



UNIVERSIDAD NACIONAL DE COLOMBIA

# Ajuste de un modelo jerárquico desde un enfoque Bayesiano

Jessica María Rojas Mora

Universidad Nacional de Colombia  
Facultad de Ciencias, Escuela de Estadística  
Medellín, Colombia  
2019



# Ajuste de un modelo jerárquico desde un enfoque Bayesiano

Jessica María Rojas Mora

Tesis o trabajo de grado presentada(o) como requisito parcial para optar al título de:  
**Magister en Ciencias Estadísticas**

Director(a):  
Ph.D., Isabel Cristina Ramírez Guevara

Universidad Nacional de Colombia  
Escuela de Estadística  
Medellín, Colombia  
2019



**Dedicatoria:**

A Dios...

A la memoria de mi padre Iván Rojas  
Álvarez

y la mejor madre del mundo Ana Berta  
Mora López



# Agradecimientos

El autor agradece a Isabel Cristina Ramírez Guevara y todos los profesores de la Escuela de Estadística de la Universidad Nacional de Colombia, sede Medellín.





## Resumen

Los modelos jerárquicos Bayesianos son utilizados en la modelación de datos en diferentes áreas en las cuales las estructuras jerárquicas se reflejan a través de efectos aleatorios. La distribución de probabilidad considerada como elección natural para el modelamiento de los efectos aleatorios es la Normal. Como distribución a priori para el parámetro de escala regularmente se utiliza Gamma-inversa  $(\epsilon, \epsilon)$  con valores de  $\epsilon$  muy pequeños y esta selección ha tenido críticas. Algunos autores comentan que se pueden obtener distribuciones posteriores inestables, lo cual ocasiona que la inferencia no sea robusta. Distribuciones como *half*-Cauchy, SBeta2 y Uniforme son consideradas como alternativas por diversos autores para modelar el parámetro de escala. En el presente trabajo de investigación se examinó el comportamiento de las estimaciones de un modelo jerárquico para el cual se asumió efectos aleatorios distribuidos *t*-Student y parámetro de escala distribuidos *half*-Cauchy, SBeta2 y Uniforme. Se llevó a cabo un estudio de simulación para evaluar el comportamiento del error de estimación de los efectos del modelo. Con base a los resultados obtenidos, y bajo los diferentes escenarios en consideración, fue posible examinar el encogimiento de los parámetros a posteriori del modelo y se pudo establecer que en presencia de valores atípicos, esta medida es menor cuando los efectos se modelan con una distribución *t* de Student comparados con los obtenidos cuando se le asocia a los efectos una distribución Normal bajo las mismas distribuciones a priori para el parámetro de escala.

**Palabras clave:** Inferencia Bayesiana, Parámetro de escala, Modelos Jerárquicos, Distribución *t* de Student.

# Abstract

Bayesian hierarchical models are used in data modeling in different areas in which hierarchical structures are reflected through random effects. The probability distribution considered as a natural choice for random effects modeling is Normal. As a priori distribution for the scale parameter Inverse-gamma  $(\epsilon, \epsilon)$  is regularly used with very small values of  $\epsilon$  and this selection has been criticized. Some authors comment that unstable later distributions can be obtained, which causes the inference not to be robust. Distributions such as half-Cauchy, SBeta2 and Uniform are considered as alternatives some authors to model the scale parameter. In the present research work we examined the behavior of the estimates of a hierarchical model for which we assumed random effects distributed Student's t and scale parameter distributed half-Cauchy, SBeta2 and Uniform. A simulation study was conducted to evaluate the behavior of the estimation error of the effects of the model. Based on the results obtained, and under the different scenarios under consideration, it was possible to examine the shrinkage of the parameters a posteriori of the model and it could be established that in the presence of atypical values, this measure is lower when the effects are modeled with a Student's t-distribution compared with those obtained when a Normal distribution is associated to the effects under the same a priori distributions for the scale parameter.

**Keywords:** Bayesian Inference, Scale Parameter, Hierarchical Models, Student's t-distribution

# Contenido

<b>Agradecimientos</b>	<b>vii</b>
<b>Resumen</b>	<b>ix</b>
<b>Abstract</b>	<b>x</b>
<b>Lista de figuras</b>	<b>xi</b>
<b>Lista de tablas</b>	<b>xv</b>
<b>1. Introducción</b>	<b>1</b>
<b>2. Marco Teórico</b>	<b>5</b>
2.1. Modelos jerárquicos . . . . .	5
2.1.1. Modelo lineal jerárquico . . . . .	5
2.2. Familia de distribuciones a priori para parámetros de escala . . . . .	7
2.3. Distribuciones de probabilidad para modelar los efectos aleatorios en un modelo jerárquico . . . . .	11
2.4. Metodología Monte Carlo por cadenas de Markov . . . . .	13
2.4.1. Aspectos generales sobre los métodos MCMC . . . . .	13
2.4.2. Chequeo de convergencia . . . . .	14
<b>3. Estudio de simulación</b>	<b>16</b>
3.1. Diseño del estudio de simulación . . . . .	16
3.2. Resultados obtenidos del estudio de simulación . . . . .	20
<b>4. Aplicación</b>	<b>41</b>
4.1. Aplicación de un modelo jerárquico normal . . . . .	41
4.1.1. Aplicación con todos los $J = 8$ colegios . . . . .	42
4.1.2. Aplicación con los tres primeros colegios . . . . .	46
<b>5. Conclusiones y Perspectivas futuras</b>	<b>49</b>
5.1. Conclusiones . . . . .	49
5.2. Perspectivas futuras . . . . .	50

<b>A. Anexo: Error global de estimación por escenario</b>	<b>51</b>
A.1. Error global de estimación para el escenario con efectos de magnitud similar	51
A.2. Error global de estimación para el escenario con algunos valores atípicos moderados . . . . .	53
A.3. Error global de estimación para el escenario de un gran valor atípico . . . . .	56
A.4. Error de estimación individual para el escenario con algunos valores atípicos moderados . . . . .	58
A.5. Error de estimación individual para todos los escenarios cuando $\nu = 5$ y se tienen $J = 8$ grupos . . . . .	59
A.6. Error de estimación individual en $S_I, S_{II}, S_{III}$ para $J = 10$ grupos . . . . .	60
<b>Bibliografía</b>	<b>61</b>

# Lista de Figuras

3-1. Error global de estimación para $J = 3$ en $S_I$ . . . . .	20
3-2. Error global de estimación para $J = 6$ en $S_I$ . . . . .	21
3-3. Error global de estimación para $J = 10$ en $S_I$ . . . . .	21
3-4. Error global de estimación para $J = 3$ en $S_I$ con grados de libertad continuos	22
3-5. Error global de estimación para $J = 10$ en $S_I$ con grados de libertad continuos	22
3-6. Traza y densidad para una de las cadenas simuladas del efecto aleatorio $\alpha$ con $J = 3$ grupos en $S_I$ . . . . .	23
3-7. Promedio móvil para una de las cadenas simuladas del efecto aleatorio $\alpha$ . .	24
3-8. Error global de estimación para $J = 3$ en $S_{II}$ . . . . .	26
3-9. Error global de estimación para $J = 4$ en $S_{II}$ . . . . .	27
3-10. Error global de estimación para $J = 5$ en $S_{II}$ . . . . .	27
3-11. Error global de estimación para $J = 3$ en el $S_{II}$ con grados de libertad continuos	28
3-12. Error global de estimación para $J = 7$ en $S_{II}$ . . . . .	28
3-13. Error global de estimación $J = 8$ en $S_{II}$ . . . . .	29
3-14. Error global de estimación $J = 10$ en $S_{II}$ . . . . .	29
3-15. Traza y densidad para una de las cadenas simuladas para $\alpha$ con $J = 3$ grupos en $S_{II}$ . . . . .	31
3-16. Promedio móvil para una de las cadenas simuladas para $\sigma_\alpha^2$ con $J = 3$ grupos en $S_{II}$ . . . . .	31
3-17. Error global de estimación para $J = 3$ en el escenario 3 . . . . .	35
3-18. Error global de estimación para $J = 4$ en $S_{III}$ . . . . .	35
3-19. Error global de estimación para $J = 5$ en $S_{III}$ . . . . .	36
3-20. Error global de estimación para $J = 3$ en $S_{III}$ . . . . .	36
3-21. Error global de estimación para $J = 4$ en $S_{III}$ con grados de libertad continuos	37
3-22. Error global de estimación para $J = 8$ en $S_{III}$ con grados de libertad continuos	37
3-23. Traza y densidad para una de las cadenas simuladas para $\alpha$ con $J = 3$ grupos en $S_{III}$ . . . . .	38
3-24. Promedio móvil para una de las cadenas simuladas para $\alpha$ con $J = 3$ grupos en $S_{III}$ . . . . .	39
4-1. Histograma de la distribución posterior de $\sigma_\alpha$ con distribución a priori para $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, 25^2)$ . . . . .	42

---

4-2.	Histograma de la distribución posterior de $\sigma_\alpha$ con distribución a priori para $\sigma_\alpha^2 \sim SBeta2(1, 1, 25^2)$ . . . . .	43
4-3.	Histograma de la distribución posterior de $\sigma_\alpha$ con distribución a priori para $\sigma_\alpha^2 \sim Uniforme(0, 100)$ . . . . .	44
4-4.	Histograma de la distribución posterior de $\sigma_\alpha$ con distribución a priori para $\sigma_\alpha^2 \sim IG(0.001, 0.001)$ . . . . .	45
4-5.	Histograma de la distribución posterior de $\sigma_\alpha$ con distribución a priori para $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, 25^2)$ para $J = 3$ . . . . .	47
4-6.	Histograma de la distribución posterior de $\sigma_\alpha$ con distribución a priori para $\sigma_\alpha^2 \sim SBeta2(1, 1, 25^2)$ para $J = 3$ . . . . .	47
4-7.	Histograma de la distribución posterior de $\sigma_\alpha$ con distribución a priori para $\sigma_\alpha^2 \sim Uniforme(0, 100)$ para $J = 3$ . . . . .	48
4-8.	Histograma de la distribución posterior de $\sigma_\alpha$ con distribución a priori para $\sigma_\alpha^2 \sim IG(0.001, 0.001)$ para $J = 3$ . . . . .	48

# Lista de Tablas

3-1. Valores para $\sigma_j^2$ . . . . .	16
3-2. Valores para los efectos $\alpha_j$ para el escenario 1 . . . . .	17
3-3. Valores para los efectos $\alpha_j$ para el escenario 2 . . . . .	18
3-4. Valores para los efectos $\alpha_j$ para el escenario 3 . . . . .	18
3-5. Autocorrelaciones por rezagos para $J = 3$ con $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, \sigma_{\alpha_j}^2)$ . . . . .	23
3-6. Comparación de errores de estimación global, $G$ en $S_I$ . . . . .	25
3-7. Variación porcentual del error de estimación global en $S_I$ . . . . .	26
3-8. Error de estimación individual para $J = 10$ en $S_{II}$ . . . . .	30
3-9. Autocorrelaciones por rezagos para $J = 10$ con $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, \sigma_{\alpha_j}^2)$ . . . . .	30
3-10. Error de estimación individual para $J = 6$ en $S_{II}$ . . . . .	32
3-11. Variación porcentual del error de estimación global de $J = 8$ en $S_{II}$ . . . . .	32
3-12. Comparación de errores de estimación global para $S_{II}$ . . . . .	34
3-13. Autocorrelaciones por rezagos para $J = 3$ con $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, 25^2)$ . . . . .	38
3-14. Comparación de errores de estimación global, $G$ en $S_{III}$ . . . . .	40
4-1. Datos colegios Gelman (2006b) . . . . .	41
4-2. Resultados del ajuste para $J = 8$ y $\sigma_\alpha^2 \sim \text{SBeta2}(1/2, 1/2, 25^2)$ . . . . .	43
4-3. Resultados del ajuste para $J = 8$ y $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, 25^2)$ . . . . .	44
4-4. Resultados del ajuste para $J = 8$ y $\sigma_\alpha^2 \sim \text{Uniforme}(0, 100)$ . . . . .	45
4-5. Resultados del ajuste para $J = 8$ y $\sigma_\alpha^2 \sim \text{IG}(0.001, 0.001)$ . . . . .	46
A-1. Error de estimación global para todas la situación $S_I$ con $J = 3$ grupos . . . . .	51
A-2. Error de estimación global para todas la situación $S_I$ con $J = 4$ grupos . . . . .	51
A-3. Error de estimación global para todas la situación $S_I$ con $J = 5$ grupos . . . . .	52
A-4. Error de estimación global para todas la situación $S_I$ con $J = 6$ grupos . . . . .	52
A-5. Error de estimación global para todas la situación $S_I$ con $J = 7$ grupos . . . . .	52
A-6. Error de estimación global para todas la situación $S_I$ con $J = 8$ grupos . . . . .	52
A-7. Error de estimación global para todas la situación $S_I$ con $J = 9$ grupos . . . . .	53
A-8. Error de estimación global para todas la situación $S_I$ con $J = 10$ grupos . . . . .	53
A-9. Error de estimación global para todas la situación $S_{II}$ con $J = 3$ grupos . . . . .	53
A-10. Error de estimación global para todas la situación $S_{II}$ con $J = 4$ grupos . . . . .	54
A-11. Error de estimación global para todas la situación $S_{II}$ con $J = 5$ grupos . . . . .	54
A-12. Error de estimación global para todas la situación $S_{II}$ con $J = 6$ grupos . . . . .	54
A-13. Error de estimación global para todas la situación $S_{II}$ con $J = 7$ grupos . . . . .	54

<b>A-14</b>	Error de estimación global para todas la situación $S_{II}$ con $J = 8$ grupos . . .	55
<b>A-15</b>	Error de estimación global para todas la situación $S_{II}$ con $J = 9$ grupos . . .	55
<b>A-16</b>	Error de estimación global para todas la situación $S_{II}$ con $J = 10$ grupos . . .	55
<b>A-17</b>	Error de estimación global para todas la situación $S_{III}$ con $J = 3$ grupos . . .	56
<b>A-18</b>	Error de estimación global para todas la situación $S_{III}$ con $J = 4$ grupos . . .	56
<b>A-19</b>	Error de estimación global para todas la situación $S_{III}$ con $J = 5$ grupos . . .	56
<b>A-20</b>	Error de estimación global para todas la situación $S_{III}$ con $J = 6$ grupos . . .	57
<b>A-21</b>	Error de estimación global para todas la situación $S_{III}$ con $J = 7$ grupos . . .	57
<b>A-22</b>	Error de estimación global para todas la situación $S_{III}$ con $J = 8$ grupos . . .	57
<b>A-23</b>	Error de estimación global para todas la situación $S_{III}$ con $J = 9$ grupos . . .	57
<b>A-24</b>	Error de estimación global para todas la situación $S_{III}$ con $J = 10$ grupos . . .	58
<b>A-25</b>	Error de estimación individual para todas la situación $S_{II}$ con $J = 6$ grupos . . .	58
<b>A-26</b>	Error de estimación individual para todas la situación $S_{II}$ con $J = 7$ grupos . . .	59
<b>A-27</b>	Error de estimación con $J = 8$ para todos los escenarios y $\nu = 5$ . . . . .	59
<b>A-28</b>	Error de estimación con $J = 8$ para todos los escenarios y $\nu \sim$ Exponencial(0,001)	
	60	
<b>A-29</b>	Error de estimación individual para $J = 10$ en todos los escenarios con $\nu = 1$ . . . . .	60



# 1. Introducción

Los modelos jerárquicos se utilizan ampliamente para modelar los datos de respuesta obtenidos a partir de diseños de medidas repetidas, estudios longitudinales y experimentos aleatorios multinivel diseñados en cuadrado latino, parcelas divididas, bloques balanceados o desbalanceados con efectos aleatorios. Estos modelos presentan dos niveles, el nivel superior (nivel 1) consiste en el modelo de la variable de respuesta, mientras que en el siguiente nivel (nivel 2) consiste en modelos para los coeficientes aleatorios (efectos aleatorios). Se debe tener en cuenta que para las respuestas obtenidas de un diseño de medidas repetidas, un estudio longitudinal o parcelas divididas, los coeficientes aleatorios explican la heterogeneidad entre los sujetos, así como la correlación entre las observaciones recopiladas del mismo sujeto en diferentes momentos, pero para los datos obtenidos de un experimento aleatorio en el que los grupos se ven como una selección aleatoria de una población de grupos, los efectos aleatorios abarcan efectos específicos del grupo, así como entre la variación del grupo (Demirhan y Kalaylioglu, 2015). En Gelman (2006a) y Browne y Draper (2006) mediante un ejemplo de predicción se ilustran las fortalezas y limitaciones cuando se utilizan modelos jerárquicos.

Desde el enfoque Bayesiano se puede definir un modelo de probabilidad directo para los parámetros de interés, el teorema de Bayes permite combinar el conocimiento acerca del parámetro y la verosimilitud de los datos (Bolstad, 2004). En Zellner (1976) se presentan diferentes aproximaciones del análisis Bayesiano, entre las cuales encontramos análisis Bayesiano objetivo, subjetivo, robusto, Bayesiano-frecuentista y cuasi-Bayesiano. En los trabajos de Berger y Strawderman (1996) y Berger (2000) se presentan los principales aspectos que se deben tener en cuenta para realizar la elección de la familia de las distribuciones que se tomarán como distribuciones a priori, la cual incorpora la creencia acerca del parámetro de interés.

En el enfoque Bayesiano de modelos jerárquicos, la estructura jerárquica se amplía para incluir otra etapa en la que se dan distribuciones a priori a los componentes de varianza de los efectos aleatorios. Generalmente no hay suficiente conocimiento a priori sobre cuál podría ser la varianza del coeficiente aleatorio, y por lo tanto se consideran distribuciones a priori no informativas dejando que la verosimilitud domine la inferencia sobre las varianzas del nivel 2 en el modelo jerárquico. De todas las distribuciones a priori para modelar componentes de varianza consideradas en la literatura, la distribución gamma con parámetros de forma

y escala pequeños denotada por  $\text{Gamma}(\epsilon, \epsilon)$  ha sido comúnmente usada como distribución a priori para la precisión. La precisión corresponde al inverso del parámetro de varianza, lo cual es equivalente a modelar el parámetro de varianza con una distribución Gamma-inversa  $(\epsilon, \epsilon)$  con  $\epsilon \rightarrow 0$ , puesto que es una distribución conjugada a la distribución normal. Autores como Natarajan y McCulloch (1998), Gelman (2006b) y Berger (2006) exponen e ilustran los principales inconvenientes de utilizarla como distribución a priori porque pueden dar lugar a distribuciones posteriori propias pero con valores incorrectos o distribuciones posteriori impropias.

En la literatura encontramos autores como Lambert, Sutton, Burton, Abrams, y Jones (2005) los cuales comparan los efectos de 13 configuraciones de distribuciones a priori para componentes de varianzas en un modelo jerárquico de efectos aleatorios a través de un estudio de simulación utilizando WinBUGS. Gelman (2006b), Fúquene, Pérez, y Pericchi (2014) y Pérez, Pericchi, y Ramírez (2017) han propuesto nuevas clases de familia de distribuciones con propiedades más atractivas con las cuales se logra un mejor desempeño en el análisis a posteriori, contribuyendo a corregir los inconvenientes presentados cuando se modela el parámetro de escala con una distribución Gamma-Inversa. En el enfoque Bayesiano, Daniels (1999) estudió la distribución a priori denominada “*uniform shrinkage*” o encogimiento uniforme para el parámetro de escala en modelos jerárquicos, la cual se puede utilizar en modelos conjugados asignando una distribución a priori uniforme al parámetro de encogimiento (*shrinkage*). La estimación del parámetro de encogimiento introduce deliberadamente sesgos en el modelo para mejorar el rendimiento general, a menudo a costa de estimaciones individuales (Efron y Hastie, 2016). Esta distribución a priori fue sugerida por primera vez por Strawderman (1971) y posteriormente generalizada por Christiansen y Morris (1997).

Posteriormente, Gelman (2006b) presenta la familia de distribuciones a priori *folded noncentral-t* para los parámetros de desviación estándar en modelos jerárquicos. Además, propone utilizar la distribución Uniforme en el intervalo  $(0, A)$  como distribución a priori no informativa, puesto que, para  $A \rightarrow \infty$  se tiene una distribución posterior adecuada y recomienda la distribución *half-Cauchy* la cual está incluida en la familia *folded noncentral-t* como distribución a priori débilmente informativa para la desviación estándar en el nivel 2 del modelo jerárquico. Este autor aconseja no utilizar la distribución Gamma-inversa  $(\epsilon, \epsilon)$  porque la inferencia posterior es sensible a la elección de  $\epsilon$ . Los autores Gustafson, Hossain, y Macnab (2006) sugieren distribuciones a priori “conservadoras” para los componentes de varianza, las cuales deliberadamente dan mayor peso a valores más pequeños y resultan apropiadas para aquellos investigadores escépticos a la presencia de variabilidad en los parámetros del segundo nivel en modelos jerárquicos, es decir, en los efectos aleatorios, evitando inferir más estructura de la que realmente está presente.

Fúquene, Cook, y Pericchi (2009) ilustran el análisis Bayesiano para datos binomiales de

---

ensayos clínicos a partir de una alternativa práctica y robusta la cual denominan el modelo Cauchy/Binomial en la escala del Log-Odds. En los trabajos de Frühwirth-Schnatter y Wagner (2010), Frühwirth-Schnatter y Wagner (2012) utilizan como distribución a priori una normal centrada en cero para la raíz cuadrada del parámetro de escala en un modelo de espacio de estado. Por medio de un estudio de simulación y una aplicación real, muestran que esta distribución a priori es menos influenciada en la inferencia posterior cuando la varianza real del proceso es cercana a cero, comparada con el ajuste cuando se utiliza Gamma-inversa como distribución a priori. Autores como Polson y Scott (2012) argumentan que la distribución *half*-Cauchy debe utilizarse como a priori en vez de la distribución Gamma-inversa para un parámetro de escala de nivel superior en los modelos jerárquicos Bayesianos cuando se requiera una distribución a priori propia. A su vez, estos autores incorporan la distribución a priori *half*-Cauchy en la clase de las distribuciones a priori beta-inversa  $(p, q)$  para el parámetro de escala tal con valores  $p = q = 0,5$  para y derivan las expresiones para los momentos posteriores, las densidades marginales y el riesgo clásico de los estimadores de Bayes que surgen a partir de esta distribución a priori. Además, proporcionan información sobre el uso de esta distribución a priori en situaciones en las que se espera que los valores del vector de parámetro de interés estén muy dispersos.

Fúquene et al. (2014) proponen una nueva clase de distribuciones a priori hipergeométricas de colas pesadas que se dan como la convolución de una densidad condicional *t*-Student para el parámetro de localización y una distribución a priori SBeta2 para el cuadrado del parámetro de escala, con la cual obtienen una distribución marginal para el parámetro de localización en términos de funciones beta e hiperparámetros, con resultados analíticos simples. Sugieren las distribuciones a priori como un análisis objetivo robusto adecuado para los modelos lineales dinámicos.

Pérez et al. (2017) presentan la distribución SBeta2 como una familia flexible y manejable para modelar los parámetros de escala. Esta familia de distribuciones puede obtenerse en forma cerrada como una mezcla de distribución Gamma. La familia SBeta2 unifica y generaliza diferentes propuestas en la literatura Bayesiana y tiene numerosas ventajas teóricas y prácticas.

Sin embargo en la mayoría de las investigaciones anteriores asumen como distribución para los efectos aleatorios una distribución normal como elección natural, por lo cual, en el presente trabajo se examinará, vía simulación, el comportamiento del error en la estimación de los efectos aleatorios a nivel de grupo cuando se utiliza una distribución *t* de Student para los efectos aleatorios y diferentes alternativas para la distribución a priori asociada al parámetro de escala como la *half*-Cauchy, SBeta2 y Uniforme en un modelo jerárquico Normal.

Este documento se encuentra organizado así: en el capítulo 2 están los principales aspectos

teóricos asociados a los modelos jerárquicos Bayesianos, distribuciones a priori para los componentes de varianza en un modelo jerárquico y propuestas para los grados de libertad en la distribución  $t$ -Student. En el capítulo 3 se presenta un extenso estudio de simulación en el que examinamos y comparamos en cada escenario en consideración la sensibilidad de la estimación del error global e individual para los estimadores posteriores asociado a los efectos aleatorios. Además, se realiza el chequeo de la convergencia de las cadenas asociadas a las distribuciones posteriori. En el capítulo 4 se presenta una aplicación de datos. Finalmente, en el capítulo 5 se presentan las principales conclusiones.

## 2. Marco Teórico

A continuación se presentan algunos conceptos importantes para el desarrollo de este trabajo relacionados con modelos jerárquicos lineales y las principales distribuciones a priori para el modelamiento del parámetro de escala.

