , por lo cual coinciden con las respectivas distribuciones de la proposición 3.1.; por otra parte, las expresiones (3.12) y (3.13) se obtienen desde (1.12) y (1.13) (sección 1.1. del capítulo 1). \square

Observación 3.3

En muchos casos en la literatura se ha denominado modelo de regresión no paramétrico a un modelo muy diferente del que hemos abordado nosotros. Nos referimos al caso de un modelo con errores normales, para el cual la función que vincula las predictoras y la variable dependiente es desconocida y debe ser estimada:

$$y_i = f(\mathbf{x}_i) + \varepsilon_i, \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2).$$

Nuestro problema, en cambio, alude a un modelo en el cual la función es conocida (lineal), pero la distribución de los errores es desconocida. En el caso uniparamétrico, por ejemplo:

$$y_i = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i, \varepsilon_i \sim \text{i.i.d. } F, F \sim \text{Dirichlet}.$$

Observación 3.4

Suponer que la función generadora de densidad elíptica h es conocida es equivalente a suponer que la distribución de G es degenerada en la f.d.a. G_0 . Este problema ya ha sido enfrentado en la literatura. Para este caso las distribuciones condicionales (3.9), (3.10) y (3.11) permanecen válidas y se obtienen los resultados de Branco et al. (2000).

Por otra parte, si la distribución G_0 corresponde a la distribución $IG(v/2; v/2)$ con $v > 0$, y se le atribuye una distribución a priori a v , se obtienen los resultados presentados por Branco et al. (1998).

Observación 3.5

Si en la *Proposición 3.2* se consideran solamente los pares de observaciones (x_i, y_i) con $i = 1, \dots, n$, se recupera el modelo de regresión semiparamétrico elíptico representable independiente, el cual no ha sido abordado en la literatura (Müller y Quintana, 2004), extendiéndose de este modo los resultados de Galea e Iglesias (2003). Es así que, la solución bayesiana semiparamétrica al modelo de regresión elíptico representable consiste en las distribuciones condicionales completas presentadas anteriormente, omitiendo la relativa a x_0 , remplazando la constante $(n + 1)$ por n y omitiendo el término $i = 0$ en las sumatorias.

Observación 3.6

En el desarrollo anterior se supuso que G y ω son independientes de β , \mathbf{x}_0 y ϕ . Se podría generalizar estos resultados relajando este supuesto y explorando, por ejemplo, los resultados asociados al modelo jerárquico:

$$\begin{aligned}\beta | \phi &\sim N_p(\mathbf{b}, \phi^{-1} \cdot \mathbf{B}) \\ \phi &\sim \text{Gama}(r_0, \lambda_0).\end{aligned}$$

Este problema fue abordado para el caso paramétrico (G degenerada en G_0) y dependiente por Arellano – Valle, Galea – Rojas e Iglesias (2000) con $r_0 = d_0/2$, $\lambda_0 = d_0 \cdot c_0$ y G_0 Gama Inversa de parámetros r_1 y λ_1 . Estos autores probaron que, bajo estos supuestos, la distribución condicional completa para β marginalizada sobre ϕ es t -student.

Aquí también la distribución a posteriori de \mathbf{x}_0 para el caso semi paramétrico es intratable:

$$\begin{aligned}\pi(\mathbf{x}_0 | \mathbf{y}) &\propto \exp[-\frac{1}{2}(\mathbf{x}_0 - \mathbf{c}_0)^t D_0^{-1}(\mathbf{x}_0 - \mathbf{c}_0)] \cdot \\ &\iint_{\mathbb{R}^+ \times \mathbb{R}^2} (\phi)^{r_0 - \frac{1}{2}(n+p+2)} \cdot \exp(-\frac{1}{2}\phi [(\beta - \mathbf{b})^t \mathbf{B}^{-1}(\beta - \mathbf{b}) + 2\lambda_0]) \\ &\cdot \prod_{i=0}^n \left[\int_0^\infty \exp((r_1 - 3/2) \ln(\omega_i)) - [\frac{1}{2}\phi \cdot (y_i - \mathbf{x}_i^t \beta)^2 + \lambda_1] / \omega_i \right] d\omega_i d\beta d\phi\end{aligned}$$

por lo cual se justifica implementar el muestreo de Gibbs. Es así como se concluye que las distribuciones condicionales completas se mantienen inalterables, a excepción del caso de los parámetros β y ϕ , para los cuales se obtiene que:

$$(3.14) \quad \beta | \mathbf{x}_0, \phi, G, \omega, \mathbf{y} \sim N_p(\mathbf{m}_\beta, S_\beta),$$

$$\text{donde} \quad \mathbf{m}_\beta = \phi S_\beta [\mathbf{X}^t \mathbf{W}^{-1} \mathbf{y} + \mathbf{V}^{-1} \mathbf{b}], S_\beta = \phi^{-1} \cdot [\mathbf{X}^t \mathbf{W}^{-1} \mathbf{X} + \mathbf{V}^{-1}]^{-1}, \mathbf{y}$$

$$(3.15) \quad \phi | \beta, \mathbf{x}_0, G, \omega, \mathbf{y} \sim \text{Gama}(r^*, \lambda^*),$$

$$\text{donde} \quad r^* = r_0 + \frac{n+1}{2} \quad \text{y} \quad \lambda^* = \lambda_0 + \frac{1}{2} \left[\sum_{i=0}^n (y_i - \mathbf{x}_i^t \beta)^2 / \omega_i + (\beta - \mathbf{b})^t \mathbf{V}^{-1} (\beta - \mathbf{b}) \right].$$

Ejemplo 3.1.

Si el modelo contempla sólo una predictora, es decir si se trata de un modelo de Regresión Simple, y si β_0 y β_1 , los coeficientes del modelo, son independientes, la posteriori del parámetro de interés \mathbf{x}_0 es la siguiente:

$$\pi(\mathbf{x}_0 | \mathbf{y}) \propto \exp[-\frac{1}{2}(\mathbf{x}_0 - \mathbf{c}_0)^t \mathbf{D}_0^{-1}(\mathbf{x}_0 - \mathbf{c}_0)] \cdot \iiint_{\mathbb{R}^+ \times \mathbb{R}^2} (\phi)^{r_0 - \frac{1}{2}(n+3)} \exp(-\frac{1}{2}[(\beta_0 - b_0)^2/v_0 + (\beta_1 - b_1)^2/v_1 + 2\lambda_0 \phi]) \\ \cdot \prod_{i=0}^n \left[\sum_{j=1}^k p_j(a_j)^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi \cdot (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / \omega_i) \cdot \delta_{a_j}(\omega_i) \right] d\beta_0 d\beta_1 d\phi.$$

Las distribuciones condicionales completas coinciden con las halladas anteriormente (3.9) a (3.13), reemplazando $\mathbf{x}_i^t \boldsymbol{\beta}$ por $\beta_0 + \beta_1 x_i$.

Si además β_0 y β_1 son independientes a priori, es decir si $\mathbf{V} = \text{diag}(v_0, v_1)$, la distribución en (3.9) se descompone en dos y las distribuciones condicionales completas se pueden reescribir como:

$$(3.16) \quad \beta_0 | \beta_1, \mathbf{x}_0, \phi, \mathbf{G}, \boldsymbol{\omega}, \mathbf{y} \sim N(m_0 = S_0^2 \cdot \left[\phi \sum_{i=0}^n \frac{(y_i - \beta_1 x_i)}{\omega_i} + \frac{b_0}{v_0} \right], S_0^2 = \left[\phi \sum_{i=0}^n \frac{1}{\omega_i} + \frac{1}{v_0} \right]^{-1})$$

$$(3.17) \quad \beta_1 | \beta_0, \mathbf{x}_0, \phi, \mathbf{G}, \boldsymbol{\omega}, \mathbf{y} \sim N(m_1 = S_1^2 \cdot \left[\phi \sum_{i=0}^n \frac{x_i \cdot (y_i - \beta_0)}{\omega_i} + \frac{b_1}{v_1} \right], S_1^2 = \left[\phi \sum_{i=1}^n \frac{x_i^2}{\omega_i} + \frac{1}{v_1} \right]^{-1}),$$

$$(3.18) \quad \mathbf{x}_0 | \boldsymbol{\beta}, \phi, \mathbf{G}, \boldsymbol{\omega}, \mathbf{y} \sim N(\mathbf{m}_{\mathbf{x}_0} = \mathbf{S}_{\mathbf{x}_0}^2 [\mathbf{D}_0^{-1} \mathbf{c}_0 + \frac{\phi}{\omega_0} (y_0 - \beta_0) \cdot \boldsymbol{\beta}_1], \mathbf{S}_{\mathbf{x}_0}^2 = \left[\frac{\phi}{\omega_0} \boldsymbol{\beta}_1^2 + \mathbf{D}_0^{-1} \right]^{-1}).$$

$$(3.19) \quad \phi | \boldsymbol{\beta}, \mathbf{x}_0, \mathbf{G}, \boldsymbol{\omega}, \mathbf{y} \sim \text{Gama} \left(r_0 + \frac{n+1}{2}, \lambda_0 + \sum_{i=0}^n (y_i - \beta_0 - \beta_1 \cdot x_i)^2 / 2\omega_i \right),$$

$$(3.20) \quad G_b^{(i)} \text{ es la distribución } Gama \text{ Inversa } (r_1 + \frac{1}{2}, \lambda_1 + \phi (y_i - \beta_0 - \beta_1 \cdot x_i)^2 / 2).$$

$$(3.21) \quad \omega_i | \boldsymbol{\beta}, \mathbf{x}_0, \phi, \mathbf{G}, \{\omega_k\}_{k \neq i}, \mathbf{y} \sim q_{0i} dG_b^{(i)}(\omega_i) + \sum_{k \neq i} q_{ki} \cdot \delta_{\omega_k}(\omega_i)$$

donde q_{0i} corresponde a la f.d.p. de la distribución $T(\beta_0 + \beta_1 \cdot x_i, 2\lambda_1/\phi, 2r_1)$ evaluada en y_i ,

y q_{ki} corresponde a la f.d.p. de la distribución $N(\beta_0 + \beta_1 \cdot x_i, \phi^{-1}\omega_i)$ evaluada en la y_i .

Por otro lado, si $\boldsymbol{\beta}$ y ϕ no son independientes (*observación 3.4*), la distribución a posteriori de \mathbf{x}_0 es:

$$\pi(\mathbf{x}_0 | \mathbf{y}) \propto \exp[-\frac{1}{2}(\mathbf{x}_0 - \mathbf{c}_0)^t \mathbf{D}_0^{-1}(\mathbf{x}_0 - \mathbf{c}_0)] \cdot \iiint_{\mathbb{R}^+ \times \mathbb{R}^2} (\phi)^{r_0 - \frac{1}{2}(n+p+2)} \cdot \exp(-\frac{1}{2} \phi [(\beta_0 - b_0)^2/v_0 + (\beta_1 - b_1)^2/v_1 + 2\lambda_0]) \\ \cdot \prod_{i=0}^n \left[\sum_{j=1}^k p_j(a_j)^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi \cdot (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / \omega_i) \cdot \delta_{a_j}(\omega_i) \right] d\beta_0 d\beta_1 d\phi,$$

y las distribuciones condicionales completas se mantienen para todos los parámetros, excepto para β_0 , β_1 y ϕ :

$$(3.22) \quad \beta_0 | \beta_1, \mathbf{x}_0, \phi, G, \boldsymbol{\omega}, \mathbf{y} \sim N(m_0 = \phi \cdot S_0^2 \cdot \left[\sum_{i=0}^n \frac{(y_i - \beta_1 x_i)}{\omega_i} + \frac{b_0}{v_0} \right], S_0^2 = \phi^{-1} \left[\sum_{i=0}^n \frac{1}{\omega_i} + \frac{1}{v_0} \right]^{-1}),$$

$$(3.23) \quad \beta_1 | \beta_0, \mathbf{x}_0, \phi, G, \boldsymbol{\omega}, \mathbf{y} \sim N(m_1 = \phi \cdot S_1^2 \cdot \left[\sum_{i=0}^n \frac{x_i \cdot (y_i - \beta_0)}{\omega_i} + \frac{b_1}{v_1} \right], S_1^2 = \phi^{-1} \left[\sum_{i=1}^n \frac{x_i^2}{\omega_i} + \frac{1}{v_1} \right]^{-1}),$$

$$(3.24) \quad \phi | \boldsymbol{\beta}, \mathbf{x}_0, G, \boldsymbol{\omega}, \mathbf{y} \sim \text{Gama} \left(r_0 + \frac{n+1}{2}, \lambda_0 + \frac{1}{2} \left[\sum_{i=0}^n (y_i - \beta_0 - \beta_1 \cdot x_i)^2 / \omega_i + \sum_{j=0}^1 (\beta_j - b_j)^2 / v_j \right] \right).$$

3.2 SOLUCIÓN BAYESIANA SEMIPARAMÉTRICA BAJO EL MODELO DE REGRESIÓN ELÍPTICO REPRESENTABLE DEPENDIENTE

En esta sección se aborda el problema de calibración bajo el modelo de regresión, pero ahora considerando el término de error elíptico dependiente y con función generadora de densidad h representable. El modelo al cual se alude es, entonces:

$$(3.25) \quad \mathbf{y} | \boldsymbol{\beta}, \mathbf{x}_0, \phi \sim EL_n(X \boldsymbol{\beta}, \phi^{-1}, h) \text{ y } h(\mathbf{u}) = \int (2\pi\omega)^{-n/2} e^{-\mathbf{u}^T / 2 \omega} dG(\omega),$$

con $h \geq 0$ y donde G es función de distribución acumulada f.d.a. tal que $G(0) = 0$.

En este caso el problema puede ser abordado considerando la especificación de (3.25) en varias etapas, como sigue (*Proposición 2.1a*, sección 2.2 del Capítulo 2):

$$(3.26) \quad \begin{aligned} \mathbf{y} | \boldsymbol{\beta}, \mathbf{x}_0, \phi, \boldsymbol{\omega} &\sim N_n(X \boldsymbol{\beta}, \phi^{-1} \boldsymbol{\omega} \cdot \mathbf{I}_n) \\ \boldsymbol{\omega} | G &\sim G \\ (G, \boldsymbol{\omega}) &\perp (\boldsymbol{\beta}, \mathbf{x}_0, \phi) \end{aligned}$$

El modelo que nos interesa une (3.25) o (3.26) a las siguientes etapas, equivalentes a (3.3):

$$(3.27) \quad \begin{aligned} \boldsymbol{\beta} &\sim N_p(\mathbf{b}, V), \\ \mathbf{x}_0 &\sim N_{p-1}(\mathbf{c}_0, D_0), \\ \phi &\sim \text{Gama}(r_0, \lambda_0). \end{aligned}$$

Al igual que en la sección 3.1, y debido a las dificultades analíticas inherentes al problema, nuestro objetivo de hallar el estimador de Bayes del parámetro de interés \mathbf{x}_0 , se transforma en la búsqueda de las distribuciones condicionales conjuntas que permitan la implementación del muestreo de Gibbs. Ahora bien, la especificación (3.25) no permite encontrar en forma explícita estas distribuciones, mientras que la especificación (3.26) si lo permite (se trata de la distribución conjunta de normales independientes).

En efecto, la función de verosimilitud asociada al modelo de interés, aún considerando G conocida, bajo la especificación (3.25), es de la forma siguiente:

$$L(\boldsymbol{\beta}, \mathbf{x}_0, \phi | \mathbf{y}, G) = (\phi / 2 \pi)^{1/2(n+1)} \left[\prod_{i=0}^n \int_0^{\infty} \omega^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi \cdot (y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2 / \omega) dG(\omega) \right],$$

Mientras que para el modelo dado por (3.26) la verosimilitud aumentada coincide con la del modelo normal homoscedástico:

$$L(\boldsymbol{\beta}, \mathbf{x}_0, \phi, \boldsymbol{\omega} | \mathbf{y}, G) = (\phi / 2 \pi \omega)^{1/2(n+1)} \cdot \exp[-\frac{1}{2} (\phi / \omega) \cdot \sum_{i=0}^n (y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2].$$

Desde un punto de vista clásico y sin conocer la distribución G , en el primer caso es imposible obtener los estimadores máximo verosímiles. Para el segundo caso, G se asume dado, pero en realidad para nosotros es desconocido. Ahora bien, la maximización de la función para obtener estimadores máximo verosímiles, nos obliga a resolver un sistema de ecuaciones análogo al del modelo independiente. En efecto, para la definición en la sección precedente de la función $k(e_i, \phi)$, con ω en lugar de ω_i , para la especificación (3.25), se llega a las mismas ecuaciones que para la especificación (3.1). Sin embargo, para el modelo dado por (3.26), las ecuaciones que se obtienen son diferentes:

$$\frac{\phi}{(n+1) \cdot \omega} \cdot \sum_{i=0}^n e_i^2 = 1; \quad \sum_{i=0}^n e_i = 0; \quad \sum_{i=0}^n x_i \cdot e_i = 0, \text{ y } e_0 = 0, \text{ con } \omega, \phi \neq 0.$$

$$\hat{\phi}^{-1} \cdot \hat{\omega} = \frac{\sum_{i=0}^n \hat{e}_i^2}{n+1} \text{ y } \hat{e}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 \cdot x_i \quad \forall i = 1, \dots, n;$$

de donde se obtiene que los estimadores máximo verosímiles de β_0 , β_1 y x_0 son soluciones del sistema:

$$\beta_0 = \bar{y} - \beta_1 \cdot \bar{x}; \quad \beta_1 = \frac{S_{XY}}{S_X^2}; \quad \omega = \frac{\sum_{i=0}^n e_i^2}{(n+1) \cdot \phi}, \text{ e } y_0 = \beta_0 + \beta_1 \cdot x_0.$$

Esto deja en evidencia que no existe solución única para el conjunto de parámetros, ya que el modelo no es identificable. En particular ϕ y ω no son estimables, pero la función $(\phi^{-1} \cdot \omega)$ sí lo es. Por otra parte, los estimadores de los coeficientes de la regresión corresponden a los estimadores de mínimos cuadrados usuales y no dependen de ω ni de ϕ .

Por otra parte, para una distribución G particular bien escogida como Gama Inversa, la estimación máximo verosímil para la especificación (3.25) coincide con aquella correspondiente a la especificación (3.1) en el caso independiente, donde, como se mostró anteriormente, se requiere una resolución mediante métodos numéricos.

Ahora bien, desde el punto de vista bayesiano, la distribución a posteriori de los parámetros en estudio para la primera especificación tienen la forma siguiente:

$$\begin{aligned} \pi(\boldsymbol{\beta}, \mathbf{x}_0, \phi | \mathbf{y}, G) &\propto L(\boldsymbol{\beta}, \mathbf{x}_0, \phi | \mathbf{y}, G) \cdot \pi(\boldsymbol{\beta}) \cdot \pi(\mathbf{x}_0) \cdot \pi(\phi) \\ &\propto (\phi)^{r_0 - 1/2(n+3)} \cdot \exp(-1/2 [(\boldsymbol{\beta}-\mathbf{b})^t \mathbf{V}^{-1}(\boldsymbol{\beta}-\mathbf{b}) + (\mathbf{x}_0-\mathbf{c}_0)^t D_0^{-1}(\mathbf{x}_0-\mathbf{c}_0)] - \lambda_0 \phi) \\ &\quad \cdot \prod_{i=0}^n \left[\int_0^{\infty} (\omega)^{-1/2} \exp(-1/2 \phi \cdot (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / \omega) dG(\omega) \right], \end{aligned}$$

cuya constante de proporcionalidad en la práctica es imposible de obtener de forma analítica.

Marginalizando se obtiene la siguiente expresión para la distribución a posteriori del parámetro de interés \mathbf{x}_0 , que, en vista de que ω es la variable de integración, coincide con la expresión análoga obtenida para el caso independiente donde las ω_i son i.i.d.:

$$\begin{aligned} \pi(\mathbf{x}_0 | \mathbf{y}) &\propto \exp[-1/2 (\mathbf{x}_0 - \mathbf{c}_0)^t D_0^{-1} (\mathbf{x}_0 - \mathbf{c}_0)] \cdot \\ &\quad \iiint_{\mathbb{R}^+ \times \mathbb{R}^2} (\phi)^{r_0 - 1/2(n+3)} \cdot \exp(-1/2 [(\boldsymbol{\beta}-\mathbf{b})^t \mathbf{V}^{-1}(\boldsymbol{\beta}-\mathbf{b}) + 2 \lambda_0 \phi]) \\ &\quad \cdot \prod_{i=0}^n \left[\int_0^{\infty} (\omega)^{-1/2} \exp(-1/2 \phi \cdot (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / \omega) dG(\omega) \right] d\boldsymbol{\beta} d\phi, \end{aligned}$$

Como se comentó en la sección anterior, esta expresión depende del proceso G y de su distribución a priori y resulta intratable analíticamente incluso para G escogida como Gama Inversa de parámetros r y λ . En este caso, entonces, el estimador de Bayes tampoco es obtenible analíticamente. Estos hechos motivan la deducción de las distribuciones condicionales completas para implementar el muestreo de Gibbs.

Ahora bien, para la segunda especificación, la posteriori asume la forma siguiente:

$$\begin{aligned} \pi(\boldsymbol{\beta}, \mathbf{x}_0, \phi, \omega | \mathbf{y}) &\propto L(\boldsymbol{\beta}, \mathbf{x}_0, \phi, \omega | \mathbf{y}, G) \cdot \pi(\boldsymbol{\beta}) \cdot \pi(\mathbf{x}_0) \cdot \pi(\phi) \cdot \pi(\omega | G), \\ &\propto (\phi)^{r_0 - \frac{1}{2}(n+3)} \cdot \exp\left[-\frac{1}{2} [(\boldsymbol{\beta}-\mathbf{b})^t V^{-1}(\boldsymbol{\beta}-\mathbf{b}) + (\mathbf{x}_0-\mathbf{c}_0)^t D_0^{-1}(\mathbf{x}_0-\mathbf{c}_0)] - \lambda_0 \phi\right] \\ &\cdot [(\omega)^{-\frac{1}{2}} \cdot dG(\omega)]^{(n+1)} \cdot \exp\left[-\frac{1}{2} (\phi/\omega) \cdot \sum_{i=0}^n (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2\right]. \end{aligned}$$

Es claro que esta expresión no es más tratable que la abordada antes y que lleva a resultados semejantes.

Ahora bien, las dos proposiciones siguientes abordan el modelo que une (3.25) y (3.27), considerando las mismas dos elecciones de la distribución a priori para la función de mezcla G utilizadas anteriormente. Tal como para el caso independiente, la distribución a posteriori para \mathbf{x}_0 es muy difícil de tratar analíticamente, pero las distribuciones condicionales completas son simples:

Proposición 3.3

Bajo las especificaciones (3.25), (3.27), y escogiendo para el parámetro de dispersión común una distribución a priori discreta $p(\omega | G) = \sum_{j=1}^{k+1} p_j \cdot \delta_{a_j}(\omega)$, cuyos pesos tienen una distribución a priori

Dirichlet $(p_1, \dots, p_k) \sim D(\alpha_1, \dots, \alpha_k)$, con a_j fijos, se obtiene la siguiente expresión para la distribución a posteriori de \mathbf{x}_0 :

$$\begin{aligned} \pi(\mathbf{x}_0 | \mathbf{y}) &\propto \exp\left[-\frac{1}{2} (\mathbf{x}_0 - \mathbf{c}_0)^t D_0^{-1}(\mathbf{x}_0 - \mathbf{c}_0)\right] \cdot \\ &\iiint_{\mathbb{R}^+ \times \mathbb{R}^2} (\phi)^{r_0 - \frac{1}{2}(n+3)} \cdot \exp\left[-\frac{1}{2} [(\boldsymbol{\beta}-\mathbf{b})^t V^{-1}(\boldsymbol{\beta}-\mathbf{b}) + 2\lambda_0 \phi]\right] \\ &\cdot \prod_{i=0}^n \left[\sum_{j=1}^k p_j (a_j)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \phi \cdot (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / \omega\right) \cdot \delta_{a_j}(\omega) \right] d\boldsymbol{\beta} d\phi \end{aligned}$$

Por otro lado, las distribuciones condicionales completas difieren de las halladas en la sección anterior sólo en que ω_i se reemplaza por ω y por lo tanto el vector $\boldsymbol{\omega}$ pasa a ser un escalar, de donde:

$$(3.28) \quad \boldsymbol{\beta} | \mathbf{x}_0, \phi, \mathbf{p}, \omega, \mathbf{y} \sim N_p(\mathbf{m}_{\boldsymbol{\beta}} = S_{\boldsymbol{\beta}}^2 [\phi \omega^{-1} X^t \mathbf{y} + V^{-1} \mathbf{b}], S_{\boldsymbol{\beta}}^2 = [\phi \omega^{-1} X^t X + V^{-1}]^{-1})$$

$$(3.29) \quad \mathbf{x}_0 | \boldsymbol{\beta}, \phi, \mathbf{p}, \omega, \mathbf{y} \sim N_{p-1}(\mathbf{m}_{\mathbf{x}_0} = S_{\mathbf{x}_0}^2 \left(\frac{\mathbf{c}_0}{D_0} + \frac{\phi}{\omega} (y_0 - \beta_0) \cdot \boldsymbol{\beta}_1 \right); S_{\mathbf{x}_0}^2 = \left[\frac{\phi}{\omega} \boldsymbol{\beta}_1 \boldsymbol{\beta}_1^t + \frac{1}{D_0} \right]^{-1})$$

$$(3.30) \quad \phi | \boldsymbol{\beta}, \mathbf{x}_0, \mathbf{p}, \omega, \mathbf{y} \sim \text{Gama} \left(r_0 + \frac{n+1}{2}, \lambda_0 + (\mathbf{y} - X \boldsymbol{\beta})^t (\mathbf{y} - X \boldsymbol{\beta}) / 2\omega \right)$$

donde \mathbf{x}_i^t es la fila i -ésima de X.

$$(3.31) \quad \mathbf{p} \mid \boldsymbol{\beta}, \mathbf{x}_0, \phi, \omega, \mathbf{y} \sim D(\boldsymbol{\alpha} + \delta_\omega), \text{ donde } \omega \text{ asume valores } a_j.$$

$$(3.32) \quad \omega \mid \boldsymbol{\beta}, \mathbf{x}_0, \phi, \mathbf{p}, \mathbf{y} \sim g(x) = \frac{\sum_{j=1}^{K+1} \frac{\exp\left\{-\sum_{i=0}^n \frac{(y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2}{2x}\right\} \cdot \phi \cdot p_j \cdot \delta_{a_j}(x)}{C \cdot \sqrt{x^{n+1}}},$$

$$\text{donde } C = \frac{\sum_{j=1}^{K+1} \frac{\exp\left\{-\sum_{i=0}^n \frac{(y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2}{2a_j}\right\} \cdot \phi \cdot p_j}{C \cdot \sqrt{a_j^{n+1}}}.$$

Demostración

Las distribuciones condicionales completas se obtienen de forma directa a partir de los lemas de la sección 2.3 del capítulo 2. □

Al igual que para el modelo independiente, una alternativa factible en este caso, en lugar de programar el muestreo desde las distribuciones condicionales completas, consiste en utilizar Winbugs.

Observación 3.7

Es claro que, tratándose de una mezcla discreta de normales, si $p_i = 1$ y $p_j = 0 \forall j \neq i$, es decir, si ω tiene una distribución degenerada en un punto, ya no se produce la mezcla, por lo cual se recupera el modelo normal multivariado, donde las distribuciones condicionales completas de los parámetros $\boldsymbol{\beta}$, ϕ y \mathbf{x}_0 se mantienen intactas, pero \mathbf{p} y ω son constantes.

Observación 3.8

Análogamente a lo expuesto para el caso independiente, omitiendo \mathbf{x}_0 , se recupera el modelo de regresión con errores mezcla finita de normales homoscedásticas, de modo que la solución bayesiana ya está dada, con sólo omitir la distribución condicional completa de este parámetro, reemplazar por n la constante $(n+1)$ y omitir el término $i = 0$ en las sumatorias.

A continuación se presenta el resultado más importante de esta sección, que corresponde a la solución bayesiana semiparamétrica del problema de calibración bajo el modelo elíptico representable dependiente.

Proposición 3.4

Bajo las especificaciones (3.25), (3.27), y considerando que la priori para el parámetro común de dispersión es un proceso Dirichlet G centrado en una distribución Gama Inversa G_0 , es decir:

$$\begin{aligned} G &\sim \mathcal{D}(c, G_0), \\ G_0 &\sim \text{Gama Inversa}(\tau_1, \lambda_1), \end{aligned}$$

la distribución a posteriori para \mathbf{x}_0 no cambia respecto del modelo independiente, lo que motiva la obtención de las distribuciones condicionales completas. Las primeras tres coinciden con las de la *Proposición 3.3*:

$$(3.33) \quad \boldsymbol{\beta} \mid \mathbf{x}_0, \phi, G, \omega, \mathbf{y} \sim N_p(\mathbf{m}_\beta = S_\beta [\phi \omega^{-1} \mathbf{X}^t \mathbf{y} + \mathbf{V}^{-1} \mathbf{b}], S_\beta = [\phi \omega^{-1} \mathbf{X}^t \mathbf{X} + \mathbf{V}^{-1}]^{-1})$$

$$(3.34) \quad \mathbf{x}_0 \mid \boldsymbol{\beta}, \phi, G, \omega, \mathbf{y} \sim N_{p-1}(\mathbf{m}_{\mathbf{x}_0} = S_{\mathbf{x}_0}^2 \left(\frac{\mathbf{c}_0}{D_0} + \frac{\phi}{\omega} (y_0 - \beta_0) \cdot \boldsymbol{\beta}_1 \right); S_{\mathbf{x}_0}^2 = \left[\frac{\phi}{\omega} \boldsymbol{\beta}_1 \boldsymbol{\beta}_1^t + \frac{1}{D_0} \right]^{-1})$$

$$(3.35) \quad \phi | \boldsymbol{\beta}, \mathbf{x}_0, G, \omega, \mathbf{y} \sim \text{Gama} \left(r_0 + \frac{n+1}{2}, \lambda_0 + (\mathbf{y} - \mathbf{X} \boldsymbol{\beta})^t (\mathbf{y} - \mathbf{X} \boldsymbol{\beta}) / 2\omega \right)$$

$$(3.36) \quad \omega | \boldsymbol{\beta}, \mathbf{x}_0, \phi, G, \mathbf{y} \sim IG \left(r_1 + \frac{n+1}{2}, \lambda_1 + \phi \sum_{i=0}^n (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 / 2 \right)$$

Demostración

Es claro que las primeras tres distribuciones condicionales completas no se ven influenciadas por la distribución a priori para G, por lo cual coinciden con las respectivas distribuciones de la proposición 3.3. Por otra parte, la expresión (3.36) se obtiene desde (3.12), con $\omega_i = \omega: \sum_{k \neq i} q_{ki} \cdot \delta_{\omega_k}(\omega_i) = 0$ y por lo tanto la distribución a posteriori de $\omega | \boldsymbol{\beta}, \mathbf{x}_0, \phi$ es proporcional a la distribución G_b , y se obtiene de modo sencillo el resultado anterior (note que $k \neq i$ implica $\delta_{\omega_k}(\omega_i) = 0$). □

Observación 3.9

Si suponemos que la función generadora de densidad h es conocida, estamos frente al caso en que G es degenerada en la f.d.a. G_0 . En este caso, permanecen válidas las expresiones (3.33), (3.34) y (3.35), y la distribución condicional completa para ω es G_b .

Observación 3.10

Si en la proposición anterior se consideran solamente los vectores (\mathbf{x}_i, y_i) para $i = 1, \dots, n$ (es decir, si se omite el caso $i = 0$), se recupera el modelo de regresión elíptico representable dependiente, donde las soluciones bayesianas semiparamétricas están dadas por considerar las distribuciones condicionales completas anteriores omitiendo aquella relativa a \mathbf{x}_0 , reemplazar $(n+1)$ por n en dichas expresiones y por omitir los términos $i = 0$ en las sumatorias.

Observación 3.11

Como se mencionó en la *Observación 3.4.*, es posible generalizar asumiendo que ϕ y $\boldsymbol{\beta}$ no son independientes, por ejemplo que $\boldsymbol{\beta} | \phi \sim N_p(\mathbf{b}, \phi^{-1} \cdot \mathbf{V})$, $\phi \sim \text{Gama}(r_0, \lambda_0)$. El caso paramétrico fue abordado por Arellano – Valle, Galea – Rojas e Iglesias (2000) con $r_0 = d_0 / 2$ y $\lambda_0 = d_0 \cdot c_0$ y G_0 Gama Inversa de parámetros r_1 y λ_1 . Marginalizando sobre ϕ , se obtiene una distribución t – student para $\boldsymbol{\beta} | \mathbf{x}_0, \omega, \mathbf{y}$.

En este caso, la posteriori para \mathbf{x}_0 no varía respecto del caso independiente, ya que ω es variable muda para la integración::

$$\pi(\mathbf{x}_0 | \mathbf{y}) \propto \exp[-1/2 (\mathbf{x}_0 - \mathbf{c}_0)^t D_0^{-1} (\mathbf{x}_0 - \mathbf{c}_0)] \cdot \iiint_{\mathbb{R}^t \times \mathbb{R}^2} (\phi)^{r_0 - 1/2(n+p+2)} \cdot \exp(-1/2 \phi [(\boldsymbol{\beta} - \mathbf{b})^t \mathbf{B}^{-1} (\boldsymbol{\beta} - \mathbf{b}) + 2 \lambda_0])$$

$$\cdot \prod_{i=0}^n \left[\int_0^\infty (\omega)^{-r-3/2} \exp(-[1/2 \phi \cdot (y_i - \mathbf{x}_i^t \boldsymbol{\beta})^2 + \lambda] / \omega) d\omega \right] d\boldsymbol{\beta} d\phi.$$

Al realizar la generalización se obtienen las siguientes distribuciones completas para β y ϕ :

$$(3.37) \quad \beta \mid \mathbf{x}_0, \phi, G, \omega, \mathbf{y} \sim N_p(\mathbf{m}_\beta, S_\beta),$$

$$\text{donde} \quad \mathbf{m}_\beta = \phi S_\beta [\omega^{-1} \mathbf{X}' \mathbf{y} + \mathbf{V}^{-1} \mathbf{b}], S_\beta = \phi^{-1} [\omega^{-1} \mathbf{X}' \mathbf{X} + \mathbf{V}^{-1}]^{-1} \quad \text{y}$$

$$(3.38) \quad \phi \mid \beta, \mathbf{x}_0, G, \omega, \mathbf{y} \sim \text{Gama}(r^*, \lambda^*),$$

$$\text{donde} \quad r^* = r_0 + \frac{n+1}{2} \quad \text{y} \quad \lambda^* = \lambda_0 + \frac{1}{2} [(\mathbf{y} - \mathbf{X} \beta)' (\mathbf{y} - \mathbf{X} \beta) / \omega + (\beta - \mathbf{b})' \mathbf{V}^{-1} (\beta - \mathbf{b})].$$

Ejemplo 3.2

Si el modelo contempla sólo una predictora, es decir si se trata de un modelo de Regresión Simple, en el cual los coeficientes β_0 y β_1 son independientes a priori, es decir $\mathbf{V} = \text{diag}(v_0, v_1)$, las distribuciones condicionales completas se pueden re-escribir como:

$$(3.39) \quad \beta_0 \mid \beta_1, \mathbf{x}_0, \phi, G, \omega, \mathbf{y} \sim N(m_0 = S_0^2 \cdot \left[\frac{\phi}{\omega} \sum_{i=0}^n (y_i - \beta_1 x_i) + \frac{b_0}{v_0} \right], S_0^2 = \left[\frac{n \cdot \phi}{\omega} + \frac{1}{v_0} \right]^{-1})$$

$$(3.40) \quad \beta_1 \mid \beta_0, \mathbf{x}_0, \phi, G, \omega, \mathbf{y} \sim N(m_1 = S_1^2 \cdot \left[\frac{\phi}{\omega} \sum_{i=0}^n x_i \cdot (y_i - \beta_0) + \frac{b_1}{v_1} \right], S_1^2 = \left[\frac{\phi}{\omega} \sum_{i=1}^n x_i^2 + \frac{1}{v_1} \right]^{-1})$$

$$(3.41) \quad \mathbf{x}_0 \mid \beta, \phi, G, \omega, \mathbf{y} \sim N(\mathbf{m}_{\mathbf{x}_0} = S_{\mathbf{x}_0}^2 [D_0^{-1} \mathbf{c}_0 + \frac{\phi}{\omega} (y_0 - \beta_0) \cdot \beta_1], S_{\mathbf{x}_0}^2 = \left[\frac{\phi}{\omega} \beta_1^2 + D_0^{-1} \right]^{-1}).$$

$$(3.42) \quad \phi \mid \beta, \mathbf{x}_0, G, \omega, \mathbf{y} \sim \text{Gama} \left(r_0 + \frac{n+1}{2}, \lambda_0 + \sum_{i=0}^n (y_i - \beta_0 - \beta_1 x_i)^2 / 2\omega \right)$$

$$(3.43) \quad \omega \mid \beta, \mathbf{x}_0, \phi, G, \mathbf{y} \sim IG \left(r = r_1 + \frac{n+1}{2}, \lambda = \lambda_1 + \frac{\phi}{2} \sum_{i=0}^n (y_i - \beta_0 - \beta_1 x_i)^2 \right).$$

Por otro lado, si β y ϕ no son independientes (*Observación 3.6.*):

$$\beta \mid \phi \sim N_p(\mathbf{b}, \phi^{-1} \cdot \mathbf{V}),$$

$$\phi \sim \text{Gama}(r_0, \lambda_0),$$

las distribuciones condicionales completas se mantienen para todos los parámetros, excepto β_0, β_1 y ϕ :

$$(3.44) \quad \beta_0 \mid \beta_1, \mathbf{x}_0, \phi, G, \omega, \mathbf{y} \sim N(m_0 = \phi \cdot S_0^2 \cdot \left[\frac{(n+1) \cdot (\bar{y} - \beta_1 \bar{x})}{\omega} + \frac{b_0}{v_0} \right], S_0^2 = \phi^{-1} \left[\frac{n+1}{\omega} + \frac{1}{v_0} \right]^{-1}),$$

$$(3.45) \quad \beta_1 \mid \beta_0, \mathbf{x}_0, \phi, G, \omega, \mathbf{y} \sim N(m_1 = m^*, S_1^2 = S^*),$$

$$\text{donde} \quad m^* = \phi \cdot S_1^2 \cdot \left[\frac{\sum_{i=0}^n x_i \cdot y_i - (n+1) \cdot \beta_0 \cdot \bar{x}}{\omega} + \frac{b_1}{v_1} \right] \quad \text{y} \quad S^* = \phi^{-1} \left[\frac{\sum_{i=1}^n x_i^2}{\omega} + \frac{1}{v_1} \right]^{-1}, \quad \text{y}$$

$$(3.46) \quad \phi \mid \beta, \mathbf{x}_0, G, \omega, \mathbf{y} \sim \text{Gama}(r^*, \lambda^*),$$

$$\text{donde} \quad r^* = r_0 + \frac{n+1}{2} \quad \text{y} \quad \lambda^* = \lambda_0 + \frac{1}{2} \left[\sum_{i=0}^n (y_i - \beta_0 - \beta_1 \cdot x_i)^2 / \omega + \sum_{j=0}^1 (\beta_j - b_j)^2 / v_j \right].$$

3.3 APLICACIONES

En esta sección se ilustran los resultados obtenidos en las secciones previas a través de simulación y conjuntos de datos considerados previamente en la literatura.

Aplicación 3.1 Medición del Porcentaje de Grasa Corporal (Branco et. al., 2000)

Un estudio conducido en la Pontificia Universidad Católica de Chile consideró 24 mujeres con edades entre 16 y 19 años, para cada una de las cuales se obtuvieron dos mediciones del porcentaje de grasa corporal. La primera medición, más costosa y precisa, corresponde a DEXA (dual – energy x – ray absorptiometry) y la segunda medición corresponde a la ecuación de Durnin (que relaciona tensión skinfold y densidad).

Tabla 3.1: Datos de Grasa Corporal, Branco et. al. (2000). La medición x corresponde a DEXA y la medición y corresponde a la ecuación de Durnin.

X	28.6	24.8	30.9	32.2	35.6	18.4	37.9	27.0	37.6	40.8	33.7	20.6
Y	28.34	25.68	25.26	29.12	35.56	18.86	35.06	32.77	34.84	33.84	30.87	18.63
X	24.7	32.7	34.2	28.9	23.8	25.6	34.9	30.7	36.7	42.4	43.5	29.7
Y	27.51	31.25	31.16	31.86	25.96	22.60	24.53	26.88	31.44	33.84	35.56	25.12

El objetivo es estimar x_0 , una nueva medición DEXA, a partir del valor de la ecuación de Durmin y_0 . Los datos, que fueron proporcionados por el Departamento de Pediatría de la institución se presentan en la *Tabla 3.1.* y el diagrama de dispersión se presenta en la *Figura 3.1.*

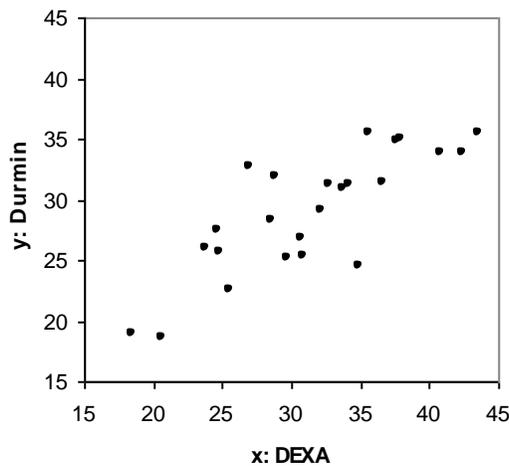


Figura 3.1. Diagrama de Dispersión para los datos de la *Tabla 3.1.:* Mediciones de grasa para 24 mujeres.

Branco et al. (2000) presentan soluciones bayesianas al problema de inferencia sobre x_0 para los casos dependiente e independiente, basada, en ambos casos, en una priori t – student con v grados de libertad, donde la priori para v es exponencial de parámetro $d = 0,10$. En esta sección relajamos dichos supuestos no asumiendo una forma conocida para la distribución elíptica representable, tanto en el caso dependiente como en el independiente.

Branco et. al. (2000) escogen los hiper parámetros de las distribuciones a priori de modo que la media de cada distribución coincida con los estimadores máximo verosímiles bajo el modelo normal. Por ejemplo, para la distribución a priori de los coeficientes de regresión, β_0 y β_1 , utilizan distribuciones normales con medias 9,8 y 0,6 y varianzas 100 y 1, respectivamente, y para el parámetro de precisión una priori gamma de parámetros de posición y escala iguales a 0,1. Para efectos comparativos, en este trabajo también se utilizan dichas prioris. Al igual que en el caso de Branco et al. (2000), se generó mediante Gibbs Sampler una cadena de largo 70.000, eliminando 10.000 iteraciones iniciales.

Para las demás distribuciones Gamma Inversa asociadas a los parámetros de dispersión ω se utilizaron dos variantes, la variante A presenta parámetros $r = 4$ y $\lambda = 4$ y la variante B presenta parámetros $r = 1$ y $\lambda = 1/3$. La primera tiene la moda y la varianza superior a la segunda, y las colas más pesadas. En este caso se asume c fijo igual a uno (Figura 3.2.).

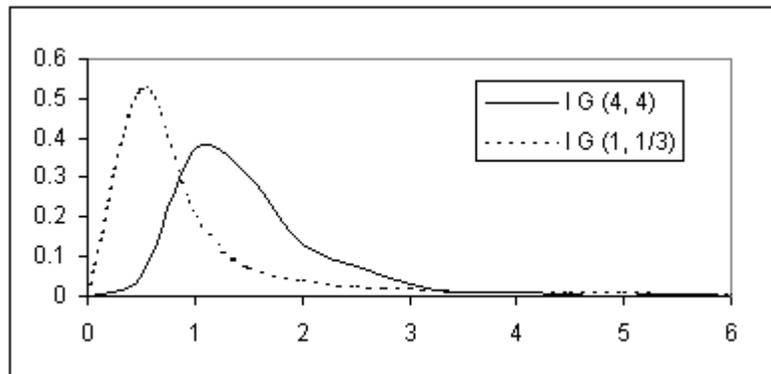


Figura 3.2. Distribuciones Gamma Inversa de parámetros (4, 4) y (1, 1/3), utilizadas como prioris (A y B) de los parámetros ω .

PRIORI I: NO INFORMATIVA

El primer análisis realizado utiliza la distribución a priori para x_0 normal con media 31,54 y varianza 52,94, que corresponden a la priori I no informativa presentada por Branco et. al. (2000). Esta distribución corresponde a una aproximación de la priori de Hoadley (1970). Los resultados se presentan en la *Tabla 3.2*.

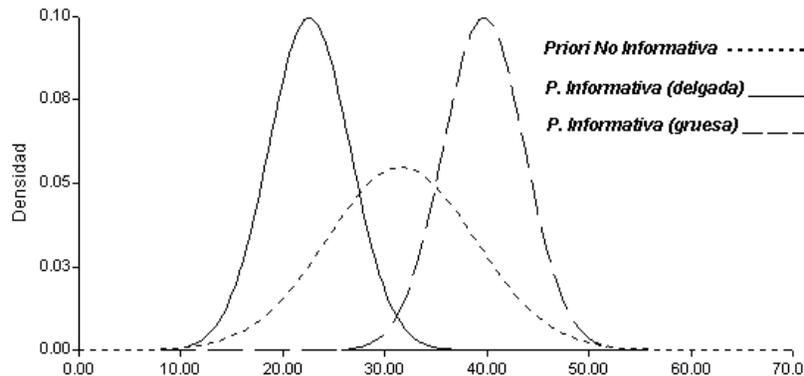


Figura 3.3. Distribuciones a priori para x_0 consideradas en el análisis: I No informativa y II Informativa.

Tabla 3.2: Estimación bayesiana de los parámetros de interés para una mujer robusta y una mujer delgada (valores de ecuación de Durmin 33.25 y 26.98 respectivamente) para la priori no informativa de Branco et. al. (2000), utilizando los modelos dependiente e independiente.

Modelo	Mujer robusta $y_0 = 33,25$				Mujer delgada $y_0 = 26,98$			
	Independiente		Dependiente		Independiente		Dependiente	
Parámetros	A	B	A	B	A	B	A	B
β_0	2.88	2.42	2.95	3.06	2.30	2.22	2.76	2.78
β_1	0.99	1.00	0.99	0.98	1.01	1.01	0.99	0.99
ϕ	0.08	0.04	0.09	0.13	0.05	0.05	0.13	0.14
x_0	30.63	30.58	30.65	30.65	26.00	25.94	25.96	25.82

PRIORI II: INFORMATIVA

El segundo análisis realizado utiliza, siguiendo el modelo de Branco et. al. (2000), una priori para x_0 normal, cuyos parámetros dependen de si la mujer está clasificada por Durmin como delgada o robusta, en cuyos casos los parámetros de las prioris para x_0 son respectivamente 22,57 y 16 ó 39,69 y 14,52. Estos valores corresponden a la priori II informativa presentada por dichos autores y que surgió de la construcción conjunta con el investigador a cargo del proyecto. Los resultados se presentan en la *Tabla 3.3*.

Tabla 3.3: Estimación bayesiana de los parámetros de interés para una mujer robusta y una mujer delgada (valores de ecuación de Durmin 33.25 y 26.98 respectivamente) para la priori informativa de Branco et. al. (2000), utilizando los modelos dependiente e independiente.

Modelo	Mujer robusta $y_0 = 33,25$				Mujer delgada $y_0 = 26,98$			
	Independiente		Dependiente		Independiente		Dependiente	
Parámetros	A	B	A	B	A	B	A	B
β_0	3.42	3.36	4.10	3.97	2.78	2.39	3.02	3.10
β_1	0,96	0,96	0.94	0.94	0.99	1.01	0.98	0.98
ϕ	0.07	0.04	0.08	0.13	0.08	0.05	0.09	0.13
x_0	35.91	35.93	36.01	35.96	23.47	23.56	23.45	23.56

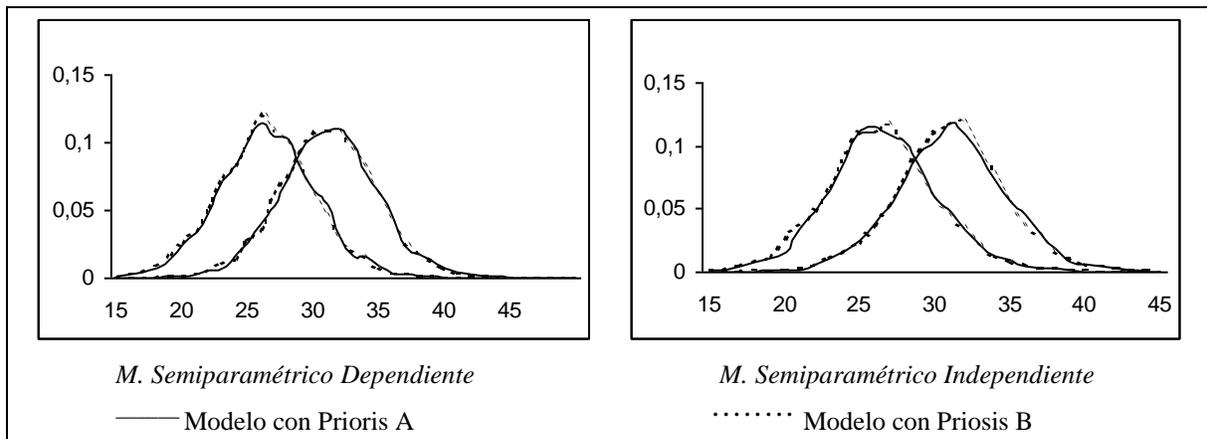
Tanto para la priori informativa como para el caso de la priori no informativa, se aprecia que en algunos casos las distribuciones a priori de los ω inciden en los estimadores, aunque en forma leve, pero es importante hacer notar que en general no inciden mayormente en la estimación del parámetro de interés x_0 , el cual se ve afectado, eso si, por el valor de y_0 (DEXA). En particular, los parámetros β_0 y ϕ se ven más afectados que β_1 , y es claro que para el valor de ecuación de Durmin correspondiente a la mujer robusta, hay mayor variación que para el caso de la mujer delgada. Por otra parte, los estimadores de x_0 no difieren mucho para los modelos dependiente e independiente, pero se aprecia que solamente el estimador de la pendiente permanece relativamente estable, ya que los demás estimadores muestran diferencias entre un modelo y otro. Ahora bien, las prioris para el parámetro x_0 (informativa y no informativa) si influyen de forma clara sobre la estimación.

Tabla 3.4: Cuadro comparativo de las estimaciones del parámetro x_0 mediante los modelos bayesianos semiparamétricos, paramétricos (Branco et. al., 2000) y clásicos para los datos en la Tabla 3.1.

Y_0	Clásico	Hoadley	Priori I No Informativa						Priori II Informativa				
			Independiente			Dependiente			Independiente			Dependiente	
			Branco	A	B	A	B	Branco	A	B	A	B	
26,98 (delgada)	28,1	29,3	26,00	25,9	26,0	25,8	26,0	23,5	23,6	23,4	23,6	23,5	
33,25 (robusta)	38,4	36,0	35,5	30,6	30,6	30,7	30,7	35,9	35,9	36,0	36,0	35,9	
ECM	27,19	17,65	16,96	9,64	9,63	9,82	9,80						

Entonces, al comparar las estimaciones de x_0 para nuestros modelos y las estimaciones de Branco et. al. (2000), podemos decir que con la priori I No Informativa, nuestras estimaciones coinciden con las de dichos autores para el caso de las mujeres delgadas, pero presenta diferencias para el caso de las mujeres de contextura gruesa. En el caso de la priori II informativa, las estimaciones son muy semejantes (Tabla 3.4.).

a) Comparación de distribuciones a posterioris de x_0 con Prioris A y B, para los modelos Semiparamétricos Elípticos (mujer delgada: izquierda, mujer robusta: derecha).



b) Comparación de distribuciones a posteriori de x_0 en los modelos Semiparamétricos Elípticos Dependiente e Independiente, con Prioris A y B (m. delgada: izquierda, m. robusta: derecha).

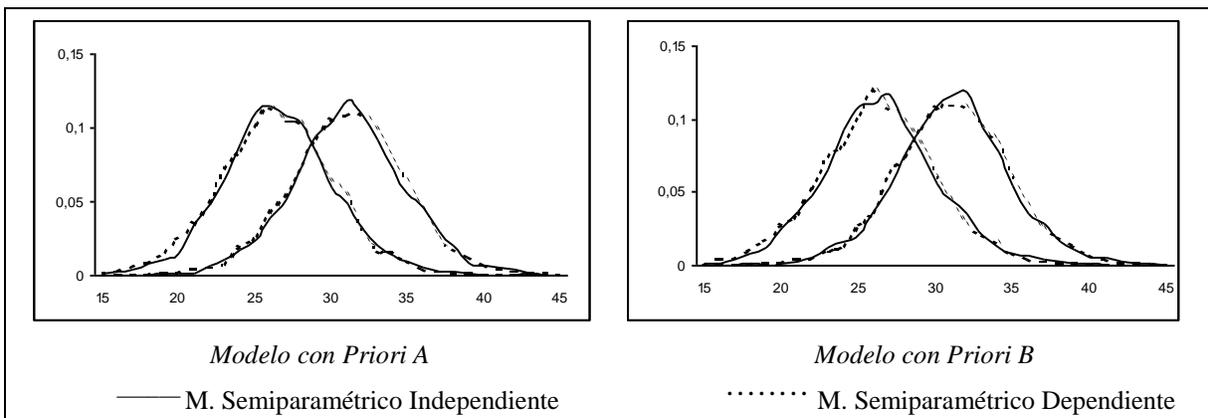


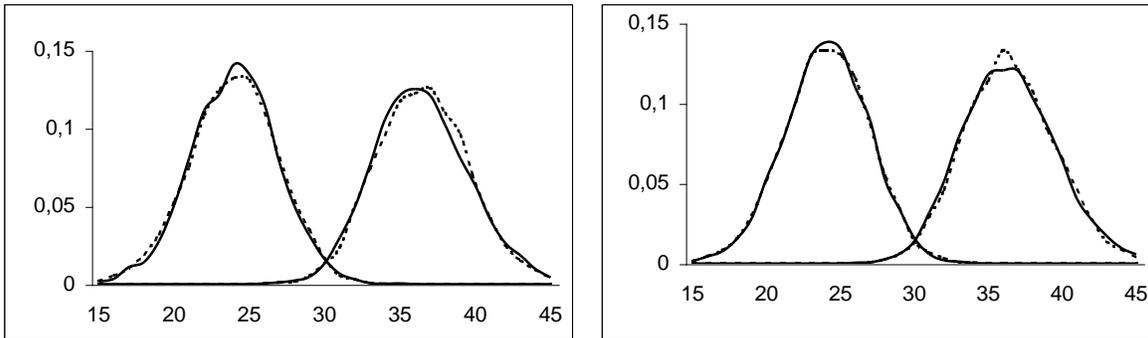
Figura 3.4. Distribuciones a posteriori para x_0 consideradas para prioti I No informativa ($c = 1$, fijo).

En fin, para el caso de la priori No informativa, se obtuvo para cada modelo el Error Cuadrático Medio ECM, en forma equivalente a la utilizada por Branco et. al. (2000), es decir, promediando las diferencias cuadráticas entre cada valor de x y su estimación, donde se estimó cada x como si dicho valor fuera x_0 desconocido, en función de un valor y_0 , que corresponde el valor de y respectivo. Estos valores se presentan en la *Tabla 3.4*. Es interesante comentar que nuestro ECM es significativamente inferior al obtenido por dichos autores, de modo que la extensión semiparamétrica mejora la estimación del parámetro de interés.

El cálculo de ECM no se puede replicar para el caso de la priori Informativa, ya que para dicho fin se requeriría una regla que permitiera clasificar a cada sujeto en estudio como de textura gruesa y delgada.

El análisis de las cadenas mediante test de Geweke reveló que existe convergencia, lo cual valida nuestros resultados. Por otra parte, la *Figura 3.2*. muestra los gráficos de las distribuciones a priori de las componentes ω del modelo (A y B), la *Figura 3.3*. presenta las distribuciones a priori para x_0 No informativa I e Informativa II y las *Figuras 3.4.* y *3.5.* presentan, respectivamente, las posteriors de x_0 para las prioris No informativa e Informativa.

a) Comparación de distribuciones a posteriors de x_0 con Prioris A y B, para los modelos Semiparamétricos Elípticos (mujer delgada: izquierda, mujer robusta: derecha).



