



Universidad Nacional de Colombia

FACULTAD DE CIENCIAS

SEGUNDO CASO DE ESTUDIO:
PRUEBA SABER 11 2022-2: UNA PERSPECTIVA MULTINIVEL

Estadística Bayesiana

Docente:

Juan Camilo Sosa

Ander Steven Cristancho Sánchez (acristanchos@unal.edu.co)

Valeria del Pilar Lopez Bosiga (vlopezbo@unal.edu.co)

Octubre 2023

Desarrollo metodológico

Los modelos presentados en los anexos se ajustan por medio de muestreadores de Gibbs con $1000 + 10 \cdot 10000 = 101000$ iteraciones. Las primeras 1000 iteraciones del algoritmo constituyen el periodo de calentamiento del algoritmo (no se tienen en cuenta para realizar inferencia). Además, con el fin de reducir la autocorrelación de la cadena después del periodo de calentamiento, se hace un muestreo sistemático de amplitud 10, de forma que la cadena para realizar inferencias acerca de la distribución posterior de los parámetros de cada modelo consta de $B = 10000$ iteraciones.

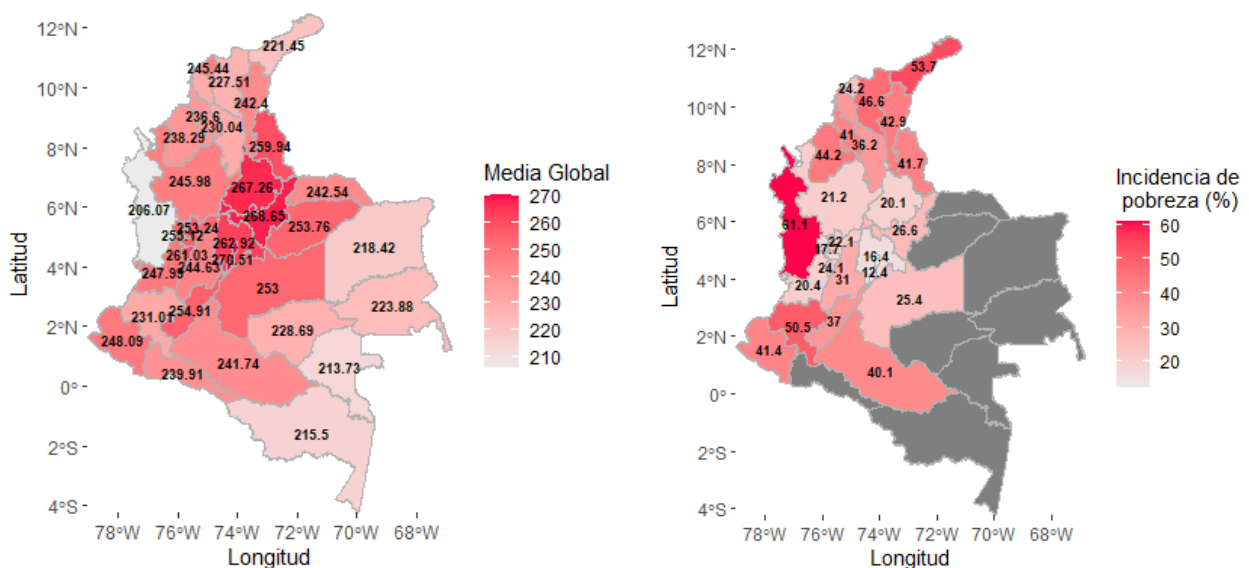
Para tal fin se emplean distribuciones previas difusas definidas por los siguientes hiperparámetros a partir de la información de la prueba:

- M_1 : $\mu_0 = 250$, $\gamma_0^2 = 50^2$, $\nu_0 = 1$, $\sigma_0^2 = 50^2$.
- M_2 : $\mu_0 = 250$, $\gamma_0^2 = 50^2$, $\eta_0 = 1$, $\tau_0^2 = 50^2$, $\nu_0 = 1$, $\sigma_0^2 = 50^2$.
- M_3 : $\mu_0 = 250$, $\gamma_0^2 = 50^2$, $\eta_0 = 1$, $\tau_0^2 = 50^2$, $\nu = 1$, $\alpha_0 = 1$, $\beta_0 = 1/50^2$.
- M_4 : $\xi_0 = 1$, $\kappa_0^2 = 50^2$, $\mu_0 = 250$, $\gamma_0^2 = 50^2$, $\eta_0 = 1$, $\tau_0^2 = 50^2$, $\nu_0 = 1$, $\sigma_0^2 = 50^2$.
- M_5 : $\xi_0 = 1$, $\kappa_0^2 = 50^2$, $\mu_0 = 250$, $\gamma_0^2 = 50^2$, $\eta_0 = 1$, $\tau_0^2 = 50^2$, $\nu = 1$, $\alpha_0 = 1$, $\beta_0 = 1/50^2$.

1. Preguntas

1. En un gráfico con dos paneles (1x2), hacer un mapa de Colombia por departamentos, donde se desplieguen los valores de la media muestral del puntaje global (panel 1, izquierda) y la incidencia de la pobreza monetaria en 2018 (panel 2, derecha). Interpretar los resultados obtenidos (máximo 100 palabras).

Nota: El archivo `pobreza monetaria.xls` disponible en la página web del curso, contiene los datos de la incidencia de la pobreza monetaria de 23 departamentos y Bogotá D.C en el periodo 2002-2018. Estos datos son de carácter público y se pueden descargar gratuitamente de la página web del DANE en este enlace.



(a) Media muestral del puntaje global

(b) Incidencia de la pobreza monetaria en 2018

Figura 1: Mapa de Colombia por departamentos

Los mapas de puntajes globales en las pruebas del ICFES y del porcentaje de la incidencia de pobreza en los departamentos de Colombia apuntan a una relación inversa. Departamentos con puntajes más elevados exhiben una menor incidencia de pobreza, mientras que aquellos con puntajes más bajos revelan una incidencia de pobreza más alta. Esto es claramente ejemplificado por departamentos como Chocó y Bogotá. Esta asociación podría sugerir que puntajes globales superiores están ligados a un mayor bienestar económico, mientras que puntajes más bajos reflejan desafíos socioeconómicos.

2. En un gráfico con dos paneles (1x2), hacer un mapa de Colombia por municipios, donde se desplieguen los valores de la media muestral del puntaje global (panel 1, izquierda) y la cobertura neta secundaria en 2022 (panel 2, derecha). Interpretar los resultados obtenidos (máximo 100 palabras).

Nota: El archivo `estadísticas educación.xls` disponible en la página web del curso, contiene estadísticas de los niveles preescolar, básica y media relacionada con indicadores sectoriales por municipio en el periodo 2011-2022. Estos datos son de carácter público y se pueden descargar gratuitamente de la página web del MEN en este enlace.