### 2.1. Modelos jerárquicos

Los modelos jerárquicos pueden ser usados para una variedad de metas inferenciales incluyendo inferencia causal, predicción y modelado descriptivo. Los modelos jerárquicos también son conocidos en la literatura bajo una gran variedad de nombres como modelos multinivel (Goldstein, 1986, 1987), modelos de coeficientes aleatorios (Longford, 1995), o como modelos de efectos mixto (Laird y Ware, 1982). La aplicación de estos modelos se puede dar en áreas, tales como: educación, biología, psicología, medicina, entre otras (Goldstein, 1987, 1995; Raudenbush y Bryk, 2002). Los modelos multinivel también se denominan jerárquicos, por dos razones diferentes: primero, por la estructura de los datos y segundo, por el modelo mismo, que tiene su propia jerarquía. Para ilustrar lo anteriormente expuesto, consideremos la siguiente situación: se requiere investigar las relaciones existentes entre alumnos y el grupo de clase al cual pertenece en una institución educativa. De acuerdo con el contexto se tiene que un alumno interactúa con el grupo de clase al cual pertenece. Así, los alumnos y los grupos de clase son definidos en niveles separados de la estructura jerárquica. De manera más general, consideramos que un modelo jerárquico es una regresión (un modelo lineal o un modelo lineal generalizado) en el que los parámetros (los coeficientes de regresión) reciben un modelo de probabilidad. Este modelo de segundo nivel tiene parámetros propios, los hiperparámetros del modelo, que también se estiman a partir de los datos. Las dos partes clave de un modelo jerárquico son coeficientes variables y un modelo para esos coeficientes variables (que pueden incluir predictores a nivel de grupo). La regresión clásica a veces puede acomodar coeficientes variables utilizando variables indicadoras. La característica que distingue los modelos jerárquicos de la regresión clásica está en el modelado de la variación entre grupos (Gelman y Hill, 2006).

#### 2.1.1. Modelo lineal jerárquico

Los modelos lineales jerárquicos constituyen una clase general de modelos que permiten modelar situaciones en las cuales se tiene datos que presentan una estructura jerárquica.

Así, en el análisis de datos que presenten estructuras jerárquicas, las técnicas estadísticas empleadas deben tomar en cuenta la estructura de tal manera que permitan simultáneamente examinar cada uno de los niveles de la estructura jerárquica, el cual puede ser representando a partir de un submodelo propio.

### Modelo jerárquico normal

En Gelman (2006b); Gelman et al. (2013) presentan un modelo de dos niveles en el cual las observaciones  $y_{ij}$  provienen de una distribución normal y se presentan efectos a nivel de grupo  $\alpha_j$ , en el cual  $y_{ij}$  representa las mediciones de los individuos  $i$ , ( $i = 1, 2, \dots, n_j$ ), de  $j$ , ( $j = 1, 2, \dots, J$ ) poblaciones diferentes:

$$\begin{aligned} y_{ij} &\sim N(\theta_j, \sigma_j^2), \quad j = 1, \dots, J \\ \theta_j &= \mu + \alpha_j, \\ \alpha_j &\sim N(0, \sigma_\alpha^2). \end{aligned} \tag{2-1}$$

con  $i = 1, \dots, n_j$  las unidades o el número de elementos asociados al  $j$ -ésimo grupo,  $\sigma_j$  denota la desviación estándar para las observaciones en el  $j$ -ésimo grupo y  $\sigma_\alpha$  denota las desviaciones a nivel de grupo. Así el modelo lineal puede expresarse como sigue:

$$y_{ij} = \mu + \underbrace{\alpha_j}_{\sim N(0, \sigma_\alpha^2)} + \underbrace{\epsilon_{ij}}_{\sim N(0, \sigma_j^2)} .$$

En el modelo (2-1) si se supone la varianza dentro de los grupos conocida, entonces podemos modelar de la siguiente manera:

$$\begin{aligned} y_j &\sim N(\theta_j, \sigma_j^2), \quad j = 1, \dots, J \\ \theta_j &= \mu + \alpha_j, \\ \alpha_j &\sim N(0, \sigma_\alpha^2). \end{aligned} \tag{2-2}$$

Con base al modelo (2-2) se realizó el estudio de simulación del presente trabajo.

En Gelman y Hill (2006, ver sección 19.6) presentan los principales aspectos que se tienen que tener en cuenta para realizar la elección de la distribución a priori para los parámetros de varianza a nivel de grupo en modelos jerárquicos, en particular presentan los principales aspectos del modelo básico de datos de dos niveles  $y_{ij}$  con coeficientes a nivel de grupo  $\alpha_j$  del modelo (2-1). A los hiperparámetros presentes en (2-1):  $\mu, \sigma_j^2, \sigma_\alpha$  se pueden asignar distribuciones a priori no informativas, por ejemplo  $p(\mu, \sigma_j) \propto 1$  o  $p(\mu, \log(\sigma_j)) \propto 1$ , con respecto a  $\sigma_\alpha$  se pueden usar distribuciones a priori sugeridas en la literatura Bayesiana como  $\sigma_\alpha^2 \sim \text{Gamma-inversa}(\epsilon, \epsilon)$ ,  $p(\sigma_\alpha^2) \propto \sigma_\alpha^{-2}$  o  $p(\sigma_\alpha) \propto \sigma_\alpha^{-1}$  entre otras. Para el desarrollo del presente trabajo se consideraron diferentes alternativas de distribuciones a priori para modelar el parámetro de escala  $\sigma_\alpha^2$ , tales como: la distribución Gamma-inversa  $(\epsilon, \epsilon)$ , la distribución *half*-Cauchy a partir de la distribución Beta 2 escalada y la distribución Uniforme.

## 2.2. Familia de distribuciones a priori para parámetros de escala

A continuación se presentan las distribuciones a priori más utilizadas para modelar parámetros de escala.

### Distribución a priori encogimiento - uniforme.

En Daniels (1999) comentan sobre lo importante y difícil que puede ser la elección de distribuciones a priori para las desviaciones en modelos jerárquicos bayesianos y de componentes de varianza. Sugiere la elección de distribuciones a priori no informativas en situaciones en las que se dispone de poca información a priori, sin embargo, se debe tener cuidado al usar estas distribuciones a priori, ya que muchas de ellas son inadecuadas y, por lo tanto, pueden llevar a distribuciones posteriores inadecuadas. Por lo anterior, el autor investigó la distribución a priori denominada “*uniform shrinkage*” o encogimiento uniforme para el parámetro de escala en modelos jerárquicos, la cual se puede utilizar en modelos conjugados asignando una distribución a priori uniforme al parámetro de encogimiento (*shrinkage*). Esta distribución a priori fue sugerida por primera vez por Strawderman (1971) y posteriormente generalizada por Christiansen y Morris (1997).

### Distribuciones folded-noncentral- $t$

Corresponden a distribuciones de mediciones absolutas y se obtienen a partir de la distribución  $t$  de Student tomando los valores absolutos de las variables (Psarakis y Panaretos, 1990). De acuerdo a la función de densidad de la distribución  $t$  de Student dada posteriormente en la **sección 2.3** para  $\mu = 0$ , la distribución *folded- $t$*  se reduce al caso especial de una distribución *half- $t$*  con función de densidad de probabilidad dada por:

$$p(x; \mu = 0, \sigma, \nu) = \frac{2\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu\sigma^2}} \left(1 + \frac{x^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}} \quad (2-3)$$

### Distribución a priori half-Cauchy.

Gelman (2006b) sugirió la *half-Cauchy* con moda tendiendo a 0 y valores para la escala establecida a un valor grande, o con la escala estimada a partir de datos en un entorno jerárquico aplicando una reparametrización multiplicativa de las expresiones en (2-1) y está dado por:

$$\begin{aligned} y_{ij} &\sim N(\mu + \xi\eta_j, \sigma_j^2), \quad i = 1, \dots, n_j, j = 1, \dots, J \\ \eta_j &\sim N(0, \sigma_\eta^2), \quad j = 1, 2, \dots, J \end{aligned} \quad (2-4)$$

con la reparametrización el parámetro  $\alpha_j$  en (2-1) es equivalente al producto  $\xi\eta_j$  en (2-4) y la desviación estándar de los efectos  $\sigma_\alpha$  en (2-1) es equivalente a  $|\xi| \sigma_\eta$  en (2-4), el cual fue propuesto para mejorar la convergencia del método Monte Carlo por cadenas de Markov. Se consideran los modelos *half*-Cauchy para los parámetros de escala que se calculan a partir de un pequeño número de grupos en los cuales las inferencias resultan ser sensibles a la elección de la distribución a priori débilmente informativa. La distribución *half-t* la cual corresponde a una distribución *t* de Student centrada en cero, para un efecto aleatorio asociado al modelo jerárquico que se está estudiando es una buena candidata para ser considerada como distribución a priori. Si el interés recaé en restringir los posibles valores del parámetro de varianza para que no tome valores grandes, Gelman (2006b) recomienda emplear la distribución *half*-Cauchy (*A*) como distribución a priori para la desviación estándar  $\sigma$ , cuya función de densidad de probabilidad está dada por:

$$p(\sigma_\alpha) \propto \left( 1 + \frac{1}{\nu} \left( \frac{\sigma_\alpha}{s_\alpha} \right)^2 \right)^{-(\nu+1)/2}$$

donde  $s_\alpha$  corresponde a la escala y  $\nu$  los grados de libertad en una distribución *t* de Student realizando una parametrización. Como caso especial tenemos la densidad uniforme impropia cuando  $\nu = -1$  y la *half*-Cauchy propia,  $p(\sigma_\alpha) \propto (\sigma_\alpha^2 + s_\alpha^2)^{-1}$  para  $\nu = 1$ .

## Distribuciones a priori beta-inversa y sus generalizaciones

En teoría de probabilidad la distribución beta-inversa también se conoce como distribución beta de segunda forma, distribución beta-prime o Pearson Tipo VI. Dubey (1970) la presenta como una distribución de probabilidad absolutamente continua. Una variable aleatoria beta-inversa es igual en distribución a la proporción de dos variables aleatorias con distribución gamma que tienen parámetros de forma  $p$  y  $q$ , respectivamente, junto con un parámetro de escala común. Consideremos la familia de distribuciones a priori beta-inversa para el cuadrado del parámetro de escala  $\sigma_\alpha^2$ :

$$p(\sigma_\alpha^2) = \frac{(\sigma_\alpha^2)^{q-1} (1 + \sigma_\alpha^2)^{-(p+q)}}{Beta(p, q)}, \quad (2-5)$$

donde  $Beta(p, q)$  denota la función beta, y  $p, q \in (0, +\infty)$ . La función beta también llamada integral de Euler de primera forma, está definida en términos de funciones Gamma como sigue:

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

La familia beta-inversa es en sí misma un caso especial de una nueva clase más amplia de distribuciones hipergeométricas beta-inversa. Considerando la familia de distribuciones a priori beta-inversa para el cuadrado del parámetro de escala,  $\sigma^2$ , la función de densidad de



probabilidad está dada por:

$$p(\sigma^2) = C^{-1}(\sigma^2)^{q-1}(\sigma^2 + 1)^{p+q} \exp \left\{ -\frac{s}{1 + \sigma^2} \right\} \left\{ \tau^2 + \frac{1 - \tau^2}{1 + \sigma^2} \right\}^{-1}$$

para  $p > 0, q > 0, \tau^2 > 0$ , y  $s \in \mathbb{R}$ . En el Apéndice 1 de Polson y Scott (2012) muestran los detalles asociados a el valor de  $C$ , el cual corresponde a

$$C = e^{-s} \text{Beta}(p, q) \Phi_1(q, 1, p + q, s, 1 - 1/\tau^2),$$

con  $\Phi_1$  como la función hipergeométrica degenerada de dos variables (Gradshteyn y Ryzhik, 1965). Además en Polson y Scott (2012) incorporan la distribución *half*-Cauchy aplicando la formula de cambio de variable en la densidad de una distribución a priori beta-inversa  $(p, q)$ . Por lo cual, una distribución a priori *half*-Cauchy para la desviación estándar corresponde a una distribución beta-inversa  $(p = 0.5, q = 0.5)$  para la varianza.

## Distribución Beta 2 Escalada (SBeta2)

La distribución SBeta2 es una versión escalada de la distribución beta-inversa con función de densidad dada en (2-5). Para la variable  $\psi = \frac{\bar{w}}{1-\bar{w}}$ , es conocida como el *odd ratio* y es tal que  $\bar{w} \sim \text{Beta}(\bar{w}|p, q)$ . La distribución beta-inversa para la variable  $\psi$  está definida como sigue:

$$\text{Beta2}(\psi|p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \frac{\psi^{q-1}}{(1+\psi)^{(p+q)}}, \quad \text{para } \psi > 0, p > 0, q > 0. \quad (2-6)$$

La distribución SBeta2 para  $\psi$  se define como sigue:

$$\text{SBeta2}(\psi|p, q, b) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)b} \frac{(\frac{\psi}{b})^{q-1}}{(1+\frac{\psi}{b})^{(p+q)}}, \quad \text{para } \psi > 0, p > 0, q > 0, b > 0. \quad (2-7)$$

El valor esperado y la varianza para la variable  $\psi$  estan dado por las siguientes expresiones:

$$E[\psi] = \frac{p}{q-1}b, \quad q > 1$$

$$\text{Var}[\psi] = \frac{p(p+q-1)}{(q-1)^2(q-2)}b^2, \quad q > 1$$

En una distribución SBeta2( $p, q, b$ ) el parámetro  $p$  controla el comportamiento en el origen, el parámetro  $q$  rige el comportamiento en la cola a derecha de la distribución de tal manera que mientras más pequeño sea más pesada es la cola y  $b$  controla la escala de la distribución. En Pérez et al. (2017), establecen que para realizar una inferencia robusta para el parámetro de varianza y precisión se puede utilizar la distribución SBeta2( $p, q, b$ ) con valores de  $q$  entre 0 y 1 dado que entre más pequeño sea dicho valor hace que la distribución sea de colas más

pesadas. Para evitar que la distribución del parámetro de escala se concentre alrededor de cero y así se presente un encogimiento hacia la media, sugieren seleccionar  $p$  entre 0.5 y 1. Adicional, sugieren valores de  $b$  en función de la varianza de los datos que se estén analizando de tal manera que  $P(\sigma^2 > a) = P(\theta > \frac{a}{a+b})$  para  $\theta \sim \text{Beta}(p, q)$  y se evalúa tal propuesta a partir de una de las propiedades de la distribución  $\text{SBeta2}(p, q, b)$  la cual establece que para  $p = q$  la mediana de la distribución  $\text{SBeta2}$  es el parámetro de escala  $b$ . Una de las razones por la que se considera la distribución  $\text{SBeta2}$  para el ajuste del parámetro de interés es por ser una distribución de colas más pesadas y por tanto se espera que la distribución posterior de la varianza no este con un pico alto cercano a cero.

### Distribución a priori Uniforme.

En los trabajos de Gelman (2006b) y Gelman y Hill (2006) establecen que para seleccionar la distribución Uniforme como distribución a priori para modelar los parámetros de varianza se debe ser explícito sobre la escala en la que se define la distribución, puesto que, para modelar datos de forma jerárquica no se puede descartar una varianza a nivel de grupo de cero. Entre otros aspectos comentan:

- Si se utiliza una distribución a priori Uniforme para  $\log(\sigma_\alpha)$ , la distribución posterior resultante tendrá una masa infinita, esto es,  $\log(\sigma_\alpha) \rightarrow -\infty$  y por tanto se tiene una distribución impropia.
- Si se define la distribución a priori en un conjunto compacto, por ejemplo en el intervalo cerrado  $[-A, A]$  para algún valor grande de  $A$ , entonces la distribución posterior dependería en gran medida del límite inferior del intervalo, es decir, de  $-A$ .
- Para  $\sigma_\alpha^2$  no recomiendan utilizar una distribución a priori no informativa uniforme, puesto que, presenta sobreestimación para valores altos de  $\sigma_\alpha$  y también requiere  $J \geq 4$  grupos para una distribución posterior adecuada.
- Una distribución a priori uniforme para  $\log(\sigma_\alpha)$  es equivalente a la distribución a priori  $p(\sigma_\alpha) \propto \sigma_\alpha^{-1}$  o  $p(\sigma_\alpha^2) \propto \sigma_\alpha^{-2}$ .

### Distribución a priori Gamma-inversa $(\epsilon, \epsilon)$ .

La familia de distribuciones Gamma-inversa es condicionalmente conjugada dado los otros parámetros en el modelo, es decir, si el parámetro  $\sigma_\alpha^2$  sigue una distribución a priori Gamma-inversa, entonces la distribución a posteriori condicional  $p(\sigma_\alpha^2 | \theta, \mu, y)$  es también Gamma-inversa Gelman (2006b, ver sección 3.1). El modelo Gamma-inversa  $(\epsilon, \epsilon)$  para  $\sigma_\alpha^2$  también puede expresarse en función de la distribución *inverse- $\chi^2$*  con escala  $s_\alpha^2 = \beta/\alpha$  y grados de libertad  $\nu_\alpha = 2\alpha$  (Gelman, 2003). La función de densidad de probabilidad Gamma-inversa

$(\epsilon, \epsilon)$  está dada por:

$$p(\sigma^2) = \frac{\epsilon^\epsilon}{\Gamma(\epsilon)} (\sigma^2)^{-(\epsilon+1)} \exp\left(\frac{-\epsilon}{\sigma^2}\right)$$

para  $\sigma > 0, \epsilon > 0$ . Esta distribución a priori para valores pequeños de sus parámetros como 1 o 0.01 o 0.001 refleja la falta de información dentro de una familia condicionalmente conjugada (Spiegelhalter, Thomas, Best, Gilks, y Lunn, 1994-2003). Una dificultad de esta distribución a priori cuando  $\epsilon \rightarrow 0$  es que la distribución posteriori produce una densidad posterior impropia, puesto que, para valores  $\epsilon$  muy pequeños, la distribución a posterior tiene un pico muy alto en cero, lo cual conlleva a que la estimación de los efectos del modelo jerárquico se van hacia la media y por tanto en presencia de un efecto que difiera mucho de la media entonces su estimación va a ser muy mala. Así, para valores bajos de  $\sigma_\alpha$  la inferencia resulta ser sensible a la elección del valor  $\epsilon$  y la distribución a priori se refleja como no informativa. Por lo cual, no se recomienda el uso de la distribución Gamma-inversa  $(\epsilon, \epsilon)$  como distribución a priori no informativa para el parámetro de varianza (Gelman, 2006b)[ver sección 4.3].

## 2.3. Distribuciones de probabilidad para modelar los efectos aleatorios en un modelo jerárquico

Entre las distribuciones de probabilidad utilizadas para los efectos aleatorios en modelos jerárquicos encontramos la distribución normal, pero esta no es la única, por ejemplo, la distribución  $t$  de Student propuesta por Student (1908), proporciona una extensión útil de la normal para el modelado estadístico de conjuntos de datos que implican errores con colas más largas de lo normal. La familia  $t$  de distribuciones  $t_\nu(\mu, \sigma)$  está caracterizada por tres parámetros: localización  $\mu$ , escala  $\sigma$  y los grados de libertad  $\nu$ , el cual determina la forma de la distribución. Las densidades  $t$  son simétricas y  $\nu$  debe pertenecer al intervalo  $(0, \infty)$ . La distribución  $t$  de Student puede ser interpretada como una mezcla de la distribución normal con una media común y varianza distribuida  $Scale-inv-\chi^2(\nu, \tau^2)$ . La función de densidad de probabilidad para una variable aleatoria  $X$  con distribución  $t$  de Student está dada como sigue:

$$p(x; \mu, \sigma, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{\pi\nu}\sigma} \left(1 + \frac{1}{\nu} \left(\frac{x - \mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}}, \quad (2-8)$$

donde  $\sigma$  establece la escala general de la distribución. Equivalentemente, la distribución puede ser expresada en términos de  $\sigma^2$ , el cuadrado del parámetro de escala como sigue:

$$p(x; \mu, \sigma^2, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{\pi\nu}\sigma^2} \left(1 + \frac{1}{\nu} \frac{(x - \mu)^2}{\sigma^2}\right)^{-\frac{\nu+1}{2}}. \quad (2-9)$$

En términos del parámetro de precisión,  $\tau = 1/\sigma^2$ , la función de densidad está definida por:

$$p(x; \mu, \tau, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{\tau}{\pi\nu}\right)^{1/2} \left(1 + \frac{\tau(x - \mu)^2}{\nu}\right)^{-\frac{\nu+1}{2}}. \quad (2-10)$$

Jackman (2009) presentan otras propiedades para la distribución  $t$  de Student las cuales se mencionan a continuación:

- Convencionalmente se escribe  $X \sim t_\nu(\mu, \sigma)$ , donde  $\mu$  es un parámetro de localización,  $\sigma > 0$  es un parámetro de escala, y  $\nu > 0$  es el parámetro de grados de libertad.
- La versión estandarizada de la desidad para la distribución  $t$  de Student se obtiene cuando  $\mu = 0$  y  $\sigma = 1$ .
- Cuando  $\nu \rightarrow \infty$ , la función de densidad tiende a la densidad de una distribución normal.
- Cuando  $\nu = 1$ , la función de densidad corresponde a la de una variable con distribución Cauchy.
- $E[X] = \mu$ , para  $\nu > 1$ .
- $Var[X] = \sigma^2 \frac{\nu}{\nu - 2}$ , para  $\nu > 2$ .

Villa y Walker (2014) presentan una distribución a priori objetiva para los grados de libertad de una distribución  $t$  de Student condicionada en los parámetros de localización y escala la cual está basada en funciones de pérdida, puesto que, distribuciones a priori incorrectas conducen a posteriores incorrectas, mientras que distribuciones a priori adecuadas pueden dominar la verosimilitud. Para el parámetro grados de libertad se pueden considerar distribuciones a priori discretas presentes en los trabajos de Jacquier, Polson, y Rossi (2004) los cuales propusieron una distribución uniforme discreta truncada dada por

$$\pi(\nu) \propto 1, \quad \text{para } 3 \leq \nu \leq 40,$$

lo cual es considerado por Fonseca, Ferreira, y Migon (2008) inapropiado, puesto que, la estimación de los grados de libertad resulta sensible al intervalo de truncamiento que se proponga. En el caso de distribuciones continuas Geweke (1993) propone una distribución a priori exponencial tal que el parámetro grados de libertad es considerado continuo y depende de un valor  $g$  estrictamente positivo, esto es,

$$\pi(\nu) \propto \exp(-g\nu),$$

con  $\nu > 0$  y Juárez y Steel (2010) propone utilizar una distribución Gamma con parámetros 2 y 1/100, esto es:

$$\pi(\nu) = \frac{\nu}{100} \exp(-\nu/100).$$

## 2.4. Metodología Monte Carlo por cadenas de Markov

En esta sección se presentan los principales aspectos de los métodos Monte Carlo por cadenas de Markov (*Markov chain Monte Carlo Methods*), los cuales denotaremos con la siguiente sigla MCMC, se utilizan ampliamente en la inferencia bayesiana y coresponden a métodos de simulación para generar muestras de distribuciones a posteriori y estimar cantidades de interés a posteriori.

### 2.4.1. Aspectos generales sobre los métodos MCMC

Los métodos MCMC fueron introducidos en la física a partir de una versión simplificada por Metropolis y Ulam (1949) y permiten a los investigadores modelos complejos y estimar con precisión las distribuciones posteriores correspondientes. En los métodos MCMC se simulan valores sucesivamente mediante un proceso iterativo, en el cual cada valor generado depende sólo del valor anterior simulado, por tanto se basan en la construcción de una cadena Markov que eventualmente converge hacia la distribución objetivo, la cual es denominada estacionaria o de equilibrio y corresponde a la distribución posterior (Gamerman y Lopes, 2006).

La idea básica de los métodos MCMC consiste en simular una cadena de Markov  $\theta^{(1)}, \theta^{(2)}, \dots$  cuya distribución estacionaria sea  $\pi(\theta|\mathbf{x})$  la cual corresponde a la distribución posterior. Si el algoritmo es implementado de forma correcta, la convergencia de la cadena de Markov está garantizada independiente de cuáles sean los valores iniciales. Se requiere simular la cadena de Markov un número elevado de iteraciones para aproximarse a la distribución estacionaria y eliminar los primeros valores simulados (iteraciones de *burn-in*) porque no están en estado estacionario.

#### Cadena de Markov

Una cadena de Markov  $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(t)}$  es una secuencia de variables aleatorias que cumplen la siguiente propiedad:

$$f(\theta^{(t+1)}|\theta^{(t)}, \theta^{(t-1)}, \theta^{(t-2)}, \dots, \theta^{(1)}) = f(\theta^{(t+1)}|\theta^{(t)}),$$

es decir, la distribución de  $\theta^{(t+1)}$  dado los valores previos  $\theta^{(t)}, \theta^{(t-1)}, \dots, \theta^{(1)}$  solo depende del estado inmediatamente anterior  $\theta^{(t)}$ .

#### Algoritmo MCMC

En Ntzoufras (2009) exponen los principales aspectos del algoritmo que se sigue para simular una cadena de Markov cuya distribución estacionaria es  $\pi(\theta|\mathbf{x})$ .

- Seleccionar un valor inicial  $\theta^{(0)}$ .

- Generar  $T$  valores hasta alcanzar la distribución de equilibrio. Puesto que, una vez que la distribución alcanza su equilibrio, un esquema de MCMC genera valores aleatorios dependientes de la distribución estacionaria correspondiente (Christian y Casella, 2004).
- Examinar la convergencia del algoritmo utilizando diagnósticos de convergencia, los cuales corresponden a pruebas estadísticas que intentan identificar los casos en los que no se logra la convergencia.
- Eliminar las primeras  $B$  observaciones.
- Considerar  $\{\theta^{(B+1)}, \theta^{(B+2)}, \dots, \theta^{(T)}\}$  como la muestra para el análisis posterior.
- Trazar la distribución posterior.
- Finalmente, obtener medidas de resumen de la distribución posterior (media, mediana, desviación estándar, cuantiles, correlaciones).