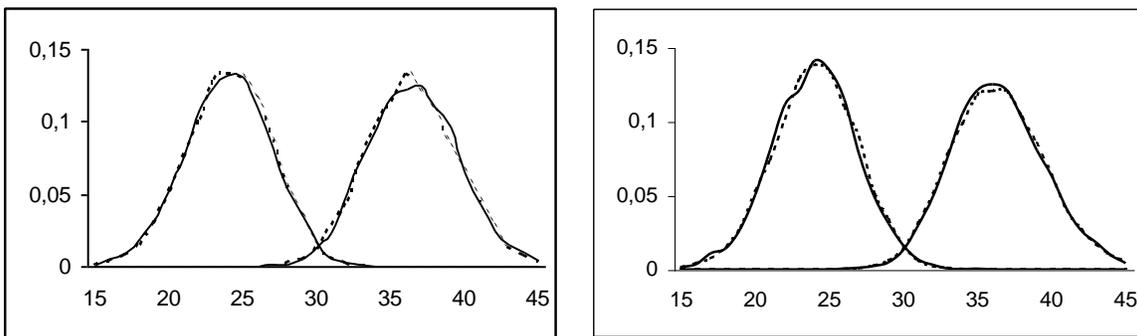
M. Semiparamétrico Dependiente

— Modelo con Prioris A

M. Semiparamétrico Independiente

..... Modelo con Prioris B

b) Comparación de distribuciones a posteriors de x_0 en los modelos Semiparamétricos Elípticos Dependiente e Independiente, con Prioris A y B (m. delgada: izquierda, m. robusta: derecha).



Modelo con Priori A

— M. Semiparamétrico Independiente

Modelo con Priori B

..... M. Semiparamétrico Dependiente

Figura 3.5. Distribuciones a posteriors para x_0 consideradas para prioris II Informativa.

Es claro desde las Figuras 3.3. y 3.4., que la priori no informativa, la cual es única para todas las mujeres en el estudio, redundante en posteriores diferentes según la contextura de cada mujer, y que la varianza se reduce notablemente. Por otra parte, es claro que las distribuciones a posteriori de x_0 no dependen ni de las prioris usadas (A y B, ver Figura 3.2.) ni de los Modelos utilizados para la estimación.

Esto confirma los resultados de la Tabla 3.5., donde se aprecia la semejanza notable en los estimadores de x_0 , aún cuando los estimadores de los otros parámetros presentan diferencias apreciables en algunos casos.

También cabe hacer notar que para las mujeres robustas las estimaciones con ambas Prioris difieren en forma notable, mientras que para las mujeres delgadas existe menor discrepancia entre las estimaciones.

No se aprecian diferencias entre las estimaciones provenientes de los Modelos Semiparamétricos Elípticos (dependiente e independiente) respecto de las Prioris I No Informativa y II Informativa.

La Tabla 3.4. confirma estos hallazgos y muestra que para las mujeres delgadas, la estimación con Priori I No Informativa se asemeja a la estimación de Branco et. al. (2000), mientras que para las mujeres robustas, la estimación de estos autores se asemeja más a nuestra estimación con Priori II Informativa.

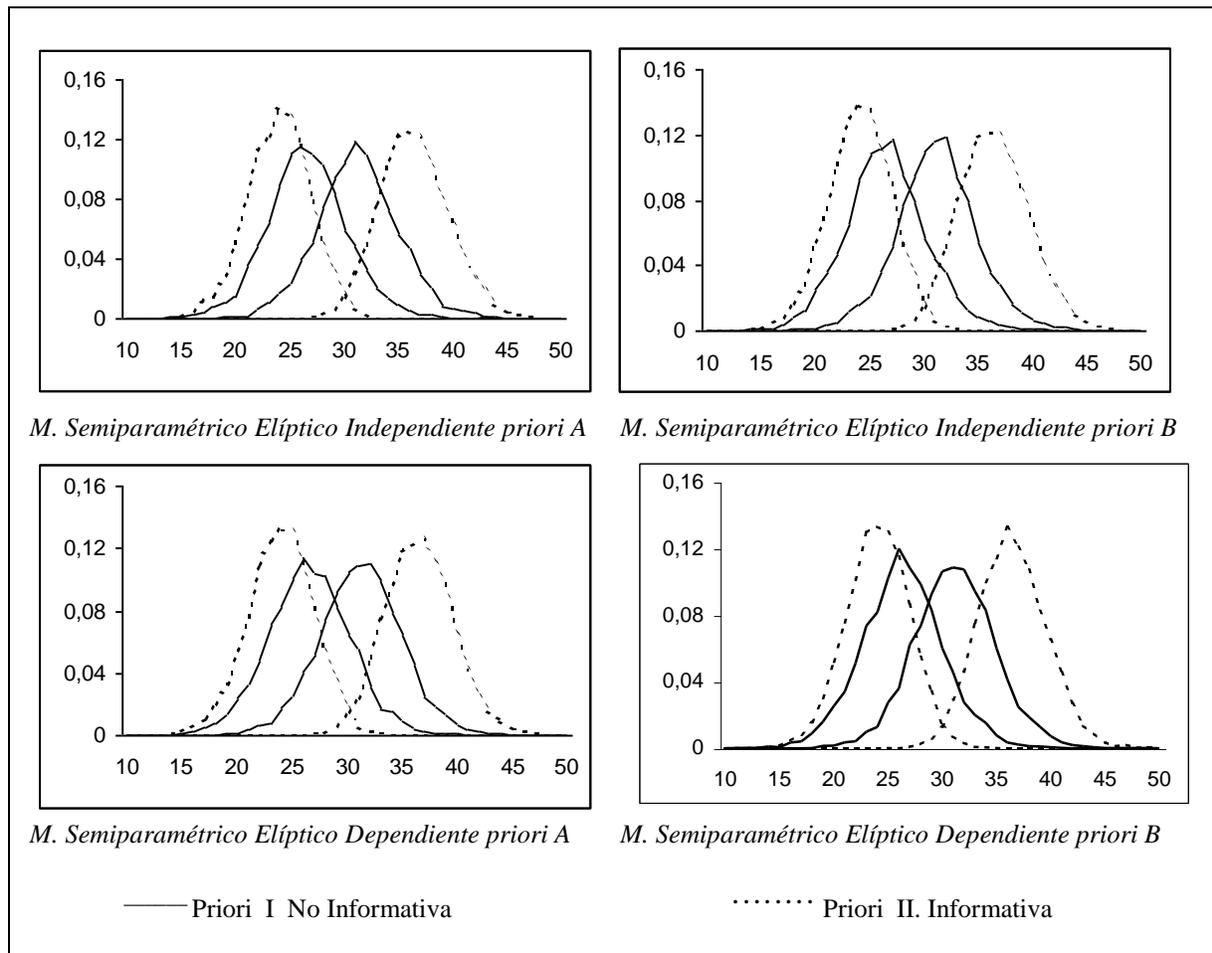


Figura 3.6. Comparación de distribuciones a Posteriores de x_0 para las Prioris I y II (para x_0), por modelos y por prioris relativas a ω (mujer delgada: izquierda, mujer robusta: derecha).

Ahora bien, en lo relativo a la Priori II Informativa, las distribuciones a priori y a posteriori de x_0 son muy semejantes, a excepción de que las prioris tienen varianza levemente mayor.

Para la Priori II Informativa (*Figura 3.5.*), al igual que en el caso de la Priori I, las posterioris son bastante gaussianas. Por otra parte, también en forma análoga a lo hallado para la Priori I, cabe destacar aquí los resultados de la *Tabla 3.3.* y confirmados gráficamente, a saber, que las posterioris para x_0 no dependen de las prioris para ω (A ó B), ni del modelo semiparamétrico elíptico utilizado para realizar la estimación (dependiente o independiente), aún cuando los estimadores de los coeficientes difieren en forma notable.

También es claro, desde la *Tabla 3.4.*, que las diferencias de estimación si dependen fuertemente de la Priori para x_0 , No Informativa (I) o Informativa (II), como queda de manifiesto en la *Figura 3.6.*

En conclusión, hemos encontrado que las estimaciones con la Priori I No Informativa son más variables. Para mujeres delgadas, la Priori I arroja estimaciones mayores a las de la Priori II, mientras que para mujeres robustas es a la inversa.

Aplicación 3.2 Mediciones de un tipo de encima en la sangre (Eno, 1999).

Aitchison & Dunsmore (1975) presentan un conjunto de datos relativos a la medición de cierta encima en la sangre, para lo cual el análisis de laboratorio es muy caro y demoroso, pero existe un método alternativo de autoanálisis muy rápido y económico. El autoanálisis se realiza con un instrumento no muy exacto y que debe ser calibrado.

Tabla 3.5. Mediciones de encima en la sangre. La medición x corresponde al resultado de laboratorio, mientras que la medición y corresponde a los valores obtenidos mediante autoanálisis.

Medición de laboratorio	Autoanálisis
3,0	2.3, 2.4, 2.5
3,4	2.6, 2.8, 2.8
3,8	3.0, 3.0, 3.1
4,2	3.2, 3.3, 3.4
4,6	3.7, 3.7, 3.7
5,0	3.9, 4.0, 4.1
5,4	4.2, 4.2, 4.3
5,8	4.6, 4.7, 4.8
6,2	4.9, 5.0, 5.2

Los investigadores que diseñaron el experimento escogieron 9 niveles de la encima, para cada uno de los cuales se obtuvieron tres muestras que presentaban dicha concentración de acuerdo con el análisis de laboratorio y para cada muestra, se realizó el autoanálisis. Los datos se presentan en la *Tabla 3.5.* y el diagrama de dispersión en la *Figura 3.7.*

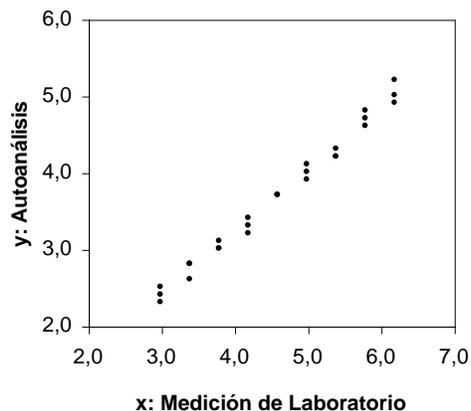


Figura 3.7. Mediciones de encima en la sangre, resultados de laboratorio y autoanálisis.

Eno (1999) en su tesis doctoral, asume que la medición mediante autoanálisis (y) depende linealmente de la medición de laboratorio (x). Este autor escoge dos valores de y_0 en los extremos del conjunto (3.7 y 5.2) y estima el verdadero valor de la medición de laboratorio x_0 . La *Tabla 3.6.* presenta el análisis bayesiano paramétrico realizado por este autor, el cual considera errores normales y prioris de referencia; también presenta los resultados que surgen de utilizar la priori de Jefreys, la priori de Hoadley y enfoque clásico.

Tabla 3.6. Estimaciones bayesianas semiparamétricas utilizando modelos dependiente e independiente, para el conjunto de datos en la Tabla 3.5., junto a los resultados presentados por Eno (1999).

Parámetro: x_0		$y_0 = 3.7$			$y_0 = 5.2$		
		Estimador	L Inf.	L Sup.	Estimador	L Inf.	L Sup.
M. dependiente	Priori Centrada (4 g.l.)	4.6221	4.3827	4.8640	6.4674	6.2099	6.7261
	Priori No Cent.(4 g.l.)	4.6270	4.3925	4.8682	6.4644	6.2144	6.7042
	Prori Centrada (30 g.l.)	4.6215	4.3869	4.8570	6.4652	6.2149	6.7074
	Priori No Cent. (30 g.l.)	4.6239	4.4158	4.8205	6.4644	6.2536	6.6806
	Prori Centrada (50 g.l.)	4.6242	4.3755	4.8498	6.4652	6.2142	6.7067
	Priori No Cent. (50 g.l.)	4.6268	4.4383	4.8309	6.4634	6.2524	6.6757
M. independiente	Priori Centrada (4 g.l.)	4.6235	4.3710	4.8823	6.4773	6.2140	6.7560
	Priori No Cent.(4 g.l.)	4.6308	4.3986	4.8818	6.4723	6.2200	6.7222
	Prori Centrada (30 g.l.)	4.6201	4.3771	4.8595	6.4728	6.2385	6.7319
	Priori No Cent. (30 g.l.)	4.6265	4.4354	4.8268	6.4668	6.2729	6.6856
	Prori Centrada (50 g.l.)	4.6237	4.4044	4.8662	6.4667	6.2167	6.7042
	Priori No Cent. (50 g.l.)	4.6280	4.4402	4.8307	6.4696	6.2661	6.6733
Eno (1999)	Priori de Hoadley	4.6225	4.4217	4.8233	6.4442	6.2322	6.6562
	P. Rerencia k = 1 de Eno	4.6227	4.4207	4.8249	6.4640	6.2547	6.6819
	P. Rerencia k = 2 de Eno	4.6227	4.4249	4.8210	6.4641	6.2591	6.6773
	P. Rerencia k =3 de Eno	4.6227	4.4288	4.8107	6.4642	6.2632	6.6562
	P. Jeffreys	4.6210	4.4300	4.8220	6.4640	6.2460	6.6850
	Enfoque clásico	4.6227	4.3534	4.9047	6.4657	6.1286	6.8206

Nuestro análisis, en cambio, considera prioris que varían los grados de libertad ($\nu = 4, 30, 50$) relativos a la distribución de los errores (t – student). Para utilizar estas especificaciones, se ha considerado una distribución a priori para ω Gama Inversa ($1/2 \nu, 1/2 \nu$).

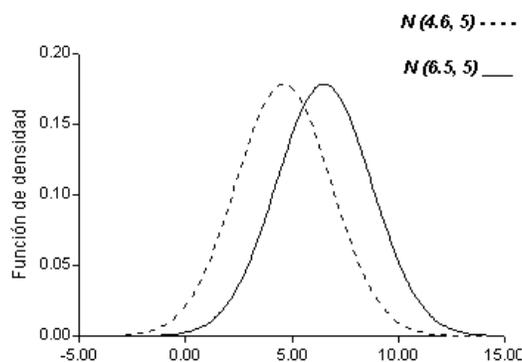


Figura 3.8. Distribuciones a priori para el parámetro de interés x_0 .

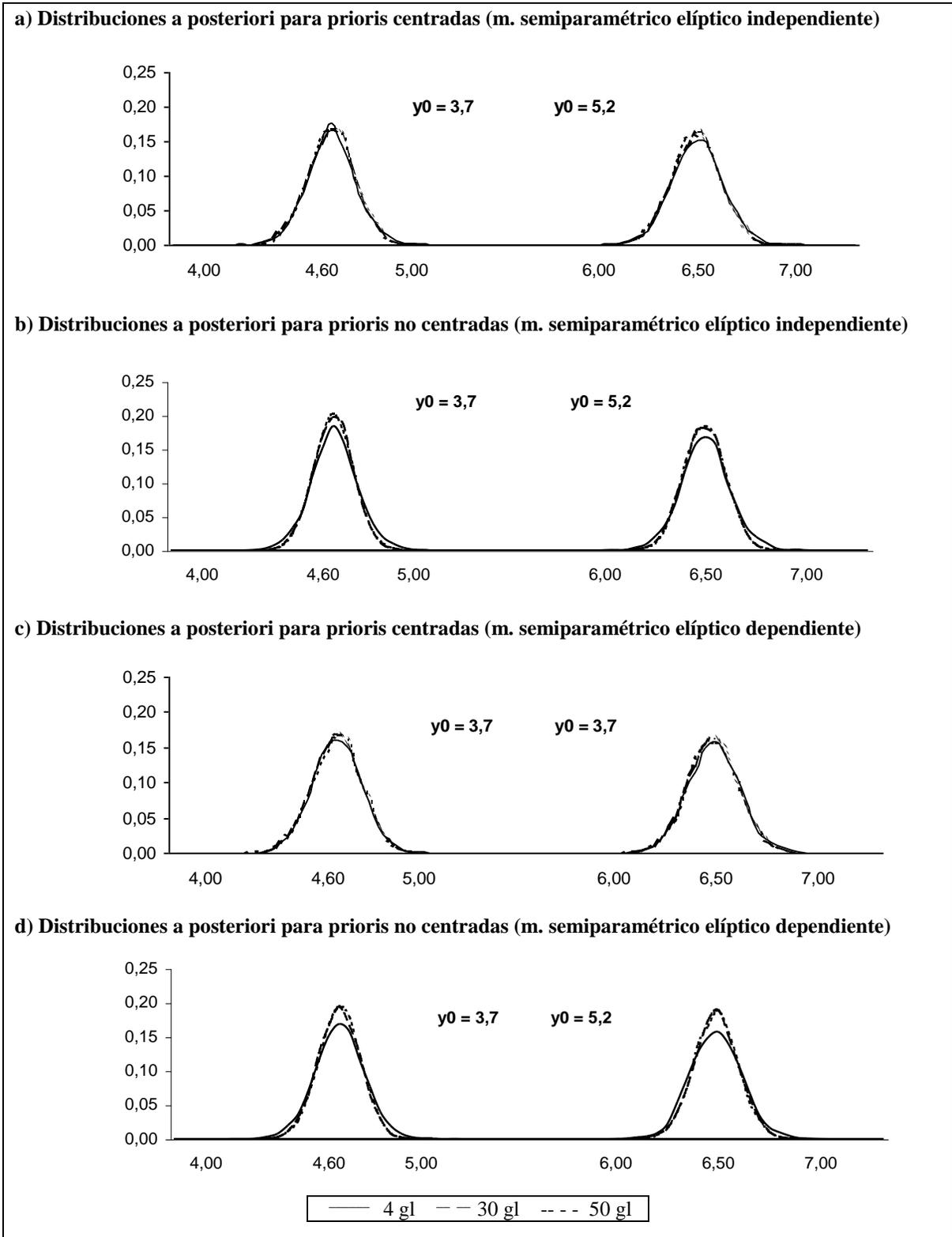


Figura 3.9. Comparación de distribuciones a posteriori del parámetro de interés x_0 , según grados de libertad, en modelos semiparamétricos elípticos.

Por otra parte, se incluyeron en nuestro análisis dos variantes para las prioris de los parámetros, la primera (centrada) considera estimadores máximo verosímiles para las medias de las distribuciones a priori, mientras que la segunda (no centrada) considera medias diferentes. Para ambos casos las prioris se consideran no informativas, ya que las varianzas se escogieron relativamente grandes.

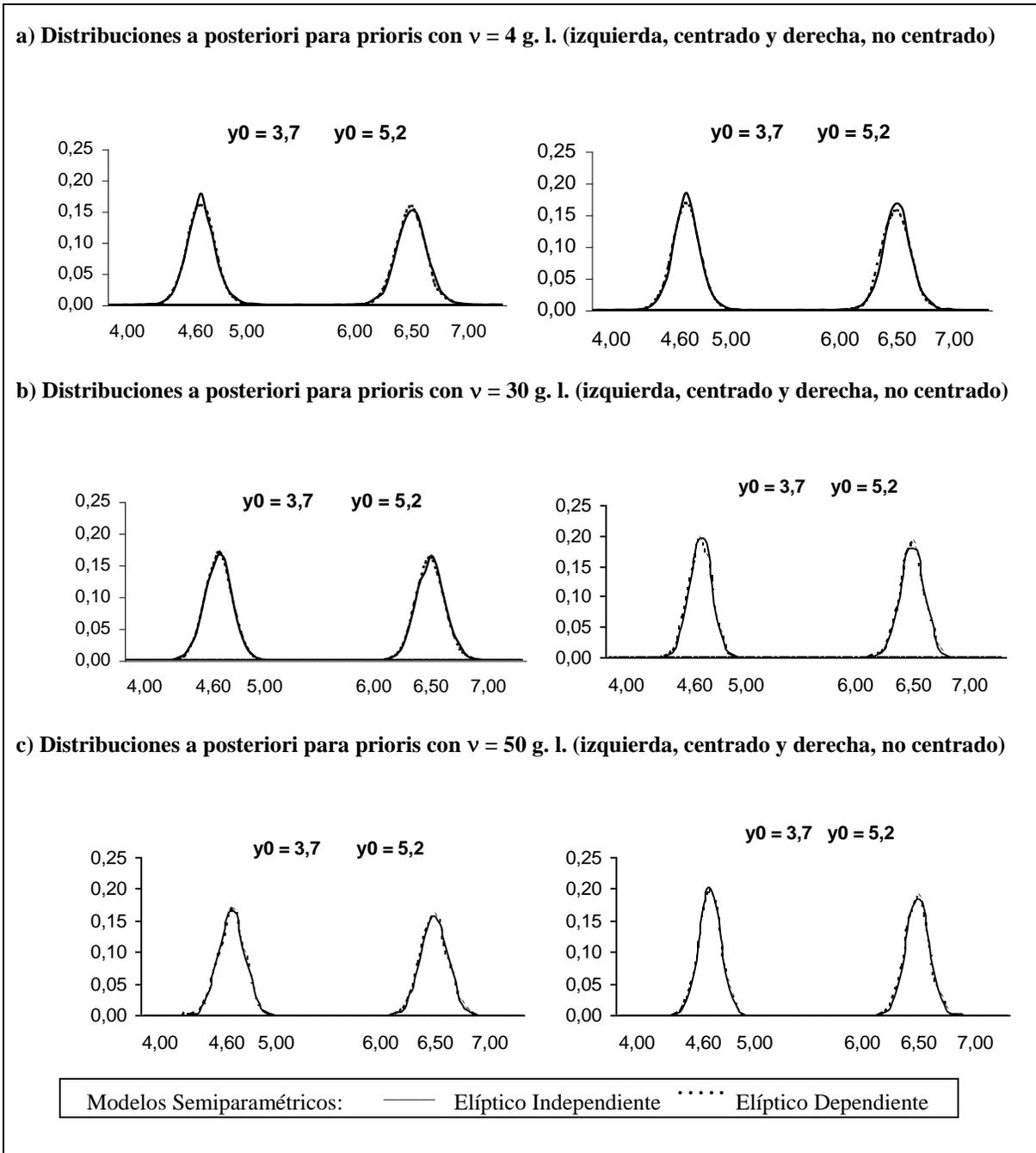


Figura 3.10. Comparación de distribuciones a posteriori del parámetro de interés x_0 , según modelo semiparamétrico elíptico, por grados de libertad.

Como muestra la *Figura 3.8.*, para el parámetro de interés, x_0 , las prioris son normales con varianza 5. Para el caso en que y_0 es 3.7, la priori centrada tiene como media el estimador clásico (EMV) 4.6, mientras que para y_0 5.2, el estimador que cumple con esa característica es 6.5. En este caso, se consideró la priori no centrada con medias contrarias, es decir, 6.5 y 4.6, respectivamente.

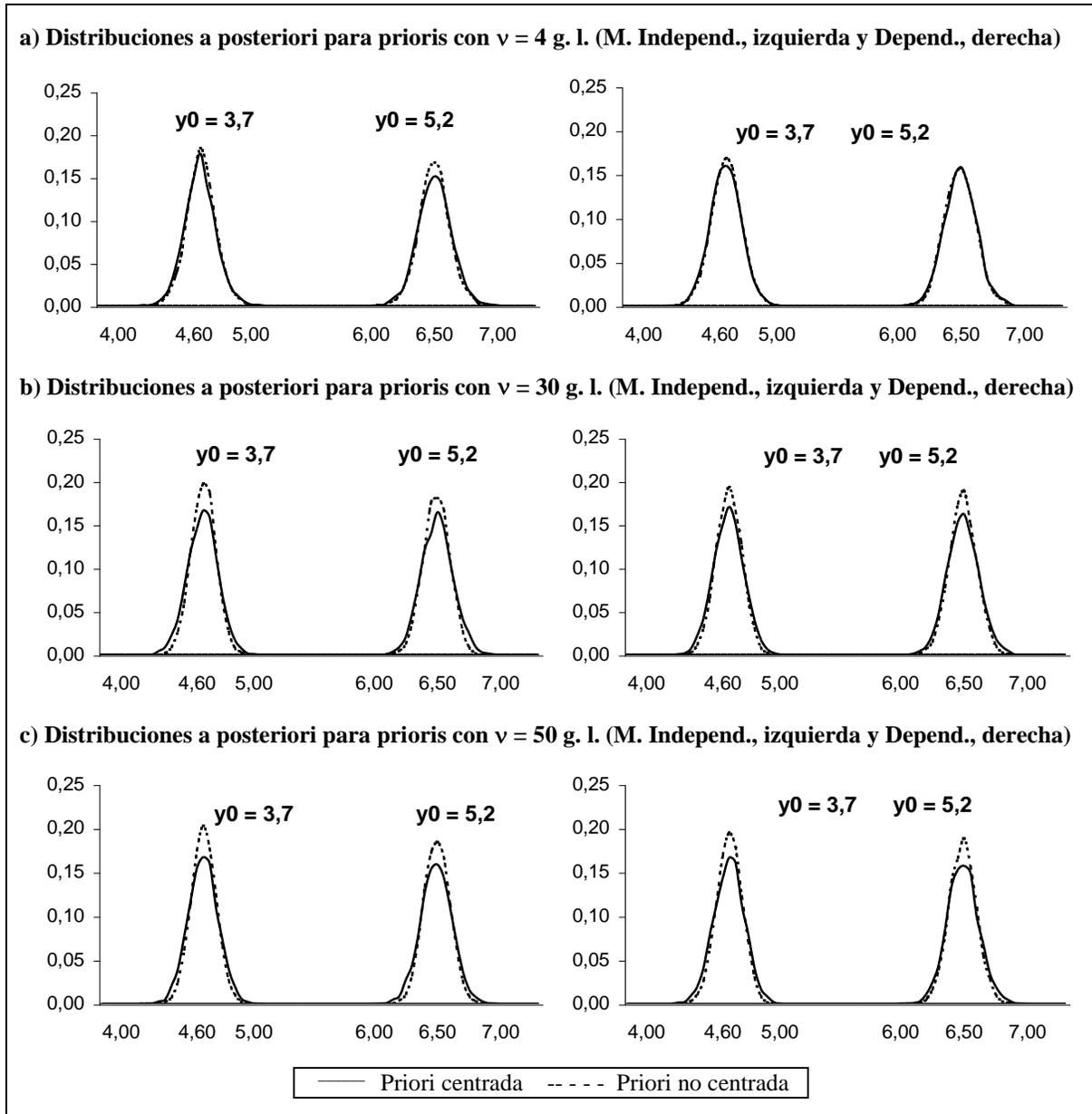


Figura 3.11. Comparación de distribuciones a posteriori del parámetro de interés x_0 , según priori centrada o no centrada, por grados de libertad.

Para el parámetro β_0 , la priori es normal con varianza 1, y se consideran dos medias a priori, la media centrada (el estimador de mínimos cuadrados -0.0624) y la media centrada (asumiendo el valor 1); análogamente, para el parámetro β_1 la priori también es normal, su varianza es 10, y las medias son, respectivamente, el estimador de mínimos cuadrados 0.8139 (media centrada) y -1 (media no centrada).

Los valores de ν considerados para nuestro estudio son 4, 30 y 50, y para el parámetro ϕ se asume una priori Gamma Inversa, tal que su coeficiente de variación sea de 17% y la media para $\sigma^2 = \phi^{-1}$ sea, en el caso centrado, aproximadamente igual 0.0088, y en el caso no centrado 0.044. esto nos lleva a que en ambos casos el parámetro de forma es $r_0 = 35$, pero el parámetro de escala asume valores respectivos $\lambda_0 = 0.3$ y $\lambda_0 = 0.15$.

Para la semilla de x_0 se usa una normal con parámetros iguales a la media y varianza muestral de x , 4.6 y 1.1077, respectivamente.

Las estimaciones y los intervalos de credibilidad se presentan en la *Tabla 3.6.*, junto a los resultados presentados por Eno. Se ve claramente que nuestras estimaciones no están lejos de las presentadas por dicho autor, tanto para el modelo dependiente como para el independiente. Por otro lado, la longitud de nuestros intervalos es inferior para las prioris no centradas, y se observa que las longitudes son levemente superiores a las de los intervalos relativos a prioris de referencia. Esto se debe a que nuestro modelo es más flexible que el utilizado por Eno.

Por otro lado, el test de Geweke revela que existe convergencia de las cadenas, lo cual valida nuestros resultados. Ahora bien, las *Figura 3.9.*, *3.10.* y *3.11.* muestran las distribuciones a posteriori obtenidas para cada modelo. Algunas conclusiones que se derivan de estos gráficos son las siguientes:

- Todas las distribuciones a posteriori para el parámetro de interés x_0 presentan una forma gaussiana.
- Tanto para el modelo semiparamétrico elíptico dependiente como para el análogo independiente, si las prioris son centradas los grados de libertad no inciden en la forma de las posterioris (las medias y varianzas se asemejan); sin embargo, para prioris no centradas, la distribución a posteriori de x_0 resulta más dispersa si los grados de libertad son menores (*Figura 3.9.*).
- Las distribuciones a posteriori para x_0 en modelos semiparamétricos elípticos son semejantes para los casos dependiente e independiente (*Figura 3.10.*).
- Las distribuciones a posteriori para x_0 son más dispersas si las prioris son centradas, esto es más notable a medida que los grados de libertad aumentan (*Figura 3.11.*).

Aplicación 3.3: Simulación análoga a Eno (1999).

Eno (1999) presenta una simulación para el modelo de calibración, la cual tiene como objetivo verificar si se puede determinar con precisión los cuantiles de la distribución a posteriori de x_0 . Este autor considera el modelo con errores normales y utiliza prioris de referencia para la estimación del parámetro de interés (ver comentarios al respecto en *Aplicación 3.2.*).

Nosotros hemos replicado esta simulación con la finalidad de validar nuestra metodología y de determinar el grado de precisión de los estimadores en función de ciertas variables, como son el tamaño muestral, la distribución de los errores, la magnitud de la pendiente y la media a priori del parámetro estudiado.

En este sentido, los tamaños de muestra asumen valores $m = 11$, $m = 31$ y $m = 101$. En una primera etapa, cada tamaño de muestra definió 24 simulaciones. Para cada simulación se escogieron un total de $n = m - 1$ observaciones de la predictora x equiespaciadas en el intervalo $[-1, 1]$ y la última observación, que corresponde a x_0 , se fija.

Los valores de β_0 y de ϕ son fijos e igual a cero y a uno, respectivamente. Las diferentes simulaciones resultan de variar x_0 (asumiendo valores 0, 0.5 y 1), β_1 (que toma los valores 0.1, 1, 5 y 10, con el fin de determinar si existe efecto de la pendiente sobre la estimación), y la distribución de los errores (que pueden ser la normal y la t – student con cuatro grados de libertad). Esto permite la generación aleatoria de m valores de la variable dependiente y como función de las componentes mencionadas anteriormente.

En una segunda etapa, los datos simulados para el modelo de calibración continua, donde los valores de x_0 se asumen desconocidos y se estiman, fueron analizados utilizando el modelo semiparamétrico elíptico dependiente e independiente. Este modelo utilizó como priori para β_0 una normal con media cero y varianza 0.5, para β_1 una normal con las medias reales y varianza 4, para x_0 una normal con media 0.5 y varianza 1 (que resulta más informativa justamente para el valor 0.5), para ϕ una gamma de parámetros 4 y 4 (centrada en el valor real 1 y con varianza 0.125) y para ω un proceso Dirichlet con distribución base gamma inversa de parámetros 2 y 2 (vinculado al modelo de cuatro grados de libertad).

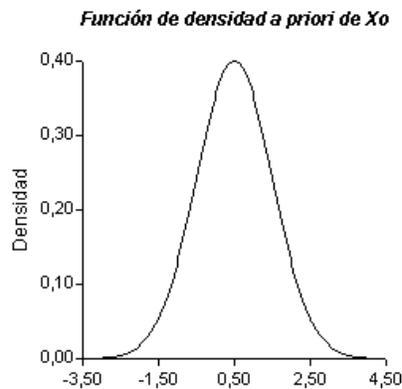


Figura 3.7. Distribución a priori del parámetro de interés x_0 .

Para la estimación de parámetros, se realizó un muestreo de Gibbs. Se realizaron 70.000 iteraciones descartando las 10.000 iniciales y muestreando una de cada 10. Las cadenas fueron analizadas mediante test de Geweke y se encuentra convergencia, lo cual valida los resultados. Las estimaciones del parámetro de interés para los modelos dependiente e independiente se presentan, respectivamente, en las *Tablas 3.7. y 3.8.*, y las gráficas de las distribuciones a posteriori para el parámetro de interés x_0 se presentan en las *Figuras 3.8. a 3.11.*

Tabla 3.7. *Análisis de los datos provenientes de una Simulación para el modelo de calibración continua, mediante el modelo Semiparamétrico elíptico dependiente.*

Tamaño de muestra n = 10	Errores Normales			Errores t – student (4)		
β_1	$x_0 = 0$	$x_0 = 0.5$	$x_0 = 1$	$x_0 = 0$	$x_0 = 0.5$	$x_0 = 1$
0.1	0.7037	0.6795	0.5377	0.6698	0.6561	-0.6563
1	0.0089	0.4789	0.4262	-1.4876	-1.4126	-1.2760
5	-0.2501	0.2576	0.9483	-0.8498	-0.3552	0.1173
10	-0.1536	0.4265	0.9365	-0.4684	0.0523	0.5473

Tamaño de muestra n = 30	Errores Normales			Errores t – student (4)		
β_1	$x_0 = 0$	$x_0 = 0.5$	$x_0 = 1$	$x_0 = 0$	$x_0 = 0.5$	$x_0 = 1$
0.1	0.4397	0.4152	0.3866	0.4909	0.5228	0.5375
1	0.0789	0.3194	1.3571	0.5213	0.7668	0.9801
5	0.3886	0.8436	1.3923	0.1491	0.6081	1.1385
10	0.1949	0.7018	1.1925	0.0715	0.5652	1.0384

Tamaño de muestra n = 100	Errores Normales			Errores t – student (4)		
β_1	$x_0 = 0$	$x_0 = 0.5$	$x_0 = 1$	$x_0 = 0$	$x_0 = 0.5$	$x_0 = 1$
0.1	0.4049	0.4017	0.3935	0.4717	0.4549	0.4893
1	0.5061	0.6961	0.8836	0.5649	0.6595	0.7693
5	0.1008	0.6035	1.1250	0.2389	0.6799	1.1610
10	0.0417	0.5578	1.0645	0.1125	0.6128	1.1007

Tabla 3.8. Análisis mediante el modelo independiente de los datos provenientes de una Simulación para el modelo de calibración.

Tamaño de muestra n = 10	Errores Normales			Errores t – student (4)		
β_1	$x_0 = 0$	$x_0 = 0.5$	$x_0 = 1$	$x_0 = 0$	$x_0 = 0.5$	$x_0 = 1$
0.1	0.7228	0.7103	0.5822	0.6737	0.6531	-0.3683
1	0.0091	0.5351	0.4768	-1.3358	-1.0531	-1.0047
5	-0.2599	0.2692	0.9676	-0.8244	-0.3363	0.1317
10	-0.1736	0.4319	0.9426	-0.4615	0.0585	0.5490

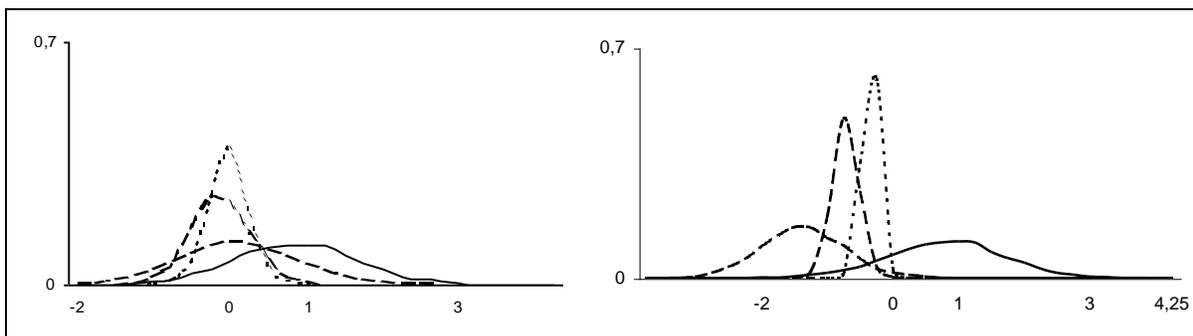
Tamaño de muestra n = 30	Errores Normales			Errores t – student (4)		
β_1	$x_0 = 0$	$x_0 = 0.5$	$x_0 = 1$	$x_0 = 0$	$x_0 = 0.5$	$x_0 = 1$
0.1	0.3649	0.3815	0.3501	0.4900	0.4635	0.5390
1	0.0689	0.2884	1.3365	0.5249	0.8017	1.0370
5	0.3846	0.8565	1.4007	0.1390	0.6087	1.1332
10	0.1939	0.7022	1.1943	0.0662	0.5686	1.0438

Tamaño de muestra n = 100	Errores Normales			Errores t – student (4)		
β_1	$x_0 = 0$	$x_0 = 0.5$	$x_0 = 1$	$x_0 = 0$	$x_0 = 0.5$	$x_0 = 1$
0.1	0.3907	0.4011	0.4198	0.4777	0.4737	0.5204
1	0.5003	0.6975	0.9078	0.6124	0.7543	0.8628
5	0.0974	0.5913	1.1225	0.2205	0.6727	1.1551
10	0.0432	0.5549	1.0637	0.1110	0.6046	1.0917

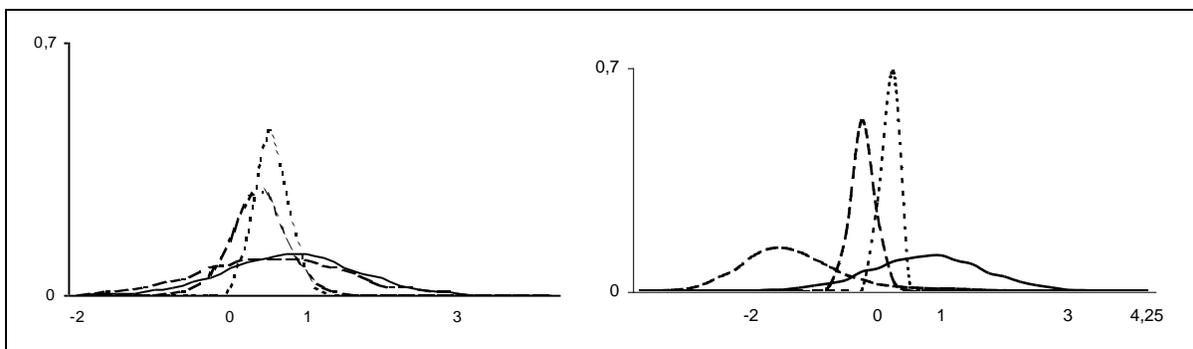
Resulta claro, a partir de las *Tablas 3.7. y 3.8.* que:

- Los estimadores obtenidos mediante el modelo dependiente son muy semejantes a los que proceden del análisis con el modelo independiente (mediante el análisis bayesiano), por otra parte, los valores de β_1 cercanos a cero hacen que la distribución a priori resulte muy influyente sobre la distribución a posteriori, de modo que los estimadores están alrededor del valor 0.5.
- También se puede decir que para tamaños de muestra mayor, se logran estimaciones mejores cuando β_1 se aleja un tanto del cero (valores de β_1 5 y 10 en nuestro caso).
- Por otra parte, para x_0 alejándose de cero (valor uno en nuestro caso) se logran mejores estimaciones, aunque la priori para el parámetro de interés esté descentrada.
- Finalmente, se puede ver que para los errores normales, que corresponden a un modelo tanto dependiente como independiente, ambos modelos resultan igualmente buenos en lo relativo al análisis; por otra parte, para los errores t – student, se ha simulado un modelo independiente, sin embargo, tanto el modelo dependiente como el independiente presentan estimaciones con características semejantes.

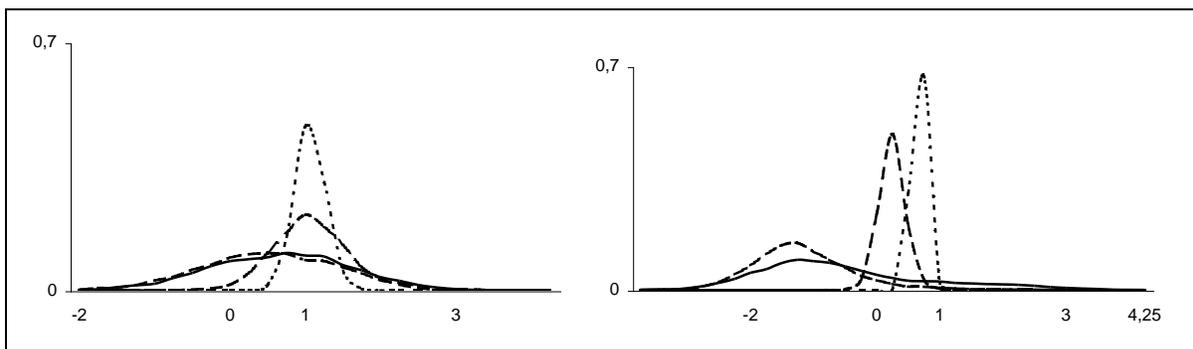
a) De izquierda a derecha, modelo de calibración continua con errores normales y con errores t – student, ambos analizados con modelo semiparamétrico elíptico dependiente con 4 g. l. para $x_0 = 0$.



b) De izquierda a derecha, modelo de calibración continua con errores normales y con errores t – student, ambos analizados con modelo semiparamétrico elíptico dependiente con 4 g. l. para $x_0 = 0.5$.



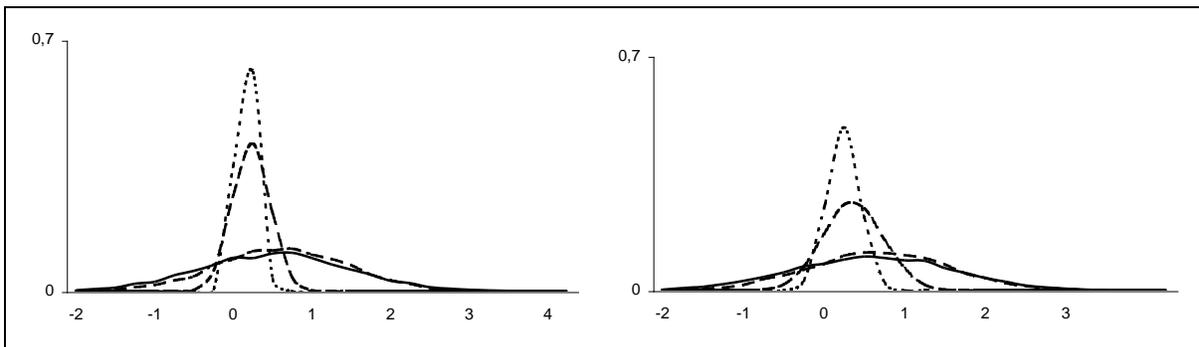
c) De izquierda a derecha, modelo de calibración continua con errores normales y con errores t – student, ambos analizados con modelo semiparamétrico elíptico dependiente con 4 g. l. para $x_0 = 1$.



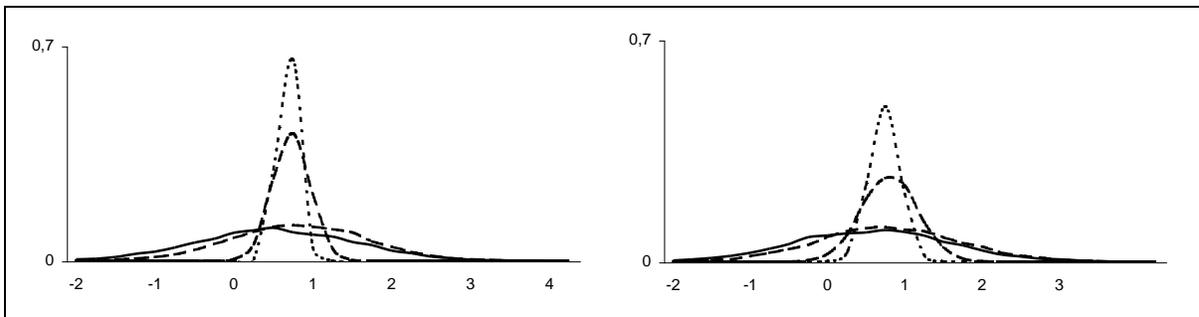
Pendiente: ————— $\beta_1 = 0.1$ - - - - - $\beta_1 = 1$ - · - · - $\beta_1 = 5$ · · · · · $\beta_1 = 10$

Figura 3.8. Distribución a posteriori del parámetro de interés x_0 para el modelo de calibración continua con errores normales y t – student, analizados con el modelo semiparamétrico elíptico dependiente ($n = 10$).

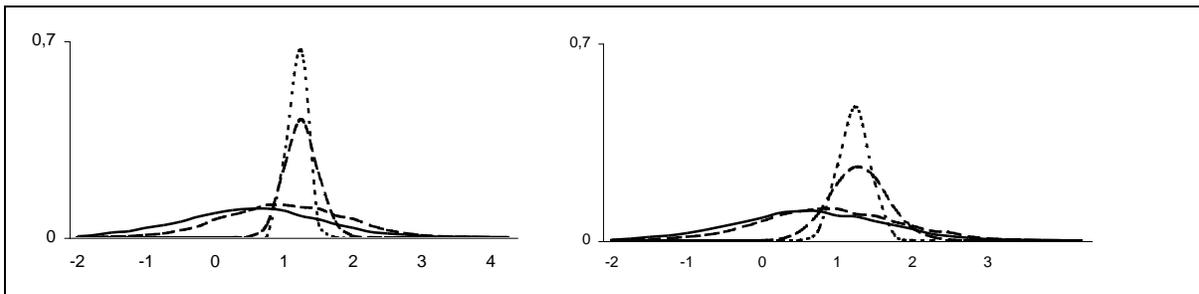
a) De izquierda a derecha, modelo de calibración continua con errores normales y con errores t – student, ambos analizados con modelo semiparamétrico elíptico dependiente con 4 g. l. para $x_0 = 0$.



b) De izquierda a derecha, modelo de calibración continua con errores normales y con errores t – student, ambos analizados con modelo semiparamétrico elíptico dependiente con 4 g. l. para $x_0 = 0.5$.



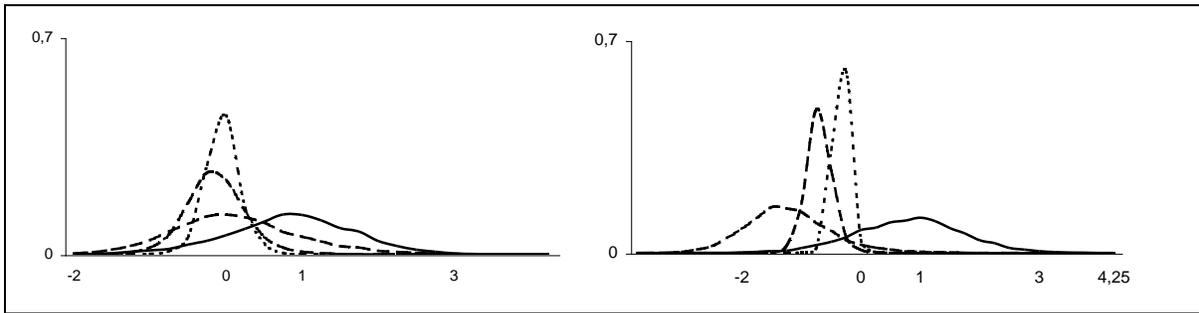
c) De izquierda a derecha, modelo de calibración continua con errores normales y con errores t – student, ambos analizados con modelo semiparamétrico elíptico dependiente con 4 g. l. para $x_0 = 1$.



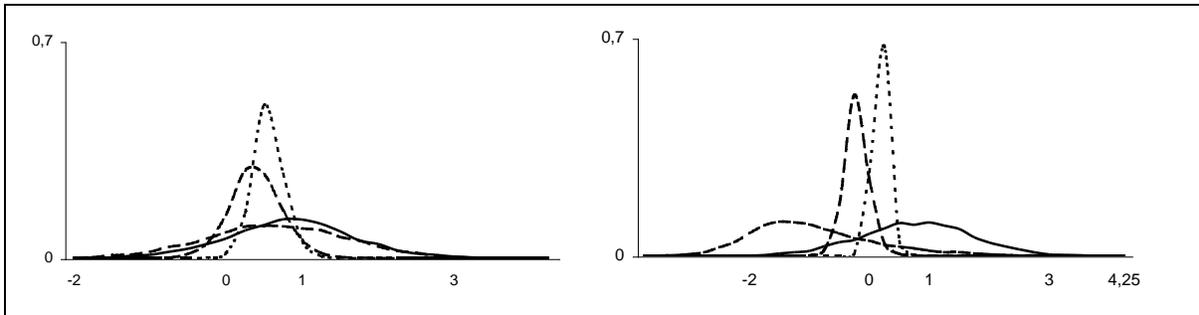
Pendiente: — $\beta_1 = 0.1$ - - - $\beta_1 = 1$ - - - $\beta_1 = 5$ - - - - $\beta_1 = 10$

Figura 3.9. Distribución a posteriori del parámetro de interés x_0 para el modelo de calibración continua con errores normales y t – student, analizados con el modelo semiparamétrico elíptico dependiente ($n = 100$).

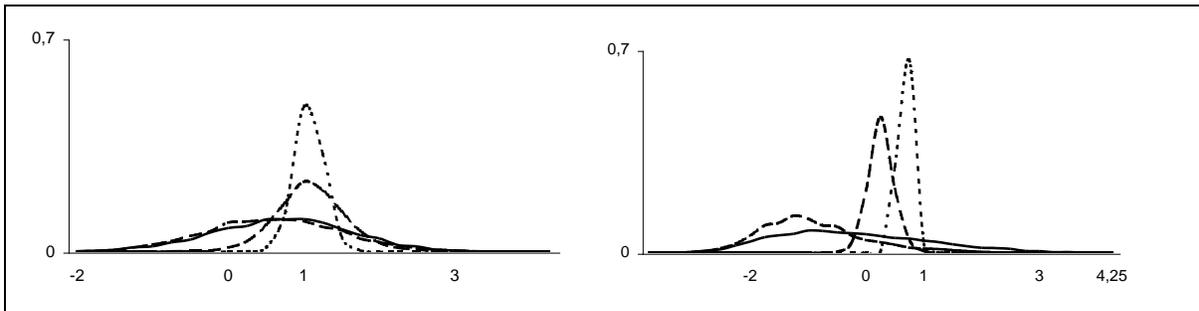
a) De izquierda a derecha, modelo de calibración continua con errores normales y con errores t – student, ambos analizados con modelo semiparamétrico elíptico independiente con 4 g. l. para $x_0 = 0$.



b) De izquierda a derecha, modelo de calibración continua con errores normales y con errores t – student, ambos analizados con modelo semiparamétrico elíptico independiente con 4 g. l. para $x_0 = 0.5$.



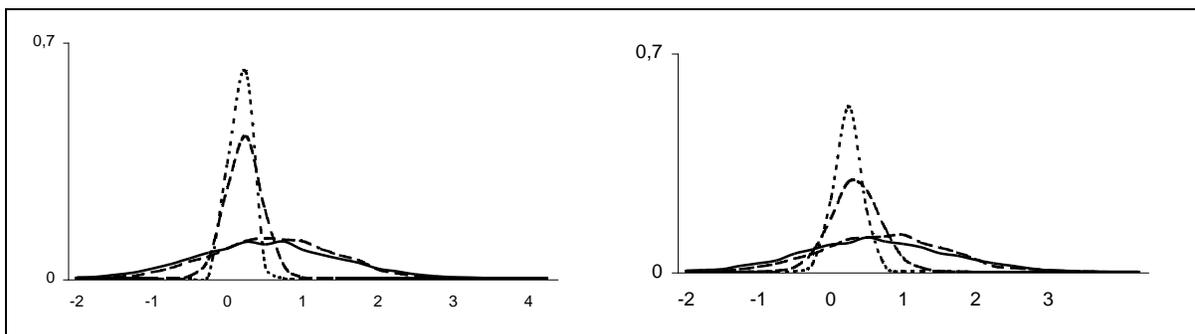
c) De izquierda a derecha, modelo de calibración continua con errores normales y con errores t – student, ambos analizados con modelo semiparamétrico elíptico independiente con 4 g. l. para $x_0 = 1$.



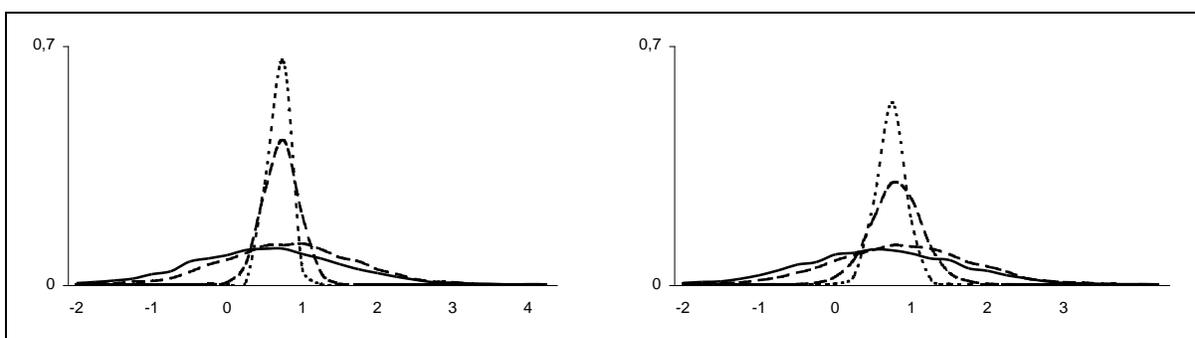
Pendiente: ——— $\beta_1 = 0.1$ - - - - $\beta_1 = 1$ - - - - $\beta_1 = 5$ - - - - $\beta_1 = 10$

Figura 3.10. Distribución a posteriori del parámetro de interés x_0 para el modelo de calibración continua con errores normales y t – student, analizados con modelo semiparamétrico elíptico independiente ($n = 10$).

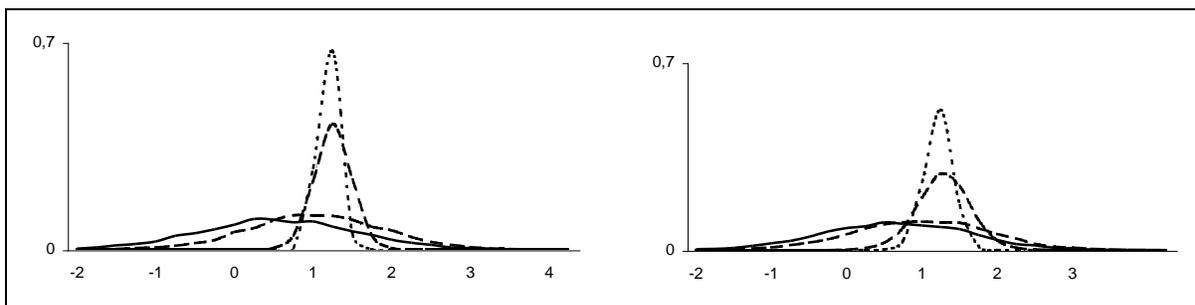
a) De izquierda a derecha, modelo de calibración continua con errores normales y con errores t – student, ambos analizados con modelo semiparamétrico elíptico independiente con 4 g. l. para $x_0 = 0$.



b) De izquierda a derecha, modelo de calibración continua con errores normales y con errores t – student, ambos analizados con modelo semiparamétrico elíptico independiente con 4 g. l. para $x_0 = 0.5$.



c) De izquierda a derecha, modelo de calibración continua con errores normales y con errores t – student, ambos analizados con modelo semiparamétrico elíptico independiente con 4 g. l. para $x_0 = 1$.



Pendiente: ————— $\beta_1 = 0.1$ - - - - $\beta_1 = 1$ - · - · $\beta_1 = 5$ · · · · · $\beta_1 = 10$

Figura 3.11. Distribución a posteriori del parámetro de interés x_0 para el modelo de calibración continua con errores normales y t – student, analizados con modelo semiparamétrico elíptico independiente ($n = 100$).

Al observar las distribuciones a priori (*Figura 3.7.*) y a posteriori (*Figuras 3.8. a la 3.11.*) del parámetro x_0 se puede apreciar que, al realizar un análisis con el modelo semiparamétrico elíptico:

- Las distribuciones a posteriori para x_0 son prácticamente iguales para los análisis con modelo semiparamétrico elíptico independiente (*Figuras 3.10. y 3.11.*) y dependiente (*Figuras 3.8. y 3.9.*), pero las primeras son levemente menos dispersas.
- Todas las distribuciones a posteriori para x_0 tienen menor varianza que la distribución a priori.
- Las distribuciones a posteriori del parámetro de interés para modelos con pendiente menor tienen mayor varianza y viceversa.
- Las distribuciones a posteriori de x_0 para modelos con pendiente menor en general son sesgadas, y a medida que la pendiente aumenta se estima con menor sesgo.
- Para tamaño de muestra menor la priori de x_0 tiende a ser más sesgada. En este caso se ve la influencia de la media de la distribución a priori para x_0 . Cuando crece el tamaño muestral, la media a priori ya no influye.
- Para el modelo simulado con errores normales las distribuciones a posteriori se hacen menos variables al aumentar el tamaño muestral. Esto no es así en el caso del modelo simulado con errores t – student, para el cual las varianzas aumentan.
- El modelo simulado con errores t – student resulta tener distribuciones a posteriori para x_0 más sesgadas para muestras pequeñas que el modelo simulado con errores normales.

CONCLUSIONES

Este capítulo aborda el análisis bayesiano semiparamétrico del modelo de calibración lineal continua.

Previo a ello, y con el fin de justificar este enfoque, se obtuvo la función de verosimilitud asociada al problema, así como las ecuaciones que se requiere resolver para la obtención de estimadores máximo verosímiles de los parámetros. Las expresiones y ecuaciones obtenidas resultan intratables analíticamente. Esto se debe al enfoque semiparamétrico asumido.

