

Análisis de las estimaciones de los efectos aleatorios de un modelo jerárquico con distribuciones a priori de colas pesadas

Analysis of the Random Effects Estimator of a Hierarchical Model With Heavy Tailed Prior Distributions

Jessica María Rojas Mora¹ y Isabel Cristina Ramírez Guevara²

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 (ε, ε) (IG) con valores de ε 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, Beta2 escalada (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 los efectos aleatorios de un modelo jerárquico con un enfoque Bayesiano. 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

Hierarchical Bayesian models are used in data modeling in different areas in which hierarchical structures are reflected through random effects. Usually the Normal distribution is used to model the random effects. The Inverse-gamma(ε, ε) distribution is used as prior distribution for scale parameters with very small ε values, this selection has been criticized, some authors comment that unstable posterior distributions can be obtained, which causes not robust inference. Distributions such as *half*-Cauchy, Scaled Beta2 (SBeta2) and Uniform are considered as alternatives by many authors to model the scale parameter. In the present research work, the behavior of the random effects estimators in a hierarchical model with a Bayesian approach was examined. It was assumed random effects distribution *t*-Student and scale parameter distributions *half*-Cauchy, SBeta2 and Uniform. A simulation study was carry on to evaluate the behavior of the random effects estimators. Based on the obtained results, and under different scenarios, it was possible to examine the shrinkage of the posterior parameters of the model. We concluded that in presence of atypical values, the shrinkage is lower when the effects are modeled with a *t*-Student distribution compared with those obtained when a Normal distribution is associated to the random effects, under the same prior distribution for the scale parameter.

Keywords: Bayesian Inference, Scale parameter, Hierarchical models, *t*-Student distribution.

Recepción: 28-Abr-2021

Aceptación: 15-Ago-2021

¹Instituto Tecnológico Metropolitano. Correo electrónico: jessicarojas@itm.edu.co

²Universidad Nacional de Colombia. Correo electrónico: isqramirezgu@unal.edu.co

1 Introducción

Los modelos jerárquicos se utilizan ampliamente para modelar los datos obtenidos a partir de diseños de medidas repetidas, estudios longitudinales y experimentos aleatorios multinivel diseñados en cuadrado latino, parcelas divididas, bloques balanceados o no balanceados con efectos aleatorios. 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 [1]. 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 efecto aleatorio, y por lo tanto se consideran distribuciones a priori no informativas dejando que la verosimilitud domine la inferencia sobre las varianzas de los efectos del modelo jerárquico. De 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 $\text{Gamma}(\varepsilon, \varepsilon)$ ha sido comúnmente usada como distribución a priori para la precisión, por lo cual, la distribución Gamma-inversa $(\varepsilon, \varepsilon)$ se utiliza para modelar el parámetro asociada a varianza. Autores como [3] y [2] exponen e ilustran los principales inconvenientes cuando $\varepsilon \rightarrow 0$, puesto que se puede dar lugar a distribuciones posteriores propias pero con valores incorrectos o distribuciones posteriores impropias. [8] comenta que la Gamma-inversa no es una distribución no informativa y conduce a varianzas muy pequeñas de los efectos y encogimientos fuertes a la media general de los mismos. Por otro lado, [4] menciona que típicamente se utiliza una Gamma-inversa con varianza grande para los componentes de varianza. Sin embargo, como su varianza se aproxima a infinito, esta distribución se aproxima a una a priori plana impropia de los componentes de varianza, lo cual conduce a una distribución posterior impropia y tiende a inflar los componentes de varianza para muestras pequeñas.

Hay autores que proponen distribuciones a priori alternativas para modelar el parámetro de escala. [4] estudió la distribución a priori encogimiento uniforme para el parámetro de escala en modelos jerárquicos. Muestra cómo la distribución posterior en modelos jerárquicos utilizando esta distribución a priori proporciona una distribución posterior propia. Para modelos jerárquicos [8] recomienda inicialmente asignar una distribución a priori Uniforme no informativa para los parámetros de la desviación estándar. Cuando se desea mayor información a priori, por ejemplo, restringir la varianza para que no tome valores grandes, el autor recomienda trabajar con la half-Cauchy(A) como distribución a priori para la desviación estándar, con un valor del parámetro de escala (A) que sea alto. Muestra que esta distribución tiene un mejor comportamiento, compara con la familia Gamma-inversa, cuando la varianza se aproxima a cero.

[6] utilizan como distribución a priori una Normal centrada en cero para la raíz cuadrada de cada uno de los parámetros de varianza de un modelo de espacio-estado. Por medio de un estudio de simulación y una aplicación real, muestran que esta distribución 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. Por otro lado, [5] presentan la distribución SBeta2 como una familia flexible y manejable para modelar los parámetros de escala. Comentan que esta familia de distribuciones puede obtenerse en forma cerrada como una mezcla de distribuciones Gamma. La familia SBeta2 unifica y generaliza diferentes propuestas de distribuciones a priori para el parámetro de escala y tiene numerosas ventajas teórica y prácticas.

