

UNIVERSIDAD  
**NACIONAL**  
DE COLOMBIA

Universidad Nacional de Colombia

## Modelos de regresión Bayesianos

Autor:

Juan Daniel Diaz Alvarez  
juadiazal@unal.edu.co

Juan Sosa  
Ph.D Estadística

Noviembre del 2023

# Índice

<b>1. Introducción</b>	<b>2</b>
<b>2. Análisis electoral: alcaldía de Bogotá 2023</b>	<b>2</b>
2.1. Ajuste del Modelo . . . . .	3
<b>3. Análisis de Diabetes</b>	<b>4</b>
3.1. Modelo Bayesianos y su estructura jerárquica . . . . .	5
3.2. Ajuste de los modelos . . . . .	7
3.3. Comparación y bondad de ajuste . . . . .	8
3.4. Conclusión . . . . .	9
<b>4. Apéndice</b>	<b>9</b>
4.1. Análisis electoral: alcaldía de Bogotá . . . . .	10
4.1.1. Distributions Condicionales Completas - DCC . . . . .	11
4.1.2. Algoritmo del muestreador . . . . .	11
4.1.3. Diagnóstico de convergencia . . . . .	12
4.2. Análisis de Diabetes . . . . .	13
4.2.1. Previa unitaria . . . . .	13
4.2.2. Previa g . . . . .	13
4.2.3. Regresión rígida . . . . .	14
4.2.4. Regresión con errores correlacionados . . . . .	15
4.2.5. Convergencia de los muestreadores . . . . .	16

# 1. Introducción

Este trabajo se divide en dos partes fundamentales. En la primera, se lleva a cabo un análisis electoral de la alcaldía de Bogotá, con el objetivo de estimar a nivel poblacional, la proporción de votos que recibirían los candidatos a dicha alcaldía, utilizando como muestra los datos recopilados por la encuestadora Invamer, dado que fue la que obtuvo menor error en las estimación de la proporción de votos de los candidatos a la alcaldía de Bogotá.

La segunda parte se centra en la construcción de un modelo predictivo para la medida de progresividad en 442 pacientes que sufren de diabetes. Para esto se proponen 4 modelos con la finalidad de seleccionar el modelo que minimiza el Error Absoluto Medio (EAM). Los dos primeros modelos consideran la previa unitaria y la previa  $g$ , mientras que los dos últimos incorporan un modelo mediante regresión rígida y otro con errores correlacionados. Por otra parte, se validará la bondad de ajuste de cada modelo usando como estadístico de prueba la media, así como los criterios de información como el DIC. En ambas apartados emplearemos el muestreador de Gibbs y Metropolis para la simulación de la distribución posterior.

## 2. Análisis electoral: alcaldía de Bogotá 2023

La Silla Vacía evaluó el desempeño de las encuestadoras en las elecciones regionales de 2023 en las cinco ciudades principales de Colombia. En el artículo, se compara las predicciones de las encuestadoras con los resultados reales de las elecciones y analiza la precisión de cada encuestadora. En este analisis, se encontró que la encuestadora **Invamer** fue la que tuvo *menor error total* para la ciudad de Bogotá (Restrepo, 2023).

Invamer llevó a cabo una encuesta entre el 17 y el 23 de octubre de 2023 para Noticias Caracol, Blu Radio y El Espectador, con el propósito de evaluar la intención de voto en las elecciones para alcaldes del 2023 en Bogotá, Medellín, Cali, Barranquilla y Bucaramanga. En Bogotá, se encuestaron 1200 personas (entre hombres y mujeres) de 18 años en adelante, de diversos estratos socioeconómicos, con derecho a voto y residencia en la ciudad. Las encuestas se realizaron personalmente en los hogares utilizando tablets, y se empleó un tarjetón para las preguntas sobre intención de voto. Los resultados específicos se encuentran detallados en la Tabla 1. (Invamer, 2023).

Candidato	Cantidad	Proporción
C. F. Galán	493	0.411
G. Bolivar	257	0.214
J. D. Oviedo	227	0.189
D. Molano	48	0.040
R. Lara	41	0.034
J. L. Vargas	38	0.032
J. E. Robledo	28	0.023
N. Ramos	11	0.009
R. A. Quintero	3	0.003
Voto en Blanco	54	0.045
<b>Total</b>	<b>1200</b>	<b>1.000</b>

Tabla 1: Intención de voto en Bogotá

Invamer utilizó un tipo particular de muestreo aleatorio simple sin reemplazo (ver “Opinión Alcaldes 2023”, [2023](#)) es factible considerar la muestra como una muestra aleatoria simple con reemplazo (IID). Esto se debe a que el tamaño total de la muestra es significativamente pequeño en comparación con el tamaño del Universo. El enfoque primordial de este apartado, tanto en el ámbito científico como político, reside en la estimación de la proporción de votos de cada candidato a nivel poblacional, utilizando los datos suministrados por Invamer.

## 2.1. Ajuste del Modelo

Ateniendo al objetivo de este apartado, en la Sección 4 se encuentra la información del modelo Bayesiano con el que se va a estimar la proporción de votos. Así pues, ajustado el modelo Bayesiano por medio de el muestreador de Gibbs y Metropolis se obtuvieron los siguientes resultados:

Candidatos	Media	CV %	IC Inf	IC Sup
C. F. Galán	40.90	3.41	38.197	43.681
G. Bolivar	21.36	5.46	19.113	23.661
J. D. Oviedo	18.87	5.87	16.723	21.096

Tabla 2: Proporción de votos estimadas

La información de la Tabla 2 presenta la estimación de la proporción de votos para los candidatos Carlos Fernando Galán, Gustavo Bolivar y Juan Daniel Oviedo. Se observa un coeficiente de variación (CV) bajo y una amplitud en los intervalos de credibilidad moderada, lo que nos lleva a concluir que existe una baja incertidumbre en las estimaciones.