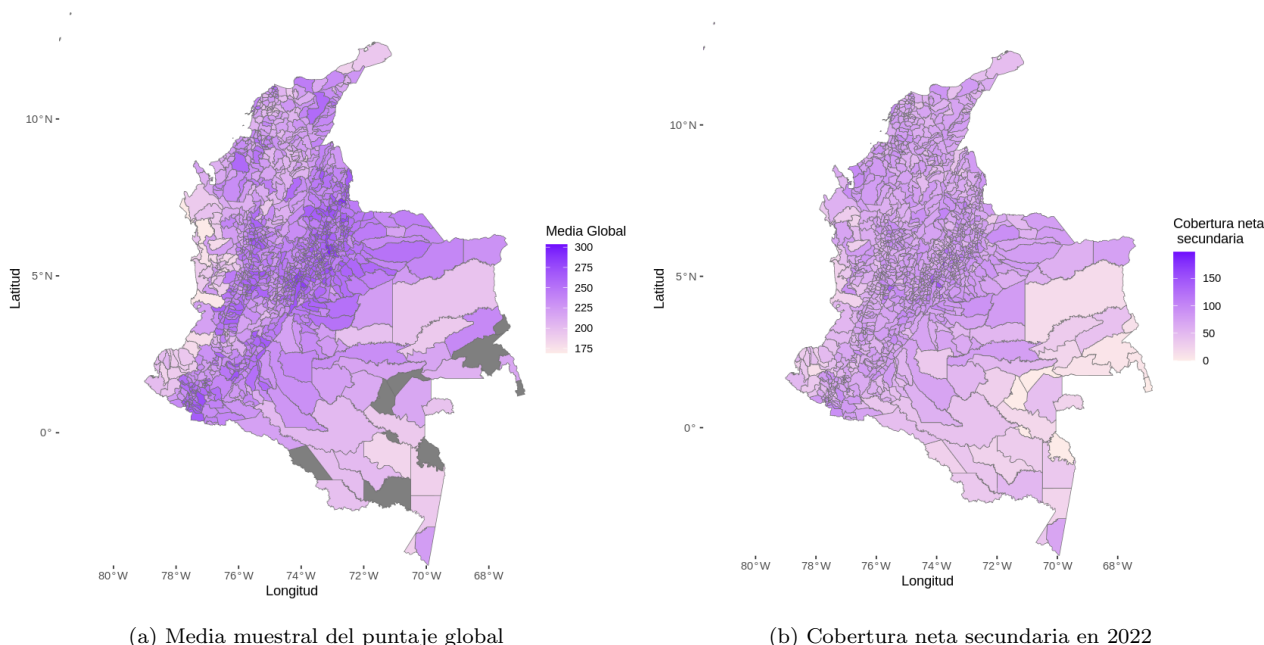


Figura 2: Mapa de Colombia por municipios

Los resultados derivados de los dos mapas de Colombia, uno representando el puntaje global del ICFES por municipios y el otro la cobertura neta en educación secundaria por municipios, apuntan a una relación directa entre estas dos variables. En otras palabras, se evidencia que aquellos municipios que exhiben puntajes más altos en las pruebas del ICFES tienden a contar con una amplia cobertura neta en educación secundaria. Esto puede indicar que un mejor rendimiento en las pruebas del ICFES está positivamente asociado con una mayor participación y acceso a la educación secundaria en estos municipios.

3. En un gráfico con cuatro paneles (2x2), hacer el DAG de M_2 (panel 1, esquina superior izquierda), M_3 (panel 2, esquina superior derecha), M_4 (panel 3, esquina inferior izquierda) y M_5 (panel 4, esquina inferior derecha).

Nota: Tomar como ejemplo la Figura 1 del artículo "Some Developments in Bayesian Hierarchical Linear Regression Modeling".

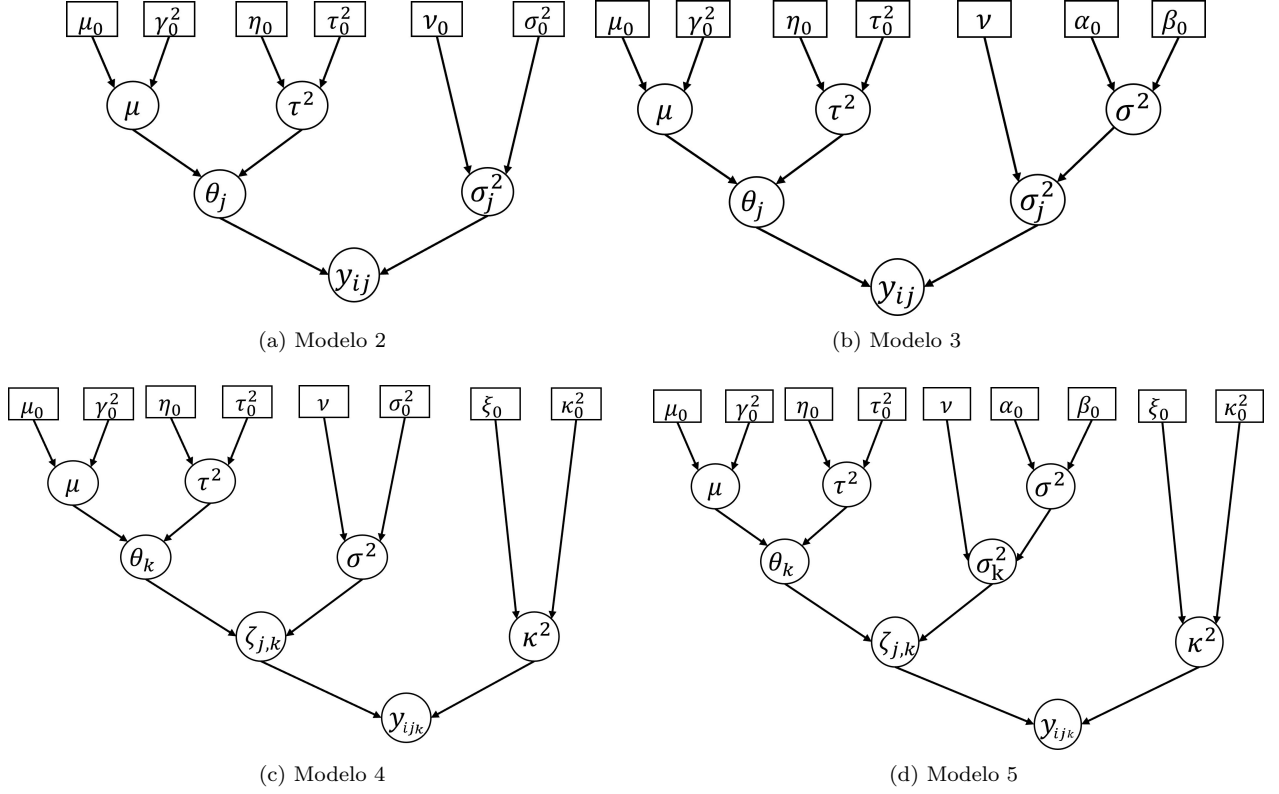


Figura 3: Representaciones de los DAG por modelo

- En un gráfico con cuatro paneles (2x2), dibujar la cadena de la log-verosimilitud de M_2 (panel 1, esquina superior izquierda), M_3 (panel 2, esquina superior derecha), M_4 (panel 3, esquina inferior izquierda) y M_5 (panel 4, esquina inferior derecha). Los gráficos deben tener la misma escala para facilitar la comparación. Interpretar los resultados obtenidos (máximo 100 palabras).

Nota: Incluir un apéndice al final del informe con las distribuciones condicionales completas (no incluir la demostración, solo cada distribución con sus respectivos parámetros) y un resumen de los coeficientes de variación de Monte Carlo de cada parámetro de cada modelo. Tomar como ejemplo la Sección 4.3 del artículo "Some Developments in Bayesian Hierarchical Linear Regression Modeling".

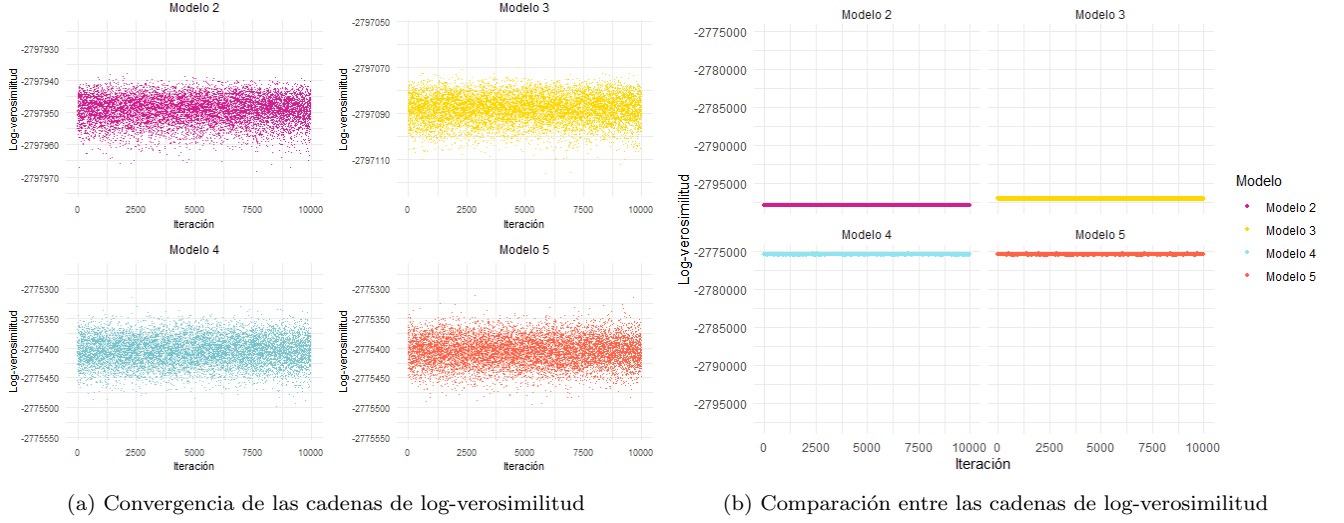


Figura 4: Cadenas de log-verosimilitud

Se puede evidenciar que las log-verosimilitudes aumentan en los modelos especificados por municipios y departamentos, es decir, en los modelos 4 y 5, siendo los que maximizan la log-verosimilitud. Las distribuciones condicionales completas para cada modelo M_k se encuentran en los anexos. Las log-verosimilitudes de los dos últimos modelos son muy parecidas, por lo que es necesario un análisis más a fondo para ver qué modelo es mejor.

