

UNIVERSIDAD DE COSTA RICA  
SISTEMA DE ESTUDIOS DE POSTGRADO

EL DILEMA DE “SHRINKAGE” O FENÓMENO DE ENCOGIMIENTO EN MODELOS  
MIXTOS, UNA COMPARACIÓN ENTRE LOS ENFOQUES BAYESIANO Y  
FRECUENTISTA

Tesis sometida a la consideración de la Comisión del Programa de Estudios de Posgrado  
en Estadística para optar al grado y título de Maestría Académica en Estadística

ANDREA DE LOS ÁNGELES CÉSPEDES SOLÍS

Cuidad Universitaria Rodrigo Facio, Costa Rica

2021

## **Dedicatoria y agradecimientos**

Este trabajo se lo dedico a mi familia por ser mi apoyo incondicional en esta travesía y aventura universitaria, especialmente a mi mamá por ser mi pilar, confiar en mí, apoyarme, amarme en todo momento y ser parte de mis decisiones, y a quien es el amor de mi corazón, a mi tía Carmen, quien fue una segunda mamá y empezó acompañarme desde el cielo hace unos meses, espero estén orgullosas; a mi esposo que ha estado siempre ahí y es el amor de mi vida, a mi pequeña bendición, Horacio esto toma más fuerza por ti.

A todos los colegas que estuvieron anuentes a colaborarme ante las dudas que me surgieron en el camino, que aportaron a mi formación profesional en este proceso de aprendizaje, a mi tutora por darme un reto académico y ser parte de esto, que además de profesionalmente, me permitió crecer personalmente. Finalmente, a los todos los que de una u otra manera formaron parte de esto y me alentaron hasta su conclusión; son parte importante de mi corazón.

“Esta tesis fue aceptada por la Comisión del Programa de Estudios de Posgrado de la Universidad de Costa Rica, como requisito parcial para optar al grado y título de Maestría Académica en Estadística”

---

Dr. Ricardo Alvarado Barrantes  
**Representante del Decano  
Sistema de Estudios de Postgrado**

---

Dra. Eiliana Montero Rojas  
**Directora de Tesis**

---

Dr. Guaner Rojas Rojas  
**Lector**

---

Dr. Juan Carlos Brenes Sáenz  
**Lector**

---

M.Sc. Johnny Madrigal Pana  
**Director del Programa  
de Posgrado en Estadística**

---

Andrea de los Ángeles Céspedes Solís  
**Sustentante**

## Tabla de contenidos

Dedicatoria.....	ii
Agradecimientos.....	ii
Hoja de aprobación.....	iii
Tabla de contenidos.....	iv
Resumen.....	vi
Abstract.....	vii
Lista de cuadros.....	viii
Lista de gráficos.....	ix
Lista de imágenes.....	xi
1. Introducción.....	1
2. Justificación y objetivos.....	7
2.1 Objetivo general.....	9
2.2 Objetivos específicos.....	9
3. Marco teórico.....	9
4. Marco metodológico.....	10
4.1 Población bajo estudio.....	10
4.2 Análisis de datos originales.....	10
4.3 Análisis de datos en esta investigación.....	27
4.4 Técnicas a emplear.....	29
4.5 Modelos analizados.....	34
4.6 Alternativas de estimación.....	40
4.6.1 Regresión LASSO.....	40
4.6.2 Laplace.....	43
4.6.3 Matriz de dispersión.....	45
4.7 Estudio de simulaciones.....	48
4.7.1 Diseño del estudio de simulaciones.....	48
4.7.2 Generación y análisis de los datos.....	50
5. Análisis de datos y resultados.....	53

5.1	Análisis descriptivo.....	53
5.2	Modelos generados.....	57
5.3	Gráficas de coeficientes estimados.....	72
5.4	Del estudio de simulaciones.....	79
6.	Conclusiones.....	91
6.1	Conclusiones relevantes del caso de estudio.....	91
6.2	Hallazgos adicionales.....	96
	Bibliografía.....	99
	Anexos.....	106
	Anexo 1: Estadísticas descriptivas.....	106
	Anexo 2: Selección del modelo.....	109
	Anexo 3: Paquetes de R utilizados.....	111
3.1	glmer.....	111
3.2	MCMCGLMM.....	112
3.3	glmmLasso.....	115
3.4	glmmSr.....	116
3.5	hglm2.....	117
3.6	glmmTBM.....	119
	Anexo 4: Estimación adicional de modelos a partir del enfoque frecuentista y bayesiano.....	121
4.1	Estimación frecuentista con R.....	121
4.2	Estimación frecuentista y bayesiana con STATA 15.....	122
	Anexo 5: Selección del valor para Lambda en la regresión LASSO.....	130
5.1	Uso de Validación cruzada.....	130
5.2	Uso de BIC y AIC.....	131
5.3	Resumen de estimaciones.....	131
	Anexo 6: Estimación de coeficientes a partir del estudio de simulaciones.....	134
6.1	Estimación del SESGO.....	134
6.2	Estimación del CME.....	137

## Resumen

En los modelos mixtos (multinivel) la especificación matemática del modelo puede impactar de manera relevante en los resultados sustantivos obtenidos.

El fenómeno de *shrinkage* o fenómeno de encogimiento provoca que los efectos fijos comiencen a reducirse al punto de no presentar significancia estadística, diversos autores señalan como los principales causantes de este dilema analítico al aumento en la cantidad de parámetros, la cantidad de elementos en los niveles ( $n$ ), la varianza entre y dentro de los conglomerados, y el enfoque de estimación. En este estudio se presentaron dos propuestas de estimadores *shrinkage*, una correspondiente a la ponderación de los parámetros del modelo y otra ajustando las matrices de varianza y covarianza. A partir de estas dos se utilizaron los siguientes métodos de estimación LASSO, Laplace y ajuste por matriz de dispersión, además de las estimaciones frecuentista y bayesiana consideradas preliminarmente. Los datos utilizados correspondieron a experimentos realizados en el Centro de Investigación en Neurociencias de la Universidad de Costa Rica. Por la estructura de los datos, se utilizó un modelo lineal generalizado multinivel con distribución Poisson de enlace logaritmo, y de tres niveles. Así mismo, se realizó un estudio de simulaciones considerando un modelo base y distintos escenarios a partir de los posibles causantes del efecto *shrinkage*, con el fin de comparar las técnicas de estimación y determinar cuál es la mejor. El análisis indica que la estimación bayesiana es la que tiene mayor estabilidad en las estimaciones de los coeficientes, también se tiene que el tamaño de muestra no está relacionado con el efecto *shrinkage*, pero, la varianza, y la cantidad de coeficientes aleatorios sí.

En el caso de las estimaciones bayesianas se utilizaron 30 000 iteraciones con un *burn-in* de 10 000. Para el análisis, generación de bases de datos y estudio de simulación, se utiliza la plataforma de programación R 3.5.2 y el paquete STATA 15.

## **Abstract**

In mixed (multilevel) models, the mathematical specification of the model can have an important impact on the substantive results obtained.

The phenomenon of shrinkage causes the fixed effects to begin to reduce to the point of not presenting statistical significance, various authors point out as the main causes of this analytical dilemma to the increase in the number of parameters, the elements in the levels ( $n$ ), the variance between and within the clusters, and the estimation approach. In this study, they present two proposals for shrinkage estimators, one corresponding to the weighting of the model parameters and the other adjusting the variance and covariance matrices. From these two, the following LASSO, Laplace, and dispersion matrix adjustment methods are used, in addition to the frequenter and Bayesian estimates previously considered. The data used correspond to experiments carried out at the Neuroscience Research Center of the University of Costa Rica. Due to the structure of the data, a multilevel generalized linear model with a Poisson distribution of logarithmic link, and three levels, was used. Likewise, a simulation study is carried out considering a base model and different scenarios from the possible causes of the shrinkage effect, to compare the estimation techniques and determine which the best is.

The analysis indicates that the Bayesian estimate is the one with the greatest stability in the estimations of the coefficients; it is also found that the sample size is not related to the shrinkage effect, but the variance, and the number of random coefficients, is.

In the case of Bayesian estimates, 30,000 iterations with a burn-in of 10,000 are used. For analysis, database generation and simulation study, the R 3.5.2 programming platform, and the STATA 15 package are used.

## Lista de cuadros

Cuadro 1: Descripción de los escenarios aplicados en las simulaciones.....	50
Cuadro 2: Estadísticos generales de las vocalizaciones emitidas según día del experimento.....	57
Cuadro 3: Estimación frecuentista para los coeficientes de los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos.....	61
Cuadro 4: Estimación bayesiana para los coeficientes de los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos.....	65
Cuadro 5: Estadísticos de resumen para los modelos estimados con R según enfoque frecuentista y bayesiano.....	66
Cuadro 6: Estimación con regresión LASSO para los coeficientes en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos.....	68
Cuadro 7: Estimación con cadenas de Markov con aproximación <i>Laplace</i> para los coeficientes de los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos.....	70
Cuadro 8: Estimación con matriz dispersa para los coeficientes de los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos.....	71
Cuadro A1: Estadísticas generales de las vocalizaciones emitidas según sujetos del grupo control.....	107
Cuadro A2: Estadísticos generales de las vocalizaciones emitidas según sujetos del grupo tratamiento.....	108
Cuadro A3: Ranking de modelos seleccionados preliminarmente para el análisis.....	110

Cuadro A4: Estimación frecuentista de coeficientes para los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, utilizando el paquete <i>glmmTMB</i> .....	121
Cuadro A5: Estimación de coeficientes fijos para los modelos de efectos fijos, según valores de <i>Lambda</i> .....	131
Cuadro A6: Estimación de los errores estándar para los coeficientes fijos en los modelos de efectos fijos, según valores de <i>Lambda</i> .....	132
Cuadro A7: Estimación de coeficientes fijos para los modelos de efectos aleatorios completos, según valores de <i>Lambda</i> .....	132
Cuadro A8: Estimación de los errores estándar para los coeficientes fijos en los modelos de efectos aleatorios completos, según valores de <i>Lambda</i> .....	133

### **Lista de gráficos**

Gráfico 1: Vocalizaciones emitidas por ratas según grupo.....	53
Gráfico 2: Vocalizaciones emitidas por día del experimento.....	54
Gráfico 3: Vocalizaciones emitidas durante el experimento por ratas del grupo control.....	55
Gráfico 4: Vocalizaciones emitidas durante el experimento por ratas del grupo tratamiento.....	56
Gráfico 5: Coeficientes del intercepto en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación.....	73
Gráfico 6: Coeficientes de DIA2 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación.....	73

Gráfico 7: Coeficientes de DIA3 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación.....	74
Gráfico 8: Coeficientes de M1 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación.....	75
Gráfico 9: Coeficientes de M4 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación.....	75
Gráfico 10: Coeficientes de M5 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación.....	76
Gráfico 11: Coeficientes de M6 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación.....	76
Gráfico 12: Coeficientes de M9 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación.....	77
Gráfico 13: Coeficientes de M10 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación.....	78
Gráfico 14: Coeficientes de M11 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación.....	78
Gráfico A1: Cantidad de vocalizaciones emitidas por los sujetos del grupo control según día.....	106
Gráfico A2: Conteo de vocalizaciones emitidas por las ratas del grupo control durante el experimento.....	106
Gráfico A3: Cantidad de vocalizaciones emitidas por los sujetos del grupo tratamiento según día.....	107
Gráfico A4: Conteo de vocalizaciones emitidas por las ratas del grupo tratamiento durante el experimento.....	108

## Lista de imágenes

Imagen 1: Diagrama de la estructura de los datos.....	28
Imagen 2: Diagrama de la estructura de los datos analizada.....	37
Imagen 3: SESGO del coeficiente DIA3 en los modelos de coeficientes fijos y efectos aleatorios, según tipo de estimación, tamaño de muestra y varianza.....	81
Imagen 4: SESGO del coeficiente M1 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	82
Imagen 5: SESGO del coeficiente M5 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	83
Imagen 6: SESGO del coeficiente M10 en los modelos de coeficientes fijos y efectos aleatorios, según tipo de estimación, tamaño de muestra y varianza.....	84
Imagen A1: Estimación frecuentista de coeficientes para el modelo de efectos fijos utilizando un modelo lineal generalizado mixto.....	122
Imagen A2: Estimación frecuentista de coeficientes para el modelo de efectos aleatorios utilizando un modelo lineal generalizado mixto.....	123
Imagen A3: Estimación frecuentista de coeficientes para el modelo de efectos fijos utilizando un modelo multinivel de regresión Poisson.....	124
Imagen A4: Estimación frecuentista de coeficientes para el modelo de efectos aleatorios utilizando un modelo multinivel de regresión Poisson.....	125
Imagen A5: Estimación bayesiana de coeficientes para el modelo de efectos fijos utilizando un modelo multinivel con regresión Poisson.....	126

Imagen A6: Estimación bayesiana de coeficientes para el modelo de efectos aleatorios utilizando un modelo multinivel con regresión Poisson.....	128
Imagen A7: SESGO del coeficiente intercepto en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	134
Imagen A8: SESGO del coeficiente DIA2 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	134
Imagen A9: SESGO del coeficiente M4 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	135
Imagen A10: SESGO del coeficiente M6 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	135
Imagen A11: SESGO del coeficiente M9 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	136
Imagen A12: SESGO del coeficiente M11 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	136
Imagen A13: CME del coeficiente intercepto en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	137

Imagen A14: CME del coeficiente M1 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	137
Imagen A15: CME del coeficiente M4 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	138
Imagen A16: CME del coeficiente M5 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	138
Imagen A17: CME: coeficiente M6 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	139
Imagen A18: CME del coeficiente M9 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	139
Imagen A19: CME del coeficiente M10 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	140
Imagen A20: CME del coeficiente M11 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	140
Imagen A21: CME del coeficiente DIA2 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....	141

Imagen A22: CME del coeficiente DIA3 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza.....141



**Autorización para digitalización y comunicación pública de Trabajos Finales de Graduación del Sistema de Estudios de Posgrado en el Repositorio Institucional de la Universidad de Costa Rica.**

Yo, Andrea de los Ángeles Céspedes Solís, con cédula de identidad 206880819, en mi condición de autor del TFG titulado El dilema de "shrinkage" o fenómeno de encojamiento en modelos mixtos. una comparación entre los enfoques frecuentista y bayesiano

Autorizo a la Universidad de Costa Rica para digitalizar y hacer divulgación pública de forma gratuita de dicho TFG a través del Repositorio Institucional u otro medio electrónico, para ser puesto a disposición del público según lo que establezca el Sistema de Estudios de Posgrado. SI  NO

\*En caso de la negativa favor indicar el tiempo de restricción: \_\_\_\_\_ año (s).

Este Trabajo Final de Graduación será publicado en formato PDF, o en el formato que en el momento se establezca, de tal forma que el acceso al mismo sea libre, con el fin de permitir la consulta e impresión, pero no su modificación.

Manifiesto que mi Trabajo Final de Graduación fue debidamente subido al sistema digital Kerwá y su contenido corresponde al documento original que sirvió para la obtención de mi título, y que su información no infringe ni violenta ningún derecho a terceros. El TFG además cuenta con el visto bueno de mi Director (a) de Tesis o Tutor (a) y cumplió con lo establecido en la revisión del Formato por parte del Sistema de Estudios de Posgrado.

**INFORMACIÓN DEL ESTUDIANTE:**

Nombre Completo: Andrea de los Ángeles Céspedes Solís

Número de Carné: A91616 Número de cédula: 206880819

Correo Electrónico: andrea.cespedes.solis@gmail.com

Fecha: 14 de mayo del 2021 Número de teléfono: 84831030

Nombre del Director (a) de Tesis o Tutor (a): Eiliana Rojas Montero

**FIRMA ESTUDIANTE**

Nota: El presente documento constituye una declaración jurada, cuyos alcances aseguran a la Universidad, que su contenido sea tomado como cierto. Su importancia radica en que permite abreviar procedimientos administrativos, y al mismo tiempo genera una responsabilidad legal para que quien declare contrario a la verdad de lo que manifiesta, puede como consecuencia, enfrentar un proceso penal por delito de perjurio, tipificado en el artículo 318 de nuestro Código Penal. Lo anterior implica que el estudiante se vea forzado a realizar su mayor esfuerzo para que no sólo incluya información veraz en la Licencia de Publicación, sino que también realice diligentemente la gestión de subir el documento correcto en la plataforma digital Kerwá.

## 1. Introducción

Los modelos de regresión son ampliamente utilizados en áreas del conocimiento como administración de empresas, economía, ingeniería, ciencias sociales, salud y biología, entre otras; este método de análisis genera para cada una de estas áreas un mayor conocimiento del fenómeno de interés y les brinda insumos para la toma de decisiones; sin embargo, la correcta aplicación de esos modelos implica una buena comprensión de la teoría que los sustenta, así como del contexto al cual se aplican (Kutner et al., 2005).

Los supuestos básicos de los modelos de regresión son: linealidad, independencia de los residuos, homocedasticidad, normalidad en la distribución de los residuos y la no colinealidad. A pesar de que el investigador cuenta en la mayoría de los casos con técnicas que le permiten realizar el análisis de los datos cuando no se cumple alguno de estos supuestos, debe considerar el uso de otro método de análisis ante el no cumplimiento de independencia entre las observaciones, ya que los modelos de regresión lineal no consideran estas semejanzas o disimilitudes que se gestan en la estructura natural de los datos (Gaviria & Castro, 2005).

La no independencia presente en los datos puede ser el reflejo de una compleja estructura de anidamiento; la cual es la principal particularidad de los modelos multinivel, y que los diferencia de los modelos de regresión lineal clásica (Gaviria & Castro, 2005).

Los modelos multinivel de dos niveles respetan la estructura jerárquica<sup>1</sup> de los datos ya que incorporan mecanismos de estimación que permiten realizar cálculos más precisos para los parámetros del modelo cuando se tienen distintos niveles. También permiten ventajas analíticas al incluir la interacción entre individuos de diferentes niveles, así como considerar más de un contexto y estimar posibles

---

<sup>1</sup> Dentro de la jerarquía se identifican los niveles, un modelo mixto debe presentar al menos dos niveles, en donde el primer nivel corresponde a la unidad mínima estudiada y los niveles siguientes contienen la información del nivel anterior. (Aparicio & Morera, 2007)

diferencias entre ellos (Gaviria & Castro, 2005). Por tanto, la no comprensión de la estructura de los datos e incorrecta especificación matemática del modelo puede impactar de manera relevante en los resultados sustantivos obtenidos, por lo cual es importante que el investigador conozca estos modelos y los utilice en sus análisis con el fin de evitar estimaciones erróneas, así como conclusiones erradas.

Dentro de la familia de los modelos multinivel se incluyen los modelos con datos longitudinales, en los cuales los conglomerados corresponden al sujeto de análisis y las unidades que conforman el conglomerado son las mediciones realizadas al sujeto a través de tiempo (Congdon, 2003, pp. 136-137). Por otra parte, los modelos de medidas repetidas hacen referencia a datos en los cuales el sujeto representa el conglomerado y las mediciones realizadas a este corresponden a las unidades del conglomerado, pero incluyendo la particularidad de que las condiciones de observación o del experimento puede someterse a un cambio (West et al., 2007).

Algunas ramas del conocimiento en las cuales las estructuras jerárquicas predominan en los datos son, por ejemplo: la educación, en la cual los centros educativos son los conglomerados y los estudiantes son los sujetos dentro del conglomerado. En el área de la salud, por ejemplo, existen dos escenarios: el primero considera a los centros de salud como posibles conglomerados conformados por los individuos y el segundo hace referencia al mismo sujeto como conglomerado a quien se le realizan varias mediciones. En el área psicológica los modelos mixtos son utilizados en los análisis conductuales de sujetos. En agronomía se utilizan en el análisis de tratamientos en cultivos, así como la efectividad de plaguicidas o herbicidas, entre otros; sin embargo, debido a la complejidad de los problemas a resolver, y a la necesidad de contar con programas estadísticos que incorporen los algoritmos necesarios, el uso de los modelos multinivel se da siempre y cuando los programas computacionales respondan a las necesidades del analista (Catalán et al, 2003, citado por Aparicio & Morera, 2007).

Además, gracias a las mejoras en los *software* en las últimas décadas la utilización de estos modelos ha ido en aumento, y es a finales de la década de los ochentas

cuando se produce una gran cantidad de publicaciones relacionadas con los modelos multinivel (Gaviria & Castro, 2005, pág. 8).

A pesar de las ventajas analíticas que ofrecen modelos multinivel, también presentan algunas problemáticas, entre las que resalta la disyuntiva entre priorizar la estimación de los efectos fijos<sup>2</sup> o la estimación de los efectos aleatorios<sup>3</sup>. Esta prioridad debe responder a los objetivos de la investigación, así como a los intereses del investigador; sin embargo, resulta importante saber que en la estimación de los efectos fijos y aleatorios del modelo se tiene un fenómeno llamado “*shrinkage*” o factor de encogimiento, el cual tiene un papel importante en las estimaciones de los efectos fijos.

Si bien al realizar la comparación estadística entre el modelo con más parámetros versus el modelo con menos parámetros, la razón de verosimilitud arrojará resultados a favor del modelo más completo, pero debido a que la significancia estadística no implica importancia práctica, es necesario buscar otras alternativas para decidir cuál es el modelo más adecuado; también resulta importante conocer otros posibles métodos en la estimación del factor de encogimiento o de los coeficientes del modelo, con el fin de no limitar al investigador a un número determinado de coeficientes aleatorios o fijos.