En la mayoría de estudios sobre modelos jerárquicos suponen como distribución para los efectos aleatorios la Normal como elección natural, sin embargo [5] comentan que seleccionar la distribución t -Student como a priori para los efectos aleatorio implicaría tener un modelo más robusto. Por lo tanto, el propósito de este estudio es examinar el comportamiento de las estimaciones en un modelo jerárquico para el cual se asumió efectos distribuidos t -Student y diferentes propuestas

de distribuciones a priori para el parámetro de escala de los efectos como la Gamma-inversa, half-Cauchy, SBeta2 y Uniforme. Para los grados de libertad de la distribución *t*-Student se consideraron dos situaciones de interés: la primera asumiendo grados de libertad discretos y la segunda dos alternativas de distribuciones de probabilidad para los grados de libertad.

Este artículo se encuentra organizado de la siguiente manera: la sección 2 se describe el modelo jerárquico Bayesiano que se analiza en este artículo. En la sección 3 se realiza una revisión sobre las distribuciones a priori del parámetro de escala. En la sección 4 se detalla el estudio de simulación llevado a cabo resaltando los principales aspectos tenidos en cuenta. Finalmente, las secciones 5 y 6 muestran los resultados de los ajustes realizados y una aplicación, respectivamente.

2 Modelos jerárquicos

Los modelos jerárquicos pueden ser usados para una variedad de metas inferenciales incluyendo inferencia causal, predicción y modelado descriptivo. La aplicación de estos modelos se puede dar en áreas tales como: educación, biología, psicología, medicina, entre otras [11, 10, 12]. Los modelos multinivel también se denominan jerárquicos por dos razones diferentes: por la estructura de los datos y por el modelo mismo, que tiene su propia jerarquía. 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 característica que distingue los modelos jerárquicos de la regresión clásica está en el modelado de la variación entre grupos [13].

2.1 Modelo jerárquico lineal

Los modelos jerárquicos lineales constituyen una clase general de modelos que permiten modelar situaciones en las cuales se tiene datos que presentan una estructura jerárquica. En [8, 9] 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 α_j , en el cual y_{ij} representa

la medición del individuo i en la j -ésima población:

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

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

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

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

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

En [13, ver sección 19.6] presentan aspectos que se deben 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 discuten el modelo básico de datos de dos niveles y_{ij} con coeficientes a nivel de grupo α_j del modelo (1). En la siguiente sección se muestran las distribuciones a priori sugeridas en la literatura Bayesiana para el parámetro de escala.

3 Familia de distribuciones a priori para parámetros de escala

Entre las distribuciones a priori más utilizadas para modelar los parámetros de escala están la Uniforme y la distribución Gamma-inversa. Esta última ha recibido algunas críticas porque valores pequeños de sus parámetros hacen que la distribución posterior para los parámetros de escala tenga valores muy pequeños y un gran encogimiento hacia la media en la estimación de los efectos del modelo. Se han propuesto como alternativas las distribuciones half-Cauchy y SBeta2. A continuación se discute con más detalle estas distribuciones.

3.1 Distribución a priori half-Cauchy

[8] sugirió la half-Cauchy como distribución a priori para la desviación estándar σ_α con parámetro de escala A . El autor recomienda un valor alto de este parámetro cuando se desea mayor información a priori, por ejemplo restringir la varianza para que no tome valores grandes. La función de densidad de probabilidad es:

$$p(\sigma_\alpha) = \frac{2}{\pi A \left(1 + \left(\frac{\sigma_\alpha}{A}\right)^2\right)}, \text{ para } \sigma_\alpha > 0, A > 0$$

Esta distribución tiene un mejor comportamiento, comparada con la familia Gamma-inversa, cuando la varianza se aproxima a cero.

3.2 Distribución Beta 2 Escalada (SBeta2)

[20] propone la distribución SBeta2 como distribución a priori para los parámetros de escala. Esta distribución es una versión escalada de la distribución Beta2, también conocida como distribución beta inversa o distribución beta prime la cual es una forma estándar de la distribución de Pearson tipo VI. La distribución Beta2 para ψ está definida como:

$$p(\psi) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \frac{\psi^{q-1}}{(1+\psi)^{(p+q)}}, \psi > 0, p > 0, q > 0. \quad (3)$$

La distribución SBeta2 para ψ se define como:

$$\text{SBeta2}(\psi|p, q, b) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)b} \frac{\left(\frac{\psi}{b}\right)^{q-1}}{\left(1 + \frac{\psi}{b}\right)^{(p+q)}} \quad (4)$$

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 derecha de la distribución de tal manera que mientras más pequeño sea más pesada es la cola y $b > 0$ controla la escala de la distribución.

4 Estudio de simulación

4.1 Diseño del estudio de simulación

Para el estudio de simulación se generan datos de acuerdo al modelo

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

donde j corresponde al número de grupos, $\mu \sim N(0, 1 + 06)$, el parámetro de varianza σ_j^2 se asume

conocido y α_j denota el efecto aleatorio en el j -ésimo grupo. Los valores de los σ_j^2 se tomaron iguales a los trabajados en [20][ver sección 4.1]. Los valores de los efectos aleatorios α_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 α_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 α_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 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 α_j se establecen en un intervalo de valores $(-0.5, 0.5)$.