5. Calcular el DIC y el WAIC de cada M_k , para $k = 1, \dots, 5$. Presentar los resultados tabularmente. Interpretar los resultados obtenidos (máximo 100 palabras).

Modelo	P _{DIC}	DIC	P _{WAIC}	WAIC
1	2.03	5636785	1.76	5636785
2	32.92	5595932	31.13	5595930
3	64.07	5594240	60.18	5594236
4	1001.78	5551814	790.44	5551603
5	988.85	5551801	780.66	5551593

Tabla 1: Comparación entre modelos por DIC y WAIC

Como se puede observar, aunque las diferencias respecto al modelo 4 y 5 sean pequeñas en términos de DIC y WAIC, estos son menores en el modelo 5, lo que significa que respecto a los otros modelos, el modelo 5 posee una mejor precisión predictiva. Es importante resaltar que el número efectivo de parámetros en el modelo 5 es menor que en el modelo 4, sin embargo, el modelo 5 tiene una mejor precisión predictiva.

6. Calcular la media posterior y el intervalo de credibilidad al 95 % basado en percentiles de μ de cada M_k , para $k = 1, \dots, 5$. Presentar los resultados tabularmente. Interpretar los resultados obtenidos (máximo 100 palabras).

Modelo	Media posterior	Q 2.5 %	Q 97.5 %
1	250.27	250.13	250.42
2	242.35	235.49	249.30
3	242.22	235.42	249.01
4	232.46	225.81	239.10
5	232.49	225.57	239.20

Tabla 2: Media posterior e intervalo de credibilidad para M_k

Se evidencia que, en los primeros modelos $\hat{\mu}$ es mayor, ya que a medida que el índice del modelo aumenta, se toma en cuenta más información, pues en los modelos 2 y 3 se consideran las medias por departamento, y en los modelos 4 y 5, se consideran además, medias por municipio. Esto también se refleja en los intervalos de credibilidad, pues el primero es un intervalo de credibilidad para un único θ , mientras los últimos son intervalos de credibilidad para la media de θ_k , es decir, la media de las medias de los 32 departamentos.

7. Usando M_5 , hacer el ranking de los departamentos basado en las medias específicas de los departamentos. Comparar los resultados con un ranking frecuentista basado en la media muestral. En un gráfico con dos paneles (1x2), hacer la visualización del ranking bayesiano (panel 1, izquierda) y el ranking frecuentista. Las visualizaciones deben incluir simultáneamente las estimaciones puntuales y los intervalos de credibilidad/confianza al 95 %. Interpretar los resultados obtenidos (máximo 100 palabras).

Nota: Usar la siguiente convención de colores: rojo oscuro para promedios significativamente inferiores a 250; negro para promedios que no difieren significativamente de 250; y verde oscuro para promedios significativamente superiores a 250.

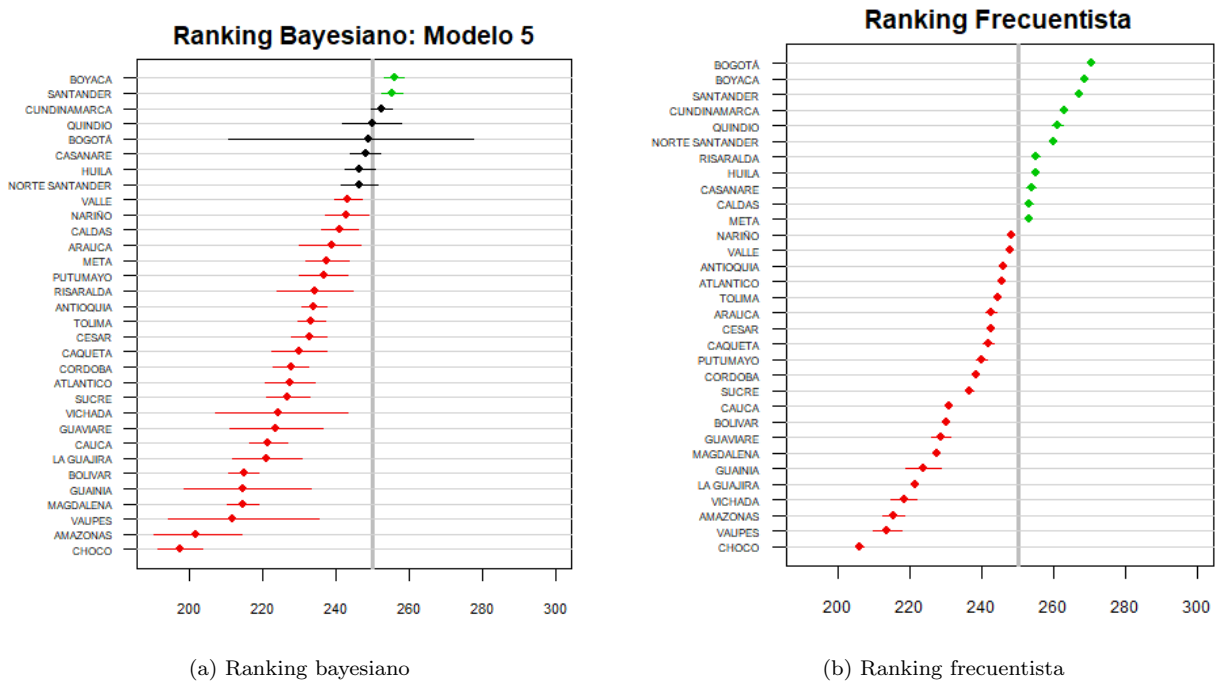


Figura 5: Ranking de los departamentos

Se destacan diferencias importantes entre los rankings, esto se debe a que el ranking bayesiano tiene en cuenta las medias y varianzas de cada municipio, mientras el frecuentista no, por lo que se ve influenciado por puntajes atípicos en los departamentos. Los intervalos de credibilidad son más amplios que los intervalos de confianza. En el ranking Bayesiano, solo dos departamentos, Boyacá y Santander, muestran promedios significativamente superiores a 250, en contraste con el ranking frecuentista, donde 11 departamentos superan este umbral, siendo Bogotá y Boyacá los líderes. En cuanto a los promedios menores a 250, en el ranking Bayesiano se registran 24 departamentos, mientras que en el enfoque frecuentista son 21. Por otro lado, en el ranking Bayesiano, 6 departamentos muestran promedios que no difieren significativamente de 250, mientras que en el enfoque frecuentista no se observa ninguno.

8. Usando M_5 , hacer una segmentación de los departamentos usando las medias específicas de los departamentos, por medio del método de agrupamiento de K -medias con cinco grupos. Presentar los resultados obtenidos visualmente a través de una matriz de incidencia organizada a partir del ranking bayesiano del numeral anterior y de un mapa que señale los departamentos que pertenecen al mismo grupo. Interpretar los resultados obtenidos (máximo 100 palabras).

Nota: Llevar a cabo la segmentación en cada iteración de la cadena de Markov asociada con M_5 , y en cada iteración, establecer los departamentos que pertenecen al mismo grupo. Tomar como ejemplo la Sección 7.5 de las notas de clase disponibles en este enlace. Para llevar a cabo la visualización del mapa se recomienda utilizar una segmentación de las medias posteriores de las medias específicas de los departamentos.