La problemática inicia, en términos matemáticos, con la presencia de correlación entre las variables predictoras. Algunas consecuencias de la colinealidad son predicciones poco fiables ya que la influencia de cada predictor se encuentra solapada por la relación con otros predictores, también se tendrían asociados “grandes” errores estándar y valores  $t$  no significativos; además, los coeficientes serán sensibles a cambios en los datos (Gujarati & Porter, 2010, pág. 327.). De esta

---

<sup>2</sup> Gaviria & Castro (2005) “Los parámetros fijos corresponden a los efectos medios en la población.” (pág. 59)

<sup>3</sup> De la Cruz (2008) menciona que los efectos aleatorios se traducen en un modelo de coeficientes aleatorios que van a tomar en cuenta la variabilidad entre agrupamientos, que a su vez pueden tomar formas simples como a través de variabilidad a nivel del intercepto, o de formas más complejas, a través de variabilidades entre niveles o contextos. (pág. 3)

manera, al especificar los efectos aleatorios en el modelo, la estimación de los efectos fijos podría reducirse al punto de no presentar significancia estadística Gaviria & Castro (2005).

Por ello, López-González (1998), Soto-Liria et al. (2000) proponen algunas alternativas para solventar este problema, entre ellas: 1- eliminar las variables problemáticas (con ello se puede cometer sesgo de especificación del modelo), 2- aumentar el tamaño de  $n$  (no siempre es posible), 3- realizar transformaciones en las variables (la interpretación de los resultados obtenidos puede ser compleja) y finalmente, 4- incorporar información externa a los datos originales, en este punto se hace referencia a un ajuste en el cálculo del coeficiente para disminuir la colinealidad de las variables mediante los llamados factores *shrinkage* o estimadores por contracción, que buscan reducir el error cuadrático medio para corregir la colinealidad. En el presente documento, se analizará el último punto señalado.

Se parte de que los principales causantes del fenómeno *shrinkage* en las estimaciones son: un aumento en la cantidad de parámetros del modelo (Royle & Link, 2002), la cantidad de elementos en los niveles, así como la varianza entre y dentro de los conglomerados (Gaviria & Castro, 2005), y que esto ocurre de manera natural tanto en estimaciones de enfoque frecuentista como bayesiano.

Existen métodos de estimación por contracción, así como de selección de variables y estimación por contracción que realizan de manera simultánea. Estos tienen un gran potencial para muchos problemas, especialmente aquellos que involucran pequeños tamaños de muestra y / o modelos altamente parametrizados (Royle-Link, 2002).

El factor *shrinkage* puede ser incorporado en la estimación de los coeficientes del modelo, de las siguientes maneras:

López (1998) indica que mediante una matriz definida positiva como en el caso de la regresión *ridge* en los modelos de regresión lineal<sup>4</sup>, el estimador contraído es incorporado en el cálculo de una matriz definida positiva mediante una constante  $k$ , dicha modificación consiste en añadir pequeñas cantidades positivas a la diagonal de la matriz  $X'X$ <sup>5</sup>, los estimadores resultantes son sesgados, pero ya que su media cuadrática es menor, son más estables por lo que no se verán afectados por las variaciones de los datos. En este caso, una de las principales ventajas del uso de estimadores *ridge* es que con un  $k$  lo suficientemente pequeño se logra reducir los efectos de la colinealidad y el estimador estaría más próximo al verdadero estimador mínimo cuadrático.

Por otra parte, se puede incluir un ponderador  $\lambda$  como lo muestra Gaviria & Castro (2005, pág. 85). Este ponderador se denomina parámetro *shrinkage* o factor de encogimiento, su cálculo corresponde a la razón que involucra la varianza de las medias de los conglomerados, y la varianza entre los conglomerados, tal como se presenta a continuación:

$$\lambda_i = \frac{\sigma_{\mu_0}^2}{\sigma_{\mu_0}^2 + \frac{\sigma_{\varepsilon}^2}{n_j}}$$

Este ponderador puede tomar valores de cero a uno siendo estos el mínimo valor posible y máximo valor posible respectivamente<sup>6</sup>. Otro ponderador utilizado en los modelos multinivel es el llamado *James-Steins*, el cual incorpora un componente

---

<sup>4</sup> La ecuación correspondiente a la estimación del parámetro en una regresión lineal es  $\hat{\beta} = (X'X)^{-1}X'Y$ .

<sup>5</sup> La ecuación de regresión con la respectiva modificación, corresponde a  $\hat{\beta} = (X'X + KQ)^{-1}X'Y$

<sup>6</sup> En el caso de  $\lambda_i = 1$ , se tienen dos escenarios: 1.  $\sigma_{\varepsilon}^2 = 0$  : en este caso los sujetos son iguales en el conglomerado. 2.  $n_j = \infty$  : la cantidad de sujetos dentro del conglomerado es infinita. Para  $\lambda_i = 0$ , se tienen dos escenarios: 1.  $\sigma_{\mu_0}^2 = 0$  : todos los conglomerados tiene la misma media. 2a.  $\sigma_{\varepsilon}^2 > \sigma_{\mu_0}^2$  : cuando las varianzas dentro de los conglomerados con mayores que las varianzas entre los conglomerados, los residuos medios de cada conglomerado no reflejan las diferencias del conglomerado respecto a la media general. 2b.  $n_j \rightarrow 0$  : si la cantidad de sujetos en el conglomerado es muy pequeña que tiende a cero, los residuos medios para estimar  $\sigma_{\mu_0}^2$  introducen mucho error. (Gaviria & Castro, 2005)

denominado  $\pi$ , que corresponde a un factor de contracción que toma valores entre cero y uno (Ejaz Ahmed et al., 2007). En la revisión de literatura se describen algunos estudios relacionados con los estimadores insesgados para modelos multinivel, propuestas para el análisis de las matrices de covarianza, así como modelos bayesianos jerárquicos y análisis de factor de *shrinkage* o de encogimiento de casos aplicados; la literatura consultada y presentada en este trabajo aporta conocimiento al tema de interés desde distintas aristas.

Finalmente, para las estimaciones de los modelos multinivel se utiliza la plataforma de programación R 3.5.2 y el programa estadístico STATA 15.

## 2. Justificación y objetivos

El objetivo del presente trabajo es comparar los enfoques bayesiano y frecuentista al estimar modelos multinivel, cuando el investigador se enfrenta al *shrinkage* o fenómeno de encogimiento. Asimismo, se desea identificar posibles alternativas a la estimación de los coeficientes en los modelos multinivel cuando se tiene este dilema analítico.

El fenómeno del *shrinkage* o encogimiento ocurre cuando en un modelo multinivel se incorpora una mayor cantidad de coeficientes aleatorios en el análisis, de modo que a mayor cantidad de coeficientes aleatorios mayor encogimiento o disminución en las estimaciones de los coeficientes fijos del modelo, de esta manera, el investigador se enfrenta ante la disyuntiva de decidir la cantidad determinada de coeficientes en el análisis o escoger variables importantes a incluir en el modelo, y de esta manera no afectar las estimaciones.

El tema toma relevancia ya que en este caso específico se pretende aplicar en el área de la neurociencia, sentando las bases para estudios más sofisticados, pues en el país se ha trabajado poco este tema, y los estudios previos se han limitado al análisis de varianza (ANOVA).

Por otra parte, el uso de los modelos multinivel adquiere protagonismo por su particularidad de considerar la estructura jerárquica de los datos para las estimaciones; por ejemplo, estos son muy utilizados en áreas como la educación, medicina y agronomía, entre otras; y, en el caso puntual de los datos recabados en neurociencias, la naturaleza de los mismos involucra una jerarquía que debe ser respetada.

Aunado a lo anterior se incorpora en el análisis a los modelos lineales generalizados multinivel, debido a que la variable de respuesta no es normal, por lo que se requieren métodos más flexibles en términos de supuestos y que generen resultados certeros.

Además, se realiza un estudio de simulaciones considerando un modelo base que se especifica en el apartado de aspectos metodológicos, asociado a distintos escenarios – en los cuales la teórica indica la posible presencia de *shrinkage* en las estimaciones – de análisis que buscarán comparar las estimaciones frecuentistas y bayesianas obtenidas de los coeficientes, además de las alternativas propuestas y con ello determinar el mejor método de ser posible, para solventar la problemática del *shrinkage* en el modelo multinivel de interés.

El estudio de simulaciones resulta útil en la generación de distintos escenarios, lo que permite al investigador comprender más el fenómeno de encogimiento; también, se considera relevante pues permite identificar y analizar los posibles cambios que se generan en las estimaciones de los efectos fijos al incrementar la cantidad de efectos aleatorios, tamaños de muestra, así como la varianza de los datos, y realizando la comparación entre los métodos frecuentista y bayesiano.

Para responder al objetivo de análisis planteado se analizarán los datos correspondientes a varios experimentos neurocientíficos realizados en el Centro de Investigación en Neurociencias de la Universidad de Costa Rica, y aplicado a 31 ratas macho de cepa Wistar, en donde se mide el número de vocalizaciones emitidas en momentos específicos durante un periodo de 10 minutos, por tres días consecutivos. Por la estructura de los datos, se plantea utilizar un modelo de medidas repetidas de tres niveles incorporando métodos de análisis con enfoques frecuentista y bayesiano.

Los objetivos planteados en la presente investigación se presentan a continuación.

## **2.1 Objetivo General**

Comparar los enfoques bayesiano y frecuentista para determinar el modelo multinivel más adecuado enfrentando el problema de *shrinkage* o fenómeno de encogimiento.

## **2.2 Objetivos específicos**

1. Presentar estrategias estadísticas y no estadísticas para enfrentar el problema de *shrinkage* o fenómeno de encogimiento en los enfoques frecuentista y bayesiano en los distintos modelos, y seleccionar aquellas que se consideren más pertinentes, más allá de la discusión en torno a la significancia estadística.
2. Aplicar los métodos de estimación LASSO, Laplace y matriz de dispersión, para enfrentar el problema de *shrinkage* en modelos mixtos de datos reales provenientes de los experimentos 723-B7-610 "Motivación y plasticidad neuroconductual: Efectos cognitivos, emocionales y sociales del enriquecimiento ambiental aleatorio e impredecible en ratas" y 742-B6-291 "Cambios en la plasticidad neuronal relacionados con la incubación de la sensibilidad por la amfetamina: evaluación molecular y estructural".
3. Comparar los métodos de estimación propuestos, así como la frecuentista y bayesiana para enfrentar el *shrinkage* a partir de escenarios simulados, e identificar el mejor método de estimación.

### 3. Marco teórico

A continuación, se presenta el análisis bibliográfico, el cual describe técnicas novedosas a considerar durante el desarrollo de la investigación, así como aplicaciones de los enfoques frecuentista y bayesiano.

Yi & Ma (2012) realizan un estudio para identificar los marcadores genéticos significativamente asociados con el estado de supervivencia de los ratones, así como estimar los efectos genéticos de estos marcadores. Los datos utilizados corresponden a 116 ratones de un cruce entre cepas (BALB/cByJ y C57BL/6ByJ), además, los sujetos fueron genotipificados en 133 marcadores genéticos que abarcan 20 cromosomas. Cada uno fue infectado con "*Listeria monocytogenes*", se tiene que el 30% de la población sobrevive hasta el final del experimento.

La particularidad de estos análisis radica en que se cuenta con muchas variables predictoras, las cuales pueden ser agrupadas de manera natural para facilitar el análisis, así mismo, esto puede provocar grupos de variables altamente correlacionadas; por lo que en el análisis es ideal incorporar una estructura jerárquica en las variables de predicción.

En este caso se propusieron dos métodos de penalización: LASSO adaptativo bayesiano y modelos jerárquicos bayesianos con dos distribuciones previas para los parámetros de varianza considerando una distribución Cauchy – que incorpora un parámetro que controla la contracción en las estimaciones de varianza –, y la distribución doble exponencial – esta cuenta con un factor que controla el encogimiento de los coeficientes –.

Los principales resultados muestran que incorporar parámetros de ajuste en las distribuciones previas para las variables y grupos permite estimaciones confiables de los parámetros y aumentan la capacidad de determinar cuáles son las variables importantes. De ambas distribuciones sobresale la doble distribución exponencial, ya que tienen mayor capacidad de encontrar las variables importantes o significativas, además su tiempo de convergencia es menor.

En relación con los modelos multinivel, López et al. (2007) realizaron un análisis con datos longitudinales de registros del peso al nacimiento de ganado criollo sanmartiniano, las variables predictoras son el sexo de la cría (hembra o macho), número de partos de la madre, época de nacimiento (invierno o verano), edad y código del padre de la cría.

El modelo planteado tiene una estructura jerárquica de cuatro niveles: (1) categoría de la edad de la madre, (2) familia a la cual pertenece, (3) padre y (4) peso del individuo. Son las variables sexo, época y número de partos los parámetros fijos, además, la variable padre se considera como factor aleatorio, y las respectivas interacciones entre estas variables también son consideradas aleatorias.

Los resultados obtenidos en este estudio muestran que utilizar el Mejor Predictor Lineal Insesgado (BLUP – por sus siglas en inglés) para las estimaciones es muy útil en modelos mixtos jerárquicos, mas no así en modelos con estructuras longitudinales.

Por otra parte, Barnard et al. (2000) plantean la modelación de una matriz de covarianza en términos de sus desviaciones estándar  $S$ , donde  $S$  es un vector  $k * 1$  y la matriz de correlación  $R$ , donde  $R$  es una matriz  $k * k$ . Los autores suponen dos escenarios: (1) usar estimadores de coeficientes *shrinkage*, (2) usar un modelo de localización general (GLOM). La alternativa de analizar esta matriz resulta relevante, ya que la estimación de la matriz de covarianza puede ser complicada debido a su dimensionalidad y restricción de ser definida no negativa.

Los supuestos considerados en el análisis son: (1) la matriz de desviaciones  $S$  y la matriz de correlaciones  $R$  son independientes, (2) no se tiene mucha información previa, (3) la información previa con la que se cuenta intencionalmente refleja débil conocimiento de la matriz  $R$ . En relación con el uso de coeficientes *shrinkage*, se parte de que no se cuenta con información previa para los parámetros  $\beta_j$ . Además,  $\beta_j \sim i. i. d. N(\beta, \Sigma)$  donde  $\beta$  representa el vector de distribuciones previas que puede ser ajustado y  $\Sigma$  corresponde a la varianza, ambos fijos y conocidos.

Los autores utilizaron poca información previa considerando un registro independiente normal de cada una de las desviaciones estándar. Para el primer coeficiente, su valor verdadero en el modelo de regresión oscila entre 1 a 2.8 y la desviación estándar de la muestra es de 0.6, a partir de ello, escogieron las previas de la siguiente manera:

Entonces, para  $s_1$  consideraron:  $\log(s_1) \sim N(-4, 0.62)$ , en donde se tiene una creencia previa de que los valores del primer coeficiente en las diez regresiones están mucho más juntos de lo que realmente están. Para  $s_2$  se considera:  $\log(s_2) \sim N(0, 1)$ , por lo tanto, los valores están más dispersos que en  $s_1$ , finalmente para  $s_3$  se replica la distribución de  $s_2$ . Se espera que los resultados muestren una mayor contracción en las estimaciones para  $s_1$  y menor contracción en  $s_2$  y  $s_3$ .

Para el segundo escenario, Barnard et al. (2000) proponen una generalización del GLOM mediante la descomposición de la matriz de covarianza en sus correlaciones y desviación estándar. Este mecanismo permite una mayor flexibilidad al especificar la estructura de covarianza condicional de las variables continuas, y la estimación de contracción de las covarianzas bajo un modelo jerárquico bayesiano.

El GLOM utilizado va acorde a la propuesta de Lui y Rubin<sup>7</sup> (1998) citado por Barnard et al. (2000). En este caso se parte de la siguiente matriz de covarianzas:

$$V_i = \text{diag}(S_i) R_i \text{diag}(S_i)$$

---

<sup>7</sup> Estos autores presentan una extensión del modelo GLOM que permite la estimación de una matriz de covarianza de la siguiente manera:  $\Sigma_i = \lambda_i \Sigma$  en la cual, la media geométrica de las  $\lambda$  se establece en 1. Este modelo supone que las normales multivariantes dentro de cada celda tienen la misma forma elipsoidal, pero posiblemente diferentes tamaños. También permiten restricciones logarítmicas adicionales sobre los coeficientes de proporcionalidad  $\lambda_i$ , lo que permite poner más estructura en términos de covariables. Esta estimación presenta como ventaja una mayor flexibilidad en las estimaciones, manteniendo la cantidad de parámetros desconocidos en un tamaño manejable.

Donde  $S_i$  es el vector de desviaciones estándar y  $R_i$  la matriz de correlaciones, además,  $i = 1, \dots, c$  donde  $c$  es cada celda de la matriz. Para este caso se asumió que  $S_i$  no tiene restricciones y  $R_i = R$  para todo  $i$ . Para los cálculos y debido a la estructura de los datos se utilizó el muestreo de Gibbs.

Los autores concluyen que la técnica de modelar la matriz de covarianza en términos de las desviaciones estándar y la matriz de correlación es una estrategia común cuando la restricción para la matriz de correlación, de ser definida positiva, es fácil de tratar utilizando el muestreo de Gibbs.

Otra propuesta de análisis para las matrices es la descomposición de Cholesky<sup>8</sup>, mencionada por varios autores, entre ellos Pan y Huang (2014). En este caso se realiza la selección de efectos aleatorios en modelos lineales generalizados mixtos mediante la función de penalización por encogimiento o contracción, para ello plantean una reparametrización de la matriz de covarianza mediante la descomposición Cholesky, en la cual se agrega un término de penalización por contracción a la función de cuasi verosimilitud penalizada (PQL<sup>9</sup>) en los componentes de varianza.

Los métodos de penalización por contracción tienen la ventaja de seleccionar los efectos aleatorios “efectivos”, es decir los que contribuyen con información al modelo, eliminando las covariables que resultan innecesarias en términos informativos; y, para estimar los parámetros asociados, esta penalización es considerada una función de las varianzas de los efectos aleatorios. La principal ventaja de los métodos de penalización por contracción es que, para las variables

---

<sup>8</sup> La factorización o descomposición de Cholesky es una matriz simétrica definida positiva, la cual se puede desagregar como el producto de una matriz triangular inferior (denominada el triángulo de Cholesky) y la traspuesta de la matriz triangular inferior. (Mora-Escobar, H.,2011)

<sup>9</sup> La cuasiverosimilitud penalizada (PQL) es propuesta por Breslow y Clayton 1993 en “Approximate inference in generalized linear mixed models”. (Pan y Huang, 2014)

significativas al modelo, la penalización es baja o casi nula. Simultáneamente, ahorran tiempos en el cálculo, y son más estables en la estimación del modelo.

En relación con el cálculo, se mencionan las ventajas en la estimación PQL, además de funciones de encogimiento como la *LASSO*<sup>10</sup> y *SCAD*<sup>11</sup>. En ambos casos estas se encargan de penalizar a las variables no informativas del modelo, de modo que reducen a cero los coeficientes de regresión asociados a dichas variables, por ello las variables significativas se mantienen en el modelo con una leve contracción en la estimación. Los autores realizan simulaciones para evaluar el método propuesto considerando cuatro distribuciones: normal<sup>12</sup>, bernoulli<sup>13</sup>, binomial<sup>14</sup> y poisson<sup>15</sup>,

---

<sup>10</sup> Este método de estimación es denominado “Least Absolute *Shrinkage* and Selection Operator” (*LASSO* por sus siglas en inglés) y fue propuesto por Robert Tibshirani (1996).

<sup>11</sup> El método de estimación es denominado “Smoothly Clipped Absolute Deviation” (*SCAD* – por sus siglas en inglés) y fue propuesto por Fan y Li (2001). Por otra parte, puede considerarse como una modificación de *LASSO*, y tiene las propiedades como la continuidad, dispersión e imparcialidad.

<sup>12</sup> Este modelo supone que las variables independientes para los efectos fijos y aleatorios siguen un orden en el tiempo. Para cada estimación se hacen 100 simulaciones. Como principales resultados se tiene que un pequeño número de efectos aleatorios es seleccionado de manera incorrecta, en las simulaciones se da una selección efectiva de los efectos entre el 97% y 98%, y las estimaciones de los parámetros tienen un sesgo pequeño.

<sup>13</sup> Consideró un modelo similar al caso de la distribución normal, con  $n = 100$  y 10 covariables dependientes en el tiempo, las respuestas son binarias distribuidas con el correspondiente logit de probabilidad de éxito. En este caso la selección de componentes aleatorios baja en comparación con la distribución normal, lo que se interpreta como un ajuste insuficiente o excesivo.

<sup>14</sup> Los ajustes de los modelos son semejantes al caso Bernoulli, excepto que esta vez la variable de respuesta provienen de distribuciones binomiales  $B(m, p_{ij})$  donde  $m$  es el número de ensayos y  $p_{ij}$  es la probabilidad de éxito de la prueba. En este caso, el método propuesto para la selección de efectos aleatorios funciona muy bien en el caso de covariables dependientes en el tiempo. En términos de comparación entre dos casos discretos de datos, el análisis con distribución binomial supera al análisis de Bernoulli, lo cual era esperado por los autores, ya que los datos binomiales proporcionan más información.

<sup>15</sup> La configuración de parámetros generales y el mecanismo de generación de datos es similar a los casos de datos normales, bernoulli y binomiales con covariables dependientes en el tiempo. Sin embargo, las respuestas son generadas según las distribuciones de Poisson, y las covariables fueron generadas a partir de una distribución normal multivariante. En este caso los autores concluyen que el método de selección de efectos

considerando las variables independientes y dependientes en el tiempo, en este caso solo se presentan los casos de variables dependientes en el tiempo ya que es similar a la estructura de datos que se desea analizar.

Cai and Dunson (2006) también utilizan la descomposición de Cholesky al realizar un análisis para conocer cuáles son los factores que afectan la fecundidad de las mujeres, además determinar si estos varían de mujer a mujer.

Para ello proponen un modelo lineal generalizado mixto (GLMM) de enfoque bayesiano, en el cual se incorpora la descomposición de Cholesky a la problemática de selección de la matriz de covarianza de los efectos aleatorios, además calculan las probabilidades marginales posteriores de incluir cada predictor en los componentes de efectos fijos y aleatorios de manera separada.