Los valores de los efectos α_j por cada escenario de interés corresponden a los trabajados en [5]. Se simulan datos para $J = 3, 4, 5, 6, 7, 8, 9, 10$. Para cada valor fijo de α_j, σ_j^2 y J se generan 1000 muestras de una distribución Normal por cada uno de los escenarios en consideración. Con estas muestras se realizó el siguiente ajuste:

$$\begin{aligned} y_j &\sim N(\mu + \alpha_j, \sigma_j^2), \quad j = 1, 2, \dots, J, \\ \alpha_j &\sim t_\nu(\mu_\alpha, \sigma_\alpha^2), \\ \sigma_\alpha^2 &\sim f(\cdot), \end{aligned} \quad (5)$$

con $\mu_\alpha \sim N(0, 1 + 06)$ y $f(\cdot)$ denota la distribución a priori para el parámetro de escala σ_α^2 . Las distribuciones a priori candidatas utilizadas para modelar el parámetro de escala σ_α^2 se asocian a las diferentes propuestas realizadas por [2], [8] y [5] las cuales están mencionadas en la sección 3 y corresponden a las distribuciones Gamma-inversa(ε, ε) con $\varepsilon = 0.001$, Uniforme($0, 1/\varepsilon$) con

$\varepsilon = 0.01$. De acuerdo a [20] para una inferencia robusta cuando se utiliza como distribución a priori para los parámetros de escala la distribución SBeta2, se sugiere seleccionar q entre 0 y 1 y p entre 0.5 y 1, por lo tanto se seleccionan como valores de estos hiperparámetros $p = q = 0.5$ y $p = q = 1$. Por otro lado, [8] estudia el modelo jerárquico Normal y propone utilizar como distribución a priori para σ_α la half-Cauchy(25), lo que corresponde a modelar σ_α^2 con una distribución SBeta2(0.5,0.5,625), entonces se selecciona como valor de $b = 625$. Adicionalmente, siguiendo a [5], también se estudian los valores del hiperparámetro b de tal manera que se cumpla el enunciado de probabilidad $p(\sigma_\alpha^2 > \text{Varianza}(\alpha_j)) = 0.5$ pues cuando $p = q$ la mediana de la distribución SBeta2 es el parámetro de escala b , entonces $b = \text{Varianza}(\alpha_j)$.

Con respecto a los grados de libertad de la distribución t -Student se trataron bajo dos escenarios, inicialmente se consideran grados de libertad discretos $\nu = 1, 2, 3, 4, 5, 10, 30$ utilizados en [9] y un segundo escenario fue considerarlos como un parámetro desconocido y utilizar las distribuciones a priori propuestas por [15] y [14], en las cuales proponen modelar los grados de libertad ν en una distribución t -Student a partir de una distribución Gamma tal que $\nu \sim \text{Gamma}(2, 0.1)$ y de una distribución Exponencial siendo $\nu \sim \exp(0.001)$. Por cada escenario de interés se obtiene como resultado por cada número de grupo J en consideración 10 000 simulaciones de las cadenas a posteriori de los parámetros del modelo jerárquico lineal a partir del método MCMC, utilizando el software **jags** [18] y la librería **R2jags** [19] del software estadístico **R** [17], estableciendo un quemado inicial de 2000. Para estudiar el efecto de las diferentes distribuciones a priori candidatas para el parámetro de escala σ_α^2 , se utilizó como medida el error global de estimación, el cual está dado por la siguiente expresión:

$$G = \sqrt{\frac{\sum_{j=1}^J \sum_{k=1}^{1000} (\alpha_j - \hat{\alpha}_{kj})^2}{1000}} \quad (6)$$

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

5 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 [5], donde se modelan los efectos aleatorios utilizando como distribución a priori la Normal.

Para el escenario 1, todos los efectos con una magnitud similar, se puede observar en las Figuras 1 y 2 el comportamiento de G cuando se utilizan grados de libertad discretos en la distribución t -Student para modelar los efectos aleatorios, para cada una de las distribuciones a priori ajustadas para el hiperparámetro σ_α^2 . En la Figura 1, cuando $J = 3$, se puede observar que el comportamiento de los valores de G se pueden clasificar en valores altos y bajos, siendo los valores más altos cuando se utiliza como distribución a priori para σ_α^2 la SBeta2(1,1,625) y Uniforme(0,100), seguido de la distribución SBeta2(0.5,0.5,625). Esto se da para todos los grados de libertad.

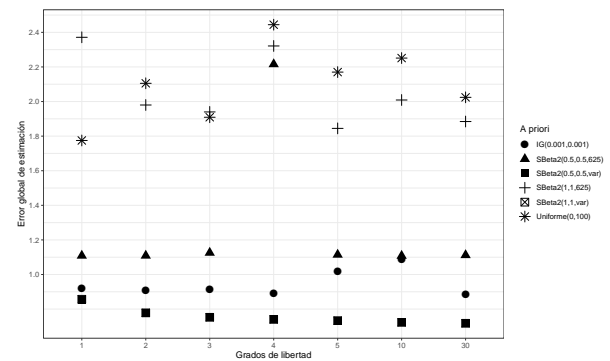


Figura 1. Error global de estimación para $J = 3$ en el escenario 1.