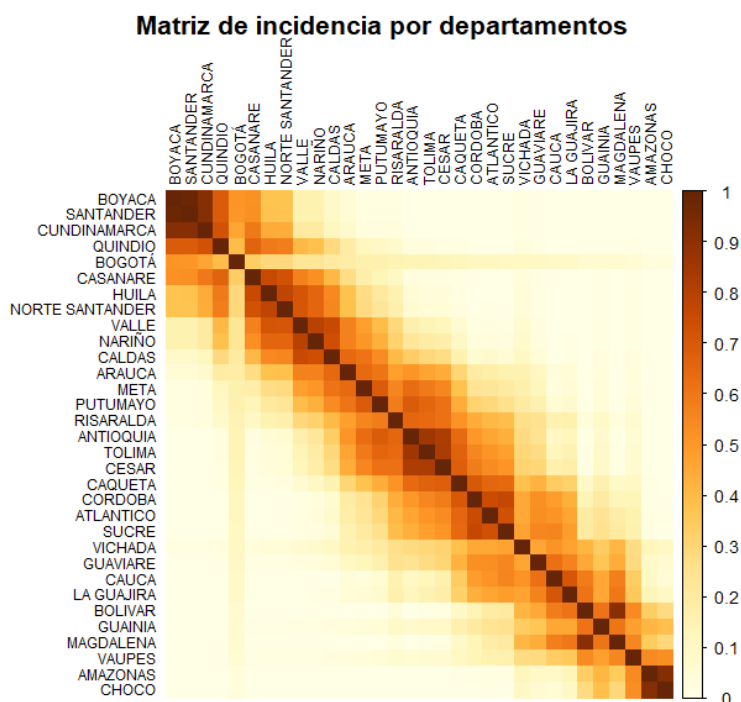


Figura 6: Matriz de incidencia por departamentos

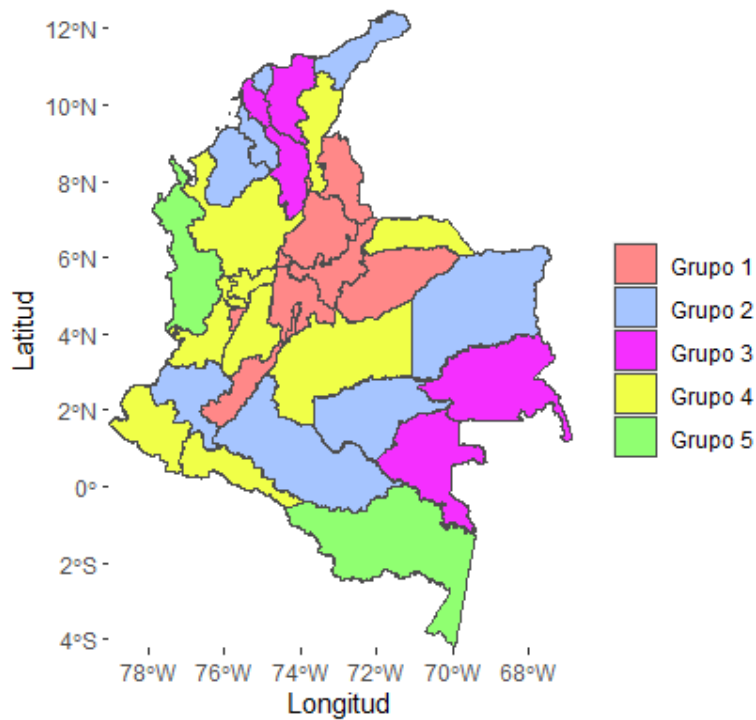


Figura 7: Mapa de departamentos que pertenecen al mismo grupo

La clasificación de departamentos mediante K-medias muestra que el Grupo 1, cuyos departamentos se ubican principalmente en la región central del país, lidera con los puntajes globales más altos, seguido en orden descendente por los grupos 4, 2 y 3. Por otro lado, el Grupo 5, formado únicamente por Amazonas y Chocó, presenta los puntajes globales más bajos. En cuanto a la matriz de incidencia se logran ver de manera difusa los patrones de los grupos.

9. Calcular la media posterior y un intervalo de credibilidad al 95 % de la incidencia de la pobreza monetaria en 2018 (IPM) para todos los departamentos que no fueron medidos por el DANE, por medio de una regresión lineal simple de la IPM frente a las medias específicas de los departamentos de M_5 . Presentar los resultados tabularmente (organizados descendente de acuerdo con la media posterior) y visualmente (por medio de un mapa usando la media posterior).

Nota: Llevar a cabo la regresión en cada iteración de la cadena de Markov asociada con M_5 , y en cada iteración, hacer las predicciones de la IPM para los departamentos a los que halla lugar. Interpretar los resultados obtenidos (máximo 100 palabras).

Departamento	Media posterior	Q 2.5 %	Q 97.5 %
Amazonas	39.96	34.42	44.64
Vaupes	37.96	32.90	43.26
Guainía	37.35	33.24	42.28
Guaviare	35.59	32.87	38.96
Vichada	35.43	31.63	39.73
Putumayo	32.93	31.44	34.42
Arauca	32.47	30.64	34.32
Casanare	30.62	28.94	32.69

Tabla 3: Incidencia de pobreza monetaria para departamentos no medidos por el DANE

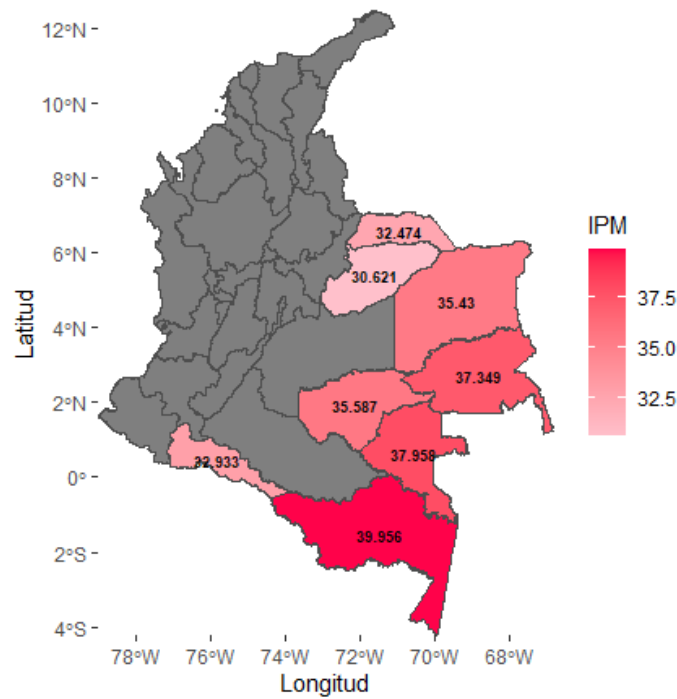


Figura 8: Incidencia de pobreza monetaria para los departamentos no medidos por el DANE

El modelo de regresión estima un IPM moderado para estos departamentos, y retomando la figura 1, parece reafirmar la asociación inversamente proporcional entre puntajes de las pruebas de estado y el IPM del departamento, pues para estos 8 departamentos, el IPM predicho es moderado y por lo mismo, sus puntajes resultan ser más bajos respecto a departamentos con un IPM más bajo.

- Usando M_5 , hacer el ranking de los municipios basado en las medias específicas de los municipios (no es preciso visualizar el ranking debido a la gran cantidad de municipios). Luego, hacer una segmentación de los municipios usando las medias específicas de los municipios, por medio del método de agrupamiento de K -medias con ocho grupos. Presentar los resultados obtenidos visualmente a través de una matriz de incidencia organizada a partir del ranking bayesiano de los municipios obtenido inicialmente y de un mapa que señale los municipios que pertenecen al mismo grupo. Interpretar los resultados obtenidos (máximo 100 palabras).