Con el objetivo de dar solución al problema, se opta por la implementación de muestreo de Gibbs, para lo cual se presentan todas las distribuciones condicionales completas de los parámetros. Los estimadores de Bayes (medias de las distribuciones a posteriori) se obtienen en base a la simulación computacional derivada de dichas distribuciones condicionales completas.

Las variantes abordadas aquí se relacionan con el modelo que contempla errores con distribución elíptica representable independiente y dependiente.

La metodología implementada se utilizó para resolver dos problemas presentes en la literatura. Medición del Porcentaje de Grasa Corporal (Branco et. al., 2000) y Mediciones de un Tipo de Encima en la Sangre (Eno, 1999). También se realizó una simulación, con el fin de determinar la influencia del tamaño muestral, de la distribución de los errores, de la magnitud de los errores, de la magnitud de la pendiente y de la media a priori del parámetro de interés sobre la precisión de las estimaciones.

Se puede comentar que a mayor pendiente la estimación de x_0 es mejor y la posteriori es menos sesgada, que el modelo (dependiente e independiente) no incide significativamente sobre la posteriori y que las estimaciones resultan mejores para muestras mayores.

Cabe mencionar que los resultados se basan en el modelo normal – gama y que se limitan, en general, al modelo de regresión lineal simple.

Se propone, como extensión, abordar el problema considerando errores skew normal y skew elíptico, considerar distribuciones a priori fuera del modelo normal – gama, y explotar nuevas posibilidades en las estructuras de dependencia entre los parámetros de interés.

CAPÍTULO 4:

SOLUCIÓN SEMIPARAMÉTRICA AL PROBLEMA DE CALIBRACIÓN LINEAL DISCRETA BAJO DISTRIBUCIONES ELÍPTICAS REPRESENTABLES

El problema de calibración lineal discreta deriva del problema de regresión discreta o, en nuestro caso, binaria, para el cual la ocurrencia o no ocurrencia de cierto fenómeno (representado por una variable binaria asumiendo, respectivamente, valores uno y cero) depende de ciertas predictoras o covariables a través de una función de enlace para la probabilidad de éxito.

En general este problema se puede plantear del modo siguiente:

$$y | \mathbf{x} \sim b(1; p(\mathbf{x})),$$

donde $p(\mathbf{x}) = H(\eta(\mathbf{x}))$ y donde la función de enlace H es una f.d.a.

El análisis de datos categóricos desde el punto de vista bayesiano se remonta al trabajo de Bayes (1763) y Laplace (1774), quienes estiman el parámetro de una distribución binomial utilizando una priori uniforme. En estos trabajos no está presente el efecto de covariables, y la posteriori de la probabilidad de éxito p es una distribución beta de parámetros $\alpha = y + 1$, y $\beta = n - y + 1$. El estimador de Bayes para p es $(y + 1) / (n + 2)$, que corresponde a un promedio ponderado entre el EMV (y / n) y la media de la distribución a priori ($1/2$).

Una clase más general de prioris para el parámetro de la binomial corresponde a la distribución beta (la uniforme es una beta de parámetros $\alpha = 1$ y $\beta = 1$). Una clase interesante de distribuciones a priori son las llamadas prioris de Jeffrey, proporcionales a la raíz cuadrada del determinante de la matriz de información de Fisher. Bernardo y Smith (1994) prueban que la priori de Jeffrey, que en este caso corresponde a una beta de parámetros $\alpha = 1/2$ y $\beta = 1/2$, es invariante, admisible y satisface otros criterios de bondad.

En general, es de interés establecer la relación entre la respuesta binaria y ciertas covariables o predictoras de tipo continuo \mathbf{x} . Leonard (1972) y Leonard y Hsu (1994) utilizan una priori logística para el parámetro de la binomial (modelo logito), donde este se define como $\text{logit}(p) = \log[p / (1 - p)]$.

Otros autores, que abordan datos binarios longitudinales, proponen escoger funciones de enlace multivariadas en una clase más general, como son las distribuciones mezcla de normales. Casos particulares son los modelos probit (H corresponde a la f.d.a. normal estándar), utilizados en este contexto por Ashford y Sowden (1970), Ameniya (1985), Chib y Greenberg (1988) y Chen y Dey (1998), y los modelos con enlace t – student multivariado, que son propuestos por Prentice (1988) y utilizados por Chen y Dey (1998).

Utilizando el enfoque bayesiano y con el objeto de modelar la función de enlace, pero considerando que η es lineal, es decir $\eta(\mathbf{x}) = \mathbf{x}_i^T \boldsymbol{\beta}$, Gelfand y Kuo (1991) y Newton et. al. (1996) usan un proceso Dirichlet para la priori de H , mientras que Mallick y Gelfand (1994) consideran mezcla discreta de f.d.a. beta, con una priori para los pesos.

En el mismo sentido, también Basu y Mukhopadhyay (2000) modelan la función de enlace, escogiendo dentro de la clase de funciones de distribución acumuladas que son mezcla de normales, obteniendo distribuciones a posteriori mezcla escala de normales truncadas (estos autores consideran enlaces simétricos y asimétricos), y generalizando así los modelos logito, probito y enlace t – student (*Proposición 2.2 (a)* de la sección 2.2.), donde la función mezcladora es un proceso Dirichlet. En su trabajo ellos presentan todas las distribuciones condicionales completas que permiten la implementación de muestreo de Gibbs. Estos autores asumen que la respuesta binaria se asocia a una variable subyacente normal no observable. Recientemente, Amewou – Atisso et. al. (2003) ampliaron los hallazgos de estos autores al considerar que η incluye términos polinómicos de mayor orden.

Nuestro trabajo se desarrolla, por una parte, siguiendo la línea de Basu y Mukhopadhyay en lo relativo a enlaces simétricos y considerando η lineal. La diferencia está en que se generalizan al modelo de calibración los resultados de estos autores relativos al modelo de regresión. En este sentido se conservan las variables subyacentes normales, pero el objetivo no es la estimación del parámetro de la binomial, sino la estimación de la covariable \mathbf{x}_0 en función de una nueva observación y_0 . Este resultado se basa en los *Lemas 2.1 y 2.7*.

El modelo abordado en este capítulo, y que se conoce como modelo de regresión binario, es el siguiente:

$$y_i \sim \text{ind Bernoulli} (\theta = H(\mathbf{x}_i^t \boldsymbol{\beta})), i = 0, \dots, n.$$

En nuestro caso, como se trata del problema de calibración en un modelo de regresión simple, contamos con vectores de observaciones (x_i, y_i) , donde las componentes x_i son fijas, donde se cuenta con una variable observable y_0 , pero donde no se cuenta con x_0 , el cual es el parámetro de interés.

Note que para este modelo,

$$P(y_i = y) = [H(\mathbf{x}_i^t \boldsymbol{\beta})]^y [1 - H(\mathbf{x}_i^t \boldsymbol{\beta})]^{1-y}; i = 0, \dots, n; y = 0, 1,$$

donde

$\boldsymbol{\beta} = (\beta_0, \beta_1)$ es el vector de Coeficientes de Regresión desconocido,

$\mathbf{x}_i = [1, x_i]$ son vectores de covariables observables y x_0 es el parámetro de interés.

Como se comentó antes, si H es la función de distribución acumulada normal, se obtiene el modelo probito; mientras que si se trata de la función de distribución acumulada Logística, se obtiene el modelo logito, y si se trata de la f.d.a. t - student, se obtiene el modelo con enlace t - student. En este caso, se usa una extensión de la función de enlace de Basu & Mukhopadhyay (2000), la cual queda expresada por:

$$H(\cdot) = \Phi_E(\cdot),$$

donde Φ_E es la función de distribución acumulada esférica. A este enlace lo llamaremos enlace elíptico y corresponde a una generalización de los enlaces logito, probito y t – student mencionados más arriba.

Luego, el modelo de regresión binario viene dado por:

$$(4.1) \quad y_i \sim \text{ind Bernoulli} (\theta = H(\mathbf{x}_i^t \boldsymbol{\beta})); i = 0, \dots, n.$$

Cabe hacer notar que el enlace t – Student (ν) corresponde al caso en que Φ_E es una función de distribución acumulada t – student con ν grados de libertad. En efecto, la expresión (4.1) es equivalente a:

$$(4.2) \quad y_i \sim \text{ind Bernoulli} (\theta = \int_0^\infty \Phi(\mathbf{x}_i^t \boldsymbol{\beta} / \sigma) dG(\sigma)); i = 0, \dots, n,$$

donde G depende del enlace específico utilizado. Por ejemplo para el enlace t – student, Φ es la función de distribución acumulada normal estándar y σ^2 tiene una distribución Gamma Inversa de parámetros de forma y escala igual a $\frac{1}{2} \nu$.

Branco et. al. (2001) abordan el modelo de calibración paramétrico con respuesta binomial

$$y_i | \mathbf{x}_i \sim b(k_i; p(\mathbf{x}_i)),$$

donde $p(\mathbf{x}) = H(\eta(\mathbf{x}))$ y η lineal, utilizando enlaces logito y probito y proponiendo el enlace t – student para el caso clásico. Para el enlace propuesto, los autores no implementan la solución bayesiana, pero entregan algunas sugerencias al respecto. Nuestro trabajo extiende estos resultados en el sentido que presenta un enlace elíptico y de este modo se obtiene un modelo más general y flexible. Los desarrollos se basan en el trabajo de Basu y Mukhopadhyay (2000).

Bajo otro enfoque, Kottas, Branco y Gelfand (2001) abordan el modelo de calibración no paramétrico con respuesta binomial. Estos autores centran su mirada en las aplicaciones a la dosimetría, es decir, el estudio de qué dosis de cierto agente (predictora) provoca tal o cual respuesta (binaria). En este sentido, junto con la dosis crece la probabilidad del evento de interés. Ellos enfrentan el problema modelando la respuesta como variables multinomiales que dependen de la dosis. Para cada multinomial existen r categorías de interés, que en su caso son tres (ningún evento, un evento y dos o más eventos). Esta restricción se justifica debido a que se trata de experiencias donde el evento es muy poco frecuente. Estos autores comparan el modelo logito con un modelo en el cual las probabilidades de las multinomiales son función de la dosis y de la categoría y las probabilidades están asociadas a un proceso Dirichlet.

Otro punto de vista consiste en considerar una función de enlace H fija y estimar η . Desde la perspectiva clásica, este problema es abordado por O’Sullivan et. al. (1986), Gu (1990), Hastie y Tibshirami (1987), Tibshirami y Hastie (1987), Staniswalis (1989), Cox y O’Sullivan (1990) y Fan et. al. (1995).

Bajo la perspectiva bayesiana, también existen trabajos enfocados a la estimación de η . Podemos nombrar entre otros a Wood y Kohn (1998) y DiMatteo et. al. (2001).

Este capítulo se divide en tres secciones. La primera sección desarrolla una metodología bayesiana semiparamétrica para el problema de calibración lineal discreta cuando la variable subyacente asociada a la respuesta binaria tiene una distribución normal, utilizando el enlace de Basu y Mukhopadhyay (2000). En esta sección se aborda, previamente, como ilustración, el modelo para el cual la variable subyacente es mezcla escala finita de normales o normal contaminada (el cual robustifica el modelo normal) y donde el vector de pesos tiene distribución Dirichlet. La extensión natural de este modelo paramétrico, y que corresponde al resultado principal de la primera sección, consiste en el caso semiparamétrico, donde los parámetros que definen la mezcla siguen un proceso Dirichlet. Esta sección también considera, en el caso paramétrico y Semiparamétrico, la solución bayesiana al modelo de calibración binomial. La segunda sección desarrolla la metodología bayesiana paramétrica para el problema de calibración lineal discreta cuando el enlace es elíptico representable independiente conocido. En particular, se presenta aquí la resolución del modelo paramétrico de calibración binaria con enlace t – student. Como subproducto de esta sección, se obtiene una metodología bayesiana paramétrica para el análisis de regresión lineal binaria con enlace t – student bajo los mismos supuestos mencionados anteriormente, también en esta sección se presenta la solución bayesiana paramétrica para el modelo de calibración binomial con enlace t - student. La resolución de los problemas en ambas secciones requiere la implementación de métodos de simulación MCMC, en particular, muestreo de Gibbs. Las prioris utilizadas se vinculan a las utilizadas por Basu y Mukhopadhyay y otras presentes en la literatura.

La tercera y última sección de este capítulo ilustra los resultados con datos utilizados previamente en la literatura y simulación.

4.1 SOLUCIÓN BAYESIANA SEMIPARAMÉTRICA AL PROBLEMA DE CALIBRACIÓN BAJO EL MODELO DE REGRESIÓN BINARIO CON VARIABLE SUBYACENTE NORMAL

En esta sección el problema de calibración discreta es abordado mediante la inclusión de variables subyacentes normales. Esto es una extensión directa de los resultados de Basu y Mukhopadhyay (2000) al modelo de calibración

El modelo jerárquico inicial para nuestro problema de calibración lineal discreta es el siguiente:

$$(4.3) \quad \begin{aligned} y_i | \beta, x_0 &\sim \text{ind Bernoulli} (H(\mathbf{x}_i^t \beta)); i = 0, \dots, n, \\ \beta &\sim \pi_1(\beta), \\ x_0 &\sim \pi_3(x_0) \\ \text{y} \quad \beta &\perp\!\!\!\perp x_0, \end{aligned}$$

donde la función de enlace se define como

$$(4.4) \quad H(z) = \int_0^{\infty} \Phi(z/\sigma) dG(\sigma) \text{ y } \Phi \text{ es la f.d.a. normal estándar.}$$

Haciendo uso de la *Proposición 2.2 (a)*, el problema puede ser abordado considerando las especificaciones en un número mayor de etapas, mediante la inclusión de variables latentes i.i.d. Z_i y ω_i , obteniéndose:

$$(4.5) \quad \begin{aligned} y_i | Z_i, \beta, x_0, \omega_i &\sim \text{ind. Bernoulli} (\theta_i = P(Z_i > 0 | \omega_i, \beta, \mathbf{x}_0)), \\ Z_i | \beta, x_0, \omega_i &\sim \text{ind } N(\mathbf{x}_i^t \beta, \omega_i), \\ \beta &\sim \pi_1(\beta), \\ x_0 &\sim \pi_3(x_0), \\ \omega_i | G &\sim \text{i.i.d } G \\ \text{y} \quad G &\sim \pi_2(G), \\ \omega &\perp\!\!\!\perp \beta | G; \beta \perp\!\!\!\perp G \perp\!\!\!\perp x_0. \end{aligned}$$

En este modelo, G es la función mezcladora, y (ω, β, x_0) y G son los parámetros de interés. El análisis requiere la especificación de π_1, π_2 y π_3 . Como ha sido usual en la literatura, se considera π_1 y π_3 normales; sin embargo π_2 , que usualmente se considera no informativa, bajo nuestro enfoque contempla otras alternativas.

Es importante destacar que la especificación (4.1) o (4.2) no permite determinar explícitamente las distribuciones condicionales completas que permiten la resolución del problema, sin embargo, la especificación (4.5) si lo permite. En efecto la verosimilitud para el modelo (4.2) es la siguiente:

$$L(\beta, x_0 | \mathbf{y}, G) = \prod_{i=0}^n \left[\int_0^{\infty} \Phi(\mathbf{x}_i^t \beta / \omega_i) \cdot dG(\omega_i) \right]^{y_i} \cdot \left[1 - \int_0^{\infty} \Phi(\mathbf{x}_i^t \beta / \omega_i) \cdot dG(\omega_i) \right]^{1-y_i},$$

mientras que la verosimilitud aumentada para el modelo (4.5) queda expresada en términos de la f.d.a. normal estándar:

$$L(\beta, x_0, \omega | \mathbf{y}, G) = \prod_{i=0}^n \left[\Phi(\mathbf{x}_i^t \beta / \omega_i) \right]^{y_i} \cdot \left[\Phi(-\mathbf{x}_i^t \beta / \omega_i) \right]^{1-y_i}.$$

Naturalmente esta segunda expresión requiere, por un lado, conocimiento respecto de la distribución G y, por otra parte, la estimación de variables latentes ω_i . La primera expresión, en cambio, aún para una G conocida, requiere la resolución del sistema definido en función de que las derivadas respecto de los parámetros sean nulas, lo cual a ojos vista es muy complicado.

Es claro que en ambos casos la estimación máximo verosímil requiere métodos numéricos; además es claro que en el primer caso se requiere conocimiento respecto de la función g y que en el segundo caso hay más parámetros que ecuaciones, por lo cual no tenemos por qué esperar una solución única.

Por otra parte, como sabemos, el estimador de Bayes de los parámetros de interés es la media de su distribución a posteriori. Para el modelo (4.2), considerando G conocida, la distribución a posteriori mencionada es de la forma siguiente:

$$\pi(\boldsymbol{\beta}, x_0 \mid \mathbf{y}, G) \propto L(\boldsymbol{\beta}, x_0 \mid \mathbf{y}, G) \cdot \pi_1(\boldsymbol{\beta}) \cdot \pi_3(x_0),$$

de donde

$$\pi(\boldsymbol{\beta}, x_0 \mid \mathbf{y}, G) \propto \pi_1(\boldsymbol{\beta}) \cdot \pi(x_0) \cdot \prod_{i=0}^n \left[\int_0^{\infty} \Phi(\mathbf{x}_i^t \boldsymbol{\beta} / \sigma) \cdot dG(\sigma) \right]^{y_i} \cdot \left[1 - \int_0^{\infty} \Phi(\mathbf{x}_i^t \boldsymbol{\beta} / \sigma) \cdot dG(\sigma) \right]^{1-y_i}.$$

Es claro que, como hemos asumido que $\boldsymbol{\beta} \perp x_0$, la posteriori de x_0 presenta la siguiente forma:

$$\pi(x_0 \mid \mathbf{y}, G) \propto \pi(x_0) \cdot \iint_{\mathbb{R}^2} \pi_1(\boldsymbol{\beta}) \cdot \prod_{i=0}^n \left[\int_0^{\infty} \Phi(\mathbf{x}_i^t \boldsymbol{\beta} / \sigma) \cdot dG(\sigma) \right]^{y_i} \cdot \left[1 - \int_0^{\infty} \Phi(\mathbf{x}_i^t \boldsymbol{\beta} / \sigma) \cdot dG(\sigma) \right]^{1-y_i} d\boldsymbol{\beta}.$$

La expresión anterior depende de G y de las distribuciones a priori de x_0 y $\boldsymbol{\beta}$, por lo cual no es posible evaluar el estimador de Bayes explícitamente.

Ahora bien, para la formulación (4.5) del modelo, en base a la verosimilitud aumentada, la distribución a posteriori de los parámetros es:

$$\pi(\boldsymbol{\beta}, x_0, \boldsymbol{\omega} \mid \mathbf{y}, G) \propto L(\boldsymbol{\beta}, x_0, \boldsymbol{\omega} \mid \mathbf{y}, G) \cdot \pi_1(\boldsymbol{\beta}) \cdot \pi_3(x_0) \cdot \prod_{i=0}^n dG(\omega_i),$$

de donde

$$\pi(\boldsymbol{\beta}, x_0, \boldsymbol{\omega} \mid \mathbf{y}, G) \propto \pi_1(\boldsymbol{\beta}) \cdot \pi_3(x_0) \cdot \prod_{i=0}^n dG(\omega_i) \cdot \prod_{i=0}^n \left[\Phi(\mathbf{x}_i^t \boldsymbol{\beta} / \omega_i) \right]^{y_i} \cdot \left[\Phi(-\mathbf{x}_i^t \boldsymbol{\beta} / \omega_i) \right]^{1-y_i}.$$

En base al resultado anterior, la distribución a posteriori de x_0 corresponde a la expresión siguiente:

$$\pi(x_0 \mid \mathbf{y}, G) \propto \pi_3(x_0) \cdot \int_{\mathbb{R}^n} \iint_{\mathbb{R}^2} \pi_1(\boldsymbol{\beta}) \cdot \prod_{i=0}^n dG(\omega_i) \cdot \prod_{i=0}^n \left[\Phi(\mathbf{x}_i^t \boldsymbol{\beta} / \omega_i) \right]^{y_i} \cdot \left[\Phi(-\mathbf{x}_i^t \boldsymbol{\beta} / \omega_i) \right]^{1-y_i} d\boldsymbol{\beta} d\boldsymbol{\omega},$$

que también depende de G y de las distribuciones a priori.

Ahora bien, si suponemos que G corresponde a la distribución Gamma Inversa de parámetros r y λ , la expresión para la distribución a posteriori de x_0 sigue siendo intratable, como se ve a continuación:

$$\pi(x_0 | \mathbf{y}, G) \propto \pi_3(x_0) \cdot \int_{R^n} \int_{IR^2} \pi_1(\boldsymbol{\beta}) \cdot \prod_{i=0}^n \left\{ \Phi(\mathbf{x}_i^t \boldsymbol{\beta} / \omega_i; \bar{y}_i) \cdot \Phi(-\mathbf{x}_i^t \boldsymbol{\beta} / \omega_i; \bar{y}_i) \cdot \exp\left(-\frac{\lambda}{\omega_i}\right) \cdot \omega_i^{-r-1} \right\} \cdot d\boldsymbol{\beta} d\boldsymbol{\omega}.$$

Estas complicadas expresiones, sin resolución analítica, explican por qué hemos optado por utilizar MCMC para implementar muestreo de Gibbs y obtener de esa forma los estimadores de Bayes de los parámetros de interés.

Para H definida por (4.4), las primeras tres etapas del modelo (4.5) se pueden re-escribir como (4.1), en base a la *Proposición 2.2. (a)*. Consideraremos dos alternativas para la elección de la distribución a priori para G :

- i) la distribución a priori para G tiene soporte finito
- ii) G se distribuye de acuerdo a un proceso Dirichlet de parámetros c y G_0 , lo cual se anota $G \sim D(c; G_0)$.

En este segundo caso, al suponer que h es conocida, G resulta degenerada en G_0 , donde G_0 es una función de distribución acumulada fija.

Proposición 4.1

Bajo las especificaciones siguientes, que corresponden a un modelo de regresión lineal simple derivado de la especificación (4.5):

$$(4.6) \quad \begin{aligned} y_i | z_i, \boldsymbol{\beta}, x_0, \omega_i &\sim \text{ind Bernoulli}(\theta_i = P(z_i > 0 | \boldsymbol{\beta}, x_0); i = 0, \dots, n, \\ z_i | \boldsymbol{\beta}, x_0, \omega_i &\sim \text{ind } N(\beta_0 + \beta_1 x_i, \omega_i), \\ \beta_0 &\sim N_1(b_0; v_0), \\ \beta_1 &\sim N_1(b_1; v_1) \\ y, x_0 &\sim N_1(c_0; d_0), \end{aligned}$$

con $p(\omega_i | G) = \sum_{j=1}^{k+1} p_j \cdot \delta_{a_j}(\omega_i)$ y $(p_1, \dots, p_k) \sim D(\alpha_1, \dots, \alpha_k)$, asumiendo k conocido a priori, la

distribución a posteriori de x_0 tiene la siguiente forma analíticamente intratable:

$$\pi(x_0 | \mathbf{y}, G) \propto \exp\left\{-\frac{c_0 - c_0}{2 \cdot d_0}\right\} \cdot \int_{R^n} \int_{IR^2} \exp\left\{-\frac{1}{2} \left[\frac{b_0 - b_0}{v_0} + \frac{b_1 - b_1}{v_1} \right]\right\} \cdot \prod_{i=0}^n \left\{ \sum_{j=1}^k \left[\Phi\left(\frac{\beta_0 + \beta_1 \cdot x_i}{\omega_i}\right) \right]^{y_i} \cdot \left[\Phi\left(-\frac{\beta_0 + \beta_1 \cdot x_i}{\omega_i}\right) \right]^{1-y_i} \cdot p_j \cdot \delta_{a_j}(\omega_i) \right\} d\boldsymbol{\beta} d\boldsymbol{\omega},$$

lo cual motiva la obtención de las siguientes distribuciones condicionales completas:

$$(4.7) \quad z_i | \boldsymbol{\beta}, x_0, \omega_i, \mathbf{p}, y_i \stackrel{\text{cond.ind.}}{\sim} NT(IR^+, \beta_0 + \beta_1 x_i; \omega_i) \cdot I(Y_i=1) + NT(IR^-; \beta_0 + \beta_1 x_i; \omega_i) \cdot I(Y_i=0),$$

donde $NT(A, \mu, \sigma^2)$ denota una distribución Normal Truncada, cuya f.d.p. es $f_{NT}(A; \mu; \sigma^2) = f_{N(\mu, \sigma^2)} \cdot I_A / P(A)$.

$$(4.8) \quad \beta_0 | \mathbf{z}, \beta_1, x_0, \omega, \mathbf{p}, \mathbf{y} \sim N_1(m_0, S_0^2), \text{ con } m_0 = S_0^2 \left[\sum_{i=0}^n \frac{z_i - \beta_1 x_i}{\omega_i} + \frac{b_0}{v_0} \right] \text{ y } S_0^2 = \left[\sum_{i=0}^n \frac{1}{\omega_i} + \frac{1}{v_0} \right]^{-1}$$

$$(4.9) \quad \beta_1 | \mathbf{z}, \beta_0, x_0, \omega, \mathbf{p}, \mathbf{y} \sim N_1(m_1, S_1^2),$$

$$\text{con } m_1 = S_1^2 \left[\sum_{i=0}^n \frac{(z_i - \beta_0) x_i}{\omega_i} + \frac{b_1}{v_1} \right] \text{ y } S_1^2 = \left[\sum_{i=0}^n \frac{x_i^2}{\omega_i} + \frac{1}{v_1} \right]^{-1}$$

$$(4.10) \quad x_0 | \mathbf{z}, \beta, \omega, \mathbf{p}, \mathbf{y} \sim N(m_2, S_2^2), \text{ con } S_2^2 = \left[\frac{\beta_1^2}{\omega_0} + \frac{1}{d_0} \right]^{-1} \text{ y } m_2 = S_2^2 \left[\frac{(Z_0 - \beta_0) \beta_1}{\omega_0} + \frac{c_0}{d_0} \right].$$

$$(4.11) \quad \omega_i | \mathbf{z}, \beta, x_0, \omega_{\cdot i}, \mathbf{p}, \mathbf{y} \sim \sum_{j=1}^{k+1} q_{ij} \cdot \delta_i(\tau_j), \text{ con } q_{ij} = \frac{\phi\left(\frac{\mathbf{z}_i - x_i \mathbf{t} \beta}{\sqrt{\tau_j}}\right) \cdot \frac{p_j}{\sqrt{\tau_j}}}{\sum_{m=1}^{k+1} \phi\left(\frac{\mathbf{z}_i - x_i \mathbf{t} \beta}{\sqrt{\tau_m}}\right) \cdot \frac{p_m}{\sqrt{\tau_m}}}, \text{ con } \tau_j \text{ los posibles}$$

valores de los ω_i y con ϕ la f.d.a. normal estándar.

$$(4.12) \quad \mathbf{p} | \mathbf{z}, \beta, x_0, \omega, \mathbf{y} \sim D(\alpha^*), \text{ con } \alpha^* = (\alpha_j^*) \text{ y } \alpha_j^* = \alpha_j + \sum_{i=0}^n \delta_{\omega_i}(\tau_j).$$

Demostración:

Las distribuciones condicionales completas se derivan directamente de los lemas de la sección 2.3. En particular, las distribuciones condicionales completas de \mathbf{z} , x_0 , β_0 , β_1 , ω y \mathbf{p} , surgen del *Ejemplo 2.5*.

Corolario 4.1.

Un problema de interés mayor al abordado anteriormente es aquel para el cual nuestra atención no se centra en las respuestas binarias de predictoras distintas, sino aquel en el cual una respuesta binomial s_i se deriva de la predictora común x_i para una muestra de n_i individuos. Este problema usualmente se conoce como dosimetría, ya que su principal aplicación se relaciona con problemas en los cuales la predictora es la dosis de cierto agente y la variable respuesta es el número de ocurrencias del fenómeno de interés en una muestra que fue expuesta a dicha dosis.

Este problema ha sido abordado en la literatura por Branco et. al. (2001) y por Cotas, Branco y Gelfand (2001). En el primer artículo los autores dejan abierto el análisis bayesiano para este problema, mientras que en el segundo artículo se utiliza un enfoque no – paramétrico.

El problema de interés está representado en el siguiente modelo jerárquico:

$$\begin{aligned}
s_i | \mathbf{z}_i, \boldsymbol{\beta}, x_0, \omega_i, \mathbf{p} &\sim \text{ind Binomial}(n_i; \theta_i = P(z_{i1} > 0 | \boldsymbol{\beta}, x_0); i = 0, \dots, n; j = 1, \dots, n_i, \\
z_{ij} | \boldsymbol{\beta}, x_0, \omega_i, \mathbf{p} &\sim \text{ind } N(\beta_0 + \beta_1 x_i, \omega_i); i = 0, \dots, n; j = 1, \dots, n_i, \\
(4.13) \quad \beta_0 &\sim N_1(\mathbf{b}_1; \mathbf{v}_1), \\
&\beta_1 \sim N_1(\mathbf{b}_1; \mathbf{v}_1) \\
\text{y } x_0 &\sim N_1(\mathbf{c}_0; \mathbf{d}), \\
\text{con } p(\omega_i | G) &= \sum_{j=1}^{k+1} p_j \cdot \delta_{a_j}(\omega_i) \text{ y } (p_1, \dots, p_k) \sim D(\alpha_1, \dots, \alpha_k).
\end{aligned}$$

Es claro que haciendo $s_i = \sum_{j=1}^{n_i} y_{ij}$, y debido a que los y_{ij} son independientes, este modelo es equivalente a la siguiente formulación, que permite derivar su solución de la *Proposición 4.1*:

$$\begin{aligned}
y_{ij} | z_{ij}, \boldsymbol{\beta}, x_0, \omega_i, \mathbf{p} &\sim \text{ind Bernoulli}(\theta_i = P(z_{ij} > 0 | \boldsymbol{\beta}, x_0); i = 0, \dots, n; j = 1, \dots, n_i, \\
z_{ij} | \boldsymbol{\beta}, x_0, \omega_i, \mathbf{p} &\sim \text{ind } N(\beta_0 + \beta_1 x_i, \omega_i); i = 0, \dots, n; j = 1, \dots, n_i, \\
(4.14) \quad \beta_0 &\sim N_1(\mathbf{b}_1; \mathbf{v}_1), \\
&\beta_1 \sim N_1(\mathbf{b}_1; \mathbf{v}_1) \\
\text{y } x_0 &\sim N_1(\mathbf{c}_0; \mathbf{d}), \\
\text{con } p(\omega_i | G) &= \sum_{j=1}^{k+1} p_j \cdot \delta_{a_j}(\omega_i) \text{ y } (p_1, \dots, p_k) \sim D(\alpha_1, \dots, \alpha_k).
\end{aligned}$$

Las distribuciones condicionales completas que permiten la utilización de muestreo de Gibbs para la estimación del parámetro de interés x_0 son las siguientes:

$$(4.15) \quad z_{ij} | \boldsymbol{\beta}, x_0, \omega_i, \mathbf{p}, y_{ij} \propto \text{cond.ind. } NT(\mathbb{R}^+; \beta_0 + \beta_1 x_i; \omega_i) \cdot \mathbf{I}(y_{ij}=1) + NT(\mathbb{R}^-; \beta_0 + \beta_1 x_i; \omega_i) \cdot \mathbf{I}(y_{ij}=0).$$

$$(4.16) \quad \beta_0 | \mathbf{z}, \beta_1, x_0, \boldsymbol{\omega}, \mathbf{p}, \mathbf{y} \sim N_1(\mathbf{m}_0, \mathbf{S}_0^2),$$

$$\text{con } \mathbf{m}_0 = \mathbf{S}_0^2 \left[\sum_{i=0}^n \frac{z_i \cdot -n_i \cdot \beta_1 \cdot x_i}{\omega_i} + \frac{\mathbf{b}_0}{\mathbf{V}_0} \right] \text{ y } \mathbf{S}_0^2 = \left[\sum_{i=0}^n \frac{n_i}{\omega_i} + \frac{1}{\mathbf{V}_0} \right]^{-1}, \text{ donde } z_i = \sum_{j=1}^{n_i} z_{ij}.$$

$$(4.17) \quad \beta_1 | \mathbf{z}, \beta_0, x_0, \boldsymbol{\omega}, \mathbf{p}, \mathbf{y} \sim N_1(\mathbf{m}_1, \mathbf{S}_1^2),$$

$$\text{con } \mathbf{m}_1 = \mathbf{S}_1^2 \left[\sum_{i=0}^n \frac{(z_i - n_i \cdot \beta_0) x_i}{\omega_i} + \frac{\mathbf{b}_1}{\mathbf{V}_1} \right] \text{ y } \mathbf{S}_1^2 = \left[\sum_{i=0}^n \frac{n_i \cdot x_i^2}{\omega_i} + \frac{1}{\mathbf{V}_1} \right]^{-1}.$$

$$(4.18) \quad x_0 | \mathbf{z}, \boldsymbol{\beta}, \boldsymbol{\omega}, \mathbf{p}, \mathbf{y} \sim N(\mathbf{m}_2, \mathbf{S}_2^2), \text{ con } \mathbf{S}_2^2 = \left[\frac{n_0 \cdot \beta_1^2}{\omega_0} + \frac{1}{\mathbf{d}} \right]^{-1} \text{ y } \mathbf{m}_2 = \mathbf{S}_2^2 \left[\frac{(Z_0 - n_0 \cdot \beta_0) \cdot \beta_1}{\omega_0} + \frac{\mathbf{c}_0}{\mathbf{d}} \right].$$

$$(4.19) \quad \omega_i | \mathbf{z}, \boldsymbol{\beta}, x_0, \omega_{-i}, \mathbf{p}, \mathbf{y} \sim g(\cdot), \quad \text{donde } g(x) = \sum_{l=1}^{k+1} q_{il} \cdot \delta_{(\tau_l)}(x)$$

$$\text{y donde } q_{il} = \frac{\exp\left\{-\frac{1}{2 \cdot \tau_l} \sum_{j=1}^{n_i} (z_{ij} - \mathbf{x}_i \mathbf{t} \boldsymbol{\beta})^2\right\} \cdot \frac{p_l}{\sqrt{\tau_l^{n_i}}}}{\sum_{m=1}^{k+1} \exp\left\{-\frac{1}{2 \cdot \tau_m} \sum_{j=1}^{n_i} (z_{ij} - \mathbf{x}_i \mathbf{t} \boldsymbol{\beta})^2\right\} \cdot \frac{p_m}{\sqrt{\tau_m^{n_i}}}}$$

$$(4.20) \quad \mathbf{p} | \mathbf{z}, \boldsymbol{\beta}, x_0, \boldsymbol{\omega}, \mathbf{y} \sim D(\boldsymbol{\alpha}^*), \quad \text{con } \boldsymbol{\alpha}^* = (\alpha_j^*) \text{ y } \alpha_j^* = \alpha_j + \sum_{i=0}^n n_i \cdot \delta_{\omega_i}(\tau_i).$$

Estos resultados permiten la implementación computacional del muestreo de Gibbs para la estimación de x_0 , sin embargo, otra alternativa consiste en utilizar Winbugs.

A continuación se presenta el principal resultado de esta sección, el cual corresponde a la solución bayesiana semiparamétrica del problema de calibración lineal discreta con variables subyacentes elípticas representables independientes.

Proposición 4.2

Bajo las especificaciones, (4.6), con $G \sim \mathcal{D}(c, G_0)$ y con G_0 la función de distribución acumulada Gamma Inversa de parámetros de posición r_1 y escala λ_1 , la distribución a posteriori de x_0 , presentada a continuación, no es tratable analíticamente:

$$\pi(x_0 | \mathbf{y}, G) \propto \exp\left\{-\frac{c_0 - c_0}{2 \cdot d_0}\right\} \cdot \int_{R^n} \int_{IR^2} \exp\left\{-\frac{1}{2} \left[\frac{b_0 - b_0}{v_0} + \frac{b_1 - b_1}{v_1} \right]\right\} \cdot \prod_{i=0}^n \left\{ \left[\Phi\left(\frac{\beta_0 + \beta_1 \cdot x_i}{\omega_i}\right) \right]^{y_i} \cdot \left[\Phi\left(-\frac{\beta_0 + \beta_1 \cdot x_i}{\omega_i}\right) \right]^{1-y_i} \cdot \frac{\exp\left(-\frac{\lambda}{\omega_i}\right)}{\omega_i^{r+1}} \right\} d\boldsymbol{\beta} d\boldsymbol{\omega},$$

lo cual hace necesario la obtención de las distribuciones condicionales completas.

En este caso, desde la Proposición anterior se derivan las expresiones para \mathbf{z} , β_0 , β_1 y x_0 , respectivamente (4.7) a (4.10) y, en base a los resultados de Escobar y West (1998), se obtiene la distribución condicional completa para ω_i :

$$(4.21) \quad \omega_i | \omega_{-i}, G, \boldsymbol{\beta}, x_0, \mathbf{z}, \mathbf{y} \sim q_{0i} \cdot G_b^{(i)}(d, \omega_i) + \sum_{k \neq i} q_{ki} \cdot \delta_{\omega_k}(\omega_i),$$

$$\text{con } q_{0i} = f_{T_i}(z_i), \quad \text{donde } T_i \sim T(\mu_i = E(z_i) = \beta_0 + \beta_1 x_i, \sigma^2 = 2\lambda_1, \gamma = 2r_1),$$

$$\text{con } q_{ki} = \phi_Z\left(\frac{z_i - \beta_0 - \beta_1 \cdot x_i}{\sqrt{\omega_k}}\right) \quad \text{y con } G_b^{(i)} \text{ la función de distribución acumulada Gamma}$$

$$\text{Inversa de parámetros de posición } r = r_1 + 1/2 \text{ y de escala } \lambda = \lambda_1 + (z_i - \beta_0 - \beta_1 x_i)^2 / 2.$$

Demostración:

Los resultados anteriores se derivan de modo análogo a los relativos al caso abordado en la *Proposición 4.1.*, a excepción de la distribución condicional de ω_i , la cual se obtiene a partir del Teorema de Blackwell y Mac Queen y de las ecuaciones (1.12) y (1.13) de la sección 1.1.

Corolario 4.2.

La extensión Semiparamétrica del problema planteado en el Corolario 4.1. considera $G \sim D$ (c G_0) y $G_0 \sim$ Gamma Inversa (r_1, λ_1). La resolución mantiene las distribuciones condicionales completas para \mathbf{z} , β_0 , β_1 y x_0 , y la distribución condicional completa para ω_i es la siguiente:

$$(4.22) \quad \omega_i / \omega_{-i}, G, \beta, x_0, \mathbf{z}, \mathbf{y} \sim q_{0i} \cdot dG_b^{(i)}(\omega_i) + \sum_{k \neq i} q_{ki} \cdot \delta_{\omega_k}(\omega_i),$$

$$\text{con } q_{0i} = \frac{\Gamma\left(\frac{1}{2}(r_1 + n_i)\right)}{\Gamma\left(\frac{1}{2}r_1\right) \pi^{n_i/2} \lambda_1^{n_i/2}} \cdot \left[\frac{\sum_{j=1}^{n_i} \mathbf{e}_{ij} - \mathbf{x}_i^t \beta}{2 \cdot \lambda_1} + 1 \right]^{-\left(\frac{r_1 + n_i}{2}\right)},$$

$$\text{con } q_{ki} = \pi \cdot \omega_i^{n_i/2} \cdot \exp\left\{-\frac{1}{2 \cdot \omega_i} \sum_{j=1}^{n_i} \mathbf{e}_{ij} - \mathbf{x}_i^t \beta\right\}$$

y con $G_b^{(i)}$ la función de distribución acumulada Gamma Inversa de parámetros de posición $r = r_1 + 1/2 \cdot n_i$ y de escala $\lambda = \lambda_1 + 1/2 \sum_{j=1}^{n_i} \mathbf{e}_{ij} - \mathbf{x}_i^t \beta$.

La implementación permite, utilizando Muestreo de Gibbs, estimar el parámetro de interés x_0 , predictor para la respuesta binomial s_0 .

4.2 SOLUCIÓN BAYESIANA PARAMÉTRICA AL PROBLEMA DE CALIBRACIÓN BAJO EL MODELO DE REGRESIÓN BINARIO CON ENLACE ELÍPTICO REPRESENTABLE INDEPENDIENTE

En esta sección el problema de calibración discreta es abordado mediante la inclusión de variables subyacentes elípticas. Esto es una aplicación de los resultados en la sección previa y que resulta en una extensión del trabajo de Branco et. al. (2001). La extensión se basa en la relación existente entre los modelos normal y elíptico, y su objetivo es dar mayor flexibilidad al modelo.

El modelo jerárquico inicial para nuestro problema de calibración lineal discreta es el siguiente:

$$(4.23) \quad \begin{aligned} y_i | \beta, x_0 &\sim \text{ind Bernoulli}(H(\mathbf{x}_i^t \beta)); i = 0, \dots, n, \\ \beta &\sim \pi_1(\beta), \\ x_0 &\sim \pi_3(x_0) \\ \mathbf{y} &\perp\!\!\!\perp x_0, \end{aligned}$$

con $H(\mathbf{x}_i^t \beta) = \Phi_E(\mathbf{x}_i^t \beta)$ y Φ_E la f.d.a. esférica.

Por la *Proposición 2.2 (b)*, y para G dado, el problema puede ser abordado considerando las especificaciones en un número mayor de etapas, mediante la inclusión de variables latentes i.i.d. Z_i ($\omega_i = 1 \forall i = 1, \dots, n$). El modelo jerárquico es, entonces:

$$\begin{aligned}
 (4.24) \quad & y_i | Z_i, \beta, x_0 && \sim \text{ind. Bernoulli} (\theta_i = P(Z_i > 0 | \beta, x_0)), \\
 & Z_i | G, \beta, x_0 && \sim \text{ind. } El_1(\mathbf{x}_i^t \beta, 1, h), \\
 & \beta && \sim \pi_1(\beta) \\
 \text{y} & x_0 && \sim \pi_3(x_0) \\
 & \beta && \perp\!\!\!\perp x_0,
 \end{aligned}$$

Ahora bien, por la *Proposición 2.1. (b)*, las expresiones segunda y tercera de (4.24) se pueden reemplazar obteniéndose el siguiente modelo jerárquico alternativo:

$$\begin{aligned}
 (4.25) \quad & y_i | Z_i, \beta, x_0, \omega_i && \sim \text{ind. Bernoulli} (\theta = P(Z_i > 0 | \omega_i, \beta, x_0)) \\
 & Z_i | \beta, x_0, \omega_i && \sim \text{ind. } N(\mathbf{x}_i^t \beta, \omega_i) \\
 & \beta && \sim \pi_1(\beta) \\
 \text{y} & x_0 && \sim \pi_3(x_0) \\
 & \omega_i | G && \sim \text{c.i.i.d. } G, \\
 & \omega && \perp\!\!\!\perp \beta; \beta \perp\!\!\!\perp x_0,
 \end{aligned}$$

de donde se recupera el modelo inicial (4.23).

El análisis de este Modelo requiere la especificación de π_1 y π_3 , y de la distribución G (la distribución elíptica representable particular de interés). Como ha sido usual en la literatura, se considera π_1 y π_3 normales.

Las siguientes posibles especificaciones para la función generadora, h_1 , h_2 y h_3 , permiten representar respectivamente los enlaces logito, probito y t – student:

$$\begin{aligned}
 (4.26) \quad & h_1(u) = \frac{\exp\left(\frac{u}{\sqrt{1+\exp(u)}}\right)}{\left(1 + \exp\left(\frac{u}{\sqrt{1+\exp(u)}}\right)\right)}, \\
 & h_2(u) = (2\pi)^{-1/2} \exp(-1/2 u) \\
 \text{y} & h_3(u) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi} \cdot \Gamma\left(\frac{\nu}{2}\right)} \left[1 + \frac{u}{\nu}\right]^{-\frac{\nu+1}{2}}.
 \end{aligned}$$

Cabe mencionar que el caso de la distribución logística no corresponden a una distribución elíptica representable.

El análisis del modelo (4.23) se puede realizar en base al modelo análogo (4.25), el cual coincide con el modelo (4.5) para G desconocido, de modo que el modelo con enlace elíptico representable es el mismo que se abordó en la sección anterior .

Observación 4.3.

Ahora bien, si la distribución elíptica particular es conocida, el modelo paramétrico presenta distribuciones condicionales completas idénticas a las obtenidas en la *Proposición 4.1.*, excepto en lo relativo a ω_i , para el cual la distribución condicional completa se puede obtener excepto una constante de proporcionalidad:

$$\pi(\omega_i \mid \mathbf{z}, \boldsymbol{\beta}, x_0,) \propto \phi((z_i - \mathbf{x}_i^t \boldsymbol{\beta}) / \omega_i) \cdot \pi(\omega_i).$$

Para el caso del enlace t – student (ν), la distribución condicional completa de ω_i es la siguiente:

$$(4.27) \quad \omega_i \sim \text{Inverse Gamma}(\frac{1}{2} \nu + \frac{1}{2}; \frac{1}{2} \nu + \frac{1}{2} (z_i - \beta_0 - \beta_1 \cdot x_i)^2).$$

Una alternativa factible en lugar de programar el muestreo de Gibbs en función de todas las distribuciones condicionales completas para el modelo paramétrico es utilizar Winbugs.

Observación 4.4.

El problema de regresión binaria con enlace t – student (ν) se resuelve utilizando las distribuciones condicionales completas de la *Proposición 4.2.*, excepto para el parámetro x_0 , el cual no existe, ya que la muestra contempla valores de la predictora x_1, \dots, x_n , y omitiendo el término $i = 0$ en las sumatorias.

Observación 4.5.

Al enfrentar el problema de calibración binomial, presentado en el *Corolario 4.1.*, pero con una distribución elíptica conocida para el enlace, las distribuciones condicionales completas de los parámetros \mathbf{z} , β_0 , β_1 y x_0 coinciden con las de la *Proposición 4.1.*, y para el caso del parámetro de escala ω_i la distribución condicional completa se puede obtener excepto por la constante de proporcionalidad:

$$\pi(\omega_i \mid \mathbf{z}, \boldsymbol{\beta}, x_0,) \propto \omega_i^{-n_i/2} \exp\left\{-\sum_{j=1}^{n_i} (z_{ij} - \mathbf{x}_i^t \boldsymbol{\beta})^2 / 2\omega_i\right\} \cdot \pi(\omega_i).$$

Aquí las variables $s_i \sim \text{binomial}(n_i, \theta_i)$, para facilitar los desarrollos, se descomponen como la suma de n_i variables $y_{ij} \sim \text{i.i.d. bernoulli}(\theta_i)$, donde θ_i depende de la predictora \mathbf{x}_i a través de una función de enlace, es decir, $\theta_i = H(\mathbf{x}_i^t \boldsymbol{\beta})$. Las variables z_{ij} , a su vez, son variables instrumentales tales que $z_{ij} \sim \text{i.i.d. } N(\mathbf{x}_i^t \boldsymbol{\beta})$, dependiendo también de la predictora \mathbf{x}_i , y se comportan como variables subyacentes a cada una de las variables aleatorias bernoulli: $y_{ij} \sim \text{i.i.d. bernoulli}(\theta_i = P(z_{ij} > 0))$. De ahí que para el caso del enlace t – student (ν), la distribución condicional completa de ω_i resulta ser:

$$(4.28) \quad \omega_i \sim \text{Inverse Gamma}(r = r_1 + \frac{1}{2} \cdot n_i; \lambda = \lambda_1 + \frac{1}{2} \sum_{j=1}^{n_i} (z_{ij} - \mathbf{x}_i^t \boldsymbol{\beta})^2).$$

Una alternativa factible en lugar de programar el muestreo de Gibbs en función de todas las distribuciones condicionales completas para el modelo paramétrico es utilizar Winbugs.

4.3 APLICACIONES

En esta sección se ilustran los resultados obtenidos en las secciones previas a través de simulación y conjuntos de datos considerados previamente en la literatura.

Aplicación 4.1 Características asociadas a la remisión del cáncer (Lee, 1974)

Lee (1974) presenta un conjunto de datos relativo a 27 pacientes y relacionado con características asociadas a la remisión del cáncer. Este estudio reveló que la principal variable explicativa (denotada x_i) para la remisión del cáncer es el índice de marcación, “labeling index” LI, que corresponde al índice de actividad proliferativa de las células o al porcentaje de células que aparecen marcadas luego de que los pacientes reciben una inyección de timidina. La variable respuesta considerada es de tipo binario, donde el valor 1 indica remisión del cáncer. La *Tabla 4.1.* presenta la información y la gráfica se puede observar en la *Figura 4.1.*

Tabla 4.1: Datos de remisión de cáncer y actividad proliferativa de las células L.I. La variable explicativa x es L.I. y la variable respuesta y es de tipo binario, con el valor 1 representando la remisión de cáncer.

LI	8	8	10	10	12	12	12	14	14	14	16	16	16	18
remision	0	0	0	0	0	0	0	0	0	0	0	0	0	1

LI	20	20	20	22	22	24	26	28	32	34	38	38	38
remision	0	1	1	0	1	0	1	1	0	1	1	1	0

El análisis clásico realizado con el modelo de regresión Logit entrega estimadores para β_0 y β_1 que asumen valores $-3,777$ y $0,145$ respectivamente. El error estándar hallado es de $0,059$ y se predice que para $LI = 26$ ($3,777/0,145$) la probabilidad de remisión es de un 50%.

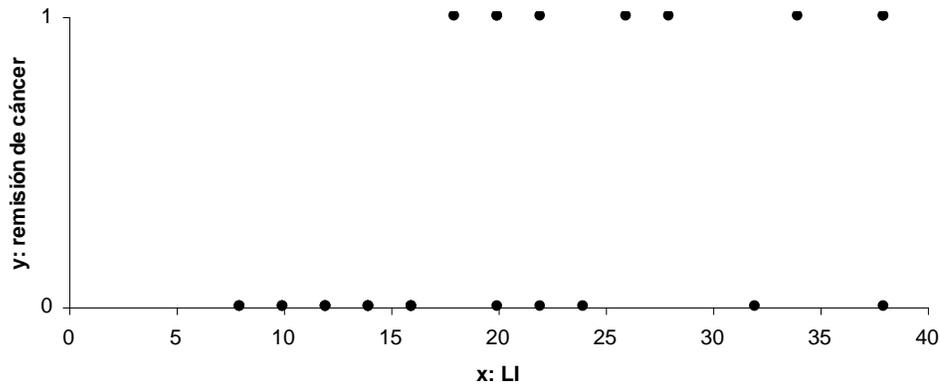


Figura 4.1: Diagrama de dispersión para los datos de la *Tabla 4.1.* relativos a remisión de cáncer.

Nuestro análisis consideró, para los coeficientes β_0 y β_1 , distribuciones a priori normales con medias $-3,777$ y $0,145$ y con varianzas 100 y $0,2$ respectivamente. Las varianzas se escogieron grandes, con el criterio de obtener distribuciones a priori no informativas.

Para la distribución de x_0 se consideró una distribución a priori normal centrada en la media de los valores LI, $20,07$ y con varianza 150 , mayor a la de las observaciones.

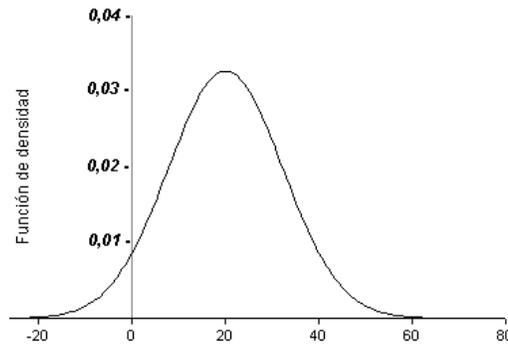


Figura 4.2: Distribución a priori para el parámetro de interés x_0 : $N(20,07; 150)$.

En lo relativo a z , se consideró que las remisiones de cáncer correspondían a la mitad de las no – remisiones, por lo cual se asignó una media negativa $\mu_x = -0,45$ (corresponde aproximadamente al valor promedio de los valores extremos de $x_i^t \beta$) y una varianza relativamente grande, $\sigma^2_x = 10$, para hacer la priori no – informativa. Para la distribución de los ω_i se consideraron valores $r_1 = \lambda_1$, asumiendo valores 2 y 25 , que vinculan el modelo Semiparamétrico elíptico al modelo t – student con 4 y 50 g. l., respectivamente. Los resultados de la estimación de x_0 se presentan en la *Tabla 4.2*.

MODELO BERNOULLI

Para el modelo bernoulli (*Proposición 4.2.*), la calibración se realiza estimando el índice de proliferación de las células $x_0 = L.I.$ que genera cada una de las respuestas ($y_0 = 1$ ó $y_0 = 0$). Los resultados en la *Tabla 4.2.* muestran que las estimaciones de los parámetros β_0 y β_1 bajo el enfoque bayesiano paramétrico y semiparamétrico no difieren demasiado de las estimaciones clásicas utilizando enlace logit.

Es claro que una estimación puntual de x_0 para cada valor de y_0 es insuficiente, ya que una amplia gama de posibles valores del índice de proliferación de las células lleva a la remisión o a la no remisión del cáncer. Esto hace necesario agregar a los estimadores de Bayes para x_0 sendas cotas de credibilidad, superiores cuando no hay remisión del cáncer e inferiores para los casos en que si hay remisión. Así se puede inferir que para un individuo que presenta remisión del cáncer, el 95% de las veces L.I. es superior a 15 ; de modo que un valor mientras que para un individuo que no presenta remisión del cáncer, el 95% de las veces L.I. es inferior a 26 (enlaces elípticos semiparamétricos). Valores entre 15 y 26 no resultan útiles para predecir un resultado.

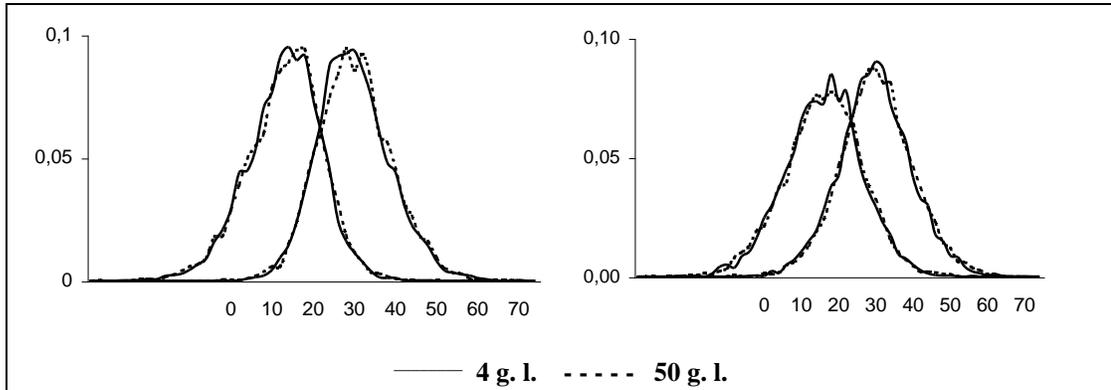
Tabla 4.2: Análisis de los datos de Lee (1974). Estimación bayesiana paramétrica y semi paramétrica del parámetro de interés para un individuo que presenta remisión de cáncer ($y_0 = 1$) y para un individuo que no la presenta ($y_0 = 0$).

Estimación x_0	Modelo	$y_0 = 0$		$y_0 = 1$	
		Estimación [cota superior 95%]	β	Estimación [cota inferior 95%]	β
Est. bayesiana semiparamétrica	4 g. l. ($r_1 = \lambda_1 = 2$)	14,48 [25,38]	$\begin{bmatrix} -3,07 \\ 0,12 \end{bmatrix}$	29,35 [15,41]	$\begin{bmatrix} -3,09 \\ 0,13 \end{bmatrix}$
	50 g. l. ($r_1 = \lambda_1 = 25$)	15,05 [25,87]	$\begin{bmatrix} -2,51 \\ 0,09 \end{bmatrix}$	29,16 [15,43]	$\begin{bmatrix} -2,50 \\ 0,10 \end{bmatrix}$
Est. bayesiana paramétrica	Enlace Probit	12,15 [25,81]	$\begin{bmatrix} -3,22 \\ 0,15 \end{bmatrix}$	28,66 [14,81]	$\begin{bmatrix} -3,12 \\ 0,15 \end{bmatrix}$
	Enlace t – student (4)	14,94 [31,49]	$\begin{bmatrix} -3,10 \\ 0,12 \end{bmatrix}$	28,81 [12,76]	$\begin{bmatrix} -3,03 \\ 0,12 \end{bmatrix}$
	Enlace t – student (50)	14,97 [31,51]	$\begin{bmatrix} -2,56 \\ 0,10 \end{bmatrix}$	29,30 [13,27]	$\begin{bmatrix} -2,49 \\ 0,10 \end{bmatrix}$

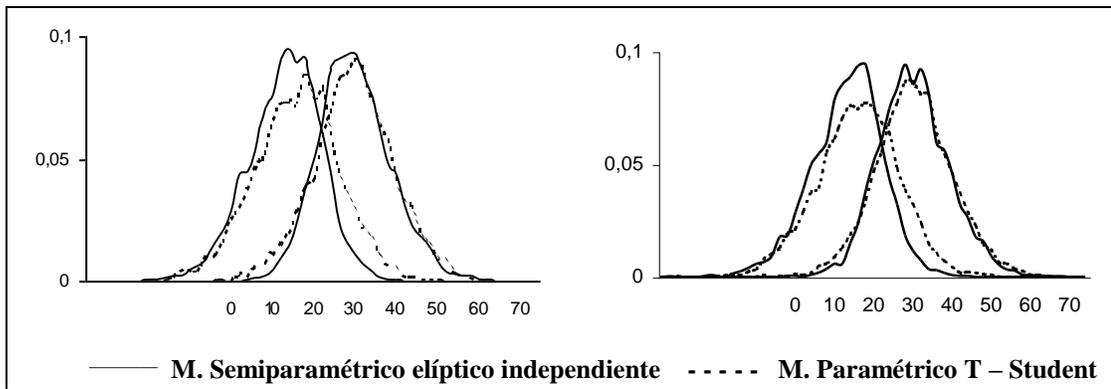
La Figura 4.2. presenta la distribución a priori de x_0 , que presenta una varianza muy grande, mientras que la Figura 4.3. presenta las distribuciones a posteriori para el mismo parámetro.