Por otro lado para $J = 6$, nuevamente en el escenario 1, en la Figura 2 se puede ver que el error cuadrático medio se puede clasificar en tres grupos: valores altos de G , valores promedios y valores de G bajos. En el primer grupo encontramos la distribución a priori SBeta2(1,1,625), el segundo grupo se da cuando la distribución a priori es la SBeta2(0.5,0.5,625) y Uniforme(0,100). Para el último grupo, valores bajos del error cuadrático medio, se dan con las distribuciones a priori

son: SBeta2(1,1,Varianza(α_j)), SBeta2(0.5,0.5, Varianza(α_j)) y Gamma-inversa(0.001,0.001).

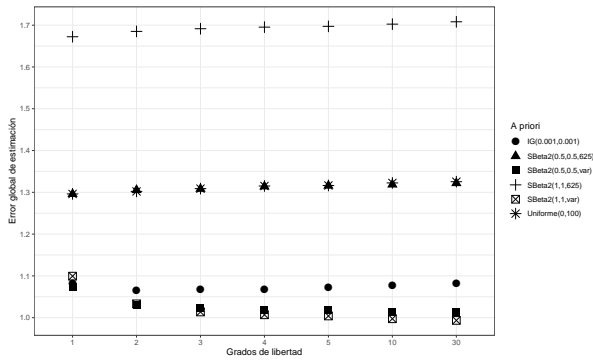


Figura 2. Error global de estimación para $J = 6$ en el escenario 1.

En general para el escenario 1, para todos los valores J y en todos los grados de libertad discretos de la distribución t -Student, los mayores valores del error cuadrático medio se dieron cuando la distribución a priori para σ_α^2 es SBeta2(1,1,625) y SBeta2(0.5,0.5,625), lo cual es razonable dado que se está asignando una distribución a priori con un parámetro de escala muy grande ($b = 625$) en un escenario donde los efectos son muy similares en magnitud y por lo tanto su varianza es pequeña.

Nuevamente en el escenario 1, cuando los grados de libertad de la t -Student son continuos, para todo J , se observa el mismo comportamiento; mayores valores de G cuando la distribución a priori es SBeta2 con $b = 625$ y Uniforme(0,100). En este escenario 1 la distribución a priori Gamma-inversa(0.001,0.001) presenta valores bajos del error cuadrático medio, tanto para los grados de libertad discretos como continuos de la t -Student.

En el escenario 2, algunos valores atípicos moderados, los resultados del estudio de simulación se van a comentar de acuerdo al número de valores atípicos en los grupos. Para $J = 3, 4$ y 5 , donde se tiene un solo valor atípico moderado, se encuentra que el error cuadrático medio es mayor cuando se utiliza como distribución a priori para σ_α^2 la SBeta2(1,1,625), Uniforme(0,100) y Gamma-inversa(0.001,0.001). Por ejemplo, en la Figura 3 se puede ver el comportamiento de G , cuando $J = 5$ y los grados de libertad de la t -Student son discretos, se observa que cuando se utiliza la distribución

Uniforme(0,100) se presenta un mayor error global de estimación, seguido de la SBeta2(1,1,625) o la Gamma-inversa(0.001,0.001). También se observa que la tendencia de los valores G es creciente a medida que aumentan los grados de libertad de la distribución t -Student. Por otro lado, cuando se utilizan grados de libertad continuos, se puede ver en la Figura 4 que la distribución que presenta mayores valores de G es la Uniforme(0,100), seguido de las distribuciones SBeta2(1,1,625) y Gamma-inversa(0.001,0.001). Esto sucede para $J = 3, 4, 5$ y coincide con el comportamiento de G cuando los grados de libertad son discretos.

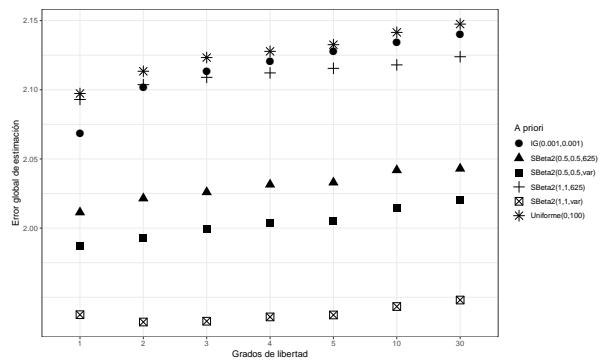


Figura 3. Error global de estimación para $J = 5$ en el escenario 2.

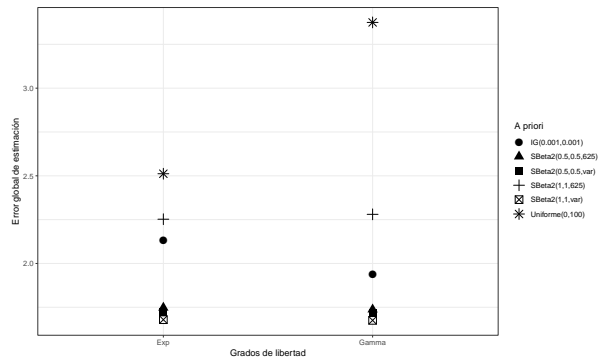


Figura 4. Error global de estimación para $J = 5$ en el escenario 2.