Los datos utilizados por los autores, corresponden al tiempo de embarazo de un grupo de 427 mujeres entre 19 y 39 años de edad que estaban registradas en *California Department of Consumer Affairs* y que proporcionaron la información. La variable de respuesta es medida como un evento discreto de sobrevivencia<sup>16</sup>. Se cuenta también con información demográfica general de la mujer, datos reproductivos, historial de anticonceptivos y otros factores relacionados con la fertilidad.

Para evaluar el comportamiento del procedimiento se realizaron simulaciones mediante la integración numérica y aproximaciones cadenas de Markov vía Monte Carlo (MCMC – por sus siglas en inglés) y el muestreo de Gibbs, para ello consideraron un total de 100 individuos con seis observaciones cada uno, y utilizaron diferentes estructuras de covarianza de los efectos aleatorios en un modelo lineal generalizado mixto con funciones de enlace: identidad, logística y

---

aleatorios propuesto funciona muy bien para los datos de distribución de Poisson con covariables dependientes del tiempo.

<sup>16</sup> En el caso de que la mujer registre el 13avo ciclo menstrual, ese registro se censura en el análisis. Cai and Dunson (2006)

logarítmica. Un *burn-in* de 2000, y 20 000 iteraciones para las estimaciones, además de un análisis de sensibilidad con cinco diferentes “semillas” o valores de inicio.

De los principales resultados obtenidos en dicho estudio, los autores concluyen que los modelos bayesianos son una buena alternativa en el análisis, la varianza de efectos aleatorios pequeña da buenas aproximaciones y los métodos resultaron robustos en el análisis de sensibilidad. Sin embargo, se hace énfasis en que el enfoque expuesto puede resultar más útil en casos con menos de 20 predictores.

Chen (2003) realiza una reparametrización para obtener una estructura condicionalmente lineal al transformar los coeficientes aleatorios en coeficientes de regresión lineal para facilitar el uso de distribuciones previas conjugadas normales. La reparametrización propuesta consiste en seleccionar un subgrupo de elementos del vector de componentes aleatorio, además suponiendo que el componente de efectos fijos es conocido. Para los elementos del vector aleatorio no seleccionados se asume que tienen varianza 0. Se elige también una matriz diagonal no negativa y otra matriz triangular con 1's en las entradas diagonales, logrando la descomposición de la matriz de covarianza de los efectos aleatorios.

Los datos utilizados por Chen (2003) son de 12 centros de salud que participan en el Proyecto Perinatal Colaborativo (CPP) entre 1959 y 1966. Se desea conocer el desarrollo psicomotor de un niño de 8 meses de edad ante la exposición prenatal a *bifenilos policlorados* (PCB) y el desarrollo motor en niños pequeños, además de investigar si existen diferencias entre los centros de salud.

La variable de respuesta es el nivel de exposición que se mide del suero materno tomado en el tercer trimestre del embarazo, las variables independientes son triglicéridos del tercer trimestre y nivel de colesterol sérico, raza de la madre (blanca, negra), educación de la madre (primaria o menos, escuela secundaria, más de secundaria), el orden de nacimiento del niño (primero o no) y un indicador de si el niño alguna vez fue amamantado.

Para la estimación se utiliza un modelo jerárquico bayesiano, y son seleccionadas distribuciones previas que facilitan las especificaciones computacionales, además, se establecieron tres cadenas con valores iniciales distintos, considerando 50 000 iteraciones iniciales y 100 000 iteraciones para las estimaciones posteriores, y se realizó también un análisis de sensibilidad.

Los principales resultados indican que la parametrización en el modelo mixto facilitó la especificación y cálculo mediante el muestreo de Gibbs, obteniendo resultados robustos. Por otra parte, se unificó el problema de selección de variables y efectos aleatorios. Finalmente, el uso de distribuciones no normales, como la distribución *t* de Student en los términos de error, o efectos aleatorios, permitió obtener resultados robustos.

Huang et al. (2006) analiza la efectividad de cierto fármaco que reciben algunos sujetos para medir el comportamiento del virus VIH. El objetivo de la investigación corresponde por una parte a la estimación de ciertos parámetros a nivel del individuo, y a un análisis en datos longitudinales a largo plazo con datos desbalanceados. Los datos utilizados corresponden a la simulación de un ensayo clínico con 20 pacientes portadores del VIH que reciben tratamiento antiviral a largo plazo, las mediciones se realizan cada 25 días desde el día 0 hasta el día 200 del seguimiento.

Además, Huang et al. (2006) mencionan que los modelos comúnmente utilizados para este tipo de análisis tienen una limitante importante: la relación que se puede establecer entre la variable de respuesta y los factores independientes es a corto plazo. Además, se sabe qué variable dependiente tiene algunos patrones similares entre los pacientes aun conservando sus características individuales, por ello los modelos jerárquicos de efectos mixtos no lineales (NLME) parecen ser los más indicados. Para el análisis se utiliza un modelo de efectos mixtos no lineal bayesiano (BNLME) debido a que los programas computacionales no eran capaces de realizar las estimaciones de NLME.

Dentro de los parámetros de este modelo se incorpora la información de marcadores de fenotipo que permiten cuantificar ciertos fármacos en el individuo a través del tiempo, los parámetros de un modelo de exposición a fármacos, la resistencia del virus, la eficacia antiviral del fármaco y la interacción entre la información de las células susceptibles a la infección del VIH, de células infectadas y del virus libre.

La especificación del modelo se realiza en tres etapas: (1) se definió la variación dentro del conglomerado, (2) se definió la variación entre conglomerados, (3) se especificaron las distribuciones previas. Por otra parte, los valores de los hiperparámetros se obtienen de estudios anteriores o de literatura de referencia, en los casos donde se carece de información son asignadas distribuciones previas no informativas con varianzas grandes. Las estimaciones son realizadas mediante las Cadenas de Markov Vía Monte Carlo (MCMC) combinando los métodos de Gibbs y Metropolis-Hastings, se consideran 30 000 iteraciones iniciales y 120 000 para la estimación, además se realiza un análisis de sensibilidad.

Dentro de los principales resultados se tiene que los modelos proporcionaban un buen ajuste a los datos, las estimaciones de sesgo para la población y los parámetros individuales fueron muy pequeñas, por su parte las correspondientes desviaciones estándar de las estimaciones se consideraron razonables. Así mismo, el modelo propuesto ajustó los datos moderadamente bien para la mayoría de los sujetos en el estudio, en menos del 10% se tuvo patrones de respuesta inusuales a la carga viral, mediciones inexactas de la exposición a fármacos y/o adherencia para estos sujetos.

Por otra parte, Kinney y Dunson (2006), realizan un análisis para la selección de efectos aleatorios en un modelo mixto logístico mediante la implementación del enfoque bayesiano con el muestreo de Gibbs, abordan principalmente dos problemáticas: definición de las distribuciones previas subjetivas, así como la ineficiencia computacional pues se tiene lentitud en las estimaciones.

Los autores se centran en “*Stochastic Search Variable Selection*” (SSVS) iniciando con el modelo que contiene todas las variables, buscando los modelos con alta probabilidad posterior, y de esta manera se logra identificar el subconjunto de predictores que tienen coeficientes no nulos o varianzas diferentes de cero para aplicar el muestreador de Gibbs. También se considera la descomposición de Cholesky para la especificación de las matrices a utilizar en las estimaciones.

Las distribuciones previas para los componentes de los efectos fijos son Bernoulli y Gamma, y en los componentes aleatorios las distribuciones son similares. Por otra parte, se menciona como posibles alternativas para el análisis los modelos logísticos de datos ordinales y probit.

La técnica se ilustra con el análisis de datos de epidemiología y datos simulados. Los datos pertenecen al Proyecto Perinatal Colaborativo (CPP), la recolección de la información se realizó entre 1959 y 1966 con un total de 5389 participantes. Se mide el efecto del DDE (Dicloro Difenil Dicloroetileno), un metabolito del DDT (Dicloro Difenil Tricloroetano), el cual es medido en el suero materno luego de una pérdida de embarazo y es considerada una variable binaria. Las covariables son la edad de la madre, el índice de masa corporal, el tabaquismo, los niveles de colesterol y triglicéridos. Para la estimación de efectos aleatorios se considera que los 12 centros de los cuales se obtuvo la información son heterogéneos entre sí.

El objetivo es seleccionar un modelo logístico de efectos mixtos que relacione los niveles de DDE y la pérdida de embarazo, en el cálculo de las estimaciones se consideran 30 000 iteraciones y un *burn-in* de 5000 iteraciones.

En la simulación se considera un modelo logístico de respuesta binaria con tres covariables de distribución uniforme para 30 observaciones de 200 sujetos, además se eligió un rango de valores posibles o realistas para las varianzas de efectos aleatorios y se realizó un análisis de sensibilidad.

Los resultados obtenidos son comparados con el enfoque de cuasi verosimilitud penalizada, ya que es muy utilizado en las estimaciones de GLMM. El enfoque

propuesto da ventajas en la selección de los efectos fijos y aleatorios pues lo realiza al mismo tiempo. Por su parte el modelo bayesiano permitió el desarrollo computacional para el cálculo de las probabilidades marginales, lo cual es de utilidad en el muestreo de Gibbs. En el análisis de sensibilidad, se obtuvo resultados robustos con poca diferencia en las estimaciones de los parámetros.

También, Bondell et al. (2010) realizan una simulación para identificar de forma simultánea los principales predictores para los componentes fijos y aleatorios en un modelo mixto. Para el análisis se usan los datos de la Red de Tendencias y Estado del Aire Limpio (CASTNet) que contiene el registro de los lugares de Estados Unidos con medidas repetidas de contaminación y variables meteorológicas. Estos datos permiten realizar simulaciones más precisas de la calidad del aire.

La propuesta consiste en realizar la selección simultánea de factores fijos y aleatorios utilizando una descomposición de Cholesky modificada, mediante una reparametrización de los modelos lineales mixtos. Para el análisis se parte del supuesto de distribución normal en la distribución condicional y la distribución de los efectos aleatorios, en caso de que este supuesto no se cumpla, se puede ocurrir falta de robustez. Para las penalizaciones se propone el método *LASSO* ajustado.

En la selección del parámetro de ajuste se considera el algoritmo de *Expectation-Maximization*<sup>17</sup> (EM) en el cual se supone que los efectos aleatorios no son observados y se le asigna un valor fijo al parámetro de regularización no negativo utilizado para estimar el modelo de regresión lineal *LASSO* ajustada.

Los resultados obtenidos de la simulación muestran que el rendimiento de los métodos típicos que seleccionan separadamente los componentes fijo y aleatorio,

---

<sup>17</sup> Utilizado por Laird y Ware en “*Random-effects models for longitudinal data*”, 1982 y Laird, Lange y Stram en “*Maximum likelihood computations with repeated measures: application of the EM algorithm*”, 1987. (Bondell et al., 2010)

como el criterio BIC o GIC<sup>18</sup> extendido, no es tan bueno como el método propuesto que selecciona simultáneamente los componentes fijos y aleatorios.

Se tiene también el caso de un ensayo clínico con datos hipotéticos, el cual tiene como objetivo analizar un nuevo medicamento para combatir la depresión. Los intereses de la investigación son enunciados a continuación: 1) Lograr evidencia de la efectividad del medicamento. 2) Establecer si el efecto del medicamento se maximiza o minimiza con el paso del tiempo. 3) Establecer si hay evidencias para afirmar que el efecto del medicamento tenderá a ser el mismo en todos los individuos con estos mismos criterios de inclusión.

Para ello se cuenta con un grupo de 24 sujetos, los cuales son asignados aleatoriamente a un grupo de control (12) y un grupo de tratamiento (12). Los sujetos del grupo de control reciben un placebo y los demás reciben el medicamento. Se realizan cuatro mediciones, utilizando una escala, la primera medición corresponde a la línea base, es decir antes de administrar el tratamiento, las demás mediciones se realizan 1 mes, 3 meses y 6 meses después de la línea base.

En análisis se realiza en dos escenarios, en el primero se considera la base de datos completa y en el segundo son eliminados algunos de los valores para realizar las estimaciones con valores faltantes, en este último caso se tienen distintas razones por las cuales los valores fueron eliminados.

Se utilizan tres modelos multinivel, el primero corresponde a un modelo de efectos fijos, el segundo modelo incorpora efectos fijos y aleatorios, y el tercer modelo incorpora además de los efectos fijos y aleatorios, las correlaciones entre los efectos aleatorios.

Los principales resultados indican que para la variable grupo de tratamiento la estimación del parámetro cambia a medida que se incorporan efectos aleatorios en

---

<sup>18</sup> El GIC es un criterio de selección similar a los criterios AIC, BIC y AIC condicional, es mencionado por Rao y Wu en "A strongly consistent procedure for model selection in regression problems", 1989. (Bondell et al., 2010)

los modelos, la estimación de la constante se mantiene similar en los tres modelos. Además, entre el segundo y tercer modelo estimado, se presentan cambios en los parámetros correspondientes a la variable tiempo y en los errores estándar de los parámetros fijos del modelo; en el caso del tercer modelo que incorpora la estimación de las covarianzas que pierde significancia estadística.

Ntzoufras y Lykou (2013) proponen un método bayesiano para lograr la reducción y selección de variables con la regresión LASSO, el problema que se trata de resolver inicialmente es la especificación del valor del parámetro  $\lambda$  a través de los factores de Bayes y utilizando el método de MCMC para realizar las estimaciones. Se le asocia a este parámetro los valores de Pearson y la correlación parcial en los límites entre significación e insignificancia según lo definido por los factores de bayes, y se realiza una comparación entre la regresión LASSO ordinaria y la bayesiana.

Se realizan dos estudios de simulación y la aplicación del modelo a un conjunto de datos reales, además se considera el escenario en el cual se tiene un número de predictores menor al número de observaciones.

Para identificar fácil y de manera clara el valor del parámetro de contracción, los autores parten del nivel de correlación identificado entre en el límite entre las variables significativas e insignificantes para el factor Bayes, de la siguiente manera:

(1) Se traza un rango de valores de correlación entre las covariables y variable de respuesta, independientemente del valor de  $\lambda$ . Así a partir de las medidas de correlación de Pearson se tendrán factores de Bayes que medirán la evidencia a favor de la adición de esta covariable al modelo (inferior a uno para todos los valores de  $\lambda$ ). Se identifica también el rango de valores de correlación donde la inclusión de esta covariable nunca es apoyada *a posteriori* con la suficiente firmeza; es decir, el factor de Bayes correspondiente nunca llega a ser más alto que un nivel específico  $\alpha$  para todos los valores de  $\lambda$ .

(2). Se especifica el valor de  $\lambda$ , partiendo de los siguientes escenarios: (2.1) considerando a  $\lambda$  como constante, en este caso se especifica  $\lambda$  definiendo los

niveles de importancia práctica (es decir, cuando los factores de bayes son iguales a uno) según las medidas de correlación producidas en el paso 1 – o variable aleatoria –, (2.2) considerando a  $\lambda$  como aleatorio: se toma el rango de correlaciones producido especificando un hiperparámetro sensible para  $\lambda$ . Se tiene que, a partir de la especificación de los hiperparámetros del paso inicial, se evita el uso de distribuciones previas no informativas para el parámetro de contracción  $\lambda$ .

Los autores mencionan que conforme aumenta el  $\lambda$  la distribución se vuelve más informativa. Y, para valores pequeños de  $\lambda$  no se implementa la contracción, mientras que  $\lambda$  aumenta todos los coeficientes se reducen a cero.

Dentro de las distribuciones previas contempladas se tiene la doble exponencial, para cada coeficiente  $\beta_j$  del modelo, con ventajas en la selección de variables ya conocidas en el enfoque bayesiano.

Se realiza un análisis de sensibilidad en los factores Bayes en la elección de los parámetros de contracción cuando se realiza la regresión de LASSO múltiple. Se exploran las ventajas de los métodos de selección de variables y encogimiento. La contracción se logra mediante el uso de un producto de distribuciones previas dobles independientes, e independientes para los coeficientes de regresión. La selección de variables se logra a través de los indicadores de inclusión de variables binarias habituales incluidos en el predictor lineal. La estimación de las distribuciones posteriores (incluido el modelo posterior y las probabilidades de inclusión variable) se logra a través de un simple esquema MCMC.

Finalmente, los resultados se comparan con una variedad de métodos: LASSO bayesiano partiendo de la correlación, LASSO bayesiano con hiperparámetros gamma, logaritmo LASSO con selección del parámetro de contracción mediante validación cruzada, y otras como normal mixto con inversa gamma. Se concluyó que los métodos LASSO bayesianos funcionan eficazmente en todos los ejemplos, rastreando los efectos importantes con altas probabilidades posteriores y eliminando las covariables no importantes con bajas probabilidades posteriores.

## **4. Marco metodológico**

### **4.1 Población bajo estudio**

Los datos para el análisis son fuente primaria de información, correspondiente a un experimento en Neurociencias, asociado a los proyectos de investigación: 723-B7-610 "Motivación y plasticidad neuroconductual: Efectos cognitivos, emocionales y sociales del enriquecimiento ambiental aleatorio e impredecible en ratas" y 742-B6-291 "Cambios en la plasticidad neuronal relacionados con la incubación de la sensibilidad por la anfetamina: evaluación molecular y estructural", estos se realizaron en el Centro de Investigación en Neurociencias de la Universidad de Costa Rica, en ambos casos es el señor Juan Carlos Brenes Sáenz el investigador principal, por lo cual en caso de alguna consulta referida a los datos, él es a quien se debe dirigir.

La población bajo análisis comprende 31 ratas machos de la cepa Wistar. Cada rata tiene aproximadamente dos meses de edad y un peso entre 200 y 210 gramos. Los animales fueron albergados en un espacio previamente adaptado con alimento y agua, tienen ciclos de luz oscura entre las 7a.m. y 7p.m., así como una temperatura que oscila entre los 23°C y los 29°C.

### **4.2 Análisis de datos originales**

Los objetivos generales asociados a los experimentos de los cuales se desprende la información utilizada en esta investigación corresponden a:

- Estudiar procesos relacionados con la plasticidad neuronal subyacentes a la incubación de la sensibilidad conductual por exposición a anfetamina, para contribuir al conocimiento sobre los cambios plásticos inducidos por drogas psicoestimulantes.
- Determinar si el incremento en el valor motivacional del enriquecimiento Ambiental (EA) en ratas potencia los efectos estimulatorios del EA sobre el aprendizaje y la memoria, la interacción y comunicación social, la sensibilidad a la anfetamina y la expresión de genes asociados a la plasticidad neural.

Este modelo se desarrolla a partir de la evidencia de que la anfetamina puede producir síntomas similares a la manía (euforia e hiperactividad psicomotora) en individuos sanos y precipitar episodios o agravar la sintomatología en pacientes con trastorno bipolar (Smith y Daves 1977, Willson et al., 2005, citado por Pereira et al. 2013); y debido a que la hiperactividad, la euforia y verbosidad excesiva son tres de los síntomas típicos de la anfetamina en humanos y de los episodios de manía dentro del trastorno bipolar (Van Kammen y Murphy, 1975; Willson et al., 2005, citado por Pereira et al. 2013), se partió del supuesto de que las vocalizaciones ultrasónicas (USVs) inducidas por la anfetamina con una frecuencia promedio superior o igual a los 50 kilohertzios (kHz) pueden ser consideradas como un marcador conductual para modelar la manía en animales de laboratorio.

Los fármacos que se desean analizar corresponden al tamoxifeno, el litio y la miricitrina, de los cuales solo el litio es un medicamento clínicamente aprobado para tratar el trastorno bipolar, los otros dos están en proceso de validación y gran parte del estudio era mostrar los potenciales efectos antimaniáticos de dichos fármacos.

Durante el experimento se realizó un tamizaje midiendo la cantidad de USVs espontáneas emitidas por las ratas al ser colocadas individualmente en una caja con borucha fresca. A partir de esta información, los sujetos fueron asignados aleatoriamente a dos grupos, pero de manera balanceada para que la variabilidad poblacional de la tasa de USVs estuviera representada lo más homogéneamente posible entre de los grupos.

El análisis de datos se realizó mediante un análisis de varianza múltiple (MANOVA), considerando una significancia de  $p < 0.05$ . Se comparó cada sustancia suministrada con el respectivo grupo de control, es decir: SAL+SAL vs. SAL+AMP, LI+SAL vs. LI+AMP y TAM+SAL vs. TAM+AMP, y se analizaron las vocalizaciones emitidas, así como la actividad locomotora.

El experimento consistió en colocar a cada uno de los sujetos de manera independiente durante tres días consecutivos en un campo abierto (CA) de acrílico

de 40 cm x 40cm x 40 cm, con el piso cubierto de borucha, bajo luz roja y con un micrófono ultrasónico colocado a 45 cm sobre el suelo de la caja por un periodo de 10 minutos cada día. Las USVs se monitorearon con un micrófono ultrasónico (UltraSoundGate Condenser MicrophoneCM16) y gravados con el *software* Avisoft Recorder 2.7. Al finalizar el tiempo del test, posterior a la extracción de cada rata y antes de ingresar a la siguiente, se cambió la borucha del CA.

En la ejecución del experimento (considerando únicamente la aplicación de las sustancias salina y anfetamina): día 1, se realizó una línea base para tener la actividad locomotora y las USVs espontánea de los animales previo a los tratamientos. En el día 2 se les suministró 0.9% de solución salina (sub cutáneo) a todos los sujetos 15 min antes de colocarlos en el CA. En el día 3 los sujetos reciben dos inyecciones previo al test: la primera inyección se aplica 25 min antes del CA, en la que se les suministró a todos los sujetos solución salina; diez min después (15 minutos antes del CA) se aplicó la segunda inyección, a un subgrupo de los sujetos se les suministró salina (16 ratas) y al otro subgrupo anfetamina (2.5mg/kg) (15 ratas).

Además, se midió la cantidad de vocalizaciones emitidas según los cuatro subtipos<sup>19</sup> en cada minuto del experimento.