- Al respecto se puede ver que los grados de libertad no tienen mayor influencia sobre las distribuciones a posteriori del parámetro de interés, ni en los modelos t – student (paramétricos), ni en modelos elípticos en general (semiparamétricos).
- Las distribuciones a posteriori del parámetro x_0 son menos dispersas para modelos semiparamétricos elípticos que para los modelos t – student, por lo cual los estimadores propuestos (semiparamétricos elípticos) resultan más precisos. Esto puede deberse a que el modelo t – Student es más restrictivo.
- Las distribuciones a posteriori de x_0 para modelos paramétricos normales y modelos semiparamétricos elípticos son semejantes.
- Las distribuciones a posteriori de x_0 en todos los modelos estudiados no parecen diferenciarse mucho de la distribución normal.
- Las distribuciones a posteriori para el parámetro de interés son un tanto más concentradas que la distribución a priori.

- a) Distribuciones a posteriori de x_0 para los modelos con enlace semiparamétrico elíptico independiente (izquierda) y paramétrico t – student independiente (derecha). Curvas: izquierda, $y_0 = 0$; derecha, $y_0 = 1$.



- b) Distribución a posteriori de x_0 para los modelos con enlaces de 4 g. l. (izquierda) y 50 g. l. (derecha). Curvas: izquierda, $y_0 = 0$; derecha, $y_0 = 1$.



- c) Comparación de distribuciones a posteriori de x_0 para los modelos con enlace paramétrico normal (enlace probito) y con enlaces semiparamétricos elípticos independientes de 4 y 50 g. l. ($y_0 = 0$, a la izquierda; $y_0 = 1$, a la derecha).

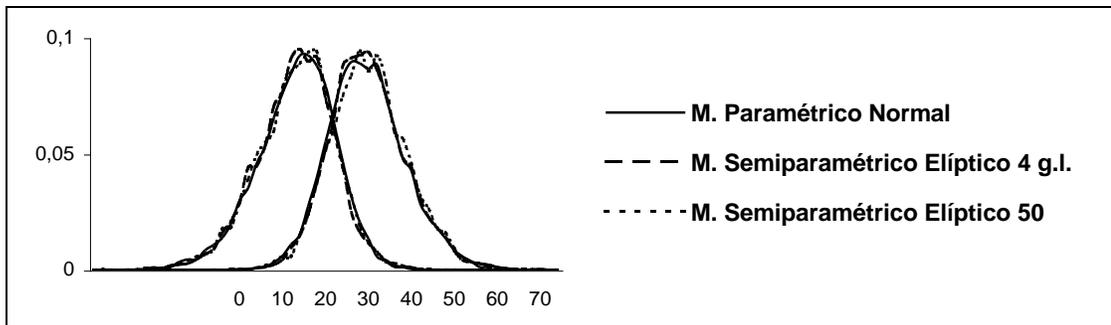


Figura 4.2: Distribuciones a posteriori para el parámetro de interés x_0

MODELO BINOMIAL

De acuerdo a lo comentado anteriormente, el modelo bernoulli no satisface las expectativas del problema de calibración. El análisis basado en el modelo de calibración binomial (*Corolarios 4.2. y 4.5.*), en cambio, permite estimar el índice de proliferación que corresponde a cierta conformación de una muestra. Los resultados de este análisis se presentan en la *Tabla 4.3.* para los distintos modelos y para resultados (s_i, n_i) presentes en la muestra.

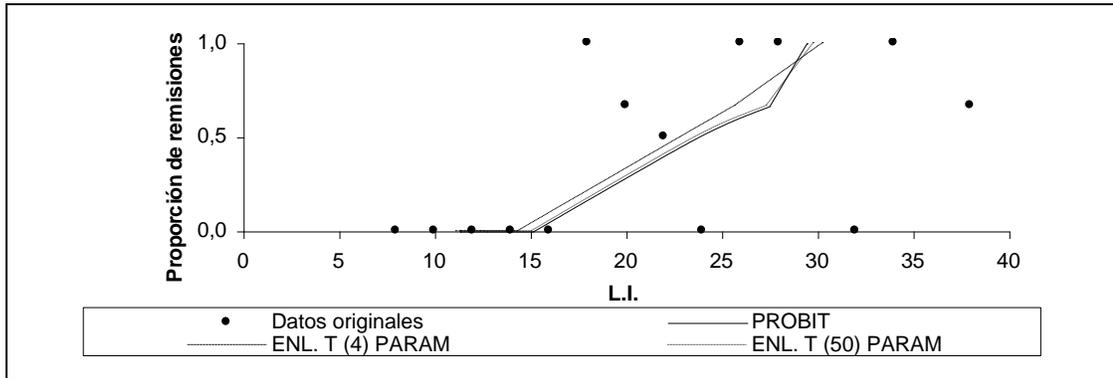
Tabla 4.3: Análisis de los datos de Lee (1974). Estimación bayesiana paramétrica y semi paramétrica del parámetro de interés mediante el modelo de calibración binomial, para los valores observados de s_i y n_i .

Estimación x_0 y β para proporción estimada s_i / n_i	Est. bayesiana semiparamétrica		Est. bayesiana paramétrica		
	4 g. l. ($r_1 = \lambda_1 = 2$)	50 g. l. ($r_1 = \lambda_1 = 25$)	Modelo Probit	Enlace t–student (4)	Enlace t–student (50)
0 / 3	$x_0 = 19.99$ $\beta^t = (-3.95, 0.12)$	$x_0 = 11.00$ $\beta^t = (-2.97, 0.11)$	$x_0 = 11.28$ $\beta^t = (-2.67, 0.10)$	$x_0 = 11.33$ $\beta^t = (-4.12, 0.16)$	$x_0 = 11.10$ $\beta^t = (-2.75, 0.10)$
0 / 2	$x_0 = 19.92$ $\beta^t = (-3.96, 0.11)$	$x_0 = 12.54$ $\beta^t = (-2.90, 0.11)$	$x_0 = 12.68$ $\beta^t = (-2.62, 0.10)$	$x_0 = 12.38$ $\beta^t = (-4.09, 0.16)$	$x_0 = 12.56$ $\beta^t = (-2.71, 0.10)$
0 / 1	$x_0 = 20.03$ $\beta^t = (-4.00, 0.11)$	$x_0 = 15.02$ $\beta^t = (-2.81, 0.11)$	$x_0 = 15.11$ $\beta^t = (-2.59, 0.10)$	$x_0 = 14.33$ $\beta^t = (-3.98, 0.16)$	$x_0 = 15.05$ $\beta^t = (-2.63, 0.10)$
1 / 2	$x_0 = 20.42$ $\beta^t = (-3.87, 0.13)$	$x_0 = 23.63$ $\beta^t = (-2.67, 0.10)$	$x_0 = 23.72$ $\beta^t = (-2.37, 0.09)$	$x_0 = 22.83$ $\beta^t = (-3.59, 0.14)$	$x_0 = 23.59$ $\beta^t = (-2.47, 0.09)$
2 / 3	$x_0 = 19.94$ $\beta^t = (-3.85, 0.13)$	$x_0 = 27.40$ $\beta^t = (-2.68, 0.10)$	$x_0 = 27.46$ $\beta^t = (-2.35, 0.09)$	$x_0 = 25.68$ $\beta^t = (-3.79, 0.15)$	$x_0 = 27.33$ $\beta^t = (-2.48, 0.10)$
1 / 1	$x_0 = 20.21$ $\beta^t = (-3.93, 0.14)$	$x_0 = 29.42$ $\beta^t = (-2.77, 0.11)$	$x_0 = 29.40$ $\beta^t = (-2.53, 0.10)$	$x_0 = 30.27$ $\beta^t = (-3.86, 0.16)$	$x_0 = 29.79$ $\beta^t = (-2.57, 0.10)$

Este análisis revela la ineficacia del modelo semiparamétrico elíptico para el caso en el cual las sub-muestras que definen cada valor de la predictora son pequeñas y el parámetro grados de libertad asociado a la priori de las variables subyacentes ω es pequeño. Los modelos semiparamétrico elíptico asociado a 50 g.l., los modelos paramétricos t – student y el modelo paramétrico normal, en cambio, discriminan adecuadamente y las estimaciones obtenidas para x_0 son semejantes. Por otra parte, las estimaciones de los coeficientes β_0 y β_1 en los modelos asociados a 50 g.l. (paramétrico y semiparamétrico) resultan semejantes entre sí y semejantes a las estimaciones provenientes del modelo paramétrico normal.

Ahora bien, la *Figura 4.3.* revela que efectivamente para nuestro conjunto de datos, relativos a remisión de cáncer, las estimaciones para modelos paramétrico y semiparamétrico con 50 g.l. se asemejan mucho a la estimación con el modelo paramétrico de enlace normal, la estimación con el modelo Semiparamétrico elíptico de 4 g.l. no discrimina entre los valores de x_0 y las estimaciones en base modelos paramétricos t – Student y normal se asemejan incluso para valores pequeños del parámetro grados de libertad.

a) Estimaciones de x_0 (L.I.) en base a modelos paramétricos con enlace normal (probit) y t – Student de 4 y 50 g.l.



b) Estimaciones de x_0 (L.I.) en base al modelo paramétrico con enlace normal (probit) y a los modelos semiparamétricos elípticos de 4 y 50 g.l.

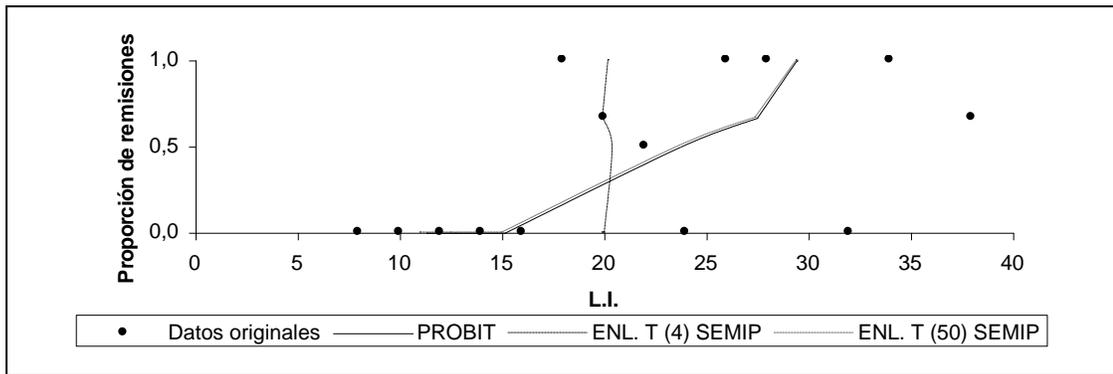


Figura 4.3: Estimación del Índice de Proliferación de las Células para tamaños de muestra n_i y número de remisiones s_i dados y presentes en la muestra, de acuerdo a la Tabla 4.3.(Lee, 1974).

La Tabla 4.4. presenta los errores cuadráticos medios obtenidos al obtener las sumas de cuadrados de las diferencias entre los valores de Índice de Proliferación de las células L.I. estimados en base al conjunto de datos en la Tabla 4.1., utilizando como información los pares (s_i, n_i) existentes en el conjunto y que se presentan en la Tabla 4.3. Aquí se comprueba la ineficacia, para este caso, del modelo semiparamétrico elíptico con 4 g.l., y la supremacía del modelo semiparamétrico elíptico con 50 g.l. sobre los modelos t – student; pero el mejor modelo resulta ser el paramétrico normal.

Tabla 4.4: Error Cuadrático Medio ECM para las estimaciones del parámetro de interés ajustado con los modelos de interés a los datos de Lee (1974).

Modelo	Paramétrico Normal	Paramétrico T – Student 4 g.l.	Paramétrico T – student 50 g.l.	Semiparamétrico 4 g.l.	Semiparamétrico 50 g.l.
ECM	757,87	832,33	772,08	1133,08	766,20

Aplicación 4.2 Simulación de un problema de Calibración binomial:

Se realizó la siguiente simulación de un problema de calibración discreta: La predictora x_i asume los enteros de 1 a 5 y se consideran tamaños de muestra $n_i = 20$ para cada x_i . El vector de coeficientes del modelo $\beta = (\beta_0, \beta_1)$ asume los valores $(-1, 0.05)$ y $(-1.5, 0.3)$; estos valores fueron escogidos con el fin de generar un modelo no significativo y significativo, respectivamente, y determinar si dicha elección incide en la estimación. Los enlaces pueden ser la f.d.a. normal estándar, t – student con 4 g.l. y logística estándar. A partir de los x_i se generan los valores $\alpha + \beta \cdot x_i$, en los cuales se evalúa la f.d.a. en cuestión. Esta evaluación da lugar a los parámetros $\theta_i = H(\alpha + \beta \cdot x_i)$. Finalmente, la variable respuesta s_i se genera aleatoriamente desde una distribución binomial de parámetros n_i y θ_i .

Tabla 4.5. Datos simulados para cada enlace de interés:

Enlace Parámetro	f.d.a. logística (logito)		f.d.a. normal (probit)		f.d.a. t – student (4)	
	x_i	s_i	x_i	s_i	x_i	s_i
$n_i = n = 20$ M. no significativo $\beta^t = (-1.0, 0.05)$	1	7	1	1	1	2
	2	5	2	4	2	7
	3	6	3	3	3	7
	4	6	4	6	4	4
	5	8	5	5	5	5
$n_i = n = 20$ M. significativo $\beta^t = (-1.5, 0.30)$	1	5	1	9	1	1
	2	8	2	3	2	8
	3	6	3	5	3	7
	4	10	4	6	4	4
	5	10	5	11	5	7

La *Tabla 4.5.* muestra los datos simulados con coeficientes significativos y no significativos y con los tres enlaces considerados; mientras que la *Tabla 4.6.* muestra los análisis para dicha información utilizando modelo semiparamétrico, modelo paramétrico con enlace t – student (4) y modelo paramétrico con enlace normal (probit), con el fin de estimar $x_3 = 3$, en base a los valores simulados de s_3 y n_3 .

Las distribuciones a priori utilizadas para el análisis son las siguientes:

- Los coeficientes β_0 y β_1 se asumen a priori normales con medias iguales a las medias reales utilizadas para simular, pero con varianzas grandes, 10 y 1 respectivamente.
- El parámetro de interés se asume a priori normal con media igual a la media real y con varianza 25 (*Figura 4.4.*).
- μ_X y σ_X asumen los valores 0 y 25, respectivamente.

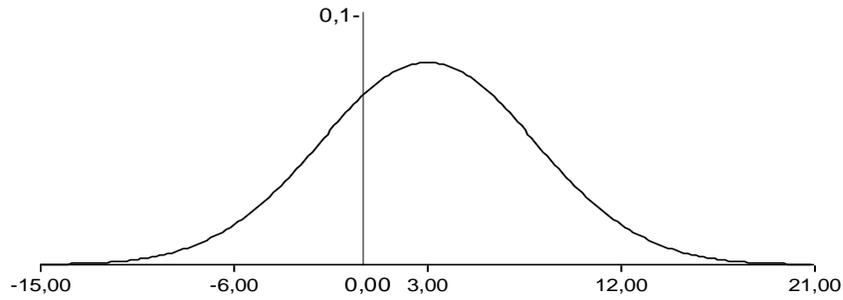


Figura 4.4: Distribuciones a priori para el parámetro de interés x_3 en el análisis de los datos simulados.

Tabla 4.6. Análisis de los datos simulados y presentados en la Tabla 4.5. mediante los modelos paramétricos y semiparamétricos de interés:

Enlace Modelo original y Modelo para Estimar	f.d.a. logística (logito)		f.d.a. normal (probit)		f.d.a. t – student (4)	
	x_3	s_3	x_3	s_3	x_3	s_3
M. original no significativo: $\beta^t = (-1.0, 0.05)$	$x_3 = 3$	$s_3 = 6$	$x_3 = 3$	$s_3 = 3$	$x_3 = 3$	$s_3 = 7$
Estimación semiparamétrica elíptica independiente de x_3 con c aleatorio	$\hat{x}_3 = 3.0181$		$\hat{x}_3 = 2.9726$		$\hat{x}_3 = 3.0267$	
Estimación semiparamétrica elíptica independiente de x_3 con c fijo igual a uno	$\hat{x}_3 = 3.0174$		$\hat{x}_3 = 2.9423$		$\hat{x}_3 = 2.9048$	
Estimación paramétrica t – student (4) de x_3	$\hat{x}_3 = 3.0992$		$\hat{x}_3 = 3.5376$		$\hat{x}_3 = 3.1333$	
Estimación paramétrica normal de x_3	$\hat{x}_3 = 2.8512$		$\hat{x}_3 = 2.1836$		$\hat{x}_3 = 3.6940$	
M. original significativo: $\beta^t = (-1.5, 0.30)$	$x_3 = 3$	$s_3 = 6$	$x_3 = 3$	$s_3 = 5$	$x_3 = 3$	$s_3 = 7$
Estimación semiparamétrica elíptica independiente de x_3 con c aleatorio	$\hat{x}_3 = 2.8618$		$\hat{x}_3 = 2.8554$		$\hat{x}_3 = 3.0222$	
Estimación semiparamétrica elíptica independiente de x_3 con c fijo igual a uno	$\hat{x}_3 = 2.9413$		$\hat{x}_3 = 2.9578$		$\hat{x}_3 = 3.1133$	
Estimación paramétrica t – student (4) de x_3	$\hat{x}_3 = 2.8701$		$\hat{x}_3 = 2.8088$		$\hat{x}_3 = 2.9132$	
Estimación paramétrica normal de x_3	$\hat{x}_3 = 1.7167$		$\hat{x}_3 = 1.6769$		$\hat{x}_3 = 4.3470$	

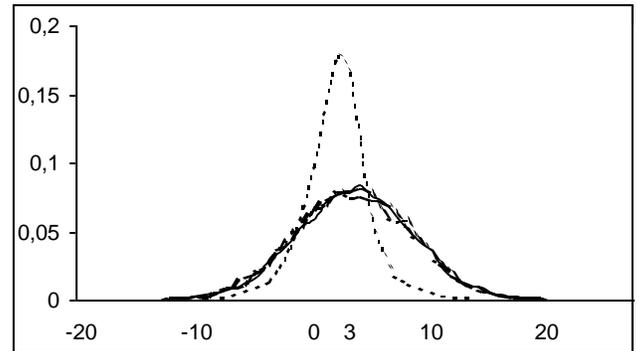
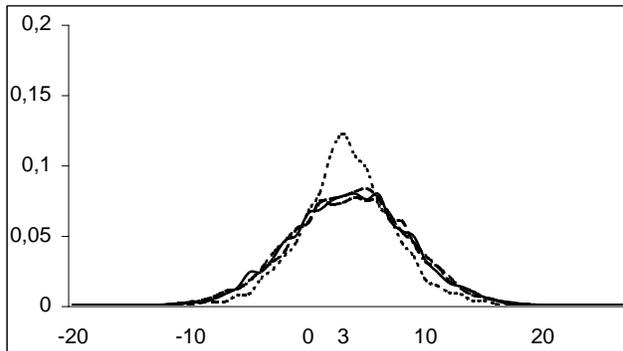
Es muy interesante verificar que:

- En general, las estimaciones para el modelo significativo (con pendiente mayor) resultan más precisas que las análogas para el modelo no significativo.
- Por otra parte, para todos los conjuntos de datos simulados con enlace f.d.a. logística estándar o logito, con enlace normal o probito y con enlace t – student, siempre la mejor estimación la entrega el modelo semiparamétrico elíptico independiente.
- Finalmente, la estimación con el modelo paramétrico t – student es mejor que la estimación con el modelo paramétrico normal.

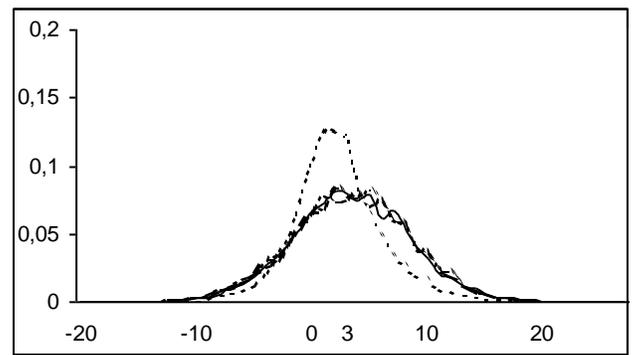
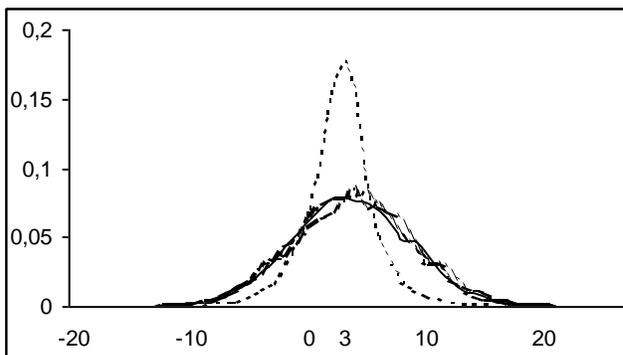
Finalmente, la *Figura 4.5.* presenta las distribuciones a posteriori del parámetro de interés x_0 bajo cada modelo de interés, a partir de la cual se confirman los hallazgos de la *Tabla 4.6.*, y se puede agregar que:

- La distribución a posteriori de x_3 para conjuntos de datos analizados con el modelo paramétrico normal es menos dispersa que las distribuciones a posterioris del parámetro de interés para conjuntos analizados con otros modelos.
- Las distribuciones a posteriori para x_3 en conjuntos de datos analizados con modelos semiparamétricos elípticos o paramétricos t – student no difieren significativamente.
- En general las distribuciones a posteriori para x_3 en conjuntos de datos analizados con modelos semiparamétricos elípticos o paramétricos t – student se presentan centradas en el verdadero valor del parámetro a estimar.
- La distribución a posteriori de x_3 para conjuntos de datos analizados con el modelo paramétrico normal es en general sesgada, principalmente para conjuntos de datos simulados con enlace t – student.

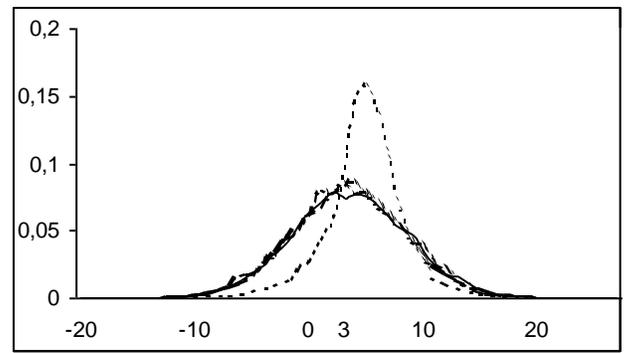
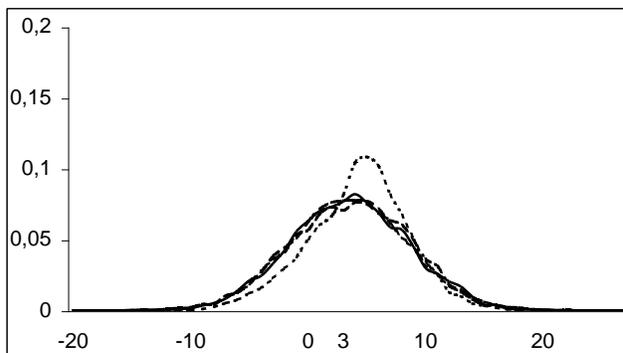
a) Modelo de calibración discreta, simulado con enlace f.d.a. logística estándar o logito (pendiente no significativa a la izquierda y significativa a la derecha)



b) Modelo de calibración discreta, simulado con enlace f.d.a. normal estándar o probito (pendiente no significativa a la izquierda y significativa a la derecha)



c) Modelo de calibración discreta, simulado con enlace f.d.a. t – student con 4 grados de libertad (pendiente no significativa a la izquierda y significativa a la derecha)



M Elíptico Semiparamétrico con c aleatorio
 M Paramétrico T – Student
 M Elíptico Semiparamétrico con c fijo
 M Paramétrico Normal

Figura 4.5: Distribuciones a posteriori para el parámetro de interés x_3 en base a los modelos simulados como significativo y no significativo, y analizados con modelos semiparamétricos elípticos y paramétricos.

Aplicación 4.3 Experimento toxicológico de mortalidad de escarabajos (Bliss, 1935):

Un experimento toxicológico está diseñado de tal modo que la variable de interés es el número de escarabajos muertos luego de cinco horas de exposición a gas carbónico en varias concentraciones distintas. La *Tabla 4.7.* y la *Figura 4.4.* presentan dicha información, obtenida por Bliss (1935) y analizada por Basu y Mukhopadhyay (2000).

Tabla 4.7. Datos relativos a un experimento toxicológico de mortalidad de escarabajos. La variable explicativa x_i es el logaritmo del nivel de concentración de gas carbónico (dosis), n_i corresponde al número de escarabajos expuestos a dicha concentración de tóxico y la variable respuesta binomial s_i es el número de escarabajos muertos para cada grupo.

Log (Dosis)	Escarabajos expuestos	Escarabajos muertos	Porcentaje de mortalidad
1,6907	59	6	10,17%
1,7242	60	13	21,67%
1,7552	62	18	29,03%
1,7842	56	28	50,00%
1,8113	63	52	82,54%
1,8369	59	53	89,83%
1,8610	62	61	98,39%
1,8839	60	60	100,00%

El análisis clásico del problema, mediante el modelo logístico concluye que la concentración del tóxico resulta significativa ($p < 0,0001$) y entrega estimadores para los coeficientes del modelo, en este caso, $\beta_0 = -60,17$ y $\beta_1 = 33,97$.

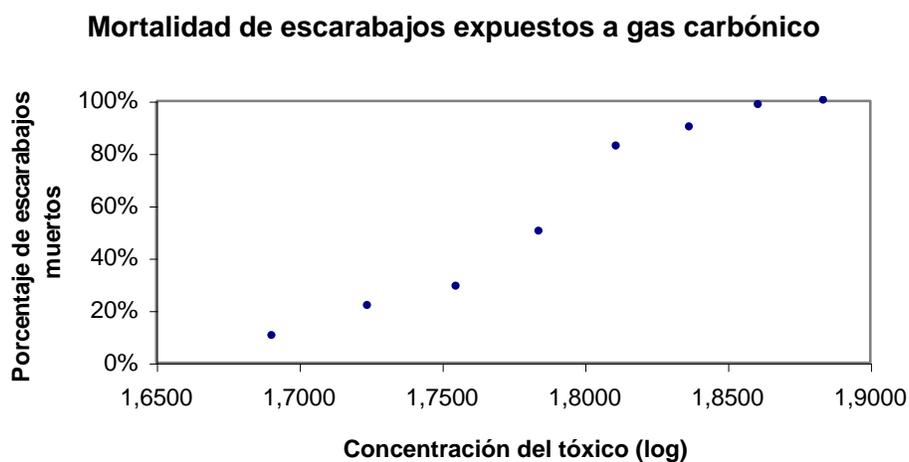


Figura 4.4. Porcentaje de Escarabajos muertos para varios niveles de concentración de gas carbónico.

Por otra parte, Basu y Mukhopadhyay (2000) realizan ajustes para el modelo de regresión con enlaces mezcla escala finita de normales, donde los pesos siguen una distribución Dirichlet. Estos autores comparan el ajuste de los modelos con enlaces simétrico y asimétrico.

El análisis de los datos se realizó utilizando el modelo de calibración binomial para los modelos bayesiano paramétrico y semi-paramétrico. Las distribuciones a priori para el modelo jerárquico consideraron para la distribución de ω los hiperparámetros λ_1 y r_1 tales que $\lambda_1 = r_1 = 2$ y $\lambda_1 = r_1 = 25$ (modelos de 4 y 50 g.l., respectivamente); mientras que para la distribución de los coeficientes, se consideraron medias semejantes a los estimadores del modelo logístico $b_0 = -60$ y $b_1 = 34$, con varianzas no pequeñas $v_0 = 400$ y $v_1 = 400$.

Por otra parte, para el parámetro de interés x_0 la media de la priori normal $c_0 = 1,8$ es aproximadamente igual al punto medio del intervalo de los datos, y la varianza es $d = 0.02$, mientras que los valores iniciales para generar los valores z_{ij} son $\mu_X = 0,8$ y $\sigma^2_X = 4$, donde la media es aproximadamente igual al valor central del conjunto de las posibles medias para los z_{ij} .

Tabla 4.8. Análisis de datos de la Tabla 4.7. Modelos bayesianos semiparamétrico, paramétrico y resultados de Basu y Mukhopadhyay (2000).

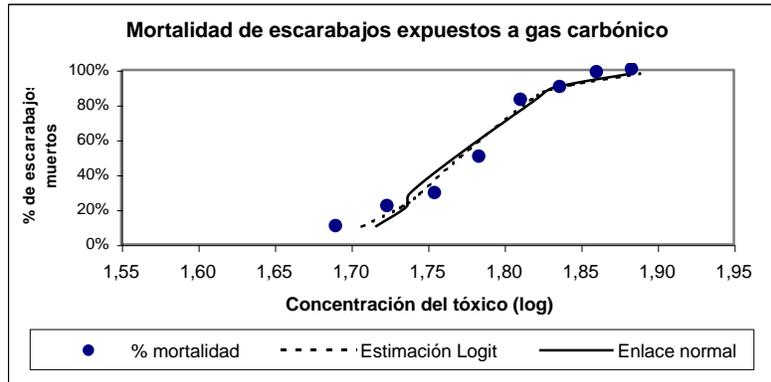
Modelo Bayesiano		SCE
Semi-paramétrico con c fijo	4 g.l.	0,0044
	50 g.l.	0,0042
Semi-paramétrico con c aleatorio	4 g.l.	0,0048
	50 g.l.	0,0040
Paramétrico	Probit	0,0119
	T – Student (4)	0,0691
	T – Student (50)	0,0699
Basu y Mukhopadhyay	Enlace Simétrico	91,32*
	Enlace Asimétrico	49,36*

* Estimación se s_i .

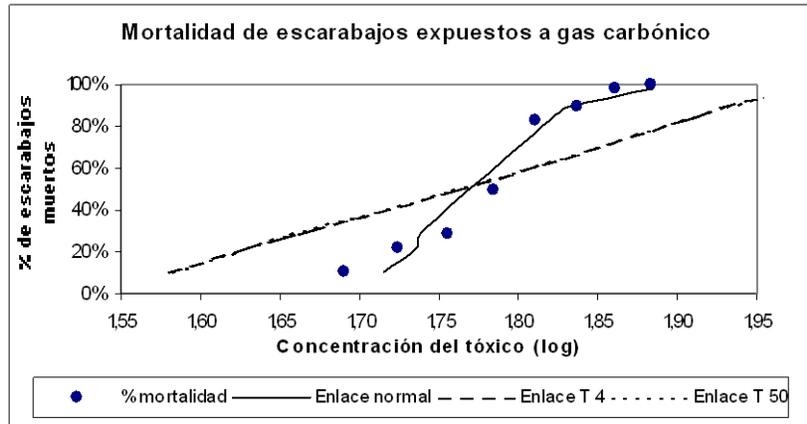
Ahora bien, la Tabla 4.8. presenta el análisis de los datos utilizando el modelo de calibración binomial para los modelos bayesiano paramétrico y semi-paramétrico (Corolario 4.2. y Corolario 4.5.), así como los resultados de los análisis semiparamétricos de Basu y Mukhopadhyay (2000), utilizando enlaces simétrico y asimétrico. La estimación se realizó eliminando sucesivamente cada grupo y utilizando los valores de escarabajos expuestos y muertos (n_0 y s_0) para estimar la dosis respectiva (x_0).

La gran diferencia en los SCE obtenidos por nosotros se debe a que la variable a estimar para los autores Basu y Mukhopadhyay es s_i (número de escarabajos muertos), mientras que para nosotros es x_i (log – dosis de tóxico), y no es factible realizar la conversión debido a que ellos asumen x_i fijos y nosotros asumimos s_i fijos.

a) Estimación de log – dosis versus porcentaje de escarabajos muertos para modelos logit (clásico) y bayesiano paramétrico con enlace normal (probit).



b) Estimación de log – dosis versus porcentaje de escarabajos muertos para modelos bayesiano paramétrico con enlace normal (probit) y con enlaces t – student de 4 y 50 g.l.



c) Estimación de log – dosis versus porcentaje de escarabajos muertos para modelos bayesiano paramétrico con enlace normal (probit) y semiparamétricos con enlaces elípticos 4 y 50 g.l.

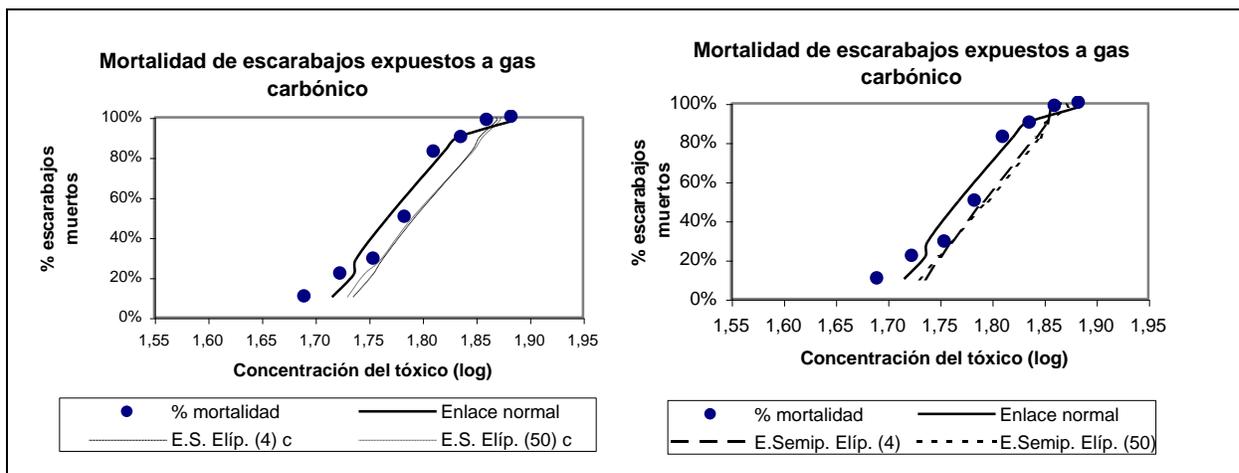


Figura 4.5. Estimación de log – dosis de tóxico vs porcentaje de escarabajos muertos en modelos de interés.

Ahora bien, por una parte es claro que los enlaces asimétricos utilizados por estos autores son mejores que los enlaces simétricos utilizados por ellos mismos (estos últimos corresponden a una priori discreta para los pesos en la mezcla). Por otra parte, respecto de los enlaces que hemos utilizado, el enlace logito no admite estimación para la última cantidad, de modo que SCE no es comparable con la de los otros enlaces, pero respecto de los enlaces t – student, probito y Semiparamétrico elíptico podemos deducir algunas conclusiones (*Tabla 4.8.*):

- El modelo paramétrico con enlace normal (probit) es más preciso que los modelos paramétricos con enlace t – student.
- Es importante enfatizar que los modelos t – student resultan ser del todo inadecuados para representar al conjunto de datos, mientras que el modelo normal es bastante robusto respecto de la asimetría presente.
- Los modelos paramétricos con enlace t – student son muy semejantes, independientemente del valor de sus grados de libertad, pero resultan muy imprecisos.
- Las mejores estimaciones las producen los modelos semiparamétricos elípticos, independientemente de que el valor del parámetro c sea fijo o aleatorio.

Ahora bien, la *Figura 4.5.* muestra las estimaciones de log – concentración del tóxico versus porcentaje de escarabajos muertos para los modelos de interés. Estos gráficos confirman los hallazgos provenientes de la *Tabla 4.8.* y llevan también a las siguientes bservaciones pertinentes:

- El análisis clásico del modelo logito entrega estimaciones muy parecidas al análisis bayesiano paramétrico con enlace normal (probit).
- Las estimaciones con modelos bayesianos de enlaces semiparamétricos elípticos sobre estiman levemente el valor del parámetro, sin embargo, entregan las estimaciones más precisas, y en este sentido también presentan robustez frente a la asimetría presente en el problema.

CONCLUSIONES

Este capítulo aborda el análisis bayesiano paramétrico y semiparamétrico del modelo de calibración lineal discreta.

Antes de abordar el problema, desde el enfoque bayesiano, se obtuvieron tanto la función de verosimilitud asociada al problema como las ecuaciones que se requiere resolver para obtener los estimadores máximo verosímiles de los parámetros. El hecho de que las expresiones y ecuaciones obtenidas resulten analíticamente inmanejables justifican el enfoque semiparamétrico asumido.

La resolución del problema se logra mediante la implementación de muestreo de Gibbs, lo cual requiere la obtención de todas las distribuciones condicionales completas de los parámetros. Los estimadores de Bayes bajo pérdida cuadrática, correspondientes a las medias de las distribuciones a posteriori, se obtienen finalmente mediante la simulación computacional, utilizando dichas distribuciones condicionales completas.

La metodología desarrollada en este capítulo se enfoque en dos variantes para la función de enlace, estas son la función de distribución normal estándar y la función de distribución elíptica representable independiente.

Los resultados obtenidos en este capítulo se basan en el lema de Basu y Mukhopadhyay (2000). Un primer aporte presentado en este capítulo consiste en la extensión de los resultados de los autores antes mencionados para el modelo de regresión lineal binaria, al modelo de calibración lineal binaria. Aquí se consideran el caso paramétrico (la distribución a priori de los pesos en la mezcla de normales es discreta) y semiparamétrico (la priori para la función de distribución de los pesos en la mezcla es un proceso Dirichlet). Un segundo aporte de trascendencia en el área de la dosimetría, consiste en la resolución del problema de calibración binomial bajo el enfoque paramétrico. Estos resultados se aplican al caso particular paramétrico en el cual el enlace corresponde a una función de distribución elíptica conocida (enlace $t - Student$), presentándose así la resolución paramétrica del problema de calibración lineal binaria y binomial.

La metodología desarrollada se utilizó para resolver dos problemas presentes en la literatura, *Características asociadas a la remisión del cáncer* (Lee, 1974) y *Experimento toxicológico de mortalidad de escarabajos* (Bliss, 1935). También se realizó una simulación con el fin de determinar si la magnitud de la pendiente del modelo incide en la estimación y para qué enlaces se obtienen mejores estimaciones.

De los resultados de la simulación surge que, efectivamente, la magnitud de la pendiente incide en la precisión de la estimación. Por otra parte, para cualquier enlace utilizado, la mejor estimación corresponde al modelo semiparamétrico elíptico independiente, aunque sus posterioris se asemejan a las del modelo paramétrico $t - Student$. También se aprecia que las posterioris del modelo paramétrico normal son las menos dispersas pero las más sesgadas, en especial cuando se considera enlace $t - Student$, y por lo tanto sus estimaciones no son tan buenas como las que surgen del modelo paramétrico $t - Student$.

Los análisis de convergencia de las cadenas arrojan buenos resultados y resultan, en general, sin autocorrelación.

En todos los problemas abordados, se utilizaron las distribuciones a priori usuales para los coeficientes del modelo y para las variables predictoras (coeficientes y predictoras normales), y se utilizaron funciones de enlace basadas en distribuciones simétricas, por lo cual se propone extender los resultados considerando otras distribuciones y enlaces asimétricos (por ejemplo skew normal y skew elípticos). Por otra parte, nuestros desarrollos consideraron coeficientes independientes entre sí, e independientes del resto de los parámetros, por lo cual una extensión interesante se relaciona con estructuras de dependencia distintas. Finalmente, nos hemos restringido al caso de una sola predictora, lo que sugiere la extensión al caso de regresión múltiple.

CAPÍTULO 5:

SOLUCIÓN SEMIPARAMÉTRICA DEL MODELO CON ERRORES EN LAS VARIABLES

La primera mención en la literatura del Modelo de Errores de medición (MEM) o Modelo de Errores en las variables, es debida a Neyman y Scout (1948), quienes lo presentan como un ejemplo de inconsistencia del estimador máximo verosímil para parámetros estructurales en presencia de infinitos parámetros de ruido. Casi una década después, Kiefer y Wolfowitz (1956) proponen que los parámetros de ruido se asuman como variables aleatorias provenientes de una distribución de probabilidad no especificada. Con este supuesto demuestran la consistencia del estimador máximo verosímil. La primera implementación computacional se debe a Laird (1978).

Fuller (1987) aborda el problema del MEM sin datos faltantes, con errores distribuidos normal. Carroll, Freedman y Pee (1997) y Carroll, Roedor y Wasserman (1999) extienden el problema al caso con datos faltantes para distribución normal y mezcla de normales, respectivamente.

Clayton (1991), Lindsay y Lesperance (1995) y Roedor, Carroll y Lindsay (1996) retoman los resultados de Kiefer y Wolfowitz (1956), aplicándolos al problema de estudios case – control. Ahora bien, Zhao y Lee (1996), Kuchenhoff y Carroll (1997), Carroll, Freedman y Pee (1997), Carroll, Roedor y Wasserman (1999) y Higdom y Schaefer (1999) prueban la eficiencia de los estimadores máximo verosímiles pero requieren de la especificación de las distribuciones. Más tarde, Schaefer (2001) presenta un algoritmo del tipo EM que no requiere de tal especificación.

Stefanski y Carroll (1990), Bolfarine y Cordani (1993) y Zhang y Karunamuni (1996) estiman los parámetros de regresión del MEM desde el punto de vista Bayesiano.

La introducción de los métodos de simulación MCMC permite aplicaciones al problema del MEM. Algunas publicaciones en este sentido, restringidas al supuesto de normalidad, son: Stephens y Dellaportas (1992), Richardson y Gilks (1993) y Dellaportas y Stephens (1995).

La extensión de los MEM para errores con distribución elíptica, desde la perspectiva clásica, se debe a Bolfarine y Arellano – Valle (1994, 1998), Arellano – Valle y Bolfarine (1995), Arellano – Valle, Bolfarine y Vilca – Labra (1997) y Vilca – Labra, Arellano – Valle y Bolfarine (1998).

Finalmente, Arellano – Valle, Galea – Rojas e Iglesias (2000) analizan el problema del MEM considerando errores Elípticos, desde la perspectiva bayesiana paramétrica.

El objetivo fundamental de este capítulo es extender estos resultados desde el punto de vista bayesiano, considerando distribuciones elípticas para los errores, pero desde una perspectiva Semiparamétrica, lo que creemos no ha sido abordado en la literatura existente.

El modelo general de errores de medición MEM tiene la forma siguiente:

$$(5.1) \quad \begin{aligned} y_i &= \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \dots, n, \\ \text{y} \quad z_i &= x_i + u_i, \quad i = 1, \dots, n, \end{aligned}$$

donde, por una parte, la relación entre la variable respuesta observable y_i y la variable predictora no observable x_i está dada por un modelo de regresión lineal simple con errores ε_i y con coeficientes β_0 y β_1 ; por otro lado, aunque la variable x_i es no observable, se observa con un error aditivo u_i , lo cual da origen a la variable z_i .

Este capítulo se divide en tres secciones. En las primeras dos secciones se desarrolla una metodología bayesiana semiparamétrica para el problema de errores de medición MEM para los casos en que los errores tienen una distribución elíptica representable independiente, es decir, cada observación posee su propio parámetro de dispersión; el caso representable dependiente, donde las observaciones tienen un parámetro de dispersión común, es tema de la segunda sección.

Para ambos casos se requiere la implementación de métodos de simulación MCMC, debido a las dificultades analíticas de un desarrollo alternativo. En este trabajo se utiliza específicamente el muestreo de Gibbs, explotando la relación entre el modelo elíptico representable y el modelo normal heteroscedástico. Para ambas secciones, con el fin de ilustrar, se comienza abordando un caso paramétrico que extiende los resultados hallados en la literatura, y que consiste en considerar que los errores son mezcla escala finita de normales o lo que se conoce como normal contaminada. Al vector de los pesos en esta mezcla se le asigna a priori una distribución Dirichlet, y de esta forma se robustifica el modelo normal. La extensión natural corresponde al caso Semiparamétrico, donde no nos amarramos a una mezcla determinada, sino que los parámetros que definen la mezcla siguen un proceso Dirichlet. Todos los casos abordados utilizan distribuciones a priori usuales en la literatura, las cuales permiten comparación de los resultados.

Finalmente, en la tercera sección los resultados son ilustrados con conjuntos de datos utilizados previamente en la literatura y con una simulación.

5.1 SOLUCIÓN BAYESIANA SEMIPARAMÉTRICA BAJO EL MODELO DE REGRESIÓN ELÍPTICO REPRESENTABLE INDEPENDIENTE CON ERRORES DE MEDICIÓN MEM

Esta sección aborda el modelo de regresión con errores de medición MEM o modelo con errores en las variables. Se considera en este caso que los términos de error son independientes e idénticamente distribuidos elípticamente, con funciones generadoras de densidad h_ε y h_U , representables. Por lo tanto el modelo se puede escribir como:

$$(5.2) \quad \begin{aligned} y_i | \beta, \mathbf{x}, \phi_\varepsilon &\sim \text{ind. } El_1(\mathbf{x}_i^T \beta, \phi_\varepsilon^{-1}, h_\varepsilon) \quad \text{y} \quad h_\varepsilon(u) = \int_0^\infty (2\pi\omega)^{-1/2} e^{-u/2\omega} dG_\varepsilon(\omega), \\ z_i | \mathbf{x}, \phi_U &\sim \text{ind. } El_1(\mathbf{x}_i, \phi_U^{-1}, h_U) \quad \text{y} \quad h_U(u) = \int_0^\infty (2\pi\omega)^{-1/2} e^{-u/2\omega} dG_U(\omega), \end{aligned}$$

con h_ε y h_U no negativas y con G_ε y G_U funciones de distribución acumuladas (llamadas mezcladoras) tales que $G_\varepsilon(0) = G_U(0) = 0$.

Los parámetros de este modelo son $(\phi_\varepsilon, \phi_U, \beta, \mathbf{x})$ y (G_ε, G_U) , todos desconocidos. Los parámetros (G_ε, G_U) son infinito dimensionales, por lo tanto los resultados que a continuación se presentan extienden los hallados en la literatura desde el modelo normal paramétrico al elíptico semiparamétrico.

El modelo jerárquico de interés, basado en las distribuciones a priori presentes en la literatura, viene dado por

$$\begin{aligned}
 \varepsilon_i \mid \omega_{\varepsilon i}, G_{\varepsilon}, \phi_{\varepsilon}, \beta, x_i &\sim \text{ind. } EL_1(0, \phi_{\varepsilon}^{-1}, h_{\varepsilon}), i=1, \dots, n, \\
 u_i \mid \omega_{u i}, G_u, \phi_u, x_i &\sim \text{ind. } EL_1(0, \phi_u^{-1}, h_u), i=1, \dots, n, \\
 (5.3) \quad x_i \mid \phi_{\varepsilon}, \phi_u, \beta &\sim \text{i.i.d. } N_1(\mu_x, \sigma_x^2), i=1, \dots, n, \\
 \phi_{\varepsilon} \mid \phi_u, \beta &\sim \text{Gamma}(r_0, \lambda_0), \\
 \phi_u \mid \beta &\sim \text{Gamma}(r_1, \lambda_1), \\
 \beta_0 &\sim N(b_0, V_0) \\
 y \quad \beta_1 &\sim N(b_1, V_1).
 \end{aligned}$$

De manera análoga a lo obtenido para el problema de calibración lineal continua (capítulo 3), los resultados para el caso en el cual los errores tienen distribución elíptica representable e independiente se obtienen a partir del caso normal, mediante la aplicación de la *Proposición 2.1*, sección 2.2.

Como se explicó anteriormente, este modelo es equivalente al modelo en el cual las primeras dos condiciones se han extendido a las cuatro condiciones equivalentes siguientes, mediante la inclusión de variables latentes que, desde la *Proposición 2.1. b*, nos permite abordar el problema considerando errores normales.

Si llamamos \mathbf{x} al vector de componentes x_i , las primeras dos etapas del modelo jerárquico (5.3) se pueden expresar equivalentemente como:

$$\begin{aligned}
 (5.4) \quad y_i \mid \omega_{\varepsilon i}, \omega_{u i}, \phi_{\varepsilon}, \phi_u, z_i, x_i, \beta &\sim \text{con. ind. } N_1(\beta_0 + \beta_1 x_i, \omega_{\varepsilon i} \phi_{\varepsilon}^{-1}), \\
 z_i \mid \omega_{\varepsilon i}, \omega_{u i}, \phi_{\varepsilon}, \phi_u, y_i, x_i &\sim \text{cond. ind. } N_1(x_i, \omega_{u i} \phi_u^{-1}), \\
 \omega_{\varepsilon i} \mid G_{\varepsilon} &\sim \text{i.i.d. } G_{\varepsilon} \text{ con } G_{\varepsilon} \text{ independiente de } \phi_{\varepsilon}, \mathbf{x} \text{ y de } \beta. \\
 y \quad \omega_{u i} \mid G_u &\sim \text{i.i.d. } G_u \text{ con } G_u \text{ independiente de } \phi_u \text{ y de } \mathbf{x}.
 \end{aligned}$$

Es claro que las especificaciones en el modelo jerárquico (5.3) no permiten determinar explícitamente las distribuciones condicionales completas requeridas para la implementación de muestreo de Gibbs, mientras que la especificación análoga basada en (5.4) si lo permite.

En efecto, la verosimilitud para el modelo (5.3) es la siguiente:

$$\begin{aligned}
 L(\beta, \mathbf{z}, \phi_{\varepsilon}, \phi_u \mid \mathbf{y}, G_{\varepsilon}, G_u) &= (\phi_{\varepsilon} \cdot \phi_u)^{n/2} (2\pi)^{-n} \left[\prod_{i=1}^n \int_0^{\infty} \omega_{\varepsilon i}^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi_{\varepsilon} \cdot (y_i - \beta_0 - \beta_1 \cdot x_i)^2 / \omega_{\varepsilon i}) dG_{\varepsilon}(\omega_{\varepsilon i}) \right] \\
 &\cdot \left\{ \int_0^{\infty} \omega_{u i}^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi_u \cdot (z_i - x_i)^2 / \omega_{u i}) dG_u(\omega_{u i}) \right\}.
 \end{aligned}$$

Ahora bien, para el modelo basado en (5.4) se obtiene la siguiente función de verosimilitud aumentada, la cual coincide con la verosimilitud de un modelo de regresión normal heteroscedástico:

$$\begin{aligned}
 L(\beta, \mathbf{z}, \phi_{\varepsilon}, \phi_u, \omega_{\varepsilon}, \omega_u \mid \mathbf{y}, G_{\varepsilon}, G_u) &= (\phi_{\varepsilon} \cdot \phi_u)^{n/2} (2\pi)^{-n} \left[\prod_{i=1}^n (\omega_{\varepsilon i} \cdot \omega_{u i})^{-\frac{1}{2}} \right] \\
 &\cdot \exp[-\frac{1}{2} (\phi_{\varepsilon} \cdot \sum_{i=0}^n \omega_{\varepsilon i}^{-1} (y_i - \beta_0 - \beta_1 \cdot x_i)^2 + \phi_u \cdot \sum_{i=0}^n \omega_{u i}^{-1} (z_i - x_i)^2)].
 \end{aligned}$$

Desde el punto de vista clásico resulta infactible realizar la estimación máximo verosímil con esta última expresión sin tener información acerca de G_ε y G_u . Ahora bien, para la primera especificación, incluso considerando G_ε y G_u convenientemente escogidas, la expresión es muy complicada de resolver, pues requiere que las derivadas de la verosimilitud $L(\beta, z, \phi_\varepsilon, \phi_u | y, G_\varepsilon, G_u)$ o de la log verosimilitud respecto de los parámetros $\beta, z, \phi_\varepsilon$ y ϕ_u , las cuales son complicadísimas, sean nulas. Por otra parte, para la segunda especificación, para la verosimilitud aumentada $L(\beta, z, \phi_\varepsilon, \phi_u, \omega_\varepsilon, \omega_u | y, G_\varepsilon, G_u)$, se requiere que G_ε y G_u sean conocidas, además de que aparecen los grupos de variables instrumentales $\omega_{\varepsilon i}$ y $\omega_{u i}$ que también requieren ser estimadas.

En efecto, si para la primera expresión suponemos que G_ε y G_u tienen asociadas funciones de densidad g_ε y g_u , respectivamente, cuya naturaleza permite intercambiar integrales respecto de $\omega_{\varepsilon i}$ o de $\omega_{u i}$ y derivadas respecto de $\beta, z, \phi_\varepsilon$ y ϕ_u , y si definimos finalmente las funciones siguientes:

$$K_\varepsilon(e_i, \phi_\varepsilon) = \left[\int_0^\infty \omega_{\varepsilon i}^{-\frac{3}{2}} \exp(-\frac{1}{2} \phi_\varepsilon \cdot e_i^2 / \omega_{\varepsilon i}) \cdot g_\varepsilon(\omega_{\varepsilon i}) d\omega_{\varepsilon i} \right] / \left[\int_0^\infty \omega_{\varepsilon i}^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi_\varepsilon \cdot e_i^2 / \omega_{\varepsilon i}) \cdot g_\varepsilon(\omega_{\varepsilon i}) d\omega_{\varepsilon i} \right]$$

y

$$K_u(e_i, \phi_u) = \left[\int_0^\infty \omega_{u i}^{-\frac{3}{2}} \exp(-\frac{1}{2} \phi_u \cdot e_i^2 / \omega_{u i}) \cdot g_u(\omega_{u i}) d\omega_{u i} \right] / \left[\int_0^\infty \omega_{u i}^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi_u \cdot e_i^2 / \omega_{u i}) \cdot g_u(\omega_{u i}) d\omega_{u i} \right],$$

los estimadores máximo verosímiles para los parámetros deben satisfacer que:

$$\begin{aligned} \frac{\phi_\varepsilon}{n} \cdot \sum_{i=1}^n e_i^2 \cdot K_\varepsilon(e_i, \phi_\varepsilon) &= 1; & \frac{\phi_u}{n} \cdot \sum_{i=1}^n a_i^2 \cdot K_u(a_i, \phi_u) &= 1; & \sum_{i=1}^n e_i \cdot k_\varepsilon(e_i, \phi_\varepsilon) &= 0; \\ \sum_{i=0}^n e_i \cdot x_i \cdot k_\varepsilon(e_i, \phi_\varepsilon) &= 0, & y \quad x_i &= \frac{k_\varepsilon(e_i, \phi_\varepsilon) \cdot \phi_\varepsilon \cdot \beta_1^2 \cdot \left(\frac{y_i - \beta_0}{\beta_1} \right) + k_u(a_i, \phi_u) \cdot \phi_u \cdot z_i}{k_\varepsilon(e_i, \phi_\varepsilon) \cdot \phi_\varepsilon \cdot \beta_1^2 + k_u(a_i, \phi_u) \cdot \phi_u}, \end{aligned}$$

donde $e_i = (y_i - \beta_0 - \beta_1 \cdot x_i)$ y $a_i = (z_i - x_i)$.

La tercera ecuación se deriva fácilmente desde la última, e implica que $\sum_{i=1}^n a_i \cdot k_u(a_i, \phi_u) = 0$. Entonces es claro que no existe solución única al sistema, de hecho ϕ_ε y ϕ_u no son estimables.

Ahora bien, para la segunda expresión y bajo los mismos supuestos, los estimadores máximo verosímiles deben satisfacer que:

$$\begin{aligned} x_i &= \left(\frac{\beta_1^2}{e_i^2} + \frac{1}{a_i^2} \right)^{-1} \left(\frac{\beta_1^2}{e_i^2} \cdot \left(\frac{y_i - \beta_0}{\beta_1} \right) + \frac{1}{a_i^2} \cdot z_i \right); & \frac{\phi_\varepsilon}{n} \cdot \sum_{i=1}^n \frac{e_i^2}{\omega_{\varepsilon i}} &= 1; & \frac{\phi_u}{n} \cdot \sum_{i=1}^n \frac{a_i^2}{\omega_{u i}} &= 1; \\ \sum_{i=1}^n \frac{e_i}{\omega_{\varepsilon i}} &= 0; & \sum_{i=1}^n \frac{x_i \cdot e_i}{\omega_{\varepsilon i}} &= 0; & \omega_\varepsilon &= \phi_\varepsilon \cdot \sum_{i=1}^n \frac{e_i^2}{n}, & y & \quad \omega_u = \phi_u \cdot \sum_{i=1}^n \frac{a_i^2}{n}; & \forall i = 1, \dots, n. \end{aligned}$$

Ahora bien, las ecuaciones segunda y tercera se deducen de las otras, de modo que el sistema no tiene solución única. Los estimadores máximo verosímiles de β_0, β_1 y x son estimables, no así los estimadores de los parámetros ϕ_ε y ϕ_u .