Cuando $J = 6, 7$ y se tienen dos valores atípicos moderados en los efectos, el mayor error cuadrático medio se tiene cuando la distribución a priori para σ_α^2 es Gamma-inversa(0.001,0.001), ver Figura 5. Esto también sucede cuando se modelan los grados de libertad continuos, el mayor error global de estimación se da cuando la distribución a priori para σ_α^2 es Gamma-inversa(0.001,0.001).

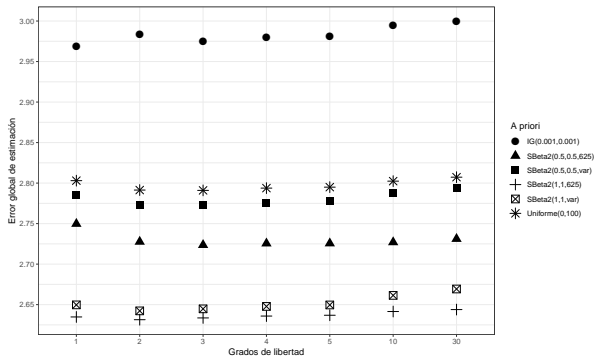


Figura 5. Error global de estimación para $J = 5$ en el escenario 2.

Para $J = 8, 9, 10$, escenario 2, donde se tienen tres valores atípicos moderados, tanto para grados de libertad discretos como continuos, los mayores valores de G se obtienen cuando la distribución a priori es SBeta2(0.5,0.5,Varianza(α_j)) y Gamma-inversa(0.001,0.001).

La Figura 6 ilustra el comportamiento de G para el tercer escenario, un efecto con un gran valor atípico, con $j = 3$ y grados de libertad discretos. Se evidencia 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 SBeta2(1/2, 1/2, 625), SBeta2(1, 1, Varianza(α_j)) y SBeta2(1/2, 1/2, Varianza(α_j)).

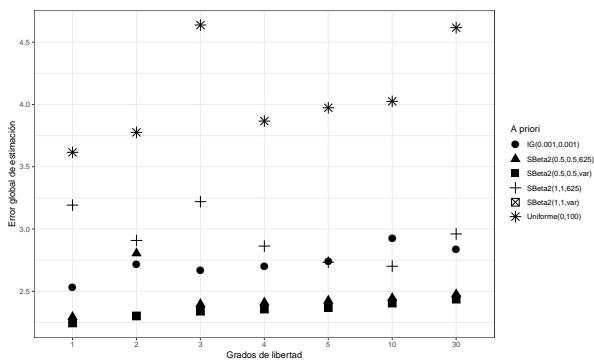


Figura 6. Error global de estimación para $J = 3$ en el escenario 3.

En el caso de $J = 3$ grupos y grados de libertad continuos para la distribución t de Student a partir de la propuesta de [15], 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 7.

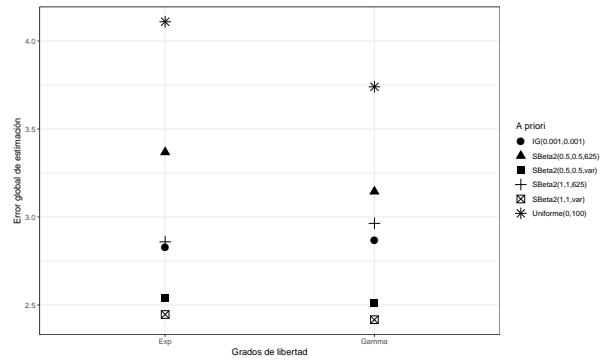


Figura 7. Error global de estimación para $J = 3$ en el escenario 3 con grados de libertad continuos.

Para $J \geq 4$, tanto grados de libertad discretos como continuos, los mayores valores de G se dan cuando $\sigma_\alpha^2 \sim \text{Gamma-inversa}(0.001, 0.001)$, ver Figuras 8 y 9.

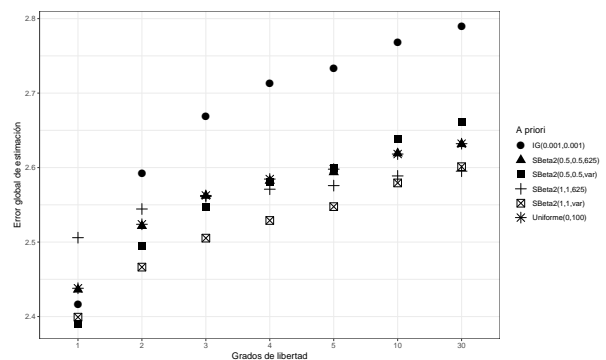


Figura 8. Error global de estimación para $J = 5$ en el escenario 3 con grados de libertad discretos.