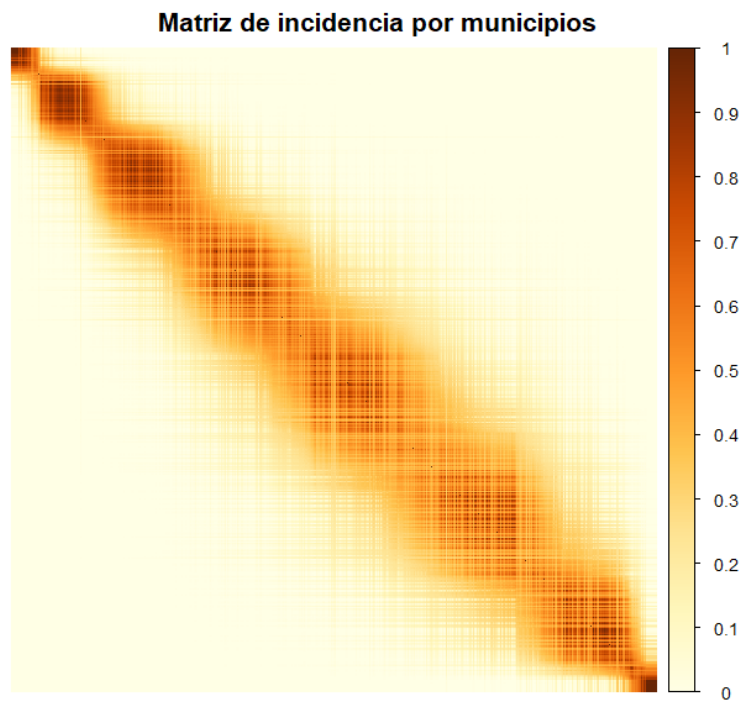


Figura 9: Matriz de incidencia por municipios

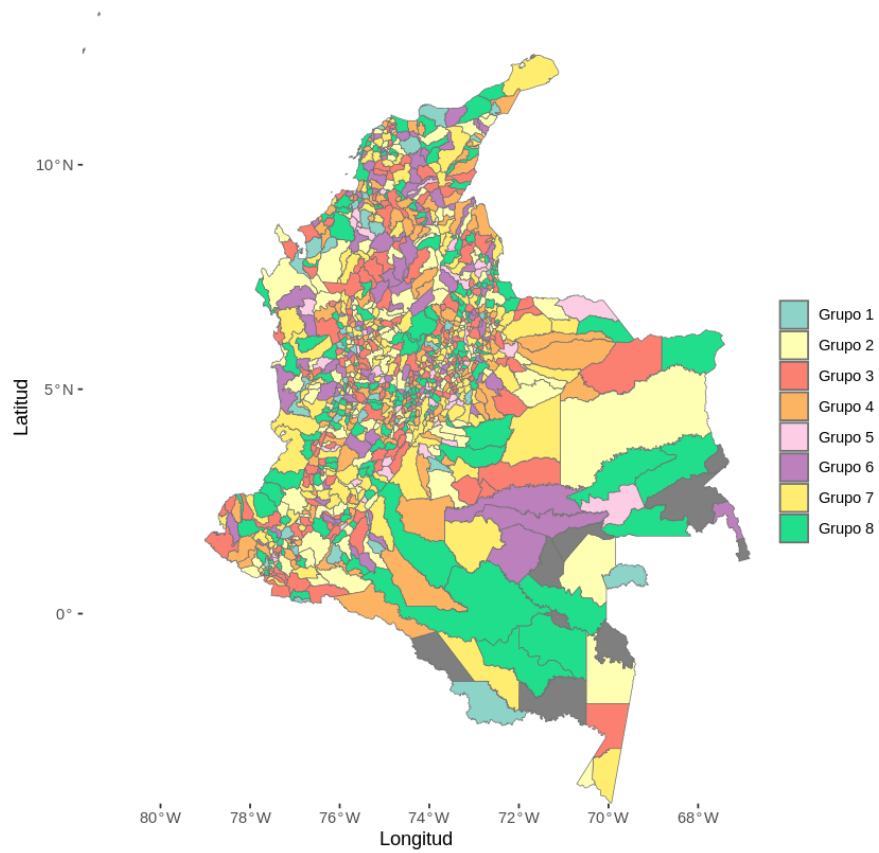


Figura 10: Mapa de municipios que pertenecen al mismo grupo

El análisis K-medias revela una clara diferencia en el desempeño global de la prueba a nivel municipal. Los municipios en el Grupo 6 son quienes lideran con los puntajes más altos, seguidos por el Grupo 8. Mientras tanto, el Grupo 2 muestra los puntajes globales más bajos. Los demás grupos ocupan posiciones intermedias en términos de puntajes globales, ilustrando una variación importante en el desempeño a nivel municipal. Los grupos se distinguen claramente en la matriz de incidencia, en donde se evidencian los 8 clusters.

11. Calcular la media posterior y un intervalo de credibilidad al 95 % de la cobertura neta secundaria en 2022 (CNS) para todos los municipios que no fueron medidos por el MEN, por medio de una regresión lineal simple de la CNS frente a las medias específicas de los municipios de M_5 . Presentar los resultados tabularmente (organizados descendente de acuerdo con la media posterior) y visualmente (por medio de un mapa usando la media posterior).

Municipio	Media posterior	Q 2.5 %	Q 97.5 %
Mapiripana	57.24	54.11	60.34
Belén de Bajirá	55.25	46.22	63.82

Tabla 4: Cobertura neta secundaria para municipios no medidos por el MEN

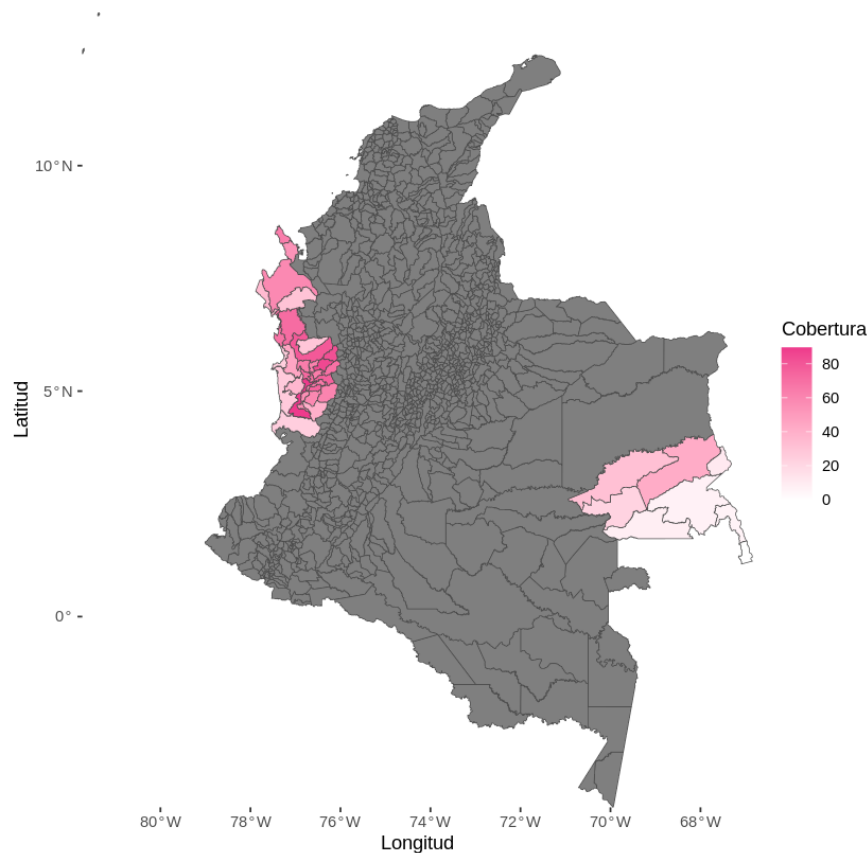


Figura 11: Cobertura neta secundaria para municipios no medidos por el MEN

Para los dos municipios no medidos, la regresión estima una cobertura neta secundaria baja, lo cual parece reafirmar lo visto en la figura 2, que hay una asociación positiva entre cobertura neta secundaria

y el puntaje global de la prueba de estado. Por lo mismo, para estos municipios que parecen tener una cobertura neta secundaria baja, los puntajes globales en las pruebas de estado no son muy altos. Los municipios no están ubicados en las coordenadas del gráfico, pero se puede notar la baja cobertura neta secundaria en las áreas adyacentes.

12. Validar la bondad ajuste de M_5 por medio de la distribución predictiva posterior en cada municipio, utilizando como estadísticos de prueba el mínimo, el máximo, el rango intercuartílico, la media, la mediana, y la desviación estándar. Presentar los resultados visualmente.

Nota: Tomar como ejemplo la Sección 8.5 del artículo "Some Developments in Bayesian Hierarchical Linear Regression Modeling."