Por otra parte, aún para una distribución G particular bien escogida, gama inversa por ejemplo, la estimación máximo verosímil requiere resolver un complicado sistema de ecuaciones. En el caso mencionado se requeriría hallar los valores de los parámetros que satisfagan las siguientes ecuaciones:

$$\frac{2r+1}{n} \sum_{i=1}^n \left(1 + \frac{2\lambda}{\phi_\varepsilon \cdot e_i^2}\right)^{-1} = 1; \quad \frac{2r+1}{n} \sum_{i=1}^n \left(1 + \frac{2\lambda}{\phi_u \cdot a_i^2}\right)^{-1} = 1; \quad \sum_{i=1}^n \frac{1}{e_i} \left(1 + \frac{2\lambda}{\phi_\varepsilon \cdot e_i^2}\right)^{-1} = 0;$$

$$\sum_{i=1}^n \frac{x_i}{e_i} \left(1 + \frac{2\lambda}{\phi_\varepsilon \cdot e_i^2}\right)^{-1} = 0, \text{ y } x_i = \frac{\frac{1}{e_i^2} \cdot \left(1 + \frac{2\lambda}{\phi_\varepsilon \cdot e_i^2}\right)^{-1} \cdot \beta_1^2 \cdot \left(\frac{y_i - \beta_0}{\beta_1}\right) + \frac{1}{a_i^2} \cdot \left(1 + \frac{2\lambda}{\phi_u \cdot a_i^2}\right)^{-1} \cdot z_i}{\frac{1}{e_i^2} \cdot \left(1 + \frac{2\lambda}{\phi_\varepsilon \cdot e_i^2}\right)^{-1} \cdot \beta_1^2 + \frac{1}{a_i^2} \cdot \left(1 + \frac{2\lambda}{\phi_u \cdot a_i^2}\right)^{-1}}.$$

Desde un punto de vista bayesiano, la verosimilitud asociada al modelo de interés bajo la especificación (5.3), aún considerando G conocida (en caso contrario las expresiones son mucho más complicadas), es de la forma siguiente:

$$\pi(\beta, \mathbf{x}, \phi_\varepsilon, \phi_u | \mathbf{y}, \mathbf{z}, G_\varepsilon, G_u) \propto L(\beta, \mathbf{x}, \phi_\varepsilon, \phi_u | \mathbf{y}, \mathbf{z}, G_\varepsilon, G_u) \cdot \pi(\beta) \cdot \pi(\mathbf{x}) \cdot \pi(\phi_\varepsilon) \cdot \pi(\phi_u),$$

es decir, la distribución a posteriori para los parámetros es corresponde a la siguiente expresión:

$$\pi(\beta, \mathbf{z}, \phi_\varepsilon, \phi_u | \mathbf{y}, G_\varepsilon, G_u) \propto (\phi_\varepsilon^{n/2+r-1} \cdot \phi_u^{n/2+r-1}) \cdot \left[\prod_{i=1}^n \int_0^\infty \omega_{\varepsilon i}^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi_\varepsilon \cdot (y_i - \beta_0 - \beta_1 \cdot x_i)^2 / \omega_{\varepsilon i}) dG_\varepsilon(\omega_{\varepsilon i}) \right]$$

$$\cdot \left\{ \int_0^\infty \omega_{u i}^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi_u \cdot (z_i - x_i)^2 / \omega_{u i}) dG_u(\omega_{u i}) \right\}$$

$$\cdot \exp \left\{ -\frac{1}{2} \left[(\beta_0 - b_0) / v_0 + (\beta_1 - b_1) / v_1 + \frac{1}{\sigma_X^2} \sum_{i=1}^n (x_i - \mu_X)^2 \right] - \lambda_0 \cdot \phi_\varepsilon - \lambda_1 \cdot \phi_u \right\},$$

donde la constante de proporcionalidad es prácticamente imposible de evaluar explícitamente.

Marginalizando, la distribución a posteriori de los x_i tienen la siguiente forma:

$$\pi(\mathbf{x} | \mathbf{y}, \mathbf{z}, G_\varepsilon, G_u) \propto \exp \left[-\frac{1}{2 \cdot \sigma_X^2} \sum_{i=1}^n (x_i - \mu_X)^2 \right]$$

$$\cdot \int_{\mathbb{R}^{2 \times n}} \int_{\mathbb{R}^{2 \times n}} \left[\phi_\varepsilon^{n/2+r-1} \cdot \phi_u^{n/2+r-1} \cdot \exp \left\{ -\frac{1}{2} \left[(\beta_0 - b_0) / v_0 + (\beta_1 - b_1) / v_1 \right] - \lambda_0 \cdot \phi_\varepsilon - \lambda_1 \cdot \phi_u \right\} \right.$$

$$\cdot \prod_{i=1}^n \int_0^\infty \omega_{\varepsilon i}^{-\frac{1}{2}} \exp \left(-\frac{\phi_\varepsilon}{2 \omega_{\varepsilon i}} (y_i - \beta_0 - \beta_1 \cdot x_i)^2 \right) dG_\varepsilon(\omega_{\varepsilon i}) \left. \right]$$

$$\cdot \prod_{i=1}^n \int_0^\infty \omega_{u i}^{-\frac{1}{2}} \exp \left(-\frac{\phi_u}{2 \omega_{u i}} (z_i - x_i)^2 \right) dG_u(\omega_{u i}) \left. \right] d\phi_\varepsilon d\phi_u d\beta,$$

expresión que depende del proceso G y de su distribución a priori.

Incluso si G_ε y G_u fueran distribuciones gama inversa de parámetros (r_0, λ_0) y (r_1, λ_1) , esta distribución resultaría igualmente complicada de deducir, ya que la expresión para las productorias de integrales respecto de ω_ε y de ω_u es proporcional a:

$$\prod_{i=0}^n [1 + \phi_\varepsilon \cdot (y_i - \beta_0 - \beta_1 \cdot x_i)^2 / 2 \lambda_0]^{-1/2(2r_0+1)} \cdot [1 + \phi_u \cdot (z_i - x_i)^2 / 2 \lambda_1]^{-1/2(2r_1+1)},$$

de donde la expresión resultante para el estimador de Bayes (la media de dicha distribución) no es obtenible analíticamente.

Por otro lado, para la especificación (5.4) la posteriori presenta una expresión semejante a la anterior y que condice a los mismos hallazgos.

Debido a lo expuesto anteriormente hemos optado por la alternativa de utilizar MCMC para implementar el muestreo de Gibbs, ya que, como se verá a continuación, las distribuciones condicionales completas requeridas para dicha simulación se pueden obtener analíticamente.

Al igual que en capítulos anteriores, consideramos dos situaciones para la elección de las distribuciones a priori de los parámetros de dispersión G_ε y G_u . La primera se incluye a modo de ejemplificación y la segunda corresponde al problema de interés:

- i) Las distribuciones a priori para G_ε y G_u tienen soporte finito.
- ii) G_ε y G_u siguen procesos Dirichlet de parámetros $(c_\varepsilon G_{0\varepsilon})$ y $(c_u G_{0u})$, respectivamente. Un caso particular para esta segunda alternativa consiste en suponer que las funciones generadoras de densidad de las distribuciones elípticas en (5.3) h_ε o h_u son conocidas, en cuyo caso G_ε o G_u resultan ser degeneradas en $G_{0\varepsilon}$ o G_{0u} , respectivamente.

Para ambas situaciones las distribuciones a posteriori son difíciles o imposibles de obtener de modo analítico, lo cual justifica la implementación de muestreo de Gibbs y, por tanto, motiva la obtención de distribuciones condicionales completas.

Proposición 5.1

Bajo las especificaciones dadas por el modelo jerárquico (5.3) y considerando además que

$$p(\omega_{\varepsilon i} | G_\varepsilon) = \sum_{j=1}^{k_\varepsilon+1} p_{\varepsilon j} \cdot \delta_{\tau_{\varepsilon j}}(\omega_{\varepsilon i}), \quad p(\omega_{ui} | G_u) = \sum_{j=1}^{k_u+1} p_{uj} \cdot \delta_{a_{uj}}(\omega_{ui}), \quad (p_{\varepsilon 1}, \dots, p_{\varepsilon k_\varepsilon}) \sim D(\alpha_{\varepsilon 1}, \dots, \alpha_{\varepsilon k_\varepsilon+1}) \text{ y}$$

$(p_{u 1}, \dots, p_{u k}) \sim D(\alpha_1, \dots, \alpha_{u k_u+1})$, *se obtienen las distribuciones condicionales completas siguientes, donde*

$\omega_\varepsilon, \omega_u, \mathbf{x}, \mathbf{y}$ *y* \mathbf{z} *corresponden a los vectores de componentes* $\omega_{\varepsilon i}, \omega_{ui}, x_i, y_i$ *y* z_i , *respectivamente:*

$$(5.5) \quad \omega_{\varepsilon i} | \omega_{\varepsilon -i}, \omega_u, \phi_\varepsilon, \phi_u, z_i, x_i, \beta, y_i \sim \sum_{j=1}^{k_\varepsilon+1} q_{\varepsilon i j} \cdot \delta_{\tau_{\varepsilon j}}(\omega_{\varepsilon i}), \text{ donde } \tau_{\varepsilon j} \text{ son los posibles valores de } \omega_{\varepsilon i} \text{ y}$$

$$\text{donde } q_{\varepsilon i j} = \frac{\phi \left(\sqrt{\frac{\phi_\varepsilon}{\tau_{\varepsilon j}}} (y_i - \beta_0 - \beta_1 \cdot x_i) \right) \cdot \frac{p_{\varepsilon j}}{\sqrt{\tau_{\varepsilon j}}}}{\sum_{m=1}^{k_\varepsilon+1} \phi \left(\sqrt{\frac{\phi_\varepsilon}{\tau_{\varepsilon m}}} (y_i - \beta_0 - \beta_1 \cdot x_i) \right) \cdot \frac{p_{\varepsilon m}}{\sqrt{\tau_{\varepsilon m}}}}$$

$$(5.6) \quad \omega_{u_i} | \omega_{u_{-i}}, \omega_{\varepsilon}, \phi_{\varepsilon}, \phi_u, z_i, x_i, \beta, y_i \sim \sum_{j=1}^{k_u+1} q_{u_{ij}} \cdot \delta_{\tau_{uj}}(\omega_{u_i}), \text{ donde } \tau_{uj} \text{ son los posibles valores de } \omega_{u_i} \text{ y}$$

$$\text{donde } q_{u_{ij}} = \frac{\phi\left(\sqrt{\frac{\phi_u}{\tau_{uj}}}(y_i - x_i)\right) \cdot \frac{p_{uj}}{\sqrt{\tau_{uj}}}}{\sum_{m=1}^{k_u+1} \phi\left(\sqrt{\frac{\phi_u}{\tau_{um}}}(y_i - x_i)\right) \cdot \frac{p_{um}}{\sqrt{\tau_{um}}}}.$$

$$(5.7) \quad \phi_{\varepsilon} | \omega_{\varepsilon}, \omega_u, G_{\varepsilon}, G_u, \phi_u, \beta, \mathbf{x}, \mathbf{y}, \mathbf{z} \sim \text{Gamma}\left(r_0 + \frac{n}{2}, \lambda_0 + \sum_{i=1}^n \frac{(y_i - \beta_0 - \beta_1 \cdot x_i)^2}{2 \cdot \omega_{\varepsilon i}}\right)$$

$$(5.8) \quad \phi_u | \omega_{\varepsilon}, \omega_u, G_{\varepsilon}, G_u, \phi_{\varepsilon}, \beta, \mathbf{x}, \mathbf{y}, \mathbf{z} \sim \text{Gamma}\left(r_1 + \frac{n}{2}, \lambda_1 + \sum_{i=1}^n \frac{(z_i - x_i)^2}{2 \cdot \omega_{ui}}\right)$$

$$(5.9) \quad \beta_0 | \omega_{\varepsilon}, \omega_u, G_{\varepsilon}, G_u, \phi_{\varepsilon}, \phi_u, \beta_1, \mathbf{x}, \mathbf{y}, \mathbf{z} \sim N(m_0, S_0^2)$$

$$\text{donde } m_0 = S_0^2 \cdot \left[\phi_{\varepsilon} \cdot \sum_{i=1}^n \frac{(y_i - \beta_1 x_i)}{\omega_{\varepsilon i}} + \frac{b_0}{V_0} \right] \text{ y } S_0^2 = \left[\phi_{\varepsilon} \cdot \sum_{i=1}^n \frac{1}{\omega_{\varepsilon i}} + \frac{1}{V_0} \right]^{-1}.$$

$$(5.10) \quad \beta_1 | \omega_{\varepsilon}, \omega_u, G_{\varepsilon}, G_u, \phi_{\varepsilon}, \phi_u, \beta_0, \mathbf{x}, \mathbf{y}, \mathbf{z} \sim N(m_1, S_1^2)$$

$$\text{donde } m_1 = S_1^2 \cdot \left[\phi_{\varepsilon} \cdot \sum_{i=1}^n \frac{x_i \cdot (y_i - \beta_0)}{\omega_{\varepsilon i}} + \frac{b_1}{V_1} \right] \text{ y } S_1^2 = \left[\phi_{\varepsilon} \cdot \sum_{i=1}^n \frac{x_i^2}{\omega_{\varepsilon i}} + \frac{1}{V_1} \right]^{-1}.$$

$$(5.11) \quad x_i | \omega_{\varepsilon}, \omega_u, G_{\varepsilon}, G_u, \phi_{\varepsilon}, \phi_u, \beta, y_i, z_i \stackrel{\text{c.i.i.d.}}{\sim} N(m_x, S_x^2),$$

$$\text{con } m_{xi} = S_x^2 \cdot \left(\phi_{\varepsilon} \cdot \beta_1 \cdot \frac{y_i - \beta_0}{\omega_{\varepsilon i}} + \frac{\phi_u \cdot z_i}{\omega_{ui}} + \frac{\mu_x}{\sigma_x^2} \right) \text{ y } S_{xi}^2 = \left(\frac{\phi_{\varepsilon} \cdot \beta_1^2}{\omega_{\varepsilon i}} + \frac{\phi_u}{\omega_{ui}} + \frac{1}{\sigma_x^2} \right)^{-1}$$

Para todos los casos anteriores se ha supuesto que los valores $\tau_{\varepsilon j}$ y τ_{uj} son fijos y conocidos.

Demostración

A partir del teorema de Escobar y West (1998), presentado en la sección 1.1 (1.12) y (1.13), se deducen las expresiones (5.5) y (5.6). Por otro lado, las expresiones (5.7) y (5.8) surgen del *Lema 2.4* con σ^2 igual a ϕ_{ε} o ϕ_u . El *Ejemplo 2.2* permite la construcción de las expresiones (5.9) y (5.10), y finalmente, la expresión (5.11) es consecuencia del *Lema 2.1*. \square

La implementación se puede realizar confeccionando un programa para muestrear desde las condicionales o utilizando Winbugs.

Observación 5.1

El caso abordado en la *proposición 5.1.* considera una mezcla discreta en parámetro escala de normales. Ahora bien, si $p_{ei} = p_{ui} = 1$ y $p_{ei} = p_{ui} = 0 \forall j \neq i$, es decir, si ω_{ei} y ω_{ui} tienen distribuciones degeneradas en un punto, las distribuciones condicionales completas de los parámetros β , x , ϕ_ε y ϕ_U se mantienen intactas, pero p_ε , p_U , ω_{ei} y ω_{ui} pasan a ser constantes, de modo que $h_\varepsilon(u) = h_U(u) = (2\pi a)^{-1/2} e^{-u/2a}$: Con esto ya no se produce una mezcla de normales, sino que se recupera el modelo de regresión lineal con errores de medición MEM normal.

El principal resultado de esta sección es presentado a continuación, y corresponde a la solución bayesiana semiparamétrica del problema de regresión lineal con errores de medición MEM considerando términos de error elípticos representables independientes.

Proposición 5.2

Bajo las especificaciones del modelo (5.1) y (5.3) con:

$$\begin{aligned}
 \omega_{ei} | G_\varepsilon &\sim \text{i.i.d. } G_\varepsilon \\
 \omega_{ui} | G_u &\sim \text{i.i.d. } G_u \\
 (5.12) \quad G_\varepsilon &\sim D(c_\varepsilon G_{\varepsilon 0}) \\
 G_u &\sim D(c_u G_{u 0}) \\
 G_{\varepsilon 0} &\sim IG(r_\varepsilon, \lambda_\varepsilon) \\
 G_{u 0} &\sim IG(r_u, \lambda_u),
 \end{aligned}$$

La distribución a posteriori de x no permite la obtención analítica del estimador de Bayes, lo cual motiva la obtención de las distribuciones condicionales completas para la implementación del muestreo de gibbs.

Las distribuciones condicionales completas resultan idénticas a las del caso discreto, excepto para ω_{ei} y ω_{ui} :

$$(5.13) \quad \omega_{ei} / \{\omega_{\varepsilon j}; j \neq i\}, G_\varepsilon, G_u, \phi_\varepsilon, \phi_u, \beta, x_i, y_i, z_i \sim q_{0i}^{(\varepsilon)} \cdot f_{G_b^{(\varepsilon)}}(\omega_{\varepsilon i}) + \sum_{k \neq i} q_{ki}^{(\varepsilon)} \cdot \delta_{\omega_{\varepsilon k}}(\omega_{\varepsilon i})$$

donde $G_b^{(\varepsilon)} = \text{Gamma Inv}(r_\varepsilon^* = r_\varepsilon + 1/2, \lambda_\varepsilon^* = \lambda_\varepsilon + (y_i - \beta_0 - \beta_1 \cdot x_i)^2 \cdot \phi_\varepsilon / 2)$,

$$q_{0i}^{(\varepsilon)} = f_{T(\mu_{\varepsilon i}^*, \sigma_{\varepsilon i}^{2*}, \gamma_{\varepsilon i}^*)}(y_i) \cdot f_{N(\mu = x_i, \sigma^2 = \omega_{ui} / \phi_u)}(z_i / \phi_u, \omega_{ui})$$

$$q_{ki}^{(\varepsilon)} = f_{N(\mu_{\varepsilon i}^*, \sigma_{\varepsilon i}^2 = \omega_{\varepsilon k} / \phi_\varepsilon)}(y_i / \phi_\varepsilon, \omega_{\varepsilon i}),$$

para $\mu_{\varepsilon i}^* = E(y_i / \beta, x_i) = \beta_0 + \beta_1 \cdot x_i$; $\sigma_{\varepsilon i}^{2*} = 2 \lambda_\varepsilon / \phi_\varepsilon$; $\gamma_{\varepsilon i}^* = 2 \cdot r_\varepsilon$.

$$(5.14) \quad \omega_{ui} / \{\omega_{uj}; j \neq i\}, G_{\varepsilon}, G_u, \phi_{\varepsilon}, \phi_u, \beta, x_i, y_i, z_i \sim q_{0i}^{(u)} \cdot f_{G_b^{(u)}}(\omega_{ui}) + \sum_{k \neq i} q_{ki}^{(u)} \cdot \delta_{\omega_{uk}}(\omega_{ui})$$

donde $G_b^{(u)}$ es la distribución $IG(r_u^* = r_u + 1/2, \lambda_u^* = \lambda_u + (z_i - x_i)^2 \cdot \phi_u / 2)$,

$$q_{0i}^{(u)} = f_{T(\mu_{ui}^*, \sigma_{ui}^{2*}, \gamma_{ui}^* | z_i)} \cdot f_{N(\mu = \beta_0 + \beta_1 \cdot x_i, \sigma^2 = \omega_{\varepsilon i} / \phi_{\varepsilon})}(y_i / \phi_u, \omega_{ui})$$

$$q_{ki}^{(u)} = f_{N(\mu_{ui}^*, \sigma_{ui}^{2*} = \omega_{uk} / \phi_u)}(z_i / \phi_u, \omega_{ui}),$$

$$\text{para } \mu_{ui}^* = E(z_i / x_i) = x_i, \sigma_{ui}^{2*} = 2 \lambda_u / \phi_u, \gamma_{ui}^* = 2 \cdot r_u.$$

Demostración:

El teorema de Escobar y West (1998), véase (1.12) y (1.13), provee los resultados para (5.13) y (5.14). \square

Ejemplo 5.1

Bajo las especificaciones (5.3) y (5.4), con $\omega_{\varepsilon i} \sim IG(r_{\varepsilon}, \lambda_{\varepsilon})$ y con $\omega_{ui} \sim IG(r_u, \lambda_u)$, se obtienen resultados análogos a los del caso anterior, excepto que las distribuciones condicionales completas para $\omega_{\varepsilon i}$ y ω_{ui} corresponden a $G_b^{(\varepsilon)}$ y $G_b^{(u)}$, respectivamente.

5.2 SOLUCIÓN BAYESIANA SEMIPARAMÉTRICA BAJO EL MODELO DE REGRESIÓN ELÍPTICO REPRESENTABLE DEPENDIENTE CON ERRORES DE MEDICIÓN MEM

En esta sección se aborda el problema de regresión con errores en las variables o con errores de medición MEM, pero en este caso considerando que los términos de error siguen una distribución elíptica dependiente y con función generadora de densidad h representable. La diferencia respecto del modelo con términos de error independientes radica en el hecho de que para este caso existe un parámetro de precisión común, es decir, $\omega_{\varepsilon i} = \omega_{\varepsilon}$ y $\omega_{ui} = \omega_u$, $\forall i \in \{1, \dots, n\}$; esto motiva que los errores sean condicionalmente i.i.d., es decir el modelo abordado en esta sección es el siguiente:

$$y_i = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i, i = 1, \dots, n,$$

$$z_i = x_i + u_i, i = 1, \dots, n,$$

donde:

$$(5.15) \quad \boldsymbol{\varepsilon} | \omega_{\varepsilon}, G_{\varepsilon}, \phi_{\varepsilon}, \boldsymbol{\beta}, \mathbf{x} \sim El_n(0, \phi_{\varepsilon}^{-1} \mathbf{I}_n, h_{\varepsilon}), \boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n), \text{ y } h_{\varepsilon}(\mathbf{u}) = \int (2\pi\omega)^{-n/2} e^{-\mathbf{u}^2/2\omega} dG_{\varepsilon}(\omega),$$

$$\mathbf{u} | \omega_u, G_u, \phi_u, \mathbf{x} \sim El_n(0, \phi_u^{-1} \mathbf{I}_n, h_u), \mathbf{u} = (u_1, \dots, u_n), \text{ y } h_{eu}(\mathbf{u}) = \int (2\pi\omega)^{-n/2} e^{-\mathbf{u}^2/2\omega} dG_u(\omega),$$

con $\boldsymbol{\varepsilon}$ y \mathbf{u} son independientes, h_{ε} y h_u no – negativas y G_{ε} y G_u funciones de distribución acumuladas tales que $G_{\varepsilon}(0) = G_u(0) = 0$.

Por otra parte, el modelo jerárquico de interés corresponde a:

$$\begin{aligned}
 \varepsilon \mid \omega_{\varepsilon i}, G_{\varepsilon}, \phi_{\varepsilon}, \beta, x_i &\sim El_n(0, \phi_{\varepsilon}^{-1} I_n, h_{\varepsilon}), i=1, \dots, n, \\
 \mathbf{u} \mid \omega_{u i}, G_u, \phi_u, x_i &\sim El_n(0, \phi_u^{-1} I_n, h_u), i=1, \dots, n, \\
 (5.16) \quad x_i \mid \phi_{\varepsilon}, \phi_u, \beta &\sim \text{i.i.d. } N_1(\mu_x, \sigma_x^2), i = 1, \dots, n, \\
 \phi_{\varepsilon} &\sim \text{Gamma}(r_0, \lambda_0), \\
 \phi_u &\sim \text{Gamma}(r_1, \lambda_1), \\
 \beta_0 &\sim N(\mathbf{b}_0, V_0) \\
 \text{y} \quad \beta_1 &\sim N(\mathbf{b}_1, V_1).
 \end{aligned}$$

Un caso interesante corresponde a aquel en el cual $\phi_{\varepsilon} = \phi_u$.

En forma análoga a lo expuesto en la sección previa, el uso de la *Proposición 2.1. a.*, permite abordar el problema considerando errores normales. Esto implica que el modelo jerárquico (5.16) es equivalente al modelo en el cual las primeras dos condiciones se han extendido a las cuatro condiciones siguientes:

$$\begin{aligned}
 (5.17) \quad y_i \mid \omega_{\varepsilon}, \omega_u, \phi_{\varepsilon}, \phi_u, z_i, x_i, \beta &\sim \text{con. ind. } N_1(\beta_0 + \beta_1 x_i, \omega_{\varepsilon} \phi_{\varepsilon}^{-1}), \\
 z_i \mid \omega_{\varepsilon}, \omega_u, \phi_{\varepsilon}, \phi_u, y_i, x_i &\sim \text{cond. ind. } N_1(x_i, \omega_u \phi_u^{-1}), \\
 \omega_{\varepsilon} \mid G_{\varepsilon} &\sim G_{\varepsilon} \text{ con } G_{\varepsilon} \text{ independiente de } \phi_{\varepsilon}, \mathbf{x} \text{ y de } \beta. \\
 \text{y} \quad \omega_u \mid G_u &\sim G_u \text{ con } G_u \text{ independiente de } \phi_u \text{ y de } \mathbf{x},
 \end{aligned}$$

donde, tanto para G_{ε} , como para G_u se escoge una distribución a priori discreta o bien se le asigna a priori un proceso *Dirichlet*.

Respecto de las distribuciones condicionales que requiere la implementación computacional, las únicas diferencias que se experimentan en este caso, respecto del caso con términos de error elípticos independientes, se relacionan con las variables subyacentes ω_{ε} y ω_u . En este sentido, para las distribuciones condicionales de las demás componentes del modelo los parámetros $\omega_{\varepsilon i}$ y $\omega_{u i}$ se reemplazan respectivamente por ω_{ε} y ω_u , y la distribución de estas últimas variables se presenta a continuación:

$$\begin{aligned}
 (5.18) \quad \omega_{\varepsilon} \mid \omega_u, G_{\varepsilon}, G_u, \phi_{\varepsilon}, \phi_u, \beta, \mathbf{x}, \mathbf{y}, \mathbf{z} &\sim I G \left(r_{\varepsilon}^* = r_{\varepsilon} + \frac{n}{2}, \lambda_{\varepsilon}^* = \lambda_{\varepsilon} + \phi_{\varepsilon} \cdot \sum_{i=1}^n \frac{(y_i - \beta_0 - \beta_1 x_i)^2}{2} \right) \\
 \omega_u \mid \omega_{\varepsilon}, G_{\varepsilon}, G_u, \phi_{\varepsilon}, \phi_u, \beta, \mathbf{x}, \mathbf{y}, \mathbf{z} &\sim I G \left(r_u^* = r_u + \frac{n}{2}, \lambda_u^* = \lambda_u + \phi_u \cdot \sum_{i=1}^n \frac{(z_i - x_i)^2}{2} \right)
 \end{aligned}$$

Al igual que en la sección 5.1, y debido a las dificultades analíticas inherentes al problema, la forma plausible de obtener el estimador de Bayes de los parámetros de interés implica la obtención de las distribuciones condicionales conjuntas que permitan implementar el muestreo de Gibbs.

En forma análoga a los desarrollos de la sección 5.1., se obtiene para el caso dependiente la verosimilitud del modelo (5.15), que corresponde a:

$$L(\boldsymbol{\beta}, \mathbf{z}, \phi_\varepsilon, \phi_u | \mathbf{y}, G_\varepsilon, G_u) = (\phi_\varepsilon \cdot \phi_u)^{n/2} (2\pi)^{-n} \prod_{i=1}^n \left\{ \int_0^\infty \omega_\varepsilon^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi_\varepsilon \cdot (y_i - \beta_0 - \beta_1 \cdot x_i)^2 / \omega_\varepsilon) dG_\varepsilon(\omega_\varepsilon) \right\} \\ \cdot \prod_{i=1}^n \left\{ \int_0^\infty \omega_u^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi_u \cdot (z_i - x_i)^2 / \omega_u) dG_u(\omega_u) \right\}.$$

Es claro que, como las variables de integración (en este caso ω_ε y ω_u) son mudas, la verosimilitud de los casos independiente y dependiente coincide.

Ahora bien, para el modelo basado en (5.17) se obtiene una función de verosimilitud aumentada que coincide con la verosimilitud de un modelo de regresión normal heteroscedástico:

$$L(\boldsymbol{\beta}, \mathbf{z}, \phi_\varepsilon, \phi_u, \omega_\varepsilon, \omega_u | \mathbf{y}, G_\varepsilon, G_u) = (\phi_\varepsilon \cdot \phi_u)^{n/2} (2\pi)^{-n} \left[\prod_{i=1}^n (\omega_\varepsilon \cdot \omega_u)^{-\frac{1}{2}} \right] \\ \cdot \exp[-\frac{1}{2} (\phi_\varepsilon \cdot \omega_\varepsilon^{-1} \cdot \sum_{i=1}^n (y_i - \beta_0 - \beta_1 \cdot x_i)^2 + \phi_u \cdot \omega_u^{-1} \cdot \sum_{i=1}^n (z_i - x_i)^2)].$$

En base a la función de verosimilitud anterior, sin contar con información acerca de G_ε ni de G_u la estimación máximo verosímil resulta infactible desde el punto de vista clásico; por otra parte, para la primera especificación e incluso para algunas G_ε y G_u particulares, la expresión requiere que las derivadas de la verosimilitud o de la log verosimilitud respecto de $\boldsymbol{\beta}$, \mathbf{z} , ϕ_ε y ϕ_u sean nulas, problema que resulta muy complicado de resolver. Por otra parte, las variables instrumentales ω_ε y ω_u también requieren ser estimadas.

Ahora bien, si asumimos G_ε y G_u dadas y k_ε y k_u definidas en forma análoga a la de la sección 5.1. (excepto por el hecho de que $\omega_{\varepsilon i}$ y $\omega_{u i}$ se reemplazan por ω_ε y ω_u , respectivamente), para la primera especificación del modelo las ecuaciones a resolver no se ven alteradas.

Por el contrario, para la segunda especificación (verosimilitud aumentada) se obtiene el siguiente sistema de ecuaciones:

$$x_i = \left(\frac{n \cdot \beta_1^2}{\sum_{i=1}^n e_i^2} + \frac{n}{\sum_{i=1}^n a_i^2} \right)^{-1} \left(\frac{n \cdot \beta_1^2}{\sum_{i=1}^n e_i^2} \cdot \left(\frac{y_i - \beta_0}{\beta_1} \right) + \frac{n}{\sum_{i=1}^n a_i^2} \cdot z_i \right); \quad \frac{\phi_\varepsilon \cdot \sum_{i=1}^n e_i^2}{n \cdot \omega_\varepsilon} = 1; \quad \frac{\phi_u \cdot \sum_{i=1}^n a_i^2}{n \cdot \omega_u} = 1; \quad \sum_{i=1}^n e_i = 0; \quad \sum_{i=1}^n x_i \cdot e_i = 0;$$

$\forall i = 1, \dots, n.$

Es claro que el sistema anterior no tiene solución única, puesto que hay menos ecuaciones que parámetros, y como consecuencia, ϕ_ε , ϕ_u , ω_ε y ω_u no son estimables.

Bajo un enfoque bayesiano, la función de verosimilitud asociada al modelo de interés bajo la especificación (5.15) coincide con la hallada en la sección precedente, excepto por el hecho de que se reemplaza $\omega_{\varepsilon i}$ y $\omega_{u i}$ por ω_ε y ω_u , respectivamente. Es por esta razón, y debido a que $\omega_{\varepsilon i}$ y $\omega_{u i}$ son variables de integración, que la distribución a posteriori de los parámetros no varía.

Proposición 5.3

Bajo el modelo jerárquico (5.16) y considerando las especificaciones $p(\omega_\varepsilon | G_\varepsilon) = \sum_{j=1}^{k_\varepsilon+1} p_{\varepsilon j} \cdot \delta_{\tau_{\varepsilon j}}(\omega_\varepsilon)$,

$$p(\omega_u | G_u) = \sum_{j=1}^{k_u+1} p_{u j} \cdot \delta_{a_{u j}}(\omega_u), p(p_{\varepsilon 1}, \dots, p_{\varepsilon k_\varepsilon}) \sim D(\alpha_{\varepsilon 1}, \dots, \alpha_{\varepsilon k_\varepsilon+1}) \text{ y } (p_{u 1}, \dots, p_{u k_u}) \sim D(\alpha_{u 1}, \dots, \alpha_{u k_u+1}),$$

se obtiene que, al igual que en el caso independiente, la distribución a posteriori de los parámetros es intratable analíticamente.

Asumiendo ahora que G_ε y G_u son distribuciones gamma inversa de parámetros (r_0, λ_0) y (r_1, λ_1) , la verosimilitud coincide con la del caso independiente.

En este caso, entonces, es necesario optar por la alternativa de utilizar MCMC para implementar el muestreo de Gibbs, ya que las distribuciones condicionales completas requeridas se pueden obtener analíticamente, como se puede ver a continuación:

$$(5.19) \quad \omega_\varepsilon | \omega_u, \phi_\varepsilon, \phi_u, \mathbf{z}, \mathbf{x}, \boldsymbol{\beta}, \mathbf{y} \sim g_\varepsilon(\omega_\varepsilon) \propto \sum_{j=1}^{k_\varepsilon+1} \exp \left\{ - \sum_{i=1}^{k_\varepsilon+1} \frac{(y_i - \beta_0 - \beta_1 \cdot x_i)^2}{2\omega_\varepsilon} \right\} \cdot \frac{\phi_\varepsilon \cdot p_{\varepsilon j} \cdot \delta_{\tau_{\varepsilon j}}(\omega_\varepsilon)}{\sqrt{\omega_\varepsilon^n}},$$

$$(5.20) \quad \omega_u | \omega_\varepsilon, \phi_\varepsilon, \phi_u, \mathbf{z}, \mathbf{x}, \boldsymbol{\beta}, \mathbf{y} \sim g_u(\omega_u) \propto \sum_{j=1}^{k_u+1} \exp \left\{ - \sum_{i=1}^{k_u+1} \frac{(z_i - x_i)^2}{2\omega_u} \right\} \cdot \frac{\phi_u \cdot p_{u j} \cdot \delta_{\tau_{u j}}(\omega_u)}{\sqrt{\omega_u^n}},$$

$$(5.21) \quad \phi_\varepsilon | \omega_\varepsilon, \omega_u, G_\varepsilon, G_u, \phi_u, \boldsymbol{\beta}, \mathbf{x}, \mathbf{y}, \mathbf{z} \sim \text{Gamma}(r_0 + \frac{n}{2}, \lambda_0 + \sum_{i=1}^n \frac{(y_i - \beta_0 - \beta_1 \cdot x_i)^2}{2 \cdot \omega_\varepsilon}),$$

$$(5.22) \quad \phi_u | \omega_\varepsilon, \omega_u, G_\varepsilon, G_u, \phi_\varepsilon, \boldsymbol{\beta}, \mathbf{x}, \mathbf{y}, \mathbf{z} \sim \text{Gamma}(r_1 + \frac{n}{2}, \lambda_1 + \sum_{i=1}^n \frac{(z_i - x_i)^2}{2 \cdot \omega_u}),$$

$$(5.23) \quad \beta_0 | \omega_\varepsilon, \omega_u, G_\varepsilon, G_u, \phi_\varepsilon, \phi_u, \beta_1, \mathbf{x}, \mathbf{y}, \mathbf{z} \sim N(m_0, S_0^2),$$

$$\text{donde } m_0 = S_0^2 \cdot \left[\phi_\varepsilon \cdot \sum_{i=1}^n \frac{(y_i - \beta_1 x_i)}{\omega_\varepsilon} + \frac{b_0}{V_0} \right] \text{ y } S_0^2 = \left[\phi_\varepsilon \cdot \sum_{i=1}^n \frac{1}{\omega_\varepsilon} + \frac{1}{V_0} \right]^{-1},$$

$$(5.24) \quad \beta_1 | \omega_\varepsilon, \omega_u, G_\varepsilon, G_u, \phi_\varepsilon, \phi_u, \beta_0, \mathbf{x}, \mathbf{y}, \mathbf{z} \sim N(m_1, S_1^2),$$

$$\text{donde } m_1 = S_1^2 \cdot \left[\phi_\varepsilon \cdot \sum_{i=1}^n \frac{x_i \cdot (y_i - \beta_0)}{\omega_\varepsilon} + \frac{b_1}{V_1} \right] \text{ y } S_1^2 = \left[\phi_\varepsilon \cdot \sum_{i=1}^n \frac{x_i^2}{\omega_\varepsilon} + \frac{1}{V_1} \right]^{-1},$$

$$(5.25) \quad x_i | \omega_\varepsilon, \omega_u, G_\varepsilon, G_u, \phi_\varepsilon, \phi_u, \boldsymbol{\beta}, y_i, z_i \sim^{c.ind.} N(m_{xi}, S_x^2),$$

$$\text{con } m_{xi} = S_x^2 \left(\phi_\varepsilon \cdot \beta_1 \cdot \frac{y_i - \beta_0}{\omega_\varepsilon} + \frac{\phi_u \cdot z_i}{\omega_u} + \frac{\mu_x}{\sigma_x^2} \right) \text{ y } S_x^2 = \left(\frac{\phi_\varepsilon \cdot \beta_1^2}{\omega_\varepsilon} + \frac{\phi_u}{\omega_u} + \frac{1}{\sigma_x^2} \right)^{-1}.$$

Demostración:

La función de verosimilitud corresponde a:

$$\begin{aligned} \pi(\boldsymbol{\beta}, \mathbf{z}, \phi_\varepsilon, \phi_u | \mathbf{y}, G_\varepsilon, G_u) \propto & (\phi_\varepsilon^{n/2+r_0-1} \cdot \phi_u^{n/2+r_1-1} \cdot \{\omega_\varepsilon \cdot \omega_u\}^{-n/2}) \cdot \prod_{i=1}^n \int_0^\infty \exp(-\phi_u \cdot (z_i - x_i)^2 / 2 \omega_u) dG_u(\omega_u) \} \\ & \cdot \prod_{i=1}^n \int_0^\infty \exp(-\phi_\varepsilon \cdot (y_i - \beta_0 - \beta_1 \cdot x_i)^2 / 2 \omega_\varepsilon) dG_\varepsilon(\omega_\varepsilon) \} \\ & \cdot \exp \left\{ -\frac{1}{2} \left[(\beta_0 - b_0) / v_0 + (\beta_1 - b_1) / v_1 + \frac{1}{\sigma_X^2} \sum_{i=1}^n (\epsilon_i - \mu_X) \right] - \lambda_0 \cdot \phi_\varepsilon - \lambda_1 \cdot \phi_u \right\}, \end{aligned}$$

y la distribución a posteriori de los x_i , obtenida marginalizando en la expresión anterior, depende del proceso G y de su distribución a priori. y tiene la siguiente forma:

$$\begin{aligned} \pi(\mathbf{x} | \mathbf{y}, \mathbf{z}, G_\varepsilon, G_u) \propto & \exp \left[-\frac{1}{2 \cdot \sigma_X^2} \sum_{i=1}^n (\epsilon_i - \mu_X) \right] \cdot \prod_{i=1}^n \int_0^\infty \omega_\varepsilon^{-\frac{1}{2}} \exp\left(-\frac{\phi_\varepsilon}{2\omega_\varepsilon} (y_i - \beta_0 - \beta_1 \cdot x_i)^2\right) dG_\varepsilon(\omega_\varepsilon) \} \\ & \cdot \prod_{i=1}^n \int_0^\infty \omega_u^{-\frac{1}{2}} \exp\left(-\frac{\phi_u}{2\omega_u} (z_i - x_i)^2\right) dG_u(\omega_u) \} \\ & \cdot \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} (\phi_\varepsilon^{n/2+r_0-1} \cdot \phi_u^{n/2+r_1-1}) \cdot \exp \left\{ -\frac{1}{2} \left[(\beta_0 - b_0) / v_0 + (\beta_1 - b_1) / v_1 \right] - \lambda_0 \phi_\varepsilon - \lambda_1 \phi_u \right\} d\phi_\varepsilon d\phi_u d\boldsymbol{\beta} \square. \end{aligned}$$

Al igual que para el modelo independiente, una opción alternativa a la programación del muestreo desde las distribuciones condicionales completas, consiste en utilizar Winbugs.

Observación 5.2.

Es claro, tratándose de una mezcla discreta de normales, que si $p_{\varepsilon_i} = 1$ y $p_{\varepsilon_j} = 0 \forall j \neq i$ (es decir, si ω_ε tiene una distribución degenerada en un punto), y si ocurre lo propio con p_{u_i} y ω_u , ya no se produce la mezcla, por lo cual se recupera el modelo de términos de error normales, donde las distribuciones condicionales completas de los parámetros $\boldsymbol{\beta}$, ϕ_ε , ϕ_u y \mathbf{x} se mantienen intactas, pero \mathbf{p}_ε , \mathbf{p}_u , ω_ε y ω_u son constantes.

A continuación se presenta el resultado más importante de esta sección, que corresponde a la solución bayesiana semiparamétrica del problema de regresión con errores en las variables o con errores de medición MEM bajo el modelo elíptico representable dependiente.

Proposición 5.4

Bajo las especificaciones del modelo (5.3) y considerando que las distribuciones a priori para los parámetros de dispersión siguen sendos procesos Dirichlet G_ε y G_u , centrados en las distribuciones Gamma Inversa siguientes:

$$(5.26) \quad \omega_\varepsilon | G_\varepsilon \sim G_\varepsilon, \text{ con } G_\varepsilon \sim \mathcal{D}(c_\varepsilon G_0^{(\varepsilon)}) \text{ y } G_0^{(\varepsilon)} \sim IG(r_\varepsilon, \lambda_\varepsilon)$$

y

$$(5.27) \quad \omega_u | G_u \sim G_u, \text{ con } G_u \sim \mathcal{D}(c_u G_0^{(u)}) \text{ y } G_0^{(u)} \sim IG(r_u, \lambda_u),$$

se mantienen las distribuciones condicionales completas de $\phi_\varepsilon, \phi_u, \beta_0, \beta_1, x_i$, (5.21) a (5.25), y las distribuciones condicionales completas de ω_ε y ω_u están dadas en (5.18).

Observación 5.3:

Si las funciones generadoras de densidad elíptica h_ε y h_u son conocidas, G_ε y G_u tienen una distribución degeneradas en $G_0^{(\varepsilon)}$ y $G_0^{(u)}$, lo cual lleva exactamente a las mismas distribuciones condicionales completas del caso general.

Observación 5.4:

No se ha considerado aquí el caso en el cual ε y u no son independientes, sino que siguen una distribución conjunta elíptica $2n -$ variada, lo cual se propone como una extensión.

5.3 APLICACIONES

En esta sección se ilustran los resultados obtenidos en las secciones previas a través de simulación y conjuntos de datos considerados previamente en la literatura.

Aplicación 5.1 Retornos COPEC e IPSA (Arellano Valle, Galea & Iglesias, 2001)

Para ejemplificar la resolución del problema de errores de medición hemos considerado, en primer lugar, las variables Retornos de COPEC (Compañía de Petróleo de Chile) e Índice de Retornos chilenos IPSA (Índice de Precios Selectivo de Acciones). La *Tabla 5.1* presenta los datos (Arellano – Valle, Galea e Iglesias, 2001) y la *Figura 5.1.* presenta el diagrama de dispersión respectivo. Este último muestra una relación de tipo lineal con pendiente positiva, mientras que el análisis del modelo de regresión arroja un coeficiente de determinación de un 54%.

Tabla 5.1: Datos de Retornos COPEC e IPSA, Arellano – Valle, Galea & Iglesias (2001). La variable dependiente y corresponde a los retornos de COPEC y el IPSA z corresponde a la variable independiente medida con error aditivo.

COPEC	IPSA	COPEC	IPSA	COPEC	IPSA	COPEC	IPSA
-0.48003103	-0.28156442	-0.19046216	0.00541771	0.06659096	0.05386072	0.07775	0.04484269
-0.01100937	-0.05153124	0.1327871	0.0307852	0.00506667	0.06404124	0.09071429	0.08617927
0.02385385	0.09275289	0.10829551	0.12807764	0.01646104	0.1119	0.06955237	-0.03249705
-0.14104706	-0.05229667	0.01	0.02967131	0.06475	0.09739281	0.06142093	0.11728146
-0.25633333	-0.05440968	0.11182927	0.06159719	0.02987225	0.04176185	0.04176953	0.0176
0.15281304	0.01706912	0.03278355	0.05608883	0.06455098	0.05541891	0.03054096	0.11314347
0.05507407	0.10997923	0.07598351	0.0801	0.00551282	0.01790165	0.02616667	-0.00259894
0.05096552	0.11398305	0.19054717	0.13843805	-0.06922613	-0.10485395	-0.06262727	-0.03726085
0.01325807	0.00988551	0.062125	0.07134417	-0.06450526	-0.04031771	-0.0243	-0.02496656
0.0435	-0.06969334	0.10094203	0.01308214	-0.14287912	-0.09895773	0.03507068	-0.04537776
-0.13664706	0.05990326	0.1828052	0.06098137	0.06893852	0.0856	-0.035981	-0.0183176
0.11553333	0.0116085	0.09579565	0.09287556	-0.01595738	0.05657222	0.02637921	-0.02388452
0.33694118	0.09180509	0.43521961	0.14646175	-0.05110363	-0.05517374	-0.02457406	-0.04069839
-0.05143044	0.05328973	-0.07560582	-0.03078277	-0.14721233	-0.12349241	-0.01351186	-0.01496382
-0.05887376	-0.00690822	-0.01994638	0.03482423	-0.07490588	-0.03887955	0.10529085	0.16916063
0.14790476	0.01650204	-0.09629197	-0.00939736	0.20793291	0.27424318	0.1454	0.11083483
0.27698361	0.06392835	0.19828571	0.13839277	0.03217931	0.03493528	0.1339202	0.1783
-0.14098413	-0.03368827	0.2856536	0.1944615	0.02410133	0.08519517	0.08447312	0.19373153
-0.12379091	-0.08917699	-0.18352915	0.1034	-0.0018537	-0.02336298	0.01799055	0.07647358
-0.15695714	-0.05334051	0.1537988	0.05723754	-0.13302675	-0.08065974	-0.0573071	0.00925976
-0.00380952	0.01839639	-0.07441026	-0.0496042	0.0119	0.03347651	0.13098571	0.09290295
-0.10246744	-0.10484066	-0.03500435	-0.05287343	0.21071103	0.1599352	-0.00666222	-0.00504463
-0.1813	-0.09898656	-0.01847514	-0.05911721	0.07135	0.11918138	-0.00822807	-0.00861558
-0.12324928	0.01200544	0.1818022	0.11106271	-0.01405714	0.10339348	0.01044472	0.03009411
0.16041936	0.0677	0.1423633	0.20732077	0.03973759	0.13388515	0.21264242	0.01176795
-0.09349315	-0.02509219	0.14607937	0.1366962	-0.08250269	-0.07972309	-0.1141099	-0.07900275
0.06323529	0.10159045	0.22839401	0.10014899	-0.05727143	-0.0678309	-0.05556593	-0.104121
-0.11288108	-0.00896671	-0.23541475	-0.21364021	0.02256471	-0.01603748	-0.01094812	0.07000732
0.07594118	0.06637494	-0.13681111	-0.12838995	-0.13678582	-0.04241289	-0.07623906	-0.08988833
-0.03733333	-0.02129235	0.08716875	0.11262278	0.11725622	0.15145042		

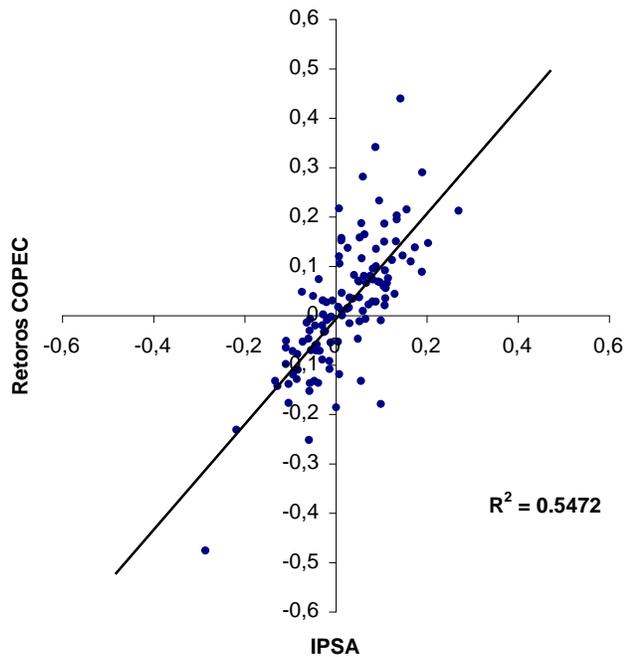


Figura 5.1. Diagrama de Dispersión para los datos de la Tabla 5.1.: Retornos de COPEC e IPSA.

La recta obtenida a partir del ajuste lineal de mínimos cuadrados basado en el modelo de regresión bajo el enfoque clásico, donde la variable independiente es IPSA, presenta intercepto cercano a cero y pendiente cercana a uno; la *Tabla 5.2.* presenta los estimadores de estos coeficientes de regresión.

Tabla 5.2: Estimadores de los coeficientes de regresión para el análisis clásico del modelo de regresión lineal simple, donde y corresponde a los Retornos COPEC y x corresponde al IPSA .

Parámetro	Estimador
β_0	0,0115
β_1	1,0681

Ahora bien, el análisis presentado por Arellano – Valle, Galea & Iglesias (2001) se basa en dos modelos jerárquicos. El primero corresponde al siguiente modelo normal:

$$\begin{aligned}
 y_i &\sim \text{i.i.d. } N(\beta_0 + \beta_1 x_i; \phi^{-1}), \\
 \beta_0 &\sim N(b_0; v_0), \\
 \beta_1 &\sim N(b_1; v_1) \\
 y \quad \phi &\sim \text{Gamma}(r_0; \lambda_0), \\
 \text{con } \beta_0 \perp \beta_1 \perp \phi.
 \end{aligned}$$

Tabla 5.3. Estimadores de Bayes para los datos de la Tabla 5.1., utilizando el modelo paramétrico normal (Arellano – Valle, Galea & Iglesias, 2001)

Parámetro	Media (Estimador de Bayes)	Desviación estandar	Mediana	Intervalo de credibilidad
β_0	- 0,0112	0,0084	- 0,0113	[- 0,0275; 0,0054]
β_1	1,066	0,0904	1,066	[0,8897; 1,244]
ϕ	134,4	17,42	133,7	[102; 170,3]

Los resultados de este primer análisis se presentan en la Tabla 5.3., y para su realización se utilizaron los valores siguientes para los hiperparámetros de las distribuciones a priori, escogidos de acuerdo con los economistas:

$$b_0 = 0, v_0 = \frac{1}{3}, b_1 = 1, v_1 = \frac{1}{4}, r_0 = \lambda_0 = 0,001.$$

El segundo modelo abordado por los autores antes mencionados es un modelo paramétrico jerárquico elíptico y corresponde al modelo t – student (v), que se vincula al modelo normal a través de variables latentes $\omega_i \sim \text{i.i.d. } G$, donde G es una distribución Gamma Inversa de parámetros ($\frac{1}{2} v; \frac{1}{2} v$). Los estimadores de Bayes para la pendiente obtenidos a partir de dicho análisis se presentan en la Tabla 5.4.

Tabla 5.4. Estimadores de Bayes de la pendiente β_1 para los datos de la Tabla 5.1., utilizando el modelo t – student con 1, 2 y 5 grados de libertad (Arellano – Valle, Galea & Iglesias, 2001).

v	Media a posteriori de β_1 (Estimador de Bayes)
2	0,9328
4	0,9724
10	1,0200

Como es lógico, a medida que los grados de libertad crecen, las estimaciones del modelo t – student se acercan a las estimaciones del modelo normal.

Nuestro análisis, basado en la *Proposición 5.2.*, y obtenido a partir de los resultados entregados por el programa computacional diseñado para este fin, utilizó parámetros análogos a los utilizados por Arellano – Valle, Galea e Iglesias (2001) para las distribuciones a priori:

$$b_0 = 0, v_0 = \frac{1}{3}, b_1 = 1, v_1 = \frac{1}{4}, r_0 = \lambda_0 = 0,001.$$

Por otro lado, se utilizó $\mu_x = 0,03$ y $\sigma_x^2 = 0,09$, valores de magnitudes semejantes a la media y varianza muestral de los valores IPSA, respectivamente.

Nuestro análisis se basa en un modelo semi paramétrico, por lo cual difiere del análisis de los autores Arellano – Valle, Galea & Iglesias (2001). En este sentido, consideramos G un proceso Dirichlet cuya distribución a priori esta centrada en la distribución Gamma Inversa utilizada por ellos.

Hemos considerado dos alternativas para los hiperparámetros de la distribución a priori de ω_{ϵ} , basándonos en los parámetros sugeridos por Arellano – Valle, Galea & Iglesias (2001). Estas alternativas surgen de la definición de los parámetros de forma y de escala de la distribución base del proceso Dirichlet, los cuales son iguales y toman los valores $r_U = \lambda_U = 2$ ($v = 4$ g. l.) y $r_U = \lambda_U = 5$ ($v = 10$ g. l.).

El objetivo de esta sección es realizar un análisis de sensibilidad, por esta razón hemos considerado diferentes combinaciones de los hiper parámetros r_1 y λ_1 de la distribución Gamma Inversa utilizada como base para la varianza de los términos de error de medición u_i :

- 101 y 0.1: error de medición poco variable (media de la variabilidad 0.001, con varianza pequeña 0.0001), es decir un modelo con poco error de medición y muy poco variable.
- 100001 y 100: error de medición poco variable en general (media de la variabilidad 0.001, con varianza 0.1 moderada), es decir un modelo con poco error en general, pero relativamente variable en ese sentido.
- 200001 y 200: error de medición poco variable en promedio, pero inestable (media de la variabilidad 0.001, con varianza grande 0.2).
- 1.04 y 0.002: error de medición medianamente variable (media de la variabilidad 0.05, con varianza pequeña 0.0001), es decir un modelo con moderada variabilidad del error de medición.
- 41 y 2: error de medición medianamente variable en promedio (media y varianza de la variabilidad moderadas, de valores 0.05 y 0.1, respectivamente), pero un tanto inestable.
- 81 y 4: error de medición medianamente variable en promedio y muy inestable (media de la variabilidad 0.05, con varianza grande 0.2).
- 1.0004 y 0.0002: error de medición muy variable (media de la variabilidad grande 0.5, con varianza pequeña 0.0001), es decir un modelo con errores muy variables.
- 1.4 y 0.2: error de medición muy variable en promedio (media de la variabilidad grande 0.5, con varianza moderada 0.1), es decir un modelo muy variable en promedio pero un tanto inestable.
- 1.8 y 0.4: error de medición muy variable en promedio (media de la variabilidad grande 0.5, con varianza grande 0.2), es decir un modelo muy variable en promedio y altamente inestable.

La *Tabla 5.5.* presenta las estimaciones de los coeficientes del modelo para todas las combinaciones de hiperparámetros mencionadas. Es claro a partir de estos resultados que las estimaciones no son sensibles a los valores de los hiperparámetros para la componente ϕ_{ii} , es decir, no son sensibles a la forma de la distribución de la componente de varianza de los errores de medición.

Tampoco se aprecia que los grados de libertad del modelo incidan sobre las estimaciones; sin embargo el modelo utilizado para el análisis tiene alguna trascendencia, en efecto, las estimaciones que surgen del análisis bayesiano semiparamétrico son levemente mayores a las obtenidas a partir del análisis bayesiano normal y a las que surgen del análisis de regresión clásico.

Tabla 5.5. Análisis de sensibilidad del modelo elíptico semi paramétrico del problema de errores de medición o de errores en las variables MEM para los datos de la Tabla 5.1.