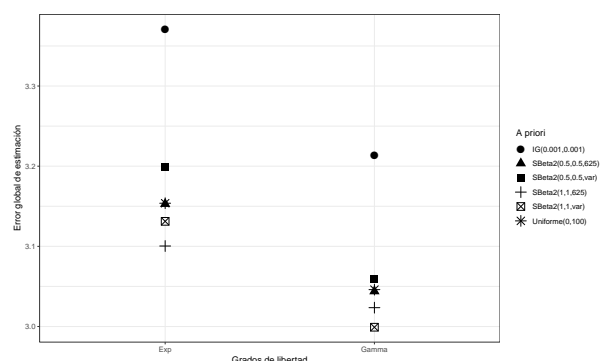


Figura 9. Error global de estimación para $J = 8$ en el escenario 3 con grados de libertad continuos.

Para evaluar el comportamiento conjunto de modelar los efectos los efectos aleatorios con una distribución t de Student teniendo diferentes distribuciones a priori para el parámetro de escala, se

comparan los resultados del error de la estimación de los efectos del modelo obtenidos vía simulación con los reportados en [5], donde los efectos aleatorios fueron modelados mediante una distribución Normal. Se realiza la comparación para $\nu=1, 2$ y 3 , valores que hace que la distribución t de Student tenga colas más pesadas, y los grados continuos. Algunos de los valores G de interés para realizar la comparación se presentan en la Tabla 1. Se observa, para el escenario 1, cuando $J=3$, los menores valores de G se dan cuando se modelan los efectos aleatorios mediante una Normal, excepto cuando la distribución a priori para σ_α^2 es $SBeta2(0.5,0.5,625)$. Cuando $J=4,5$ y 6 , los menores valores de G se dan cuando se ajusta σ_α^2 con las distribuciones $SBeta2(0.5,0.5,625)$ y $SBeta2(1,1,625)$ y los efectos aleatorios son modelos con una distribución t de Student. En el caso de $J=7, 8, 9, 10$ se obtienen valores de G menores a los reportados bajo el ajuste de los efectos con una distribución Normal cuando la distribución a priori de σ_α^2 es $Gamma\text{-inversa}(0.001,0.001)$, $SBeta2(0.5,0.5,625)$ y $SBeta2(1,1,625)$.

Los resultados de la comparación del escenario 2 con el estudio [5] se observan en la Tabla 2. Para $J=3, 4$ en los cuales hay un solo valor atípico moderado, los mayores valores G se dan cuando la distribución a priori para los efectos aleatorios siguen una distribución Normal, excepto cuando se realiza el ajuste $\sigma_\alpha^2 \sim SBeta2(1,1,625)$ y cuando $\sigma_\alpha^2 \sim Gamma\text{-inversa}(0.001,0.001)$ con grados de libertad continuos para la distribución t -Student. Cuando $J=5$ todos los valores G son menores comparados con los reportados por [5]. Para $J \geq 6$ y 7 el ajuste de los efectos aleatorios con una distribución Normal presenta menores valores G .

Para el escenario 3, un gran valor atípico, los resultados de la comparación se aprecian en la Tabla 3. Se observa que para todos los grupos los valores G son menores cuando los efectos aleatorios siguen una distribución t -Student. La distribución a priori para σ_α^2 con los valores más altos es $Gamma\text{-inversa}(0.001,0.001)$. Así, podemos establecer que bajo las condiciones del escenario 3 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 los escenarios 1 y 2 se

presentan valores más altos para G y una tendencia a disminuir a medida que aumentaban los grados de libertad.

6 Aplicación de un modelo jerárquico Normal

Consideramos el modelo jerárquico lineal dado en (2) el cual se aplicó a los datos del ejemplo de los colegios analizado en [8] en el nivel 1 y distribuciones a priori para σ_α^2 en el nivel 2 las listadas en la sección 2.2. Los datos corresponden al efecto relativo de un programa de entrenamiento sobre una prueba de actitud escolar en ocho colegios, en el modelo los parámetros α_j para $j=1, 2, \dots, 8$ representan el efecto relativo del programa de entrenamiento en el j -ésimo colegio y σ_α 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, σ_α se define como límite superior de 100.

De la Figura 10 a la Figura 13 se presentan los histogramas de frecuencias para la distribución posterior del parámetro σ_α ajustado por cuatro diferentes alternativas para la distribuciones a priori a partir de métodos Monte Carlo por cadenas de Markov basado en 10 000 simulaciones del modelo ajustado en **jags** bajo el paquete **R2jags** con un quemado inicial de 2000 iteraciones.

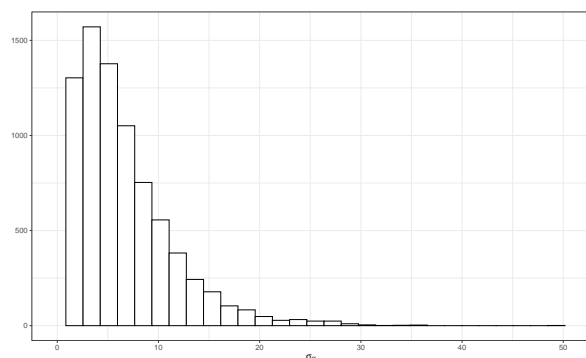


Figura 10. Histograma de la distribución posterior de σ_α con distribución a priori para $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, 25^2)$.