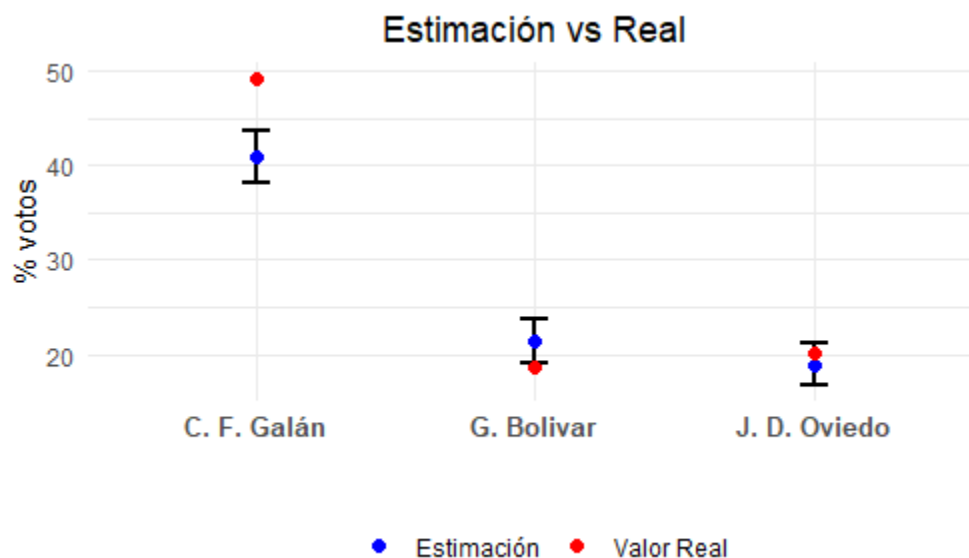


Figura 1: Estimación vs Valor real

En la Figura 1 se comparan la estimación de la proporción de votos con las que se obtuvieron en los resultados de las elecciones el pasado 29 de Octubre de 2023. Se observa que para el candidato Galán y Oviedo se subestimó la proporción de votos, mientras que para el candidato Bolivar se sobrestimó. Sin embargo se logró estimar que el ganador de la alcaldía de Bogotá iba a ser el candidato Galán. Cabe mencionar que el intervalo de credibilidad para el candidato Oviedo contiene el valor real, contrario a los demás candidatos.

### 3. Análisis de Diabetes

Considere la base de datos sobre diabetes proporcionada en la Sección 9.3 de Hoff, 2009. En este conjunto de datos se relaciona 10 medidas iniciales  $x_1, \dots, x_{10}$  en un grupo de 442 pacientes diabéticos, así como una medida de progresión de la enfermedad  $y$  tomada un año después de las medidas iniciales. Los datos pueden se pueden descargar de este [enlace](#) en los archivos llamados  $yX.diabetes.train$  y  $yX.diabetes.test$ .

El objetivo principal es construir un modelo predictivo para  $y$  basado en  $x_1, \dots, x_{10}$  (donde tanto  $y$  como las  $x_j$  están estandarizadas). Además, se sospecha que la relación entre  $y$  y las  $x_j$  podría no ser lineal, por lo tanto se añaden termino de segundo orden en la forma de la  $x_j^2$  y  $x_j x_k$ . En este contexto, las variables regresoras incluyen diez efectos principales  $x_j$ , 45 interacciones  $x_j x_k$ , y nueve términos cuadráticos  $x_j^2$  (no es necesario incluir  $x_2^2$  en el modelo debido a que  $x_2 = \text{sexo}$  es binario, por lo que  $x_2^2 = x_2$ ). Esto resulta en un total de  $p = 64$  variables regresoras. No se considera el intercepto ya que las variables se encuentran estandarizadas.

Para la construcción del modelo predictivo se proponen 4 modelos con la finalidad de seleccionar el modelo con mayor capacidad predictiva, es decir, elegir el modelo que minimiza el Error Absoluto Medio (EAM). Para lograr esto, el conjunto de datos está dividido en dos partes: entrenamiento y testeo. De esta manera se ajustarán los modelos con el conjunto de entrenamiento y se validará en los conjunto de testeo.

### 3.1. Modelo Bayesianos y su estructura jerárquica

Consideraremos 4 modelos de regresión Bayesianos. Para los primeros 3 modelos consideremos la misma distribución muestral dada por:

$$\mathbf{y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2 \sim \mathbf{N}_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I}_n)$$

Con  $\mathbf{y}_{(n \times 1)}$  es el vector de respuesta,  $\boldsymbol{\beta}_{(p \times 1)}$  vector de parámetros a estimar,  $\mathbf{X}_{(n \times p)}$  matriz de diseño y  $\mathbf{I}_n$  matriz identidad de  $n \times n$ . Además, para los primeros 2 modelos consideraremos la distribución previa dada por:

$$\begin{aligned} \boldsymbol{\beta} &\sim \mathbf{N}_p(\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0) \\ \sigma^2 &\sim \text{Gl} \left( \frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2} \right) \end{aligned}$$

1. **Previa unitaria:** esta propone usar los datos para asignarles valores a los hiperparámetros. De manera que:

$$\boldsymbol{\beta}_0 = \hat{\boldsymbol{\beta}}_{\text{ols}}, \quad \boldsymbol{\Sigma}_0 = n \hat{\sigma}_{\text{ols}}^2 (\mathbf{X}^\top \mathbf{X})^{-1}, \quad \nu_0 = 1, \quad \sigma_0^2 = \hat{\sigma}_{\text{ols}}^2.$$

Donde **ols** hace referencia a Ordinary Least Squares (Mínimos Cuadrados Ordinarios).