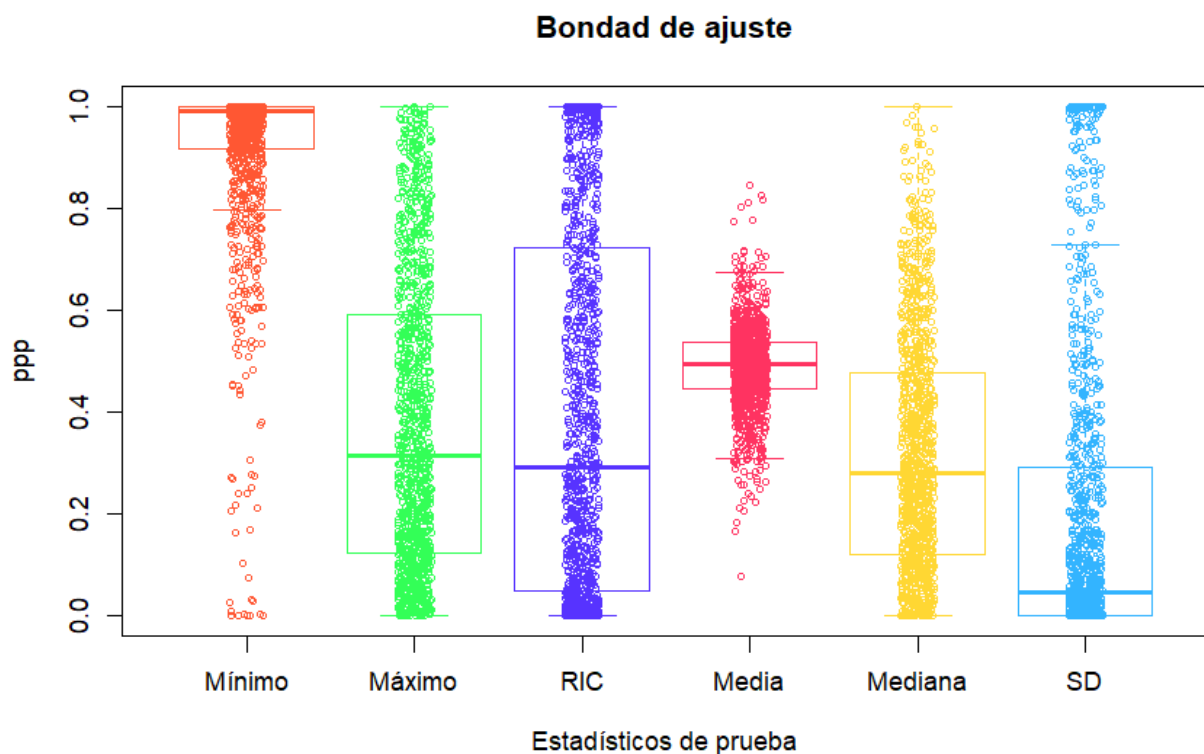


Figura 12: PPP's para los estadísticos de prueba

El modelo parece fallar al capturar el comportamiento del mínimo de cada grupo, subestimándolo. Con relación al máximo, la mediana y el rango intercuartil, gran parte de los municipios parece capturarlos de manera adecuada, a pesar de que para algunos hay falla en la captura, ya que la región de densidad se encuentra al rededor de 0.5. En el caso de la desviación estándar, el modelo la sobreestima. Por último, para la media, el modelo parece capturar muy bien su comportamiento en cada municipio, pues apenas hay ppp's extremos (cercanos a 0 o 1) y la altura de su caja es muy pequeña.

Anexos

Modelos

M1: Modelo Normal

Distribución muestral:

$$y_{i,j}|\theta, \sigma^2 \text{iid} \sim N(\theta, \sigma^2), \text{ para } i = 1, \dots, n_j \text{ y } j = 1, \dots, m,$$

donde $y_{i,j}$ es el puntaje global del estudiante i en el departamento j y $N(\theta, \sigma^2)$ denota la distribución Normal con media θ y varianza σ^2 .

Distribución previa:

$$\begin{aligned}\theta &\sim N(\mu_0, \gamma_0^2), \\ \sigma^2 &\sim \text{GI}\left(\frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2}\right),\end{aligned}$$

donde μ_0 , γ_0^2 , ν_0 , y σ_0^2 son los hiperparámetros del modelo, y $\text{GI}(\alpha, \beta)$ denota la distribución Gamma-Inversa con media $\frac{\beta}{\alpha-1}$ para $\alpha > 1$ y varianza $\frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ para $\alpha > 2$.

Distribución posterior

$$\begin{aligned}p(\boldsymbol{\theta} \mid \mathbf{y}) &\propto p(\mathbf{y} \mid \boldsymbol{\theta})p(\boldsymbol{\theta}) \\ &= \prod_{i=1}^n N(y_i \mid \theta, \sigma^2) \times N(\theta \mid \mu_0, \tau_0^2) \times \text{GI}\left(\sigma^2 \mid \frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2}\right)\end{aligned}$$

Distribuciones condicionales completas

$$\begin{aligned}\theta \mid \text{resto} &\sim N\left(\frac{\frac{1}{\tau_0^2}\mu_0 + \frac{n}{\sigma^2}\bar{y}}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}}, \frac{1}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}}\right) \\ \sigma^2 \mid \text{resto} &\sim \text{GI}\left(\frac{\nu_0 + n}{2}, \frac{\nu_0 \sigma_0^2 + \sum_{i=1}^n (y_i - \theta)^2}{2}\right)\end{aligned}$$

Coefficientes de variación de Monte Carlo

θ	σ^2
0.00039	0.002

Tabla 5: Coeficientes de variación de Monte Carlo en porcentaje para el modelo 1

M2: Modelo Normal con medias específicas por departamento

Distribución muestral:

$$y_{i,j} | \theta_j, \sigma^2 \stackrel{ind}{\sim} N(\theta_j, \sigma^2).$$

Distribución previa:

$$\begin{aligned} \theta_j | \mu, \tau^2 &\stackrel{ind}{\sim} N(\mu, \tau^2), & \mu &\sim N(\mu_0, \gamma_0^2), \\ \tau^2 &\sim \text{GI}\left(\frac{\eta_0}{2}, \frac{\eta_0 \tau_0^2}{2}\right), & \sigma^2 &\sim \text{GI}\left(\frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2}\right), \end{aligned}$$

Donde $\mu_0, \gamma_0^2, \eta_0, \tau_0^2, \nu_0, \sigma_0^2$ son los hiperparámetros del modelo.

Distribución posterior

$$\begin{aligned} p(\boldsymbol{\theta} | \mathbf{y}) &\propto p(\mathbf{y} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) \\ &= \prod_{j=1}^m \prod_{i=1}^{n_j} N(y_{i,j} | \theta_j, \sigma^2) \times \prod_{j=1}^m N(\theta_j | \mu, \tau^2) \times N(\mu | \mu_0, \gamma_0^2) \\ &\quad \times \text{GI}\left(\tau^2 | \frac{\eta_0}{2}, \frac{\eta_0 \tau_0^2}{2}\right) \times \text{GI}\left(\sigma^2 | \frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2}\right) \end{aligned}$$

Distribuciones condicionales completas

$$\begin{aligned} \theta_j | \text{resto} &\sim N\left(\left(\frac{\mu}{\tau^2} + \frac{n_j \bar{y}_j}{\sigma^2}\right) \left(\frac{1}{\tau^2} + \frac{n_j}{\sigma^2}\right)^{-1}, \left(\frac{1}{\tau^2} + \frac{n_j}{\sigma^2}\right)^{-1}\right) \\ \sigma^2 | \text{resto} &\sim \text{GI}\left(\frac{\nu_0 + \sum_{j=1}^m n_j}{2}, \frac{\nu_0 \sigma_0^2 + \sum_{j=1}^m \sum_{i=1}^{n_j} (y_{i,j} - \theta_j)^2}{2}\right) \\ \mu | \text{resto} &\sim N\left(\left(\frac{\mu_0}{\gamma_0^2} + \frac{m \bar{\theta}}{\tau^2}\right) \left(\frac{1}{\gamma_0^2} + \frac{m}{\tau^2}\right)^{-1}, \left(\frac{1}{\gamma_0^2} + \frac{m}{\tau^2}\right)^{-1}\right) \\ \tau^2 | \text{resto} &\sim \text{GI}\left(\frac{\eta_0 + m}{2}, \frac{\eta_0 \tau_0^2 + \sum_{j=1}^m (\theta_j - \mu)^2}{2}\right) \end{aligned}$$

Donde n_j es el número de individuos en el departamento j , \bar{y}_j es la media muestral de los puntajes en el departamento j y $\bar{\theta}$ es la media muestral de los θ_j para $j = 1, \dots, m$.