---

<sup>19</sup> Los subtipos de vocalizaciones son definidos por Pereira et al. (2014), de acuerdo al tono, frecuencia y forma. Dicha clasificación se realizó manualmente a partir de espectrogramas. Los tipos de vocalización son los siguientes:

- a. Flat calls (planas): cuando los cambios de frecuencia dentro de un único elemento de cada USV eran iguales o inferiores a 5-kHz, pero la diferencia en la frecuencia de inicio y final sí pudo ser superior a 5-kHz.
- b. Step-calls. (escalonadas): son USV de frecuencia modulada (FM) y ocurren cuando la vocalización flat tiene al menos un elemento plano corto superpuesto al inicio y/o al final de la misma.
- c. Trills (zigzag): son USV de FM con cambios de frecuencia superiores a 5-kHz o con dos o más cambios de frecuencia en direcciones opuestas
- d. Mixed calls (mixtas): son USV de FM que no se incluyeron en las categorías anteriores y que pueden tener uno o más componentes planos y/o zigzag superpuestos.

Finalmente, por medio del programa Avisoft Bioacustics se midió la frecuencia (en Hercios), intensidad (en decibeles) y duración promedio (en segundos) para cada vocalización.

#### ***4.3 Análisis de datos en esta investigación***

Para el presente análisis se debe aclarar que, por un tema de confidencialidad de los proyectos y la información recolectada en los mismos, solo se tuvo acceso a los datos provenientes de animales tratados con solución salina-salina y salina-anfetamina.

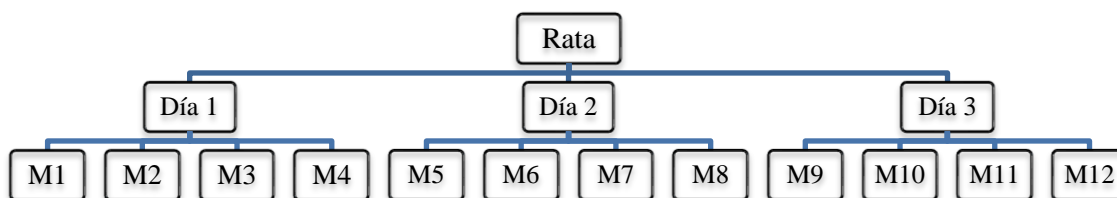
Además, no se incluye la variable tipo de vocalización, densidad ni duración de la misma, solamente la información correspondiente a la cantidad de vocalizaciones. Debido al interés del estudio, no se consideró la información de cada registro de vocalización, sino que se hizo un conteo de las vocalizaciones emitidas para cada uno de los sujetos en cada minuto de los diez que duró el experimento, por otra parte, por criterio y recomendación del experto, se utilizó la información de los minutos 1, 4, 7 y 10 del CA en los tres días.

Por lo tanto, para cada sujeto se tienen mediciones cada día en el que se desarrolló el experimento, y a su vez, para cada día del experimento, se tiene la información de los cuatro minutos seleccionados, es decir, cada rata es un conglomerado.

Así, se tiene una estructura de tres niveles, en la cual el primer nivel corresponde a las mediciones de las vocalizaciones en los minutos indicados anteriormente y se cuenta con un total de 372 registros. El segundo nivel corresponde a las mediciones en los tratamientos recibidos por los sujetos, en los tres días, para un total de 93 registros; finalmente, el tercer nivel corresponde a las 31 ratas o conglomerados.

La estructura de los datos se presenta en la imagen 1.

Imagen 1  
Diagrama de la estructura de los datos



Día 1: sin tratamiento

Día 2: solución salina

Día 3: solución salina o anfetamina

M1: medición de vocalizaciones en el minuto 1 – día 1

M2: medición de vocalizaciones en el minuto 4 – día 1

M3: medición de vocalizaciones en el minuto 7 – día 1

M4: medición de vocalizaciones en el minuto 10 – día 1

M5: medición de vocalizaciones en el minuto 1 – día 2

M6: medición de vocalizaciones en el minuto 4 – día 2

M7: medición de vocalizaciones en el minuto 7 – día 2

M8: medición de vocalizaciones en el minuto 10 – día 2

M9: medición de vocalizaciones en el minuto 1 – día 3

M10: medición de vocalizaciones en el minuto 4 – día 3

M11: medición de vocalizaciones en el minuto 7 – día 3

M12: medición de vocalizaciones en el minuto 10 – día 3

Se debe considerar que debido a que en el día tres del experimento los sujetos son asignados a distintos tratamientos – 16 reciben salina y 15 anfetamina – el diseño del conjunto de datos es desbalanceado en el tercer día.

Para las variables categóricas tratamiento y minuto se utiliza la notación de Gaviria & Castro (2005) para especificar el modelo mediante la codificación completa que emplea tantas variables dummies como categorías de la variable predictora.

A continuación, se detallan las variables a utilizar en el análisis:

Variable de respuesta:

- Número de vocalizaciones: cantidad de vocalizaciones emitidas.

Variables predictoras:

- Día del tratamiento: día 1 – sin tratamiento, día 2 – salina, día 3 – salina o anfetamina.
- Medición: en total se tienen 12 mediciones y corresponden al minuto en el cual se registró la variable de respuesta durante la ejecución el experimento 1, 4, 7 y 10, cada día.

#### ***4.4 Técnicas a emplear***

Para predecir la cantidad de vocalizaciones emitidas por las ratas, se utilizaron los modelos multinivel considerando estimaciones frecuentistas y bayesianas, en este caso puntual, el empleo de los modelos mixtos responde a la misma naturaleza jerárquica de los datos, pues las vocalizaciones son mediciones en distintos y determinados momentos del experimento, que a su vez pertenecen a una misma rata; en este caso específico corresponde a un modelo de medidas repetidas.

En relación con los modelos mixtos, estos son modelos estadísticos para variables de respuesta continua, y en los cuales los residuos se distribuyen normalmente. A diferencia de los modelos de regresión clásica pueden no tener varianza constante o no ser independientes. Se tiene que, en casos donde los estudios tienen datos agrupados como clúster, diseños experimentales, estudios longitudinales o de medidas repetidas, son recomendados estos modelos pues le brindan al investigador una herramienta más flexible y poderosa para el análisis de los datos (West et al., 2007).

En los modelos mixtos se pueden diferenciar dos tipos de parámetros, por una parte, están los fijos que corresponden a los efectos medios de la población y los aleatorios que corresponde a las varianzas y covarianzas de todos los niveles del modelo

(Gaviria & Castro, 2005, pág. 59). Estos permiten conocer diferencias entre los sujetos, ya sea considerando variables del conglomerado o de la unidad de análisis que conforma el conglomerado, por ende, omitir el uso de modelos mixtos bajo estas condiciones provocaría estimaciones erradas, tal como lo mencionan Gaviria & Castro (2005):

Al ignorar la estructura de los datos, el problema que se produce es que eliminamos toda la varianza interna de los grupos, que puede llegar a ser el 80% o del 90%. Así las relaciones parecerán como muy fuertes, y pueden ser de hecho muy distintas de los resultados con las variables desagregadas. (pág. 12)

También, los modelos multinivel resuelven el dilema de no independencia de los datos, que, al ignorar esta condición, puede mostrar resultados significativos cuando no lo son, por ende, las conclusiones estarían erradas. Los modelos mixtos: “constituyen una estrategia analítica que permite la formulación jerárquica de las fuentes de variación y con capacidad para dar cuenta de esta estructura” (Gaviria & Castro, 2005, pág. 14).

Finch et al. (2014, pág. 37) señalan como supuestos de los modelos multinivel de dos niveles los siguientes: 1. En el segundo nivel los residuales son independientes entre clústers. 2. En el segundo nivel los interceptos y coeficientes son asumidos como independientes de los errores de primer nivel. 3. En el primer nivel los residuales están distribuidos normalmente y tienen varianza constante. 4. En el nivel dos el intercepto y las pendientes deben tener una distribución normal multivariada con una matriz de covarianza constante.

Sin embargo, los modelos multinivel suponen que la distribución de la variable de respuesta es normal, por ende, para este caso de estudio al tener un variable de respuesta Poisson deben utilizarse los modelos lineales generalizados, los cuales permiten la estimación de modelos cuando en estos la variable de respuesta no tiene distribución normal.

Para la estimación frecuentista del modelo, se consideró la máxima verosimilitud restringida (RELM – por sus siglas en inglés) y no la conocida máxima verosimilitud (MLE – por sus siglas en inglés). Ello porque la RELM en el contexto multinivel incorpora un ajuste en los grados de libertad, ya que los considera en la estimación de los componentes de varianza, por ende, las estimaciones serán más apropiadas (Finch, 2014).

Por otra parte, en relación con el enfoque bayesiano, su principal característica es incorporar en las estimaciones información previa que involucra el conocimiento del experto, o puede obtenerse también de estudios preliminares que tengan relación con el tema de interés. Esta información en conjunto con la información que generan los datos, conforma la denominada distribución a posteriori de la estimación bayesiana. La distribución previa puede ser informativa o no informativa: la primera hace referencia a información sustentada en resultados de investigaciones que logran identificar atributos de variables de interés, la no informativa que hace referencia a las situaciones en las cuales no se conoce del fenómeno lo suficiente. (Carlin & Louis, 2000)

Las estimaciones del método bayesiano representan una alternativa a los métodos frecuentistas, se realizan mediante MCMC, el cual es un proceso iterativo donde la distribución a previa se combina con la información de la muestra actual y se genera la distribución *a posteriori* para cada uno de los parámetros del modelo. (Finch et al., 2014, pág. 169). Mediante un proceso repetitivo en el cual se actualiza la información en las cadenas de Markov, cuando cada cadena se estabiliza en un valor indicado se dice que la cadena converge y dicho valor corresponde a la estimación del parámetro en el modelo.

Además, las aproximaciones de las distribuciones *a posteriori* resultan más precisas en la implementación del método de MCMC (Barnard, McCulloch & Meng, 2000), ya que la estimación mejora con el número de iteraciones; también hay una mayor flexibilidad al comparar modelos con una cantidad diferente de componentes de

varianza. Pero, este enfoque implica una mayor carga computacional para el investigador (Kinney y Dunson, 2006).

Algunas de las ventajas de utilizar estadística bayesiana en la estimación de modelos multinivel son: (1). No supone una distribución de los datos, la estimación de los intervalos de credibilidad no se afecta por distribuciones asimétricas de los datos. (2). La estadística bayesiana puede ser de mucha utilidad cuando el modelo a estimar es muy complejo, pues en algunos casos las estimaciones frecuentistas pueden no lograr la convergencia. (3). Pueden estimar de manera precisa los parámetros con muestras pequeñas. (4). Las estimaciones bayesianas pueden ser utilizadas en los casos donde las estimaciones frecuentistas también funcionan. (Finch et al., 2014, pág. 168)

Por otra parte, se debe indicar que en el contexto de experimentos neurocientíficos los tamaños de muestra tienden a ser pequeños, lo cual podría desfavorecer las estimaciones en los modelos estadísticos frecuentistas, no así en los modelos de enfoque bayesiano.

Algunos ejemplos de estudios con tamaños de muestra reducidos son: Brenes et al. (2016) en "*Differential Effects of Social and Physical Environmental Enrichment on Brain Plasticity, Cognition, and Ultrasonic Communication in Rats*" realizan los análisis en 48 sujetos, por otra parte Brenes & Schwarting (2014) en su investigación "*Attribution and Expression of Incentive Salience Are Differentially Signaled by Ultrasonic Vocalizations in Rats*" tienen una población de estudio de 30 sujetos, otros casos aplicados que se relacionan con el tema son: "*The Rat Pup Study*" en Mixed-Effects Models in S and S-PLUS realizado por Pinheiro & Bates (2000) en el cual los datos fueron recolectados de 30 ratas, finalmente, Douglas et al (2004) en "*Pontine and basal forebrain cholinergic interaction: implications for sleep and breathing*" cuentan únicamente con cinco sujetos. Todos los anteriores relacionados en el área médica y con sujetos sometidos a análisis.

Respecto a la especificación de los modelos multinivel, el enfoque bayesiano es muy similar al enfoque frecuentista, los valores de las covarianzas y los errores pueden ser especificados en cada nivel y ser una fuente potencial acumulativa para la explicación del error en la variable de respuesta (Congdon, 2003, pág. 138). Para la estimación es importante identificar el tipo de variable dependiente en el análisis (continua, discreta, dicotómica, ordinal), pues a partir de ello se establecerán los lineamientos necesarios en la especificación del modelo que será utilizado. En este análisis, específicamente en la estimación bayesiana del modelo multinivel, se utilizaron previas no informativas.

En relación con el fenómeno de encogimiento en los coeficientes fijos de los modelos mixtos, en la literatura se conocen como estimadores *shrinkage* a los ponderadores o valores utilizados para ajustar las estimaciones de los coeficientes.

Así mismo, Royle & Link (2002) indican que los estimadores de contracción que surgen como consecuencia de la estimación (o predicción) de un mayor número de coeficientes aleatorios ya sea bajo consideraciones bayesianas o frecuentistas, lo que buscan es minimizar el error cuadrático medio de la predicción. También mencionan que, los coeficientes de contracción o encogimiento pueden mejorar la estimación de los coeficientes del modelo en los siguientes casos: donde se tenga un gran número de parámetros en el modelo, o donde se tiene tamaños de muestra pequeños.

Gaviria & Castro (2005) señalan que el tamaño del *shrinkage* – considerado como un ponderador – es determinado por la cantidad de elementos en los niveles, de esta manera “conforme aumenta el número de unidades del primer nivel ( $n_j$ ), este término tiende a uno. Del mismo modo, cuando el número de unidades individuales en cada contexto desciende, este factor tiende a cero.” (pp. 71-72).

Gaviria & Castro (2005) también mencionan que la varianza también es determinante para el factor *shrinkage*, pues si dentro de los conglomerados ésta es pequeña, es decir los sujetos son muy similares o hay muchos sujetos en el

conglomerado el factor *shrinkage* tiende a uno. Sin embargo, cuando los conglomerados tienen medias similares, las varianzas de los conglomerados son mayores que entre los conglomerados, es decir los grupos no son homogéneos, o la cantidad de sujetos es pequeña el factor de encogimiento tiende a cero.

#### **4.5 Modelos analizados**

Inicialmente se realizó un análisis descriptivo de los datos, así como un análisis preliminar del fenómeno de *shrinkage* considerando distintos escenarios mediante un estudio de simulaciones.

En el análisis descriptivo inicial, para conocer el comportamiento de los datos (anexo 1), se desagrega la variable cantidad de vocalizaciones por día y sustancia suministrada en el experimento, así como para cada sujeto. También, se creó una variable denominada “grupo” para identificar el comportamiento de los sujetos que recibieron anfetamina y los que no, y con ello generar insumos que permitieron determinar el mejor modelo que permita analizar los datos.

A partir de las estadísticas descriptivas y, considerando que la variable de respuesta es un conteo mayor o igual a cero con distribución asimétrica positiva, aunado a ello, debido a que la varianza es mayor a la media, es decir, se tiene sobre dispersión se decidió utilizar un modelo lineal generalizado con distribución QuasiPoisson y función de enlace logaritmo natural.

Los principales componentes de un modelo lineal generalizado son: (1) el componente sistemático es expresado como una relación lineal, (2) una función denominada función de enlace o *link* que indica la relación entre el valor esperado y el vector de covariables, (3) cuenta con un componente aleatorio que corresponde a la distribución condicional entre la variable de respuesta y el vector de covariables.

En este caso puntual se tiene que la función de verosimilitud de la Poisson corresponde a:

$$L(\mu) = f(\mu|X) = \frac{e^{-\mu} * \mu^x}{x!}$$

Aplicando el logaritmo y la función se tiene:

$$LL(\mu) = -\ln(x!) + \ln(e^{-\mu}) + \ln(\mu^x)$$

$$LL(\mu) = -\ln(x!) - \mu + x * \ln(\mu)$$

La función de enlace en el caso Poisson es  $\ln(\mu)$ . Es decir:

$$\ln(\mu) = \boldsymbol{\beta}'\mathbf{x}$$

$$\mu = \exp(\boldsymbol{\beta}'\mathbf{x})$$

Finamente se tendría que la log-verosimilitud del modelo es:

$$LL(\mu) = -\ln(x!) - \exp(\boldsymbol{\beta}'\mathbf{x}) + x * (\boldsymbol{\beta}'\mathbf{x})$$

Considerando todos los parámetros del modelo y de manera general, el predictor lineal del modelo puede ser detallado como sigue:

$$\eta_i = \beta_0 + \sum_{j=1}^{J_0} x_{ij}\beta_j + \sum_{k=1}^K \sum_{j \in G_k} z_{ij}\beta_j = \mathbf{X}_i\boldsymbol{\beta}$$

Donde  $\beta_0$  es el intercepto,  $x_{ij}$  y  $z_{ij}$  representan los valores observados para las variables no agrupadas y agrupadas respectivamente,  $\beta_j$  son los coeficientes del modelo. La notación  $j \in G_k$  indica el grupo de la variable  $j$ . Además,  $X_i$  contiene todas las variables,  $\boldsymbol{\beta}$  es el vector que contiene todos los coeficientes, incluyendo el intercepto.

De manera resumida, se tiene que  $X_i = (1, x_{i1}, x_{i2}, \dots, x_{iJ})$  y  $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_J)'$ , donde  $J = \sum_{g=0}^K J_k$  corresponde al total de variables.

La media de la variable de respuesta es relatada mediante la función de enlace  $g$ :

$$E(y_i|X_i) = g(X_i\beta)$$

Finalmente, la distribución de los datos puede ser representada de la siguiente manera:

$$p(y|X\beta, \emptyset) = \prod_{i=1}^n p(y_i|X_i\beta, \emptyset)$$

Donde  $\emptyset$  es el parámetro de dispersión, en el caso de las distribuciones de Poisson y Binomial, dicho parámetro toma valor 1.

Para la selección de las variables que se incluyen en el modelo, se parte del conocimiento del experto en el tema de neurociencias así como algunos requerimientos para conservar la estructura jerárquica de los datos (ver anexo 2), para el análisis se utilizó la plataforma de programación R 3.5.2, los paquetes *glmer*<sup>20</sup> Bates et al. (2015) para la estimación frecuentista y el paquete *MCMCglmm*<sup>21</sup> Jarrod D Hadfield (2010) en el caso de la estimación bayesiana, además del programa STATA 15. (Para mayor detalle ver anexo 3.1 y 3.2 respectivamente).

Para identificar la presencia de *shrinkage* en el modelo seleccionado, e iniciando con la estimación frecuentista, se estima un modelo de efectos fijos para las variables día y medición, seguido, se estima un modelo con tres coeficientes aleatorios<sup>22</sup> y finalmente otro modelo con los siete efectos aleatorios de las variables

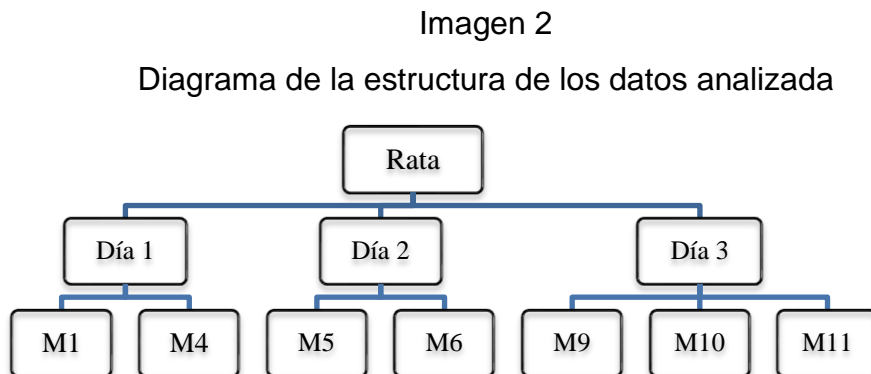
---

<sup>20</sup> El *glmer* se utiliza en la estimación frecuentista de Modelos Mixtos Lineales Generalizados, utiliza el algoritmo de mínimos cuadrados penalizados.

<sup>21</sup> Este paquete realiza la estimación en Modelos Lineales Generalizados Mixtos utilizando las técnicas de Cadenas de Markov Vía Monte Carlo, y utiliza como información previa a la distribución Wishart inversa para las (co)varianzas y una distribución normal para los efectos fijos. Utiliza un *burn-in* de 3000 iteraciones y 10 000 iteraciones para la estimación de los coeficientes.