2. **Previa g:** al igual que la previa unitaria, propone usar los datos para asignares valores a los hiperparámetros. Lo que las diferencia es que en este caso  $\beta$  depende del valor de  $\sigma^2$ . De manera que:

$$\beta_0 = \mathbf{0}_p, \quad \Sigma_0 = g\sigma^2(\mathbf{X}^\top \mathbf{X})^{-1}$$

En  $\Sigma_0$  se observa la dependencia de  $\beta$  en  $\sigma^2$ . Los parámetros para la distribución de  $\sigma^2$  se mantienen como en la previa unitaria.

3. **Regresión rígida:** Consideremos la siguiente distribución previa:

$$\beta \sim N_p\left(\mathbf{0}_p, \frac{\sigma^2}{\lambda} \mathbf{I}_p\right), \quad \sigma^2 \sim \text{Gl}\left(\frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2}\right), \quad \lambda \sim G(a_\lambda, b_\lambda)$$

Con  $\nu_0 = 1$ ,  $\sigma_0^2 = \hat{\sigma}_{\text{OLS}}^2$ ,  $a_\lambda = 1$  y  $b_\lambda = 2$ .

4. **Regresión con errores correlacionados:** consideramos la siguiente distribución muestral:

$$\mathbf{y} \mid \mathbf{X}, \beta, \sigma^2, \rho \sim N_n(\mathbf{X}\beta, \sigma^2 \mathbf{C}_\rho)$$

Siendo  $\mathbf{C}_\rho$  es una matriz con estructura autoregresiva de primer orden:

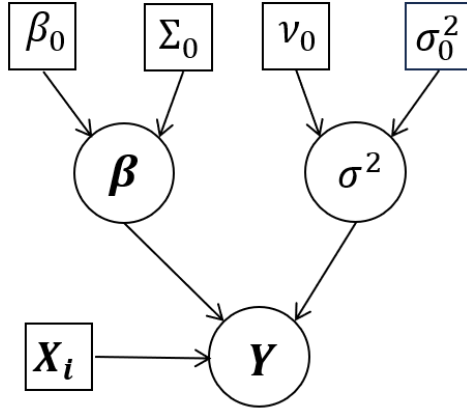
$$\mathbf{C}_\rho = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{bmatrix}$$

Por otro lado, se propone la siguiente distribución previa:

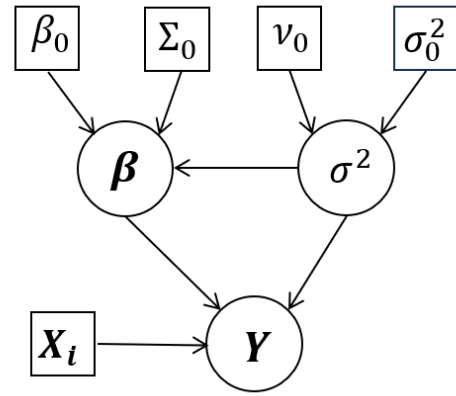
$$\beta_j \stackrel{\text{iid}}{\sim} N(0, \tau_0^2), \quad \sigma^2 \sim \text{Gl}\left(\frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2}\right), \quad \rho \sim U(a_\rho, b_\rho)$$

Con  $\tau_0^2 = 50$ ,  $\nu_0 = 1$ ,  $\sigma^2 = \sigma_{\text{ols}}^2$ ,  $a_\rho = 0$  y  $b_\rho = 1$ .

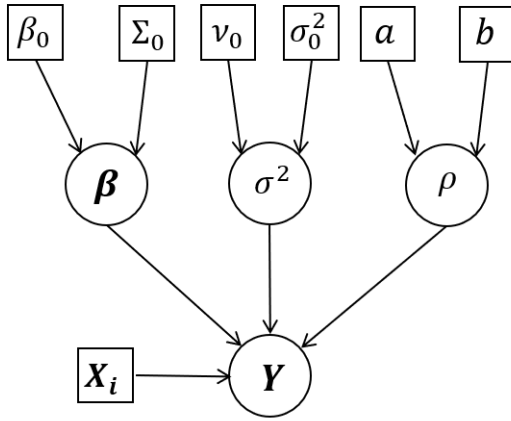
Los 4 modelos descritos anteriormente, se ajustarán usando el muestreador de Gibss y el algoritmo de Metropolis, usando 21.000 iteraciones, con 1.000 muestras de calentamiento y un muestreo sistemático con una amplitud de 2. En la Figura 2 se presentan visualmente los Grafos Acíclicos Dirigidos (DAG, por sus siglas en inglés) correspondientes a los modelos previamente discutidos.



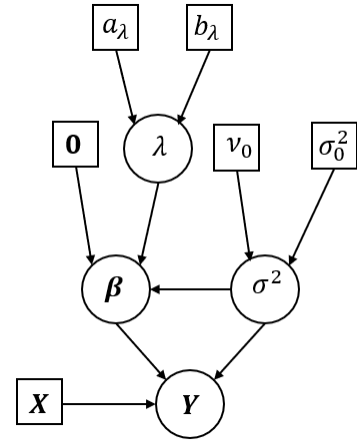
(a) DAG previa unitaria



(b) DAG previa g



(c) DAG errores correlacionados



(d) DAG regresión rígida

Figura 2: DAGs

### 3.2. Ajuste de los modelos

Sea  $\hat{y}_{\text{test}} = \mathbf{X}_{\text{test}} \hat{\beta}$ , donde  $\hat{\beta}$  es la estimación de los coeficientes de regresión usando el conjunto de entrenamiento. Así, el Error Absoluto Medio, EAM =  $\frac{1}{n} |\hat{y}_{\text{test}} - y_{\text{test}}|$ . No obstante, ajustados los modelos de la forma descrita en la sección anterior y calculados los EAM se obtuvieron los resultados descritos en la Figura 3. En la que claramente se observa que el modelo con el menor EAM es el modelo de regresión rígida, sin embargo la previa unitaria y la previa g tienen el mismo EAM, y la regresión con errores correlacionados es un poco menor a los dos anteriores.