Los dos métodos de MCMC más populares son: el algoritmo Metropolis-Hastings Hastings (1970); Metropolis, Rosenbluth, Rosenbluth, Teller, y Teller (1953) y el muestreador de Gibbs Geman y Geman (1984).

### 2.4.2. Chequeo de convergencia

Dado que en esta tesis se utiliza la metodología Monte Carlo por cadenas de Markov para obtener las distribuciones posteriores de los parámetros, es importante examinar si los valores han convergido aproximadamente a la distribución estacionaria. Para ello es recomendable:

- Examinar a partir de gráficos como el algoritmo MCMC está explorado el espacio de estados. Para esto, se pueden graficar los valores simulados del parámetro de interés  $\theta^{(t)}$  versus  $t$  y analizar si se producen desviaciones de la estacionariedad. Se puede utilizar la traza de  $\theta$  para realizar la exploración de los estados.
- Verificar la convergencia de las medias de los valores simulados en el MCMC.
- Analizar si los valores simulados son aproximadamente una muestra de valores independientes e idénticamente distribuidos.

Cowles y Carlin (1996) presentan una revisión de algunos procedimientos propuestos en la literatura para estudiar la convergencia de la cadena. Una posibilidad es ejecutar el algoritmo varias veces comenzando en distintos valores iniciales y comprobar si el comportamiento de la distribución estacionaria es la misma. Para examinar si la cadena está explorando bien el espacio de estados, se grafican los valores simulados del parámetro de interés y se analizan si se presentan desviaciones de la estacionariedad.

El procedimiento utilizado para examinar la convergencia de las cadenas corresponde al planteado por Barrera y Correa (2008) en la sección de chequeo de convergencia, para el cual se utiliza el estadístico  $LM$  desarrollado por Kwiatkowski, Phillips, y Schmidt (1992) asociado a la prueba KPSS (Kwiatkowski-Phillips-Schmidt-Shin), el cual está implementado e incorporado en el paquete **tseries** (Trapletti y Hornik, 2018) del lenguaje R (Team, 2018). La prueba KPSS se usa para contrastar la hipótesis asociada a la estacionariedad alrededor de una tendencia determinista de una serie de tiempo observable. Este test contrasta las siguientes hipótesis:

$$\begin{cases} H_0 : & \text{La cadena de Markov ha alcanzado la distribución estacionaria} \\ H_1 : & \text{La cadena de Markov no ha alcanzado la distribución estacionaria} \end{cases}$$

Para el chequeo de convergencia de la cadena de Markov generada de la distribución condicional del parámetro de interés se procede como sigue:

- Dada la distribución condicional del parámetro de interés, se generan 10000 muestras a partir de esta distribución y los primeros 2000 valores simulados, (iteraciones de *burn-in*), se eliminan porque no están en el estado estacionario.
- Se calcula el valor del nivel del test KPSS y se observa el valor del parámetro de truncamiento y el  $p$ -valor asociado. Con base a estos aspectos se examina si se rechaza o no la hipótesis nula  $H_0$  de que la cadena de Markov haya alcanzado la distribución estacionaria.
- Si  $H_0$  no se rechaza, se verifica la correlación existente entre los valores generados para el parámetro de interés con distintos rezagos. Además, se realizan diagnósticos gráficos a partir del gráfico de promedios móviles y de densidad para la cadena de valores generados de la distribución condicional del parámetro de interés.

## 3. Estudio de simulación

En este capítulo se realizó el análisis de las estimaciones de los efectos aleatorios de un modelo jerárquico lineal en términos del impacto de diferentes distribuciones a priori para el parámetro de escala vía simulación.

### 3.1. Diseño del estudio de simulación

Se inició generando datos a partir de la siguiente expresión:

$$y_j \sim N(\mu + \alpha_j, \sigma_j^2), \quad j = 1, 2, \dots, J,$$

donde  $J$  corresponde al número de grupos, el parámetro de varianza  $\sigma_j^2$  se asume conocido con valores en el  $j$ -ésimo grupo dados en la Tabla **3-1**,  $\mu \sim N(0, 1+06)$  y  $\alpha_j$  denota el efecto aleatorio en el  $j$ -ésimo grupo. Los valores de los  $\sigma_j^2$  dados en la Tabla **3-1** se tomaron iguales a los trabajados en Ramírez (2016)[ver sección 4.1].

**Tabla 3-1.:** Valores para  $\sigma_j^2$

		Número de grupos $J$							
$j$	3	4	5	6	7	8	9	10	
1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	
2	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	
3	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	
4		1.8	1.8	1.8	1.8	1.8	1.8	1.8	
5			1.3	1.3	1.3	1.3	1.3	1.3	
6				1.4	1.4	1.4	1.4	1.4	
7					1.7	1.7	1.7	1.7	
8						1.5	1.5	1.5	
9							1.2	1.2	
10								1.6	

Los valores de los efectos aleatorios  $\alpha_j$  se consideraron fijos de acuerdo a los siguientes escenarios:



- (a) Escenario 1: todos los efectos con una magnitud similar. En este escenario, los valores de los  $\alpha_j$  se consideran fijos en el siguiente intervalo  $[-0.5, 0.5]$ .
- (b) Escenario 2: algunos valores atípicos moderados. En este escenario, los valores de los  $\alpha_j$  se establecen en tres intervalos de valores diferentes entre los cuales encontramos:  $(-0.5, 0.5)$ ,  $(-2, 2)$  y  $(-3, 3)$ . El número de valores atípicos asignados por grupos  $J$  varía. Para  $J = 3, 4, 5$  grupos, se asigna un solo valor atípico moderado. Cuando  $J = 6, 7$  se asignan dos valores atípicos moderados y para  $J \geq 8$  se asignan tres valores atípicos moderados.
- (c) Escenario 3: un efecto con un gran valor atípico. En este escenario, se considera solo un efecto tal que su valor absoluto supere 5 unidades, el resto de los valores de los  $\alpha_j$  se establecen en un intervalo de valores  $(-0.5, 0.5)$ .

Los valores de los efectos  $\alpha_j$  por cada escenario de interés corresponden a los trabajados en Pérez et al. (2017) y están dados en las tablas **3-2**, **3-3** y **3-4** respectivamente con sus varianzas asociadas. Además, los valores de los  $\sigma_j^2$  dados en la Tabla **3-1** no varían en los escenarios considerados. Posteriormente, teniendo en cuenta el número de grupos experimentales  $J = 3, 4, 5, 6, 7, 8, 9, 10$  se simulan  $m = 1000$  datos u observaciones para cada valor fijo de  $\alpha_j, \sigma_j^2$  y  $J$  por cada uno de los escenarios en consideración a partir de una distribución normal.

**Tabla 3-2.:** Valores para los efectos  $\alpha_j$  para el escenario 1

$j$	Número de grupos $J$							
	3	4	5	6	7	8	9	10
1	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3
2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
4		-0.1	0.4	0.4	0.4	0.4	0.4	-0.4
5			-0.5	-0.3	-0.5	-0.5	-0.5	-0.5
6				-0.3	-0.3	-0.3	-0.3	-0.3
7					-0.4	-0.4	-0.4	-0.4
8						0.5	0.5	0.5
9							-0.1	-0.1
10								0.2
$\sigma_{\alpha_j}^2$	0.0933	0.0667	0.1500	0.1350	0.1300	0.1564	0.1375	0.1782

**Tabla 3-3.:** Valores para los efectos  $\alpha_j$  para el escenario 2

$j$	Número de grupos $J$							
	3	4	5	6	7	8	9	10
1	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3
2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
3	2.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3
4		-2.1	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2
5			2.5	2.5	0.5	0.5	0.5	0.5
6				-2.6	2.5	2.5	0.2	0.2
7					-2.6	-2.6	2.5	-0.4
8						2.8	-2.6	2.5
9							2.8	-2.6
10								2.8
$\sigma_{\alpha_j}^2$	1.8033	1.2000	1.3320	2.6466	2.2461	2.8755	2.5200	2.2987

**Tabla 3-4.:** Valores para los efectos  $\alpha_j$  para el escenario 3

$j$	Número de grupos $J$							
	3	4	5	6	7	8	9	10
1	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3
2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
3	5.2	0.3	0.3	0.3	0.3	0.3	0.3	0.3
4		-5.1	0.4	0.4	0.4	0.4	0.4	0.4
5			-5.5	-0.4	-0.4	-0.4	-0.4	-0.4
6				-5.5	0.2	0.2	0.2	0.2
7					5.2	-0.1	-0.1	-0.1
8						-5.6	-0.2	-0.2
9							5.5	0.5
10								-5.6
$\sigma_{\alpha_j}^2$	9.4033	6.6500	6.4000	5.1800	3.8780	4.0392	3.4361	3.2898

Con las muestras obtenidas en cada uno de los escenarios de interés se realizó el siguiente ajuste, a partir del modelo dado por:

$$\begin{aligned}
 y_j &\sim N(\mu + \alpha_j, \sigma_j^2), \quad j = 1, 2, \dots, J. \\
 \alpha_j &\sim t_\nu(\mu_\alpha, \sigma_\alpha^2)
 \end{aligned}
 \tag{3-1}$$

Note que en este caso para los efectos  $\alpha_j$  se utiliza como distribución a priori la distribución  $t$  de Student con parámetro de localización  $\mu_\alpha$ , parámetro de escala  $\sigma_\alpha^2$  y  $\nu$  grados de libertad. Las distribuciones a priori candidatas utilizadas para modelar el parámetro de escala,  $\sigma_\alpha^2$  se asocian a las diferentes propuestas realizadas por Berger (2006), Gelman (2006b) y Pérez et al. (2017) las cuales están detalladas en la sección **2.2** y corresponden a las distribuciones Gamma-inversa( $\epsilon, \epsilon$ ) con  $\epsilon = 0.001$ , Uniforme( $0, 1/\epsilon$ ) con  $\epsilon = 0.01$ , y SBeta2( $p, q, b$ ) con valores de hiperparámetros dados por  $p = q = 0.5, p = q = 1$ . El parámetro de escala  $b$ , se asigna estableciendo la probabilidad  $p(\sigma^2 > a) = c$ , donde los valores de  $a$  y  $c$  se determinan de acuerdo a la varianza de los datos que se están modelando, de tal manera que la inferencia sea robusta Pérez et al. (2017)[**ver sección 2.3**]. Gelman (2006b) propone utilizar como distribución a priori para  $\sigma_\alpha$  en el modelo jerárquico normal la distribución *half*-Cauchy de parámetro 25, lo cual es equivalente a modelar  $\sigma_\alpha^2$  con una distribución SBeta2( $p = 0.5, q = 0.5, b = 25^2$ ) Pérez et al. (2017)[**ver sección 2.2**].

El modelo jerárquico lineal estudiado está dado por:

$$\begin{aligned} y_j &\sim N(\mu + \alpha_j, \sigma_j^2), \quad j = 1, 2, \dots, J. \\ \alpha_j &\sim t_\nu(\mu_\alpha, \sigma_\alpha^2), \\ \sigma_\alpha^2 &\sim f(\cdot), \end{aligned} \tag{3-2}$$

con  $\mu_\alpha \sim N(0, 1 + 06)$  y  $f(\cdot)$  denota la distribución a priori para el parámetro de escala  $\sigma_\alpha^2$  utilizadas y corresponden a:

- Gamma-inversa(0.001, 0.001).
- SBeta2(0.5, 0.5,  $25^2$ ).
- SBeta2(1, 1,  $25^2$ ).
- SBeta2(0.5, 0.5,  $\sigma^2(\alpha_j)$ ).
- SBeta2(1, 1,  $\sigma^2(\alpha_j)$ ).
- Uniforme(0, 100).

Con respecto a los grados de libertad se trataron bajo dos puntos de vistas diferentes, inicialmente se consideran grados de libertad discretos  $\nu = 1, 2, 3, 4, 5, 10, 30$  utilizados en Gelman et al. (2013) y el segundo punto de vista es considerarlos como un parámetro desconocido y utilizar las distribuciones a priori propuestas por Juárez y Steel (2010) y Geweke (1993), en las cuales proponen modelar los grados de libertad  $\nu$  en una distribución  $t$ -Student a partir de una distribución Gamma,  $\nu \sim \text{Gamma}(2, 0.1)$  y de una distribución Exponencial,  $\nu \sim \exp(0.001)$  respectivamente. Por cada escenario de interés se obtiene como resultado por cada número de grupo  $J$  en consideración 10000 simulaciones de las cadenas a posteriori de los parámetros del modelo jerárquico lineal a partir del método MCMC, utilizando

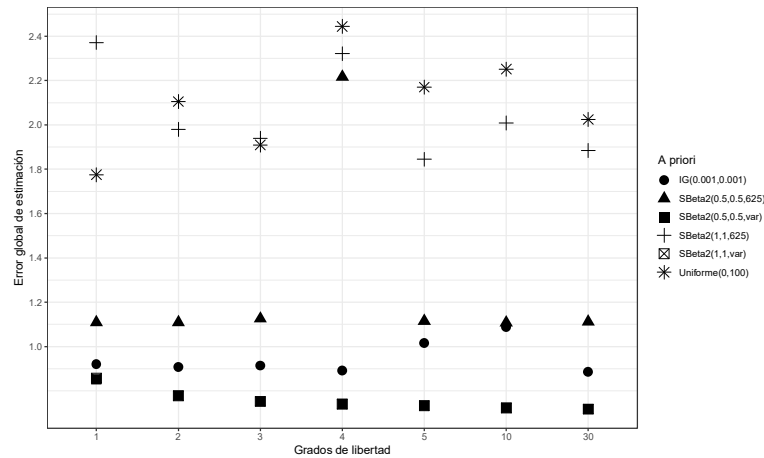
el software **jags** (Plummer, 2003) y la librería **R2jags** (Su y Yajima, 2015) del software estadístico Team (2018), estableciendo un quemado inicial de 2000. Para estudiar el efecto de las diferentes distribuciones a priori candidatas para el parámetro de escala,  $\sigma_\alpha^2$  se utilizó como medida el error global de estimación, el cual está dado por la siguiente expresión:

$$G = \sqrt{\frac{\sum_{j=1}^J \sum_{k=1}^{1000} (\alpha_j - \hat{\alpha}_{kj})^2}{1000}}, \quad (3-3)$$

donde  $\hat{\alpha}_{kj}$  denota la mediana de la distribución posterior de  $\alpha_j$  para la  $k$ -ésima muestra.

## 3.2. Resultados obtenidos del estudio de simulación

Por cada escenario de interés se presenta el comportamiento de la medida de error global de estimación,  $G$ , asociado a los efectos aleatorios estimados. Adicionalmente se compararon los valores de  $G$  obtenidos con los presentados en Pérez et al. (2017). Para el escenario 1 el cual denotaremos por  $S_I$ , se puede observar en las figuras **3-1**, **3-2** y **3-3** respectivamente el comportamiento de los valores de  $G$ , los cuales varían de acuerdo a la elección de la distribución a priori candidata, asumiendo grados de libertad discretos para realizar el ajuste del componente de varianza asociado al efecto aleatorio,  $\sigma_\alpha^2$ . Note que, para  $J = 3$  en  $S_I$  el comportamiento del error global de estimación ilustrado en la Figura **3-1** se puede clasificar en dos grupos: valores de  $G$  mayores o altos y valores de  $G$  menores o bajos. En el primer grupo encontramos las distribuciones a priori SBeta2(1, 1, 625) y Uniforme(0, 100).



**Figura 3-1.:** Error global de estimación para  $J = 3$  en  $S_I$

En el caso de  $J = 6$  en  $S_I$  ilustrado en la Figura **3-2**, el comportamiento de  $G$  de acuerdo con los ajustes de las distribuciones a priori candidatas para el parámetro de escala se puede clasificar en tres grupos: valores de  $G$  mayores o altos, valores de  $G$  promedios

y valores de  $G$  menores o bajos. En el primer grupo encontramos la distribución a priori  $SBeta2(1, 1, 625)$ , en el segundo grupos con valores muy parecidos encontramos las distribuciones  $SBeta2(1/2, 1/2, 625)$ , y  $Uniforme(0, 100)$ . Para el tercer grupo encontramos las distribuciones  $SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$ ,  $SBeta2(1, 1, \sigma_{\alpha_j}^2)$  y  $IG(0.001, 0.001)$ .

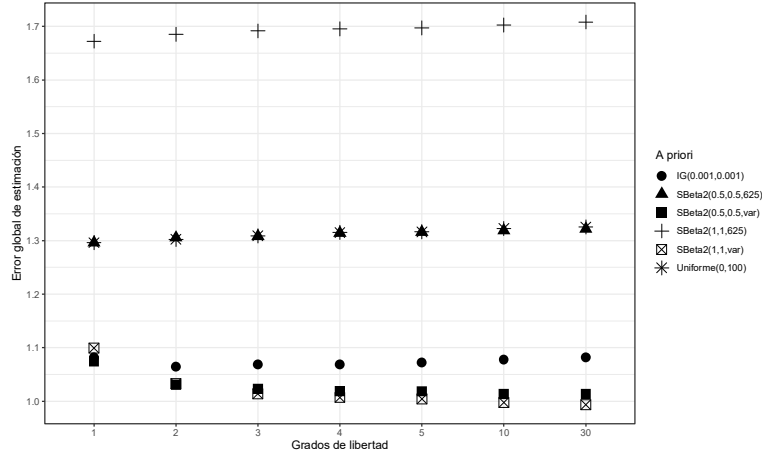


Figura 3-2.: Error global de estimación para  $J = 6$  en  $S_I$

Para  $J = 10$  grupos en  $S_I$ , se puede observar en la Figura 3-3 el comportamiento de  $G$ . Al igual que en el caso de  $J = 6$ , se enmarca en tres grupos, con la diferencia que los valores más bajos de  $G$  se obtienen cuando se realiza el ajuste  $\sigma_{\alpha}^2 \sim IG(0.001, 0.001)$ , pero la elección de esta distribución a priori no informativa pese a que conduce a una distribución posterior propia, es muy sensible para realizar inferencias y es comentado por Gelman (2006b).

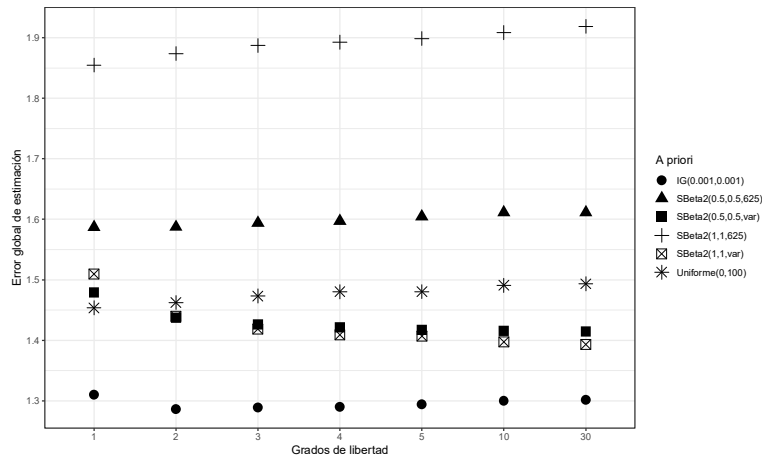
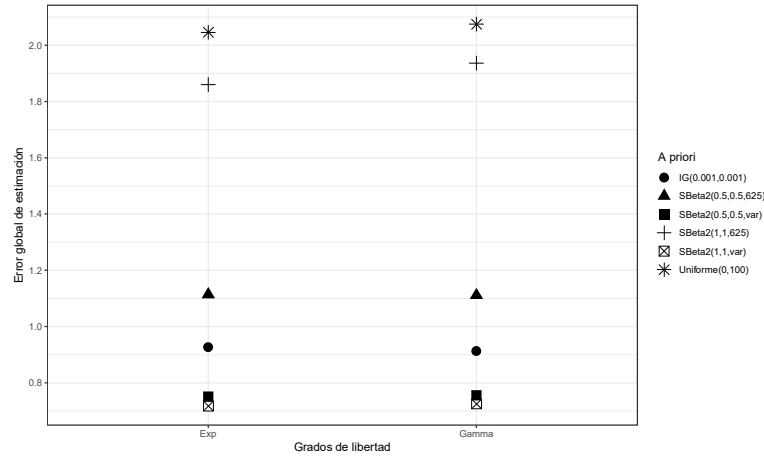


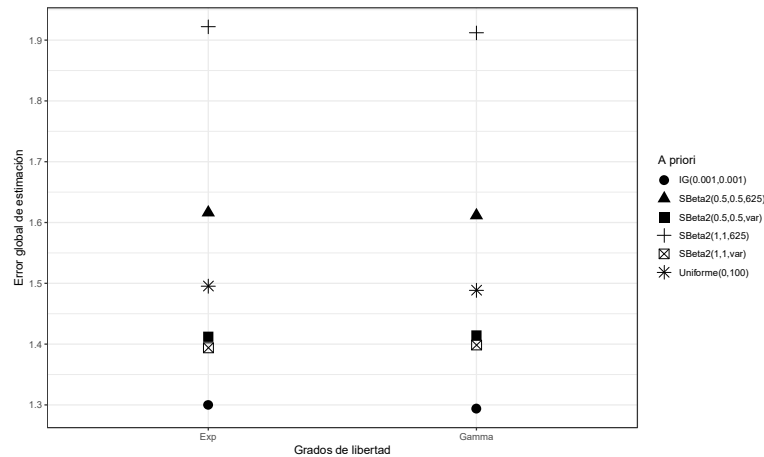
Figura 3-3.: Error global de estimación para  $J = 10$  en  $S_I$

Los valores más altos de  $G$  se presentan bajo el ajuste  $\sigma_{\alpha}^2 \sim SBeta2(1, 1, 25^2)$  seguido de  $\sigma_{\alpha}^2 \sim SBeta2(1/2, 1/2, 25^2)$ , lo cual se puede presentar debido a que bajo las condiciones

como fue construido este escenario de simulación el valor de la varianza muestral de los efectos es muy pequeña comparada con el parámetro de escala de la distribución SBeta2 el cual corresponde a 625. Con respecto a los valores de  $G$  cuando se asumen para los grados de libertad las propuestas presentadas por Juárez y Steel (2010) y Geweke (1993) las figuras 3-4 y 3-5 ilustran el comportamiento para  $J = 3$  y  $J = 10$  grupos experimentales respectivamente, donde se puede apreciar dos grupos, los valores de  $G$  mayores y menores respectivamente. En el primer grupo encontramos el ajuste con las distribuciones a priori SBeta2(1, 1, 625) y Uniforme(0, 100) para  $J = 3$  grupos, pero cuando  $J = 10$  el grupo con los valores de  $G$  más altos queda definido por los valores del ajuste  $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, 625)$ .



**Figura 3-4.:** Error global de estimación para  $J = 3$  en  $S_I$  con grados de libertad continuos



**Figura 3-5.:** Error global de estimación para  $J = 10$  en  $S_I$  con grados de libertad continuos

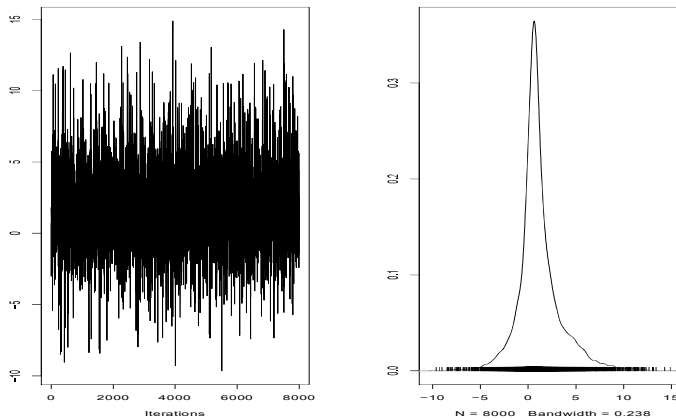
Se realizó el chequeo de convergencia para verificar el supuesto asociado a que las cadenas alcanzan la distribución estacionaria evaluando la autocorrelación existente entre los valores

generados del parámetro en distintos rezagos los cuales están dados en la Tabla **3-5** y con ellos realizar el test KPSS (**ver sección 2.3**). Tomando las cadenas obtenidas para  $J = 3$  grupos el nivel KPSS es de 0.12345 con parámetro de truncamiento 11 y un  $p$ -valor de 0.1 de donde se sigue que, no existe suficiente evidencia muestral para rechazar la hipótesis nula la cual establece  $H_0$  : la cadena de Markov ha alcanzado la distribución estacionaria.