Tabla 1. Comparación de errores de estimación global, G , para el escenario 1.

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

Tabla 2. Comparación de errores de estimación global para el escenario 2.

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

Tabla 3. Comparación de errores de estimación global, G en el escenario 3.

J		IG (0.001, 0.001)	SBeta2 (1/2, 1/2, 25 ²)	SBeta2 (1, 1, 25 ²)	SBeta2 (1/2, 1/2, $\sigma_{\alpha_i}^2$)	SBeta2 (1, 1, $\sigma_{\alpha_i}^2$)
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

En la Figura 10 se presenta el comportamiento cuando se utiliza como distribución a priori para σ_α^2 una $SBeta2(1/2, 1/2, 25^2)$, la cual es equivalente a realizar el ajuste con una *half-Cauchy*(25) para σ_α . Se evidencia que a partir de valores de σ_α superiores a 15, se empieza a extender una pequeña cola, por lo cual, los valores para σ_α 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 11 una pequeña cola a derecha se empieza a extender a partir de valores de σ_α mayores de 20 considerados como admisibles y, comparando con los resultados del ajuste con $\sigma_\alpha^2 \sim SBeta2(1/2, 1/2, 25^2)$, se presenta una media con mayor cuantía.

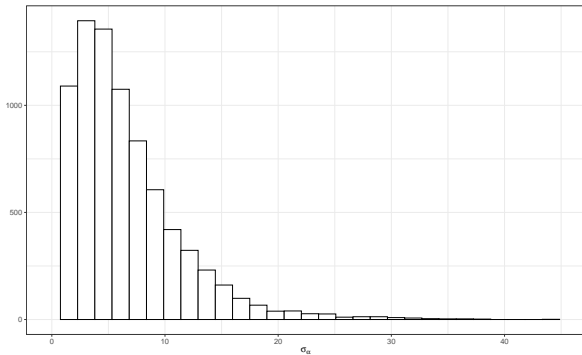


Figura 11. Histograma de la distribución posterior de σ_α con distribución a priori para $\sigma_\alpha^2 \sim SBeta2(1, 1, 25^2)$.

El ajuste con una distribución a priori Uniforme presentado en la Figura 12, la cual [8] recomienda para $J \geq 5$ grupos, tenemos que la cola a derecha se empieza a extender a partir de valores de σ_α 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 σ_α en un rango admisible.

En contraposición a lo presentado con el ajuste $\sigma_\alpha^2 \sim Uniforme(0, 100)$ tenemos el comportamiento del ajuste $\sigma_\alpha^2 \sim Gamma\text{-inversa}(0.001, 0.001)$, el cual se puede apreciar en la Figura 13, donde 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

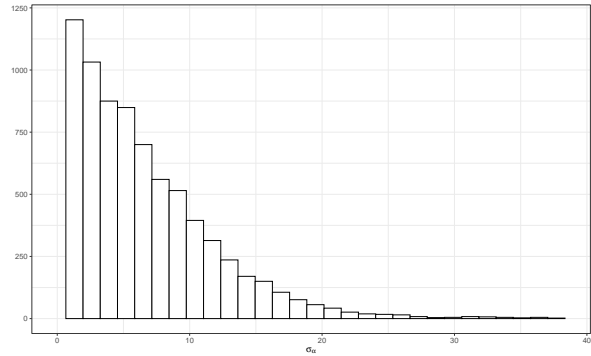


Figura 12. Histograma de la distribución posterior de σ_α con distribución a priori para $\sigma_\alpha^2 \sim Uniforme(0, 100)$.

a partir de valores de σ_α superiores a 5, por lo tanto podemos establecer que la inferencia se va a ver afectada debido al encogimiento de los efectos en cada grupo, α_j .

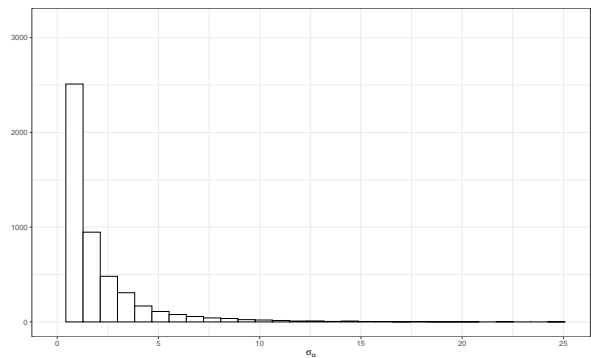


Figura 13. Histograma de la distribución posterior de σ_α con distribución a priori para $\sigma_\alpha^2 \sim IG(0.001, 0.001)$.

Por tanto, podemos establecer con el análisis gráfico de las distribuciones posteriori que, se obtienen valores de σ_α admisibles, menores que los reportados en [8], [9] y [5] ajustando los efectos aleatorios a una distribución *t-Student*.