Coefficientes de variación de Monte Carlo

Parámetro	σ^2	μ	τ^2
CV (%)	0.002	0.014	0.268

Tabla 6: Coeficientes de variación de Monte Carlo para el modelo 2

Parámetro	Mínimo	Q 25 %	Mediana	Media	Q 75 %	Máximo
θ_j	0.0007	0.0014	0.0018	0.0030	0.0033	0.0112

Tabla 7: Resumen de coeficientes de variación de Monte Carlo para θ_j

M3: Modelo Normal con medias y varianzas específicas por departamento

Distribución muestral:

$$y_{i,j} | \theta_j, \sigma_j^2 \stackrel{ind}{\sim} N(\theta_j, \sigma_j^2).$$

Distribución previa:

$$\begin{aligned} \theta_j | \mu, \tau^2 &\stackrel{ind}{\sim} N(\mu, \tau^2), & \mu &\sim N(\mu_0, \gamma_0^2), & \nu &= \text{Constante}, \\ \tau^2 &\sim \text{GI} \left(\frac{\eta_0}{2}, \frac{\eta_0 \tau_0^2}{2} \right), & \sigma_j^2 | \nu, \sigma^2 &\stackrel{ind}{\sim} \text{GI} \left(\frac{\nu}{2}, \frac{\nu \sigma^2}{2} \right), & \sigma^2 &\sim G \left(\frac{\alpha_0}{2}, \frac{\beta_0}{2} \right), \end{aligned}$$

Donde $\mu_0, \gamma_0^2, \eta_0, \tau_0^2, \nu, \alpha_0^2, \beta_0^2$ son los hiperparámetros del modelo, y $G(\alpha, \beta)$ denota la distribución Gamma con media $\frac{\alpha}{\beta}$ y varianza $\frac{\alpha}{\beta^2}$.

Distribución posterior

$$\begin{aligned} p(\boldsymbol{\theta} | \mathbf{y}) &\propto p(\mathbf{y} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) \\ &= \prod_{j=1}^m \prod_{i=1}^{n_j} N(y_{i,j} | \theta_j, \sigma_j^2) \times \prod_{j=1}^m N(\theta_j | \mu, \tau^2) \times \prod_{j=1}^m \text{GI} \left(\sigma_j^2 | \frac{\nu}{2}, \frac{\nu \sigma^2}{2} \right) \times N(\mu | \mu_0, \gamma_0^2) \\ &\times \text{GI} \left(\tau^2 | \frac{\eta_0}{2}, \frac{\eta_0 \tau_0^2}{2} \right) \times G \left(\sigma^2 | \frac{\alpha_0}{2}, \frac{\beta_0}{2} \right) \end{aligned}$$

Distribuciones condicionales completas

$$\begin{aligned}
\theta_j \mid \text{resto} &\sim \text{N} \left(\left(\frac{\mu}{\tau^2} + \frac{n_j \bar{y}_j}{\sigma_j^2} \right) \left(\frac{1}{\tau^2} + \frac{n_j}{\sigma_j^2} \right)^{-1}, \left(\frac{1}{\tau^2} + \frac{n_j}{\sigma_j^2} \right)^{-1} \right) \\
\sigma_j^2 \mid \text{resto} &\sim \text{GI} \left(\frac{\nu + n_j}{2}, \frac{\nu \sigma^2 + \sum_{i=1}^{n_j} (y_{i,j} - \theta_j)^2}{2} \right) \\
\sigma^2 \mid \text{resto} &\sim \text{G} \left(\frac{\alpha_0 + m\nu}{2}, \frac{\beta_0 + \nu \sum_{j=1}^m \frac{1}{\sigma_j^2}}{2} \right) \\
\mu \mid \text{resto} &\sim \text{N} \left(\left(\frac{\mu_0}{\gamma_0^2} + \frac{m\bar{\theta}}{\tau^2} \right) \left(\frac{1}{\gamma_0^2} + \frac{m}{\tau^2} \right)^{-1}, \left(\frac{1}{\gamma_0^2} + \frac{m}{\tau^2} \right)^{-1} \right) \\
\tau^2 \mid \text{resto} &\sim \text{GI} \left(\frac{\eta_0 + m}{2}, \frac{\eta_0 \tau_0^2 + \sum_{j=1}^m (\theta_j - \mu)^2}{2} \right)
\end{aligned}$$

Coefficientes de variación de Monte Carlo

Parámetro	σ^2	μ	τ^2
CV (%)	0.014	0.252	0.243

Tabla 8: Coeficientes de variación de Monte Carlo para el modelo 3

Parámetro	Mínimo	Q 25 %	Mediana	Media	Q 75 %	Máximo
θ_j	0.0007	0.0013	0.0018	0.0028	0.0029	0.0111
σ_j^2	0.0050	0.0107	0.0127	0.0202	0.0214	0.0720

Tabla 9: Resumen de coeficientes de variación de Monte Carlo en porcentaje para θ_j y σ_j^2 del modelo 3

M4: Modelo Normal con medias específicas por municipio y departamento

Distribución muestral:

$$y_{i,j,k} \mid \zeta_{j,k}, \kappa^2 \stackrel{\text{ind}}{\sim} \text{N}(\zeta_{j,k}, \kappa^2),$$

para $i = 1, \dots, n_{j,k}$, $j = 1, \dots, n_k$, y $k = 1, \dots, m$, donde $y_{i,j,k}$ es el puntaje global del estudiante i en el municipio j del departamento k .

Distribución previa:

$$\begin{aligned}
\zeta_{j,k} \mid \theta_k, \sigma^2 &\stackrel{\text{ind}}{\sim} \text{N}(\theta_k, \sigma^2), \\
\kappa^2 &\sim \text{GI} \left(\frac{\xi_0}{2}, \frac{\xi_0 \kappa_0^2}{2} \right), \\
\theta_k \mid \mu, \tau^2 &\stackrel{\text{ind}}{\sim} \text{N}(\mu, \tau^2),
\end{aligned}$$

$$\tau^2 \sim \text{GI} \left(\frac{\eta_0}{2}, \frac{\eta_0 \tau_0^2}{2} \right), \quad \mu \sim \text{N}(\mu_0, \gamma_0^2), \quad \sigma^2 \sim \text{GI} \left(\frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2} \right),$$

Donde $\xi_0, \kappa_0^2, \mu_0, \gamma_0^2, \eta_0, \tau_0^2, \nu_0, \sigma_0^2$ son los hiperparámetros del modelo.

Distribución posterior

$$\begin{aligned} p(\boldsymbol{\theta} \mid \mathbf{y}) &\propto p(\mathbf{y} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) \\ &= \prod_{k=1}^m \prod_{j=1}^{n_k} \prod_{i=1}^{n_{j,k}} \text{N}(y_{i,j,k} \mid \zeta_{j,k}, \kappa^2) \times \prod_{k=1}^m \prod_{j=1}^{n_k} \text{N}(\zeta_{j,k} \mid \theta_k, \sigma^2) \times \prod_{k=1}^m \text{N}(\theta_k \mid \mu, \tau^2) \times \text{N}(\mu \mid \mu_0, \gamma_0^2) \\ &\times \text{GI} \left(\tau^2 \mid \frac{\eta_0}{2}, \frac{\eta_0 \tau_0^2}{2} \right) \times \text{GI} \left(\kappa^2 \mid \frac{\xi_0}{2}, \frac{\xi_0 \kappa_0^2}{2} \right) \times \text{GI} \left(\sigma^2 \mid \frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2} \right) \end{aligned}$$