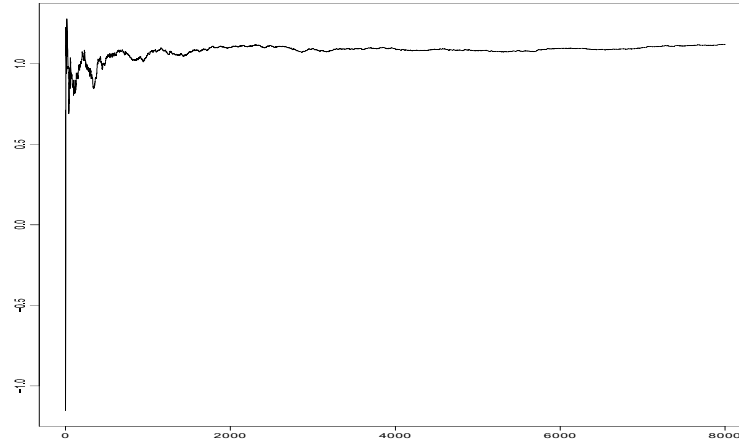
**Tabla 3-5.:** Autocorrelaciones por rezagos para  $J = 3$  con  $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, \sigma_{\alpha_j}^2)$

$j$	Rezago	Autocorrelación de $\sigma_\alpha^2$ con $j$ rezagos
1	1 Rezago	0.0068795549
5	5 Rezagos	0.0007616035
10	10 Rezagos	-0.0132159653
50	50 Rezagos	0.0011158273

Como parte del diagnóstico asociado a la prueba de hipótesis anterior tenemos los valores de las autocorrelaciones a diferentes rezagos reportados en la Tabla **3-5** en donde se puede apreciar que la asociación entre los valores del parámetro generado con los diferentes rezagos es muy baja. En las figuras **3-6** y **3-7** se muestran la traza, la densidad y los promedios móviles para la cadena de valores generados con los diferentes rezagos de la distribución a posteriori con  $J = 3$  grupos para  $S_I$ , en las cuales se puede apreciar que después de realizar el *burn-in* de las primeras 2000 cadenas se eliminan las autocorrelaciones no estables en la cadena y la distribución a posterior del parámetro de interés sigue un comportamiento estacionario, muestrándose de la distribución límite.



**Figura 3-6.:** Traza y densidad para una de las cadenas simuladas del efecto aleatorio  $\alpha$  con  $J = 3$  grupos en  $S_I$



**Figura 3-7.:** Promedio móvil para una de las cadenas simuladas del efecto aleatorio  $\alpha$

En síntesis para  $S_I$ , se sigue que, con base a los resultados obtenidos vía simulación para todos los grupos desde  $J = 3$  hasta 10 en  $S_I$ , los valores más altos de la medida de error global de estimación se presentan cuando se realiza el ajuste del componente de varianza  $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, 25^2)$  seguido del ajuste  $\sigma_\alpha^2 \sim \text{Uniforme}(0, 100)$ , con un comportamiento muy característico para  $J = 3$  grupos experimentales porque no se tiene un patrón de tendencia a incrementar o decrecer con respecto a los valores de los grados de libertad, lo cual puede asociarse a que al tener un número de grupos pequeño la estimación de la variación entre grupos puede resultar difícil de estimar conduciendo a que el modelado jerárquico no impacte mucho en la estimación de la medida de interés. Los menores valores del error global de estimación se obtienen cuando se realizan los ajustes  $\sigma_\alpha^2 \sim \text{SBeta2}(1/2, 1/2, \sigma_{\alpha_j})$  y  $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, \sigma_{\alpha_j})$ . Para evaluar el comportamiento conjunto del modelamiento de los efectos aleatorios a partir de la distribución  $t$  de Student asociándole al parámetro de escala diferentes distribuciones a priori, comparamos los resultados del error en la estimación de los efectos del modelo obtenidos vía simulación para  $\nu = 1, 2, 3$  con los resultados presentados en Pérez et al. (2017), los cuales modelan los efectos aleatorios en un modelo jerárquico con una distribución normal. Los valores de  $G$  de interés para realizar la comparación se presentan en Tabla 3-6. Notemos que, cuando se tienen  $J = 3$  grupos, solo para el ajuste  $\sigma_\alpha^2 \sim \text{SBeta2}(1/2, 1/2, 25^2)$  se obtienen valores de la medida de error de estimación global menores a los reportados bajo el ajuste de los efectos con una distribución Normal, en cambio para  $J = 4, 5, 6$  grupos corresponden al ajuste de  $\sigma_\alpha^2 \sim \text{SBeta2}(1/2, 1/2, 25^2)$  y  $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, 25^2)$ , los cuales consisten en la propuesta para  $\sigma_\alpha$  de Gelman (2006b) a partir de la distribución  $\text{SBeta2}$  estudiada en Pérez et al. (2017) como alternativa para garantizar una inferencia robusta. En el caso de  $J = 7, 8, 9, 10$  grupos se obtienen valores de  $G$  menores a los reportados bajo el ajuste de los efectos con una distribución Normal cuando se realizó el ajuste con  $\sigma_\alpha^2 \sim \text{IG}(0.001, 0.001)$ ,  $\sigma_\alpha^2 \sim \text{SBeta2}(1/2, 1/2, 25^2)$  y  $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, 25^2)$ .



**Tabla 3-6.:** Comparación de errores de estimación global,  $G$  en  $S_I$ 

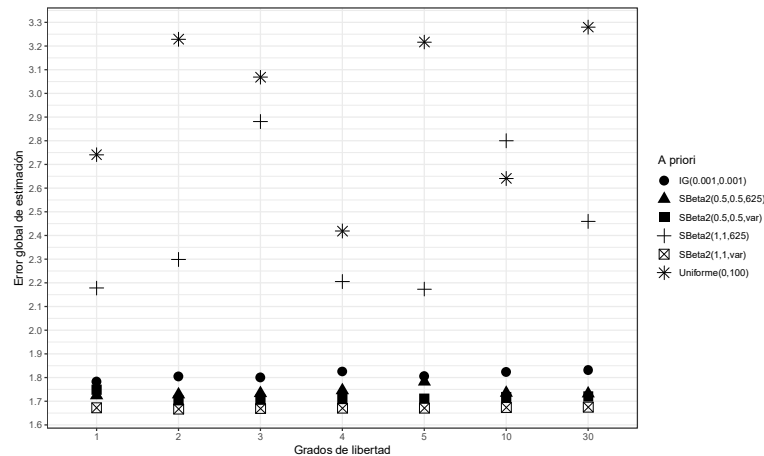
$J$		$IG$ (0.001, 0.001)	SBeta2 (1/2, 1/2, 25 <sup>2</sup> )	SBeta2 (1, 1, 25 <sup>2</sup> )	SBeta2 (1/2, 1/2, $\sigma_{\alpha_j}^2$ )	SBeta2 (1, 1, $\sigma_{\alpha_j}^2$ )
3	$t$ -Student( $\nu = 1$ )	0.9206	<b>1.1088</b>	2.3714	0.8523	0.8568
	$t$ -Student( $\nu = 2$ )	0.9084	<b>1.1090</b>	1.9800	0.7952	0.7782
	$t$ -Student( $\nu = 3$ )	0.9146	<b>1.1257</b>	1.9399	0.7765	0.7526
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	0.9133	<b>1.1110</b>	1.9365	0.7559	0.7248
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	0.9275	<b>1.1137</b>	1.8602	0.7513	0.7176
	Normal	0.6521	1.1304	1.4502	0.5060	0.4574
4	$t$ -Student( $\nu = 1$ )	0.8433	<b>1.1580</b>	<b>1.5832</b>	0.7913	0.7988
	$t$ -Student( $\nu = 2$ )	0.8391	<b>1.1666</b>	<b>1.5903</b>	0.7602	0.7479
	$t$ -Student( $\nu = 3$ )	0.8347	<b>1.1696</b>	<b>1.5910</b>	0.7514	0.7359
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	0.8365	<b>1.1758</b>	<b>1.5965</b>	0.7453	0.7247
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	0.8273	<b>1.1759</b>	1.5977	0.7392	0.7231
	Normal	0.6947	1.2017	1.6316	0.5266	0.4741
5	$t$ -Student( $\nu = 1$ )	1.0320	<b>1.2618</b>	<b>1.6479</b>	1.0244	1.1275
	$t$ -Student( $\nu = 2$ )	1.0211	<b>1.2715</b>	<b>1.6544</b>	0.9891	1.0546
	$t$ -Student( $\nu = 3$ )	1.0251	<b>1.2723</b>	<b>1.6594</b>	0.9790	1.0327
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	1.0334	<b>1.2814</b>	<b>1.6680</b>	0.9675	1.0053
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	1.0281	<b>1.2814</b>	<b>1.6696</b>	0.9652	0.9999
	Normal	0.8856	1.2833	1.6934	0.8156	0.7855
6	$t$ -Student( $\nu = 1$ )	1.0819	<b>1.2959</b>	<b>1.6723</b>	1.0745	1.0994
	$t$ -Student( $\nu = 2$ )	1.0646	<b>1.3054</b>	<b>1.6851</b>	1.0306	1.0331
	$t$ -Student( $\nu = 3$ )	1.0689	<b>1.3082</b>	<b>1.6915</b>	1.0230	1.0137
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	1.0789	<b>1.3187</b>	<b>1.7049</b>	1.0129	0.9978
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	1.0816	<b>1.3246</b>	<b>1.7096</b>	1.0117	0.9926
	Normal	0.9519	1.3288	1.7373	0.8818	0.8514
7	$t$ -Student( $\nu = 1$ )	<b>1.1461</b>	<b>1.3411</b>	<b>1.7095</b>	1.1429	1.1747
	$t$ -Student( $\nu = 2$ )	<b>1.1245</b>	<b>1.3500</b>	<b>1.7272</b>	1.1027	1.1031
	$t$ -Student( $\nu = 3$ )	<b>1.1296</b>	<b>1.3551</b>	<b>1.7333</b>	1.0892	1.0814
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	<b>1.1399</b>	<b>1.3704</b>	<b>1.7532</b>	1.0790	1.0613
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	<b>1.1434</b>	<b>1.3760</b>	<b>1.7598</b>	1.0777	1.0583
	Normal	2.8659	3.4144	3.8001	0.9765	0.9453
8	$t$ -Student( $\nu = 1$ )	<b>1.2734</b>	<b>1.4350</b>	<b>1.7640</b>	1.2777	1.3098
	$t$ -Student( $\nu = 2$ )	<b>1.2553</b>	<b>1.4412</b>	<b>1.7812</b>	1.2383	1.2370
	$t$ -Student( $\nu = 3$ )	<b>1.2654</b>	<b>1.4459</b>	<b>1.7919</b>	1.2239	1.2152
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	<b>1.2629</b>	<b>1.4627</b>	<b>1.8142</b>	1.2118	1.1960
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	<b>1.2686</b>	<b>1.4642</b>	<b>1.8197</b>	1.2106	1.1895
	Normal	3.4815	3.9480	4.2805	1.0964	1.0640
9	$t$ -Student( $\nu = 1$ )	<b>1.2807</b>	<b>1.4402</b>	<b>1.7581</b>	1.2853	1.3216
	$t$ -Student( $\nu = 2$ )	<b>1.2619</b>	<b>1.4457</b>	<b>1.7772</b>	1.2402	1.2433
	$t$ -Student( $\nu = 3$ )	<b>1.2623</b>	<b>1.4520</b>	<b>1.7867</b>	1.2272	1.2180
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	<b>1.2668</b>	<b>1.4692</b>	<b>1.8121</b>	1.2120	1.1954
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	<b>1.2717</b>	<b>1.4752</b>	<b>1.8194</b>	1.2121	1.1883
	Normal	3.9462	4.4790	4.8496	1.1066	1.0797
10	$t$ -Student( $\nu = 1$ )	<b>1.3105</b>	<b>1.5871</b>	<b>1.8548</b>	1.4790	1.5094
	$t$ -Student( $\nu = 2$ )	<b>1.2865</b>	<b>1.5873</b>	<b>1.8738</b>	1.4375	1.4397
	$t$ -Student( $\nu = 3$ )	<b>1.2893</b>	<b>1.5941</b>	<b>1.8871</b>	1.4263	1.4188
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	<b>1.2939</b>	<b>1.6117</b>	<b>1.9121</b>	1.4144	1.3984
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	<b>1.3001</b>	<b>1.6166</b>	<b>1.9221</b>	1.4119	1.3937
	Normal	3.1022	3.6381	4.0216	1.1333	1.1025

Para los otros ajustes se presentan las variaciones porcentuales en la Tabla **3-7**, estos es, se toman los registros del ajuste de  $\sigma_\alpha^2 \sim \text{SBeta2}(1/2, 1/2, \sigma_{\alpha_j}^2)$  y  $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, \sigma_{\alpha_j}^2)$  con un grado de libertad y como referencia el valor de error de estimación global reportado en Pérez et al. (2017), arrojando resultados que muestran una tendencia creciente para la variación porcentual a medida de que aumenta el número de grupos.

**Tabla 3-7.:** Variación porcentual del error de estimación global en  $S_I$

Grupos	Distribución para $\sigma_\alpha^2$	
	$\text{SBeta2}(1/2, 1/2, \sigma_{\alpha_j}^2)$	$\text{SBeta2}(1, 1, \sigma_{\alpha_j}^2)$
$J = 7$	17.04 %	24.26 %
$J = 8$	16.53 %	23.10 %
$J = 9$	16.14 %	22.40 %
$J = 10$	30.50 %	36.90 %

Para el escenario en el cual se presentan algunos efectos con valores atípicos moderados, denotado por  $S_{II}$ , los resultados principales del estudio de simulación se comentan a continuación de acuerdo con las asignaciones del número de valores atípicos. Para  $J = 3$  grupos en la Figura **3-8** se puede apreciar que cuando se utiliza como distribución a priori para la varianza  $\text{SBeta2}(1, 1, 25^2)$  y  $\text{Uniforme}(0, 100)$ , los valores de  $G$  son mayores comparados con las otras distribuciones a priori. Adicionalmente se observa que los valores de  $G$  bajo el ajuste de una  $IG(0.001, 0.001)$  es mayor que las otras tres distribuciones para  $\sigma_\alpha^2$  estudiadas:  $\text{SBeta2}(1/2, 1/2, 25^2)$ ,  $\text{SBeta2}(1/2, 1/2, \sigma_{\alpha_j}^2)$  y  $\text{SBeta2}(1, 1, \sigma_{\alpha_j}^2)$ .



**Figura 3-8.:** Error global de estimación para  $J = 3$  en  $S_{II}$

En cambio, para  $J = 4, 5$  las figuras **3-9** y **3-10** respectivamente, ilustran un comportamiento en el cual se puede establecer como el valor del ajuste  $\sigma_\alpha^2 \sim IG(0.001, 0.001)$  incrementa.

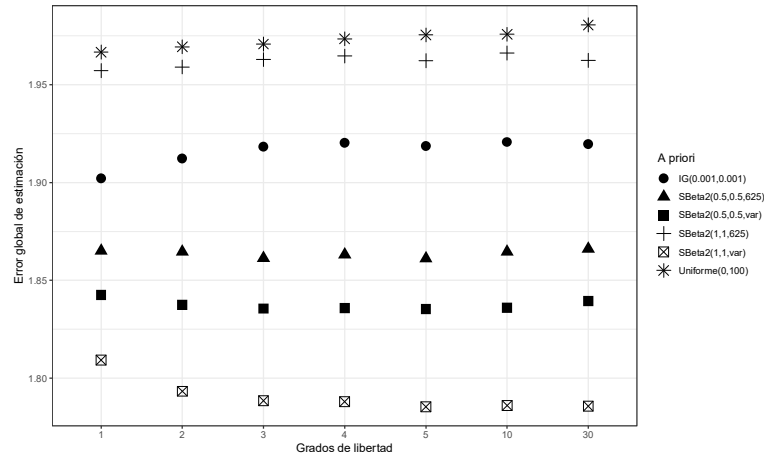


Figura 3-9.: Error global de estimación para  $J = 4$  en  $S_{II}$

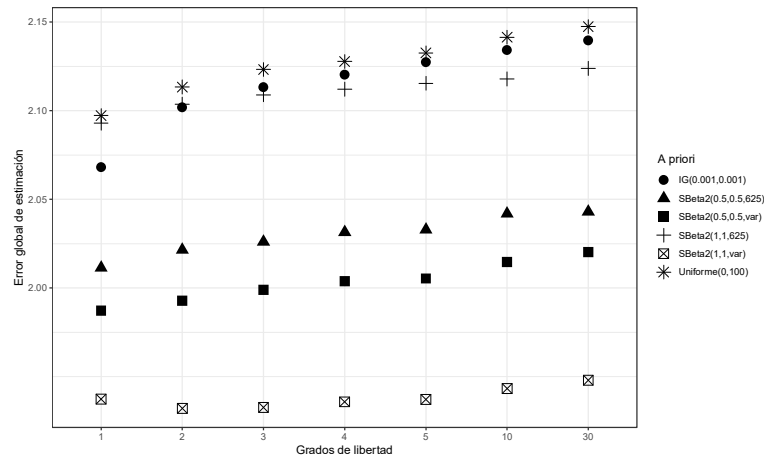
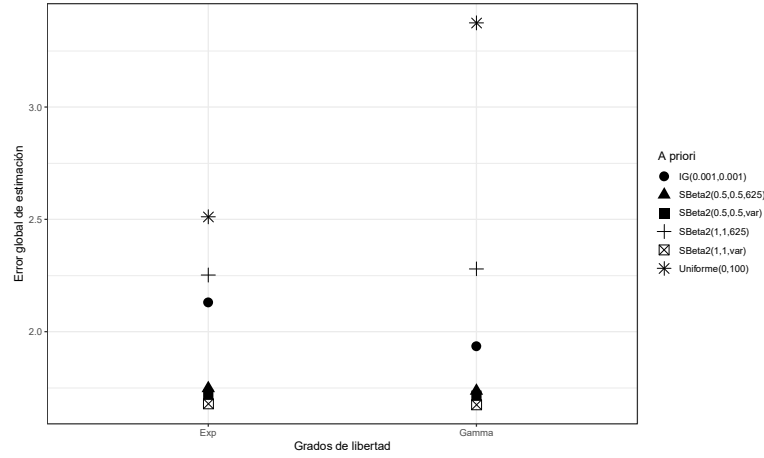


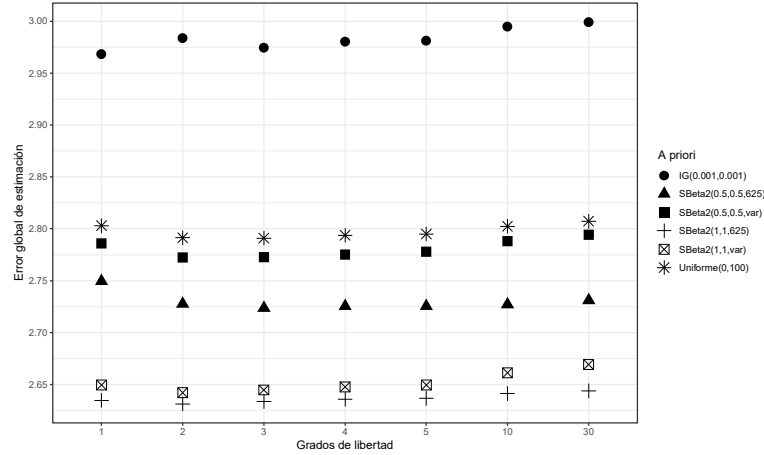
Figura 3-10.: Error global de estimación para  $J = 5$  en  $S_{II}$

Por lo anterior, podemos establecer que para  $J = 3, 4, 5$  grupos los cuales tienen solo un valor atípico moderado se presenta que el mayor valor de  $G$  con el ajuste  $\sigma_\alpha^2 \sim \text{Uniforme}(0, 100)$  seguido del ajuste  $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, \sigma_{\alpha_j}^2)$ . La tendencia de los valores de  $G$  es creciente en términos del aumento de los valores de los grados de libertad que se están analizando. Desde el punto de vista de grados de libertad con la propuesta de Juárez y Steel (2010) y Geweke (1993) para  $J = 3$  grupos, la Figura 3-11 presenta el comportamiento de los valores de  $G$ , los cuales presentan mayores valores bajo el ajuste  $\sigma_\alpha^2 \sim \text{Uniforme}(0, 100)$



**Figura 3-11.:** Error global de estimación para  $J = 3$  en el  $S_{II}$  con grados de libertad continuos

Para  $J = 6, 7$  grupos se les asigna dos valores atípicos moderado y se puede apreciar en la Figura 3-12 que el mayor valor de  $G$  se presenta con el ajuste  $\sigma_\alpha^2 \sim IG(0.001, 0.001)$  seguido de  $\sigma_\alpha^2 \sim Uniforme(0, 100)$ . Además se observa que los valores con los ajustes  $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$  y  $\sigma_\alpha^2 \sim SBeta2(1, 1, \sigma_{\alpha_j}^2)$  son muy parecidos.



**Figura 3-12.:** Error global de estimación para  $J = 7$  en  $S_{II}$

Con respecto a  $J = 8$  la Figura 3-13 presenta la tendencia de los valores de  $G$ , los cuales disminuyen a medida que aumentan los grados de libertad discretos para todas las distribuciones a priori candidatas y registrando los mayores valores cuando se realiza el ajuste  $\sigma_\alpha^2 \sim IG(0.001, 0.001)$ .

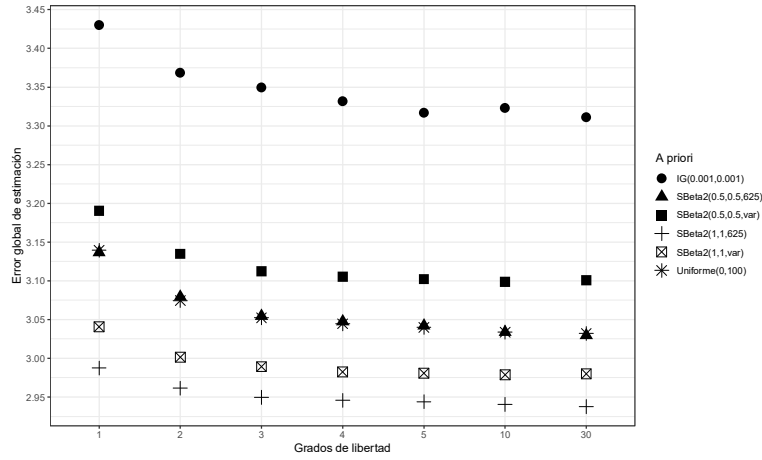


Figura 3-13.: Error global de estimación  $J = 8$  en  $S_{II}$

En el caso de  $J = 10$  grupos, en la Figura 3-14 se observa que los valores mayores de  $G$  se observan para  $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$  seguido del ajuste  $\sigma_\alpha^2 \sim IG(0.001, 0.001)$  y los valores menores bajo  $\sigma_\alpha^2 \sim SBeta2(1, 1, 25^2)$ ,  $\sigma_\alpha^2 \sim SBeta2(1, 1, \sigma_{\alpha_j}^2)$ .

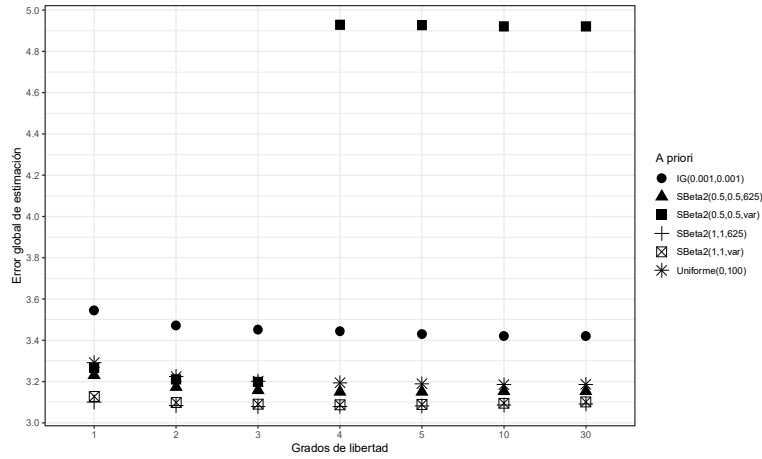


Figura 3-14.: Error global de estimación  $J = 10$  en  $S_{II}$

Además, el comportamiento del error de estimación individual, el cual corresponde a los valores obtenidos a partir de la expresión  $\frac{\sum_{k=1}^{1000} (\alpha_j - \alpha_{kj})}{1000}$ , cuando se realiza el ajuste  $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$  para  $J = 10$  grupos,  $\nu = 1$  en  $S_{II}$  es presentado en la Tabla 3-8. Comparando los errores de estimación individual registrados con los valores de  $G$  para  $J = 10$  en  $S_{II}$  se puede establecer que en los efectos que representan el valor atípico y los cuales corresponden a  $\alpha_8, \alpha_9$  y  $\alpha_{10}$  (ver Tabla 3-3, sección 3.1) presentan valores menores a  $G$  bajo las mismas consideraciones en  $S_{II}$ .