<sup>22</sup> Los coeficientes aleatorios corresponden las variables M1, M5 y M11. Las variables M1 y M5 corresponden a las primeras mediciones que son registradas al iniciar el experimento el día 1 y 2 respectivamente, estas toman relevancia ya que corresponde al registro de las primeras reacciones de adaptación del sujeto al CA, así las más recientes posterior a la aplicación de la inyección; por otra parte, la variable M11 corresponde a mediciones realizadas el tercer día del experimento, la que también resulta valiosa en el análisis pues es la que

día y medición, esto se repite en la estimación bayesiana. La estructura de los datos corresponde a:



Donde se tiene que:

Día 1: sin tratamiento

Día 2: solución salina

Día 3: solución salina o anfetamina

M1: medición de vocalizaciones en el minuto 1 – día 1

M4: medición de vocalizaciones en el minuto 10 – día 1

M5: medición de vocalizaciones en el minuto 1 – día 2

M6: medición de vocalizaciones en el minuto 4 – día 2

M9: medición de vocalizaciones en el minuto 1 – día 3

M10: medición de vocalizaciones en el minuto 4 – día 3

M11: medición de vocalizaciones en el minuto 7 – día 3

Los modelos analizados son detallados a continuación:

### Modelo 1: de efectos fijos

---

refleja el comportamiento del sujeto luego del proceso de adaptación al CA así como de las inyecciones, de igual manera corresponde a mediciones en el día que se aplica el tratamiento o sustancia de interés.

$$y_{ijk} = \beta_{0jk} + \beta_{1jk}X_{1ijk} + \beta_{2jk}X_{2ijk} + \beta_{3jk}X_{3ijk} + \beta_{4jk}X_{4ijk} + \beta_{5jk}X_{5ijk} + \beta_{6jk}X_{6ijk} \\ + \beta_{7jk}X_{7ijk} + \beta_8X_{8jk} + \beta_9X_{9jk} + \varepsilon_{ijk}$$

**Modelo 2: incorpora coeficientes de efectos aleatorios para tres variables**

$$y_{ijk} = \beta_{0jk} + \beta_{1jk}X_{1ijk} + \beta_{2jk}X_{2ijk} + \beta_{3jk}X_{3ijk} + \beta_{4jk}X_{4ijk} + \beta_{5jk}X_{5ijk} + \beta_{6jk}X_{6ijk} \\ + \beta_{7jk}X_{7ijk} + \beta_8X_{8jk} + \beta_9X_{9jk} + \varepsilon_{ijk}$$

$$y_{ijk} = (\beta_0 + V_{0j} + V_{0jk}) + (\beta_1 + V_{1jk})X_{1ijk} + \beta_2X_{2ijk} + (\beta_3 + V_{3jk})X_{3ijk} + \beta_{4jk}X_{4ijk} \\ + \beta_{5jk}X_{5ijk} + \beta_{6jk}X_{6ijk} + (\beta_7 + V_{7jk})X_{7ijk} + \beta_8X_{8jk} + \beta_9X_{9jk} + \varepsilon_{ijk}$$

$$y_{ijk} = \beta_0 + \beta_1X_{1ijk} + \beta_2X_{2ijk} + \beta_3X_{3ijk} + \beta_{4jk}X_{4ijk} + \beta_{5jk}X_{5ijk} + \beta_{6jk}X_{6ijk} + \beta_7X_{7ijk} \\ + \beta_8X_{8jk} + \beta_9X_{9jk} + (V_{0j} + V_{0jk} + V_{1jk}X_{1ijk} + V_{3jk}X_{3ijk} + V_{7jk}X_{7ijk}) \\ + \varepsilon_{ijk}$$

Igualmente, en el segundo nivel se calcula la varianza correspondiente al intercepto  $\sigma_{V_0}^2$  únicamente, para el tercer nivel se presenta la estructura de la matriz de variancias y covariancias para los efectos aleatorios del modelo 2:

$$\begin{vmatrix} \sigma_{V_0}^2 & - & - & - \\ \sigma_{V_1V_0} & \sigma_{V_1}^2 & - & - \\ \sigma_{V_3V_0} & \sigma_{V_3V_1} & \sigma_{V_3}^2 & - \\ \sigma_{V_7V_0} & \sigma_{V_7V_1} & \sigma_{V_7V_3} & \sigma_{V_7}^2 \end{vmatrix}$$

**Modelo 3: completo de efectos aleatorios**

$$y_{ijk} = \beta_{0jk} + \beta_{1jk}X_{1ijk} + \beta_{2jk}X_{2ijk} + \beta_{3jk}X_{3ijk} + \beta_{4jk}X_{4ijk} + \beta_{5jk}X_{5ijk} + \beta_{6jk}X_{6ijk} \\ + \beta_{7jk}X_{7ijk} + \beta_8X_{8jk} + \beta_9X_{9jk} + \varepsilon_{ijk}$$

$$\begin{aligned}
y_{ijk} = & (\beta_0 + V_{0j} + V_{0jk}) + (\beta_1 + V_{1jk})X_{1ijk} + (\beta_2 + V_{2jk})X_{2ijk} + (\beta_3 + V_{3jk})X_{3ijk} \\
& + (\beta_4 + V_{4jk})X_{4ijk} + (\beta_5 + V_{5jk})X_{5ijk} + (\beta_6 + V_{6jk})X_{6ijk} \\
& + (\beta_7 + V_{7jk})X_{7ijk} + \beta_8 X_{8jk} + \beta_9 X_{9jk} + \varepsilon_{ijk}
\end{aligned}$$

$$\begin{aligned}
y_{ijk} = & \beta_0 + \beta_1 X_{1ijk} + \beta_2 X_{2ijk} + \beta_3 X_{3ijk} + \beta_4 X_{4ijk} + \beta_5 X_{5ijk} + \beta_6 X_{6ijk} + \beta_7 X_{7ijk} \\
& + \beta_8 X_{8jk} + \beta_9 X_{9jk} \\
& + (V_{0j} + V_{0jk} + V_{1jk}X_{1ijk} + V_{2jk}X_{2ijk} + V_{3jk}X_{3ijk} + V_{4jk}X_{4ijk} \\
& + V_{5jk}X_{5ijk} + V_{6jk}X_{6ijk} + V_{7jk}X_{7ijk}) + \varepsilon_{ijk}
\end{aligned}$$

Para el segundo nivel se calcula la varianza correspondiente al intercepto  $\sigma_{V_0}^2$  únicamente, y en el tercer nivel para el modelo completo la estructura de la matriz de variancias y covariancias para los efectos aleatorios corresponde a:

$$\begin{vmatrix}
\sigma_{V_0}^2 & - & - & - & - & - & - & - \\
\sigma_{V_1V_0} & \sigma_{V_1}^2 & - & - & - & - & - & - \\
\sigma_{V_2V_0} & \sigma_{V_2V_1} & \sigma_{V_2}^2 & - & - & - & - & - \\
\sigma_{V_3V_0} & \sigma_{V_3V_1} & \sigma_{V_3V_2} & \sigma_{V_3}^2 & - & - & - & - \\
\sigma_{V_4V_0} & \sigma_{V_4V_1} & \sigma_{V_4V_2} & \sigma_{V_4V_3} & \sigma_{V_4}^2 & - & - & - \\
\sigma_{V_5V_0} & \sigma_{V_5V_1} & \sigma_{V_5V_2} & \sigma_{V_5V_3} & \sigma_{V_5V_4} & \sigma_{V_5}^2 & - & - \\
\sigma_{V_6V_0} & \sigma_{V_6V_1} & \sigma_{V_6V_2} & \sigma_{V_6V_3} & \sigma_{V_6V_4} & \sigma_{V_6V_5} & \sigma_{V_6}^2 & - \\
\sigma_{V_7V_0} & \sigma_{V_7V_1} & \sigma_{V_7V_2} & \sigma_{V_7V_3} & \sigma_{V_7V_4} & \sigma_{V_7V_5} & \sigma_{V_7V_6} & \sigma_{V_7}^2
\end{vmatrix}$$

Donde:

$$\varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$$

$y_{ijk}$  = cantidad de vocalizaciones

$X_1$  = medición minuto 1 – día 1

$X_2$  = medición minuto 10 – día 1

$X_3$  = medición minuto 1 – día 2

$X_4$  = medición minuto 4 – día 2

$X_5$  = medición minuto 1 – día 3

$X_6$  = medición minuto 4 – día 3

$X_7$  = medición minuto 7 – día 3

$X_8$  = día 2

$X_9$  = día 3

## **4.6 Alternativas de estimación**

### 4.6.1 Regresión LASSO

La regresión de LASSO es una alternativa de estimación para los modelos en los cuales se desea minimizar los cuadrados mínimos.

La regresión LASSO además de realizar una ponderación, realiza la selección de las variables que matemáticamente aportan información al modelo, por ello en algunos casos los coeficientes de contracción pueden ser próximos o iguales a cero, ésta es la principal diferencia de la regresión en Cresta o *ridge* utilizada en modelo modelos de regresión más simples.

LASSO se encarga de realizar la respectiva ponderación a los coeficientes del modelo, mientras que en la regresión *ridge* los componentes de la varianza son iguales y se mantienen fijos en alguna constante. Tibshirani (2011) indica que en LASSO los componentes de la varianza se pueden distribuir de forma adaptativa a los datos para maximizar la probabilidad, y es de esta manera en la cual se pueden especificar más variación y por ende obtener una menor contracción en los predictores que son importantes al modelo.

Groll & Tutz (2014) mencionan que en estudios previos la predicción utilizando LASSO es más baja y tan buena como la regresión en cresta a excepción cuando el modelo verdadero tiene muchos coeficientes distintos de cero. Tibshirani (1996) muestra que la técnica LASSO es una buena alternativa en comparación con la regresión *ridge*, además tiene una leve ventaja de generar un modelo de

interpretación más simple. Por otra parte, Ahmed & Fallahpour (2012) mencionan en sus principales resultados numéricos se tiene que el estimador de LASSO es mejor cuando hay un gran número de predictores en el modelo y solo algunos de ellos son irrelevantes.

En términos de ajuste Lykou & Ntzoufras (2013), el método de LASSO es más rápido, eficiente y estable que los modelos de selección paso a paso, computacionalmente y bajo condiciones apropiadas es más factible su uso (pág. 362).; y, en relación con el parámetro de contracción, la regresión LASSO es claramente mejor que la regresión de la cresta en términos de contracción ya que para la selección de las variables importantes al modelo, que los coeficientes pequeños se reducen hacia cero más rápido, mientras que se aplica menos contracción a los grandes coeficientes (pág. 363).

Ahora, también se deben de considerar algunas de las desventajas en este método son: (1) no permite estimar la incertidumbre del modelo – lo cual es importante dentro del marco bayesiano, especialmente si la meta es la predicción –. (2) la selección del valor para el parámetro de contracción  $\lambda$ , pues si se selecciona un valor pequeño no se realizará la contracción (o será leve), pero si se selecciona un valor muy alto, todos los coeficientes se reducirán a cero, para ello, es recomendable realizar el método de validación cruzada pues es un buen método que proporciona un mecanismo para seleccionar el valor de  $\lambda$ .

La validación cruzada también es utilizada por Chen, Liao & Schorfheide (2015) la definir el valor del  $\lambda$  para un modelo LASSO adaptativo en un estudio de panel que incluye factores latentes.

Matemáticamente la regresión LASSO tiene la siguiente representación:

$$y = X\beta + \varepsilon$$

Donde  $y$  es el vector de variable de respuesta,  $X$  corresponde al vector de variables predictoras,  $\beta$  el vector de parámetros y  $\varepsilon$  representa los errores. A partir de lo anterior, se tiene que los errores pueden ser presentados de la siguiente manera:

$$y - X\beta = \varepsilon$$

Así, la regresión LASSO realiza una penalización mediante el parámetro  $\lambda$  en los  $\beta$ 's de manera que se minimiza el cuadrado medio, tal como se presenta a continuación:

$$\hat{\beta}_{lasso} = \sum_{i=1}^N \left( y_i - \sum_j x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j|$$

La estimación LASSO también puede ser utilizada bajo el enfoque bayesiano, por ello es necesario que sea especificado las distribuciones de probabilidad previas de los parámetros que se desea estimar. Así mismo, puede obtenerse un valor del ponderador para cada variable predictiva, sin embargo, en este caso puntual se tiene un único valor del  $\lambda$ .

Para la selección del valor del parámetro de penalización *Lambda* ( $\lambda$ ) en la regresión LASSO multinivel, se realizó un estudio de simulación utilizando dos criterios: (1) la validación cruzada y (2) los valores AIC<sup>23</sup> y BIC<sup>24</sup> generados por el modelo.

En ambos casos se parte de un vector de posibles valores de *Lambda* que van desde 0 hasta 500, con intervalos de 5 unidades. (Ver anexo 5.1 y 5.2), para las

---

<sup>23</sup> AIC denominado Criterio de Información de Akaike, se calcula de la siguiente manera:  $AIC = 2k - 2\ln(L)$  donde  $k$  corresponde al número de parámetros y  $L$  es el valor máximo de la función de verosimilitud. Este estadístico de resumen mide la bondad de ajuste a partir de la máxima verosimilitud del modelo.

<sup>24</sup> BIC denominado Criterio de Información Bayesiano, se calcula de la siguiente manera:  $BIC = k * \ln(n) - \ln(L)$ , similar al caso anterior,  $k$  es el número de parámetros y  $L$  es el valor de la máxima verosimilitud y  $n$  es el número de datos, en este caso se agrega una penalización, con el tamaño de muestra.

estimaciones se utiliza el paquete `glmLasso` de la plataforma de programación R – los detalles del mismo pueden ser consultados en el Anexo 3.3 –.

El valor óptimo obtenido utilizando la validación cruzada y comparación de los valores AIC y BIC es  $\lambda=370$ , pero, dado que la regresión LASSO realiza una ponderación y selección de variables de forma simultánea, se tiene el caso de una variable cuyo coeficiente ponderado y estimado como 0; y, puesto que el interés de este estudio es proponer mecanismos para la estimación de los coeficientes de efectos fijos cuando se tiene *shrinkage*, mas no así la selección de variables pues estas se asocian a la teoría y conocimiento del investigador, se realizó nuevamente la estimación considerando el criterio de experto para varios modelos con distintos valores para *Lambda*: 0, 50, 70, 75, 80, 100 y 370. (Ver anexo 5.3)

Se selecciona  $\lambda=75$ , pues genera los errores estándar más bajos en la mayoría de los coeficientes de efectos fijos del modelo, además en ninguno de los casos se tiene coeficientes iguales a cero.

En este caso se indica en los componentes de la fórmula del modelo que se tiene sobre dispersión, dentro de las especificaciones del algoritmo de la estimación, denominado “control”, el siguiente corresponde al modelo son efectos aleatorios:

```
glmLasso(fix = y ~ DIA2 + DIA3 + M1 + M4 + M5 + M6 + M9 + M10 +
M11, rnd = list(Sujeto = ~ ID_DIA), data = base, family =
poisson(link = "log"), lambda = 75, switch.NR = FALSE, final.re =
TRUE, control = list(overdispersion = TRUE, maxIter = 1000))
```

Finalmente, se generan tres modelos a partir de los datos de la base madre y son considerados los valores base o de comparación para los obtenidos en el análisis de simulaciones que se realizó.

#### 4.6.2 Laplace

La distribución Laplace en algunos casos en mencionada como la versión bayesiana del método LASSO, sin embargo es llamada de esta manera porque en la mayoría

de los casos es la distribución Laplace la utilizada en las distribuciones previas. La versión bayesiana de la LASSO hereda todas sus características, las diferencias primordiales entre ambos es que en este caso se incorporan distribuciones previas a los parámetros y el método MCMC; por otra parte, no tiene la propiedad de establecer coeficientes no importantes iguales a cero, ello porque las medidas posteriores se ubican en el centro de la inferencia bayesiana y por ende no cuantifica la importancia de cada variable, además la distribución tiene colas un poco más gruesas, lo que provoca un menor efecto de encogimiento en los coeficientes grandes (Lykou & Ntzoufras, 2013, pág. 363).

Algunas de las principales ventajas del método de Laplace son indicadas por Carlin y Louis (2000): (1) es un procedimiento computacionalmente rápido, (2) utiliza diferenciación numérica en lugar de la integración numérica, el cual es más simple y estable, (3) es un algoritmo determinista, (4) reduce la complejidad computacional. Finalmente Lui (1993) menciona que la aproximación de Laplace indica ser muy precisa.

Ahora bien, la función de densidad de la doble exponencial o Laplace se detalla a continuación:

$$f(\beta|\mu, b) = \frac{1}{2b} \exp\left(-\frac{|\beta - \mu|}{b}\right)$$

Con media  $\mu$  y varianza  $2b^2$ . Como la normal, esta distribución es simétrica y unimodal, pero tiene colas más pesadas, siendo estrictamente cóncavas en ambos lados de  $\mu$ .

Las distribuciones previas para los parámetros corresponden a:  $\beta_j \sim DE(0, \sigma^2/\lambda)$  y se supone que los  $\beta_j$  son independientes. En el caso de la estimación del parámetro  $\lambda$ , a diferencia de la regresión LASSO, este se incorpora en la distribución previa, y se calcula de manera conjunta.

Se utiliza el paquete `glmmst` (Ogden, 2018) permite la aproximación de la verosimilitud mediante el cálculo de Laplace, son especificados los valores de los parámetros que se muestran a continuación, en los casos de lo que no se especifican es porque se decidió respetar los que de manera automática utiliza el paquete. Para mayor detalle ver anexo 3.4.

El siguiente corresponde al modelo sin efectos aleatorios:

```
glmm(y ~ DIA2 + DIA3 + M1 + M4 + M5 + M6 + M9 + M10 + M11 +
(1|ID_DIA) + (1|ID_SUJ), data = base, family = poisson, method
= "Laplace")
```

De igual manera que en el caso de la estimación LASSO, se generan tres modelos a partir de los datos de la base madre, los coeficientes son considerados como los valores de comparación para los obtenidos en el análisis de simulaciones realizado.

#### 4.6.3 Matriz de dispersión

El siguiente método propuesto corresponde al ajuste de la estimación en los modelos lineales generalizados jerárquicos mediante el modelado de la dispersión, lo cual consiste en agregar características adicionales a la estimación de los componentes del modelo mediante ajustes en la varianza y covarianza de los estimadores.

En Lee, Nelder y Pawitan (2006), se indica que a los modelos lineales generalizados multinivel se les puede agregar extensiones que permitan estructuras bien definidas para los parámetros de dispersión que se incorporan en la estimación; pues, si las variaciones desiguales no se modelan adecuadamente, pueden causar análisis no válidos. Por ello, son considerados los siguientes componentes que se presentan a continuación:

Considerando el componente sistemático del modelo lineal generalizado,

$$\eta = g(\mu) = X\beta + Zv.$$

Con una función de enlace, se incorporar un parámetro de dispersión al modelo  $\phi$  en la varianza, tal como se muestra:

$$E(y|u) = \mu$$

$$var(y|u) = \phi V(\mu)$$

Donde  $V(\cdot)$  es una función conocida,  $X$  y  $Z$  las respectivas matrices del modelo con las variables de interés,  $v = g_m(\mu)$  alguna función monótona de efectos aleatorios y  $\beta$  la matriz de efectos fijos del modelo.

Para la matriz de los efectos aleatorios, se repite lo anterior de manera que se incorpora una función de enlace en la cual para la función de varianza se tiene:

$$E(\varphi_M) = u$$

$$var(\varphi_M) = \lambda V_M(u)$$

Donde  $\lambda$  es el parámetro de dispersión para los efectos aleatorios  $u$  y  $\varphi_M$  es una constante que forma parte de fórmula de cálculo de la log verosimilitud para los efectos aleatorios.  $M$ , indica que el efecto aleatorio aparece en el predictor para la media.

A partir de lo anterior y para cada uno de los parámetros de dispersión se incorporan las dispersiones estructuradas  $(\phi, \lambda)$  mediante los siguientes modelos:

$$\xi = h(\phi) = G_\gamma$$

$$\xi_M = h_M(\lambda) = G_{M\gamma M}$$

Donde,  $h_M$  y  $h$  son funciones de enlace donde  $\gamma$  y  $\gamma M$  son los efectos fijos de los  $\phi$  y  $\lambda$ , respectivamente.

Dicha especificación está dada en el paquete `hglm`<sup>25</sup> y `hglm2` (una versión mejorada del paquete anterior) de la plataforma de programación R. (Para mayor detalle ver anexo 3.5)

Una de las principales características de este paquete es que permite especificar una distribución para los efectos aleatorios distinta a la normal, específicamente proveniente de la familia exponencial conjugada (Gaussiana, Gamma, Beta o Gamma inversa); además, incorpora una corrección en la estimación de la quasi probabilidad extendida mediante el método EQL1 – el cual mejora la estimación en los GLMM especialmente para los modelos de Poisson con un gran número de niveles en los efectos aleatorios (Ronnegard, Alam & Shen, 2010, pág. 26).

En este caso únicamente se ajustan los comandos de manera que son indicadas las matrices de los efectos fijos, el vector de respuesta y la matriz de efectos aleatorios, la familia con la respectiva función de enlace, el método de ajuste antes indicado (EQL1) y se indica que sea utilizada la matriz de dispersión mediante el comando “`sparse=TRUE`”. En este caso es importante mencionar que se respeta la especificación que R utiliza de manera automática para las funciones de enlace de los parámetros de dispersión.

El siguiente corresponde al modelo sin efectos aleatorios:

```
hglm2(y1 ~ DIA2 + DIA3 + M1 + M4 + M5 + M6 + M9 + M10 + M11 +
(1|ID_DIA) + (1|ID_SUJ), family = poisson(link = "log"), data =
base10, method = "EQL1", sparse = TRUE, maxit = 500)
```

---

<sup>25</sup> En GLMM, se asume que los efectos aleatorios son gaussianos, mientras que la función `hglm()` también extiende el algoritmo de ajuste del paquete `dglm` (Dunn y Smyth, 2009) al incluir efectos aleatorios en el predictor lineal para la media, es decir, extiende el algoritmo para que pueda manejar modelos mixtos. Además, la especificación del modelo en `hglm()` se puede dar como una fórmula o, alternativamente, en términos de `y`, `X`, `Z` y `X.disp`. Aquí `y` es el vector de respuestas observadas, `X` y `Z` son las matrices de diseño para el fijo y el aleatorio.

A nivel general se tiene que los componentes de dispersión son estimados a partir de h-verosimilitud, la cual fue adecuada para inferencia en modelos lineales generalizados, a partir de la corrección los *leverages* en la parte media del modelo.

Finalmente, Lee y Nelder (2006) mencionan que la h-verosimilitud desempeña un papel importante, además, da un algoritmo estadístico numéricamente eficiente que puede ser utilizado en otros modelos y no requiere la especificación de distribuciones previas.

En este caso también se generan tres modelos a partir de los datos de la base madre, los coeficientes son considerados como valores de comparación para los obtenidos en el análisis de simulaciones realizado.

#### **4.7 Estudio de simulaciones**

Se realizó un estudio de simulaciones con el fin de determinar, por una parte, la presencia del fenómeno de *shrinkage* en el modelo de interés, y, en segundo lugar, analizar los escenarios indicados por la teoría en los cuales se menciona puede estar presente el encogimiento de coeficientes.

##### 4.7.1 Diseño del estudio de simulaciones

Para el estudio de simulaciones se proponen cuatro escenarios, en los cuales los aspectos considerados fueron: el tamaño de la muestra, la varianza de las observaciones, el tipo de estimación y la cantidad de coeficientes aleatorios en el modelo.