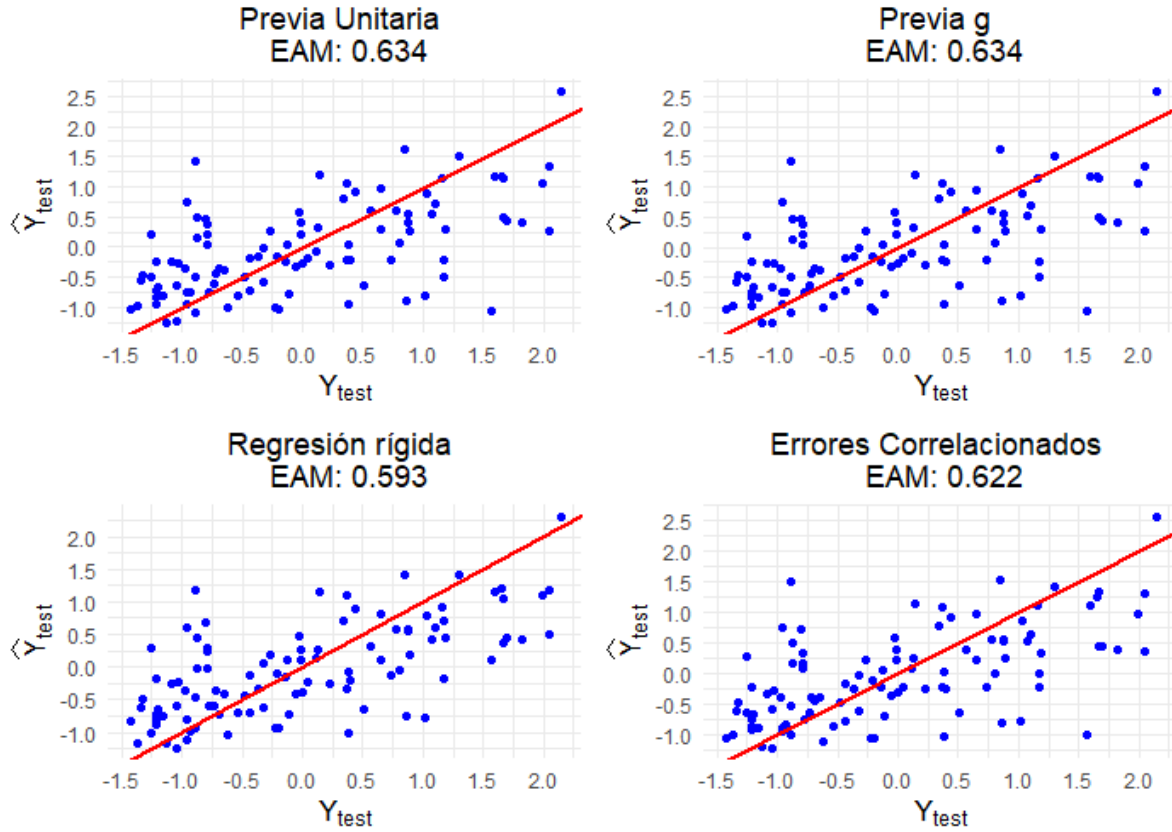


Figura 3:  $\hat{y}_{test}$  vs  $y_{test}$

### 3.3. Comparación y bondad de ajuste

Atendiendo al Criterio de Información de la Devianza (DIC, por sus siglas en inglés), se comparan los modelos y se obtienen los resultados a continuación:

Modelo	DIC
Previa Unitaria	784.655
Previa g	777.928
Rígida	759.255
E. Corr.	782.172

Tabla 3: Criterio de información

Con la información de la Tabla 3 se puede apreciar que el modelo de tipo regresión rígida sobresale en comparación a los demás modelos, dado que atiende al menor DIC por lo tanto, la precisión predictiva de este modelo es mayor que los demás.

Por otro lado, se evaluó la bondad de ajuste para los 4 modelos, usando como estadístico de prueba la media, así se calculó el valor p predictivo posterior (PPP) y se tuvieron los resultados descritos en Figura 4.