**Tabla 3-8.:** Error de estimación individual para  $J = 10$  en  $S_{II}$ 

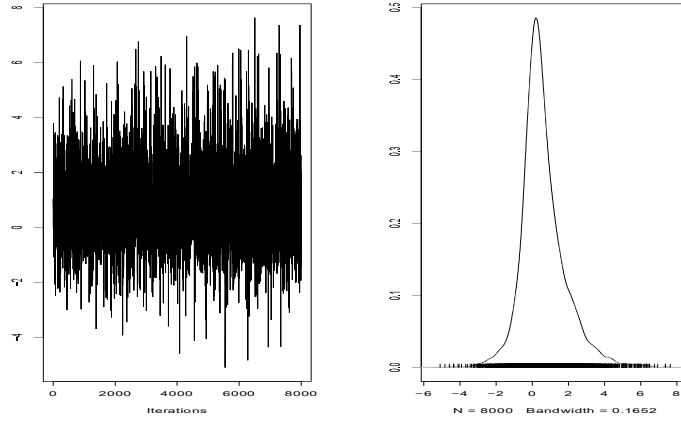
$\nu$	$j$									
	1	2	3	4	5	6	7	8	9	10
1	0.2597	0.2441	0.2915	0.3262	0.3110	0.2458	0.4049	2.7872	2.5042	3.3010
2	0.2899	0.2987	0.3471	0.3768	0.3585	0.3012	0.4258	2.5474	2.3234	3.0568
3	0.3093	0.3302	0.3820	0.4075	0.3859	0.3318	0.4401	2.4393	2.2497	2.9576
4	0.3235	0.3488	0.4014	0.4271	0.4037	0.3515	0.4501	2.3828	2.2080	2.9052
5	0.3298	0.3598	0.4152	0.4369	0.4151	0.3614	0.4588	2.3551	2.1882	2.8783
10	0.3466	0.3824	0.4404	0.4625	0.4361	0.3860	0.4767	2.3004	2.1608	2.8233
30	0.3587	0.4005	0.4599	0.4775	0.4532	0.4047	0.4876	2.2693	2.1431	2.7993

El chequeo de convergencia para verificar el supuesto asociado a que las cadenas alcanzan la distribución estacionaria en  $S_{II}$ , se ilustra para  $J = 10$  grupos a partir de los valores de la autocorrelación existente entre los valores generados del parámetro de interés en distintos rezagos, los cuales están dados en la Tabla **3-9** y el test KPSS con el cual se obtiene un nivel KPSS de 0.32543, con parámetro de truncamiento 12 y un  $p$ -valor de 0.1 de donde se sigue que, no existe suficiente evidencia muestral para rechazar la hipótesis nula la cual establece  $H_0$ : la cadena de Markov ha alcanzado la distribución estacionaria.

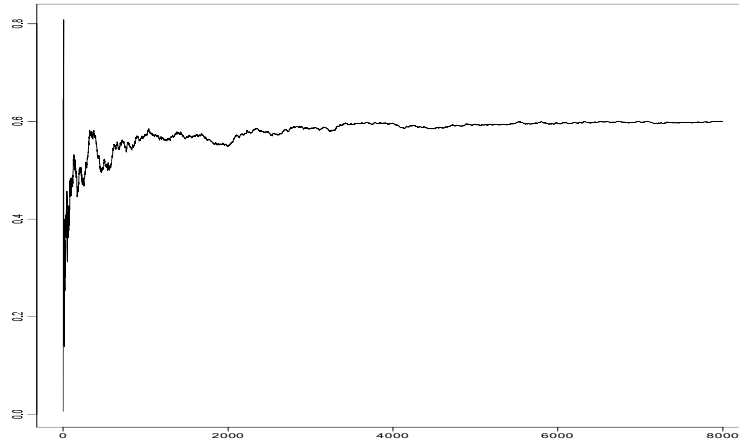
**Tabla 3-9.:** Autocorrelaciones por rezagos para  $J = 10$  con  $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, \sigma_{\alpha_j}^2)$ 

$j$	Rezago	Autocorrelación de $\sigma_\alpha^2$ con $j$ rezagos
1	1 Rezago	0.0051922584
5	5 Rezagos	0.0007036689
10	10 Rezagos	-0.0069672322
50	50 Rezagos	-0.0038706726

En la Tabla **3-9** se puede apreciar que la asociación entre los valores del parámetro generado con los diferentes rezagos es muy baja. En las figuras **3-15** y **3-16** se ilustra la traza, densidad y los promedios móviles respectivamente para una de las cadenas de valores generados con los diferentes rezagos con  $J = 10$  grupos para  $S_{II}$ , en las cuales se puede apreciar que después de realizar el *burn-in* de las primeras 2000 cadenas se eliminan las autocorrelaciones no estables en la cadena y la distribución a posterior del parámetro de interés sigue un comportamiento estacionario, muestreándose de la distribución límite.



**Figura 3-15.:** Trazo y densidad para una de las cadenas simuladas para  $\alpha$  con  $J = 3$  grupos en  $S_{II}$



**Figura 3-16.:** Promedio móvil para una de las cadenas simuladas para  $\sigma_\alpha^2$  con  $J = 3$  grupos en  $S_{II}$

Análogo a lo realizado con el primer escenario, se comparan los resultados del error en la estimación de los efectos del modelo obtenidos vía simulación para  $\nu = 1, 2, 3$  con los resultados presentados en Pérez et al. (2017). Los resultados se comentaran en función de la asignación de valores atípicos por  $J$  grupos y se presentan en Tabla 3-12. Para  $J = 3, 4$  grupos en los cuales existe un solo valor atípico, los valores para  $G$  son mayores cuando la a priori para los efectos aleatorios sigue una distribución Normal, excepto cuando se realiza el ajuste  $\sigma_\alpha^2 \sim SBeta2(1, 1, 25^2)$  y cuando la a priori corresponde a  $IG(0.001, 0.001)$  con grados de libertad continuos. Para  $J = 5$  grupos todos los valores son menores a los reportados en Pérez et al. (2017), en opuesto a lo anterior tenemos los registros de  $J = 6, 7$  grupos en el cual el ajuste de los efectos aleatorios bajo la normal son menores a los obtenidos en el estudio de simulación, por lo cual, se observó el comportamiento del error de estimación individual

para  $J = 6$  cuando se realiza el ajuste con una distribución a priori  $SBeta2(1, 1, \sigma_{\alpha_j}^2)$  y están reportados en la Tabla **3-10**, la cual se evidencia que el mayor valor se presenta en el valor atípico moderado de los dos asignados para el grupo.

**Tabla 3-10.:** Error de estimación individual para  $J = 6$  en  $S_{II}$

$\nu$	$j$					
	1	2	3	4	5	6
1	0.3048	0.3201	0.4559	0.3550	2.3284	2.6114
2	0.3319	0.3507	0.4831	0.3948	2.2236	2.5022
3	0.3501	0.3682	0.5003	0.4231	2.1750	2.4461
4	0.3581	0.3796	0.5109	0.4364	2.1588	2.4297
5	0.3668	0.3860	0.5176	0.4484	2.1509	2.4230
10	0.3812	0.4025	0.5349	0.4724	2.1383	2.4040
30	0.3933	0.4150	0.5456	0.4852	2.1299	2.3959

Para  $J = 8$  grupos los valores del ajuste  $\sigma_{\alpha}^2 \sim IG(0.001, 0.001)$  para grados de libertad discretos son mayores que los obtenidos bajo el ajuste de los efectos con una distribución normal seguido de los obtenidos bajo el ajuste  $\sigma_{\alpha}^2 \sim SBeta2(1/2, 1/2\sigma_{\alpha_j}^2)$ , para los demás ajustes los valores del error global de estimación presentan poca variación porcentual para grados de libertad discretos y continuos excepto cuando se tiene un grado de libertad, las variaciones porcentuales se presentan en la Tabla **3-11**

**Tabla 3-11.:** Variación porcentual del error de estimación global de  $J = 8$  en  $S_{II}$

$\nu$	Distribución para $\sigma_{\alpha}^2$		
	SBeta2(1/2, 1/2, 25 <sup>2</sup> )	SBeta2(1, 1, 25 <sup>2</sup> )	SBeta2(1, 1, $\sigma_{\alpha_j}^2$ )
$\nu = 1$	4.56 %	0.37 %	2.99 %
$\nu = 2$	2.66 %	-0.49 %	1.66 %
$\nu = 3$	1.83 %	-0.89 %	1.24 %
( $\nu \sim \text{Exp}(0.001)$ )	1.09 %	-1.24 %	0.93 %
$t\text{-Student}(\nu \sim \text{Exp}(0.001))$	1,0101 %	-1.24 %	1.07 %

Con  $J = 9$  grupos el ajuste  $\sigma_{\alpha}^2 \sim IG(0.001, 0.001)$  presenta mayor error de estimación global cuando los efectos aleatorios se ajustan con una distribución  $t\text{-Student}$  excepto cuando los grados de libertad son continuos con un porcentaje de variación en ambas propuestas de 0.3078 %. Para grados de libertad discretos con valor de la unidad el error global de de estimación para los ajustes  $\sigma_{\alpha}^2 \sim SBeta2(1/2, 1/2, 25^2)$  y  $\sigma_{\alpha}^2 \sim SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$ , son mayores que los obtenidos bajo los efectos con distribución normal con una variación



---

porcentual de 0,2089 % y 0,5292 % respectivamente. En el caso de  $J = 10$  grupos el valor más alto se obtiene con el ajuste de  $\sigma_\alpha^2 \sim \text{SBeta2}(1/2, 1/2, 25^2)$  seguidos de  $\sigma_\alpha^2 \sim \text{IG}(0.001, 0.001)$ .

**Tabla 3-12.:** Comparación de errores de estimación global para  $S_{II}$ 

$J$		$IG$ (0.001, 0.001)	SBeta2 (1/2, 1/2, 25 <sup>2</sup> )	SBeta2 (1, 1, 25 <sup>2</sup> )	SBeta2 (1/2, 1/2, $\sigma_{\alpha_j}^2$ )	SBeta2 (1, 1, $\sigma_{\alpha_j}^2$ )
3	$t$ -Student( $\nu = 1$ )	<b>1.7834</b>	<b>1.7247</b>	2.1779	<b>1.7491</b>	<b>1.6725</b>
	$t$ -Student( $\nu = 2$ )	<b>1.8055</b>	<b>1.7285</b>	2.2990	<b>1.7034</b>	<b>1.6673</b>
	$t$ -Student( $\nu = 3$ )	<b>1.8010</b>	<b>1.7333</b>	2.8813	<b>1.7069</b>	<b>1.6689</b>
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	1.9367	<b>1.7376</b>	2.2803	<b>1.7155</b>	<b>1.6755</b>
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	2.1317	<b>1.7479</b>	2.2525	<b>1.7204</b>	<b>1.6794</b>
	Normal	1.9008	1.8538	1.9158	1.8347	1.7986
4	$t$ -Student( $\nu = 1$ )	1.9022	<b>1.8651</b>	<b>1.9572</b>	<b>1.8425</b>	<b>1.8092</b>
	$t$ -Student( $\nu = 2$ )	1.9124	<b>1.8646</b>	<b>1.9591</b>	<b>1.8375</b>	<b>1.7933</b>
	$t$ -Student( $\nu = 3$ )	1.9184	<b>1.8613</b>	<b>1.9629</b>	<b>1.8356</b>	<b>1.7885</b>
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	1.9202	<b>1.8653</b>	<b>1.9635</b>	<b>1.8383</b>	<b>1.7858</b>
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	1.9202	<b>1.8675</b>	<b>1.9708</b>	<b>1.8397</b>	<b>1.7872</b>
	Normal	1.8966	1.9197	2.0588	1.8489	1.8202
5	$t$ -Student( $\nu = 1$ )	<b>2.0682</b>	<b>2.0114</b>	<b>2.0930</b>	<b>1.9873</b>	<b>1.9374</b>
	$t$ -Student( $\nu = 2$ )	<b>2.1020</b>	<b>2.0215</b>	<b>2.1037</b>	<b>1.9928</b>	<b>1.9321</b>
	$t$ -Student( $\nu = 3$ )	<b>2.1134</b>	<b>2.0260</b>	<b>2.1090</b>	<b>1.9990</b>	<b>1.9327</b>
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	<b>2.1363</b>	<b>2.0397</b>	<b>2.1220</b>	<b>2.0145</b>	<b>1.9448</b>
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	<b>2.1449</b>	<b>2.0453</b>	<b>2.1260</b>	<b>2.0229</b>	<b>1.9495</b>
	Normal	2.1835	2.1581	2.2710	2.1134	2.0755
6	$t$ -Student( $\nu = 1$ )	2.7145	2.4905	2.4174	2.5250	2.4127
	$t$ -Student( $\nu = 2$ )	2.7255	2.4709	2.4174	2.5073	2.3979
	$t$ -Student( $\nu = 3$ )	2.7185	2.4691	2.4194	2.5026	2.3942
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	2.7274	2.4698	2.4250	2.5151	2.4048
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	2.7319	2.4729	2.4315	2.5241	2.4151
	Normal	2.5672	2.3760	2.3671	2.4033	2.3188
7	$t$ -Student( $\nu = 1$ )	2.9684	2.7497	2.6348	2.7859	2.6497
	$t$ -Student( $\nu = 2$ )	2.9839	2.7276	2.6314	2.7725	2.6423
	$t$ -Student( $\nu = 3$ )	2.9746	2.7236	2.6336	2.7727	2.6448
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	2.9960	2.7272	2.6412	2.7873	2.6633
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	2.9920	2.7287	2.6458	2.7989	2.6740
	Normal	2.6535	2.4811	2.4775	2.4979	2.4238
8	$t$ -Student( $\nu = 1$ )	3.4303	3.1365	2.9874	3.1904	3.0407
	$t$ -Student( $\nu = 2$ )	3.3687	3.0795	<b>2.9614</b>	3.1348	3.0014
	$t$ -Student( $\nu = 3$ )	3.3497	3.0546	<b>2.9496</b>	3.1122	2.9892
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	3.3166	3.0322	<b>2.9390</b>	3.1012	2.9800
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	3.3110	3.0298	<b>2.9389</b>	3.1010	2.9840
	Normal	3.1622	2.9995	2.9761	3.0213	2.9523
9	$t$ -Student( $\nu = 1$ )	3.4324	3.1656	3.0292	3.2099	3.0654
	$t$ -Student( $\nu = 2$ )	<b>3.3812</b>	<b>3.1129</b>	<b>3.0134</b>	<b>3.1606</b>	<b>3.0370</b>
	$t$ -Student( $\nu = 3$ )	3.3656	<b>3.0985</b>	<b>3.0097</b>	<b>3.1462</b>	<b>3.0283</b>
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	<b>3.3359</b>	<b>3.0907</b>	<b>3.0091</b>	<b>3.1446</b>	<b>3.0341</b>
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	<b>3.3359</b>	<b>3.0902</b>	<b>3.0157</b>	<b>3.1477</b>	<b>3.0407</b>
	Normal	3.3462	3.1590	3.1204	3.1930	3.1186
10	$t$ -Student( $\nu = 1$ )	3.5456	<b>3.2310</b>	<b>3.1009</b>	<b>3.2670</b>	<b>3.1282</b>
	$t$ -Student( $\nu = 2$ )	3.4723	<b>3.1739</b>	<b>3.0838</b>	<b>3.2113</b>	<b>3.0978</b>
	$t$ -Student( $\nu = 3$ )	3.4521	<b>3.1577</b>	<b>3.0790</b>	<b>3.1986</b>	<b>3.0904</b>
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	<b>3.4267</b>	<b>3.1518</b>	<b>3.0874</b>	4.9229	<b>3.0949</b>
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	<b>3.4222</b>	<b>3.1568</b>	<b>3.0923</b>	4.9220	<b>3.1043</b>
	Normal	3.4352	3.2333	3.1895	3.2758	3.1949

Para el tercer escenario el cual considera un efecto con un gran valor atípico y el cual denotaremos por  $S_{III}$ , la Figura 3-17 ilustra el comportamiento de  $G$ , los cuales evidencian que los valores más altos se presentan con el ajuste  $\sigma_\alpha^2 \sim \text{Uniforme}(0, 100)$ , los menores valores se presentan con los ajustes  $\text{SBeta2}(1/2, 1/2, 25^2)$ ,  $\text{SBeta2}(1, 1, \sigma_{\alpha_j}^2)$  y  $\text{SBeta2}(1/2, 1/2, \sigma_{\alpha_j}^2)$  respectivamente.

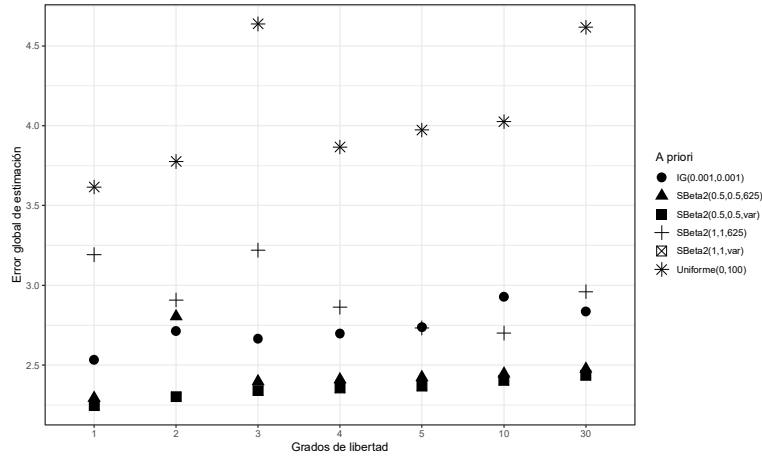


Figura 3-17.: Error global de estimación para  $J = 3$  en el escenario 3

Para  $J = 4, 5$  grupos, las figuras 3-18 y 3-19, ilustran el comportamiento de  $G$  y de las cuales se aprecia que, los mayores valores de  $G$  se presentan cuando se realiza el ajuste  $\sigma_\alpha^2 \sim \text{IG}(0.001, 0.001)$  seguido del ajuste  $\sigma_\alpha^2 \sim \text{SBeta2}(1/2, 1/2, )$  y además los  $\sigma_{\alpha_j}^2$  valores de  $G$  para los ajustes de  $\sigma_\alpha^2$  con las distribución  $\text{SBeta2}(1, 1, 25^2)$  y  $\text{Uniforme}(0, 100)$  son muy parecidos.

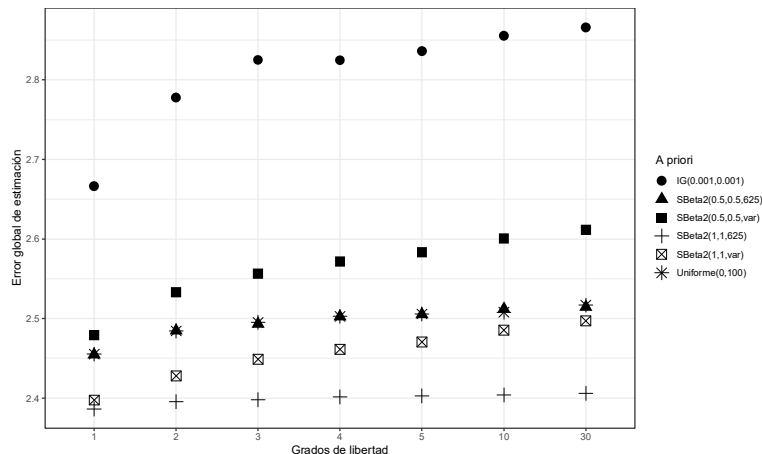


Figura 3-18.: Error global de estimación para  $J = 4$  en  $S_{III}$

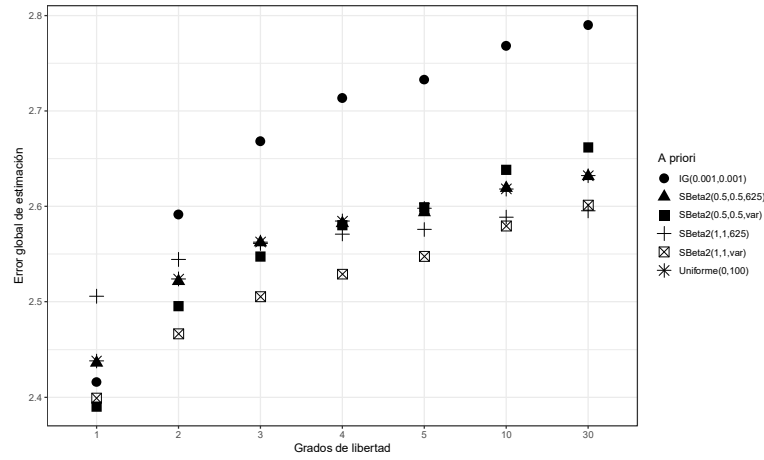


Figura 3-19.: Error global de estimación para  $J = 5$  en  $S_{III}$

En el caso de  $J = 3$  grupos y grados de libertad continuos para la distribución  $t$  de Student a partir de la propuesta de Juárez y Steel (2010), el ajuste  $\sigma_\alpha^2 \sim \text{Uniforme}(0, 100)$  presentan los mayores de  $G$  con respecto a las otras distribuciones a priori y se puede evidenciar en la Figura 3-20.

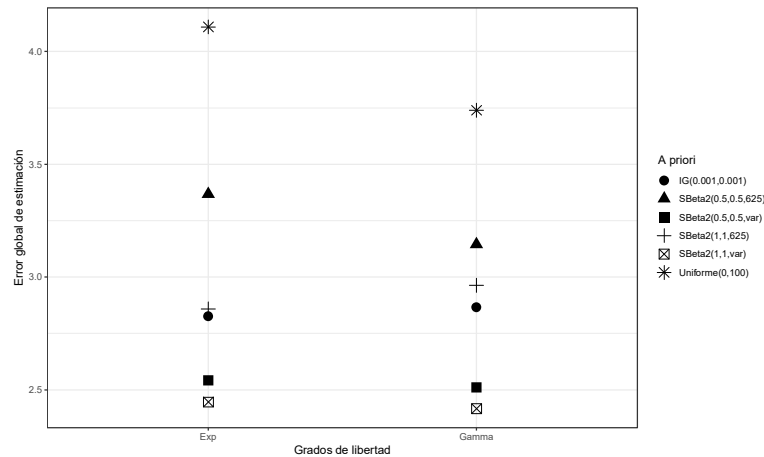


Figura 3-20.: Error global de estimación para  $J = 3$  en  $S_{III}$

Para  $J \geq 4$  grupos con grados de libertad continuos, se presentan los mayores de  $G$  para el ajuste  $\sigma_\alpha^2 \sim IG(0.001, 0.001)$ , lo cual se puede apreciar en las figuras 3-21 y 3-22 para  $J = 3$  y  $J = 8$  grupos respectivamente.

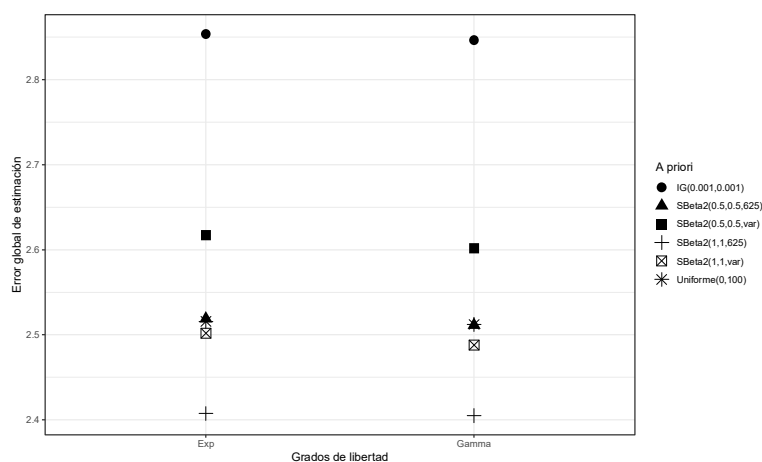


Figura 3-21.: Error global de estimación para  $J = 4$  en  $S_{III}$  con grados de libertad continuos

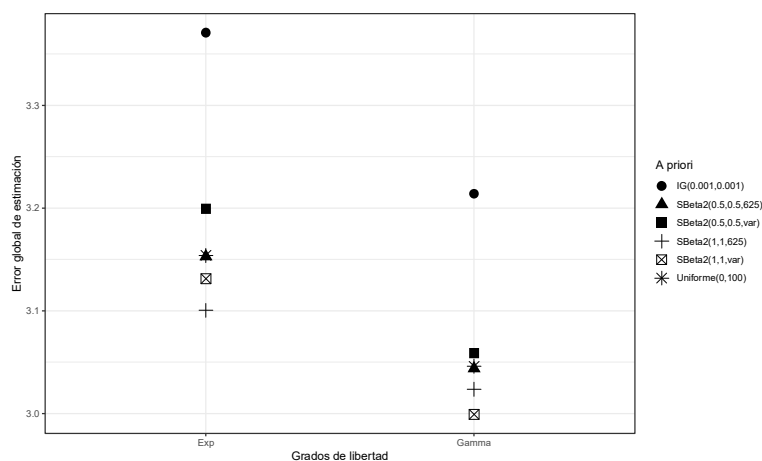


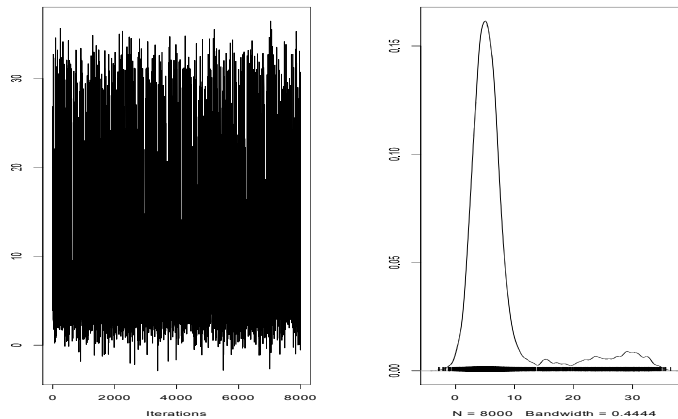
Figura 3-22.: Error global de estimación para  $J = 8$  en  $S_{III}$  con grados de libertad continuos

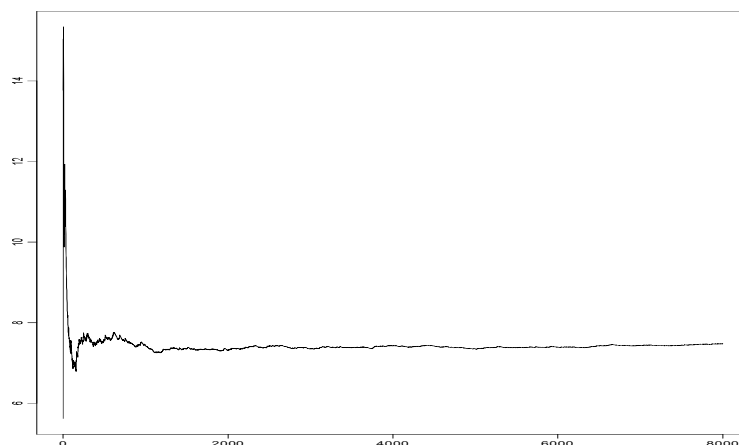
La revisión del supuesto de convergencia para verificar el supuesto asociado a que las cadenas alcanzan la distribución estacionaria en  $S_{III}$  presenta a continuación para  $J = 3$  grupos, a partir de los valores de la autocorrelación existente entre los valores generados del parámetro de interés en distintos rezagos, los cuales están dados en la Tabla 3-13 y el test KPSS con el cual se obtiene un nivel KPSS de 0.14446, con parámetro de truncamiento 20 y un  $p$ -valor de 0.1 de donde se sigue que, no existe suficiente evidencia muestral para rechazar la hipótesis nula la cual establece  $H_0$ : la cadena de Markov ha alcanzado la distribución estacionaria.