	ESTIMADORES				Hiperparámetros distribución de ϕ_U		Interpretación hiper parámetros. Errores de medición con:	
	β_1		β_0		r_1	λ_1	Esperanza	Varianza
	4 g. l.	10 g. l.	4 g. l.	10 g. l.				
M. Calibración. Bayesiano Elíptico Semiparamétrico	1.0725	1.0738	-0.0116	-0.0114	101	0.1	Pequeña (0.001)	Pequeña (0.0001)
	1.0719	1.0718	-0.0114	-0.0116	100001	100		Mediana (0.1)
	1.0736	1.0720	-0.0115	-0.0115	200001	200		Grande (0.2)
	1.0827	1.0781	-0.0117	-0.0116	1.04	0.002	Mediana (0.05)	Pequeña (0.0001)
	1.0902	1.0892	-0.0121	-0.0120	41	2		Mediana (0.1)
	1.0863	1.0862	-0.0119	-0.0120	81	4		Grande (0.2)
	1.0710	1.0701	-0.0112	-0.0113	1.0004	0.0002	Grande (0.5)	Pequeña (0.0001)
	1.0953	1.0958	-0.0121	-0.0123	1.4	0.2		Mediana (0.1)
	1.0804	1.0774	-0.0117	-0.0117	1.8	0.4		Grande (0.2)
	0,9724	1,0200			M. Calib. Bayesiano Elíptico Param.			
	1,0681		0,0115		Modelo de Regresión Clásico			
	1,066		- 0,0112		M. Calib. Bayesiano Normal			

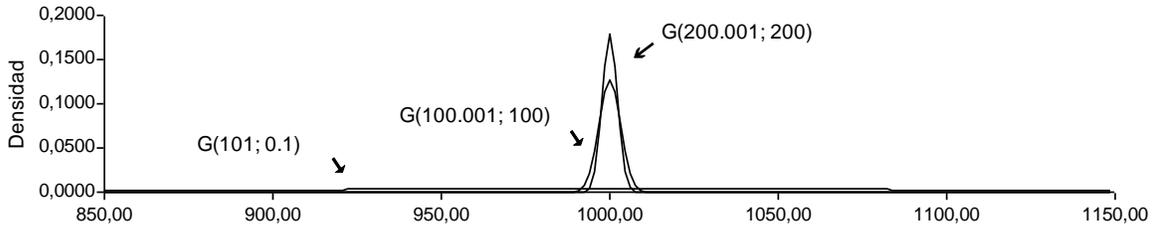
La *Figura 5.2.* presenta todas las alternativas consideradas como distribución a priori para el parámetro ϕ_u , mientras que las *Figuras 5.3.* y *5.4.* presentan las correspondientes distribuciones a posteriori considerando los modelos con 4 y 10 grados de libertad, respectivamente.

Observando las *Figuras 5.3.* y *5.4.* no se aprecian diferencias, de hecho las medias y las varianzas de las cadenas son prácticamente idénticas. Es claro, entonces, que las variaciones consideradas en los grados de libertad no inciden sobre las distribuciones a posteriori.

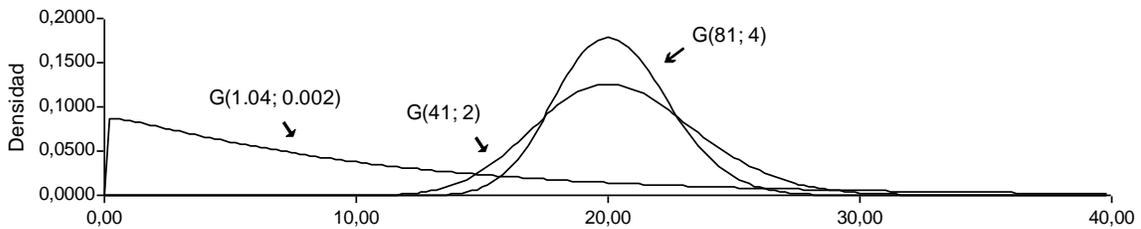
Como hemos mencionado antes, la obtención analítica de las distribuciones a posteriori para los parámetros no resulta factible, lo cual hace importante indagar respecto de la forma de dicha distribución. En forma anecdótica se incluyen los ajustes normales y gama para las distribuciones a posteriori de ϕ , junto a los resultados del test no paramétrico de Kolmogorov – Smirnov (*Tabla 5.6.*)

Comparando las *Figuras 5.2.*, *5.3.* y *5.4.*, y considerando la información en la *Tabla 5.6.*, es claro que cuando las distribuciones a priori de las componentes de precisión de los errores de medición ϕ_U inducen una varianza con media pequeña (ϕ_U a priori muy variable), las distribuciones a posteriori del mismo parámetro resultan muy parecidas a las distribuciones a priori, pero la varianza se reduce un tanto. En ese caso también se obtiene que las distribuciones a posteriori son simétricas y presentan un buen ajuste normal. Ahora bien, para el caso en que la varianza a priori de ϕ_U es muy grande, la estimación bayesiana de la componente ϕ_U no resulta confiable, debido a que su distribución a posteriori también es muy variable.

- a) Distribuciones a priori para ϕ_U asociadas una distribución para la varianza de los errores de medición con media pequeña y varianzas pequeña, mediana y grande (de izquierda a derecha).



- b) Distribuciones a priori para ϕ_U asociadas una distribución para la varianza de los errores de medición con media moderada y varianzas pequeña, mediana y grande (de izquierda a derecha).



- c) Distribuciones a priori para ϕ_U asociadas una distribución para la varianza de los errores de medición con media grande y varianzas pequeña, mediana y grande (de izquierda a derecha).

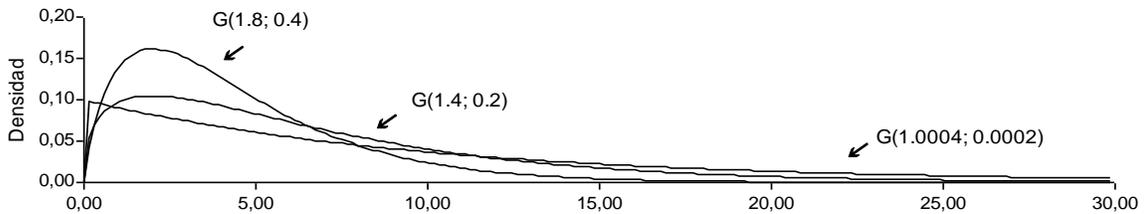
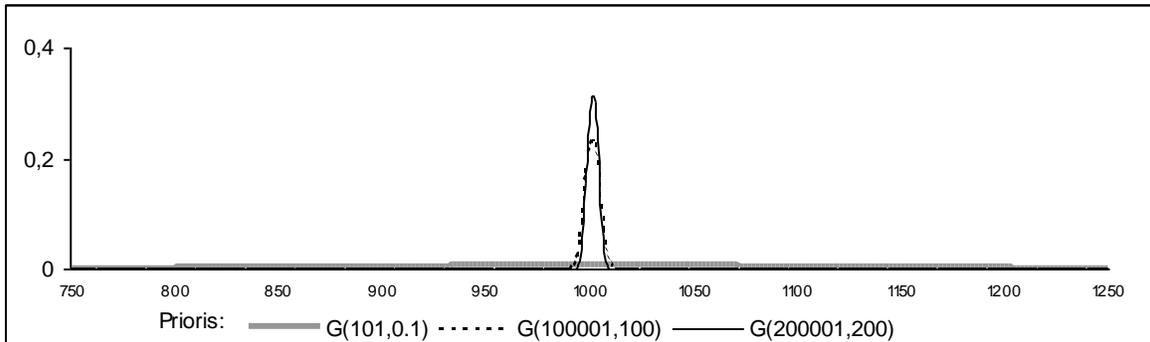
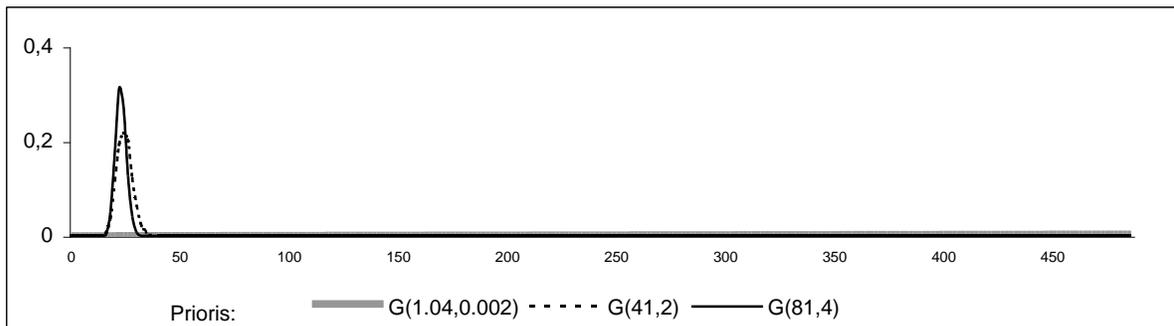


Figura 5.2. Distribuciones a priori para el parámetro ϕ_U .

a) Distribuciones a posteriori para ϕ_U donde las prioris presentaban una distribución para la varianza de los errores con media pequeña y varianzas pequeña, mediana y grande, respectivamente.



b) D Distribuciones a posteriori para ϕ_U donde las prioris presentaban una distribución para la varianza de los errores con media mediana y varianzas pequeña, mediana y grande, respectivamente.



c) Distribuciones a posteriori para ϕ_U donde las prioris presentaban una distribución para la varianza de los errores con media grande y varianzas pequeña, mediana y grande, respectivamente.

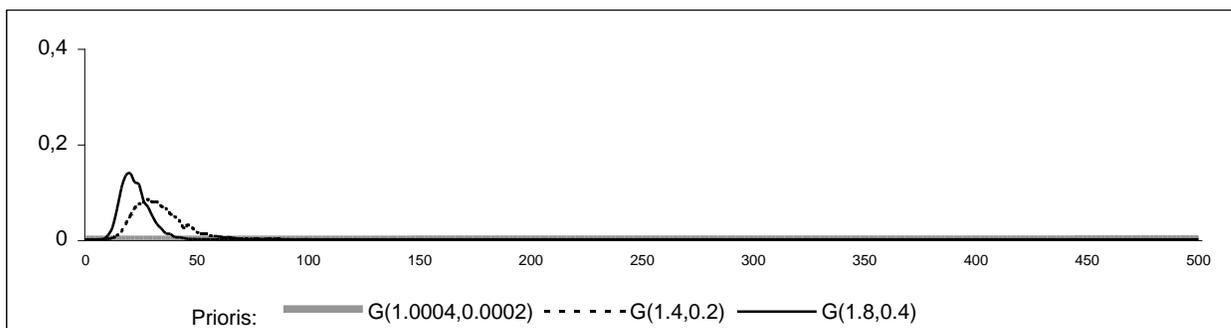
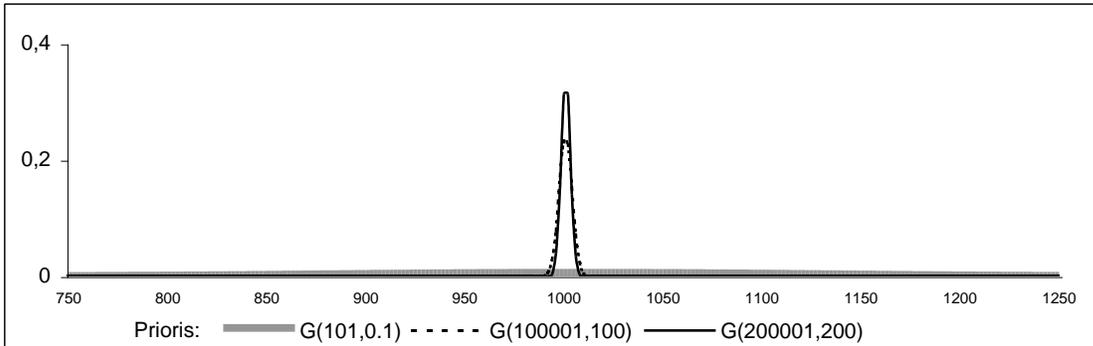
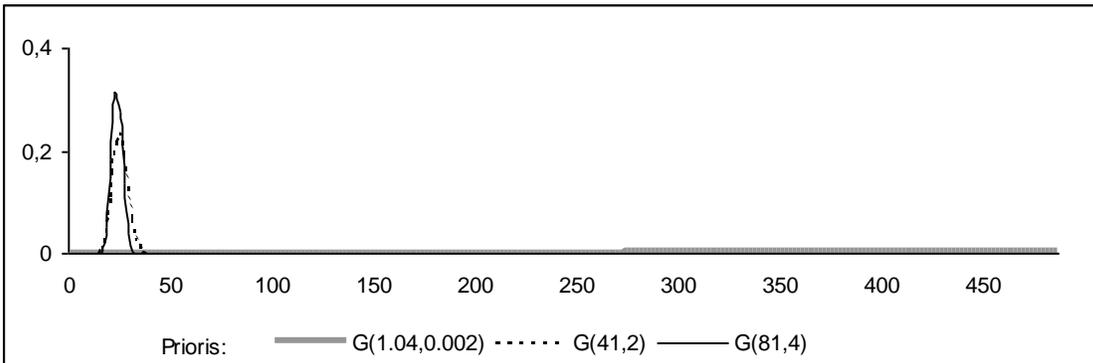


Figura 5.3. Distribuciones a posteriori para el parámetro ϕ_U con el modelo con errores de medición semiparamétrico elíptico de 4 grados de libertad.

a) Distribuciones a posteriori para ϕ_U donde las prioris presentaban una distribución para la varianza de los errores con media pequeña y varianzas pequeña, mediana y grande, respectivamente.



b) Distribuciones a posteriori para ϕ_U donde las prioris presentaban una distribución para la varianza de los errores con media mediana y varianzas pequeña, mediana y grande, respectivamente.



c) Distribuciones a posteriori para ϕ_U donde las prioris presentaban una distribución para la varianza de los errores con media grande y varianzas pequeña, mediana y grande, respectivamente.

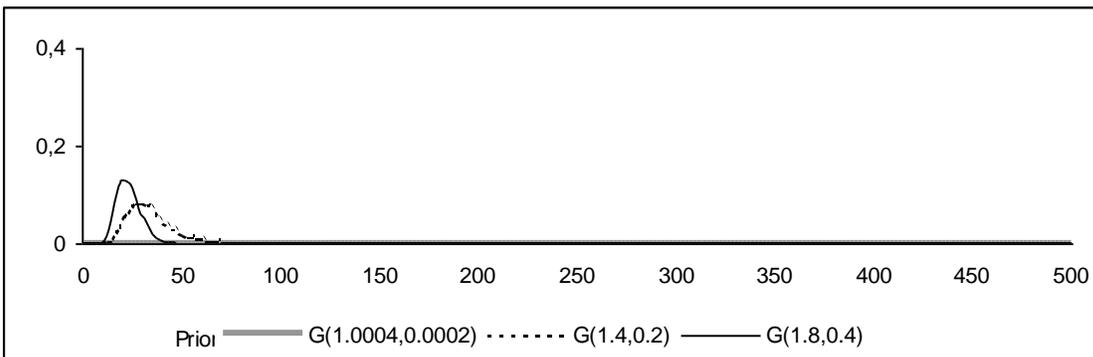


Figura 5.4. Distribuciones a posteriori para el parámetro ϕ_U con el modelo con errores de medición semiparamétrico elíptico de 10 grados de libertad.

Tabla 5.6. Ajuste normal y gamma para las posteriores de ϕ_U , mediante estadístico de Kolmogorov – Smirnov y basado en los análisis para los datos de la Tabla 5.1.

Hiperparámetros de la distribución a priori de ϕ_U		Ajuste Normal				Ajuste Gama			
		v = 4 g. l.		v = 10 g. l.		v = 4 g. l.		v = 10 g. l.	
r_1	λ_1	Estimadores	Valor p	Estimadores	Valor p	Estimadores	Valor p	Estimadores	Valor p
101	0.1	(1005; 14872)	0.0652	(1010; 14738)	0.1051	(67.42; 0.07)	<0.0001	(68.61;0.07)	<0.0001
100001	100	(1000; 10.4)	0.8638	(1000; 10.6)	0.9992	----	----	----	----
200001	200	(1000; 5.2)	0.9655	(1000; 5.1)	0.9825	----	----	----	----
1.04	0.002	----	----	----	----	(1.58;0.002)	<0.0001	(1.60;0.002)	<0.0001
41	2	----	----	----	----	(45.41;1.94)	0.7742	(44.68;1.90)	0.8839
81	4	(21.79; 5.9)	0.0517	(21.79; 5.8)	0.0978	(80.14;3.68)	0.0948	(82.06;3.77)	0.0472
1.0004	0.0002	----	----	----	----	(1; 0.00018)	<0.0001	(1; 0.00018)	<0.0001
1.4	0.2	----	----	----	----	(8.988;0.08)	0.0188	(8.876;0.28)	0.0010
1.8	0.4	----	----	----	----	(12.01;0.57)	0.0024	(11.88;0.56)	0.0159

El diagnostico de Geweke mostró convergencia de las cadenas simuladas y los resultados del análisis de sensibilidad y la comparación con los análisis mencionados antes, llevan a concluir que las estimaciones resultan altamente estables.

A continuación se presentan algunas conclusiones importantes que derivan del análisis:

- Si la distribución a priori para ϕ_U induce una varianza de los errores de medición cuya media asume un valor moderado y cuyas varianzas son medianas o grandes (ϕ_U a priori no muy variable), las distribuciones a priori y a posteriori son aproximadamente simétricas, pero la variabilidad de las distribuciones a posteriori es menor y la media de las distribuciones a posteriori es superior a la de las distribuciones a priori.
- Si la distribución a priori para ϕ_U induce una varianza de los errores de medición cuya media asume un valor grande y cuyas varianzas son moderadas o grandes (ϕ_U a priori no muy variable), las distribuciones a priori y a posteriori son asimétricas, pero la variabilidad de las distribuciones a posteriori es menor y la media de las distribuciones a posteriori es superior a la de las distribuciones a priori. En el caso en que la varianza es moderada, la distribución a posteriori presenta un buen ajuste gamma, y los parámetros en este ajuste se asemejan mucho a los de la distribución a priori. Para el caso en que la varianza es grande, la distribución a posteriori no resulta significativamente distinta de la normal.
- Si ϕ_U a priori tiene una distribución muy variable, a posteriori esta variabilidad aumenta.
- El único caso en el cual se conserva la media en la distribución a posteriori de ϕ_U , es aquel en el cual la media de la variabilidad de los errores de medición es pequeña, y por ende los errores de medición son uniformemente pequeños. Para los otros casos la distribución a posteriori de la componente ϕ_U presenta no sólo una enorme variabilidad, sino también una media muy grande.
- A pesar de todas estas consideraciones y de que las distribuciones a posteriori de ϕ_U presenten tal variedad, la estimación de los coeficientes β_0 y β_1 no se ve afectada por los hiperparámetros considerados para la distribución a priori de ϕ_U , como se verificó anteriormente.

Aplicación 5.2 Comparación de modelos normal y elíptico para datos de resistencia de concreto (Wellman & Gunst, 1991).

Arellano – Valle, Iglesias & Vidal (2002a) presentan el análisis bayesiano del modelo normal con errores de medición para datos de resistencia de concreto a la compresión. La resistencia fue medida en dos momentos: el segundo (z_i) y el vigésimo octavo día (y_i) después de verterlo; en ambas oportunidades la medición fue afectada por varias fuentes de error, por lo cual la verdadera resistencia del concreto al segundo día x_i es desconocida. La *Tabla 5.7.* presenta los datos, para los cuales se asumió que la relación existente entre la resistencia medida al octavo día con error y la verdadera resistencia al segundo día es lineal. El problema corresponde a un modelo con errores de medición MEM, para el cual, asumiendo que los errores de medición se distribuyen normalmente, los autores determinaron los estimadores de Bayes para los coeficientes de la relación lineal.

Tabla 5.7. Resistencia de concreto medida con error al segundo y vigésimo octavo día (Arellano – Valle, Iglesias y Vidal, 2002a).

observación	día 2	día 28	observación	día 2	día 28	observación	Día 2	día 28
1	2830	4470	15	2985	4690	29	3335	4650
2	3295	4740*	16	3135	4880	30	3800	4680
3	2710	5115	17	2750	3425	31	2680	5165
4	2855	4880	18	3205	4265	32	3760	5075
5	2980	4445	19	3000	4485	33	3605	4710
6	3065	4080	20	3035	5220	34	2005	4200
7	3765	5390	21	4245	7695	35	2495	4645
8	3265	4045	22	1635	3330	36	3205	4725
9	3170	4370	23	2270	4065	37	2060	4695
10	2895	4955*	24	2895	4715	38	3425	5470
11	2630	3835	25	2845	4735	39	3315	4330
12	2830	4290	26	2205	3605	40	3825	4950
13	2935	4600	27	3590	4670	41	3160	4460
14	3115	4605	28	3080	4720			

Como consecuencia de dicho análisis los autores antes mencionados detectaron seis observaciones influyentes (observaciones 17, 21, 22, 26, 34 y 37). La *Figura 5.5.* muestra el diagrama de dispersión con los outliers destacados, que deja de manifiesto la existencia de una significativa correlación entre las mediciones de resistencia del concreto al segundo y al vigésimo octavo día del vertido.

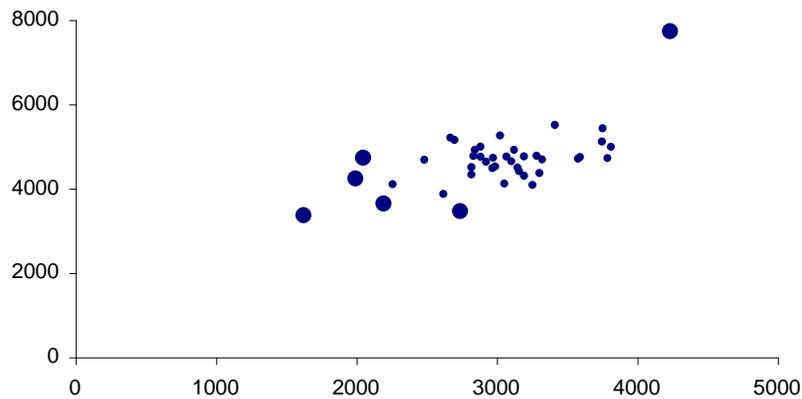


Figura 5.5. Diagrama de Dispersión de las observaciones de resistencia de concreto presentadas en la Tabla 5.7. (Arellano – Valle, Iglesias y Vidal, 2002a).

Esto es confirmado por el análisis clásico de regresión simple (p - valor $< 0,0001$), el cual es presentado en la Tabla 5.8.

Tabla 5.8. Análisis clásico de regresión simple para los datos de resistencia de concreto (Tabla 5.7.).

F.V.	SC	gl	CM	F	p-valor
Modelo	6937188,87	1	6937188,87	23,13	<0,0001
Error	11697022,11	39	299923,64		
Total	18634210,98	40			
N = 41	$R^2 = 37,23\%$				
Coefficientes	Estimador	Error Estándar	Estadístico T – Student	p-valor	
Intercepto β_0	2250,71	503,29	4,47	0,0001	
Pendiente β_1	0,79	0,16	4,81	<0,0001	

Como alternativa al análisis realizado por Arellano – Valle, Iglesias & Vidal (2002a), hemos abordado el problema considerando errores elípticos, utilizando para la inferencia modelos semiparamétricos dependientes e independientes, donde las variables subyacentes siguen un proceso Dirichlet con distribución base Gamma Inversa asociada a 4 y a 50 grados de libertad. Nuestro objetivo es comparar estos resultados con aquellos derivados del modelo paramétrico normal, en lo relativo a las observaciones influyentes.

Los hiperparámetros utilizados en los análisis se presentan a continuación y se determinan de modo que las medias de las distribuciones de los coeficientes y de los x_i se asemejen a las estimaciones para el modelo clásico de regresión (*Tabla 5.8.*) y las varianzas sean grandes.

- β_0 tiene una distribución a priori normal con media 2.000 y desviación estándar 1.000.
- β_1 tiene una distribución a priori normal con media y varianza igual a 1.
- ϕ_ε y ϕ_u tienen distribución gamma de parámetro de forma 3 y parámetro de escala 9.
- $\mu_x = 3.000$ y $\sigma_x^2 = 350.000$.
- ω_ε y ω_u tienen la misma distribución gamma inversa de parámetros de forma y escala 2 (para el modelo asociado a 4 grados de libertad) y 25 (para el modelo asociado a 50 grados de libertad).

El criterio utilizado aquí para verificar si el ajuste con el modelo semiparamétrico elíptico es mejor en el sentido de que logra que ya no existan observaciones influyentes, consiste en calcular la diferencia entre el estimador obtenido a partir de la muestra completa y el estimador obtenido a partir de la muestra sin el dato influyente y el porcentaje de reducción obtenida. Los resultados se presentan en la *Tabla 5.9.*

Es claro desde la *Tabla 5.9.* que los modelos semiparamétricos elípticos capturan adecuadamente estas observaciones que bajo el modelo paramétrico normal resultaban influyentes. Las reducciones relativas en la diferencia de estimaciones son notables, en efecto, respecto de la pendiente son todas superiores a un 51%, y en lo relativo al intercepto, superiores a un 56%. Además la reducción es mucho mayor en aquellas observaciones que eran más influyentes, principalmente la observación 21 y luego la 37 y la 17.

Tabla 5.9. Comparación de modelos bayesiano paramétrico normal y semiparamétricos elípticos dependiente e independiente en el ajuste para las observaciones influyentes pertenecientes al conjunto de datos de resistencia de concreto (Arellano – Valle, Iglesias & Vidal, 2002a).

Diferencia entre estimadores con la muestra completa y sin la observación influyente.

Dato	Modelo		$\beta_0 - \beta_{0(i)}$	Reducción	$\beta_1 - \beta_{1(i)}$	Reducción
17	Semiparamétrico Elíptico (4 g.l.)	Dependiente	26.060	84%	-0.008	82%
		Independiente	1.344	99%	-0.001	99%
	Semiparamétrico Elíptico (50 g.l.)	Dependiente	9.572	94%	-0.004	92%
		Independiente	-12.657	92%	0.005	90%
	Paramétrico Normal		-166.682		0.047	
21	Semiparamétrico Elíptico (4 g.l.)	Dependiente	27.124	98%	-0.009	98%
		Independiente	-22.121	98%	0.007	99%
	Semiparamétrico Elíptico (50 g.l.)	Dependiente	0.497	100%	-0.001	100%
		Independiente	-24.940	98%	0.008	98%
	Paramétrico Normal		1385.400		0.480	
22	Semiparamétrico Elíptico (4 g.l.)	Dependiente	-10.411	91%	0.003	93%
		Independiente	-35.894	71%	0.012	70%
	Semiparamétrico Elíptico (50 g.l.)	Dependiente	20.819	83%	-0.007	81%
		Independiente	-16.374	87	0.005	87%
	Paramétrico Normal		-121.964		0.039	
26	Semiparamétrico Elíptico (4 g.l.)	Dependiente	14.427	86%	-0.006	81%
		Independiente	-2.379	98%	0.001	97%
	Semiparamétrico Elíptico (50 g.l.)	Dependiente	43.140	57%	-0.015	52%
		Independiente	-20.872	79%	0.007	76%
	Paramétrico Normal		-99.600		0.030	
34	Semiparamétrico Elíptico (4 g.l.)	Dependiente	19.935	85%	-0.007	82%
		Independiente	20.498	84%	-0.007	82%
	Semiparamétrico Elíptico (50 g.l.)	Dependiente	18.980	85%	-0.007	83%
		Independiente	-40.713	69%	0.013	66%
	Paramétrico Normal		129.919		-0.040	
37	Semiparamétrico Elíptico (4 g.l.)	Dependiente	18.269	92%	-0.007	89%
		Independiente	1.383	99%	0.000	100%
	Semiparamétrico Elíptico (50 g.l.)	Dependiente	-3.817	98%	0.001	99%
		Independiente	-30.515	87%	0.010	85%
	Paramétrico Normal		232.779		-0.069	

Aplicación 5.3 Simulación de un modelo con errores de medición MEM y posterior análisis con los modelos semiparamétricos elípticos (independiente y dependiente) y paramétrico normal.

Se realizó una simulación para el modelo con errores de medición MEM donde la regresión contempla un intercepto $\beta_0 = 2$ y una pendiente $\beta_1 = 1$. Los valores de los otros parámetros se varían en tres sentidos: en primer lugar, de tal manera que el cociente entre $\sigma_\varepsilon = \phi_\varepsilon^{-1/2}$, la desviación estándar del error ε en la regresión y $\sigma_u = \phi_u^{-1/2}$, la desviación estándar del error de medición u asuma valores 0.1, 1 y 5, es decir, el error de medición con dispersión mayor, igual y menor a la del error de la regresión; en segundo lugar, la varianza del error de medición $\sigma_u^2 = \phi_u^{-1}$ varía asumiendo valores 0.05, 0.55, 1.05 y 1.55, y en tercer lugar, los errores se asumen elípticos independientes y elípticos dependientes (en particular se utilizó errores t – student con dos grados de libertad). Todos los tamaños de muestra son $n = 20$.

El análisis utiliza las siguientes distribuciones a priori:

- β_0 y β_1 normales con varianza 0.2 y medias 2 y 1 respectivamente.
- ϕ_ε y ϕ_u gamma con parámetros tales que la varianza sea igual a 1 y la media corresponda a los valores asignados a estos parámetros en la simulación.
- ω_ε y ω_u ambas gamma inversa de parámetros de forma y escala iguales a 2. Según si el modelo es dependiente o independiente, se trata de una variable aleatoria unidimensional o n-dimensional de componentes gamma inversa, respectivamente.

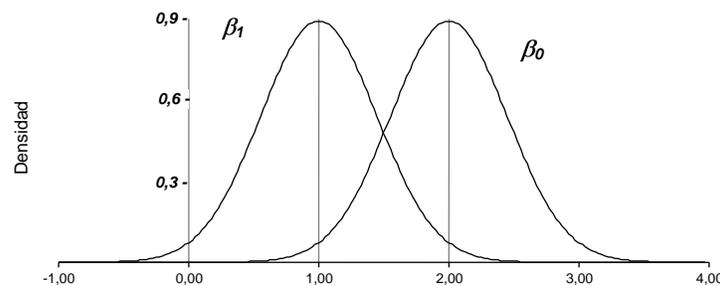


Figura 5.6. Distribución a priori de los coeficientes β_0 y β_1 .

Las Figuras 5.6 y 5.7 presentan las distribuciones a priori y a posteriori, de los coeficientes β_0 y β_1 , respectivamente.

Para cada conjunto simulado se realizó la estimación de parámetros con los modelos semiparamétrico elíptico dependiente, Semiparamétrico elíptico independiente y paramétrico normal. El objetivo de esta sección es comparar los modelos semiparamétricos elíptico dependiente e independiente entre ellos y determinar sus bondades respecto del modelo paramétrico normal.

La Tabla 5.10. presenta el análisis de los datos simulados a partir de errores independientes t – student con 2 grados de libertad mediante los tres modelos aludidos y la Tabla 5.11. hace lo propio con los datos simulados a partir de errores t – student dependientes con 2 grados de libertad.

Tabla 5.10. Análisis de los datos simulados para un modelo con errores de medición MEM donde los errores son t – student independientes con dos grados de libertad. Los modelos utilizados para analizar los datos son semiparamétrico elíptico dependiente, independiente y normal.

	$\delta = \sigma_{\varepsilon} / \sigma_U = 0,1$	$\delta = \sigma_{\varepsilon} / \sigma_U = 1$	$\delta = \sigma_{\varepsilon} / \sigma_U = 5$
$\sigma_U^2 =$ $1/\phi_U = 0,05$	M Dep: $\beta_0 = 2.02, \beta_1 = 0.94$ M. Ind: $\beta_0 = 2.01, \beta_1 = 0.86$ M. Nor: $\beta_0 = 2.00, \beta_1 = 0.90$	M Dep: $\beta_0 = 2.11, \beta_1 = 0.93$ M. Ind: $\beta_0 = 2.17, \beta_1 = 0.67$ M. Nor: $\beta_0 = 2.06, \beta_1 = 0.93^*$	M Dep: $\beta_0 = 1.79, \beta_1 = 1.25$ M. Ind: $\beta_0 = 1.75, \beta_1 = 1.11$ M. Nor: $\beta_0 = 2.00, \beta_1 = 1.23$
$\sigma_U^2 =$ $1/\phi_U = 0,55$	M Dep: $\beta_0 = 1.91, \beta_1 = 0.83$ M. Ind: $\beta_0 = 1.92, \beta_1 = 0.87$ M. Nor: $\beta_0 = 2.18, \beta_1 = 0.88$	M Dep: $\beta_0 = 1.84, \beta_1 = 0.90$ M. Ind: $\beta_0 = 1.86, \beta_1 = 0.93^*$ M. Nor: $\beta_0 = 1.82, \beta_1 = 0.89$	M Dep: $\beta_0 = 2.08, \beta_1 = 1.46$ M. Ind: $\beta_0 = 2.04, \beta_1 = 1.26$ M. Nor: $\beta_0 = 2.17, \beta_1 = 1.19$
$\sigma_U^2 =$ $1/\phi_U = 1,05$	M Dep: $\beta_0 = 2.12, \beta_1 = 0.93$ M. Ind: $\beta_0 = 2.11, \beta_1 = 0.95^*$ M. Nor: $\beta_0 = 1.86, \beta_1 = 0.95$	M Dep: $\beta_0 = 1.84, \beta_1 = 1.02$ M. Ind: $\beta_0 = 1.87, \beta_1 = 1.02^*$ M. Nor: $\beta_0 = 1.89, \beta_1 = 1.13$	M Dep: $\beta_0 = 1.75, \beta_1 = 0.54$ M. Ind: $\beta_0 = 1.75, \beta_1 = 0.91$ M. Nor: $\beta_0 = 2.10, \beta_1 = 0.82$
$\sigma_U^2 =$ $1/\phi_U = 1,55$	M Dep: $\beta_0 = 1.80, \beta_1 = 0.84$ M. Ind: $\beta_0 = 1.80, \beta_1 = 0.85$ M. Nor: $\beta_0 = 1.96, \beta_1 = 1.03^*$	M Dep: $\beta_0 = 2.13, \beta_1 = 0.89$ M. Ind: $\beta_0 = 2.14, \beta_1 = 0.91$ M. Nor: $\beta_0 = 2.05, \beta_1 = 0.85$	M Dep: $\beta_0 = 1.78, \beta_1 = 0.51$ M. Ind: $\beta_0 = 1.78, \beta_1 = 0.91$ M. Nor: $\beta_0 = 1.82, \beta_1 = 0.60$

* : La mejor estimación.

La generación de las muestras requirió, para el caso independiente, variables normales para cada observación simulada, las cuales dan origen a las variables t – student utilizando una variable aleatoria Chi – Cuadrado diferente en cada caso. Para el caso dependiente, sólo la variable normal cambia, pero la variable Chi – Cuadrado es única.

Tabla 5.11. Análisis de los datos simulados para un modelo con errores de medición MEM donde los errores son t – student dependientes con dos grados de libertad. Los modelos utilizados para analizar los datos son semiparamétrico elíptico dependiente, Semiparamétrico elíptico independiente y paramétrico normal.

	$\delta = \sigma_{\varepsilon} / \sigma_U = 0,1$	$\delta = \sigma_{\varepsilon} / \sigma_U = 1$	$\delta = \sigma_{\varepsilon} / \sigma_U = 5$
$\sigma_U^2 =$ $1/\phi_U = 0,05$	M Dep: $\beta_0 = 2.01, \beta_1 = 0.77$ M. Ind: $\beta_0 = 2.02, \beta_1 = 0.82$ M. Nor: $\beta_0 = 2.00, \beta_1 = 0.88^*$	M Dep: $\beta_0 = 2.01, \beta_1 = 1.03^*$ M. Ind: $\beta_0 = 2.09, \beta_1 = 0.94$ M. Nor: $\beta_0 = 2.05, \beta_1 = 0.92$	M Dep: $\beta_0 = 1.97, \beta_1 = 1.25$ M. Ind: $\beta_0 = 2.04, \beta_1 = 1.10$ M. Nor: $\beta_0 = 2.01, \beta_1 = 1.23$
$\sigma_U^2 =$ $1/\phi_U = 0,55$	M Dep: $\beta_0 = 2.26, \beta_1 = 0.88$ M. Ind: $\beta_0 = 2.32, \beta_1 = 0.80$ M. Nor: $\beta_0 = 2.17, \beta_1 = 0.88^*$	M Dep: $\beta_0 = 1.82, \beta_1 = 0.89$ M. Ind: $\beta_0 = 1.81, \beta_1 = 0.91$ M. Nor: $\beta_0 = 1.82, \beta_1 = 0.89$	M Dep: $\beta_0 = 2.18, \beta_1 = 1.15$ M. Ind: $\beta_0 = 2.18, \beta_1 = 1.08^*$ M. Nor: $\beta_0 = 2.19, \beta_1 = 1.14$
$\sigma_U^2 =$ $1/\phi_U = 1,05$	M Dep: $\beta_0 = 1.87, \beta_1 = 0.95$ M. Ind: $\beta_0 = 1.88, \beta_1 = 0.96$ M. Nor: $\beta_0 = 1.87, \beta_1 = 0.95$	M Dep: $\beta_0 = 1.89, \beta_1 = 1.13$ M. Ind: $\beta_0 = 1.93, \beta_1 = 1.11$ M. Nor: $\beta_0 = 1.90, \beta_1 = 1.13$	M Dep: $\beta_0 = 2.10, \beta_1 = 0.82$ M. Ind: $\beta_0 = 2.10, \beta_1 = 0.96^*$ M. Nor: $\beta_0 = 2.11, \beta_1 = 0.81$
$\sigma_U^2 =$ $1/\phi_U = 1,55$	M Dep: $\beta_0 = 1.97, \beta_1 = 1.03^*$ M. Ind: $\beta_0 = 1.93, \beta_1 = 1.04$ M. Nor: $\beta_0 = 1.97, \beta_1 = 1.03^*$	M Dep: $\beta_0 = 2.04, \beta_1 = 0.85$ M. Ind: $\beta_0 = 2.02, \beta_1 = 0.84$ M. Nor: $\beta_0 = 2.03, \beta_1 = 0.86$	M Dep: $\beta_0 = 1.81, \beta_1 = 0.60$ M. Ind: $\beta_0 = 1.81, \beta_1 = 0.92^*$ M. Nor: $\beta_0 = 1.81, \beta_1 = 0.59$

* : La mejor estimación.

Algunas observaciones que surgen de la Tabla 5.10. y que se pueden realizar respecto de esta simulación y del posterior análisis de los datos, son las siguientes:

- En general las estimaciones obtenidas mediante los modelos semiparamétricos se asemejan. En gran cantidad de casos resultan ser las más precisas.
- Generalmente, la estimación del intercepto es semejante para los modelos semiparamétricos elípticos y resulta más precisa y además coincidente para todos los modelos, si la varianza del error de medición es muy pequeña y relativamente muy superior a la otra varianza. Ahora bien, si los errores de medición tienen variabilidad σ_u^2 pequeña (0.05 en nuestro caso), la estimación de los interceptos mediante el modelo paramétrico normal (PN) es bastante precisa.
- Las mejores estimaciones de la pendiente para el modelo semiparamétrico elíptico independiente (SEI) se obtienen si la varianza del error de medición es grande (1,05) y coincide con la otra varianza.
- Cuando ambas varianzas coinciden o la varianza del error de medición es superior, la estimación de la pendiente mejora a medida que la varianza del error de medición crece, excepto si es demasiado grande (1,55).
- La mejor estimación de ambos coeficientes se debe al modelo paramétrico normal (PN) y se para el caso en que la varianza del error de medición asume su mayor valor y la otra varianza asume su menor valor relativo respecto de ella.
- En general, el modelo SEI no presenta un compartamiento óptimo en la estimación para errores t – student independientes.

De la Tabla 5.11. también surgen algunas observaciones:

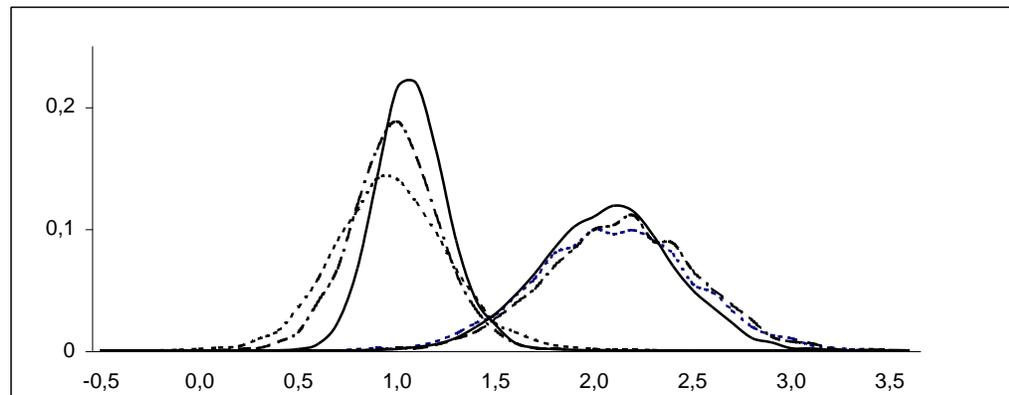
- En general las estimaciones obtenidas mediante el modelo PN y el modelo semiparamétrico elíptico dependiente (SED) son muy semejantes y resultan más precisas que las estimaciones obtenidos a partir del modelo SEI.
- En general la estimación del intercepto es semejante para todos los modelos y resulta más precisa si la varianza del error de medición es muy pequeña (independientemente del valor de la otra varianza) o si es muy grande pero la otra varianza asume un valor menor o igual a ella.
- Si ambas varianzas coinciden los estimadores de los tres modelos se asemejan bastante.
- Si la varianza del error de medición es muy superior a la otra varianza, las estimaciones de la pendiente mejoran a medida que la varianza del error de medición crece, pero cuando ambas varianzas coinciden, la estimación de la pendiente se desmejora a medida que la varianza del error de medición crece. El modelo SEI supera en la estimación de la pendiente a los otros modelos cuando la varianza de los errores en el modelo de regresión es muy grande y comparativamente muy superior a la varianza del error de medición.
- Una estimación precisa de ambos coeficientes se logra para el modelo SED en el caso en que la varianza de ambos errores coincide y es muy pequeña o en el caso en que la varianza del error de medición es muy grande y la otra varianza es comparativamente muy pequeña.
- En general, el modelo SED es adecuado para el caso del respectivo modelo con errores elípticos dependientes (en nuestro caso t – student dependientes).

En resumen, desde las Tablas 5.10. y 5.11. se desprende que:

- El modelo SED se comporta de forma adecuada, no así el modelo SEI.
- Para errores t – student dependientes, el modelo respectivo se comporta semejante al modelo PN, mientras que para errores t – student independientes, los modelos semiparamétricos se comportan en forma semejante en lo relativo a estimación.

Ahora bien, al comparar las Figuras 5.6. y 5.7., es claro que las distribuciones a posterioris de los coeficientes no distan mucho de la normalidad, presente en las prioris, pero también es claro que las varianzas, que para las distribuciones a priori (Figura 5.6.) eran iguales para ambos coeficientes, en el caso de las distribuciones a posteriori (Figura 5.7.) presentan variaciones, reduciéndose respecto de su valor inicial para ambos coeficientes, pero mostrando más dispersión en el caso de los interceptos (lo que lleva a estimaciones menos precisas en dicho caso).

a) Errores elípticos (t – student) dependientes (β_1 a la izquierda y β_0 a la derecha)



b) Errores elípticos (t - student) independientes (β_1 a la izquierda y β_0 a la derecha)

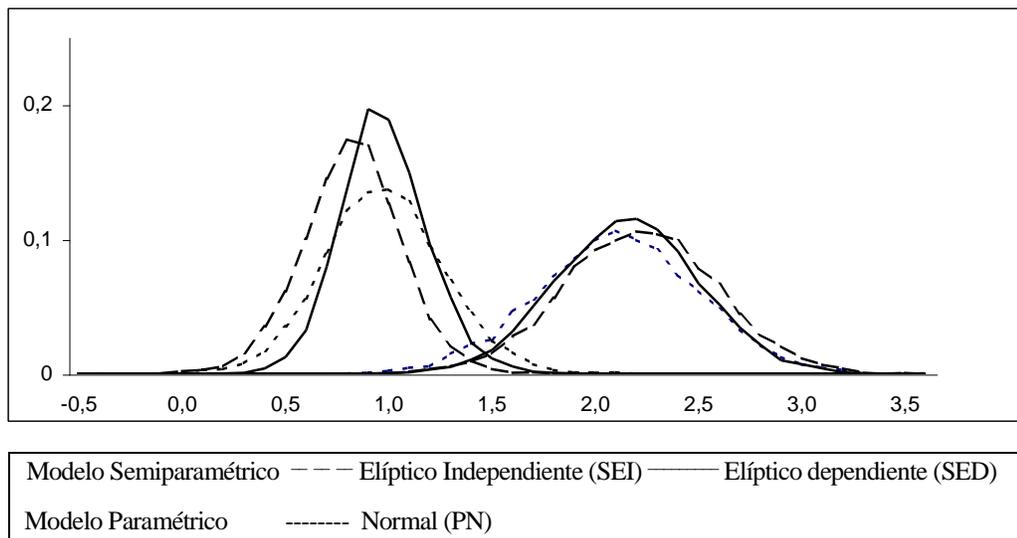


Figura 5.7. Distribuciones a posteriori de los coeficientes β_0 (a la derecha) y β_1 (a la izquierda) para los distintos modelos, donde las varianzas de los errores coinciden y son pequeñas.

La Tabla 5.12. presenta las medidas de tendencia central para las cadenas que surgen del análisis (distribución a posteriori de los coeficientes), así como un pequeño análisis de sesgos.

Tabla 5.12. Medidas de Tendencia central para los estimadores de los coeficientes del modelo.

Distribución a posteriori del Intercepto β_0						
	Errores t - Student Dependientes			Errores t - Student Independientes		
MEDIDAS	M Normal	M S Elíptico Dep	M S Elíptico Ind	M Normal	M S Elíptico Dep	M S Elíptico Ind
Media	2,0537	2,0150	2,0918	2,0575	2,1080	2,1665
Mediana	2,0537	2,0212	2,0951	2,0588	2,1075	2,1689
Moda	2,1788	2,1154	2,1774	2,1012	2,1675	2,2278
Sesgo	positivo	positivo	positivo	positivo	positivo	Positivo
Distribución a posteriori de la Pendiente β_1						
	Errores t - Student Dependientes			Errores t - Student Independientes		
MEDIDAS	M Normal	M S Elíptico Dep	M S Elíptico Ind	M Normal	M S Elíptico Dep	M S Elíptico Ind
Media	0,9240	1,0280	0,9443	0,9257	0,9279	0,7715
Mediana	0,9146	1,0224	0,9435	0,9222	0,9197	0,7749
Moda	0,9487	1,0567	0,9929	0,9706	0,9381	0,8361
Sesgo	negativo	app nulo	negativo	negativo	negativo	Negativo

Desde la *Tabla 5.12.* y la *Figura 5.7.*, se obtienen las siguientes conclusiones:

- Para errores elípticos dependientes, la estimación de la pendiente mediante el modelo respectivo (SED) resulta aproximadamente insesgada, mientras que los otros modelos subestiman este coeficiente (sesgo negativo). La varianza menor se da también para el modelo correspondiente al tipo de errores y la varianza mayor se da para el modelo PN.
- Para errores elípticos dependientes todas las estimaciones del intercepto presentan un leve sesgo positivo (el cual es menor en el caso del modelo respectivo SED) y las varianza del modelo respectivo y del modelo PN no difieren mucho, resultando más dispersas las estimaciones provenientes del análisis mediante el modelo SEI.
- Para errores elípticos independientes, los modelos PN y SED subestiman levemente la pendiente, mientras que con el modelo respectivo (SEI) se presenta un sesgo negativo mayor. La varianza mayor se da para el modelo PN, mientras que las varianzas de los modelos semiparamétricos elípticos no presentan diferencias notables.
- Para errores elípticos independientes, los modelos PN y SED sobreestiman levemente al intercepto y, contradictoriamente, se presenta el sesgo positivo mayor en el caso del modelo respectivo y el menor para el modelo PN. La mayor varianza se da para el modelo PN, mientras que las varianzas de los modelos semiparamétricos elípticos no presentan diferencias notables.

En consecuencia, el modelo semiparamétrico elíptico dependiente (SED) presenta un comportamiento mucho más adecuado que el modelo semiparamétrico elíptico independiente (SEI) para la inferencia relativa a los coeficientes del modelo.

CONCLUSIONES

Este Capítulo aborda el análisis bayesiano semiparamétrico del modelo de regresión con errores de medición MEM o con errores en las variables.

Con el fin de justificar el enfoque bayesiano utilizado, se obtuvo la función de verosimilitud asociada al problema y las ecuaciones a resolver para la obtención de estimadores máximo verosímiles de los parámetros, resultando las expresiones y ecuaciones intratables analíticamente.

Con el objeto de dar solución conveniente al problema, se implementa el muestreo de Gibbs, para lo cual se deducen todas las distribuciones condicionales completas de los parámetros y, en base a simulación, se obtienen los estimadores de Bayes bajo pérdida cuadrática.

El capítulo aborda dos casos, de acuerdo a si el modelo contempla errores con distribución elíptica representable independiente o dependiente.

Se resuelven dos problemas presentes en la literatura mediante la metodología implementada, “*Retornos COPEC e IPSA*” (Arellano – Valle, Galea e Iglesias, 2001) y “*Resistencia de Concreto*” (Wellman y Gunst, 1991). También se realizó una simulación con el fin de determinar como inciden en la precisión de las estimaciones las magnitudes de los parámetros de precisión para el error de medición y para el error de la regresión ϕ_u y ϕ_e , respectivamente, y la relación entre ellas.

Algunos comentarios se desprenden del análisis de los datos simulados. En primer lugar, el modelo semiparamétrico elíptico dependiente resulta más preciso que el independiente; en segundo lugar, si la distribución de los errores es t – Student dependiente, el modelo respectivo (semiparamétrico elíptico dependiente) estima en forma semejante al modelo paramétrico normal, mientras que si la distribución de los errores es t – Student independiente, los dos modelos semiparamétricos se comportan en forma semejante en lo relativo a la estimación.

Algunas limitaciones de los resultados obtenidos se relacionan con que estos se basan en el modelo normal – gama; se limitan, en general al modelo de regresión lineal simple; los coeficientes se asumen independientes entre sí e independientes de los otros parámetros, y los errores son elípticos. Por esto se propone como extensiones considerar prioris fuera del modelo normal – gama, extender los resultados al modelo de regresión lineal múltiple, explorar nuevas posibilidades en las estructuras de dependencia en los parámetros de interés y considerar errores skew – normal y skew – elípticos.

CAPÍTULO 6:

ANÁLISIS SEMIPARAMÉTRICO DEL MODELO DE CALIBRACIÓN COMPARATIVA

Los modelos de calibración comparativa históricamente se han utilizado para comparar varios instrumentos, cuya función es obtener mediciones de una misma cantidad desconocida. Estos modelos tienen aplicación en diversas áreas, es por esta razón que en la literatura encontramos variadas aplicaciones. En efecto, Grubbs (1948, 1973), Barnett (1969) y Carter (1981) comparan instrumentos de laboratorio, mientras que Kelly (1984, 1985), Chipkevitch et al. (1996) y Lu et al. (1997) realizan aplicaciones al área médica. Por otra parte, Fuller (1987) considera aplicaciones a la agricultura y Dunn (1992) aborda aplicaciones a la psicología y educación.

En general en la literatura, al abordar el problema de calibración comparativa, se asume que los términos de error son normales. Bolfarine y Galea – Rojas (1995, 1996) relajan el supuesto de normalidad abordando el modelo de calibración comparativa estructural con errores t – student y elípticos. Posteriormente, Galea et al. (1997) desarrollan inferencias para el modelo de calibración comparativa con errores elípticos y recientemente Galea et al. (2002a, 2002b) aplican el método de influencia local a modelos de calibración comparativa estructurales y funcionales, considerando respectivamente errores normales y elípticos. Mpor otra parte, Bolfarine, Cabral y Paula (2002) desarrollan test de distancia para estos modelos. Todas estas referencias han considerado un enfoque clásico en la realización de la inferencia y análisis estadístico.

Ahora bien, el modelo de calibración comparativa se puede expresar como:

$$(6.1) \quad y_{ij} = \beta_{0j} + \beta_{1j} x_i + \varepsilon_{ij}, \quad i = 1, \dots, n; j = 1, \dots, p.$$
$$e \quad y_{i0} = x_i + \varepsilon_{i0}, \quad i = 1, \dots, n,$$

donde existe una relación dada por un modelo de regresión lineal entre la variable predictora no observable x_i y p variables respuesta observables y_{ij} .

Los coeficientes del modelo de regresión lineal son β_0 y β_1 y los errores aleatorios asociados a este modelo se denotan por ε_{ij} . Al igual que para el caso del MEM en el capítulo 5, la variable x_i no es observable, pero se observa con un error aditivo ε_{i0} , lo cual da origen a la variable y_{i0} .

Tal como se expuso en el capítulo 2, el Modelo de Calibración Comparativa puede ser visto como una extensión de dos Modelos. En primer lugar, generaliza el Modelo con Errores en las Variables, para el cual las predictoras observadas z_i se denotan por y_{i0} , y se cuenta con p diferentes variables observadas y_{ij} para cada x_i . En segundo lugar, generaliza al modelo de Calibración, para el cual se cuenta con un instrumento exacto, cuya medición y_{i0} tiene un error aditivo, y con p instrumentos inexactos en lugar de uno; es así como p mediciones inexactas y_{ij} están relacionadas linealmente con x_i , más un error aditivo, y x_i es el verdadero valor de la medición.

El objetivo fundamental de este capítulo es realizar un análisis bayesiano semiparamétrico del modelo de calibración comparativo, donde la distribución a priori para el parámetro de dispersión se selecciona de acuerdo con un cierto proceso Dirichlet.

El capítulo se divide en tres secciones. Las primeras dos secciones desarrollan una metodología bayesiana semiparamétrica para el problema de calibración comparativa donde los términos de error siguen una distribución elíptica representable independiente y dependiente, respectivamente.

Para ambos casos se requiere la implementación de métodos de simulación MCMC, debido a la inafectabilidad de una resolución analítica del problema; específicamente, se utiliza en este trabajo el muestreo de Gibbs, explotando, como en capítulos previos, la relación existente entre los modelos normal y elíptico representable.

Con el fin de ilustrar e introducir los principales resultados, ambas secciones abordan en principio el caso paramétrico, el cual considera errores mezcla finita de normales o distribuidos como una normal contaminada, con vector de pesos de la mezcla distribuidos Dirichlet. El caso semiparamétrico se obtiene como una extensión natural, donde no se utiliza una mezcla específica, puesto que los parámetros que la definen siguen un proceso Dirichlet.

Las distribuciones a priori utilizadas para los desarrollos en ambas secciones se han escogido semejantes a las presentes en la literatura, con el fin de permitir la comparación con resultados de dichos autores..

La tercera sección y final ilustra los resultados de las secciones previas mediante conjuntos de datos presentes en la literatura y mediante simulación.

6.1 SOLUCIÓN BAYESIANA SEMIPARAMÉTRICA DEL MODELO DE CALIBRACIÓN COMPARATIVO CON ERRORES ELÍPTICOS REPRESENTABLES INDEPENDIENTES

En esta sección abordamos el problema de calibración comparativa, considerando que los términos de error son independientes e idénticamente distribuidos elípticos, con función generadora de densidad h representable:

$$(6.2) \quad y_{ij} | \beta, \phi_j, \mathbf{x} \sim \text{ind. } EL_1(\beta_{0j} + \beta_{1j} x_i, \phi_j^{-1}, h_j) \text{ y } h_j(u) = \int_0^\infty (2\pi\omega)^{-1/2} e^{-u/2\omega} dG_j(\omega),$$

$$e \quad y_{i0} | \phi_0, \mathbf{x} \sim \text{ind. } EL_1(x_i, \phi_0^{-1}, h_0) \text{ y } h_0(u) = \int_0^\infty (2\pi\omega)^{-1/2} e^{-u/2\omega} dG_0(\omega),$$

con $h_j \geq 0$ y donde G_j es función de distribución acumulada tal que $G_j(0) = 0, j = 0, \dots, n$.

Ahora bien, las funciones G_j se denominan funciones mezcla o mezcladoras. Los parámetros $(\phi, \beta, \mathbf{x})$ y $\{G_j\}_{j=0, \dots, n}$ son desconocidos, y el hecho de que los G_j sean infinito dimensionales es lo que imprime el carácter semiparamétrico a este enfoque, y es también lo que distingue los resultados que se presentan a continuación de los presentes en la literatura.

En forma análoga a lo desarrollado en capítulos anteriores, la *Proposición 2.1*. permite enfrentar el modelo de calibración comparativa semiparamétrico con errores elípticos en base al modelo paramétrico con errores normales, mediante la inclusión de variables latentes.

El modelo jerárquico que nos interesa viene dado por:

$$(6.3) \quad \begin{aligned} \varepsilon_{ij} | \phi_j &\sim \text{cond. ind. } EL_1(0, \phi_j^{-1}, h_j), & i = 1, \dots, n; j = 0, \dots, p, \\ \phi_j &\sim \text{Gamma}(r_{2j}, \lambda_{2j}), & j = 0, \dots, p, \\ \beta_{0j} &\sim N(\mathbf{b}_{0j}, \mathbf{V}_{0j}), & j = 1, \dots, p, \\ \beta_{1j} &\sim N(\mathbf{b}_{1j}, \mathbf{V}_{1j}), & j = 1, \dots, p \\ \text{y} \quad x_i &\sim \text{i.i.d. } N(\mu_x, \sigma_x^2), & i = 1, \dots, n. \end{aligned}$$

El análisis se basa en la equivalencia entre este modelo y el modelo paramétrico normal independiente, obtenido al reemplazar la primera etapa del modelo anterior por las siguientes dos:

$$(6.4) \quad \begin{aligned} \varepsilon_{ij} | \omega_{ij}, \phi_j &\sim \text{cond. ind. } N(0, \phi_j^{-1} \cdot \omega_{ij}), \\ y &\quad \omega_{ij} \quad \sim G_j, \end{aligned}$$

con $\omega_{ij} \perp\!\!\!\perp \phi_j, \mathbf{x}$, para $i \in \{0, \dots, n\}$ y $j \in \{1, \dots, p\}$.