Distribuciones condicionales completas

$$\begin{aligned} \zeta_{j,k} \mid \text{resto} &\sim \text{N} \left(\left(\frac{\theta_k}{\sigma^2} + \frac{n_{j,k} \bar{y}_{j,k}}{\kappa^2} \right) \left(\frac{1}{\sigma^2} + \frac{n_{j,k}}{\kappa^2} \right)^{-1}, \left(\frac{1}{\sigma^2} + \frac{n_{j,k}}{\kappa^2} \right)^{-1} \right) \\ \mu \mid \text{resto} &\sim \text{N} \left(\left(\frac{\mu_0}{\gamma_0^2} + \frac{m \bar{\theta}}{\tau^2} \right) \left(\frac{1}{\gamma_0^2} + \frac{m}{\tau^2} \right)^{-1}, \left(\frac{1}{\gamma_0^2} + \frac{m}{\tau^2} \right)^{-1} \right) \\ \tau^2 \mid \text{resto} &\sim \text{GI} \left(\frac{\eta_0 + m}{2}, \frac{\eta_0 \tau_0^2 + \sum_{k=1}^m (\theta_k - \mu)^2}{2} \right) \\ \sigma^2 \mid \text{resto} &\sim \text{GI} \left(\frac{\nu_0 + \sum_{k=1}^m n_k}{2}, \frac{\nu_0 \sigma_0^2 + \sum_{k=1}^m \sum_{j=1}^{n_k} (\zeta_{j,k} - \theta_k)^2}{2} \right) \\ \kappa^2 \mid \text{resto} &\sim \text{GI} \left(\frac{\xi_0 + n}{2}, \frac{\xi_0 \kappa_0^2 + \sum_{k=1}^m \sum_{j=1}^{n_k} \sum_{i=1}^{n_{j,k}} (y_{i,j,k} - \zeta_{j,k})^2}{2} \right) \\ \theta_k \mid \text{resto} &\sim \text{N} \left(\left(\frac{\mu}{\tau^2} + \frac{n_k \bar{\zeta}_k}{\sigma^2} \right) \left(\frac{1}{\tau^2} + \frac{n_k}{\sigma^2} \right)^{-1}, \left(\frac{1}{\tau^2} + \frac{n_k}{\sigma^2} \right)^{-1} \right) \end{aligned}$$

Donde $n_{j,k}$ es el número de individuos en el municipio k del departamento j , $\bar{y}_{j,k}$ es la media muestral de los puntajes en el municipio k del departamento j y $\bar{\zeta}_k$ es la media muestral de los $\zeta_{j,k}$, con $j = 1, \dots, n_k$ en el departamento k .

Coefficientes de variación de Monte Carlo

Parámetro	κ^2	σ^2	μ	τ^2
CV (%)	0.002	0.048	0.015	0.272

Tabla 10: Coeficientes de variación de Monte Carlo para el modelo 4

Parámetro	Mínimo	Q 25 %	Mediana	Media	Q 75 %	Máximo
$\zeta_{j,k}$	0.0007	0.0117	0.0169	0.0182	0.0237	0.0582
θ_j	0.0059	0.0103	0.0130	0.0168	0.0176	0.0473

Tabla 11: Resumen de coeficientes de variación de Monte Carlo en porcentaje para $\zeta_{j,k}$ y θ_j del modelo 4

M5: Modelo Normal con medias específicas por municipio y departamento

Distribución muestral:

$$y_{i,j,k} | \zeta_{j,k}, \kappa^2 \stackrel{ind}{\sim} N(\zeta_{j,k}, \kappa^2).$$

Distribución previa:

$$\begin{aligned} \zeta_{j,k} | \theta_k, \sigma_k^2 &\stackrel{ind}{\sim} N(\theta_k, \sigma_k^2), & \theta_k | \mu, \tau^2 &\stackrel{ind}{\sim} N(\mu, \tau^2), & \sigma_k^2 | \nu, \sigma^2 &\stackrel{ind}{\sim} \text{GI}\left(\frac{\nu}{2}, \frac{\nu\sigma^2}{2}\right), \\ \kappa^2 &\sim \text{GI}\left(\frac{\xi_0}{2}, \frac{\xi_0\kappa_0^2}{2}\right), & \mu &\sim N(\mu_0, \gamma_0^2), & \nu &= \text{Constante}, \\ \tau^2 &\sim \text{GI}\left(\frac{\eta_0}{2}, \frac{\eta_0\tau_0^2}{2}\right), & \sigma^2 &\sim \text{G}\left(\frac{\alpha_0}{2}, \frac{\beta_0}{2}\right), \end{aligned}$$

Donde $\xi_0, \kappa_0^2, \mu_0, \gamma_0^2, \eta_0, \tau_0^2, \nu, \alpha_0, \beta_0$ son los hiperparámetros del modelo.

Distribución posterior

$$\begin{aligned} p(\boldsymbol{\theta} | \mathbf{y}) &\propto p(\mathbf{y} | \boldsymbol{\theta})p(\boldsymbol{\theta}) \\ &= \prod_{k=1}^m \prod_{j=1}^{n_k} \prod_{i=1}^{n_{j,k}} N(y_{i,j,k} | \zeta_{j,k}, \kappa^2) \times \prod_{k=1}^m \prod_{j=1}^{n_k} N(\zeta_{j,k} | \theta_k, \sigma_k^2) \times \prod_{k=1}^m N(\theta_k | \mu, \tau^2) \times N(\mu | \mu_0, \gamma_0^2) \\ &\times \prod_{k=1}^m \text{GI}\left(\sigma_k^2 | \frac{\nu}{2}, \frac{\nu\sigma^2}{2}\right) \times \text{GI}\left(\tau^2 | \frac{\eta_0}{2}, \frac{\eta_0\tau_0^2}{2}\right) \times \text{GI}\left(\kappa^2 | \frac{\xi_0}{2}, \frac{\xi_0\kappa_0^2}{2}\right) \times \text{GI}\left(\sigma^2 | \frac{\nu_0}{2}, \frac{\nu_0\sigma_0^2}{2}\right) \\ &\times \text{G}\left(\sigma^2 | \frac{\alpha_0}{2}, \frac{\beta_0}{2}\right) \end{aligned}$$

Distribuciones condicionales completas

$$\begin{aligned}
\zeta_{j,k} \mid \text{resto} &\sim \text{N} \left(\left(\frac{\theta_k}{\sigma_k^2} + \frac{n_{j,k} \bar{y}_{j,k}}{\kappa^2} \right) \left(\frac{1}{\sigma_k^2} + \frac{n_{j,k}}{\kappa^2} \right)^{-1}, \left(\frac{1}{\sigma_k^2} + \frac{n_{j,k}}{\kappa^2} \right)^{-1} \right) \\
\mu \mid \text{resto} &\sim \text{N} \left(\left(\frac{\mu_0}{\gamma_0^2} + \frac{m \bar{\theta}}{\tau^2} \right) \left(\frac{1}{\gamma_0^2} + \frac{m}{\tau^2} \right)^{-1}, \left(\frac{1}{\gamma_0^2} + \frac{m}{\tau^2} \right)^{-1} \right) \\
\tau^2 \mid \text{resto} &\sim \text{GI} \left(\frac{\eta_0 + m}{2}, \frac{\eta_0 \tau_0^2 + \sum_{k=1}^m (\theta_k - \mu)^2}{2} \right) \\
\theta_k \mid \text{resto} &\sim \text{N} \left(\left(\frac{\mu}{\tau^2} + \frac{n_k \bar{\zeta}_k}{\sigma_k^2} \right) \left(\frac{1}{\tau^2} + \frac{n_k}{\sigma_k^2} \right)^{-1}, \left(\frac{1}{\tau^2} + \frac{n_k}{\sigma_k^2} \right)^{-1} \right) \\
\kappa^2 \mid \text{resto} &\sim \text{GI} \left(\frac{\xi_0 + n}{2}, \frac{\xi_0 \kappa_0^2 + \sum_{k=1}^m \sum_{j=1}^{n_k} \sum_{i=1}^{n_{j,k}} (y_{i,j,k} - \zeta_{j,k})^2}{2} \right) \\
\sigma^2 \mid \text{resto} &\sim \text{G} \left(\frac{\alpha_0 + mv}{2}, \frac{\beta_0 + v \cdot \sum_{k=1}^m \frac{1}{\sigma_k^2}}{2} \right) \\
\sigma_k^2 \mid \text{resto} &\sim \text{GI} \left(\frac{n_k + \nu}{2}, \frac{\nu \sigma^2 + \sum_{j=1}^{n_k} (\zeta_{j,k} - \theta_k)^2}{2} \right)
\end{aligned}$$

Parámetro	κ^2	μ	τ^2	σ^2
CV (%)	0.002	0.014	0.284	0.275

Tabla 12: Coeficientes de variación de Monte Carlo para el modelo 5

Parámetro	Mínimo	Q 25 %	Mediana	Media	Q 75 %	Máximo
$\zeta_{j,k}$	0.0007	0.0116	0.0168	0.0180	0.0233	0.0645
θ_j	0.0055	0.0094	0.0125	0.0177	0.0185	0.0707
σ_j^2	0.1336	0.2446	0.3135	0.7927	0.5158	9.7532

Tabla 13: Resumen de coeficientes de variación de Monte Carlo en porcentaje para $\zeta_{j,k}$, θ_j y σ_j^2 del modelo 5