**Tabla 3-13.:** Autocorrelaciones por rezagos para  $J = 3$  con  $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, 25^2)$ 

$j$	Rezago	Autocorrelación de $\sigma_\alpha^2$ con
1	1 Rezago	0.002590061
2	5 Rezagos	-0.002876783
3	10 Rezagos	0.021183279
4	50 Rezagos	-0.004754381

En la Tabla **3-13** se puede apreciar que la asociación entre los valores del parámetro generado con los diferentes rezagos es muy baja. En las figuras **3-23** y **3-24** muestran la traza, densidad y promedios móviles respectivamente para una de las cadenas de valores generados con los diferentes rezagos de la distribución a posteriori del efecto aleatorio con  $J = 3$  grupos para  $S_{III}$ , en la cual se evidencia el comportamiento estacional de la cadena luego de realizar el *burn-in* y la distribución a posterior del parámetro de interés sigue un comportamiento estacionario. Los resultados obtenidos para todos los escenarios de interés están presentes en el Anexo A.

**Figura 3-23.:** Traza y densidad para una de las cadenas simuladas para  $\alpha$  con  $J = 3$  grupos en  $S_{III}$



**Figura 3-24.:** Promedio móvil para una de las cadenas simuladas para  $\alpha$  con  $J = 3$  grupos en  $S_{III}$

Comparando los resultados vía simulación bajo las condiciones de  $S_{III}$  con los reportados en Pérez et al. (2017) los cuales se pueden apreciar en la Tabla **3-14**, se sigue, para todos los grupos los valores de  $G$  son menores cuando los efectos aleatorios siguen una distribución  $t$  de Student. La distribución a priori para  $\sigma_\alpha^2$  con los valores más altos es  $IG(0.001, 0.001)$ . Así, podemos establecer que bajo las condiciones de  $S_{III}$  se presentan los menores valores de  $G$ , los cuales incrementan a medida que aumenta el valor de los grados de libertad, en cambio que, en  $S_I$  y  $S_{II}$  se presentan valores más altos para  $G$  y una tendencia a disminuir a medida que aumentaban los grados de libertad.

**Tabla 3-14.:** Comparación de errores de estimación global,  $G$  en  $S_{III}$ 

$J$		$IG$ (0.001, 0.001)	SBeta2 (1/2, 1/2, 25 <sup>2</sup> )	SBeta2 (1, 1, 25 <sup>2</sup> )	SBeta2 (1/2, 1/2, $\sigma_{\alpha_j}^2$ )	SBeta2 (1, 1, $\sigma_{\alpha_j}^2$ )
3	$t$ -Student( $\nu = 1$ )	2.5341	2.2945	3.1920	2.3195	2.2461
	$t$ -Student( $\nu = 2$ )	2.7142	2.8051	2.9083	2.4139	2.3022
	$t$ -Student( $\nu = 3$ )	2.6659	2.3979	3.2199	2.4264	2.3405
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	2.8672	3.1445	2.9635	2.5108	2.4161
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	2.8263	3.3680	2.8585	2.5415	2.4456
	Normal	3.6400	3.3783	3.3187	3.4457	3.3858
4	$t$ -Student( $\nu = 1$ )	2.6692	2.4229	2.3787	2.4482	2.3574
	$t$ -Student( $\nu = 2$ )	2.8288	2.5050	2.4261	2.5461	2.4346
	$t$ -Student( $\nu = 3$ )	2.8752	2.5436	2.4499	2.6028	2.4810
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	2.9673	2.6250	2.5075	2.7160	2.5945
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	3.0365	2.6530	2.5369	2.7616	2.6425
	Normal	3.5496	3.3040	3.2371	3.3694	3.2948
5	$t$ -Student( $\nu = 1$ )	2.2107	2.2284	2.3339	2.1806	2.1945
	$t$ -Student( $\nu = 2$ )	2.4418	2.3671	2.4170	2.3303	2.3063
	$t$ -Student( $\nu = 3$ )	2.5537	2.4426	2.4634	2.4193	2.3769
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	2.7405	2.5916	2.5728	2.6012	2.5473
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	2.8167	2.6506	2.6060	2.6781	2.6187
	Normal	3.3458	3.2175	3.1934	3.2409	3.1962
6	$t$ -Student( $\nu = 1$ )	2.3431	2.3519	2.4571	2.3097	2.3199
	$t$ -Student( $\nu = 2$ )	2.5987	2.5232	2.5684	2.4876	2.4566
	$t$ -Student( $\nu = 3$ )	2.7445	2.6172	2.6318	2.5952	2.5434
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	2.9680	2.7964	2.7634	2.8172	2.7508
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	3.0839	2.8677	2.8141	2.9116	2.8394
	Normal	3.5963	3.4181	3.3693	3.4566	3.3953
7	$t$ -Student( $\nu = 1$ )	2.6662	2.5606	2.5910	2.5458	2.4957
	$t$ -Student( $\nu = 2$ )	2.9432	2.7306	2.7142	2.7282	2.6441
	$t$ -Student( $\nu = 3$ )	3.0735	2.8265	2.7821	2.8402	2.7411
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	3.3007	3.0128	2.9310	3.0705	2.9670
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	3.3970	3.0824	2.9802	3.1643	3.0588
	Normal	3.4577	3.3549	3.3463	3.3593	3.3154
8	$t$ -Student( $\nu = 1$ )	2.3465	2.3814	2.5242	2.3352	2.3771
	$t$ -Student( $\nu = 2$ )	2.6964	2.6196	2.7013	2.5766	2.5661
	$t$ -Student( $\nu = 3$ )	2.8693	2.7642	2.8055	2.7360	2.6928
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	3.2140	3.0439	3.0235	3.0587	2.9991
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	3.3708	3.1529	3.1004	3.1992	3.1312
	Normal	3.8349	3.6049	3.5295	3.6620	3.5857
9	$t$ -Student( $\nu = 1$ )	2.0600	2.1978	2.4147	2.1378	2.2328
	$t$ -Student( $\nu = 2$ )	2.3796	2.4602	2.6179	2.3813	2.4203
	$t$ -Student( $\nu = 3$ )	2.6007	2.6241	2.7397	2.5498	2.5579
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	2.9885	2.9560	3.0049	2.9116	2.8914
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	3.1780	3.0961	3.1039	3.0902	3.0532
	Normal	3.6869	3.5509	3.5244	3.5707	3.5177
10	$t$ -Student( $\nu = 1$ )	2.5227	2.5223	2.6282	2.4934	2.5153
	$t$ -Student( $\nu = 2$ )	2.8793	2.7790	2.8406	2.7481	2.7239
	$t$ -Student( $\nu = 3$ )	3.1173	2.9516	2.9702	2.9287	2.8761
	$t$ -Student( $\nu \sim \text{Gamma}(2, 0.1)$ )	3.4846	3.2965	3.2572	3.3218	3.2516
	$t$ -Student( $\nu \sim \text{Exp}(0.001)$ )	3.7063	3.4409	3.3642	3.5086	3.4224
	Normal	3.9907	3.7542	3.6676	3.8117	3.7275



## 4. Aplicación

### 4.1. Aplicación de un modelo jerárquico normal

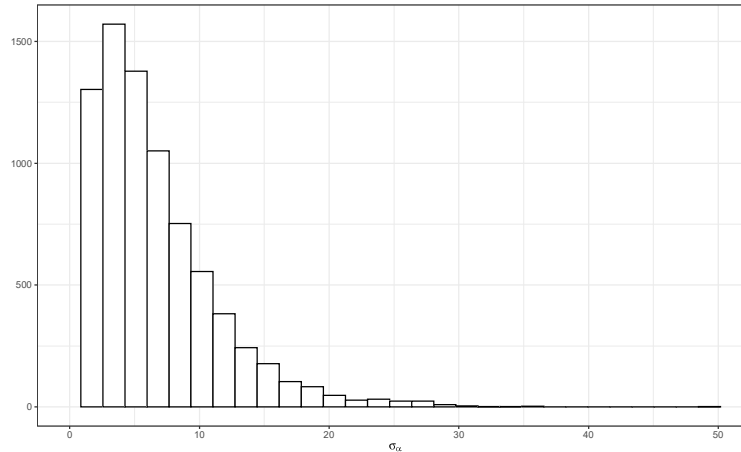
Consideramos el modelo jerárquico lineal dado en (3-2) el cual se aplicó al ejemplo de los colegios analizado en Gelman (2006b), asociado al efecto relativo de un programa de entrenamiento sobre una prueba de actitud escolar en ocho colegios los cuales se presentan en la Tabla 4-1. En el modelo, los parámetros  $\alpha_j$  para  $j = 1, 2, \dots, 8$  representan el efecto relativo del programa de entrenamiento en el  $j$ -ésimo colegio y  $\sigma_\alpha$  la desviación estándar de dichos efectos entre los colegios. Los efectos son medidos como puntaje en la prueba y están entre 200 y 800 con un promedio aproximado de 500, por lo cual, el mayor efecto está alrededor de los 300 puntos y para la desviación estándar entre estos efectos,  $\sigma_\alpha$  se define como límite superior de 100. Gelman (2006b) expone e ilustra las principales dificultades que se presentan cuando se realiza el ajuste de los efectos aleatorios con una distribución normal y la distribución a priori para  $\sigma_\alpha$  corresponde a  $IG(0.001, 0.001)$ . Entre los cuales resalta, los inconvenientes cuando se tiene un número pequeño de grupos ( $J < 5$ ) y las alternativas para restringir el parámetro  $\sigma_\alpha$  como lo son distribuciones pertenecientes a la familia *folded-noncentral-t* como la *half-t* y la *half-Cauchy* presentes en Gelman y Hill (2006) (ver sección 1.2.1) .

Tabla 4-1.: Datos colegios Gelman (2006b)

	Colegio	Estimación	Desviación estándar
1	A	28.39	14.90
2	B	7.94	10.20
3	C	-2.75	16.30
4	D	6.82	11.00
5	E	-0.64	9.40
6	F	0.63	11.40
7	G	18.01	10.40
8	H	12.16	17.60

### 4.1.1. Aplicación con todos los $J = 8$ colegios

Ajustando el modelo (3-2) en el nivel 1 con los datos dados en la Tabla **4-1** y utilizando como distribuciones a priori para  $\sigma_\alpha^2$  para el nivel 2 las listadas en la sección **2.2**. En las figuras **4-1** hasta **4-4** se presentan los histogramas de frecuencias para la distribución posterior del parámetro  $\sigma_\alpha$  ajustado por cuatro diferentes alternativas para la distribuciones a priori a partir de métodos Monte Carlo por cadenas de Markov basado en 10000 simulaciones del modelo ajustado en **jags** bajo el paquete **R2jags** con un quemado inicial de 2000 iteraciones.



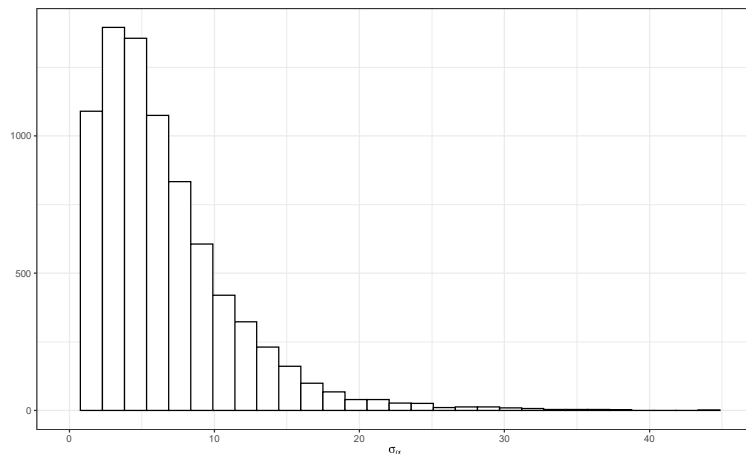
**Figura 4-1.:** Histograma de la distribución posterior de  $\sigma_\alpha$  con distribución a priori para  $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, 25^2)$

En la Figura **4-1** donde la distribución a priori utilizada en el ajuste para  $\sigma_\alpha^2$  corresponde a  $SBeta2(1/2, 1/2, 25^2)$  y la cual es equivalente a realizar el ajuste con una *half*-Cauchy de parámetro 25 para  $\sigma_\alpha$  se evidencia que, a partir de valores de  $\sigma_\alpha$  superiores a 15 se empieza a extender una pequeña cola, por lo cual, los valores para  $\sigma_\alpha$  son admisibles en términos de la cota superior, esto es,  $\sigma_\alpha < 100$ . Las estimaciones a posteriori obtenidas del ajuste del modelo (3-2) con los datos asociados a la Tabla **4-1** y  $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, 25^2)$  se presentan a continuación:

**Tabla 4-2.:** Resultados del ajuste para  $J = 8$  y  $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, 25^2)$ 

Parámetro	Estimación	Desviación	2.5 %	97.5 %
$\mu$	3.774	2.435	-0.560	8.980
$\sigma_\alpha$	13.487	23.414	0.000	86.308
$\alpha_1$	10.956	9.118	-1.212	35.551
$\alpha_2$	7.657	5.922	-3.131	20.778
$\alpha_3$	6.149	7.145	-9.747	19.706
$\alpha_4$	7.456	6.135	-3.998	21.138
$\alpha_5$	5.671	6.031	-7.726	17.238
$\alpha_6$	6.356	6.210	-6.995	18.728
$\alpha_7$	9.576	6.703	-0.832	25.781
$\alpha_8$	7.657	6.613	-5.025	21.863
<i>deviance</i>	60.474	2.095	57.177	65.688

En el caso del ajuste  $\sigma_\alpha^2 \sim SBeta2(1, 1, 25^2)$  presentado en la Figura 4-2 una pequeña cola a derecha se empieza a extender a partir de valores de  $\sigma_\alpha$  menores de 20 considerados como admisibles y comparando con los resultados del ajuste  $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, 25^2)$  se presenta una media con mayor cuantía.

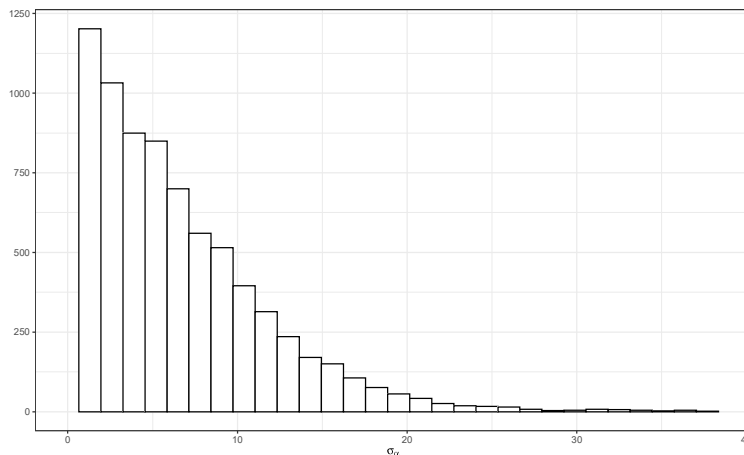
**Figura 4-2.:** Histograma de la distribución posterior de  $\sigma_\alpha$  con distribución a priori para  $\sigma_\alpha^2 \sim SBeta2(1, 1, 25^2)$ 

Las estimaciones a posteriori obtenidas del ajuste del modelo (3-2) con los datos asociados a la Tabla 4-1 y  $\sigma_\alpha^2 \sim SBeta2(1, 1, 25^2)$  se presentan a continuación:

**Tabla 4-3.:** Resultados del ajuste para  $J = 8$  y  $\sigma_\alpha^2 \sim \text{SBeta2}(1, 1, 25^2)$ 

Parámetro	Estimación	Desviación	2.5 %	97.5 %
$\mu$	3.774	2.435	-0.560	8.980
$\sigma_\alpha$	3.959	2.801	-1.482	9.524
$\alpha_1$	64.189	131.522	0.471	361.290
$\alpha_2$	13.986	10.510	-2.946	39.847
$\alpha_3$	8.013	7.046	-5.963	22.532
$\alpha_4$	4.942	9.627	-17.256	22.079
$\alpha_5$	7.464	7.337	-7.948	22.337
$\alpha_6$	5.432	7.852	-11.867	19.909
$\alpha_7$	11.798	7.855	-1.778	29.186
$\alpha_8$	9.116	9.542	-8.782	30.402
<i>deviance</i>	60.611	2.568	56.847	66.888

Continuando, tenemos que cuando se realizó el ajuste con una distribución a priori Uniforme presentado en la Figura 4-8, la cual Gelman (2006b) recomienda para  $J \geq 5$  grupos, tenemos que la cola a derecha se empieza a extender a partir de valores de  $\sigma_\alpha$  superiores a cuatro y se presenta una gran concentración en el intervalo  $(0, 4)$  pero no con picos altos alrededor de cero, es decir, esta distribución a priori captura desviaciones pequeñas para los  $J = 8$  grupos que se están comparando con valores de  $\sigma_\alpha$  en un rango admisible.

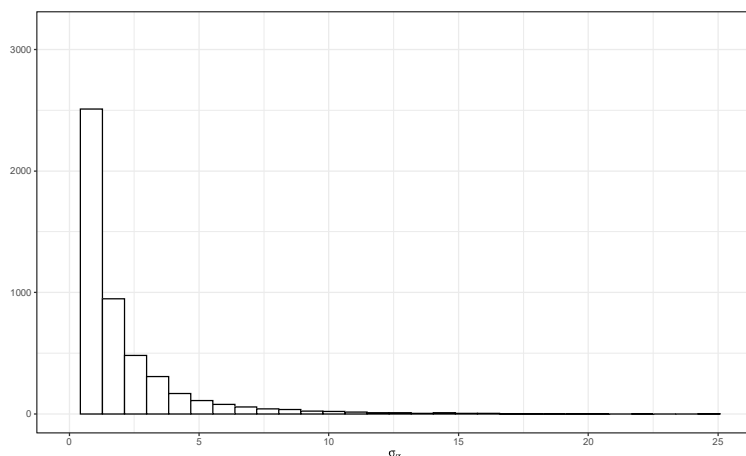
**Figura 4-3.:** Histograma de la distribución posterior de  $\sigma_\alpha$  con distribución a priori para  $\sigma_\alpha^2 \sim \text{Uniforme}(0, 100)$ 

Las estimaciones a posteriori obtenidas del ajuste del modelo (3-2) con los datos asociados a la Tabla 4-1 y  $\sigma_\alpha^2 \sim \text{Uniforme}(0, 100)$  se presentan a continuación:

**Tabla 4-4.:** Resultados del ajuste para  $J = 8$  y  $\sigma_\alpha^2 \sim \text{Uniforme}(0, 100)$ 

Parámetro	Estimación	Desviación	2.5 %	97.5 %
$\mu$	7.815	5.584	-3.434	18.412
$\sigma_\alpha$	6.597	5.871	0.159	22.032
$\alpha_1$	11.665	8.754	-2.338	32.972
$\alpha_2$	7.804	6.474	-5.338	20.383
$\alpha_3$	6.068	8.143	-12.361	20.386
$\alpha_4$	7.637	6.834	-6.378	20.988
$\alpha_5$	5.305	6.691	-9.383	17.124
$\alpha_6$	6.030	7.163	-10.091	18.613
$\alpha_7$	10.576	7.076	-1.859	26.384
$\alpha_8$	8.345	8.138	-8.321	25.020
<i>deviance</i>	60.599	2.274	57.055	66.203

En contraposición a lo presentado con el ajuste  $\sigma_\alpha^2 \sim \text{Uniforme}(0, 100)$  tenemos el comportamiento del ajuste  $\sigma_\alpha^2 \sim IG(0.001, 0.001)$  el cual se puede apreciar en **4-4** en la cual el rango de valores se encuentra concentrado en un intervalo pequeño, con un pico muy alto alrededor de cero y una pequeña cola a derecha extendiéndose a partir de valores de  $\sigma_\alpha$  superiores a 5, con lo cual podemos establecer que la inferencia se va a ver afectada debido al encogimiento de los efectos en cada grupo,  $\alpha_j$ . Por tanto, podemos establecer con el análisis gráfico de las distribuciones posteriori que, se obtienen valores de  $\sigma_\alpha$  admisibles, menores que los reportados en Gelman (2006b), Gelman et al. (2013) y Pérez et al. (2017) ajustando los efectos aleatorios a una distribución  $t$  de Student y adicional tiene en cuenta presencia de valores de desviaciones cercanos a cero.

**Figura 4-4.:** Histograma de la distribución posterior de  $\sigma_\alpha$  con distribución a priori para  $\sigma_\alpha^2 \sim IG(0.001, 0.001)$

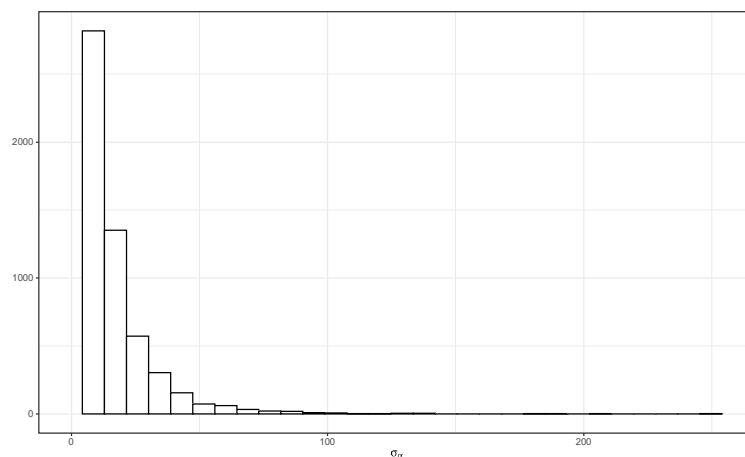
**Tabla 4-5.:** Resultados del ajuste para  $J = 8$  y  $\sigma_\alpha^2 \sim IG(0.001, 0.001)$ 

Las estimaciones a posteriori obtenidas del ajuste del modelo (3-2) con los datos asociados a la Tabla 4-1 y  $\sigma_\alpha^2 \sim IG(0.001, 0.001)$  están dadas por:

Parámetro	Estimación	Desviación	2.5 %	97.5 %
$\mu$	3.614	2.240	-0.583	8.054
$\sigma_\alpha$	4.876	24.457	0.001	39.254
$\alpha_1$	8.748	7.055	-1.828	28.959
$\alpha_2$	7.131	5.525	-2.563	17.161
$\alpha_3$	6.431	5.663	-6.494	16.395
$\alpha_4$	7.223	4.792	-1.830	16.403
$\alpha_6$	6.975	4.798	-2.254	16.336
$\alpha_7$	8.230	5.518	-1.286	21.118
$\alpha_8$	8.011	6.285	-1.582	23.665
<i>deviance</i>	60.507	1.810	57.603	65.210

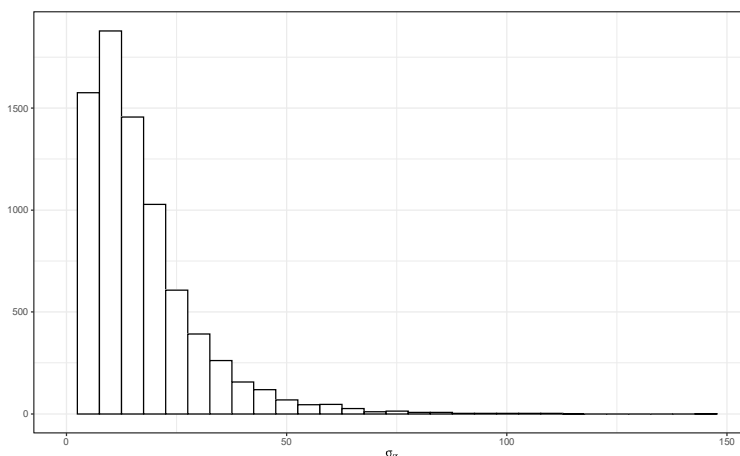
#### 4.1.2. Aplicación con los tres primeros colegios

Siguiendo lo realizado en Gelman (2006b) y análogamente a lo realizado en la sección anterior, se ajusta el modelo (3-2) en el nivel 1 con los datos dados en la Tabla 4-1 y utilizando como distribuciones a priori para  $\sigma_\alpha^2$  para el nivel 2 las listadas en la sección 2.3 para los tres primeros grupos A, B y C. En las figuras 4-5 hasta 4-6 se presentan los histogramas de frecuencias para la distribución posterior del parámetro  $\sigma_\alpha$  ajustado por cuatro diferentes alternativas para la distribuciones a priori a partir de métodos MCMC basado en 10000 simulaciones del modelo ajustado en **jags** bajo el paquete **R2jags** con un quemado inicial de 2000 iteraciones.