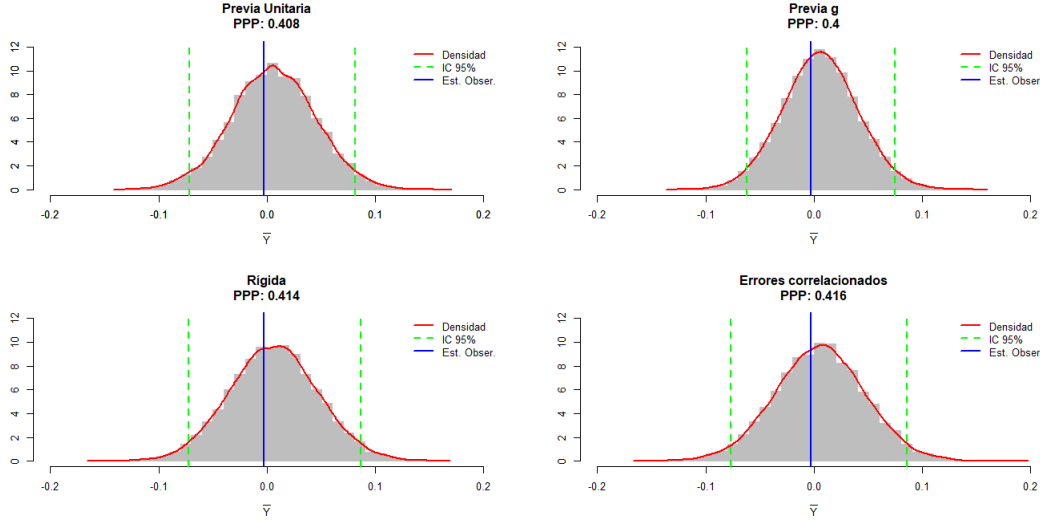


Figura 4: Bondad de ajuste

Así que para los 4 modelos la media está siendo estimada correctamente, dado que el PPP está cercano a 0.5 además que la estadística observada (Est. Obser.) se encuentra dentro de los Intervalos de Confianza (IC).

### 3.4. Conclusión

Los resultados revelaron que el **modelo de regresión rígida** sobresalió en términos de capacidad predictiva para las medidas de progresividad de los pacientes que padecen de diabetes, minimizando el Error Absoluto Medio (EAM). Esta consistencia se reflejó en el Criterio de Información de la Devianza (DIC), donde el modelo de regresión rígida exhibió el menor DIC, indicando una mejor precisión predictiva en comparación con los modelos alternativos. Además, la evaluación de la bondad de ajuste respaldó aún más la elección de este modelo.

## 4. Apéndice

Se presentan algunos resultados de la distribución posterior, Distribuciones Condicionales Completas (DCC), algoritmo usado para la generación de

las muestras, así como resúmenes de la Cadena de Markov de Monte Carlo - CMCM, tales como: Gráfica de la log verosimilitud en función de las iteraciones, Error de Monte Carlo - EMC, Coeficiente de Variación de Monte Carlo - CVMC, y tamaños efectivos de muestras - neff para cada uno de los parámetros del modelo que lo requieran.

#### 4.1. Análisis electoral: alcaldía de Bogotá

Para este análisis consideraremos el siguiente modelo Bayesiano, teniendo en cuenta que  $\mathbf{n} = (n_1, \dots, n_k)$ ,  $\Theta = (\theta_1, \dots, \theta_k)$ . Además,  $\sum_{j=1}^k n_j = n$  y  $\sum_{j=1}^k \theta_j = 1$ .

1. **Distribución muestral:**  $\mathbf{n}|n, \Theta \sim \text{Multinomial}(n, \Theta)$ .

$$p(\mathbf{n}|n, \Theta) = \frac{n!}{\prod_{j=1}^k n_j!} \prod_{j=1}^k \theta_j^{n_j}$$

2. **Distribución previa:**  $\Theta|\alpha \sim \text{Dirichlet}(\alpha \mathbf{1}_k)$  y  $\alpha \sim \text{Gamma}(a, b)$

$$p(\Theta|\alpha) = \frac{\Gamma(k\alpha)}{\Gamma(\alpha)^k} \prod_{j=1}^k \theta_j^{\alpha-1} \quad p(\alpha) = \frac{b^a}{\Gamma(a)} \alpha^{a-1} \exp(-b\alpha)$$

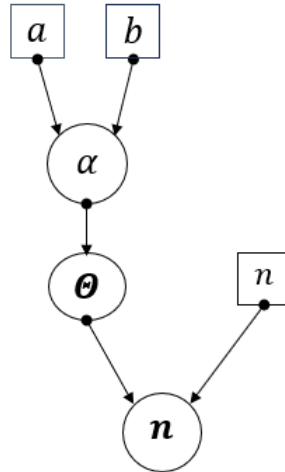


Figura 5: DAG

Obteniendo así que el vector de parámetros a estimar es  $\Phi = (\Theta, \alpha)$ . Por otro lado, se considera  $\alpha$  como un escalar, dado que para la distribución considera que cada  $\theta_j$  tiene el mismo valor de  $\alpha$ , lo que nos lleva a una distribución de **Dirichlet central**. Por lo tanto se tiene la siguiente distribución posterior:

$$p(\Phi|\mathbf{n}) \propto p(\mathbf{n}|n, \Theta) \times p(\Theta|\alpha) \times p(\alpha|a, b)$$

#### 4.1.1. Distributions Condicionales Completas - DCC

A)  $\Theta| - \sim \text{Dirichlet}(\mathbf{n} + \alpha)$

B) En este caso la DCC para  $\alpha$  no es conocida, por lo tanto la distribución tendría la siguiente expresión:

$$p(\alpha| -) \propto \frac{\Gamma(k\alpha)}{\Gamma(\alpha)^k} \prod_{j=1}^k \theta_j^{\alpha-1} \times \alpha^{a-1} \exp(-b\alpha)$$

Para obtener valor de  $\alpha$  usaremos el algoritmo de Metropolis, considerando la siguiente transformación  $\gamma = \log(\alpha)$ . Lo que nos lleva a la siguiente expresión:

$$p(\gamma| -) \propto \frac{\Gamma(ke^\gamma)}{\Gamma(e^\gamma)^k} \prod_{j=1}^k \theta_j^{e^\gamma-1} \times \exp[\gamma(a-1) - be^\gamma + \gamma] \quad (1)$$

#### 4.1.2. Algoritmo del muestreador

En este apartado utilizamos un muestreador de gibbs, dado que la distribución condicional de  $\Theta$  se conoce, mientras que la distribución de  $\alpha$  no tenía forma conocida, lo que se optó por usar el algoritmo de Metropolis para  $\alpha$ , lo que llevó a la transformación dada en (1). Así el algoritmo es el siguiente:

1. Establezco valores iniciales  $\Theta^{(0)}$  y  $\gamma^{(0)}$ .
2. Dada el estado actual  $\Theta^{(b)}, \gamma^{(b)}$ .