7 Conclusiones

En el presente trabajo se propone e implementa al análisis conjunto de modelar los efectos aleatorios de una modelo jerárquico Normal a partir de una distribución *t-Student*, desde un enfoque Bayesiano, utilizando diferentes distribuciones a priori para el parámetro de escala. A partir de un estudio de simulación, se analizó el comportamiento de las estimaciones de los efectos aleatorios del modelo

jerárquico, teniendo como medida de desempeño el error global de estimación. Se utilizaron diversas propuestas 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 global de estimación a medida que aumentan los grados de libertad asociados a la distribución t -Student, para todas las distribuciones a priori para el parámetro de escala que fueron analizadas.

De acuerdo con los escenarios analizados se observó en el escenario 1 que los menores valores del error global de estimación se obtienen cuando se realizan los ajustes de σ_{α}^2 con las distribuciones SBeta2(0.5,0.5,Varianza(α_j)) y SBeta2(1,1,Varianza(α_j)). En los escenarios 2 y 3, valores atípicos moderados y un gran valor atípico, respectivamente, los valores más altos del error global de estimación se dieron cuando la distribución a priori para σ_{α}^2 es Gamma-inversa(0.001,0.001), en la mayoría de los casos.

Los resultados de la comparación del escenario 3, un gran valor atípico, con el estudio [5] muestran que para todos los grupos los valores G son menores cuando los efectos aleatorios siguen una distribución t -Student. La distribución a priori para σ_{α}^2 con los valores más altos es Gamma-inversa(0.001,0.001). Así, podemos establecer que bajo las condiciones de este escenario 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 los escenarios 1 y 2 se presentan valores más altos para G y una tendencia a disminuir a medida que aumentaban los grados de libertad.

Referencias

- [1] H. Demirhan, and Z. Kalaylioglu, “Joint prior distributions for variance parameters in Bayesian analysis of normal hierarchical models”, *Journal of Multivariate Analysis*, vol. 135, no. 2, pp. 163-174, 2015.
- [2] J. Berger, “The Case for Objective Bayesian Analysis”, *Bayesian Analysis*, vol. 1, no. 3, pp. 385-402, 2006.
- [3] Ranjini Natarajan and Charles E. McCulloch, “Gibbs Sampling with Diffuse Proper Priors: A Valid Approach to Data-Driven Inference?”, *Journal of Computational and Graphical Statistics*, vol. 7, no. 3 pp. 267-277, 1998.
- [4] M. Daniels, “A prior for the variance in hierarchical models”, *The Canadian Journal of Statistics*, vol. 27, no. 3 pp. 567-578, 1999.
- [5] M. Pérez, and L. Pericchi, and I. Ramírez, “The Scaled Beta2 distribution as a robust prior for scales”, *Bayesian Analysis*, vol. 12, no. 3 pp. 615-637, 2017.
- [6] S. Frühwirth-Schnatter, and H. Wagner, “Stochastic model specification search for gaussian and partial non-gaussian state space models”, *Journal of Econometrics*, vol. 154, no. 1 pp. 85-100, 2010.
- [7] N. Polson, and J. Scott, “On the half-cauchy prior for a global scale parameter”, *Bayesian Analysis*, vol. 4, pp. 771-1052, 2012.
- [8] A. Gelman, “Prior Distributions for Variance Parameters in Hierarchical Models” *Bayesian Analysis*, vol. 1, no. 3 pp. 515-533, 2006.
- [9] A. Gelman, and J. B. Carlin, and Hal S. Stern and David B. Dunson and Aki Vehtari and Donald B. Rubin, “Bayesian Data Analysis”, Chapman and Hall/CRC, 2013.
- [10] S. W. Raudenbush, and A. S. Bryk, “Hierarchical Linear Models: Applications and Data Analysis Methods”, Sage Publications, 2002.
- [11] H. Goldstein, “Multilevel Models in Educational and Social Research”, Griffin, 1987.
- [12] H. Goldstein, “Multilevel Statistical Models”, Halsted Press, 1995.
- [13] A. Gelman, and J. Hill, “Data Analysis Using Regression and Multilevel/Hierarchical Models”, Cambridge University Press, 2006.
- [14] J. Geweke, “Bayesian treatment of the independent student-t linear model” *Journal of Applied Electrochemistry*, vol. 8, no. 1 pp. 19-40, 1993.

- [15] M. Juárez, and M. Steel, “Model-based clustering of non-Gaussian panel data based on skew-t distributions” *Journal of Business & Economic Statistics*, vol. 28, no. 1 pp. 52-66, 2010.
- [16] S. Psarakis and J. Panaretos, “The folded t distribution” *Communications in Statistics - Theory and Methods*, vol. 19, no. 7 pp. 2717-2734, 1990.
- [17] R Core Team, “R: A Language and Environment for Statistical Computing”, R Foundation for Statistical Computing, Vienna, Austria, 2018.
- [18] Plummer, Martyn, “JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling”, 2003.
- [19] Yu Sung Su and Masanao Yajima, “2jags: Using R to Run JAGS”, R package version 0.5-7. <https://CRAN.R-project.org/package=R2jags>, 2015
- [20] I. Ramírez, “Distribución Beta 2 Escalada como Distribución a priori para los parámetros de escala”, Universidad Nacional de Colombia, Tesis de Doctorado, 2016.