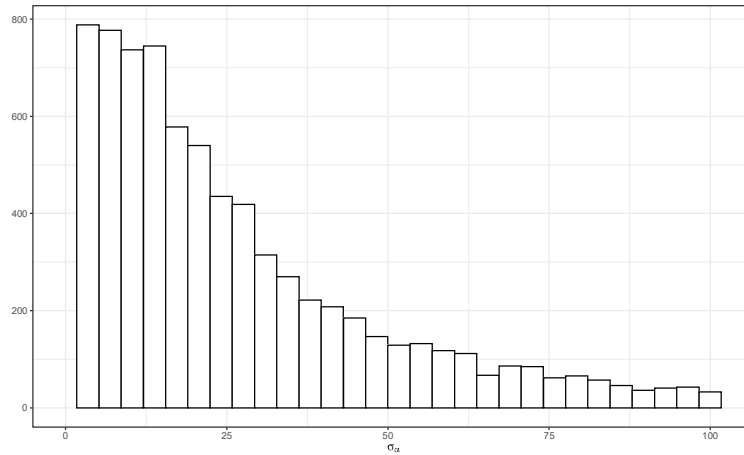


**Figura 4-5.:** Histograma de la distribución posterior de  $\sigma_\alpha$  con distribución a priori para  $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, 25^2)$  para  $J = 3$

En la Figura 4-5 donde la distribución a priori utilizada en el ajuste para  $\sigma_\alpha^2$  es equivalente a realizar el ajuste con una *half*-Cauchy de parámetro 25 para  $\sigma_\alpha$ , evidencia que, a partir de valores de  $\sigma_\alpha$  superiores a 50 se empieza a extender una pequeña cola, por lo cual, los valores para  $\sigma_\alpha$  son admisibles en términos de la cota superior, esto es,  $\sigma_\alpha < 100$ . En el caso del ajuste  $\sigma_\alpha^2 \sim SBeta2(1, 1, 25^2)$  presentado en la Figura 4-6, se observa una pequeña cola a derecha se empieza a extender a partir de valores de  $\sigma_\alpha$  menores de 50 considerados como admisibles y comparando con los resultados del ajuste  $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, 25^2)$  no se presenta tanta concentración en el intervalo pequeño (0, 5) aproximadamente.

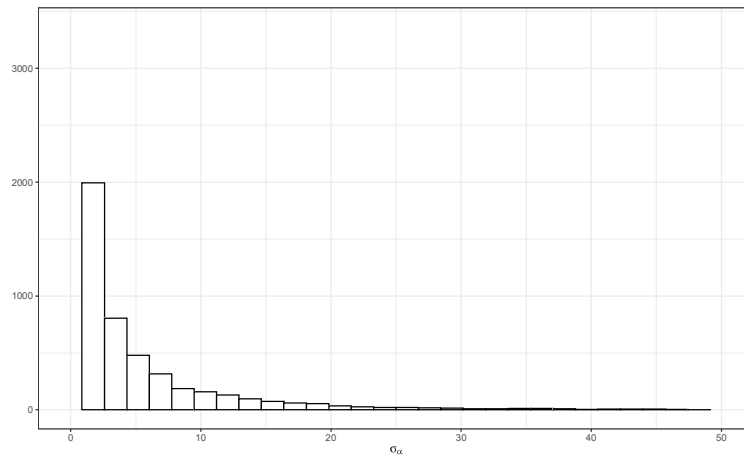


**Figura 4-6.:** Histograma de la distribución posterior de  $\sigma_\alpha$  con distribución a priori para  $\sigma_\alpha^2 \sim SBeta2(1, 1, 25^2)$  para  $J = 3$



**Figura 4-7.:** Histograma de la distribución posterior de  $\sigma_\alpha$  con distribución a priori para  $\sigma_\alpha^2 \sim \text{Uniforme}(0, 100)$  para  $J = 3$

Continuando, tenemos que cuando se realiza el ajuste con una distribución a priori Uniforme presentado en la Figura 4-7, la cual Gelman (2006b) recomienda para  $J \geq 5$  grupos porque cuando el número de grupos es pequeño los datos proporcionan poca información acerca de la varianza entre grupos, observamos que se presenta una alta concentración en el intervalo de valores para  $\sigma_\alpha$  de  $(0, 10)$  y sin presencia de cola a derecha extendiéndose a partir de valores mayores a 10, lo cual nos da indicio que ajustando los efectos a partir de una distribución  $t$  de Student se permite tener presente en el modelamiento jerárquico poca variación en un número pequeño de grupos y así  $\sigma_\alpha$  posee valores en un rango admisible.



**Figura 4-8.:** Histograma de la distribución posterior de  $\sigma_\alpha$  con distribución a priori para  $\sigma_\alpha^2 \sim \text{IG}(0.001, 0.001)$  para  $J = 3$

El ajuste  $\sigma_\alpha^2 \sim \text{IG}(0.001, 0.001)$  para  $J = 3$  presenta alta concentración en intervalo pequeño y una cola extensa a derecha a partir de valores de  $\sigma_\alpha > 20$ .



# 5. Conclusiones y Perspectivas futuras

## 5.1. Conclusiones

El presente trabajo propone e implementa el análisis conjunto de modelar los efectos aleatorios en un modelo jerárquico normal a partir de una distribución  $t$ -Student desde un enfoque bayesiano utilizando diferentes distribuciones a priori candidatas asociadas al parámetro de escala. Los resultados obtenidos vía simulación estadística nos permitieron examinar el comportamiento de las estimaciones de las medidas de error global de estimación y la raíz del error cuadrático medio para los efectos aleatorios del modelo jerárquico lineal teniendo en cuenta diferentes variaciones para el parámetro grados de libertad de la distribución  $t$ -Student.

A partir de los resultados obtenidos del estudio de simulación se evidencia la tendencia de incremento del error cuadrático de estimación a partir del incremento de los grados de libertad asociados a la distribución  $t$ -Student para todas las distribuciones a priori candidatas analizadas. De acuerdo con las situaciones de interés planteadas con respecto a la presencia de valores atípicos moderados y un gran valor atípico, los valores más altos para las medidas de error global de estimación y raíz de error cuadrático medio en la mayoría de las veces se presentaban en el ajuste cuando se utilizaba una distribución Gamma-Inversa (0.001, 0.001) como distribución a priori para el parámetro de varianza.

Con el estudio de simulación se muestra que el error cuadrático de estimación de los efectos que representa un gran valor atípico es mucho menor cuando se ajustan los efectos aleatorios a partir de una distribución  $t$ -Student para todas las alternativas de distribuciones a priori para el parámetro de escala en consideración que los reportados cuando los efectos siguen una distribución normal y en general el ajuste con la elección natural para los parámetros de escala, como lo es la distribución Gamma-inversa  $(\epsilon, \epsilon)$  con  $\epsilon$  pequeño presenta el mayor de la medida de error cuadrático medio presente a nivel global e individual. Adicionalmente, en el escenario con valores atípicos moderados en función del número de grupos, dependiendo de los grados de libertad en la distribución  $t$ -Student, la variación porcentual con los resultados reportados en Pérez et al. (2017) cuando los efectos siguen una distribución normal es baja en función del número de grupos que se esté contrastando.

Los resultados obtenidos con la aplicación de los datos de los colegios ilustran que para valores pequeños de desviaciones para los efectos en términos de asociar la distribución  $t$ -Student para los efectos y las diferentes alternativas de distribuciones a priori al parámetro

de escala se obtienen valores de  $\sigma_\alpha$  admisibles tanto para un número pequeño de grupos ( $J = 3$ ) como para un número de grupos considerable ( $J = 8$ ).

## 5.2. Perspectivas futuras

Algunas propuestas de extensión que se pueden considerar viables a partir del presente trabajo de investigación corresponden a:

- En Villa y Walker (2014) encontramos diferentes propuestas de distribuciones a priori para los grados de libertad de una distribución  $t$ -Student, por lo cual, se pueden considerar y realizar un análisis similar en términos del comportamiento conjunto de modelar los efectos aleatorios en el modelo jerárquico a partir de una distribución  $t$ -Student con diferentes alternativas a priori para modelar el parámetro de escala.
- Realizar una análisis similar pero con distribuciones  $t$ -Student sesgadas y se puede trabajar con los valores analizados en Juárez y Steel (2010) para el parámetro de sesgo.

# A. Anexo: Error global de estimación por escenario

## A.1. Error global de estimación para el escenario con efectos de magnitud similar

Tabla A-1.: Error de estimación global para todas la situación  $S_I$  con  $J = 3$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	0.9206	1.1088	2.3714	0.8523	0.8568	1.7749
2	0.9084	1.1090	1.9800	0.7952	0.7782	2.1054
3	0.9146	1.1257	1.9399	0.7765	0.7526	1.9092
4	0.8918	2.2158	2.3213	0.7681	0.7405	2.4443
5	1.0160	1.1145	1.8448	0.7668	0.7332	2.1702
10	1.0884	1.1080	2.0089	0.7584	0.7233	2.2513
30	0.8860	1.1120	1.8844	0.7531	0.7179	2.0243
$\nu \sim exp(0.001)$	0.9275	1.1137	1.8602	0.7513	0.7176	2.0461
$\nu \sim gamma(2, 0.1)$	0.9133	1.1110	1.9365	0.7559	0.7248	2.0750

Tabla A-2.: Error de estimación global para todas la situación  $S_I$  con  $J = 4$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	0.8433	1.1580	1.5832	0.7913	0.7988	1.1603
2	0.8391	1.1666	1.5903	0.7602	0.7479	1.1699
3	0.8347	1.1696	1.5910	0.7514	0.7359	1.1868
4	0.8459	1.1728	1.5920	0.7479	0.7308	1.1748
5	0.8329	1.1730	1.5992	0.7439	0.7281	1.1793
10	0.8402	1.1738	1.5939	0.7403	0.7238	1.1776
30	0.8336	1.1760	1.5977	0.7373	0.7226	1.1837
$\nu \sim exp(0.001)$	0.8273	1.1759	1.5977	0.7392	0.7231	1.1859
$\nu \sim gamma(2, 0.1)$	0.8365	1.1758	1.5965	0.7453	0.7247	1.1798

**Tabla A-3.:** Error de estimación global para todas la situación  $S_I$  con  $J = 5$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	1.0320	1.2618	1.6479	1.0244	1.1275	1.2628
2	1.0211	1.2715	1.6544	0.9891	1.0546	1.2707
3	1.0251	1.2723	1.6594	0.9790	1.0327	1.2759
4	1.0233	1.2767	1.6602	0.9764	1.0218	1.2755
5	1.0265	1.2845	1.6625	0.9735	1.0151	1.2798
10	1.0287	1.2812	1.6678	0.9695	1.0084	1.2831
30	1.0312	1.2819	1.6697	0.9688	1.0005	1.2857
$\nu \sim exp(0,001)$	1.0281	1.2814	1.6696	0.9652	0.9999	1.2851
$\nu \sim gamma(2, 0.1)$	1.0334	1.2814	1.6680	0.9675	1.0053	1.2825

**Tabla A-4.:** Error de estimación global para todas la situación  $S_I$  con  $J = 6$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	1.0819	1.2959	1.6723	1.0745	1.0994	1.2961
2	1.0646	1.3054	1.6851	1.0306	1.0331	1.3022
3	1.0689	1.3082	1.6915	1.0230	1.0137	1.3088
4	1.0688	1.3140	1.6953	1.0189	1.0069	1.3151
5	1.0723	1.3159	1.6972	1.0185	1.0041	1.3164
10	1.0781	1.3186	1.7025	1.0134	0.9974	1.3225
30	1.0823	1.3220	1.7081	1.0133	0.9935	1.3254
$\nu \sim exp(0,001)$	1.0816	1.3246	1.7096	1.0117	0.9926	1.3227
$\nu \sim gamma(2, 0.1)$	1.0789	1.3187	1.7049	1.0129	0.9978	1.3221

**Tabla A-5.:** Error de estimación global para todas la situación  $S_I$  con  $J = 7$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	1.1461	1.3411	1.7095	1.1429	1.1747	1.3421
2	1.1245	1.3500	1.7272	1.1027	1.1031	1.3514
3	1.1296	1.3551	1.7333	1.0892	1.0814	1.3595
4	1.1290	1.3616	1.7388	1.0879	1.0743	1.3618
5	1.1414	1.3640	1.7435	1.0825	1.0694	1.3666
10	1.1387	1.3706	1.7519	1.0769	1.0619	1.3744
30	1.1383	1.3763	1.7585	1.0753	1.0601	1.3761
$\nu \sim exp(0,001)$	1.1434	1.3760	1.7598	1.0777	1.0583	1.3778
$\nu \sim gamma(2, 0.1)$	1.1399	1.3704	1.7532	1.0790	1.0613	1.3696

**Tabla A-6.:** Error de estimación global para todas la situación  $S_I$  con  $J = 8$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	1.2734	1.4350	1.7640	1.2777	1.3098	1.4351
2	1.2553	1.4412	1.7812	1.2383	1.2370	1.4406
3	1.2654	1.4459	1.7919	1.2239	1.2152	1.4466
4	1.2609	1.4514	1.7984	1.2225	1.2073	1.4518
5	1.2675	1.4551	1.8022	1.2194	1.2030	1.4550
10	1.2604	1.4584	1.8126	1.2150	1.1929	1.4617
30	1.2688	1.4638	1.8180	1.2121	1.1889	1.4653
$\nu \sim exp(0,001)$	1.2686	1.4642	1.8197	1.2106	1.1895	1.4677
$\nu \sim gamma(2, 0.1)$	1.2629	1.4627	1.8142	1.2118	1.1960	1.4603

**Tabla A-7.:** Error de estimación global para todas la situación  $S_I$  con  $J = 9$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	1.2807	1.4402	1.7581	1.2853	1.3216	1.4398
2	1.2619	1.4457	1.7772	1.2402	1.2433	1.4452
3	1.2623	1.4520	1.7867	1.2272	1.2180	1.4513
4	1.2688	1.4559	1.7966	1.2221	1.2087	1.4581
5	1.2612	1.4572	1.7942	1.2209	1.2093	1.4606
10	1.2725	1.4645	1.8142	1.2149	1.1966	1.4693
30	1.2646	1.4686	1.8174	1.2152	1.1923	1.4739
$\nu \sim exp(0.001)$	1.2717	1.4752	1.8194	1.2121	1.1883	1.4734
$\nu \sim gamma(2, 0.1)$	1.2668	1.4692	1.8121	1.2120	1.1954	1.4703

**Tabla A-8.:** Error de estimación global para todas la situación  $S_I$  con  $J = 10$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	1.3105	1.5871	1.8548	1.4790	1.5094	1.4537
2	1.2865	1.5873	1.8738	1.4375	1.4397	1.4624
3	1.2893	1.5941	1.8871	1.4263	1.4188	1.4732
4	1.2905	1.5972	1.8925	1.4216	1.4092	1.4803
5	1.2945	1.6045	1.8987	1.4174	1.4068	1.4804
10	1.3003	1.6111	1.9086	1.4158	1.3973	1.4909
30	1.3019	1.6111	1.9188	1.4145	1.3931	1.4934
$\nu \sim exp(0.001)$	1.3001	1.6166	1.9221	1.4119	1.3937	1.4954
$\nu \sim gamma(2, 0.1)$	1.2939	1.6117	1.9121	1.4144	1.3984	1.4885

## A.2. Error global de estimación para el escenario con algunos valores atípicos moderados

**Tabla A-9.:** Error de estimación global para todas la situación  $S_{II}$  con  $J = 3$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	1.7834	1.7247	2.1779	1.7491	1.6725	2.7414
2	1.8055	1.7285	2.2990	1.7034	1.6673	3.2290
3	1.8010	1.7333	2.8813	1.7069	1.6689	3.0687
4	1.8261	1.7462	2.2056	1.7097	1.6701	2.4188
5	1.8060	1.7825	2.1735	1.7115	1.6714	3.2171
10	1.8241	1.7353	2.8006	1.7172	1.6741	2.6410
30	1.8317	1.7323	2.4595	1.7199	1.6754	3.2800
$\nu \sim exp(0.001)$	2.1317	1.7479	2.2525	1.7204	1.6794	2.5119
$\nu \sim gamma(2, 0.1)$	1.9367	1.7376	2.2803	1.7155	1.6755	3.3751

**Tabla A-10.:** Error de estimación global para todas la situación  $S_{II}$  con  $J = 4$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	1.9022	1.8651	1.9572	1.8425	1.8092	1.9667
2	1.9124	1.8646	1.9591	1.8375	1.7933	1.9694
3	1.9184	1.8613	1.9629	1.8356	1.7885	1.9708
4	1.9204	1.8632	1.9647	1.8358	1.7880	1.9735
5	1.9188	1.8612	1.9623	1.8353	1.7854	1.9755
10	1.9208	1.8646	1.9662	1.8360	1.7860	1.9758
30	1.9198	1.8661	1.9625	1.8394	1.7857	1.9806
$\nu \sim exp(0,001)$	1.9202	1.8675	1.9708	1.8397	1.7872	1.9788
$\nu \sim gamma(2, 0.1)$	1.9202	1.8653	1.9635	1.8383	1.7858	1.9777

**Tabla A-11.:** Error de estimación global para todas la situación  $S_{II}$  con  $J = 5$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	2.0682	2.0114	2.0930	1.9873	1.9374	2.0973
2	2.1020	2.0215	2.1037	1.9928	1.9321	2.1134
3	2.1134	2.0260	2.1090	1.9990	1.9327	2.1233
4	2.1205	2.0315	2.1122	2.0039	1.9358	2.1278
5	2.1275	2.0330	2.1155	2.0054	1.9372	2.1326
10	2.1343	2.0419	2.1180	2.0147	1.9433	2.1415
30	2.1398	2.0430	2.1239	2.0203	1.9480	2.1475
$\nu \sim exp(0,001)$	2.1449	2.0453	2.1260	2.0229	1.9495	2.1486
$\nu \sim gamma(2, 0.1)$	2.1363	2.0397	2.1220	2.0145	1.9448	2.1415

**Tabla A-12.:** Error de estimación global para todas la situación  $S_{II}$  con  $J = 6$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	2.7145	2.4905	2.4174	2.5250	2.4127	2.5661
2	2.7255	2.4709	2.4174	2.5073	2.3979	2.5601
3	2.7185	2.4691	2.4194	2.5026	2.3942	2.5604
4	2.7128	2.4673	2.4196	2.5047	2.3968	2.5617
5	2.7159	2.4698	2.4212	2.5085	2.3970	2.5655
10	2.7286	2.4693	2.4253	2.5166	2.4037	2.5719
30	2.7234	2.4736	2.4303	2.5229	2.4092	2.5783
$\nu \sim exp(0,001)$	2.7319	2.4729	2.4315	2.5241	2.4151	2.5802
$\nu \sim gamma(2, 0.1)$	2.7274	2.4698	2.4250	2.5151	2.4048	2.5706

**Tabla A-13.:** Error de estimación global para todas la situación  $S_{II}$  con  $J = 7$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	2.9684	2.7497	2.6348	2.7859	2.6497	2.8030
2	2.9839	2.7276	2.6314	2.7725	2.6423	2.7914
3	2.9746	2.7236	2.6336	2.7727	2.6448	2.7910
4	2.9804	2.7255	2.6359	2.7752	2.6478	2.7938
5	2.9813	2.7256	2.6369	2.7779	2.6496	2.7950
10	2.9950	2.7270	2.6415	2.7882	2.6614	2.8024
30	2.9991	2.7311	2.6439	2.7942	2.6693	2.8073
$\nu \sim exp(0,001)$	2.9920	2.7287	2.6458	2.7989	2.6740	2.8092
$\nu \sim gamma(2, 0.1)$	2.9960	2.7272	2.6412	2.7873	2.6633	2.8023

**Tabla A-14.:** Error de estimación global para todas la situación  $S_{II}$  con  $J = 8$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	3.4303	3.1365	2.9874	3.1904	3.0407	3.1394
2	3.3687	3.0795	2.9614	3.1348	3.0014	3.0745
3	3.3497	3.0546	2.9496	3.1122	2.9892	3.0523
4	3.3319	3.0475	2.9457	3.1053	2.9823	3.0446
5	3.3171	3.0421	2.9440	3.1022	2.9808	3.0399
10	3.3233	3.0339	2.9406	3.0986	2.9787	3.0337
30	3.3113	3.0295	2.9375	3.1008	2.9800	3.0320
$\nu \sim exp(0.001)$	3.3110	3.0298	2.9389	3.1010	2.9840	3.0302
$\nu \sim gamma(2, 0.1)$	3.3166	3.0322	2.9390	3.1012	2.9800	3.0324

**Tabla A-15.:** Error de estimación global para todas la situación  $S_{II}$  con  $J = 9$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	3.4324	3.1656	3.0292	3.2099	3.0654	3.1661
2	3.3812	3.1129	3.0134	3.1606	3.0370	3.1123
3	3.3656	3.0985	3.0097	3.1462	3.0283	3.0987
4	3.3555	3.0899	3.0078	3.1415	3.0279	3.0939
5	3.3468	3.0937	3.0087	3.1405	3.0275	3.0908
10	3.3368	3.0920	3.0098	3.1402	3.0308	3.0879
30	3.3333	3.0895	3.0132	3.1494	3.0360	3.0892
$\nu \sim exp(0.001)$	3.3359	3.0902	3.0157	3.1477	3.0407	3.0905
$\nu \sim gamma(2, 0.1)$	3.3359	3.0907	3.0091	3.1446	3.0341	3.0902

**Tabla A-16.:** Error de estimación global para todas la situación  $S_{II}$  con  $J = 10$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	3.5456	3.2310	3.1009	3.2670	3.1282	3.2916
2	3.4723	3.1739	3.0838	3.2113	3.0978	3.2243
3	3.4521	3.1577	3.0790	3.1986	3.0904	3.2022
4	3.4452	3.1504	3.0797	4.9292	3.0879	3.1941
5	3.4310	3.1509	3.0824	4.9274	3.0886	3.1896
10	3.4215	3.1531	3.0868	4.9207	3.0946	3.1852
30	3.4210	3.1537	3.0914	4.9215	3.1015	3.1866
$\nu \sim exp(0.001)$	3.4222	3.1568	3.0923	4.9220	3.1043	3.1878
$\nu \sim gamma(2, 0.1)$	3.4267	3.1518	3.0874	4.9229	3.0949	3.1888

### A.3. Error global de estimación para el escenario de un gran valor atípico

**Tabla A-17.:** Error de estimación global para todas la situación  $S_{III}$  con  $J = 3$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	2.5341	2.2945	3.1920	2.3195	2.2461	3.6158
2	2.7142	2.8051	2.9083	2.4139	2.3022	3.7760
3	2.6659	2.3979	3.2199	2.4264	2.3405	4.6377
4	2.6988	2.4100	2.8630	2.9772	2.3578	3.8666
5	2.7397	2.4255	2.7334	2.4764	2.3697	3.9734
10	2.9288	2.4459	2.7016	2.5153	2.4056	4.0252
30	2.8379	2.4767	2.9607	2.5361	2.4373	4.6165
$\nu \sim exp(0,001)$	2.8263	3.3680	2.8585	2.5415	2.4456	4.1091
$\nu \sim gamma(2, 0,1)$	2.8672	3.1445	2.9635	2.5108	2.4161	3.7395

**Tabla A-18.:** Error de estimación global para todas la situación  $S_{III}$  con  $J = 4$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	2.6666	2.4545	2.3862	2.4791	2.3975	2.4553
2	2.7778	2.4848	2.3956	2.5332	2.4278	2.4844
3	2.8252	2.4931	2.3979	2.5564	2.4488	2.4952
4	2.8248	2.5025	2.4015	2.5717	2.4611	2.5027
5	2.8363	2.5053	2.4028	2.5832	2.4704	2.5057
10	2.8556	2.5120	2.4039	2.6006	2.4853	2.5079
30	2.8661	2.5144	2.4059	2.6114	2.4970	2.5168
$\nu \sim exp(0,001)$	2.8537	2.5189	2.4075	2.6171	2.5018	2.5157
$\nu \sim gamma(2, 0,1)$	2.8465	2.5115	2.4049	2.6016	2.4878	2.5122