Se plantean todas las combinaciones posibles en los escenarios y se realizaron las respectivas estimaciones.

Debido al interés de este trabajo de investigación se considera relevante la comparación entre los métodos de estimación frecuentista y bayesiano.

En el caso de la varianza, se definió como varianza pequeña la generada por las observaciones de los sujetos a los cuales se les suministró solución salina el día tres del experimento – es decir los sujetos del grupo control –, y varianza grande es la generada por las observaciones de los sujetos a los cuales se les suministró anfetamina el día tres del experimento.

El tamaño de la muestra hace referencia a la cantidad de sujetos sometidos al experimento, se proponen por criterio de experto tres casos, uno similar al real con 30 sujetos, otro con menos sujetos ( $n=10$ ), y un caso con más posibles sujetos (50).

La cantidad de coeficientes aleatorios se desagrega en tres escenarios, el primer modelo no incorpora efectos aleatorios, otro caso es un modelo con todos los efectos aleatorios posibles a estimar, el cual permite determinar la presencia de *shrinkage*; se decidió incluir un modelo “intermedio” con algunos coeficientes aleatorios, y poder contar con un punto de comparación entre los modelos antes indicados.

En el siguiente cuadro se muestra de manera detallada la desagregación por aspecto considerado:

Cuadro 1  
Descripción de los escenarios aplicados en las simulaciones

Aspecto	Escenario
Tipo de estimación	<ul style="list-style-type: none"> <li>• Frecuentista</li> <li>• Bayesiana</li> <li>• LASSO</li> <li>• Laplace</li> <li>• Matriz de dispersión</li> </ul>
Varianza	<ul style="list-style-type: none"> <li>• Varianza pequeña</li> <li>• Varianza grande</li> </ul>
Tamaño de muestra	<ul style="list-style-type: none"> <li>• <math>n = 10</math></li> <li>• <math>n = 30</math></li> <li>• <math>n = 50</math></li> </ul>
Cantidad de parámetros aleatorio	<ul style="list-style-type: none"> <li>• Modelo de efectos fijos.</li> <li>• Modelo de efectos fijos y tres efectos aleatorios el en 2do nivel M1, M5 y M11, denominado <i>modelo de efectos aleatorios</i>.</li> <li>• Modelo con todos los efectos aleatorios posibles de la variable medición, denominado <i>modelo de efectos aleatorios completo</i>.</li> </ul>

#### 4.7.2 Generación y análisis de los datos

Para el estudio de simulaciones se generaron bases de datos a partir de la base de datos original o base madre, que corresponde a la base que incorpora los registros de las vocalizaciones por sujeto, los tres días del experimento en los cuatro minutos anteriormente definidos.

Inicialmente, se parte de la consideración de la varianza generada por los sujetos según el tipo de tratamiento que recibieron, y se trabaja de manera separada. Por lo cual, se obtiene una base de datos para cada tipo de varianza, es decir, una nueva base de datos con “varianza grande” y otra con “varianza pequeña”, en cada caso se generan tres nuevas bases de datos de acuerdo al tamaño de muestra de interés, obteniendo así seis nuevas bases de datos.

Debido a las especificaciones iniciales para las nuevas bases, en la generación de la variable cantidad de vocalizaciones se realizó un muestreo considerando una selección sin reemplazo para  $n=10$ , y utilizando selección con reemplazo para  $n=30$  y  $n=50$ . También, se conservó el orden en la asignación de los días 1 y 2, así como las dummies de la variable medición especificada previamente en la base de datos.

Por ello, la distribución de la variable de respuesta puede representarse de la siguiente manera:

$$y \sim \text{Poisson}(\lambda | \text{sustancia asignada el día tres \& día del experimento})$$

Para la asignación del tratamiento en el tercer día del experimento, se seleccionó aleatoriamente la mitad de los sujetos que componen la muestra y se les asignó el tratamiento anfetamina (código uno), a los demás sujetos se les asignó solución salina (código cero), y para cada una de las seis bases de datos se realizó un análisis gráfico para determinar si bajo este método se logró dentro de lo posible replicar de manera similar el comportamiento de los datos de la base madre.

En cada base nueva se realizó la estimación de los tres modelos antes indicados considerando los métodos frecuentista y bayesiano, es decir, en cada base se estimaron seis modelos.

Incorporando los escenarios, se tiene un total de 36 modelos para predecir la cantidad de vocalizaciones emitidas por las ratas y cada uno fue replicado 100 veces.

Para el análisis de los modelos generados a partir de las bases de datos que surgieron de los cuatro escenarios propuestos en el cuadro 1, en los cuales se incorporan diferentes condiciones a las variables o factores de diseño, se calculó para cada coeficiente fijo el Error Cuadrático Medio<sup>26</sup> y el valor absoluto del

---

<sup>26</sup> Su forma de cálculo corresponde a  $MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$  y es una función de la varianza del estimador y sus sesgo. (Wackerly et al., 2010, p. 393)

SESGO<sup>27</sup>, con el fin de comparar las predicciones de los coeficientes en los respectivos modelos y determinar a partir de esto la presencia de *shrinkage*.

Para facilitar la comprensión del análisis fueron graficadas las estimaciones de los coeficientes fijos en cada escenario, ello también se generó la gráfica para los valores de SESGO y CME.

Finalmente, como mecanismo de control en la generación de las bases de datos, y con el fin de garantizar que cada escenario replicara de manera cercana el comportamiento de la base madre, se realizaron varias pruebas piloto en algunas de las muestra de bases generadas, en las cuales se graficó la variable de respuesta según grupo, día y sujeto por grupo asignado, y se comparó el comportamiento con el de la base madre obtenido al graficar las mismas variables (Este detalle no es presentado en el documento).

---

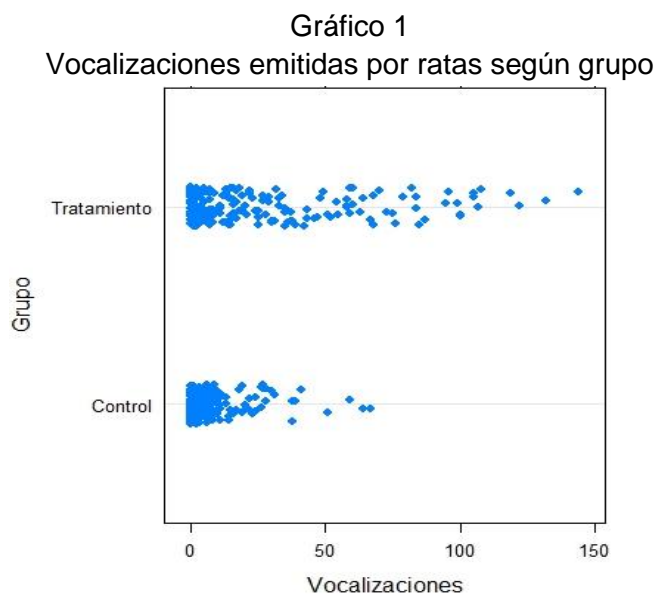
<sup>27</sup> Es sesgo de un estimador está dado por  $B(\hat{\theta}) = E|(\hat{\theta}) - \theta|$  y corresponde a la diferencia de la estimación del parámetro y es valor real del mismo. (Wackerly et al., 2010, p. 393)

## 5. Análisis de datos y resultados

### 5.1 Análisis descriptivo

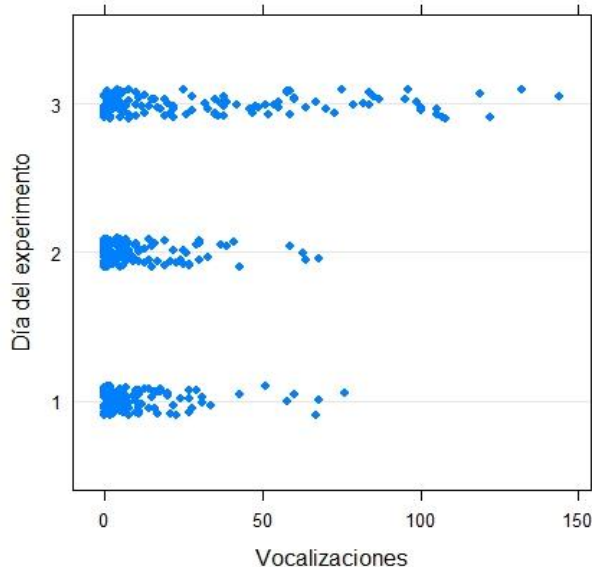
Inicialmente se muestra un análisis descriptivo de la base de datos madre, para comprender mejor el comportamiento mismo de los datos, así como el porqué de ciertas técnicas seleccionadas. En la base madre se incluyó una variable identificadora que permite diferenciar a los sujetos del grupo control y los del grupo tratamiento.

En el gráfico 1 se tiene que los sujetos que recibieron el tratamiento de amfetamina emitieron mayor cantidad de vocalizaciones que el grupo control, además las vocalizaciones están más dispersas en ese grupo, es decir, la varianza es mayor.



Aunado a lo anterior, en el gráfico 2 se presentan desagregadas las vocalizaciones emitidas por los sujetos en cada día del experimento. Específicamente, se tiene que el día tres fueron emitidas una mayor cantidad de vocalizaciones, que en los días uno y dos; además, en el día tres a nivel de inspección visual se observa una mayor dispersión en las vocalizaciones.

Gráfico 2  
Vocalizaciones emitidas por día del experimento



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

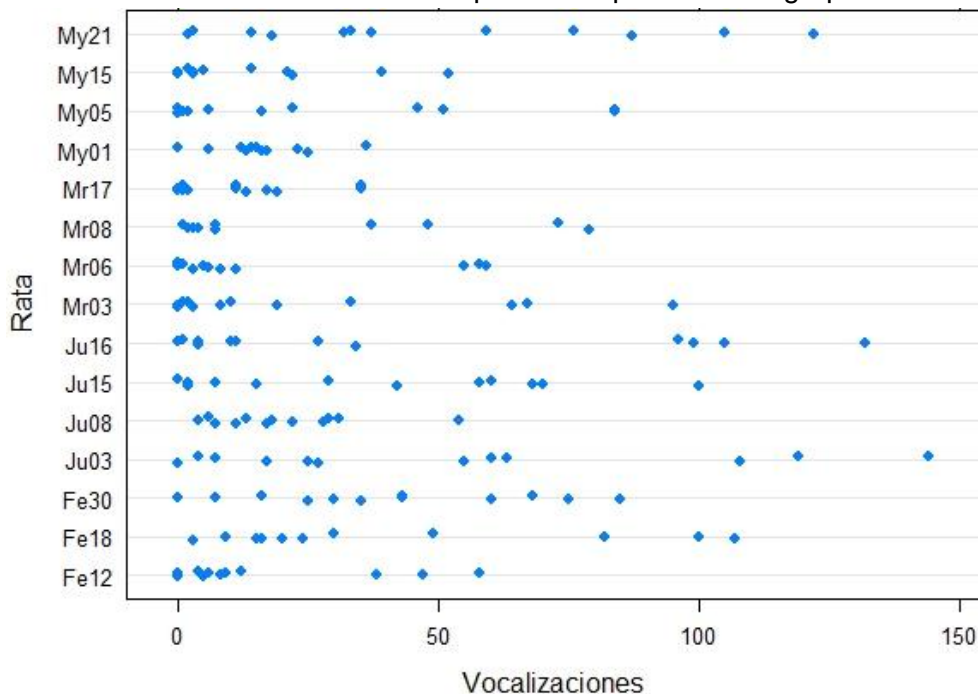
Considerando la variable grupo, se tiene que el comportamiento de los sujetos es similar, es decir, es homogénea la distribución de las vocalizaciones entre grupos y no dentro del grupo, ello porque la asignación de los tratamientos bajo criterio de experto se realizó a partir del comportamiento de los sujetos bajo estudio en los días uno y dos del experimento, de manera que lo deseado era obtener grupos similares entre sí, a fin de aplicar el tratamiento a sujetos que fuesen dentro de los posible similares.

En los gráficos 3 y 4 se presenta la cantidad de vocalizaciones que emitió cada sujeto en los tres días del experimento. Los sujetos del grupo de control se muestran homogéneos, algunos presentan un máximo de 60 vocalizaciones (gráfico 3); en general los sujetos del grupo control emitieron menor cantidad de vocalizaciones.



Gráfico 4

Vocalizaciones emitidas durante el experimento por ratas del grupo tratamiento



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

En el cuadro 2 se presenta desagregado por día del experimento, la cantidad total de vocalizaciones emitidas por todas las ratas, su promedio y varianza. En términos generales se observa un aumento en las vocalizaciones para el tercer día, así como una diferencia entre los sujetos que recibieron la anfetamina y los que no, además de una varianza y promedio mayor, ello nos da indicios de que los datos presentan sobre dispersión, lo cual se consideró en la especificación de los modelos utilizados.

Cuadro 2

Estadísticas generales de las vocalizaciones emitidas según día del experimento

Día	Total de vocalizaciones emitidas	Total de registros	Promedio de vocalizaciones emitidas	Varianza de vocalizaciones emitidas
Día 1	1376	124	11,10	219,70
Día 2	1248	124	10,06	200,68
Día 3				
Sin tratamiento	542	64	8,47	153,40
Con tratamiento	3540	60	59,00	1232,27
<b>Total</b>	<b>6706</b>	<b>372</b>	<b>18,03</b>	<b>685,89</b>

Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Adicionalmente se tiene que: las vocalizaciones de los sujetos del grupo control son similares en cada día del experimento (ver anexo 1, gráfico A1), la mayor frecuencia de vocalizaciones se concentra en valores bajos, teniendo una distribución asimétrica positiva (ver anexo 1, gráfico A2); y, el detalle de la cantidad de vocalizaciones por sujeto del grupo control, así como su promedio y varianza se muestra en cuadro A1 del anexo 1, se observa que la varianza es superior al promedio. Para los sujetos del grupo tratamiento se tiene que las vocalizaciones son similares en los días uno y dos del experimento, pero en el día tres se observa un leve aumento en la cantidad de vocalizaciones así como en la varianza (ver anexo 1, gráfico A3), la distribución de las vocaciones emitidas presenta una asimetría positiva (ver anexo 1, gráfico A4); y, el detalle de la cantidad de vocalizaciones por sujeto del grupo control, así como su promedio y varianza se muestra en cuadro A2 del anexo 1. Finalmente, se tiene que la varianza es mayor al promedio de vocalizaciones emitidas.

## 5.2 Modelos generados

A partir de la base de datos madre, se estimaron los modelos con enfoque frecuentista y bayesiano, considerando tres modelos con distinta cantidad de parámetros aleatorios, los resultados se presentan en los cuadros 3 y 4, respectivamente.

En el caso frecuentista, se presenta el error estándar para la estimación de los efectos fijos y la desviación estándar para la estimación de los efectos aleatorios. En el caso de la estimación bayesiana, se muestran los intervalos de credibilidad con los límites inferior y superior al 95%.

En las estimaciones frecuentistas fueron generadas con el paquete `glmer` – presentadas en el cuadro 3 – a partir del segundo modelo se tuvieron problemas de convergencia al ir incorporando paulatinamente coeficientes aleatorios, por ello a partir de consultas a expertos sobre el paquete utilizado, se adiciona en la especificación del modelo el siguiente parámetro “`glmerControl`”, en el cual se define un optimizador denominado “`bobyqa`”<sup>28</sup> y un elemento “`Maxfun`”<sup>29</sup>.

Sin embargo, surge la inquietud de si al realizar estos ajustes se podría estar obligando a los modelos a lograr la convergencia, pues ello podría generar resultados cuestionables, por lo anterior, se decide de manera alternativa utilizar dos paquetes más de la plataforma de programación R para realizar la estimación frecuentista: (1) `glmmPQL` presentado por Finch, W. et al (2014) como una alternativa para la estimación de modelos multinivel con variables dependientes de distribución Poisson; y, (2) `glmmTMB` utilizado para el ajuste de los modelos multinivel, según el autor este cuenta con ventajas ante presencia de ceros y mayor velocidad cuando se tienen gran cantidad de parámetros Bolker (2019).

El paquete `glmmPQL` también presentó problemas de convergencia a partir del segundo modelo en el cual se incorporan paulatinamente los coeficientes aleatorios, por lo cual se descartó la posible comparación con este, ya que al ajustar las especificaciones en el modelo para facilitar la convergencia se podría incurrir en la misma disyuntiva del paquete `glmer`. Sin embargo, es importante mencionar que los resultados del modelo de efectos fijos en ambos casos mostraron valores

---

<sup>28</sup> El propósito de “`bobyqa`” es minimizar la función de muchas variables mediante un método de región de confianza que forma modelos cuadráticos por interpolación. Se permiten restricciones de cuadro (límites) en los parámetros.

<sup>29</sup> Corresponde al número máximo permitido de evaluaciones de funciones. Si se supera esto, el método terminará.

similares con diferencias a partir del tercer decimal. (Los resultados de las estimaciones no se incluyen en este documento).

En el caso de los resultados obtenidos con el paquete `glmmTMB`, con este se logró la estimación de los tres modelos, el de efectos fijos, el que incluye efectos aleatorios y el de efectos aleatorios completo. Al comparar los resultados obtenidos con la información del cuadro 3, se perciben diferencias a partir del tercer decimal en algunos de los casos únicamente, manteniendo así la dirección de los coeficientes, la magnitud y consistencia con las variables que no resultan significativas al modelo con un 5%, los resultados se presentan en el cuadro A4 del anexo 4.1.

A partir de los resultados anteriores, se decide mantener los resultados obtenidos con el paquete `glmer`, con el ajuste de los parámetros que facilitaron la convergencia.

De manera paralela, se realiza la estimación de los modelos de efectos fijos y con efectos aleatorios utilizando el paquete STATA 15, sin embargo, debido a la duración de las estimaciones, así como la capacidad en la convergencia, se logró la estimación del modelo de efectos fijos y el modelo que incorporaba algunos coeficientes aleatorios. Los modelos obtenidos en STATA 15 presentaron resultados similares en dirección a los obtenidos con la plataforma de programación R, con diferencias desde el primer decimal, mismos que se muestran en el anexo 4.2, imágenes A1, A2, A3 y A4.

En la estimación frecuentista presentada en el cuadro 3, se tiene que la estimación de los efectos fijos cambia levemente a medida que se agregan efectos aleatorios al modelo. Puntualmente, todas las variables presentaron diferencias iguales o superiores al 0.10, exceptuando M10 con 0.06, al comparar las estimaciones en el modelo de únicamente efectos fijos y el modelo de efectos aleatorios completo; las variables M6 y M9 presentaron diferencias iguales o mayores a 0.4, M4 una diferencia de 0.28, así como DIA2 y DIA3 con diferencias de 0.2 respectivamente.

Por otra parte, se tiene que el error estándar presentó en todos los efectos fijos un aumento gradual a medida que se agregan coeficientes aleatorios al modelo, lo que podría ser indicio de la presencia de *shrinkage*. Los errores estándar que presentaron un mayor cambio son los de las variables M4, M6 y M9, con una diferencia igual o superior a 0.19.

Para la estimación de los efectos aleatorios se presentaron cambios en las estimaciones a medida que se agregan en el modelo, pero, no se determina un patrón definido pues en algunas de las variables se tienen aumentos y en otras disminuciones.

Finalmente, en términos de significancia estadística y considerando un 5%, se tiene que las variables M6, M10 y M11 pierden significancia en el modelo con mayor cantidad de efectos aleatorios, ello, al rechazarse la hipótesis nula de que son distintos de cero.

**Cuadro 3**  
**Estimación frecuentista para los coeficientes de los modelos de efectos fijos,**  
**efectos aleatorios y efectos aleatorios completos**

Variable	Modelo de efectos fijos		Modelo con efectos aleatorios		Modelo completo de efectos aleatorios	
	Coeficiente	Error estándar	Coeficiente	Error estándar	Coeficiente	Error estándar
Intercepto	2.0120	0.1252	2.0581	0.1281	1.7111	0.1575
DIA2	-0.6971	0.1289	-0.8266	0.1396	-0.4923	0.1893
DIA3	1.4892	0.1621	1.4001	0.1780	1.7248	0.2476
M1	0.7268	0.0550	0.4537	0.2143	0.5630	0.2213
M4	-1.2377	0.1113	-1.2509	0.1108	-1.5226	0.3012
M5	1.3306	0.0685	1.2253	0.2115	1.2060	0.2518
M6	0.6649	0.0789	0.6649	0.0786	0.2286 <sup>a/</sup>	0.2716
M9	0.2829	0.0464	0.2935	0.0466	0.7868	0.2897
M10	0.3852	0.0454	0.3958	0.0456	0.3258 <sup>a/</sup>	0.1807
M11	0.3204	0.0461	-0.3016 <sup>a/</sup>	0.2021	0.2135 <sup>a/</sup>	0.1297
Efectos aleatorios	Varianza	Desviación estándar	Varianza	Desviación estándar	Varianza	Desviación estándar
ID_SUJ	0.2971	0.5451	0.2882	0.5368	0.3111	0.5578
ID_DIA	0.2132	0.4618	0.2493	0.4993	0.4895	0.6997
M1			1.2266	1.1075	1.2923	1.1368
M5			1.0885	1.0433	1.5782	1.2563
M11			0.8736	0.9347	0.1621	0.4027
M4					1.6494	1.2843
M6					1.5289	1.2365
M9					2.3572	1.5353
M10					0.6492	0.8057

<sup>a/</sup> Coeficientes en los cuales se rechaza la hipótesis nula de que son distintos de cero, es decir, no resultaron significativos al modelo, con  $\alpha=0.05$ .

En caso de la estimación bayesiana, se realizan una revisión general de los resultados, y de manera preliminar se comparan tentativamente con los resultados frecuentistas, teniendo conocimiento de que estos pueden diferir debido a la misma naturaleza de cada enfoque y de los datos considerados.