La especificación (6.3) no permite explicitar las distribuciones condicionales completas para el modelo, mientras que la especificación (6.4) si lo permite, por tratarse la verosimilitud de una productoria de normales; las expresiones, por tanto, se facilitan en dicho caso. Esto permite la deducción de las distribuciones condicionales completas que, en base al lema de Blackwell y McQueen (1973) y los resultados de Escobar y West (1998), admite la utilización de muestreo de Gibbs para hallar los estimadores de Bayes de los parámetros.

La distribución a posteriori de los parámetros de interés, por supuesto, son las mismas en ambas formulaciones del modelo, pero las expresiones respectivas son analíticamente imposibles de obtener. En efecto, la función de verosimilitud asociada al modelo (6.3) es la siguiente:

$$L(\beta, \mathbf{x}, \phi | \mathbf{y}, G_1, \dots, G_p) = \left[\prod_{j=0}^p \phi_j \right] (2\pi)^{-(p+1)n} \cdot \left[\prod_{j=0}^p \prod_{i=1}^n \int_0^\infty \omega_{ij}^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi_j \cdot (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2 / \omega_{ij}) dG(\omega_{ij}) \right],$$

donde $\beta_{00} = 0$ y $\beta_{10} = 1$; mientras que para el modelo (6.4), con la misma definición de β_{00} y β_{10} , se obtiene una función de verosimilitud aumentada que coincide con la verosimilitud de un modelo de calibración comparativa normal heteroscedástico:

$$L(\beta, \mathbf{x}, \phi, \omega | \mathbf{y}, G_1, \dots, G_p) = \left[\prod_{j=0}^p \phi_j \right] (2\pi)^{-(p+1)n} \cdot \left[\prod_{j=0}^p \prod_{i=1}^n \omega_{ij}^{-\frac{1}{2}} \right] \cdot \exp[-\frac{1}{2} \sum_{j=0}^p \sum_{i=1}^n \frac{\phi_j}{\omega_{ij}} (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2].$$

Bajo el enfoque clásico no resulta plausible realizar la estimación máximo verosímil con estas expresiones. Esto se debe a que para la primera especificación, aunque se consideraran distribuciones G_j particulares, la expresión resultaría muy complicada de resolver, y para la última expresión la estimación no es factible sin contar con información acerca de las funciones de distribución G_j .

Supongamos, para la primera expresión, que las funciones de distribución G_j tienen asociadas funciones de densidad g_j y que la naturaleza de esas funciones permite intercambiar la integral respecto de ω_{ij} y la derivada respecto de β, \mathbf{x} y ϕ_j . Entonces los estimadores máximo verosímiles deben satisfacer que:

$$\frac{\phi_j}{n} \cdot \sum_{i=1}^n e_{ij}^2 \cdot k(e_{ij}, \phi_j) = 1, j = 0, \dots, p; \quad \sum_{j=1}^p e_{ij} \cdot \beta_{1j} \cdot k(e_{ij}, \phi_j) = (x_i - y_{i0}) \cdot k(e_{i0}, \phi_0), i = 1, \dots, n;$$

$$\sum_{i=1}^n e_{ij} \cdot k(e_{ij}, \phi_j) = 0 \quad \text{y} \quad \sum_{i=1}^n e_{ij} \cdot x_i \cdot k(e_{ij}, \phi_j) = 0, j = 1, \dots, p,$$

$$\text{donde } k(e_{ij}, \phi_j) = \left[\int_0^\infty \omega_{ij}^{-\frac{1}{2}} \exp(-\frac{1}{2} \phi_j \cdot e_{ij}^2 / \omega_{ij}) \cdot g(\omega_{ij}) \cdot d\omega_{ij} \right] / \left[\int_0^\infty \omega_{ij}^{-\frac{3}{2}} \exp(-\frac{1}{2} \phi_j \cdot e_{ij}^2 / \omega_{ij}) \cdot g(\omega_{ij}) \cdot d\omega_{ij} \right] \text{ y}$$

$$\text{con } e_{ij} = (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i).$$

Ahora bien, para la segunda expresión, bajo los mismos supuestos, los estimadores máximo verosímiles deben satisfacer que:

$$\omega_{ij} = \phi_j \cdot e_{ij}^2, i = 1, \dots, n, j = 0, \dots, p; (x_i - y_{i0}) \cdot \sum_{j=1}^p \frac{\beta_{1j}}{e_{ij}} = 1, i = 1, \dots, n; \sum_{i=1}^n \frac{1}{e_{ij}} = 0 \text{ y } \sum_{i=1}^n \frac{x_i}{e_{ij}} = 0; j = 1, \dots, p.$$

Es claro en ambos casos que no existe solución única para este sistema, ya que hay menos ecuaciones que parámetros a estimar. Además, los ϕ_j no son estimables.

Por otra parte, para distribuciones G_j particulares bien escogidas como gamma inversa de parámetros r_j y λ_j , respectivamente, la expresión $k(e_{ij}, \phi_j)$ está dada por:

$$k(e_{ij}, \phi_j) = \frac{2 \cdot r_j + 1}{\phi_j \cdot e_{ij}^2} \cdot \left(\frac{2 \cdot \lambda_j}{\phi_j \cdot e_{ij}^2} + 1 \right)^{-1},$$

de donde la estimación máximo verosímil requiere determinar los valores de los parámetros que satisfagan las siguientes ecuaciones:

$$\frac{2r_j + 1}{n} \sum_{i=1}^n \left(1 + \frac{2\lambda_j}{\phi_j \cdot e_{ij}^2} \right)^{-1} = 1; j = 0, \dots, p; \sum_{i=1}^n \frac{1}{e_{ij}} \left(1 + \frac{2\lambda_j}{\phi_j \cdot e_{ij}^2} \right)^{-1} = 0 \text{ y } \sum_{i=1}^n \frac{x_i}{e_{ij}} \left(1 + \frac{2\lambda_j}{\phi_j \cdot e_{ij}^2} \right)^{-1} = 0, j = 1, \dots, p;$$

$$\sum_{j=1}^p \frac{(2 \cdot r_j + 1) \cdot \beta_{1j}}{\phi_j \cdot e_{ij}} \cdot \left(1 + \frac{2\lambda_j}{\phi_j \cdot e_{ij}^2} \right)^{-1} = \frac{(2 \cdot r_0 + 1)}{\phi_0 \cdot (x_i - y_{i0})} \cdot \left(1 + \frac{2\lambda_0}{\phi_0 \cdot e_{i0}^2} \right)^{-1}, i = 1, \dots, n.$$

Es claro que en el caso del problema de calibración comparativa bajo el modelo de regresión lineal simple, $e_i = (y_i - \beta_{0j} - \beta_{1j} \cdot x_i)$, pero bajo el modelo de regresión lineal múltiple, $e_i = (y_i - \beta_{0j} - \beta_{1j} \cdot x_{i1} - \dots - \beta_{pj} \cdot x_{ip})$. Salta a la vista que la resolución analítica del sistema planteado arriba no es posible, por lo cual bajo el enfoque clásico se requiere una resolución por métodos numéricos. Bolfarine y Galea – Rojas (1995) utilizan algoritmo EM para realizar la estimación máximo verosímil de los parámetros para un modelo de calibración comparativa estructural.

La función de distribución particular utilizada por estos autores es $G_j \sim \text{Gama Inversa } (v/2, v/2)$, obteniéndose así el modelo t – Student.

Bajo el enfoque bayesiano, por otra parte, la distribución a posteriori asociada al modelo de interés bajo la especificación (6.2), aún considerando las funciones de distribución G_j conocidas (en caso contrario las expresiones son mucho más complicadas), es de la forma siguiente:

$$\pi(\beta, \mathbf{x}, \phi | \mathbf{y}, G_1, \dots, G_p) \propto L(\beta, \mathbf{x}, \phi | \mathbf{y}, G_1, \dots, G_p) \cdot \pi(\beta) \cdot \pi(\mathbf{x}) \cdot \pi(\phi),$$

es decir, la distribución a posteriori para los parámetros se puede expresar como:

$$\pi(\beta, \mathbf{x}, \phi | \mathbf{y}, G_1, \dots, G_p) \propto \prod_{j=0}^p \phi_j^{r_j - 1/2} \cdot \prod_{i=1}^n \prod_{j=0}^p \int_0^\infty [(\omega_{ij})^{-1/2} \cdot \exp\{-1/2 \phi_j \cdot (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2 / \omega_{ij}\}] dG_j(\omega_{ij})$$

$$\cdot \exp\left\{-\frac{1}{2} \left(\sum_{j=1}^p \left[\mathbf{1}_{0j}^{-1} \cdot \left(\mathbf{1}_{0j} - b_{0j} \right) \right] + v_{1j}^{-1} \cdot \left[\mathbf{1}_{1j} - b_{1j} \right] \right) \sum_{i=1}^n \sigma_X^{-2} \cdot \left(\mathbf{1}_i - \mu_X \right) \right\} - \sum_{j=0}^p \lambda_j \cdot \phi_j \},$$

donde la constante de proporcionalidad es prácticamente imposible de evaluar explícitamente.

Incluso si las funciones de distribución G_j fueran distribuciones gamma inversa de parámetros r_j y λ_j , esta expresión resultaría igualmente complicada, ya que la expresión para la productoria de integrales respecto de ω_{ij} es proporcional a:

$$\prod_{i=1}^n \prod_{j=0}^p \int_0^{\infty} [(\omega_{ij})^{-1/2} \cdot \exp\{-1/2 \phi_j \cdot (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2 / \omega_{ij}\}] dG_j(\omega_{ij}) \propto \prod_{i=1}^n \prod_{j=0}^p [1 + \phi_j (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2 / 2\lambda_j]^{-1/2(r_j+1)},$$

de donde el estimador de Bayes (la media de la distribución resultante al marginalizar) no es obtenible analíticamente.

Para la especificación (6.4), la expresión resultante es semejante y conduce a los mismos hallazgos.

Como nuestro enfoque es semiparamétrico, no nos amarramos a una distribución G_j particular, es así como los comentarios anteriores justifican nuestra decisión de optar por utilizar MCMC en la implementación de muestreo de Gibbs. Esto resulta factible puesto que las distribuciones condicionales completas requeridas para dicha simulación son obtenibles analíticamente. Con esto en mente, las siguientes dos proposiciones abordan el modelo (6.3) o (6.4), contemplando sendas elecciones para las distribuciones a priori de las funciones de mezcla G_j .

Proposición 6.1:

Bajo las especificaciones (6.3) y (6.4), se escoge para los parámetros de dispersión una distribución a priori discreta $p(\omega_{ij} | \mathbf{p}) = \sum_{l=1}^{k_j+1} p_{lj} \cdot \delta_{\tau_j}(\omega_{ij})$, $j = 0, \dots, p$, cuyos pesos, a su vez, tienen una distribución a priori Dirichlet $(p_{1j}, \dots, p_{k_j j}) \sim \mathcal{D}(\alpha_{1j}, \dots, \alpha_{k_j j+1})$. En este caso la distribución a posteriori de los parámetros está dada por la siguiente expresión analíticamente intratable:

$$\pi(\boldsymbol{\beta}, \mathbf{x}, \boldsymbol{\phi} | \mathbf{y}, G_1, \dots, G_p) \propto \prod_{j=0}^p \phi_j^{r_j-1/2} \cdot \prod_{i=1}^n \prod_{j=0}^p \sum_{l=1}^{k_j+1} p_{lj} [(\tau_j)^{-1/2} \cdot \exp\{-1/2 \phi_j \cdot (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2 / \tau_j\}] \cdot \delta_{\tau_j}(\omega_{ij}) \\ \cdot \exp\left\{-\frac{1}{2} \left(\sum_{j=1}^p \left[\mathbf{1}_{oj} \cdot \left(\sum_{i=1}^n \mathbf{1}_{ij} \cdot (y_{ij} - b_{0j}) \right) + \mathbf{v}_{1j}^{-1} \cdot \left(\sum_{i=1}^n \mathbf{1}_{ij} \cdot (y_{ij} - b_{1j} \cdot x_i) \right) \right] \sum_{i=1}^n \sigma_X^{-2} \cdot \left(\sum_{i=1}^n \mathbf{1}_{ij} \cdot (y_{ij} - \mu_X) \right) \right) - \sum_{j=0}^p \lambda_j \cdot \phi_j \right\}.$$

Ahora bien, las distribuciones condicionales completas requeridas en la implementación del muestreo de Gibbs son las siguientes:

$$(6.5) \quad \omega_{ij} | \{ \omega_{\cdot(i,j)}, \phi_j, \mathbf{x}, \boldsymbol{\beta}_j, y_{ij} \} \sim g_j(\omega_{ij}) \propto \sum_{l=1}^{k_j+1} \exp\left\{-\frac{(y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2}{2\omega_{ij}}\right\} \cdot \frac{\phi_j \cdot p_{lj} \cdot \delta_{\tau_j}(\omega_{ij})}{\sqrt{\omega_{ij}^n}},$$

$$(6.6) \quad \phi_j | \mathbf{W}, \mathbf{x}, \boldsymbol{\beta}_j, y_{ij} \sim \text{ind. Gamma}(r = r_{2j} + 1/2 n, \lambda = \lambda_{2j} + \sum_{i=0}^n \frac{(y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2}{2 \cdot \omega_{ij}}),$$

$$(6.7) \quad \beta_{0j} | \mathbf{W}, \phi_j, x_i, \beta_{1j}, y_{ij} \sim \text{ind. N}(m_{0j}, S_{0j}^2),$$

$$\text{donde } m_{0j} = \phi_j \cdot \left[\sum_{i=1}^n \frac{(y_{ij} - \beta_{1j} x_i)}{\omega_{ij}} + \frac{b_{0j}}{v_{0j}} \right] \text{ y } S_{0j}^2 = \left[\phi_j \cdot \left[\sum_{i=1}^n \frac{1}{\omega_{ij}} + \frac{1}{v_{0j}} \right] \right]^{-1},$$

$$(6.8) \quad \beta_{1j} | W, z_i, \phi_j, x_i, \beta_{0j}, y_{ij} \sim \text{ind. } N(m_{1j}, S_{1j}^2),$$

$$\text{donde } m_{1j} = S_{1j}^2 \cdot \left[\phi_j \cdot \sum_{i=1}^n \frac{x_i \cdot (y_{ij} - \beta_{0j})}{\omega_{ij}} + \frac{b_{1j}}{v_{1j}} \right] \text{ y } S_{1j}^2 = \left[\phi_j \cdot \sum_{i=1}^n \frac{x_i^2}{\omega_{ij}} + \frac{1}{v_{1j}} \right]^{-1},$$

$$(6.9) \quad x_i | W, \phi_j, \beta_j, y_{ij} \sim \text{i.i.d. } N(m_{xi}, S_{xi}^2),$$

$$\text{con } m_{xi} = S_{xi}^2 \cdot \left[\sum_{j=0}^p \frac{\phi_j \cdot \beta_{1j} \cdot (y_{ij} - \beta_{0j})}{\omega_{ij}} + \frac{\mu_x}{\sigma_x^2} \right] \text{ y } S_{xi}^2 = \left[\sum_{j=0}^p \frac{\phi_j \cdot \beta_{1j}^2}{\omega_{ij}} + \frac{1}{\sigma_x^2} \right]^{-1}$$

y donde W es la matriz de componentes ω_{ij} .

Note que, dado ω , los parámetros β_{0j} , β_{1j} y ϕ_j son independientes de G_j .

La implementación para los resultados anteriores se puede realizar confeccionando un programa para muestrear desde las distribuciones condicionales completas o, en este caso, utilizando Winbugs.

Observación 6.1

El caso abordado en la *Proposición 6.1*. considera una mezcla discreta en parámetro escala de normales. Ahora bien, si $p_l = 1$ y $p_m = 0 \forall l \neq m$, es decir, si ω_{ij} tiene una distribución degenerada en un punto, no se produce una mezcla de normales, sino que se recupera el modelo de calibración comparativa lineal normal ($h(u) = (2\pi\tau)^{-1/2} e^{-u^2/2\tau}$). Las distribuciones condicionales completas de los parámetros β , x y ϕ se mantienen intactas, pero p y ω_{ij} pasan a ser constantes.

A continuación se presenta la solución bayesiana semiparamétrica del problema de calibración comparativa bajo el modelo elíptico representable independiente, que corresponde al principal resultado de esta sección.

Proposición 6.2:

Bajo las especificaciones (6.3) y (6.4), se asume que las distribuciones a priori para los parámetros de dispersión son procesos Dirichlet G_j centrados en distribuciones gamma inversa G_{0j} , es decir $\omega_{ij} | G_j \sim \text{ind. } G_j$, $G_j \sim \text{ind. } \mathcal{D}(c_j G_{0j})$, y $G_{0j} \sim \text{ind. } I G(r_{1j}, \lambda_{1j})$, $j = 0, \dots, p$.

Entonces, la distribución a posteriori de los parámetros es analíticamente inmanejable, por lo cual los estimadores de Bayes no son obtenibles; sin embargo, no es difícil obtener las distribuciones condicionales completas, que coinciden con las obtenidas para la *Proposición 6.1*. en lo relativo a los parámetros ϕ_j , β_{0j} , β_{1j} y x_i , (6.6) a (6.9) y difieren en lo relativo al parámetro ω_{ij} , cuya distribución corresponde a la siguiente expresión:

$$(6.10) \quad \omega_{ij} | \omega_{k \neq i}, G_j, \phi_j, x_i, \beta, y_{ij} \propto q_{0i}^{(i)} \cdot f_{G_{jb}^{(i)}}(\omega_{ij}) + \sum_{k \neq i} q_{ki}^{(j)} \cdot \delta_{\omega_{kj}}(\omega_{ij}),$$

donde $f_{G_{jb}^{(i)}}$ es la función de densidad de probabilidad Gamma Inversa con parámetros $r = r_{1j} + 1/2$

$$\text{y } \lambda = \lambda_{1j} + 1/2 (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2 \cdot \phi_j,$$

$q_{0i}^{(j)}$ es la función de densidad $T(\beta_{0j} + \beta_{1j} \cdot x_i; \lambda_{1j} / \phi_j; 2 r_{1j})$ evaluada en y_{ij} y

$q_{ki}^{(j)}$ es la función de densidad $N(\beta_{0j} + \beta_{1j} \cdot x_i; \omega_{jk} / \phi_j)$ evaluada en y_{ij} .

con $\beta_{00} = 0$ y $\beta_{10} = 1$, ($i = 1, \dots, n$; $j = 0, \dots, p$)

y con \mathbf{y}_i el vector de componentes y_{ij} ($j = 0, \dots, p$) cuya distribución condicional completa corresponde a:

$$(6.11) \quad \mathbf{y}_i | \omega_{ij}, G_j, \phi_j, \beta, x_i \sim N_{p+1}(\boldsymbol{\mu}_{0i}, \Sigma_{0i}),$$

con $\boldsymbol{\mu}_{0i} = (0, \beta_{01}, \dots, \beta_{0p})^t + (1, \beta_{11}, \dots, \beta_{1p})^t \cdot x_i$ y $\Sigma_{0i} = \text{diag}(\omega_{i0} \cdot \phi_0^{-1}, \dots, \omega_{ip} \cdot \phi_p^{-1})$.

Ejemplo 6.1:

Si las funciones generadoras de densidad h_j son conocidas, se tiene que G_j es degenerada en G_{0j} y por lo tanto $G_j \sim \text{ind. } I G(r_{1j}, \lambda_{1j})$, $j = 1, \dots, p$. En este caso, las distribuciones condicionales completas de los parámetros ϕ_j , β_{0j} , β_{1j} y x_i , están dadas por (6.6) a (6.9), respectivamente; mientras que la distribución de ω_{ij} coincide con $G_{jb}^{(i)}$ en (6.10).

6.2 SOLUCIÓN BAYESIANA SEMIPARAMÉTRICA DEL MODELO DE CALIBRACIÓN COMPARATIVO CON ERRORES ELÍPTICOS REPRESENTABLES DEPENDIENTES

En esta sección se aborda el problema de calibración comparativa bajo el modelo de regresión, considerando los términos de error distribuidos elípticos dependientes, con funciones generadoras de densidad h_j representables.

Si se define $\mathbf{y}^{(j)} = (y_{1j}, \dots, y_{nj})^t$ y $\boldsymbol{\beta}^{(j)} = (\beta_{0j}, \beta_{1j})^t$, el modelo al cual se alude es:

$$(6.12) \quad \mathbf{y}^{(j)} | \boldsymbol{\beta}^{(j)}, \phi_j, \mathbf{x} \sim \text{ind. } E_{I_n}(\beta_{0j} \cdot \mathbf{1}_n + \beta_{1j} \cdot \mathbf{x}, \phi_j^{-1} \cdot I_n, h_j) \text{ y } h_j(u) = \int_0^\infty (2\pi\omega)^{-1/2} e^{-u/2\omega} dG_j(\omega), j = 1, \dots, p$$

$$\mathbf{y}^{(0)} | \boldsymbol{\beta}^{(0)}, \phi_0, \mathbf{x} \sim \text{ind. } E_{I_n}(\mathbf{x}, \phi_0^{-1} \cdot I_n, h_0) \text{ y } h_0(u) = \int_0^\infty (2\pi\omega)^{-1/2} e^{-u/2\omega} dG_0(\omega),$$

con $h_j \geq 0$ y donde G_j es función de distribución acumulada f.d.a. tal que $G_j(0) = 0$, $j = 0, \dots, p$.

Ahora bien, el modelo jerárquico con errores elípticos dependientes que interesa abordar se expresa como:

$$\begin{aligned}
 \boldsymbol{\varepsilon}_j \mid \phi_j &\sim \text{cond. ind. } El_n(\mathbf{0}, \phi_j^{-1} \cdot \mathbf{I}_n, h_j), j = 0, \dots, p, \\
 \phi_j &\sim \text{Gamma}(r_{2j}, \lambda_{2j}), \quad j = 0, \dots, p, \\
 \beta_{0j} &\sim N(\mathbf{b}_{0j}, \mathbf{V}_{0j}), \quad j = 1, \dots, p, \\
 \beta_{1j} &\sim N(\mathbf{b}_{1j}, \mathbf{V}_{1j}), \quad j = 1, \dots, p \\
 y \quad x_i &\sim \text{i.i.d. } N(\mu_x, \sigma_x^2), \quad i = 1, \dots, n
 \end{aligned}
 \tag{6.13}$$

De forma análoga a lo desarrollado en la sección anterior, la *Proposición 2.1 a.*, plantea la equivalencia entre este modelo y el modelo normal dependiente, previa inclusión de variables latentes. El análisis, entonces, se basa en el modelo equivalente, que reemplaza la primera etapa del modelo (6.13) por las siguientes dos:

$$\begin{aligned}
 \boldsymbol{\varepsilon}_j \mid \omega_j, \phi_j &\sim \text{cond. ind. } N_n(\mathbf{0}, \phi_j^{-1} \cdot \omega_j \cdot \mathbf{I}_n) \\
 y \quad \omega_j &\sim G_j, \\
 \text{con } \omega_j &\perp\!\!\!\perp \phi_j, \mathbf{x}, j \in \{0, \dots, p\}.
 \end{aligned}
 \tag{6.14}$$

Al igual que en la sección precedente y debido a las dificultades analíticas inherentes al problema, el objetivo de hallar estimadores de Bayes de los parámetros de interés se transforma en la búsqueda de las distribuciones condicionales completas que permitan la implementación de muestreo de Gibbs.

Entonces la función de verosimilitud asociada al modelo de interés, aún considerando funciones de distribución G_j conocidas y bajo la especificación (6.12), es de la forma siguiente:

$$L(\boldsymbol{\beta}, \mathbf{x}, \boldsymbol{\phi} \mid \mathbf{y}, G_1, \dots, G_p) = \left[\left(\prod_{j=0}^p \phi_j \right) (2\pi)^{-(p+1)n} \right] \cdot \left[\prod_{j=0}^p \int_0^\infty \prod_{i=1}^n \omega_j^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \phi_j \cdot (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2 / \omega_j\right) dG(\omega_j) \right],$$

mientras que para el modelo basado en (6.14), la función de verosimilitud aumentada que se presenta a continuación, coincide con la del modelo normal homoscedástico:

$$L(\boldsymbol{\beta}, \mathbf{x}, \boldsymbol{\phi}, \boldsymbol{\omega} \mid \mathbf{y}, G_1, \dots, G_p) = \left[\left(\prod_{j=0}^p \phi_j \right) (2\pi)^{-(p+1)n} \right] \cdot \left[\prod_{j=0}^p \omega_j^{-\frac{n}{2}} \right] \cdot \exp\left[-\frac{1}{2} \sum_{j=0}^p \frac{\phi_j}{\omega_j} \sum_{i=1}^n (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2\right].$$

Es claro que, si estamos interesados en la función de verosimilitud, resulta preferible trabajar con la expresión relativa a la formulación (6.14).

Un caso particular no abordado aquí es aquel para el cual $G_j = G$ para todo j .

Para la primera formulación, las funciones de distribución G_j son desconocidas, por lo cual la resolución analítica no es posible, mientras que para la segunda formulación, las distribuciones G_j se asumen dadas y no se presentan explícitamente en la función de verosimilitud aumentada, pero en realidad son desconocidas y la resolución depende de su forma, puesto que las variables instrumentales ω_j que surgen de dichas distribuciones requieren ser estimadas.

En efecto, la maximización de la función de verosimilitud requiere la resolución de un sistema de ecuaciones análogo al del modelo independiente. Para la formulación basada en el modelo (6.13) las ecuaciones coinciden con las obtenidas en la sección anterior, salvo que la función $k(e_{ij}, \phi_j)$ se especifica con ω_j en lugar de ω_{ij} . Ahora bien, para la formulación basada en el modelo (6.14) las ecuaciones que se obtienen son las siguientes:

$$\omega_j = \frac{\phi_j}{n} \cdot \sum_{i=1}^n e_{ij}^2, j = 0, \dots, p; \quad \sum_{j=1}^p \beta_{1j} \cdot e_{ij} = -e_{i0}, i = 1, \dots, n; \quad \sum_{i=1}^n e_{ij} = 0 \quad \text{y} \quad \sum_{i=1}^n x_i \cdot e_{ij} = 0, j = 1, \dots, p,$$

por lo cual la solución del sistema respectivo no es única. Los parámetros ω_j y ϕ_j no son estimables (aunque si lo son los cocientes ω_j / ϕ_j), pero los estimadores de β_{0j} y de β_{1j} no dependen de estos parámetros y son semejantes a los estimadores mínimo cuadráticos usuales.

Desde el punto de vista clásico, y en vista de los problemas analíticos inherentes al modelo, como se mencionó antes, Bolfarine y Galea – Rojas (1995) resuelven el problema de estimación máximo verosímil vía la utilización del algoritmo EM – considerado un modelo estructural.

Ahora bien, las dos proposiciones que a continuación se presentan abordan el modelo (6.13) ó (6.14), considerando elecciones de las distribuciones a priori para las funciones de mezcla G_j análogas a las utilizadas en la precedente sección. Al igual que para el caso independiente, las distribuciones a posteriori de los parámetros de interés son muy difíciles de tratar analíticamente, sin embargo las distribuciones condicionales completas resultan ser simples:

Proposición 6.3:

Bajo las especificaciones (6.13) y escogiendo para el parámetro de dispersión común una distribución a priori discreta $p(\omega_j | \mathbf{p}) = \sum_{l=1}^{k_j+1} p_{lj} \cdot \delta_{\tau_j}(\omega_j)$, $j = 0, \dots, p$; cuyos pesos tienen una distribución a priori Dirichlet $(p_{l1}, \dots, p_{lk_j}) \sim \mathcal{D}(\alpha_{l1}, \dots, \alpha_{lk_j+1})$, se obtiene la siguiente expresión para la distribución a posteriori de los parámetros:

$$\pi(\beta, \mathbf{x}, \phi | \mathbf{y}, G_1, \dots, G_p) \propto \prod_{j=0}^p \phi_j^{r_j-1/2} \cdot \prod_{i=1}^n \sum_{l=1}^{k_j+1} p_{lj} [(\tau_j)^{-1/2} \cdot \exp\{-1/2 \phi_j \cdot (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2 / \tau_j\}] \cdot \delta_{\tau_j}(\omega_j) \cdot \exp\left\{-\frac{1}{2} \left(\sum_{j=1}^p \left[\mathbf{1}_{\omega_j}^{-1} \cdot \left(\mathbf{1}_{\omega_j} - \mathbf{b}_{0j} \right) + v_{1j}^{-1} \cdot \left(\mathbf{1}_{1j} - \mathbf{b}_{1j} \right) \right] \sum_{i=1}^n \sigma_X^{-2} \cdot \left(\mathbf{1}_i - \mu_X \right) \right) - \sum_{j=0}^p \lambda_j \cdot \phi_j \right\}.$$

Los estimadores de Bayes se obtienen a partir de las funciones de distribución a posteriori marginales, las cuales resultan no ser obtenibles analíticamente, no así las distribuciones condicionales completas (requeridas en la implementación del Muestreo de Gibbs).

Las distribuciones condicionales completas de ϕ_j , β_{0j} , β_{1j} y x_i coinciden con las expresiones (6.6) a (6.9), excepto por el hecho de que se reemplaza ω_{ij} por ω_j , es decir:

$$(6.15) \quad \phi_j | \omega, G_j, \mathbf{x}, \beta_j, y_{ij} \sim^{\text{cond.ind.}} \text{Gamma}(r=r_{2j} + \frac{1}{2} \mathbf{n}, \lambda = \lambda_{2j} + \frac{\sum_{i=0}^n (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2}{2 \cdot \omega_j}).$$

$$(6.16) \quad \beta_{0j} | \omega, G_j, \phi_j, x_i, \beta_{1j}, y_{ij} \sim^{\text{cond. ind.}} \text{N}(m_{0j}, S_{0j}^2),$$

$$\text{donde } m_{0j} = S_{0j}^2 \cdot \left[\frac{\phi_j}{\omega_j} \cdot \left(\sum_{i=1}^n y_{ij} - \beta_{1j} \cdot \bar{x} \right) + \frac{b_{0j}}{V_{0j}} \right] \text{ y } S_{0j}^2 = \left[\frac{\mathbf{n} \cdot \phi_j}{\omega_j} + \frac{1}{V_{0j}} \right]^{-1}.$$

$$(6.17) \quad \beta_{1j} | W, G_j, \phi_j, x_i, \beta_{0j}, y_{ij} \sim^{\text{cond. ind.}} \text{N}(m_{1j}, S_{1j}^2),$$

$$\text{donde } m_{1j} = S_{1j}^2 \cdot \left[\frac{\phi_j}{\omega_j} \cdot \left(\sum_{i=1}^n x_i \cdot y_{ij} - \beta_{0j} \cdot \bar{x} \right) + \frac{b_{1j}}{V_{1j}} \right] \text{ y } S_{1j}^2 = \left[\frac{\phi_j}{\omega_j} \cdot \sum_{i=1}^n x_i^2 + \frac{1}{V_{1j}} \right]^{-1}.$$

$$(6.18) \quad x_i | W, G_j, \phi_j, \beta_j, y_{ij} \sim^{\text{c.i.i.d.}} \text{N}(m_{xi}, S_{xi}^2),$$

$$\text{con } m_{xi} = S_{xi}^2 \cdot \left[\sum_{j=0}^p \frac{\phi_j}{\omega_j} \cdot \beta_{1j} \cdot (y_{ij} - \beta_{0j}) + \frac{\mu_x}{\sigma_x^2} \right] \text{ y } S_{xi}^2 = \left[\sum_{j=0}^p \frac{\phi_j}{\omega_j} \cdot \beta_{1j}^2 + \frac{1}{\sigma_x^2} \right]^{-1}.$$

Finalmente, la distribución condicional completa de ω_j es discreta y corresponde a:

$$(6.19) \quad \omega_j | \{\omega_k: k \neq j\}, \phi_j, \mathbf{x}, \beta_j, y_{ij} \sim g_j(\omega_j) \propto \sum_{l=1}^{k_j+1} \exp \left\{ - \sum_{i=1}^{k_j+1} \frac{(y_{ij} - \beta_0 - \beta_1 \cdot x_i)^2}{2\omega_j} \right\} \cdot \frac{\phi_j \cdot p_{lj} \cdot \delta_{\tau_j}(\omega_j)}{\sqrt{\omega_j^n}}.$$

Observación 6.2

Es claro que si $p_i = 1$ y $p_j = 0 \forall j \neq i$, es decir si los ω_j tienen distribución degenerada en un punto, ya no se produce la mezcla de normales, sino que se recupera el modelo normal. En este caso las distribuciones condicionales completas de los parámetros β , ϕ y \mathbf{x} se mantienen intactas, pero \mathbf{p} y ω_j son constantes.

A continuación se presenta el resultado más importante de esta sección, que corresponde a la solución bayesiana semiparamétrica del problema de calibración comparativa bajo el modelo elíptico representable dependiente.

Proposición 6.4:

Bajo las especificaciones (6.13), y considerando que las prioris para los parámetros de dispersión son procesos Dirichlet G_j , es decir, $\omega_j | G_j \sim \text{ind. } G_j$, y que los procesos están centrados en distribuciones gamma inversa G_{0j} , o sea, $G_j \sim \text{ind. } \mathcal{D}(c G_{0j})$, con $G_{0j} \sim \text{ind. } IG(r_j, \lambda_j)$, $j = 1, \dots, p$, la distribución a posteriori para los parámetros del modelo no cambia respecto del caso independiente, lo que motiva la obtención de las distribuciones condicionales completas.

Para los parámetros β_{0j} , β_{1j} , ϕ_j y x_i las distribuciones condicionales completas corresponden a las expresiones (6.15) a (6.18), respectivamente, mientras que para los parámetros ω_j , las distribuciones condicionales completas corresponden a la siguiente expresión:

$$(6.20) \quad \omega_j | \phi_j, \mathbf{x}, \beta_j, y_{ij} \sim IG \left(r_{1j} + \frac{1}{2} n; \lambda_{1j} + \phi_j \cdot \sum_{i=1}^n (y_{ij} - \beta_{0j} - \beta_{1j} \cdot x_i)^2 / 2 \right).$$

Ejemplo 6.2:

Análogamente a lo obtenido para el *Ejemplo 6.1*, h_j conocida lleva a que G_j sea degenerada en G_{0j} , es decir:

$$(6.21) \quad G_j \sim \text{ind. } IG(r_{1j}, \lambda_{1j}), j = 1, \dots, p.$$

Para este caso las distribuciones condicionales completas coinciden exactamente con aquellas derivadas de la *Proposición 6.4*.

6.3 APLICACIONES

En esta sección se ilustran los resultados obtenidos en las secciones previas a través de simulación y conjuntos de datos considerados previamente en la literatura.

Aplicación 6.1: *Medición de volumen testicular de 42 adolescentes mediante 5 métodos. Chipkevitch, Nishimura, Tu & Galea – Rojas (1996).*

Chipkevitch, Nishimura, Tu y Galea – Rojas (1996) analizan las mediciones de volumen testicular derecho de 42 adolescentes, mediante 5 métodos. Los datos se presentan en la *Tabla 6.1.* y los diagramas de dispersión de todos los pares de variables en la *Figura 6.1.* La medición de referencia es el ultrasonido (US). Los otros cuatro métodos son: método gráfico propuesto por los autores (m1), medición dimensional (m2), orquidómetro de Prader (m3), orquidómetro de anillo (m4). El objetivo de los autores es calibrar los instrumentos m1, m2, m3, m4 en base a la medición de ultrasonido (US).

Tabla 6.1: *Mediciones de volumen testicular de 42 adolescentes. El método base es ultra sonido (US), en base al cual se quiere calibrar los otros instrumentos. (Chipkevitch, Nishimura, Tu & Galea – Rojas, 1996)*

US	M1	m2	m3	m4	US	m1	M2	m3	m4	US	m1	m2	m3	m4
5,0	7,5	5,9	5,0	9,0	16,4	15,0	17,3	20,0	20,0	6,7	7,5	7,2	10,0	8,0
5,7	5,0	4,8	6,0	10,0	17,6	15,0	17,3	20,0	22,5	5,3	5,0	5,9	8,0	10,0
7,4	5,0	6,8	9,0	12,0	10,0	7,5	7,9	12,0	12,0	20,0	20,0	16,3	25,0	22,5
2,6	3,5	3,1	4,0	4,0	4,1	3,5	4,4	4,0	6,0	18,8	15,0	16,3	20,0	25,0
5,7	5,0	5,0	6,0	7,0	12,7	10,0	11,4	12,0	12,0	13,9	12,5	12,2	15,0	17,5
6,1	5,0	4,4	7,0	5,0	2,7	3,5	4,1	2,5	6,0	9,4	10,0	10,3	12,0	13,5
6,2	5,0	6,0	8,0	9,0	10,2	10,0	11,1	12,0	13,5	9,1	7,5	10,8	12,0	12,0
10,4	10,0	8,8	10,0	10,0	16,5	10,0	15,3	15,0	15,0	14,1	15,0	13,0	13,5	15,0
9,1	7,5	7,9	10,0	11,0	4,5	3,5	3,9	6,0	7,0	9,3	10,0	8,4	10,0	10,0
14,8	10,0	13,0	12,0	15,0	5,6	5,0	4,5	4,5	6,0	20,9	20,0	22,1	25,0	25,0
16,4	12,5	10,3	17,5	17,5	11,0	7,5	9,7	9,0	11,0	11,5	10,0	10,6	15,0	13,5
9,6	7,5	8,2	10,0	11,0	9,2	10,0	11,3	12,0	13,5	9,7	10,0	9,7	11,0	12,0
15,7	15,0	19,8	20,0	20,0	8,5	7,5	8,8	12,0	12,0	13,7	12,5	11,6	17,5	15,0
3,0	2,0	2,0	3,0	4,0	5,4	5,0	6,1	8,0	8,0	8,9	10,0	8,1	12,0	12,0

Los autores antes mencionados realizan un análisis del problema de calibración comparativa para el conjunto de datos en la *Tabla 6.1.*, considerando un modelo estructural con errores normales y utilizando las mediciones de volumen testicular al cubo. Como se aprecia en la *Figura 6.1.* esta transformación resuelve el problema de heterocedasticidad presentes en los modelos individuales de regresión, donde la predictora es la medición alternativa y la variable dependiente es el ultrasonido. Por otra parte, también resuelven el problema de normalidad. Anecdóticamente podemos mencionar que la prueba de normalidad de

Shapiro – Wilks concluye que, en general, las mediciones de volumen testicular difieren significativamente de la normalidad ($p < 0,05$), no así las mediciones cúbicas ($p > 0,15$).

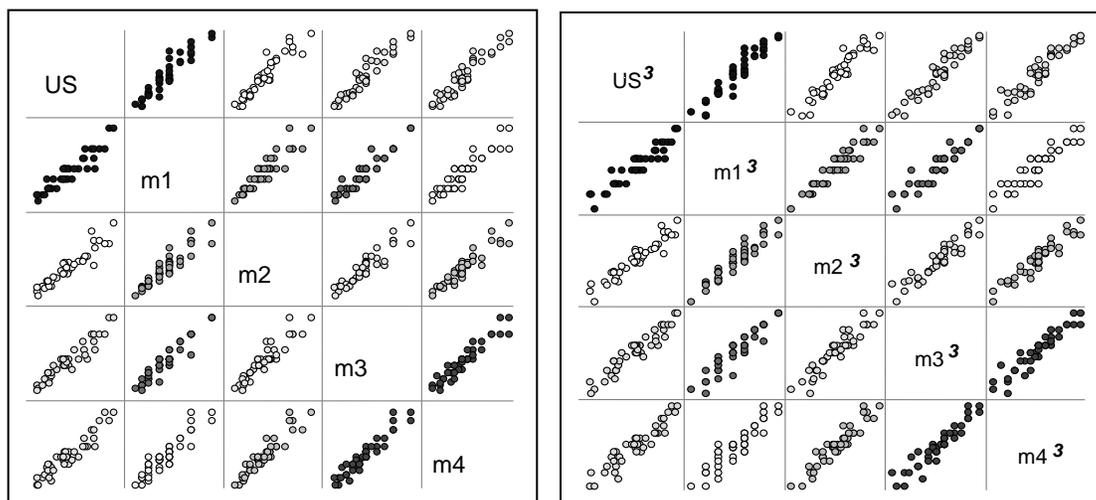


Figura 6.1. Diagrama de Dispersión para todos los pares de mediciones de Volumen testicular. Los instrumentos son: Ultrasonido (US), medición de referencia, e instrumentos alternativos (m1, m2, m3 y m4). A la izquierda, mediciones originales; a la derecha, mediciones al cubo.

Las especificaciones consideradas para nuestro análisis son las siguientes:

Todos los coeficientes y la variable subyacente x se asumen normales a priori. Los hiperparámetros considerados en las distribuciones de x y de los coeficientes fueron asignados de modo que la media se asemeje al estimador de mínimos cuadrados en el modelo de regresión simple respectivo y de modo que la varianza fuese comparativamente muy grande, para hacerla no informativa. Esto conduce a la siguiente selección de hiperparámetros:

- β_{01} con media 0.6 y varianza 3, β_{11} con media 0.8 y varianza 6, β_{02} con media 0.5 y varianza 2, β_{12} con media 0.9 y varianza 7, β_{03} con media 0.5 y varianza 2, β_{13} con media 1.1 y varianza 11, β_{04} con media 2.0 y varianza 36, y β_{14} con media 1 y varianza 9.
- x es normal con media 10 y varianza 30, distribución presentada en la *Figura 6.2*.

La distribución base del proceso Dirichlet, que corresponde a la distribución a priori de los ω_j se fijan de modo de asegurar que la distribución asociada a los términos de error sea una distribución elíptica lejana a la normal, en este caso con 4 g. l., para lo cual se considera una distribución base del proceso Gamma Inversa de parámetros 2 y 2.

Los hiperparámetros vinculados a las distribuciones de los parámetros de precisión ϕ_j , se centran en un valor cercano a su estimador clásico ($\phi_j = 1/\text{CME} \approx 1/3$) y se les asigna una varianza grande, esto lleva a la distribución Gamma de parámetros de escala $\lambda_{2j} = 0.3333$ y de forma $r_{2j} = 0.1111$, respectivamente.

Los modelos considerados son dos: el modelo semiparamétrico elíptico representable dependiente y el modelo semiparamétrico elíptico representable independiente. Para este último hemos considerado dos casos, en primer lugar, el modelo para el cual la ponderación de la distribución base del proceso Dirichlet es fija e igual a uno y en segundo lugar, el modelo para el cual es aleatoria.

Tabla 6.2: Estimación de parámetros realizada por Galea – Rojas et. al. (2002a, 200b) mediante un modelo estructural y estimación bayesiana para los modelos paramétrico normal, semiparamétrico elíptico dependientes y semiparamétrico elíptico independiente, este último con ponderación de la distribución base Dirichlet fija ($C=1$) y aleatoria.

Parámetros	MODELOS				
	Semip. Elíptico Independiente		Semiparamétrico Elíptico Dependiente	Paramétrico Normal	Modelo Estructural
	C = 1	C aleatorio			
β_{01}	0,280	0,255	0,284	0,270	0,0684
β_{11}	0,867	0,870	0,868	0,869	0,9320
β_{02}	0,057	0,023	0,186	0,154	0,0305
β_{12}	0,952	0,956	0,933	0,935	0,9686
β_{03}	0,172	0,145	0,087	0,082	0,0332
β_{13}	1,131	0,134	1,134	1,134	1,0319
β_{04}	1,676	1,665	1,525	1,503	0,3876
β_{14}	1,062	1,063	1,079	1,081	0,8973
ϕ_0	1,139	0,819	1,053	0,780	0,0054
ϕ_1	1,070	1,104	0,941	0,686	
ϕ_2	1,219	1,271	0,847	0,597	
ϕ_3	1,023	1,065	1,013	0,728	
ϕ_4	0,923	0,991	0,845	0,600	

La *Tabla 6.2.* presenta las estimaciones de los coeficientes β_{0j} y β_{1j} , y de los parámetros de precisión ϕ_j . Nuestras estimaciones se asemejan, en general, a las halladas por los autores antes mencionados, excepto en lo relativo a los intrreceptos β_{01} , β_{04} y ϕ_j .

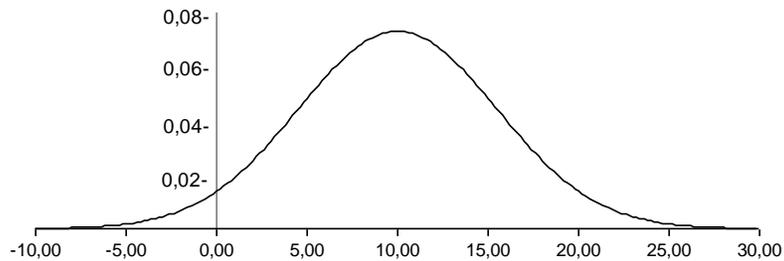


Figura 6.2. Función de densidad de distribución a priori de la verdadera medición de volumen testicular x .

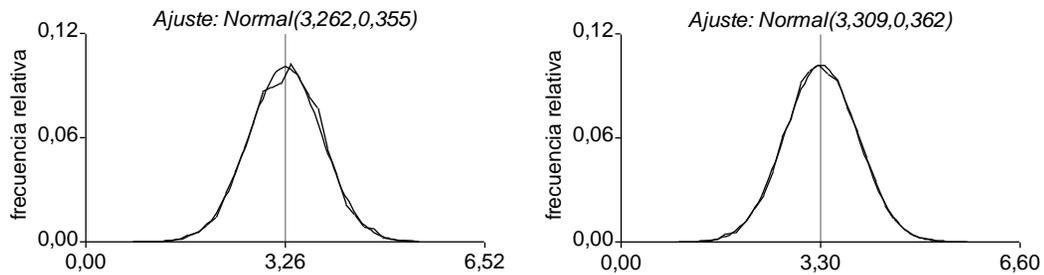
Con el conjunto de datos presentado en la *Tabla 6.1*, los autores antes mencionados realizaron un análisis de influencia, detectando que la observación 20 corresponde a un outlier. Nuestros modelos semiparamétricos elípticos pretenden capturar dicha observación; es así como la *Tabla 6.3* muestra las diferencias que se producen en las estimaciones de los coeficientes y de los parámetros de precisión al eliminar la observación influyente.

Tabla 6.3: Análisis de influencia para la observación 20 de los datos de volumen testicular (*Tabla 6.1*.) Utilizando modelos semiparamétricos elípticos independiente, dependiente y normal.

Diferencia de Estimadores	MODELOS			
	Semip. Elíptico Independiente		Semiparamétrico Elíptico Dependiente	Paramétrico Normal
	C = 1	C aleatorio		
$\beta_{01} - \beta_{01(-20)}$	0,051	0,074	0,056	0,042*
$\beta_{11} - \beta_{11(-20)}$	-0,004	-0,006	-0,004	-0,004
$\beta_{02} - \beta_{02(-20)}$	0,138	0,140	0,122	0,088*
$\beta_{12} - \beta_{12(-20)}$	-0,011	-0,012	-0,010	-0,007*
$\beta_{03} - \beta_{03(-20)}$	-0,082	-0,080*	-0,096	-0,105
$\beta_{13} - \beta_{13(-20)}$	0,006	0,005*	0,007	0,007
$\beta_{04} - \beta_{04(-20)}$	0,123*	0,144	0,145	0,120
$\beta_{14} - \beta_{14(-20)}$	-0,009*	-0,012	-0,012	-0,010
$\phi_0 - \phi_{0(-20)}$	0,029	0,007*	0,015	0,013
$\phi_1 - \phi_{1(-20)}$	0,035	0,032	0,024	0,017*
$\phi_2 - \phi_{2(-20)}$	-0,028	-0,017	0,009	0,008*
$\phi_3 - \phi_{3(-20)}$	-0,048	-0,067	0,011*	-0,019
$\phi_4 - \phi_{4(-20)}$	0,007	0,018	0,023	0,002*

La *Figura 6.3* muestra las distribuciones a posteriori para x_{20} (verdadero valor de volumen testicular para el vigésimo adolescente en el estudio, identificado como outlier por los autores, valor destacado en la *Tabla 6.1*.) bajo los modelos semiparamétricos propuestos. Estas gráficas muestran que, independientemente del modelo, la distribución a posteriori de x_{20} y, por supuesto, muy diferente de la distribución a priori (*Figura 6.2*.) En general, las distribuciones a posteriori para los modelos paramétrico normal y semiparamétrico elíptico dependiente se asemejan bastante, también se observa una gran semejanza para las distribuciones a posteriori de los dos modelos semiparamétricos elípticos independientes. Las gráficas incluyen el ajuste normal, puesto que la curva muestral no se asemeja a otras distribuciones.

- a) Distribución a posteriori de x_{20} para el modelo semiparamétrico elíptico dependiente (a la izquierda) y distribución a posteriori de x_{20} para el modelo paramétrico normal (a la derecha).



- b) Distribución a posteriori de x_{20} para los modelos semiparamétricos elípticos independientes con ponderación de la distribución base del proceso Dirichlet aleatoria (a la izquierda) y fija igual a uno (a la derecha)

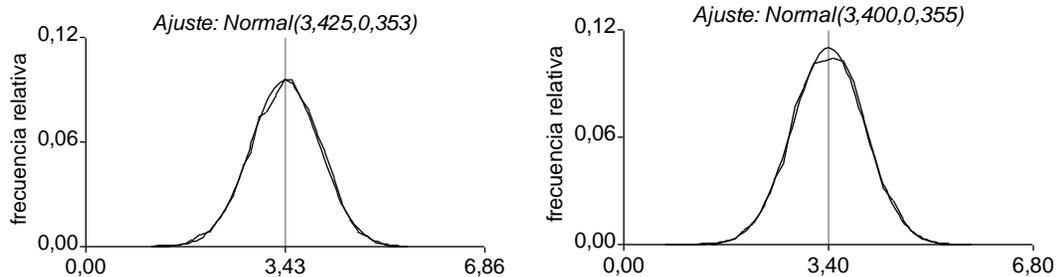


Figura 6.3. Función de densidad de distribución a posteriori de la verdadera medición de volumen testicular para el adolescente 20 basada en los modelos propuestos, con su respectivo ajuste normal.

Para cada uno de los métodos utilizados (paramétrico normal, semiparamétrico elíptico dependiente, semiparamétrico elíptico independiente con c fijo y c aleatorio) se estimaron para todos los individuos los valores de medición de ultrasonido en base a cada una de las mediciones alternativas, utilizando para ello los coeficientes estimados. La *Tabla 6.4.* muestra que, en base a la suma de cuadrados del error (SCE), definida aquí como la suma de cuadrados de las diferencias entre las observaciones y las estimaciones, el mejor modelo (entre los que hemos propuesto) es el semiparamétrico elíptico representable independiente con c fijo e igual a uno. Además cabe hacer notar que las estimaciones de los modelos elípticos y paramétrico normal son semejantes, mientras que las estimaciones de los modelos semiparamétricos elípticos independientes se asemejan bastante entre ellas.

Entonces, la distribución a posteriori para la verdadera medición del volumen testicular del adolescente 20, es aproximadamente normal con media 3,4 y varianza 0,355 (*Figura 6.3. b*).

Tabla 6.4: Sumas de los cuadrados del error SCE de las estimaciones de volumen testicular utilizando los estimadores de Bayes de los valores de volumen testicular real \mathbf{x} como estimación de los valores de ultra sonido \mathbf{y}_0 .

Modelos	SCE
Modelo semiparamétrico dependiente	40,18
Modelo semiparamétrico independiente	39,40
Modelo semiparamétrico independiente con C aleatoria	43,82
Modelo paramétrico Normal	40,58

Finalmente es claro que:

- Las distribuciones a posteriori para el parámetro de interés x_{20} difieren significativamente de la distribución a priori, por lo cual se ve que no resultan influenciadas notablemente por dicha distribución.
- Las distribuciones a posteriori para x_{20} basadas en el modelo semiparamétrico elíptico independiente con c fijo y en el modelo paramétrico normal presentan formas más gaussianas que los otros modelos.
- La distribución a posteriori más concentrada es la relativa al modelo semiparamétrico elíptico independiente con c aleatorio (capítulo 1, sección 1.2) y la más dispersa es la relativa al modelo paramétrico normal.
- Todas las distribuciones a posteriori son mucho menos dispersas que la distribución a priori.

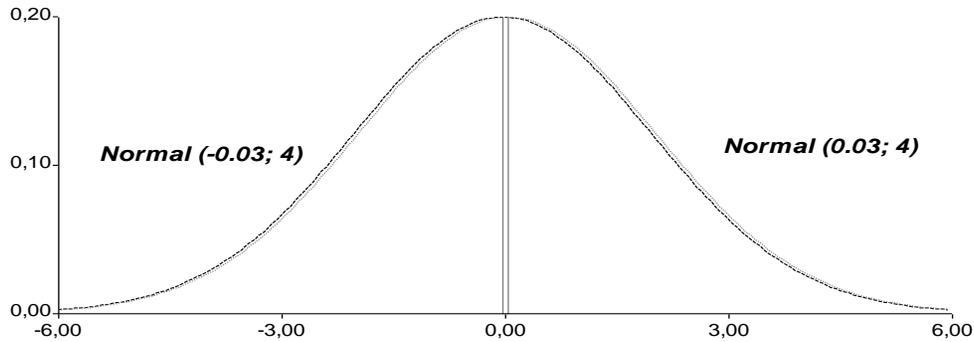
Aplicación 6.2: Simulación

Se realizó una simulación de datos para calibración comparativa con tres instrumentos. Se consideraron dos tamaños de muestra ($n = 10$ y $n = 50$), valores de $\beta_{01} = 0.03$, $\beta_{02} = -0.03$, $\beta_{11} = 0.97$ y $\beta_{12} = 1.03$, todos estos valores escogidos de modo que los instrumentos presentan buena calibración. Los valores de la predictora x se generan equiespaciados entre -1 y 1 , los errores son normales o t – student (independientes) con cuatro grados de libertad. El parámetro de precisión $\phi = 1/\sigma^2$ asume valores 100 y 64 .

El objetivo de esta aplicación es evaluar la calidad de los estimadores en función de la relación de los parámetros de precisión.

El análisis considera las distribuciones a priori de los coeficientes normales, con medias nulas para los interceptos y con medias de las pendientes unitarias. Estas distribuciones a priori contemplan varianzas grandes (4). La *Figura 6.4.* muestra las distribuciones a priori de los coeficientes.

a) Distribuciones a priori de β_{01} (media 0.03) y de β_{02} (media -0.03)



b) Distribuciones a priori de β_{11} (media 0.97) y de β_{12} (media 1.03)

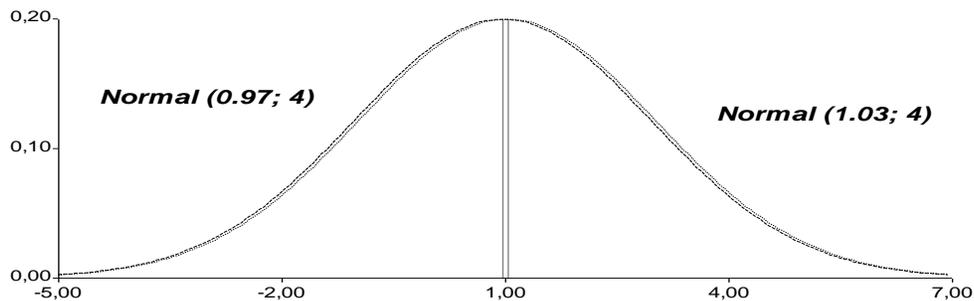


Figura 6.4. Funciones de densidad de las distribución a priori de los coeficientes del modelo de calibración comparativa con dos instrumentos a calibrar.

Tabla 6.6: Inferencia bayesiana para simulación de modelos de calibración comparativa con errores normales y *t*-student, mediante modelos elípticos semiparamétricos dependiente e independiente y mediante modelo paramétrico normal. Se adjuntan también estimadores clásicos de las regresiones univariadas. Tamaño de muestra $n = 10$.