**Tabla A-19.:** Error de estimación global para todas la situación  $S_{III}$  con  $J = 5$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	2.4161	2.4362	2.5059	2.3901	2.3992	2.4380
2	2.5917	2.5216	2.5444	2.4954	2.4664	2.5239
3	2.6683	2.5621	2.5611	2.5474	2.5053	2.5626
4	2.7135	2.5823	2.5710	2.5802	2.5290	2.5846
5	2.7329	2.5938	2.5758	2.5991	2.5476	2.5979
10	2.7683	2.6191	2.5888	2.6382	2.5794	2.6182
30	2.7901	2.6315	2.5952	2.6617	2.6012	2.6321
$\nu \sim exp(0,001)$	2.8108	2.6360	2.5987	2.6684	2.6094	2.6370
$\nu \sim gamma(2, 0,1)$	2.7710	2.6189	2.5904	2.6399	2.5823	2.6191



### A.3 Error global de estimación para el escenario de un gran valor atípico 57

**Tabla A-20.:** Error de estimación global para todas la situación  $S_{III}$  con  $J = 6$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	2.4955	2.5267	2.6154	2.4812	2.4930	2.5266
2	2.7212	2.6523	2.6817	2.6172	2.5873	2.6513
3	2.8427	2.7110	2.7141	2.6972	2.6452	2.7128
4	2.8865	2.7471	2.7333	2.7409	2.6816	2.7483
5	2.9208	2.7683	2.7446	2.7717	2.7066	2.7693
10	2.9842	2.8096	2.7680	2.8290	2.7598	2.8099
30	3.0229	2.8316	2.7833	2.8685	2.7965	2.8318
$\nu \sim exp(0.001)$	3.0443	2.8409	2.7876	2.8823	2.8091	2.8422
$\nu \sim gamma(2, 0.1)$	2.9768	2.8124	2.7709	2.8314	2.7657	2.8112

**Tabla A-21.:** Error de estimación global para todas la situación  $S_{III}$  con  $J = 7$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	2.7774	2.7049	2.7499	2.6853	2.6537	2.7079
2	3.0108	2.8377	2.8322	2.8305	2.7616	2.8388
3	3.1451	2.9085	2.8768	2.9204	2.8304	2.9094
4	3.1955	2.9522	2.9028	2.9715	2.8783	2.9515
5	3.2358	2.9777	2.9180	3.0048	2.9082	2.9780
10	3.3031	3.0311	2.9518	3.0793	2.9769	3.0288
30	3.3525	3.0593	2.9734	3.1283	3.0267	3.0592
$\nu \sim exp(0.001)$	3.3497	3.0702	2.9797	3.1442	3.0393	3.0691
$\nu \sim gamma(2, 0.1)$	3.2849	3.0313	2.9566	3.0829	2.9851	3.0332

**Tabla A-22.:** Error de estimación global para todas la situación  $S_{III}$  con  $J = 8$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	2.3465	2.3814	2.5242	2.3352	2.3771	2.3818
2	2.6964	2.6196	2.7013	2.5766	2.5661	2.6220
3	2.8693	2.7642	2.8055	2.7360	2.6928	2.7657
4	3.0129	2.8524	2.8701	2.8357	2.7839	2.8515
5	3.0810	2.9140	2.9111	2.9065	2.8443	2.9115
10	3.2406	3.0384	3.0080	3.0608	2.9881	3.0388
30	3.3184	3.1212	3.0730	3.1623	3.0909	3.1238
$\nu \sim exp(0.001)$	3.3708	3.1529	3.1004	3.1992	3.1312	3.1538
$\nu \sim gamma(2, 0.1)$	3.2140	3.0439	3.0235	3.0587	2.9991	3.0460

**Tabla A-23.:** Error de estimación global para todas la situación  $S_{III}$  con  $J = 9$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	2.0600	2.1978	2.4147	2.1378	2.2328	2.1990
2	2.3796	2.4602	2.6179	2.3813	2.4203	2.4611
3	2.6007	2.6241	2.7397	2.5498	2.5579	2.6259
4	2.7322	2.7292	2.8170	2.6658	2.6530	2.7305
5	2.8299	2.7308	2.8174	2.6649	2.6532	2.8031
10	3.0169	2.9540	2.9876	2.9216	2.8879	2.9549
30	3.1512	3.0558	3.0696	3.0479	3.0077	3.0576
$\nu \sim exp(0.001)$	3.1780	3.0961	3.1039	3.0902	3.0532	3.0954
$\nu \sim gamma(2, 0.1)$	2.9885	2.9560	3.0049	2.9116	2.8914	2.9569

Tabla A-24.: Error de estimación global para todas la situación  $S_{III}$  con  $J = 10$  grupos

$\nu$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$U(0, 100)$
1	2.5227	2.5223	2.6282	2.4934	2.5153	2.5211
2	2.8793	2.7790	2.8406	2.7481	2.7239	2.7786
3	3.1173	2.9516	2.9702	2.9287	2.8761	2.9508
4	3.2488	3.0621	3.0535	3.0547	2.9844	3.0580
5	3.3669	3.1338	3.1120	3.1416	3.0607	3.1352
10	3.5379	3.2959	3.2398	3.3321	3.2463	3.2932
30	3.6708	3.4022	3.3298	3.4658	3.3769	3.3999
$\nu \sim exp(0,001)$	3.7063	3.4409	3.3642	3.5086	3.4224	3.4429
$\nu \sim gamma(2, 0.1)$	3.4846	3.2965	3.2572	3.3218	3.2516	3.2941

#### A.4. Error de estimación individual para el escenario con algunos valores atípicos moderados

Tabla A-25.: Error de estimación individual para todas la situación  $S_{II}$  con  $J = 6$  grupos

$\nu$	$\alpha_j$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$
1	$\alpha_5$	2.9722	2.0745	1.4885	2.3284	1.9124
	$\alpha_6$	3.2690	2.3534	1.7619	2.6114	2.1914
2	$\alpha_5$	2.9192	1.9260	1.4063	2.2236	1.8434
	$\alpha_6$	3.1921	2.1931	1.6642	2.5022	2.1173
3	$\alpha_5$	2.8439	1.8624	1.3655	2.1750	1.8131
	$\alpha_6$	3.1097	2.1294	1.6241	2.4461	2.0801
4	$\alpha_5$	2.7994	1.8249	1.3472	2.1588	1.7996
	$\alpha_6$	3.0747	2.0878	1.5987	2.4297	2.0680
5	$\alpha_5$	2.7679	1.8121	1.3272	2.1509	1.7900
	$\alpha_6$	3.0563	2.0682	1.5807	2.4230	2.0556
10	$\alpha_5$	2.7810	1.7659	1.3002	2.1383	1.7795
	$\alpha_6$	3.0513	2.0249	1.5484	2.4040	2.0447
30	$\alpha_5$	1.7446	1.2814	2.7287	2.1299	1.7747
	$\alpha_6$	1.9991	1.5316	2.9808	2.3959	2.0372
Normal	$\alpha_5$	<b>2.3916</b>	<b>1.5632</b>	<b>1.2267</b>	<b>1.8852</b>	<b>1.6382</b>
	$\alpha_6$	<b>2.7024</b>	<b>1.7900</b>	<b>1.4080</b>	<b>2.1606</b>	<b>1.8993</b>

**Tabla A-26.:** Error de estimación individual para todas la situación  $S_{II}$  con  $J = 7$  grupos

$\nu$	$\alpha_j$	$IG(0,001, 0,001)$	$SBeta2(1/2, 1/2, 25^2)$	$SBeta2(1, 1, 25^2)$	$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$SBeta2(1, 1, \sigma_{\alpha_j}^2)$
1	$\alpha_6$	3.5142	2.6266	1.9144	2.8776	2.3718
	$\alpha_7$	4.0573	3.0662	2.2779	3.3387	2.7788
2	$\alpha_7$	3.4427	2.4517	1.8055	2.7658	2.3073
	$\alpha_6$	4.0404	2.8821	2.1625	3.2399	2.7298
3	$\alpha_6$	3.3659	2.3775	1.7557	2.7201	2.2850
	$\alpha_7$	3.9520	2.8095	2.1091	3.2046	2.7136
4	$\alpha_6$	3.3385	2.3399	1.7264	2.6970	2.2712
	$\alpha_7$	3.9328	2.7798	2.0807	3.1925	2.7048
5	$\alpha_6$	3.3261	2.3230	1.7116	2.6910	2.2618
	$\alpha_7$	3.9227	2.7586	2.0600	3.1900	2.7031
10	$\alpha_6$	3.3096	2.2753	1.6815	2.6818	2.2640
	$\alpha_7$	3.9192	2.7082	2.0227	3.1941	2.7179
30	$\alpha_6$	3.2978	2.2614	1.6517	2.6715	2.2598
	$\alpha_7$	3.9055	2.6926	1.9961	3.1923	2.7283
Normal	$\alpha_6$	<b>2.6794</b>	<b>1.8598</b>	<b>1.4855</b>	<b>1.1902</b>	<b>1.9608</b>
	$\alpha_7$	<b>2.5897</b>	<b>1.6631</b>	<b>1.2576</b>	<b>2.0209</b>	<b>1.7535</b>

## A.5. Error de estimación individual para todos los escenario cuando $\nu = 5$ y se tienen $J = 8$ grupos

**Tabla A-27.:** Error de estimación con  $J = 8$  para todos los escenarios y  $\nu = 5$

Distribución	Escenario	$j$							
		1	2	3	4	5	6	7	8
$IG(0.001, 0.001)$	$S_I$	0.1487	0.1021	0.1697	0.2322	0.2626	0.1479	0.2087	0.3345
	$S_{II}$	0.4012	0.3384	0.3816	0.4904	0.4052	2.4814	3.3432	3.1617
	$S_{III}$	0.3792	0.4539	0.6300	0.5794	0.4834	0.4953	0.5293	5.9422
$SBeta2(1/2, 1/2, 25^2)$	$S_I$	0.2174	0.1874	0.2385	0.2817	0.3080	0.2345	0.2819	0.3681
	$S_{II}$	0.4814	0.4682	0.5496	0.6682	0.5426	1.8772	2.3598	2.3076
	$S_{III}$	0.4880	0.5631	0.7774	0.6954	0.5426	1.8772	2.3598	2.3076
$SBeta2(1, 1, 25^2)$	$S_I$	0.3390	0.3317	0.4016	0.4301	0.4300	0.3924	0.4409	0.4823
	$S_{II}$	0.5721	0.5910	0.7421	0.8558	0.6766	1.5523	1.8442	1.8332
	$S_{III}$	0.5832	0.6705	0.9420	0.8282	0.7487	0.7344	0.8639	3.1034
$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$S_I$	0.1347	0.0887	0.1587	0.2197	0.2464	0.1337	0.1905	0.3145
	$S_{II}$	0.4212	0.3860	0.4321	0.5443	0.4572	2.0806	2.6907	2.6114
	$S_{III}$	0.4236	0.4882	0.6606	0.6000	0.5349	0.5269	0.5828	4.6306
$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$S_I$	0.1324	0.0879	0.1537	0.2142	0.2399	0.1330	0.1864	0.2997
	$S_{II}$	0.4430	0.4276	0.4860	0.5849	0.5012	1.8477	2.3090	2.2861
	$S_{III}$	0.4558	0.5189	0.6967	0.6250	0.5739	0.5565	0.6265	4.0367
$U(0, 100)$	$S_I$	0.2191	0.1855	0.2417	0.2805	0.3088	0.2341	0.2799	0.3675
	$S_{II}$	0.4841	0.4701	0.5535	0.6706	0.5445	1.8692	2.3510	2.2977
	$S_{III}$	0.4884	0.5629	0.7827	0.6906	0.6200	0.6138	0.6998	4.0184

**Tabla A-28.:** Error de estimación con  $J = 8$  para todos los escenarios y  $\nu \sim \text{Exponencial}(0,001)$

Distribución	Escenario	$j$							
		1	2	3	4	5	6	7	8
$IG(0.001, 0.001)$	$S_I$	0.1526	0.1057	0.1723	0.2322	0.2600	0.1503	0.2052	0.3309
	$S_{II}$	0.4375	0.3920	0.4385	0.5578	0.4616	2.3641	3.2721	3.0395
	$S_{III}$	0.4900	0.5668	0.7665	0.7013	0.6167	0.6142	0.6775	6.9289
$SBeta2(1/2, 1/2, 25^2)$	$S_I$	0.2246	0.1962	0.2419	0.2857	0.3101	0.2383	0.2824	0.3648
	$S_{II}$	0.5171	0.5209	0.6120	0.7231	0.5940	1.7788	2.2438	2.1901
	$S_{III}$	0.5994	0.6710	0.9132	0.8115	0.5940	1.7788	2.2438	2.1901
$SBeta2(1, 1, 25^2)$	$S_I$	0.3526	0.3469	0.4093	0.4370	0.4315	0.4006	0.4475	0.4858
	$S_{II}$	0.6050	0.6388	0.8041	0.9105	0.7317	1.4739	1.7333	1.7401
	$S_{III}$	0.6846	0.7712	1.0634	0.9339	0.8670	0.8423	0.9911	3.4593
$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$S_I$	0.1326	0.0869	0.1553	0.2182	0.2449	0.1298	0.1875	0.3104
	$S_{II}$	0.4463	0.4302	0.4799	0.5847	0.5022	2.0058	2.6345	2.5327
	$S_{III}$	0.5293	0.5901	0.7834	0.7086	0.6577	0.6374	0.7182	5.6103
$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$S_I$	0.1267	0.0828	0.1518	0.2137	0.2373	0.1252	0.1814	0.2960
	$S_{II}$	0.4649	0.4666	0.5248	0.6187	0.5393	1.7871	2.2699	2.2328
	$S_{III}$	0.5551	0.6125	0.8020	0.7202	0.6867	0.6568	0.7467	5.0247
$U(0, 100)$	$S_I$	0.2271	0.1968	0.2440	0.2903	0.3094	0.2385	0.2840	0.3641
	$S_{II}$	0.5163	0.5212	0.6169	0.7306	0.6015	1.7737	2.2365	2.1857
	$S_{III}$	0.5996	0.6746	0.9124	0.8134	0.7510	0.7321	0.8465	4.6166

## A.6. Error de estimación individual en $S_I, S_{II}, S_{III}$ para $J = 10$ grupos

**Tabla A-29.:** Error de estimación individual para  $J = 10$  en todos los escenarios con  $\nu = 1$

Distribución a priori $\sigma_\alpha^2$	Escenario	$j$									
		1	2	3	4	5	6	7	8	9	10
$IG(0.001, 0.001)$	$S_I$	0.1321	0.0743	0.1503	0.2140	0.2620	0.1556	0.2062	0.3216	0.0944	0.1068
	$S_{II}$	0.2189	0.1840	0.2359	0.8494	0.2780	0.1941	0.3453	3.2825	3.0435	3.9396
	$S_{III}$	0.1965	0.1521	0.2679	0.3076	0.2196	0.1746	0.1370	0.1652	0.3317	4.4117
$SBeta2(1/2, 1/2, 25^2)$	$S_I$	0.1530	0.1325	0.2222	0.6239	0.2609	0.1810	0.2254	0.3947	0.1503	0.1748
	$S_{II}$	0.2864	0.2886	0.3471	0.3829	0.3496	0.2923	0.4417	2.6229	2.3418	3.0860
	$S_{III}$	0.2729	0.2340	0.3903	0.4080	0.2992	0.2748	0.2552	0.2881	0.3709	3.5687
$SBeta2(1, 1, 25^2)$	$S_I$	0.2441	0.2393	0.3278	0.6264	0.3505	0.2948	0.3392	0.4569	0.2692	0.2922
	$S_{II}$	0.3622	0.4191	0.5284	0.5513	0.4740	0.4324	0.5543	2.1066	1.7932	2.3942
	$S_{III}$	0.3685	0.3434	0.5594	0.5453	0.4099	0.4121	0.4203	0.4476	0.4450	2.9559
$SBeta2(1/2, 1/2, \sigma_{\alpha_j}^2)$	$S_I$	0.1196	0.0918	0.1861	0.6404	0.2284	0.1410	0.1820	0.3658	0.1010	0.1314
	$S_{II}$	0.2590	0.2436	0.2901	0.3284	0.3107	0.2468	0.4058	2.7854	2.5024	3.3010
	$S_{III}$	0.2482	0.2090	0.3505	0.3743	0.2758	0.2426	0.2172	0.2527	0.3523	3.6941
$SBeta2(1, 1, \sigma_{\alpha_j}^2)$	$S_I$	0.1319	0.1065	0.1954	0.6150	0.2376	0.1586	0.1982	0.3662	0.1222	0.1467
	$S_{II}$	0.2905	0.3069	0.3664	0.4013	0.3677	0.3092	0.4476	2.4147	2.0803	2.8013
	$S_{III}$	0.3043	0.2690	0.4369	0.4411	0.3341	0.3139	0.3063	0.3369	0.3859	3.1985
$U(0, 100)$	$S_I$	0.1661	0.1196	0.1930	0.2427	0.2908	0.1998	0.2473	0.3494	0.1489	0.1558
	$S_{II}$	0.2820	0.2937	0.3631	0.8433	0.3662	0.3022	0.4297	2.5959	2.3236	3.0349
	$S_{III}$	0.2728	0.2353	0.3890	0.4064	0.2983	0.2740	0.2540	0.2880	0.3715	3.5667

# Bibliografía

- Barrera, C., y Correa, J. (2008). Distribución predictiva bayesiana para modelos de pruebas de vida vía mcmc. *Revista Colombiana de Estadística*, 31, 145-155.
- Berger, J. (2000). Bayesian analysis: A look at today and thoughts of tomorrow. *Journal of the American Statistical Association*, 95(452), 1269-1276.
- Berger, J. (2006). The case for objective bayesian analysis. *Bayesian Analysis*, 1(3), 385-402.
- Berger, J., y Strawderman, W. (1996). Choice of hierarchical priors: admissibility in estimation of normal means. *Ann. Statist*, 24(3), 931-951.
- Bolstad, W. (2004). *Introduction to bayesian statistics*. New York: John Wiley & Sons.
- Browne, W., y Draper, D. (2006). A comparison of bayesian and likelihood-based methods for fitting multinivel models. *Bayesian Analysis*, 1(3), 473-514.
- Christian, R., y Casella, G. (2004). *Monte carlo statistical methods*. New York: Springer.
- Christiansen, C., y Morris, C. (1997). Hierarchical poisson regression modeling. *Journal of the American Statistical Association*, 92(438), 618-632.
- Cowles, M. K., y Carlin, B. P. (1996). Markov chain monte carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association*, 91, 883-904.
- Daniels, M. (1999). A prior for the variance in hierarchical models. *The Canadian Journal of Statistics*, 27(3), 567-578.
- Demirhan, H., y Kalaylioglu, Z. (2015). Joint prior distributions for variance parameters in bayesian analysis of normal hierarchical models. *Journal of Multivariate Analysis*, 135(2), 163-174.
- Dubey, S. D. (1970). Compound gamma, beta and f distributions. *Metrika*, 16, 27-31. doi: <https://doi.org/10.1007/BF02613934>
- Efron, B., y Hastie, T. (2016). *Computer age statistical inference: Algorithms, evidence, and data science*. New York: Cambridge University Press.
- Fonseca, T. C. O., Ferreira, M. A. R., y Migon, H. S. (2008). Objective bayesian analysis for the student-t regression model. *Biometrika*, 95(2), 325-333.
- Frühwirth-Schnatter, S., y Wagner, H. (2010). Stochastic model specification search for gaussian and partial non-gaussian state space models. *Journal of Econometrics*, 154(1), 85-100.
- Frühwirth-Schnatter, S., y Wagner, H. (2012). Bayesian variable selection for random intercept modeling of gaussian and non-gaussian data. *Bayesian Statistics 9*.
- Fúquene, J., Cook, J., y Pericchi, L. (2009). A case for robust bayesian priors with applications to clinical trials. *Bayesian Analysis*, 4(4), 817-846.

- Fúquene, J., Pérez, M., y Pericchi, L. (2014). An alternative to the inverted gamma for the variances to modelling outliers and structural breaks in dynamic models. *Brazilian Journal of Probability and Statistics*, 28(2), 288-299.
- Gamerman, D., y Lopes, H. (2006). *Markov chain monte carlo* (Second Edition ed.). New York: Chapman & Hall.
- Gelman, A. (2003). *Bugs.r: functions for calling bugs from r*.
- Gelman, A. (2006a). Multilevel (hierarchical) modeling: What it can and cannot do. *Technometrics*, 48(3), 432-435.
- Gelman, A. (2006b). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1(3), 515-533.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., y Rubin, D. B. (2013). *Bayesian data analysis* (3.<sup>a</sup> ed.). Reading, Massachusetts: Chapman and Hall/CRC.
- Gelman, A., y Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
- Geman, S., y Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Journal IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.
- Geweke, J. (1993). Bayesian treatment of the independent student-t linear model. *Journal of Applied Electrochemistry*, 8(S1), S19-S40.
- Goldstein, H. (1986). Efficient statistical modeling of longitudinal data. *Annals of Human Biology*, 132(2), 129-141.
- Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Griffin,.
- Goldstein, H. (1995). *Multilevel statistical models. second edition*. (Second Edition ed.). New York: Halsted Press,.
- Gradshteyn, I., y Ryzhik, I. (1965). *Table of integrals, series, and products. academic press*.
- Gustafson, P., Hossain, S., y Macnab, Y. C. (2006). Conservative prior distributions for variance parameters in hierarchical models. *Canadian Journal of Statistics*, 34(3), 377-390. doi: 10.1002/cjs.5550340302
- Hastings, W. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57, 97-109.
- Jackman, S. (2009). *Bayesian analysis for the social sciences*. United Kingdom: Wiley Series in Probability and Statistics.
- Jacquier, E., Polson, N. G., y Rossi, P. E. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, 122, 185-212.
- Juárez, M., y Steel, M. (2010). Model-based clustering of non-gaussian panel data based on skew-t distributions. *Journal of Business & Economic Statistics*, 28, 52-66.
- Kwiatkowski, D., Phillips, P., y Schmidt, P. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54, 159-178.
- Laird, N. M., y Ware, J. H. (1982). Efficient statistical modeling of longitudinal data. *Biometrics*, 38, 963-974.

- Lambert, P. C., Sutton, A. J., Burton, P. R., Abrams, K. R., y Jones, D. R. (2005). How vague is vague? a simulation study of the impact of the use of vague prior distributions in mcmc using winbugs. *Statistics in Medicine*, 24(15), 2401-2428. doi: <https://doi.org/10.1002/sim.2112>
- Longford, N. T. (1995). *Random coefficient models. in handbook of statistical models for the social and behavioral sciences*. New York: Plenum Press.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A., y Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21, 1087-1092.
- Metropolis, N., y Ulam, S. (1949). The monte carlo method. *Journal of the American Statistical Association*, 44, 335-341.
- Natarajan, R., y McCulloch, C. E. (1998). Gibbs sampling with diffuse proper priors: A valid approach to data-driven inference? *Journal of Computational and Graphical Statistics*, 7(3), 267-277. doi: 10.1080/10618600.1998.10474776
- Ntzoufras, L. (2009). *Bayesian modeling using winbugs*. New Jersey: John Wiley & Sons.
- Pérez, M., Pericchi, L., y Ramírez, I. (2017). The scaled beta2 distribution as a robust prior for scales. *Bayesian Analysis*, 12(3), 615-637.
- Plummer, M. (2003). “jags: A program for analysis of bayesian graphical models using gibbs sampling.”. (Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003).)
- Polson, N., y Scott, J. (2012). On the half-cauchy prior for a global scale parameter. *Bayesian Analysis*, 4, 771-1052.
- Psarakis, S., y Panaretos, J. (1990). The folded t distribution. *Communications in Statistics - Theory and Methods*, 19(7), 2717-2734. doi: 10.1080/03610929008830342
- Ramírez, I. (2016). *Distribución beta 2 escalada como distribución a priori para los parámetros de escala* (Tesis de Doctorado). Universidad Nacional de Colombia, Medellín.
- Raudenbush, S. W., y Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. NewburyPark: Sage Publications.
- Spiegelhalter, D. J., Thomas, A., Best, N. G., Gilks, W. R., y Lunn, D. (1994-2003). *Bugs: Bayesian inference using gibbs sampling*. Descargado de [England. www.mrc-bsu.cam.ac.uk/bugs/](http://www.mrc-bsu.cam.ac.uk/bugs/)
- Strawderman, W. (1971). Proper bayes minimax estimators of the multivariate normal mean. *Ann. Math. Statist.*, 42(1), 385-388.
- Student. (1908). The probable error of a mean. *Biometrika*, 6(1), 1-25.
- Su, Y. S., y Yajima, M. (2015). *R2jags: Using r to run “jags”*. Descargado de <https://CRAN.R-project.org/package=R2jags> (R package version 0.5-7)
- Team, R. C. (2018). *R: A language and environment for statistical computing*. Vienna, Austria. Descargado de <https://www.R-project.org/>
- Trapletti, A., y Hornik, K. (2018). *tseries: Time series analysis and computational finance*. Descargado de <https://CRAN.R-project.org/package=tseries> (R package version

0.10-46.)

Villa, C., y Walker, S. G. (2014). Objective prior for the number of degrees of freedom of a t distribution. *Bayesian Analysis*, 9(1), 197-220.

Zellner, A. (1976). Bayesian and non-bayesian analysis of the regression model with multivariate student-t error terms. *Journal of the American Statistical Association*, 71, 400-405.