Ahora, si bien se sabe que las estimaciones bayesianas a diferencia de las frecuentistas no serán las mismas en cada “corrida” del modelo – pues los valores de los coeficientes son los puntos en los cuales las estimaciones de las cadenas de Markov se estabilizan – es de esperar que estos no difieran en gran medida pues el

punto de convergencia tiende o debería ser similar; si ello ocurriera podría ser indicio de que algo no está bien con el modelo.

Puntualmente se tiene que el valor del intercepto de la variable día difiere mucho en ambos enfoques, por lo que se decidió repetir en varias ocasiones la estimación bayesiana para el modelo de efectos fijos que incluye los interceptos para los niveles de día y sujeto.

Se logra identificar que en la estimación del intercepto día el parámetro “*eff.samp*” (efecto de la muestra) cambia mucho, por ejemplo: en algunas corridas se tienen valores de 1000 y al repetir el ejercicio el valor cae a 0, y en una nueva corrida sube a 103, en diversas consultas realizadas a colegas de estadística y matemática, los profesionales mencionaron que éstos cambios podrían estar provocados por altas correlaciones internas de los datos, o porque no se tiene suficiente *burn-in*.

Por ello, se decide realizar un análisis para el modelo inicial modificando la especificación de la parte aleatoria, a partir de las funciones disponibles en el paquete *MCMCglmm* (ver detalle en anexo 3.2) se aumentó la cantidad de las iteraciones a 30 000, así como el *burn-in*, se decide no ajustar la información de las distribuciones previas, pues se considera que no es apropiado buscar un buen ajuste sin tener suficiente conocimiento del fenómeno al cual corresponden los datos, además de que un cambio en la información previa puede cambiar también las conclusiones obtenidas (Los resultados no se muestran en este documento).

Se decidió utilizar para los análisis 30 000 iteraciones con 10 000 de *burn-in*, y distribuciones previas para los hiperparámetros no informativas, pues las estimaciones obtenidas bajo estos parámetros resultaron levemente más estables en los escenarios considerados.

Además, se utilizó el programa STATA 15 para realizar la estimación bayesiana, los resultados se presentan en el anexo 4.2, específicamente en las imágenes A5 y A6. Al comparar ambos resultados, se observa que las estimaciones difieren a partir del primer decimal, además que en STATA son más estables las estimaciones de los

efectos aleatorios que las obtenidas en la plataforma de programación R, inclusive siendo estas muy similares a las estimaciones frecuentistas, es importante indicar que en este caso tampoco se logró la convergencia del modelo de efectos aleatorios completo.

También, se debe indicar que el programa STATA 15 muestra una nota aclaratoria mencionando que es posible que exista una alta correlación entre los datos, lo cual podría afectar las estimaciones, y que se tenga alta correlación para algunas estimaciones al ajustar un modelo multinivel con antecedentes no informativos.

Por ello, se decide seguir trabajando con la plataforma de programación R pues permite estimar todos los modelos de manera rápida y hacer la comparación de los posibles cambios en los coeficientes de los efectos fijos, que es uno de los intereses de este análisis.

Los resultados en el cuadro 4 muestran que, para los coeficientes fijos se tienen diferencias menores o iguales a 0.04 al incorporar los parámetros aleatorios al modelo, es decir, las estimaciones bayesianas son bastante estables.

En términos de significancia estadística, para cada efecto fijo se analizan los intervalos de credibilidad, los casos en que se incluye al valor 0 en los intervalos, se concluye que no son significativos al modelo, puntualmente las variables M10 y M11 presentan valores de 0 entre las estimaciones de los intervalos de credibilidad; es decir, no son significativas en ninguno de los modelos estimados bajo el enfoque bayesiano.

En ambos casos, las estimaciones frecuentistas y bayesianas muestran similitudes en el tamaño de los coeficientes, así como en la dirección de estos, y en ambos casos las variables M10 y M11 pierden significancia estadística a medida que se incorporan efectos aleatorios. También, se tiene que en ambos casos los resultados muestran posible presencia del fenómeno *shrinkage*.

Es importante aclarar que para las alternativas se utilizó la plataforma de programación R, pues es la que logró en la convergencia de las estimaciones en la totalidad de los modelos de interés.

Cuadro 4

Estimación bayesiana para los coeficientes de los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos

Variable	Modelo de efectos fijos			Modelo con efectos aleatorios			Modelo completo de efectos aleatorios		
	Coeficiente	IC – inf	IC – sup	Coeficiente	IC – inf	IC – sup	Coeficiente	IC – inf	IC - sup
Intercepto	1.6070	1.3032	1.9494	1.6057	1.2948	1.9059	1.6054	1.2567	1.9152
DIA2	-0.7642	-1.1443	-0.3629	-0.7615	-1.1612	-0.3803	-0.7479	-1.1388	-0.3556
DIA3	1.9091	1.5073	2.3395	1.9066	1.4664	2.3220	1.9253	1.5449	2.4041
M1	0.9146	0.4595	1.3508	0.9201	0.4108	1.3438	0.9063	0.4528	1.3534
M4	-1.1421	-1.6469	-0.6132	-1.1493	-1.6278	-0.6018	-1.1401	-1.6493	-0.5789
M5	1.6118	1.1466	2.1439	1.6108	1.1313	2.1430	1.6020	1.1470	2.0936
M6	0.6933	0.1945	1.2237	0.6963	0.1932	1.2171	0.6833	0.2108	1.1915
M9	0.5654	0.0886	1.0095	0.5679	0.0602	1.0046	0.5688	0.0662	1.0554
M10	0.0705 <sup>a/</sup>	-0.4055	0.5332	0.0773 <sup>a/</sup>	-0.4135	0.5482	0.0569 <sup>a/</sup>	-0.4441	0.4907
M11	-0.1509 <sup>a/</sup>	-0.6686	0.3111	-0.1431 <sup>a/</sup>	-0.6059	0.3697	-0.1558 <sup>a/</sup>	-0.6691	0.2756
<b>Efectos aleatorios</b>	<b>Varianza</b>	<b>IC – inf</b>	<b>IC – sup</b>	<b>Varianza</b>	<b>IC – inf</b>	<b>IC – sup</b>	<b>Varianza</b>	<b>IC – inf</b>	<b>IC - sup</b>
ID_SUJ	0.3270	0.1267	0.5787	0.2781	0.0000	0.5398	0.0224	0.0000	0.1680
ID_DIA	0.0016	0.0000	0.0008	0.0027	0.0000	0.0152	0.0050	0.0000	0.0300
M1				0.0382	0.0000	0.3145	0.0679	0.0000	0.3287
M5				0.0033	0.0000	0.0175	0.0035	0.0000	0.0155
M11				0.0120	0.0000	0.0688	0.0012	0.0000	0.0022
M4							0.1026	0.0000	0.3969
M6							0.0020	0.0000	0.0040
M9							0.1740	0.0003	0.5050
M10							0.0147	0.0000	0.1045

IC = Intervalo de credibilidad.

<sup>a/</sup> Coeficientes que incluyen el valor cero en el intervalo de credibilidad.

En el cuadro 5, se presentan los estadísticos de resumen AIC, BIC y DIC<sup>30</sup> para los modelos de efectos fijos, así como los que incorporan efectos aleatorios respectivamente.

Es el modelo de efectos aleatorios completo el que presenta los valores más bajos, lo cual intuitivamente se esperaba puesto que la razón de verosimilitud arroja resultados a favor del modelo más completo. En el caso del DIC los valores son similares en los tres modelos estimados.

Es importante mencionar que, en este caso puntual, considerar los valores de estos estadísticos no es un criterio que permita determinar cuál es el mejor modelo, sin embargo, se muestran como información adicional a la investigación.

Cuadro 5  
Estadísticos de resumen para los modelos estimados con R según enfoque frecuentista y bayesiano

Modelo	Estadísticos de resumen		
	AIC	BIC	DIC
De efectos fijos	4 146.6	4 193.7	2 029.2
De efectos aleatorios	3 645.0	3 724.2	2 029.7
Completo de efectos aleatorios	2 879.5	3 063.2	2 031.2

En los cuadros 6, 7 y 8 se presentan los resultados de las estimaciones obtenidas al considerar los ajustes LASSO, Laplace y ajuste de matriz respectivamente, generadas a partir de la base de datos madre, mismas que corresponden a las alternativas propuestas para la estimación del modelo bajo estudio y determinar si estas son buenas alternativas para contrarrestar el fenómeno de *shrinkage* en los modelos multinivel.

<sup>30</sup> DIC denominado Criterio de Información de Deviancia, corresponde a una versión generalizada de los estadísticos AIC y BIC, pero que incorpora procedimientos y la lógica de la estadística bayesiana. Y se calcula de la siguiente manera El DIC utiliza cadenas de Monte Carlo para calcular la distribución *a posteriori* de los parámetros. Luego trabaja con estas distribuciones (no solo con un valor puntual del parámetro) para medir la complejidad y la bondad de ajuste. Primero, la bondad de ajuste será el promedio de la devianza total del modelo:  $\bar{D}(\theta) = -2 \int \log(p(y|\theta))d\theta$ , donde  $y$  corresponde a los datos,  $\theta$  los parámetros desconocidos en el modelo, y  $p(y|\theta)$  la verosimilitud. El criterio de Información de Deviancia se calcula como:  $DIC = D(\bar{\theta}) + 2p_D$ , donde  $p_D$  es el número de parámetros del modelo.

La primera propuesta corresponde a la estimación LASSO, como se indicó anteriormente, en este caso se cuenta con un parámetro denominado *Lambda*, debido a que se desconoce cuál es el valor que permitirá un mejor ajuste en este análisis, y a partir la validación cruzada y uso de AIC y BIC, se propone seis valores para *Lambda*: 0, 50, 75, 80, 100 y 370 (siendo este el recomendado por ambas técnicas). Se presentan las estimaciones de los coeficientes, así como los errores estándar únicamente en los modelos de efectos fijos y el modelo de efectos aleatorio completo.

En el modelo de efectos fijos (cuadro A5 del anexo 5), se tiene los mismos valores del error estándar considerando *Lambda* 0, 50, 75, 80 y 100, al aumentar el valor de *Lambda* a 370 se perciben diferencias a partir del segundo decimal en algunas de las variables, además para M6 se tiene que el ponderador es cero (cuadro A6 del anexo 5), lo cual indica que matemáticamente dicha variable pareciera no aportar información al modelo; sin embargo, todos son significativos al modelo.

En los modelos de efectos aleatorios completos (cuadro A7 del anexo 5), los errores estándar muestran diferencias a partir del segundo decimal, lo que se logra identificar un comportamiento específico, sin embargo, a medida que se aumentó el valor del *Lambda* los errores tienen a mostrar un cambio. Además, para todos los valores de *Lambda* (excepto  $\lambda=100$ ) las variables M6 y M11 no son significativas el modelo. Además, para  $\lambda=370$  el ponderado de M6 es cero (cuadro A8 del anexo 5), es decir, no debería ser seleccionado en el modelo.

A partir de los resultados anteriores y de la *expertise* de investigador, se decide utilizar para las estimaciones un *Lambda* de 75, ello porque en el modelo de efectos fijos todos los coeficientes son estimados, recordando que el interés es buscar un modelo que facilite la estimación, mas no realizar selección de variables.

En el cuadro 6 se tiene que la estimación utilizando el método de LASSO, en los efectos fijos cambia levemente a medida que se agregan efectos aleatorios al modelo.

Puntualmente, la mayoría de los coeficientes presentan diferencias desde el primer decimal, las diferencias son iguales o superiores al 0.18, exceptuando M1 con 0.05 y M4 con 0.12.

Además, se tiene que el error estándar presentó en todos los efectos fijos un aumento gradual a medida que se agregan coeficientes aleatorios al modelo, los que podría ser indicio de la presencia de *shrinkage* – no se considera el leve descenso del intercepto –. Los errores estándar que presentaron un mayor cambio son los de las variables M1, M6 y M9, con una diferencia igual o superior a 0.15.

Cuadro 6  
Estimación con regresión LASSO para los coeficientes en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos

Variable	Modelo de efectos fijos		Modelo con efectos aleatorios		Modelo completo de efectos aleatorios	
	Coeficiente	Error estándar	Coeficiente	Error estándar	Coeficiente	Error estándar
Intercepto	1.9248	0.1051	2.0474	0.1098	1.6393	0.1032
DIA2	-0.4739	0.0711	-0.5581	0.0745	-0.1961	0.0849
DIA3	1.6401	0.1061	1.3809	0.1253	1.9756	0.1534
M1	0.7283	0.0551	0.5354	0.1982	0.7733	0.2051
M4	-1.2363	0.1117	-1.2548	0.1117	-1.1125	0.2462
M5	1.3306	0.0688	1.0583	0.1759	0.8104	0.1838
M6	0.6649	0.0793	0.6649	0.0793	-0.0476 <sup>a/</sup>	0.2332
M9	0.2818	0.0464	0.2967	0.0467	0.4663 <sup>a/</sup>	0.2447
M10	0.3841	0.04538	0.3990	0.0457	0.1281 <sup>a/</sup>	0.1328
M11	0.3192	0.0460	-0.2054 <sup>a/</sup>	0.1711	0.1243 <sup>a/</sup>	0.0924
Efectos aleatorios	Varianza	Desviación estándar	Varianza	Desviación estándar	Varianza	Desviación estándar
ID_SUJ	1.1990	1.0950	1.3450	1.1597	4.6152	2.1483
ID_DIA	0.0177	0.1329	0.0221	0.1488	0.1212	0.3481
M1			0.0101	0.1005	0.0084	0.0919
M5			0.0033	0.0573	0.0061	0.0781
M11			0.0030	0.0543	0.0002	0.0131
M4					0.0111	0.1055
M6					0.0104	0.1021
M9					0.0318	0.1783
M10					0.0024	0.0488

<sup>a/</sup> Coeficientes que no resultaron significativos al modelo, con  $\alpha=0.05$ .

Para la estimación de los efectos aleatorios se presentaron cambios en las estimaciones a medida que se agregan en el modelo, pero, no se determina un

patrón definido pues en algunas de las variables se tienen aumentos y en otras disminuciones.

Finalmente, en términos de significancia estadística en el modelo completo de efectos aleatorios y considerando un 5%, se tiene que las variables M6, M9, M10 y M11 pierden significancia, ello, al rechazarse la hipótesis nula de que son distintas de cero.

En el cuadro 7 se muestran los resultados de la estimación con aproximación Laplace, los ajustes de los parámetros que se incorporan en los modelos son ajustados de manera automática por la plataforma de programación R. En este caso los efectos fijos cambian levemente a medida que se agregan efectos aleatorios al modelo, y se perciben diferencias en los coeficientes estimados desde el primer decimal. Todas las variables presentaron diferencias superiores al 0.17, al comparar las estimaciones en el modelo de únicamente efectos fijos y el modelo con más efectos aleatorios; a excepción de la variable M10 con una diferencia de 0.07. Las variables con una mayor diferencia corresponden a DIA3 con 0.62, seguida por M6 con 0.54 y M4 con 0.49.

En el caso del error estándar, todos los efectos fijos muestran un aumento gradual a medida que se agregan coeficientes aleatorios al modelo completo, las diferencias son superiores a 0.09, el error estándar que presentó mayor cambio corresponde al de la variable DIA2 con 0.33.

Para la estimación de los efectos aleatorios en los modelos completos, no se identifica un patrón claro y definido, puesto que en algunos casos algunas variables presentan aumentos y otros decrecimientos. Además, en el caso de los errores estándar, las estimaciones en el modelo completo muestran aumentos con excepción de la varianza de la variable M11.

En términos de significancia estadística, y considerando el modelo completo, son las variables DIA2, M6 y M11 las que tiene un valor p menor a 0.05, es decir, estas

no son significativas al modelo, pues se rechaza la hipótesis de que son distintas de cero.

Cuadro 7

Estimación con cadenas de Markov con aproximación Laplace para los coeficientes de los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos

Variable	Modelo de efectos fijos		Modelo con efectos aleatorios		Modelo completo de efectos aleatorios	
	Coefficiente	Error estándar	Coefficiente	Error estándar	Coefficiente	Error estándar
Intercepto	2.1031	0.1098	2.1565	0.2153	1.7677	0.4091
DIA2	-0.6208	0.0649	-0.6910	0.0651	-0.3277 <sup>a/</sup>	0.3973
DIA3	1.3288	0.0462	1.1662	0.0524	1.9503	0.1318
M1	0.7698	0.0515	0.4111	0.1885	0.4175	0.2090
M4	-1.1948	0.1100	-1.2646	0.1102	-1.6837	0.2833
M5	1.3303	0.0688	1.0714	0.1741	1.0423	0.1879
M6	0.6650	0.0792	0.6649	0.0793	0.1212 <sup>a/</sup>	0.2219
M9	0.2498	0.0433	0.3048	0.0443	0.7311	0.2516
M10	0.3526	0.0422	0.4071	0.0433	0.2784	0.1406
M11	0.2876	0.0429	-0.1491 <sup>a/</sup>	0.1662	0.1206 <sup>a/</sup>	0.1416
Efectos aleatorios	Varianza	Error estándar	Varianza	Error estándar	Varianza	Error estándar
ID_SUJ	0.5827	0.0765	0.5292	0.2311	0.0006	0.8176
ID_DIA	0.0002	0.0306	0.0000	0.0226	0.3170	0.1476
M1			0.6865	0.1040	0.7623	0.1128
M5			0.5959	0.0955	0.6475	0.1050
M11			0.5429	0.1025	0.4053	0.0867
M4					0.8703	0.1547
M6					0.7149	0.1290
M9					0.9417	0.1283
M10					0.4410	0.0849

<sup>a/</sup> Coeficientes que no resultaron significativos al modelo, con  $\alpha=0.05$ .

En el cuadro 8 se muestran las estimaciones utilizando el método de matriz dispersa, para los efectos fijos, estos presentan diferencias desde el primer decimal y superiores a 0.20 excepto por las variables M4, M5, M10 y M11 que muestran diferencias menores a 0.10.

En el caso del error estándar, no se logra identificar un patrón definido pues en algunas de las variables se observa en el error estándar un aumento gradual y en otros decrecimientos, es el error estándar de la variable DIA2 el que presenta mayor diferencia (0.39); en los demás casos las diferencias son iguales o inferiores a 0.10.

Para la estimación de los efectos aleatorios se perciben cambios, pero no se identifica un patrón definido de aumento o decrecimiento.

Al analizar la significancia estadística, y considerando el modelo completo, para los coeficientes fijos son las variables DIA2 y M6 las que presentan un valor p menor a 0.05, es decir, estas no son significativas al modelo, pues se rechaza la hipótesis de que son distintas de cero.

**Cuadro 8**  
Estimación con matriz dispersa para los coeficientes de los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos

Variable	Modelo de efectos fijos		Modelo con efectos aleatorios		Modelo completo de efectos aleatorios	
	Coefficiente	Error estándar	Coefficiente	Error estándar	Coefficiente	Error estándar
Intercepto	2.0948	0.1576	2.0910	0.1727	1.8739	0.3752
DIA2	-0.5741	0.2414	-0.5671	0.2704	-0.3384 <sup>a/</sup>	0.6339
DIA3	1.4875	0.1499	1.5127	0.1557	2.0351	0.1829
M1	0.7723	0.1649	0.6516	0.1872	0.5650	0.1875
M4	-1.1923	0.3439	-1.2082	0.3342	-1.2453	0.2984
M5	1.3306	0.2136	1.2822	0.2164	1.2377	0.1982
M6	0.6649	0.2461	0.6649	0.2389	0.4307 <sup>a/</sup>	0.2473
M9	0.2483 <sup>a/</sup>	0.1375	0.2601 <sup>a/</sup>	0.1364	0.5076	0.2326
M10	0.3506	0.1343	0.3624	0.1333	0.3810	0.1233
M11	0.2857	0.1363	0.2905	0.1369	0.3134	0.1257
Efectos aleatorios	Varianza	Desviación estándar	Varianza	Desviación estándar	Varianza	Desviación estándar
ID_SUJ	0.2566	0.2901	0.2492	0.2914	0.2186	0.2942
ID_DIA	0.0122	2.4275	0.0237	1.9759	0.2467	1.4764
M1			0.2028	0.4956	0.3326	0.3997
M5			0.0831	0.6982	0.1622	0.4979
M11			0.0058	1.7261	0.0064	1.5048
M4					0.0150	2.8281
M6					0.3777	0.4778
M9					0.9478	0.3099
M10					0.0039	1.8153

<sup>a/</sup> Coeficientes que no resultaron significativos al modelo, con  $\alpha=0.05$ .

En general las alternativas propuestas muestran resultados similares a las estimaciones frecuentista y bayesiana, en magnitud de los coeficientes de los parámetros fijos, así como la dirección de estos, incluso coinciden algunas de las variables que pierden significancia en los modelos de efectos aleatorios completos.

Por otra parte, resalta el caso de las variables M6, M10 y M11 que no muestran significancia estadística en la mayoría de las alternativas consideradas, en el caso de la variable M6, a medida que se agregan coeficientes aleatorios al modelo esta deja de ser significativa independientemente de la alternativa que se utilice, para M11 resulta significativa al considerar el método de matriz de dispersión. Se debe tener presente que esta variable es de las últimas medidas que son tomadas a los sujetos en el experimento, además de que se realizan en el tercer día cuando se aplica la anfetamina, por lo que es importante para el investigador.

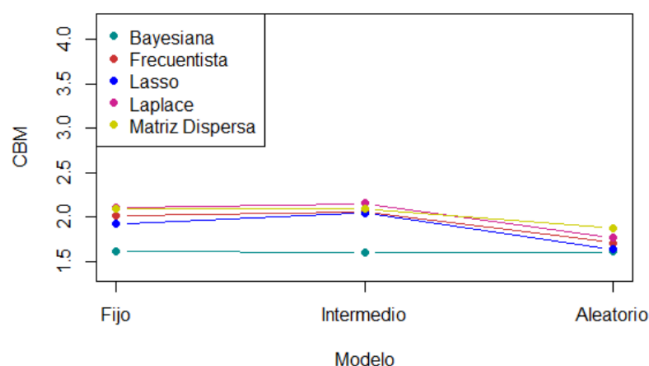
Se debe resaltar que en todos los coeficientes se tienen cambios en los valores de sus estimaciones, así como en el error estándar lo que sugiere que se tiene *shrinkage* o fenómeno de encogimiento en este caso de estudio.