- a) Simulo  $\Theta^{(b+1)} \sim \text{Dirichlet}(\mathbf{n} + \exp[\gamma^{(b)}])$
- b) Simulo  $\gamma^{(b+1)}$  usando Metropolis como sigue:

1)  $\gamma^* \sim N(\gamma^{(b)}, \delta)$ .

2) Calculo la tasa de aceptación:

$$r = \exp \left( \log p(\gamma^* | -) - \log p(\gamma^{(b)} | -) \right)$$

3) Genero  $u \sim U(0, 1)$ . Si  $u < r$  entonces  $\gamma^{(b+1)} = \gamma^*$ . Si no  $\gamma^{(b+1)} = \gamma^{(b)}$

3. Almaceno  $\Theta^{(b+1)}$  y  $\exp(\gamma^{(b+1)})$ .

#### 4.1.3. Diagnóstico de convergencia

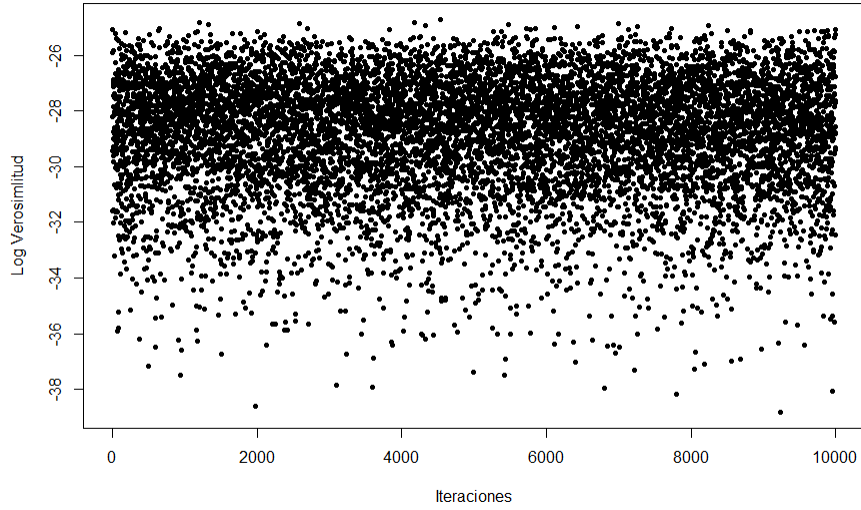


Figura 6: Convergencia del muestreador

$\Phi$	EMC	CVMC %	neff
$\theta_1$	0.00014	0.03479	9617.203
$\theta_2$	0.00012	0.05519	9953.427
$\theta_3$	0.00011	0.06014	9718.607
$\alpha$	0.00258	0.37206	10000

Tabla 4: Diagnóstico

## 4.2. Análisis de Diabetes

### 4.2.1. Previa unitaria

Se tiene que el vector de parámetros a estimar  $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2)$ . Teniendo así como distribución posterior:

$$p(\boldsymbol{\theta} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2) \times p(\boldsymbol{\beta} \mid \boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0) \times p(\sigma^2 \mid \nu_0, \sigma_0^2) \quad (2)$$

A) **Distribuciones Condicionales Completas:**

$$\begin{aligned} \boldsymbol{\beta} \mid - &\sim N_p \left( (\boldsymbol{\Sigma}_0^{-1} + \frac{1}{\sigma^2} \mathbf{X}^T \mathbf{X})^{-1} (\boldsymbol{\Sigma}_0^{-1} \boldsymbol{\beta}_0 + \frac{1}{\sigma^2} \mathbf{X}^T \mathbf{y}), (\boldsymbol{\Sigma}_0^{-1} + \frac{1}{\sigma^2} \mathbf{X}^T \mathbf{X})^{-1} \right) \\ \sigma^2 \mid - &\sim \text{Gl} \left( \frac{\nu_0 + n}{2}, \frac{\nu_0 \sigma_0^2 + \text{SSR}(\boldsymbol{\beta})}{2} \right) \end{aligned}$$

Siendo  $\text{SSR}(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$

B) **Algoritmo del muestreador:** utilizamos el muestreado de Gibbs dada que las DCC son conocidas, teniendo así el siguiente algoritmo:

- 1) Establezco los valores iniciales para  $\boldsymbol{\beta}^{(0)}$  y  $\sigma^{2(0)}$ .
- 2) Dado el estado actual  $\boldsymbol{\beta}^{(b)}$  y  $\sigma^{2(b)}$ :
  - a) Simulo  $\boldsymbol{\beta}^{(b+1)} \sim p(\boldsymbol{\beta} \mid \sigma^{2(b)}, \mathbf{y})$
  - b) Simulo  $\sigma^{2(b+1)} \sim p(\sigma^2 \mid \boldsymbol{\beta}^{(b+1)}, \mathbf{y})$
- 3) Almacenar  $\boldsymbol{\beta}^{(b+1)}, \sigma^{2(b+1)}$ .