	CALIBR. COMPARATIVA MOD. INDEP. C ALEATORIO				CALIBR. COMPARATIVA MOD. INDEPENDIENTE			
N=10, Err. normales	β_{01}	β_{11}	β_{02}	β_{12}	β_{01}	β_{11}	β_{02}	β_{12}
$\phi_0 = \phi_1 = \phi_2 = 100$	0,066	1,125	0,028	1,148	0,066	1,134	0,032	1,156
$\phi_0 = \phi_1 = \phi_2 = 64$	-0,008	1,037	-0,085	1,218	-0,009	1,037	-0,082	1,216
$100 = \phi_0 > \phi_1 = \phi_2 = 64$	-0,03	0,891	-0,057	1,128	-0,026	0,889	-0,052	1,119
$64 = \phi_0 < \phi_1 = \phi_2 = 100$	0,05	1,045	-0,022	1,118	0,055	1,064	-0,018	1,135
N=10, Err. t-student	β_{01}	β_{11}	β_{02}	β_{12}	β_{01}	β_{11}	β_{02}	β_{12}
$\phi_0 = \phi_1 = \phi_2 = 100$	0,038	0,912	-0,002	0,917	0,036	0,916	-0,003	0,920
$\phi_0 = \phi_1 = \phi_2 = 64$	-0,097	1,184	0,005	1,331	-0,094	1,140	0,004	1,273
$100 = \phi_0 > \phi_1 = \phi_2 = 64$	0,012	0,966	-0,099	1,147	0,014	0,996	-0,096	1,176
$64 = \phi_0 < \phi_1 = \phi_2 = 100$	0,035	0,88	-0,058	0,978	0,030	0,851	-0,065	0,945
	CALIBR. COMPARATIVA MOD. DEPENDIENTE				CALIBR. COMPARATIVA MOD. NORMAL			
N=10, Err. normales	β_{01}	β_{11}	β_{02}	β_{12}	β_{01}	β_{11}	β_{02}	β_{12}
$\phi_0 = \phi_1 = \phi_2 = 100$	0,059	1,132	0,030	1,152	0,059	1,124	0,030	1,145
$\phi_0 = \phi_1 = \phi_2 = 64$	-0,004	1,032	-0,085	1,210	-0,005	1,044	-0,085	1,225
$100 = \phi_0 > \phi_1 = \phi_2 = 64$	-0,028	0,897	-0,037	1,122	-0,026	0,908	-0,037	1,139
$64 = \phi_0 < \phi_1 = \phi_2 = 100$	0,047	1,060	-0,026	1,137	0,047	1,049	-0,026	1,126
N=10, Err. t-student	β_{01}	β_{11}	β_{02}	β_{12}	β_{01}	β_{11}	β_{02}	β_{12}
$\phi_0 = \phi_1 = \phi_2 = 100$	0,039	0,912	-0,031	0,978	0,037	0,932	-0,033	1,000
$\phi_0 = \phi_1 = \phi_2 = 64$	-0,093	1,150	0,001	1,267	-0,093	1,192	-0,001	1,338
$100 = \phi_0 > \phi_1 = \phi_2 = 64$	0,000	1,013	-0,094	1,170	-0,001	1,016	-0,096	1,153
$64 = \phi_0 < \phi_1 = \phi_2 = 100$	0,042	0,863	-0,064	0,954	0,042	0,915	-0,066	1,016
	MODELO REGRESIÓN CLÁSICO UNIVARIADO							
N=10, Err. normales	β_{01}	β_{11}	β_{02}	β_{12}				
$\phi_0 = \phi_1 = \phi_2 = 100$	0,060	1,130	0,030	1,151				
$\phi_0 = \phi_1 = \phi_2 = 64$	-0,004	1,037	-0,083	1,215				
$100 = \phi_0 > \phi_1 = \phi_2 = 64$	-0,026	0,889	-0,037	1,112				
$64 = \phi_0 < \phi_1 = \phi_2 = 100$	0,047	1,048	-0,026	1,131				
N=10, Err. t-student	β_{01}	β_{11}	β_{02}	β_{12}				
$\phi_0 = \phi_1 = \phi_2 = 100$	0,038	0,906	-0,033	0,959				
$\phi_0 = \phi_1 = \phi_2 = 64$	-0,095	1,195	0,000	1,302				
$100 = \phi_0 > \phi_1 = \phi_2 = 64$	0,000	0,989	-0,095	1,140				
$64 = \phi_0 < \phi_1 = \phi_2 = 100$	0,044	0,842	-0,061	0,899				

Tabla 6.7: Inferencia bayesiana para simulación de modelos de calibración comparativa con errores normales y *t* – student, mediante modelos elípticos semiparamétricos dependiente e independiente y mediante modelo paramétrico normal. Se adjuntan también estimadores clásicos de las regresiones univariadas. Tamaño de muestra $n = 50$.

	CALIBR. COMPARATIVA MOD. INDEP. C ALEATORIO				CALIBR. COMPARATIVA MOD. INDEPENDIENTE			
N=50, Err. normales	β_{01}	β_{11}	β_{02}	β_{12}	β_{01}	β_{11}	β_{02}	β_{12}
$\phi_0 = \phi_1 = \phi_2 = 100$	0,018	0,932	-0,040	0,984	0,018	0,996	-0,041	1,049
$\phi_0 = \phi_1 = \phi_2 = 64$	0,020	0,941	-0,001	1,004	0,020	0,898	0,000	0,954
$100 = \phi_0 > \phi_1 = \phi_2 = 64$	-0,023	0,958	-0,057	0,986	-0,024	0,896	-0,057	0,924
$64 = \phi_0 < \phi_1 = \phi_2 = 100$	0,063	0,916	-0,023	0,941	0,061	0,914	-0,025	0,937
N=50, Err. t-student	β_{01}	β_{11}	β_{02}	β_{12}	β_{01}	β_{11}	β_{02}	β_{12}
$\phi_0 = \phi_1 = \phi_2 = 100$	0,065	0,918	-0,021	0,919	0,066	0,937	-0,021	0,936
$\phi_0 = \phi_1 = \phi_2 = 64$	0,022	0,817	-0,062	0,823	0,022	0,817	-0,062	0,823
$100 = \phi_0 > \phi_1 = \phi_2 = 64$	0,055	0,899	0,007	1,006	0,056	0,833	0,008	0,930
$64 = \phi_0 < \phi_1 = \phi_2 = 100$	0,011	0,951	-0,037	0,991	0,012	0,949	-0,037	0,985
	CALIBR. COMPARATIVA MOD. DEPENDIENTE				CALIBR. COMPARATIVA MOD. NORMAL			
N=50, Err. normales	β_{01}	β_{11}	β_{02}	β_{12}	β_{01}	β_{11}	β_{02}	β_{12}
$\phi_0 = \phi_1 = \phi_2 = 100$	0,017	0,979	-0,041	1,027	0,016	0,942	-0,042	0,985
$\phi_0 = \phi_1 = \phi_2 = 64$	0,014	0,892	-0,009	0,950	0,014	0,953	-0,008	1,016
$100 = \phi_0 > \phi_1 = \phi_2 = 64$	-0,020	0,901	-0,059	0,938	-0,021	0,960	-0,060	1,001
$64 = \phi_0 < \phi_1 = \phi_2 = 100$	0,064	0,907	-0,020	0,930	0,064	0,925	-0,021	0,949
N=50, Err. t-student	β_{01}	β_{11}	β_{02}	β_{12}	β_{01}	β_{11}	β_{02}	β_{12}
$\phi_0 = \phi_1 = \phi_2 = 100$	0,064	0,907	-0,019	0,922	0,065	0,924	-0,018	0,927
$\phi_0 = \phi_1 = \phi_2 = 64$	0,022	0,814	-0,053	0,805	0,021	0,887	-0,053	0,880
$100 = \phi_0 > \phi_1 = \phi_2 = 64$	0,081	0,899	0,009	0,963	0,081	0,951	0,009	1,021
$64 = \phi_0 < \phi_1 = \phi_2 = 100$	0,011	0,945	-0,037	0,979	0,010	0,966	-0,038	0,994
	MODELO REGRESIÓN CLÁSICO UNIVARIADO							
N=50, Err. normales	β_{01}	β_{11}	β_{02}	β_{12}				
$\phi_0 = \phi_1 = \phi_2 = 100$	0,016	0,933	-0,042	0,986				
$\phi_0 = \phi_1 = \phi_2 = 64$	0,014	0,946	-0,060	1,006				
$100 = \phi_0 > \phi_1 = \phi_2 = 64$	-0,020	0,954	-0,060	0,991				
$64 = \phi_0 < \phi_1 = \phi_2 = 100$	0,064	0,912	-0,021	0,938				
n=50, Err. t-student	β_{01}	β_{11}	β_{02}	β_{12}				
$\phi_0 = \phi_1 = \phi_2 = 100$	0,064	0,900	-0,019	0,918				
$\phi_0 = \phi_1 = \phi_2 = 64$	0,021	0,867	-0,054	0,853				
$100 = \phi_0 > \phi_1 = \phi_2 = 64$	0,081	0,918	0,008	0,986				
$64 = \phi_0 < \phi_1 = \phi_2 = 100$	0,010	0,954	-0,037	0,988				

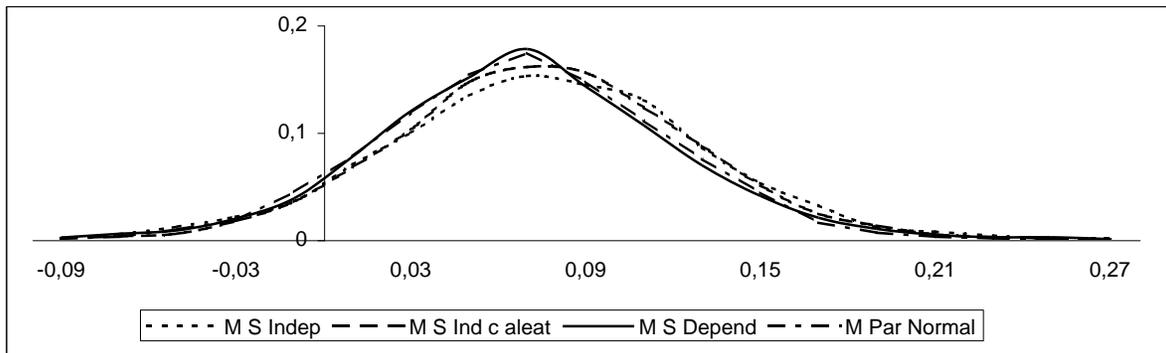
La distribución base del proceso Dirichlet es gamma inversa con parámetros 2 y 2, lo cual genera un modelo con presenta 4 g.l., y la distribución a priori para los parámetros de precisión ϕ_j es gamma con hiperparámetros asumiendo valores tales que la media corresponde al verdadero valor y la varianza es moderada, es decir, los valores de (r, λ) se escogieron como (100, 1) y (40.96, 0.64), respectivamente. Los valores de los hiperparámetros para la distribución a priori normal de x son cero y ocho.

La inferencia es presentada en las *Tablas 6.6. y 6.7.*, donde se comparan las estimaciones para los modelos bayesianos semiparamétricos elípticos dependiente SED (definido por la *Proposición 6.4*) e independiente SEI (definido por la *Proposición 6.2.*), paramétrico normal PN y las estimaciones del modelo de regresión simple clásico RC.

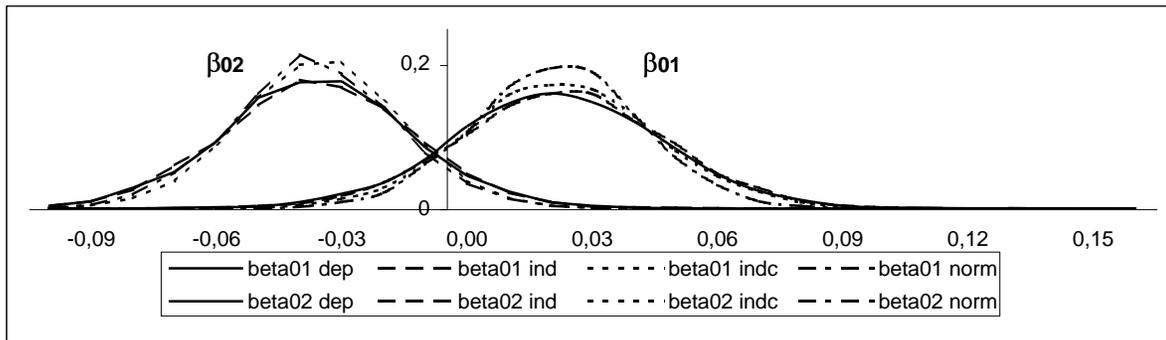
De las *Tablas 6.6. y 6.7.*, podemos deducir que:

- Si los errores de las mediciones son normales, para ambos modelos semiparamétricos considerados (SED y SEI), en el caso de varianzas iguales, la estimación mejora con el tamaño de muestra, pero no se aprecia lo mismo para varianzas distintas.
- Si los errores de medición siguen una distribución t – student, se da el fenómeno antes descrito sólo en el caso en que las varianzas son grandes e iguales.
- En general, para tamaños de muestra pequeños ($n = 10$) y errores normales el modelo SED es levemente superior al modelo SEI. Esta diferencia no se mantiene consistentemente al aumentar el tamaño de muestra ($n = 50$), ya que en dicho caso el modelo SED supera al modelo SEI cuando todas las varianzas son pequeñas e iguales, pero en general, si las varianzas son mayores e iguales, ocurre a la inversa.
- Independiente del tamaño de la muestra, si las varianzas son pequeñas e iguales, ambos modelos semiparamétricos estiman mejor. Por otro lado, para tamaños de muestra pequeño se aprecia que si la varianza de las mediciones del instrumento de referencia y de los otros instrumentos no son iguales, para ambos modelos semiparamétricos, en general, las estimaciones mejoran respecto del caso en el cual todas las varianzas son iguales a la menor. Para muestras grandes y para ambos modelos semiparamétricos, en cambio, si la varianza de las mediciones del instrumento de referencia es menor que la de los otros instrumentos, se estima mejor que si la varianza del instrumento de referencia es la mayor.
- Para muestra pequeña el modelo PN entrega estimaciones semejantes a las del modelo SED y a las estimaciones obtenidas del modelo RC.
- Cuando la muestra es mayor, el modelo SED supera al modelo PN y el modelo SEI lo supera cuando todas las varianzas son pequeñas.
- Para muestras pequeñas ($n = 10$) y errores t – student, el modelo SED es, en general, superior al modelo SEI. Ahora bien, cuando la muestra crece ($n = 50$), el modelo SEI es superior al modelo SED para el caso en que todas las varianzas son pequeñas, y los modelos no difieren mucho si las varianzas son iguales y mayores; sin embargo, cuando la mayor varianza la presenta el instrumento de referencia, estima mejor el modelo SEI.
- En general, independientemente del tamaño de la muestra, ambos modelos semiparamétricos estiman mejor para varianzas más pequeñas; consecuentemente, para muestras pequeñas, si la varianza de las mediciones del instrumento de referencia y de los otros instrumentos no son iguales, para ambos modelos semiparamétricos las estimaciones son mejores que para el caso en el cual todas las varianzas son iguales a la mayor. En fin, para muestras grandes y varianzas pequeñas, el modelo SEI supera a los modelos SED y PN, y estos últimos dos se asemejan.
- Para muestras pequeñas, los estimadores del modelo RC funcionan bien cuando todas las varianzas son pequeñas y el modelo PN funciona mejor; pero para muestras grandes, esta relación se mantiene cuando el instrumento de referencia tiene varianza menor y esas estimaciones se asemejan en precisión a las del modelo PN.
- Finalmente, el modelo que considera la ponderación de la distribución base del proceso Dirichlet aleatoria SEIA es más preciso que el modelo que la considera fija SEIF, pero la diferencia se hace más evidente para tamaños de muestra grandes.

a) Distribuciones a posteriori de β_{01} con $\phi_j = 100$ ($j = 0, 1, 2$), con errores normales y tamaño de muestra $n = 10$, para los modelos de interés.



b) Distribuciones a posteriori de β_{01} y β_{02} , para errores normales, tamaño de muestra $n = 50$ y $\phi_j = 100$, en los diferentes modelos.



c) Distribuciones a posteriori de β_{01} y β_{02} , para errores t – student, tamaño de muestra $n = 50$ y $\phi_j = 100$, en los diferentes modelos.

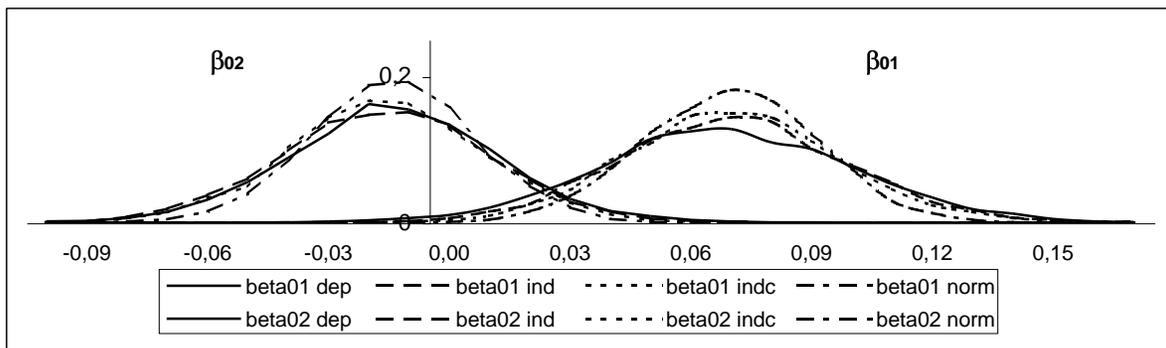
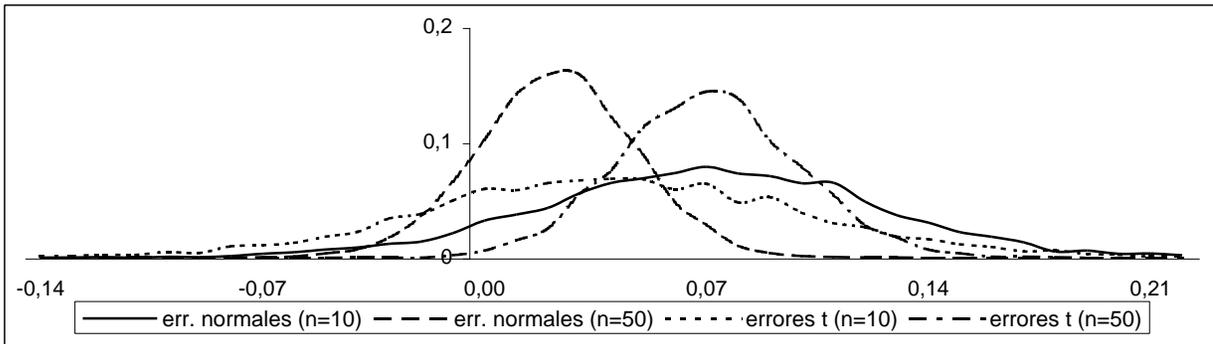
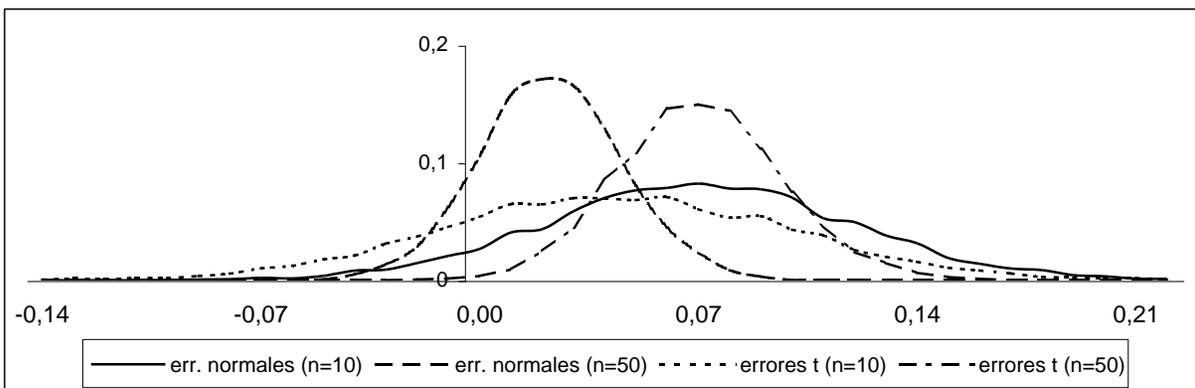


Figura 6.5. Comparación de funciones de densidad de las distribución a posteriori de los coeficientes del modelo de calibración comparativa con dos instrumentos a calibrar, para los diferentes modelos utilizados y errores normales o t – student.

- a) Distribuciones a posteriori de β_{01} con $\phi_j = 100$ ($j = 0, 1, 2$), con errores normales y t – student y con tamaños de muestra $n = 10$ y $n = 50$, para el modelo independiente con $c = 1$.



- b) Distribuciones a posteriori de β_{01} con $\phi_j = 100$ ($j = 0, 1, 2$), con errores normales y t – student y con tamaños de muestra $n = 10$ y $n = 50$, para el modelo independiente con c aleatorio.



- c) Distribuciones a posteriori de β_{01} con $\phi_j = 100$ ($j = 0, 1, 2$), con errores normales y t – student y con tamaños de muestra $n = 10$ y $n = 50$, para el modelo dependiente.

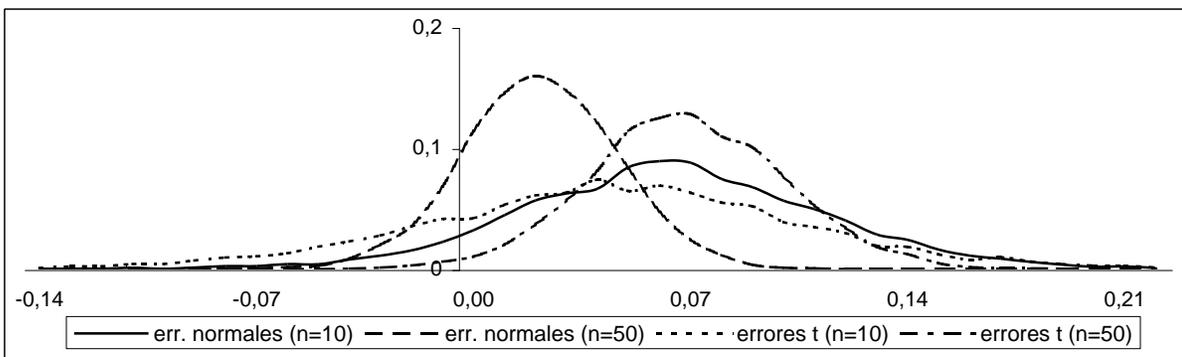
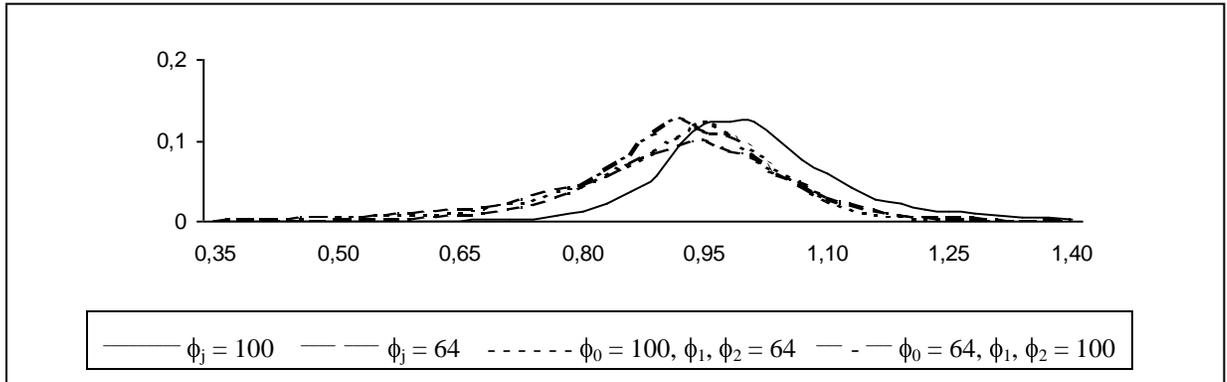
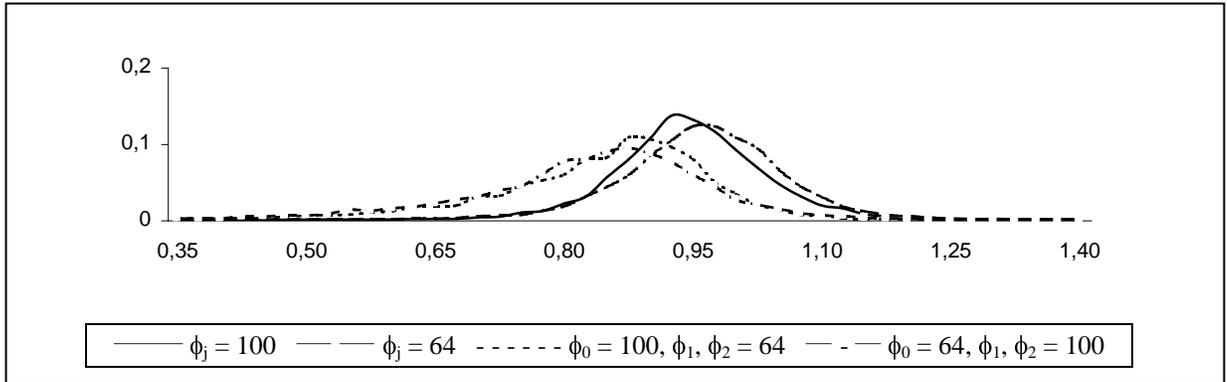


Figura 6.6. Comparación de funciones de densidad de las distribución a posteriori de los coeficientes del modelo de calibración comparativa con dos instrumentos a calibrar, para diferentes combinaciones de tamaños de muestra y distribuciones para los errores.

a) Distribuciones a posteriori de β_{11} , para el modelo independiente con c fijo y errores normales, para un tamaño de muestra $n = 50$ y para distintas combinaciones de los ϕ_j .



b) Distribuciones a posteriori de β_{11} , para el modelo independiente con c fijo y errores t – student, para un tamaño de muestra $n = 50$ y para distintas combinaciones de los ϕ_j .



c) Distribuciones a posteriori de β_{11} , para el modelo independiente con c aleatorio y errores t – student, para un tamaño de muestra $n = 50$ y para distintas combinaciones de los ϕ_j .

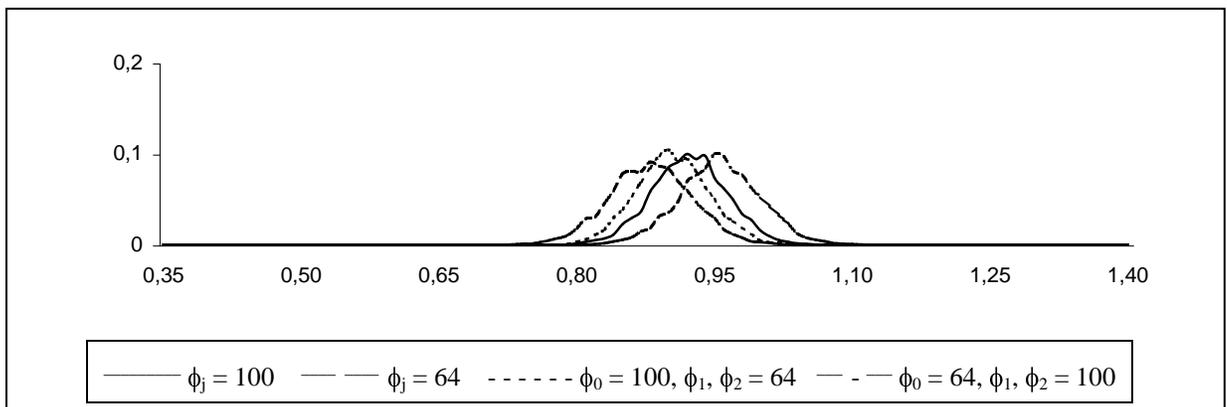


Figura 6.7. Comparación de funciones de densidad de las distribución a posteriori de la pendiente del modelo de calibración comparativa con dos instrumentos a calibrar, con tamaños de muestra $n = 50$, para diferentes combinaciones de magnitudes de las varianzas y distribuciones para los errores.

Las Figuras 6.5., 6.6. y 6.7. muestran algunas de las distribuciones a posteriori de coeficientes del modelo de calibración comparativa. Algunas observaciones interesantes se desprenden de estos gráficos.

Bajo el modelo de calibración comparativa con errores normales, muestras pequeñas ($n = 10$) y varianzas pequeñas e iguales, no existen diferencias muy grandes en las distribuciones a posteriori del intercepto para los modelos analizados; sin embargo todos los modelos entregan distribuciones positivamente sesgadas. El problema del sesgo se corrige con el aumento del tamaño muestral. Ahora bien, si los errores son t – student, una muestra grande ($n = 50$) de todas formas arroja estimaciones sesgadas y en este caso el sesgo también es de tipo positivo. En general, el análisis con el modelo PN arroja una distribución a posteriori con varianza menor a la de los otros modelos (aunque su sesgo es significativo) y le sigue el modelo SEIA. Luego, si los errores son normales, le sigue el modelo SED y si los errores son t – student, el modelo SEI, aunque estos dos presentan distribuciones a posteriori parecidas. En fin, cabe hacer notar que las distribuciones a posterioris para el intercepto se asemejan a una distribución normal, especialmente para muestra grande (Figura 6.5.).

También se aprecia que si los errores son normales, independientemente del modelo que se utilice para estimar el intercepto, se obtienen estimadores mejor centrados en el verdadero valor del parámetro para tamaño de muestra grande; al contrario, si los errores son t – student, se logra menor sesgo en la estimación para muestra pequeña. Como es natural, las distribuciones a posteriori presentan mayor variabilidad cuando el tamaño muestral es menor (Figura 6.6.).

Por otra parte, para varianzas de los errores del medición del instrumento de referencia y varianzas de los modelos de regresión para los otros instrumentos todas iguales, las distribuciones a posteriori son aproximadamente simétricas; mientras que en caso contrario son negativamente sesgadas (Figura 6.7.). Además, para el modelo SEI, si los errores son normales, se estima con mejor precisión si las varianzas son iguales, pero en el caso de errores distribuidos t – student, este modelo estima con mejor precisión cuando la varianza del error de medición es mayor a las otras varianzas. Finalmente, para el caso en que c se considera aleatorio (modelo SEIA, capítulo 1, sección 1.2), todas las distribuciones a posteriori son relativamente simétricas.

CONCLUSIONES

Este capítulo aborda el análisis bayesiano semiparamétrico del modelo de calibración comparativa.

Con el propósito de justificar la utilización del enfoque bayesiano, se presenta la función de verosimilitud asociada al problema, así como las ecuaciones a resolver para la obtención de estimadores máximo verosímiles para los parámetros. Dichas expresiones y ecuaciones resultan ser intratables analíticamente.

Para dar solución en forma adecuada al problema de interés, se incluyen las distribuciones condicionales completas de todos los parámetros, puesto que a partir de ellas, implementando el muestreo de Gibbs, se simulan los estimadores de Bayes (bajo pérdida cuadrática).

Los modelos considerados en este capítulo consideran errores con distribución elíptica representable independiente y dependiente.

Se utiliza la metodología implementada para dar solución a un problema presente en la literatura, “*Medición de Volumen Testicular de 42 adolescente mediante 5 métodos*” (Chipkevitch, Nishimura, Tu y Galea – Rojas, 1996). También se realiza una simulación del modelo con el fin de evaluar la calidad de los estimadores en función de la magnitud de los parámetros de precisión ϕ_j y de la relación entre aquel relativo al error de medición y aquellos relacionados con las regresiones para los instrumentos alternativos.

El análisis de los datos simulados brinda algunas conclusiones. En primer lugar, respecto de la distribución de los errores, tenemos que para el caso en que los errores son normales la estimación mejora con el tamaño de muestras y las varianzas (ϕ_j^{-1}) son iguales; mientras que si son *t* – Student se requiere además que sean grandes o bien que la varianza del error de medición sea mayor. Además, si los errores son normales para muestras grandes, se reduce el sesgo; mientras que si son *t* – Student, el sesgo disminuye para muestras pequeñas.

En segundo lugar, y respecto del modelo utilizado para la estimación, se tiene que para los modelos semiparamétricos y considerando varianzas iguales, el modelo dependiente estima mejor para varianzas pequeñas y el independiente para varianzas grandes. Ahora bien, si el tamaño de muestra es pequeño, las estimaciones que entrega el modelo paramétrico normal, e incluso las surgidas de regresiones simples individuales para cada instrumento alternativo, se asemejan a las estimaciones que provee el modelo semiparamétrico elíptico dependiente, mientras que para tamaños de muestras mayores se aprecia la supremacía de este último modelo y se corrigen los problemas de sesgo en la estimación. Por otra parte, los modelos semiparamétricos con *c* aleatorio estiman mejor que aquellos con *c* fijo.

En tercer lugar, en lo relativo a la magnitud de las varianzas, si aquella relativa al error de medición difiere de las otras, se estima mejor para el caso en que la primera es menor. Finalmente varianzas iguales llevan a posterioris simétricas, mientras que varianzas distintas llevan a posterioris asimétricas.

Los resultados obtenidos de este capítulo se limitan a *prioris* dentro del modelo normal – gama, a regresiones lineales simples, a coeficientes asumidos independientes entre sí e independientes del resto de los parámetros y a errores elípticos; de modo que se propone considerar las siguientes extensiones: *prioris* fuera del modelo normal – gama, más de un apredictora en el modelo, estructuras de dependencia entre los parámetros de interés y errores skew normal o skew elípticos.

ANEXO:

GLOSARIO DE NOTACIONES

A continuación se presenta la notación relativa a las distribuciones utilizada en este trabajo.

- $N_n(\boldsymbol{\mu}, \Sigma)$ → Distribución Normal n – variada con vector de medias $\boldsymbol{\mu}$ y matriz de varianzas y covarianzas Σ .
- Bernoulli (θ) → Distribución de Bernoulli de parámetro θ .
- $T(\gamma)$ → Distribución T – Student central con γ grados de libertad.
- $T(\boldsymbol{\mu}, \sigma^2, \gamma)$ → Distribución T – Student no central con media $\boldsymbol{\mu}$, parámetro de escala σ^2 y γ grados de libertad.
- Gamma (r, λ) → Distribución Gamma con parámetro de escala λ y parámetro de forma r .
- $IG(r, \lambda)$ → Distribución Gamma Inversa con parámetro de escala λ y de forma r .
- $El_1(\boldsymbol{\mu}, \sigma^2, h)$ → Distribución Elíptica uni – variada con media $\boldsymbol{\mu}$, parámetro de escala σ^2 y función generadora de densidad h .
- $El_n(\boldsymbol{\mu}, \Sigma, h)$ → Distribución Elíptica n – variada con parámetro de medias $\boldsymbol{\mu}$, Matriz de varianzas y covarianzas Σ y función generadora de densidad h .
- $D(\boldsymbol{\alpha})$ → Distribución Dirichlet de vector de parámetros $\boldsymbol{\alpha}$.
- $D(c G_0)$ → Proceso Dirichlet con parámetro de precisión c y distribución base o central G_0 .
- Φ → Función de distribución acumulada normal estándar.
- Φ_E → Función de distribución acumulada elíptica estándar o esférica.
- $x_i \sim^{i.i.d.} F$ → Las variables aleatorias x_i son independientes y su distribución es F idéntica.
- $x_i \sim^{ind.} F_i$ → Las variables aleatorias x_i son independientes con distribución F_i .
- $x_i | \theta_i \sim^{c.i.i.d.} F$ → Las variables aleatorias x_i son condicionalmente independientes dadas las variables θ_i y su distribución F es común.
- $x_i | \theta_i \sim^{cond.ind.} F_i$ → Las variables aleatorias x_i son condicionalmente independientes dadas las variables θ_i y la distribución de $x_i | \theta_i$ es F_i .
- $x \perp y$ → las variables x e y son independientes
- $x \perp y | z$ → las variables x e y son condicionalmente independientes, dado z .

CONCLUSIONES

En el presente trabajo se realizó un análisis bayesiano semiparamétrico de cuatro modelos: Calibración Lineal Discreta, calibración Lineal Continua, Modelo con Errores de Medición y Calibración Comparativa.

Para cada modelo se obtuvo la función de verosimilitud y las ecuaciones a resolver para la obtención de los estimadores máximo verosímiles. En todos los casos, las expresiones para las funciones de verosimilitud y los sistemas de ecuaciones a resolver resultan ser intratables analíticamente debido a que los modelos que nos interesan son semiparamétricos. Esto justifica la determinación de todas las distribuciones condicionales completas, realizada con el fin de implementar el Muestreo de Gibbs y obtener, de este modo, estimadores de Bayes desde las densidades a posteriori de los parámetros de interés. Para cada modelo abordado se programó computacionalmente dicha implementación y se utilizaron los programas para realizar inferencia bayesiana respecto de conjuntos de datos presentes en la literatura y comparar los resultados con los de otros autores. También se utilizaron los programas para realizar inferencia en conjuntos de datos simulados, determinando de esta forma las bondades y limitaciones de las técnicas propuestas.

En el caso de Calibración Lineal Discreta, se realizó un aporte consistente en la extensión del Lema de Basu y Mukhopadhyay (2000). Este lema vinculaba un modelo de regresión binaria con el modelo normal mediante la inclusión de una variable subyacente normal y una muestra de variables instrumentales i.i.d. Estos autores también consideraron una función de enlace diferente de las usuales (logito y probito) y que corresponde a una mezcla en parámetro de escala de funciones de distribución acumulada normales. Una primera extensión realizada consiste en implementar la resolución al problema de calibración lineal binaria manteniendo las funciones de enlace utilizadas por ellos, es así como surgen los casos paramétrico y semiparamétrico, dependiendo de si la distribución a priori de los pesos en la mezcla de normales es discreta o es un proceso de Dirichlet, respectivamente. Una segunda extensión que implica un aporte muy significativo debido a su importancia en el campo de la dosimetría, se relaciona con la resolución del problema de calibración binomial o la inferencia bayesiana para los parámetros de una distribución multinomial, bajo el enfoque paramétrico (distribución a priori de los parámetros que definen la mezcla discreta) y semiparamétrico. Las extensiones tercera y cuarta se relacionan, respectivamente, con la resolución del problema de regresión y de calibración binaria, considerando enlace elíptico representable independiente conocido. Estos modelos son paramétricos y los resultados se aplican al caso en que el enlace es t – student. La quinta y última extensión al trabajo de los autores antes mencionados consiste en la resolución paramétrica del problema de calibración binomial considerando función de enlace elíptica conocida; naturalmente, al igual que en los casos anteriores, los resultados se aplican al caso de enlace t – student.

Para los problemas de Calibración Continua, Modelo con Errores de Medición y Calibración Comparativa, se extienden los resultados paramétricos presentes en la literatura, y que consideran términos de error siguiendo una distribución normal, una distribución elíptica representable conocida, o mezclas de distribuciones elípticas en las cuales los parámetros que definen la mezcla tienen una distribución a priori discreta. La extensión considerada para los modelos de interés contempla la resolución bayesiana Semiparamétrica, donde no nos amarramos a una mezcla particular de distribuciones elípticas, sino que la priori de la función mezcladora sigue un proceso Dirichlet. Para estos tres problemas se consideran dos casos, el modelo elíptico representable independiente y el modelo elíptico representable dependiente. En todos los casos, resultaron de gran utilidad los resultados de Arellano – Valle, Galea e Iglesias (1999), ya que permiten relacionar los modelos con errores elípticos representables con sus análogos de errores normales, mediante la inclusión ciertos parámetros de dispersión como variables latentes en el modelo jerárquico. La ventaja de establecer esta relación, está en que se puede recuperar modelos conocidos y extender así sus resultados. El carácter semiparamétrico del modelo se lo da el hecho de que la priori para estos parámetros latentes es un proceso Dirichlet en el caso continuo.

En todos los problemas abordados, se utilizaron las distribuciones a priori usuales para los coeficientes del modelo, para los parámetros de dispersión y para las variables predictoras (coeficientes normales, parámetro de dispersión γ , predictoras normales), de modo que los resultados se basan en el modelo normal – γ .

Las aplicaciones utilizadas para ejemplificar el uso de las técnicas propuestas se relacionan con conjuntos de datos presentes en la literatura, de modo que permitieron la comparación en el ámbito de la precisión con otros modelos de uso tradicional.

Los modelos abordados están relacionados entre sí, de modo que las proposiciones y lemas generales presentados en el capítulo 2 permitieron abordar cada uno de los problemas de forma satisfactoria, por otro lado, el algoritmo de Escobar y West (1998) fue utilizado para la construcción de los programas computacionales. La importancia de este algoritmo está en que permite hallar las distribuciones condicionales completas de las variables latentes con priori Dirichlet que fueron incluidas en los modelos. Los programas fueron construidos en SAS System y presentaron una velocidad razonable (aproximadamente 20 minutos para los casos independiente, que incluyen las variables latentes, por lo cual el número de parámetros es altísimo, y aproximadamente 30 segundos para los casos dependientes, que no requieren el uso del algoritmo).

Los análisis de convergencia de las cadenas arrojan buenos resultados y resultan en general sin autocorrelación.

PROYECCIONES FUTURAS

Las limitaciones presentes en este trabajo inducen futuras proyecciones, por lo cual es importante mencionarlas:

- La función de enlace utilizada para el modelo de Calibración Lineal Discreta es simétrica.
- Para los problemas de Calibración Lineal Discreta, Modelo con Errores de Medición y Calibración Comparativa, el modelo de regresión utilizado es simple.
- Las distribuciones a priori utilizadas se restringen, en general, al modelo normal – gamma.
- Los coeficientes del modelo se asumen independientes entre sí e independientes de los otros parámetros.
- Las distribuciones de los errores en los modelos son elípticas.

Algunas extensiones que se pueden proponer, entonces, son las siguientes:

- Para el modelo de Calibración Discreta se propone implementar la resolución bayesiana semiparamétrica para
 - El modelo de regresión discreta binaria
 - con función de enlace asimétrica (extensión de Basu y Mukhopadhyay, 2000), por ejemplo skew normal y skew elípticos,
 - con más de una predictora.
 - El modelo de regresión discreta no binaria (distribución multinomial, distribución binomial)
 - con función de enlace asimétrica (extensión de Basu y Mukhopadhyay, 2000), por ejemplo skew normal y skew elípticos,
 - con más de una predictora.
- Para los modelos de Calibración Lineal Continua, Errores de Medición y Calibración Comparativa se propone implementar la resolución bayesiana semiparamétrica
 - considerando errores skew normal y skew elíptico,
 - considerando distribuciones a priori fuera del modelo normal – gamma,
 - considerando modelos jerárquicos con estructuras de dependencia entre los parámetros de interés.

REFERENCIAS

1. Aitchison J. y Dunsmore I. (1975). *Statistical prediction Analysis*. Cambridge University Press.
2. Agresti A. (2002), *Categorical Data Analysis*, Second Edition, New York, John Wiley y Sons.
3. Amewou – Atisso, M., Ghosal, S., Ghosh, J. K. y Ramamoorthi, R. V. (2003). Posterior consistency for semiparametric regression problems. *Bernoulli* **9**, 291 – 312.
4. Arellano – Valle R. B., Bolfarine, H. e Iglesias, P. (1994). A predictivistic interpretation to the multivariate T distribution. *Test* **3** (2), 221 – 236.
5. Arellano – Valle R. B. y Bolfarine, H. (1995). On some characterizations of the t – distribution. *Statistics and Probability Letters* **25**, 79 – 85.
6. Arellano-Valle R. B. y Bolfarine, H. (1996). Elliptical structural models. *Communications in Statistics: Theory and Methods* **25**, 2319-2341.
7. Arellano – Valle R. B., Bolfarine, H. y Vilca – Labra (1998). Elliptical functional models. *Journal of Multivariate Analysis* **65**, 36-57.
8. Arellano – Valle, R. B., Galea, M. y Iglesias, P. (1999). Bayesian analysis in elliptical linear regression models. *Journal of the Chilean Statistical Society* **16** y **17**, 59 – 104.
9. Arellano – Valle, R. B., Galea – Rojas, M. y Iglesias, P. (2000). Bayesian sensitivity analysis in elliptical linear regression models. *Journal of Statistical Planning and Inference* **86**, 175 – 199.
10. Arellano – Valle, R. B., Galea, M. y Iglesias, P. (2001). Bayesian analysis in elliptical CAPM in the Chilean Stock Market
11. Arellano – Valle, R. B., Iglesias, P. y Vidal, I. (2002a). Bayesian inference for elliptical linear models: Conjugate analysis and model comparison. *Bayesian Statistics* **7**.
12. Arellano – Valle, R. B., Del Pino, G. y Iglesias, P. (2002b). Bayesian analysis for the spherical linear model. Preprint.
13. Arellano – Valle, R. B., Bolfarine, H. y Iglesias, P. (2003). Systematic Risk Estimator using Comparative Calibration Models: An Application to the Chilean Stock Market bayesian analysis. Preprint.
14. Barnett V. D. (1969) Simultaneous pair wise linear structural relationships. *Biometrics* **25**, 129 – 142.
15. Basu, S. y Mukhopadhyay, S. (2000). Bayesian analysis of binary regression using symmetric and asymmetric links. *The Indian Journal of statistics* **62 B**, 372 – 387.
16. Blackwell, D. y McQueen, J. B. (1973). Ferguson distributions via Pólya urn schemes, *The Annals of Statistics* **1**, 353 – 355.
17. Bolfarine, H. y Cordani L. K. (1993). Estimation of a structural linear regression model with a known reliability ratio *Annals of the Institute of Statistical Mathematics* **3**, 531 – 540.

18. Bolfarine, H. y Arellano – Valle R. B. (1994). Robust modeling in measurement error models using the Student – t distribution. *Brazilian Journal of Probability and Statistics* **8**, 67 – 84.
19. Bolfarine, H. y Galea – Rojas, M (1995). Structural comparative calibration using the EM algorithm. *Journal of Applied Statistics* **22**, 277 – 292.
20. Bolfarine, H. y Galea – Rojas M. (1996). On Structural Comparative Calibration under a t-Model. *Computational Statistics* **11**, 63-85.
21. Bolfarine, H. y Arellano – Valle R. B. (1998). Weak nondifferential measurement error models. *Statistics and Probability Letters* **40**, 279 – 287.
22. Bolfarine, H., Cabral, C. R. B. y Paula, G. A. (2002). Distance tests under nonregular conditions: applications to the comparative calibration model. *Journal of Statistical Computation and Simulation* **72**, 231-244.
23. Box, G. E. P. y Tiao, G. C. (1973). Bayesian inference in statistical analysis. Reading, M A: Addison – Wesley.
24. Branco, M. D., Bolfarine, H. y Iglesias, P. (1998). Bayesian calibration under a Student – t model. *Computational statistics* **13**, 319 – 338.
25. Branco, M. D., Bolfarine, H., Iglesias, P. y Arellano –Valle, R. B. (2000). Bayesian analysis of the calibration problem under elliptical distributions. *Journal of Statistical Planning and Inference* **90**, 69 – 85.
26. Branco, M. D., Bolfarine, H., Iglesias, P. y Arellano –Valle, R. B. (2003). Bayesian and classical solutions for binomial cytogenetic dosimetry problem. *Journal of Data Science* **1**, 65 – 82.
27. Brown, P. J. (1982). Multivariate Calibration. *Journal of the Royal Statistical Society B*. **44**, 287 – 321.
28. Brown, P. J. (1993). Measurement, regression and calibration. Oxford, Oxford University Press.
29. Bush C. y MacEachern S. (1996). A semiparametric bayesian model for randomized block design. *Biometrika* **83**, 275 – 285.
30. Carroll, R.J., Freedman, L. S. y Pee, D. (1997). Design aspects of calibration studies in nutrition with analysis of missing data in linear measurement error models. *Biometrics* **53**, 1444 – 1451.
31. Carroll, R. J., Roeder, K. y Wasserman, L. (1999). Flexible parametric measurement error models. *Biometrics* **53**, 44 – 54.
32. Carter, R. (1981). Restricted maximum likelihood estimation of bias and reliability in the comparison of several measuring methods. *Biometrics* **37**, 733 – 741.
33. Chib, S., Jammalamadaka, S. R. y Tiwari, R. C. (1988). Bayes prediction in regressions with Elliptical errors. *Journal of Econometrics* **38**, 349 – 360.
34. Chipkevitch, E., Nishimura, R., Tu, D. y Galea – Rojas, M. (1996). Clinical measurement of testicular volume in adolescents: Comparison of reliability of 5 methods. *Journal of Urology* **156**, 2050 – 2053.
35. Clayton, D. G. (1991). A Monte Carlo method for bayesian inference in frailty models. *Biometrics* **47**, 467 – 485.
36. Dellaportas, P. y Stephens, D. A. (1995). Bayesian analysis of errors in variables regression models. *Biometrics* **51**, 1085 – 1095.

37. Chen, M. H. y Dey, D. (1998). Bayesian modeling of correlated binary responses via scale mixture of multivariate normal link functions. *Sankia. The Indian Journal of Statistics* **60**, 322 – 343.
38. Diebolt y Robert, C. (1994). Estimation of finite mixtures distributions through bayesian sampling. *Journal of the Royal Statistical Society* **56**, 363 – 375.
39. Dunn, G. (1992). *Design and Analysis of Reliability: The statistical evaluation of measurement errors*. Edward Arnold. New York.
40. Eisenhart, G. (1939). The interpretation of certain regression methods and their use in biological and industrial research. *Annals of Mathematical Statistics*. **10**, 162 – 186.
41. Eno D. (1999), *Noninformative Prior Bayesian Analysis for Statistical Calibration Problems*, Doctoral tesis, Virginia Polytechnic Institute and State University.
42. Escobar, M. D. y West, M. (1995). Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association* **85**, 378 – 409.
43. Escobar, D. E. y West, M. (1998). Computing nonparametric hierarchical models. In *Practical Nonparametric and Semiparametric Bayesian Statistics* (P.Müller, B. Vidakovic), Springer, New York, p 1 – 22.
44. Fang, K. T.; Kotz, S. y Ng, K. W. (1990). *Symmetric multivariate and related distributions*. London: Chapman and Hall.
45. Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *The Annals of Statistics* **1**, 209 – 230.
46. Fernández, C. y Steel, M. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association* **93**, 359 – 371.
47. Fuller, W. A. (1987). *Measurement error models*. New York, John Wiley y Sons.
48. Galea – Rojas, M., Paula, G. A. y Bolfarine, H. (1997). Local influence in elliptical linear regression models. *The Statistician* **46**, 71 – 79.
49. Galea – Rojas, M., Bolfarine, H. y de Castro, M. (2002a). Local influence in comparative calibration models. *Biometrical Journal* **44**, 59 – 81.
50. Galea – Rojas, M., Bolfarine, H. y Vilca, F. (2002b). Influence in comparative calibration models. *Biometrical Journal* **44**, 59 – 81.
51. Galea – Rojas, M. e Iglesias, P. (2003). Inference and Diagnostics in elliptical comparative calibration models. Preprint.
52. Gelfand A. y Smith, A. F. M. (1990) Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association* **85**, 398-409.
53. Geweke, J. (1993). Bayesian treatment of the independent Student – t linear model. *Journal of Applied Econometrics* **8**, 519 – 540.
54. Grubbs, F. E. (1948). On estimating precision of measuring instruments and product variability. *Journal of the American Statistical Association* **43**, 243 – 264.
55. Grubbs, F. E. (1973). Errors of measurements, precision, accuracy and the statistical comparison of measuring instruments. *Technometrics* **15**, 53 – 66.
56. Gu, C. (1990). Adaptive Spline smoothing in non-gaussian regression models. *Journal of the American Statistical Association* **85**, 801 – 807.
57. Higdom, R. y Schafer, D. W. (1999). Maximum likelihood computations for regression with measurement error. *Statistical Computing and Data Analysis* **35**, 283 – 299.
58. Hoadley, B. (1970). A bayesian look at inverse linear regression. *Journal of the American Statistical Association* **65**, 356 – 369.

59. Hunter, W. G. y Lamboy, W. F. (1981). A bayesian analysis of the linear calibration. *Technometrics* **23**, 323 – 350.
60. Jammalamadaka, S. R., Tiwari, R. C. y Chib, S. (1987). Bayes prediction in the linear model with Spherically symmetric errors. *Statistics and Probability Letters* **24**, 39 – 44.
61. Kelker, D. (1970). Distribution theory of spherical distributions and location – scale parameter generalization. *Sankhya A* **32**, 419 – 430.
62. Kelly, G. (1984). The influence function in the errors in variables problem. *The Annals of Statistics* **12**, 87 – 100.
63. Kelly, G. (1985). Use of the structural equations model in assessing the reliability of a new measurement technique. *Applied Statistics* **34**, 258 – 263.
64. Kiefer, J. y Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many nuisance parameters. *Annals of Mathematical Statistics* **27**, 886 – 906.
65. Krutchkoff, R. C. (1967). Classical and inverse regression methods of calibration. *Technometrics* **9**, 425 – 439.
66. Küchenhoff, H. y Carroll, R. J. (1997). Biased in segmented regression with errors in predictors. *Statistics in Medicine* **16**, 169 – 188.
67. Laird, N. (1978). Nonparametric maximum likelihood estimation of a mixing distribution. *Journal of the American Statistical Association* **73**, 805 – 811.
68. Lindley, D. V. y Smith, A. F. M. (1972). Bayes estimates for the linear model (with discussion). *Journal of the Royal Statistical Society B* **34**, 1 – 42.
69. Lindsay, B. G. y Lesperance, M. L. (1995). A review of semiparametric mixtures models. *Journal of Statistical Planning and Inference* **47**, 29 – 39.
70. Lu, Y., Ye, K., Mathur, A., Hui, S., Fuerst, T. y Genant, H. (1997). Comparative calibration without a gold standard. *Statistics in Medicine* **16**, 1889 – 1905.
71. MacEachern, S. (1994). Estimating normal means with a conjugate style Dirichlet process prior, *Communications in Statistics: Simulation and Computation* **23**, 727 – 741.
72. Mengersen, K. y Robert, C. (1993). Testing for mixtures: a Bayesian entropic approach. Technical Report # 9340. Institut National de la Statistique et des Etudes Economique, Universite de Rouen.
73. Müller, P., Erkanli, A. y West, M. (1996). Bayesian curve fitting using multivariate normal mixtures. *Biometrika* **83**, 67 – 79.
74. Müller, P. and Quintana, F.A. (2004). Nonparametric Bayesian Data Analysis. *Statistical Science*, **19(1)**, 95-110.
75. Neyman, J. y Scott, E.L. (1948). Consistent estimates based on partially consistent observations. *Econometrica* **16**, 1 – 32.
76. Nobile, A. (1994). Bayesian analysis of finite mixture distributions. PhD Dissertation. Department of Statistics, Carnegie Mellon University.
77. Osborne, C. (1991). Statistical calibration: A review. *International Statistical Review*. **59**, 309 – 336.
78. Osiewalski, J. And Stell, M. (1993). Robust bayesian inference in Elliptical regression models. *Journal of Econometrics* **57**, 345 – 363.

79. Raftery, A.E. (1995). Hypothesis testing and model selection via posterior simulation. W. Gilks, S. Richardson and D. J. Spiegelhalter eds. Practical Markov Chain Monte Carlo. London. Chapman and Hall.
80. Richardson, S. y Gilks, W. R. (1993). A bayesian approach to measurement error problems in epidemiology using conditional independence models. *American Journal of Epidemiology* **138**, 430 – 442.
81. Roeder, K. y Wasserman, L. (1995). Practical bayesian density estimation using mixtures of normal. Technical report # 633. Department of Statistics, Carnegie Mellon University.
82. Roeder, K., Carroll, R. J. y Lindsay, B. G. (1996). A non parametric mixture approach to case – control studies with errors in variables. *Journal of the American Statistical Association* **91**, 722 – 732.
83. Salini, S., Tiano, A. y Zirilli, A. (2002). Statistical Calibration by means of Kalman Filter. Convenio SIS 2002. Dipartimento di Informatica e Sistemistica, Universidad Cattolica del Sacro Cuore di Milano.
84. Salini, S., Zirilli, A. y Tiano, A. (2002). Multivariate Calibration by means of Kalman Filter. Convenio SIS 2002. Dipartimento di Informatica e Sistemistica, Universidad Cattolica del Sacro Cuore di Milano.
85. Schafer, D. W. (2001). Semiparametric maximum likelihood for measurement error model regression. *Biometrics* **57**, 53 – 61.
86. Skoog, West, Holler y Crouch. (2001). 7ª Ed. McGraw-Hill. Madrid. *Química Analítica*.
87. Staniswalis, J. G. (1989). On the kernel estimate of a regression function in likelihood based models. *Journal of the American Statistical Association: Theory and Methods*, **84** (405), 276 – 283.
88. Stephanski, L. A. y Carroll, R. J. (1990). Score test in generalized linear measurement error models. *Journal of the Royal Statistical Society B*, **52**, 345 – 359.
89. Stephens, D. A. y Dellaportas, P. (1992). Bayesian analysis for generalized linear models with covariate measurement error. *Bayesian Statistics 4*, Ed. J. M. Bernardo, J. Berger, A. P. Dawid y A. F. M. Smith. 813 – 820. Oxford. Oxford University Press.
90. Sundberg, R. (1999). Multivariate calibration - direct and indirect regression methodology (with discussion). *Scandinavian Journal of Statistics* **26**, 2,161-207.
91. Vidal, I., Iglesias, P. y Galea, M. (2005). Influential Observations in Functional Measurement Error Models. Preprint.
92. Vilca – Labra, R. B., Arellano – Valle, R. B. y Bolfarine, H. (1998). Elliptical functional models. *Journal of Multivariate Analysis* **65**, 36 – 57.
93. Wellman, M. J. y Gunst, R. F. (1991). Influence diagnostic for linear measurement errors models. *Biometrika* **78**(2), 373 – 380.
94. West, M. (1992) Modelling with mixtures (with discussion). J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith Eds. Oxford. Oxford University press. *Bayesian Statistics 4*, 503 – 524.
95. West, M. y Turner, D. A. (1994). Deconvolution of mixtures in analysis of neural synaptic transmission. *Statistician* **43**, 31 – 43.
96. Zellner, A. (1976). Bayesian and non – Bayesian analysis of the regression model with multivariate Student – t error term. *Journal of the American Statistical Association* **71**, 400 - 405.

97. Zhang, S. y Karunamuni, R. J. (1997). Bayes and empirical bayes estimation with errors in variables. *Statistics and Probability Letters* **33**, 23 – 34.
98. Zhao, Y. y lee, A. H. (1996). A simulation study of estimators for generalized linear measurement error models. *Journal of Statistical Computation and Simulation* **54**, 55 – 74.