### 4.2.2. Previa g

Se tiene el vector de parámetros a estimar  $\boldsymbol{\eta} = (\boldsymbol{\beta}, \sigma^2)$ . Teniendo así que la distribución posterior está dada por:

$$p(\boldsymbol{\eta} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2) \times p(\boldsymbol{\beta} \mid \sigma^2, \mathbf{X}, \mathbf{y}) \times p(\sigma^2 \mid \mathbf{X}, \mathbf{y})$$

A) **Distribuciones Condicionales Completas:**

$$\boldsymbol{\beta} \mid - \sim N_p \left( \frac{g}{g+1} (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}, \frac{g}{g+1} \sigma^2 (\mathbf{X}^\top \mathbf{X})^{-1} \right)$$

$$\sigma^2 \mid - \sim \text{Gl} \left( \frac{\nu_0 + n}{2}, \frac{\nu_0 \sigma_0^2 + \text{SSR}_g}{2} \right)$$

$$\text{Con } \text{SSR}_g = \mathbf{y}^\top \left( \mathbf{I}_n - \frac{g}{g+1} \mathbf{X} (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \right) \mathbf{y}$$

B) **Algoritmo del muestreador:** dada las característica de este previa, el muestreo consiste en simulación de Monte Carlo directo, lo que nos lleva a:

1) Simulo  $\sigma^2 \sim p(\sigma^2 \mid -)$ .

2) Simulo  $\boldsymbol{\beta} \sim p(\boldsymbol{\beta} \mid -)$ .

Por lo tanto no se necesitan diagnóstico de convergencia, dado de que no es un MCMC.

### 4.2.3. Regresión rígida

Se tiene que el vector de parámetros a estimar  $\boldsymbol{\Gamma} = (\boldsymbol{\beta}, \sigma^2, \lambda)$ . Teniendo así como distribución posterior:

$$p(\boldsymbol{\Gamma} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2) \times p(\boldsymbol{\beta} \mid \mathbf{0}_p, \frac{\sigma^2}{\lambda} \mathbf{I}_p) \times p(\sigma^2 \mid \frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2}) \times p(\lambda \mid a_\lambda, b_\lambda)$$

A) **Distribuciones Condicionales Completas:**

$$\boldsymbol{\beta} \mid - \sim N_p \left( (\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I}_p)^{-1} \mathbf{X}^\top \mathbf{y}, (\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I}_p)^{-1} \right)$$

$$\sigma^2 \mid - \sim \text{Gl} \left( \frac{\nu_0 + p + n}{2}, \frac{\nu_0 \sigma_0^2 + \lambda \boldsymbol{\beta}^\top \boldsymbol{\beta} + (\mathbf{y} - \mathbf{X} \boldsymbol{\beta})^\top (\mathbf{y} - \mathbf{X} \boldsymbol{\beta})}{2} \right)$$

$$\lambda \mid - \sim \text{G} \left( a_\lambda + \frac{p}{2}, b_\lambda + \frac{\boldsymbol{\beta}^\top \boldsymbol{\beta}}{2\sigma^2} \right)$$

B) **Algoritmo del muestreador:** utilizamos el muestreado de Gibbs dada que las DCC son conocidas, teniendo así el siguiente algoritmo:

- 1) Establezco los valores iniciales para  $\boldsymbol{\lambda}^{(0)}, \boldsymbol{\beta}^{(0)}$  y  $\sigma^{2(0)}$ .
- 2) Dado el estado actual  $\boldsymbol{\lambda}^{(b)}, \boldsymbol{\beta}^{(b)}$  y  $\sigma^{2(b)}$ :
  - a) Simulo  $\lambda^{(b+1)} \sim p(\lambda \mid \beta^{(b)} \sigma^{2(b)})$
  - b) Simulo  $\beta^{(b+1)} \sim p(\boldsymbol{\beta} \mid \lambda^{(b+1)} \sigma^{2(b)}, \mathbf{X}\mathbf{y})$
  - c) Simulo  $\sigma^{2(b+1)} \sim p(\sigma^2 \mid \beta^{(b+1)} \lambda^{(b+1)}, \mathbf{X}\mathbf{y})$
- 3) Almacenar  $\boldsymbol{\beta}^{(b+1)}, \sigma^{2(b+1)}, \lambda^{(b+1)}$ .

#### 4.2.4. Regresión con errores correlacionados

Se tiene que el vector de parámetros a estimar  $\boldsymbol{\Omega} = (\boldsymbol{\beta}, \sigma^2, \rho)$ . Teniendo así como distribución posterior:

$$p(\boldsymbol{\Omega} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{X}, \boldsymbol{\beta}, \sigma^2, \rho) \times \prod_{j=1}^p p(\beta_j \mid 0, \tau_0^2) \times p(\sigma^2 \mid \nu_0, \sigma_0^2) \times U(\rho \mid a_\rho, b_\rho)$$

##### A) Distribuciones Condicionales Completas:

$$\begin{aligned} \boldsymbol{\beta} \mid - &\sim N(\boldsymbol{\beta}_n, \boldsymbol{\Sigma}_n) \\ \sigma^2 \mid - &\sim \text{IG} \left( \frac{\nu_0 + n}{2}, \frac{\nu_0 \sigma_0^2 + \text{SSR}_\rho}{2} \right) \\ p(\rho \mid -) &\propto |\sigma^2 \mathbf{C}_\rho|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top \mathbf{C}_\rho^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right\} \end{aligned}$$

Con  $\boldsymbol{\Sigma}_n = (\boldsymbol{\Sigma}_0^{-1} + \frac{1}{\sigma^2} \mathbf{X}^\top \mathbf{C}_\rho \mathbf{X})^{-1}$ ,  $\text{SSR}_\rho = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top \mathbf{C}_\rho^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$  y  $\boldsymbol{\beta}_n = \boldsymbol{\Sigma}_n (\boldsymbol{\Sigma}_0^{-1} \boldsymbol{\beta}_0 + \frac{1}{\sigma^2} \mathbf{X}^\top \mathbf{C}_\rho^{-1} \mathbf{y})$ . Además, la DCC de  $\rho$  no se conoce, así que se propone usar Metropolis para las muestras de  $\rho$ .

##### B) Algoritmo del muestreador:

- 1) Establezco los valores iniciales para  $\boldsymbol{\beta}^{(0)}, \sigma^{2(0)}$  y  $\rho^{(0)}$ .
- 2) Dado el estado actual  $\boldsymbol{\beta}^{(b)}, \sigma^{2(b)}$  y  $\rho^{(b)}$ :
  - a) Simulo  $\boldsymbol{\beta}^{(b+1)} \sim p(\boldsymbol{\beta} \mid -)$
  - b) Simulo  $\sigma^{2(b+1)} \sim p(\sigma^2 \mid -)$
  - c) Simulo  $\rho^{(b+1)}$  de la siguiente manera:



I)  $\rho \sim U(\rho \mid \rho^{(b)} - \delta, \rho^{(b)} + \delta).$

Establezco  $\rho^* = \min\{|\rho|, 2 - |\rho|\}$

II) Calculo la tasa de aceptación:

$$r = \exp \left( \log p(\rho^* \mid -) - \log p(\rho^{(b)} \mid -) \right)$$

III) Genero  $u \sim U(0, 1).$

Si  $u < r$  entonces  $\rho^{(b+1)} = \rho^{(b)} = \rho^*$ . Si no  $\rho^{(b+1)} = \rho^{(b)}$

3) Almacenar  $\beta^{(b+1)}, \sigma^{2(b+1)}$  y  $\rho^{(b+1)}$ .

#### 4.2.5. Convergencia de los muestreadores

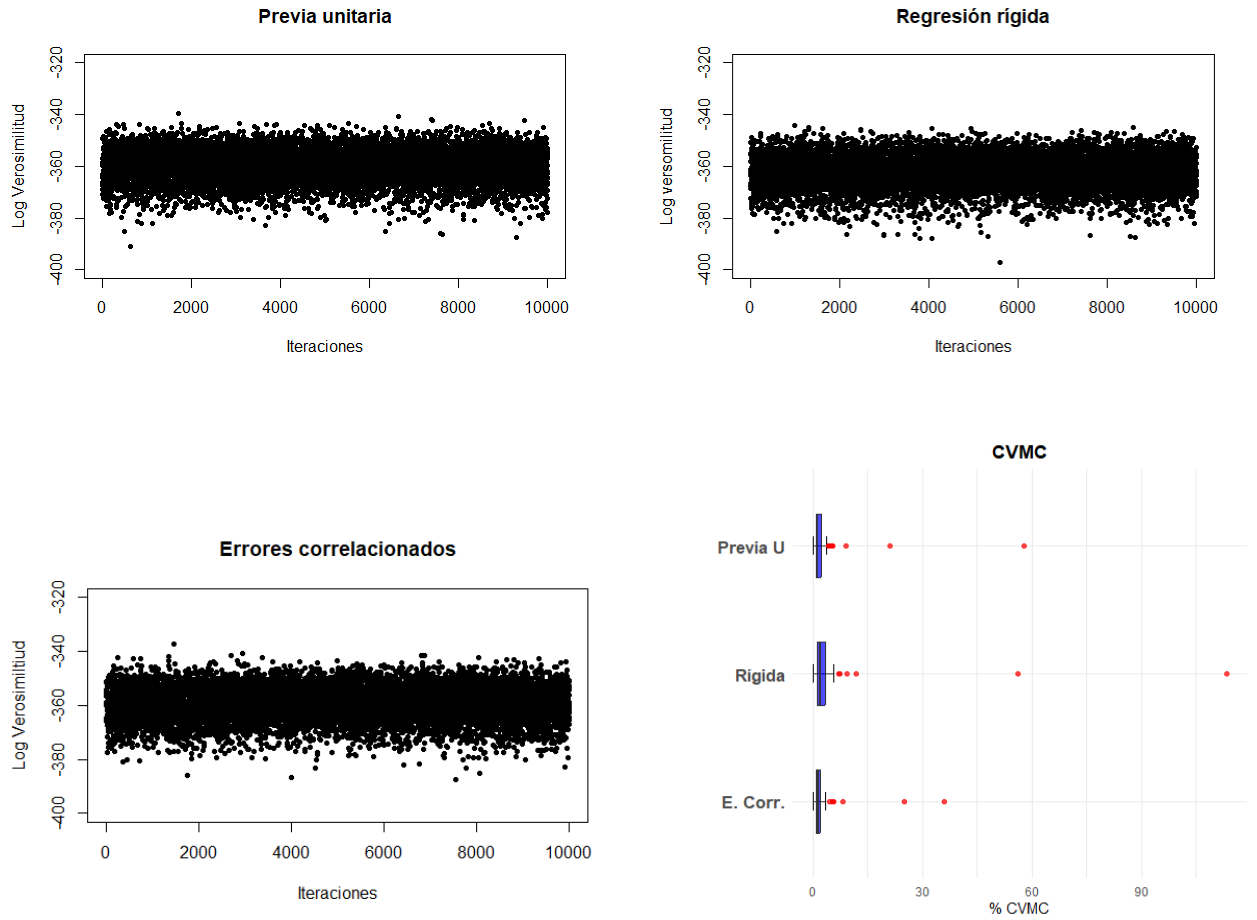


Figura 7: Convergencia de la MCMC

## Referencias

- Hoff, P. (2009). *A First Course in Bayesian Statistical Methods*. Springer.
- Invamer. (2023). Opinión Alcaldes 2023 [En línea; accedido el 14-November-2023]. <https://drive.google.com/drive/u/0/folders/1ttej1l3q-5LdX0Z8TPRorLMMEzxOXq->
- Restrepo, M. J. (2023). Así les fue a las encuestas a las encuestas en las elecciones de grandes alcaldías [En línea; accedido el 14-November-2023]. <https://www.lasillavacia.com/silla-nacional/asi-les-fue-a-las-encuestas-en-las-elecciones-de-grandes-alcaldias/>