### **5.3 Gráficas de coeficientes estimados**

A continuación, se muestran los gráficos correspondientes a las estimaciones de los coeficientes fijos para cada modelo considerando la información de la base madre (CBM).

En el gráfico 5 se tienen las estimaciones del intercepto, la estimación bayesiana es la que muestra visualmente una mayor estabilidad y menor valor en relación con las demás, en las otras estimaciones se tiene un comportamiento similar, y una disminución a medida que se agregan efectos aleatorios.

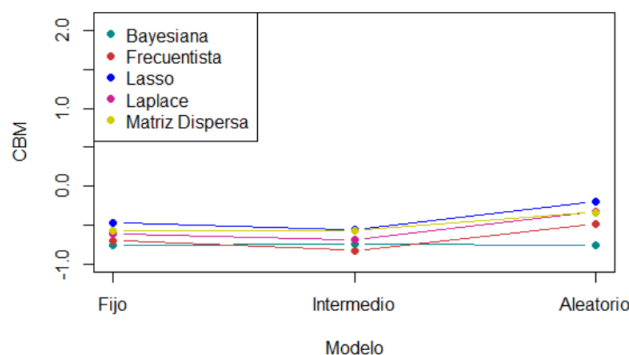
Gráfico 5  
Coeficientes del intercepto en los modelos de coeficientes fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

En el caso de la variable del DIA2, la estimación bayesiana muestra mayor estabilidad en los tres modelos presentados, las demás estimaciones consideradas muestran un comportamiento similar muy estable con un leve aumento en el modelo con mayor cantidad de efectos aleatorios.

Gráfico 6  
Coeficientes de DIA2 en los modelos de coeficientes fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación

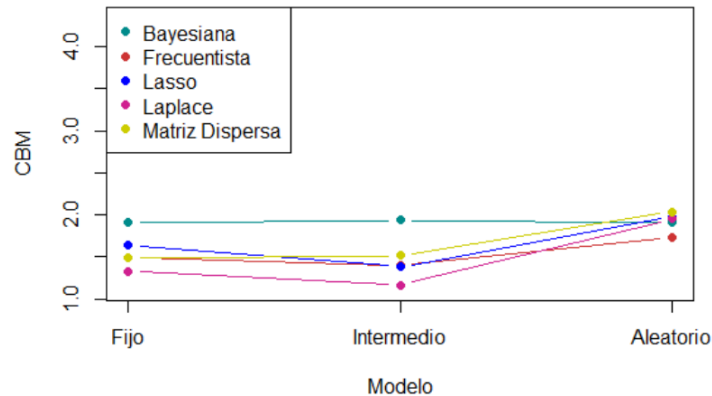


Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

En el gráfico 7 se presenta caso de la variable del DIA3, la estimación bayesiana muestra mayor estabilidad en los tres modelos presentados, pero visualmente un valor levemente mayor en los coeficientes, las demás estimaciones consideradas

muestran un comportamiento similar con un leve aumento en el modelo con mayor cantidad de efectos aleatorios.

Gráfico 7  
Coeficientes de DIA3 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación

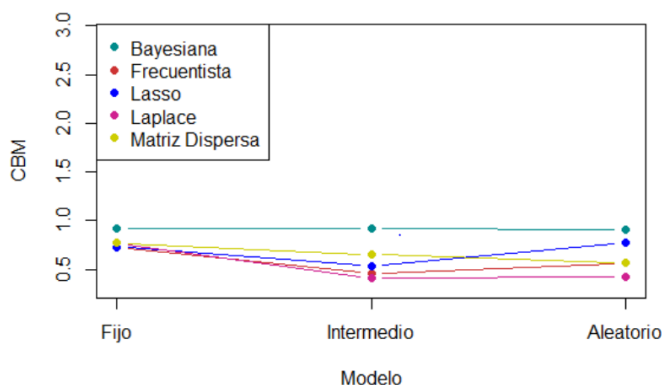


Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Para la variable M1, en el gráfico 8, la estimación bayesiana muestra mayor estabilidad en los tres modelos presentados, pero visualmente con un valor levemente mayor en el coeficiente; las demás estimaciones consideradas muestran un comportamiento similar con leves diferencias en el modelo con mayor cantidad de efectos aleatorios.

En este caso resalta que los valores del coeficiente del modelo con efectos fijos son similares en todas las estimaciones, pero los valores de los coeficientes a los cuales se les incorpora una mayor cantidad de efectos aleatorios no, y se ven un poco más separados.

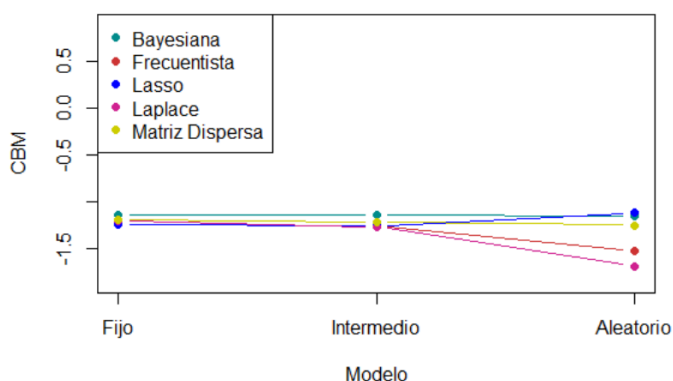
Gráfico 8  
Coeficientes de M1 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

En el caso de M4, las estimaciones frecuentista y Laplace presentan una disminución en el modelo con mayor cantidad de efectos aleatorios, en los demás casos – bayesiana, LASSO y con matriz de dispersión – se observa un comportamiento muy estable en los tres modelos estimados.

Gráfico 9  
Coeficientes de M4 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación

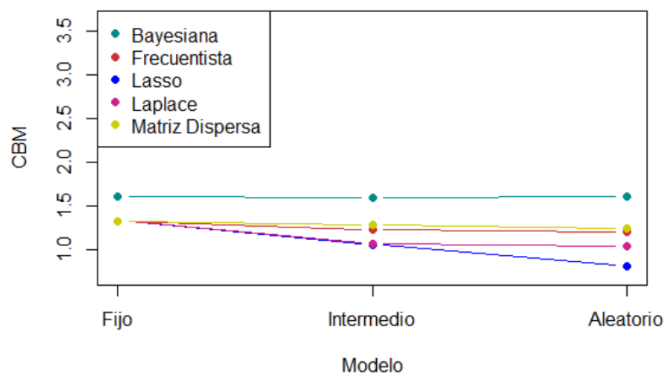


Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Para la variable M5, las estimaciones de los coeficientes bayesiana – visualmente presente el valor del coeficiente más alto –, frecuentista y matriz de dispersión se muestran muy estables, en el caso de Laplace y LASSO muestra una leve

disminución en la estimación en el modelo con la mayor cantidad de coeficientes aleatorios.

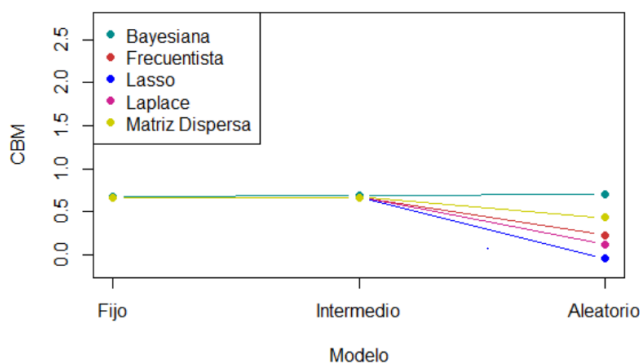
Gráfico 10  
Coeficientes de M5 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

En el caso de la variable M6, las estimaciones son similares para el modelo de efectos fijos y el que incorpora alguno de los coeficientes de efectos aleatorios, las estimaciones bayesianas son estables en los tres modelos, en los demás casos se tiene una encogimiento o disminución en el valor del coeficiente en el modelo con todos los coeficientes aleatorios.

Gráfico 11  
Coeficientes de M6 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación

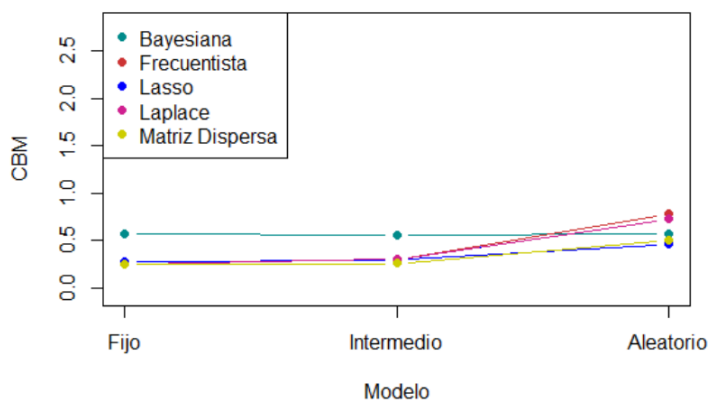


Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

En el gráfico 12 se presenta el comportamiento de los coeficientes de la variable M9, la estimación bayesiana muestra una mayor estabilidad en los modelos presentados; en el caso de las demás estimaciones hay un aumento en el valor del coeficiente estimado en el modelo de efectos aleatorios completo.

Además, la estimación de los coeficientes del modelo de efectos fijo es muy similar en todos los casos exceptuando la estimación bayesiana.

Gráfico 12  
Coeficientes de M9 en los modelos de coeficientes fijos y efectos aleatorios, según tipo de estimación

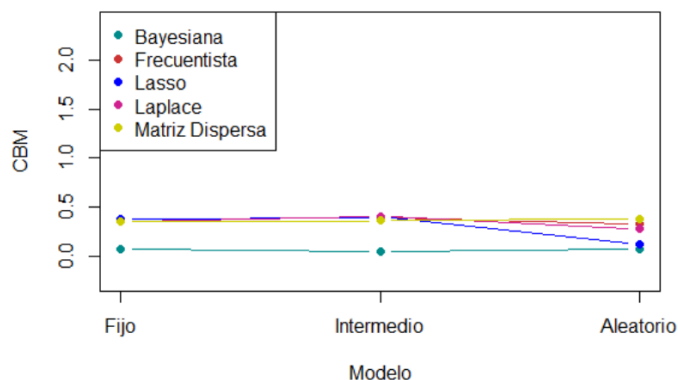


Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

En el caso de la variable M10, la estimación bayesiana se mantiene estable en los tres modelos de interés y visualmente muestra un valor inferior a las demás alternativas en los modelos de efectos fijos y los que incorporan algunos efectos aleatorios, en los demás casos se muestra una leve diferencia entre el modelo de efectos fijos y el modelo con todos los coeficientes aleatorios.

Gráfico 13

Coefficientes de M10 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación

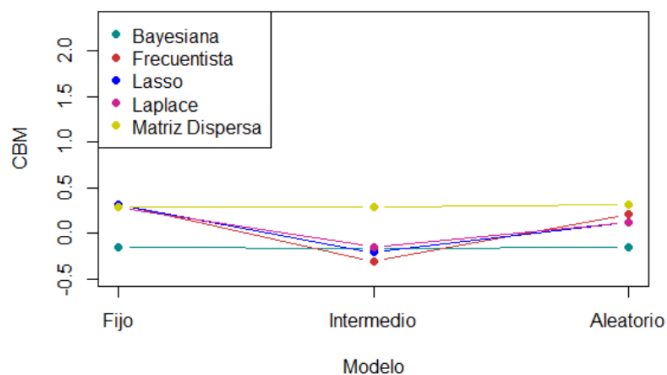


Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Finalmente, en el caso de la variable M11, se tiene que las estimaciones bayesianas y la que considera la matriz de dispersión muestran una mayor estabilidad, pero siendo la bayesiana con un valor visualmente menor y la matriz de dispersión con un valor visualmente mayor, en los casos frecuentista, LASSO y Laplace se tiene una leve disminución en el modelo que incorpora algunos efectos aleatorios, y un leve aumento en el coeficiente para el modelo de efectos aleatorios completo.

Gráfico 14

Coefficientes de M11 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

En general se tiene que los coeficientes presentan cambios a medida que se agregan coeficientes aleatorios al modelo; la estimación bayesiana es la que mostró una mayor estabilidad en las variables analizadas.

Los gráficos anteriores además de los resultados presentados en los cuadros 4, 5, 6, 7 y 8 permiten visualizar de manera más simple el posible fenómeno de *shrinkage* en el caso de estudio.

#### **5.4 Del estudio de simulaciones**

Para el presente análisis se realizan dos ejercicios de simulación, el primero se realizó para seleccionar el mejor valor de *Lambda* a utilizar en el modelo LASSO, mismo que permitió a partir del criterio de experto y con base en los objetivos del estudio determinar el mejor valor.

Como se ha mencionado en el presente documento, utilizando las técnicas de validación y los estadísticos AIC y BIC, se concluye que, para el modelo de interés, es el valor de  $\lambda=75$  el que se debe utilizar.

El segundo análisis utilizando simulaciones, corresponde al detallado en la metodología, mismo que desagrega en distintos escenarios el análisis de los modelos, este, resulta de utilidad pues permite identificar en cuáles casos, mencionados por diversos autores, se presenta el fenómeno de *shrinkage*; además gráficamente le brinda al investigador mayor información para que este pueda determinar si las estimaciones son estables o no, es decir, si hay o no cambios que pueden ser considerados importantes en los valores.

Por otra parte, el estudio de simulaciones puede también dar indicios de que sean requeridas más repeticiones para una conclusión más acertada.

Los resultados obtenidos en las simulaciones para los estadísticos SESGO y CME se presentan gráficamente en el anexo 6.

A continuación, se presenta para las variables DIA3, M1, M5 y M10 el análisis, así como la representación gráfica de las estimaciones de los efectos fijos, en los distintos escenarios de interés.

La variable DIA3 representa en el experimento, el momento en el cual los sujetos reciben la sustancia anfetamina, además de que es cuando se diferencian los sujetos en dos grupos. Es de interés determinar posibles cambios en las estimaciones de los coeficientes de esta variable.

En la imagen 3 se presenta el comportamiento de los coeficientes para la variable DIA3, este se muestra similar en los escenarios que involucran cambio en la variabilidad en los datos y distintos tamaños de muestra.

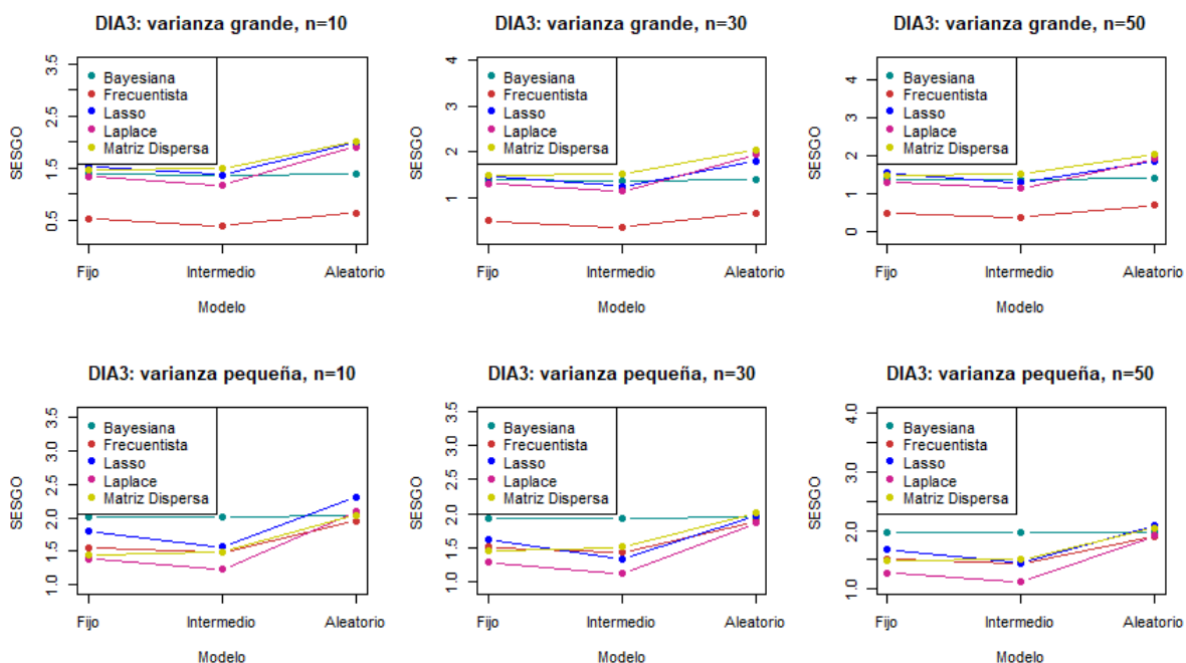
En todos los escenarios analizados, la estimación bayesiana es la que muestra una mayor estabilidad y valores más altos del sesgo cuando se considera una varianza pequeña.

Al considerar varianza pequeña (gráficas inferiores), las estimaciones frecuentista y bayesiana presentan mayor estabilidad en los coeficientes y por ende en los valores del sesgo, sin embargo, las otras alternativas presentan un comportamiento similar, con un aumento en los sesgos del modelo que incluye todos los efectos aleatorios. Al considerar varianza grande (gráficas superiores) son las estimaciones frecuentista y bayesiana las que presentan una mayor estabilidad en los modelos estimados, visualmente la estimación frecuentista para las demás alternativas se percibe un aumento en el sesgo a medida que se incorporan coeficientes aleatorios, pero menos pronunciada que cuando se tiene una varianza pequeña.

Resalta el caso de que los distintos tamaños de muestra parecieran no presentar diferencias entre los valores del sesgo, pues las gráficas muestran comportamientos similares.

Imagen 3

SESGO del coeficiente DIA3 en los modelos de coeficientes fijos y efectos aleatorios, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

En el caso de la variable M1, corresponde al primer minuto del experimento, donde el sujeto es colocado de manera aislada en un lugar ajeno y nuevo para él. Aquí el sujeto inicia el proceso de adaptación con el entorno, por lo que resulta de interés analizar esta variable.

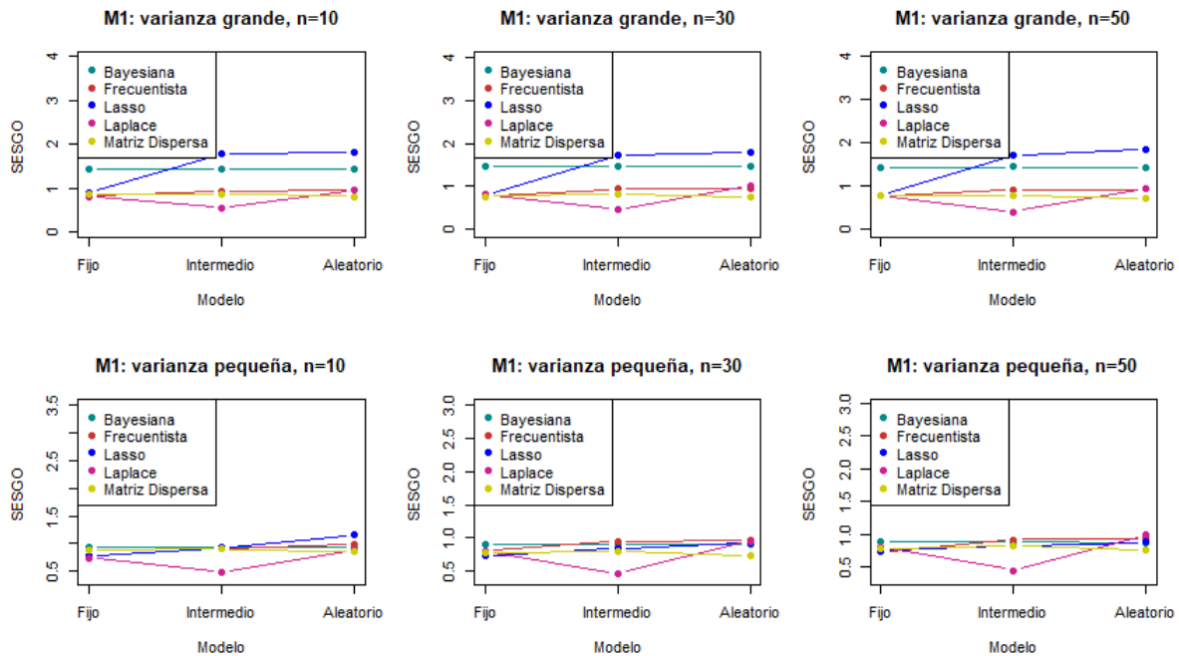
En la imagen 4 se muestra el comportamiento de la variable M1, en general se observa una mayor estabilidad para las alternativas bayesiana, matriz de dispersión y frecuentista.

Al considerar el escenario de varianza pequeña, las estimaciones muestran mayor estabilidad y valores similares, con excepción de la estimación Laplace, que en el modelo que incorpora efectos aleatorios, arroja valores más bajos en los tres tamaños de muestra considerados. En el caso de varianza grande se visualizan los valores de sesgo más separados entre las alternativas con un aumento en el modelo de coeficientes aleatorios completo. Para la alternativa Laplace se repite la disminución en el segundo modelo estimado.

Los gráficos muestran comportamiento similar en los diferentes tamaños de muestra considerados.

Imagen 4

SESGO del coeficiente M1 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

La variable M5 es el primer minuto del segundo día del experimento, nuevamente el sujeto se enfrenta al entorno al que ya había sido expuesto, pero con la variante que se le colocó una inyección con solución salina, por ende, resulta de interés para el investigador considerar el comportamiento de esta variable.

La siguiente imagen presenta el comportamiento de la variable M5, este es similar en los escenarios que involucran distintos tamaños de muestra y variabilidad en los datos.

En general se tiene que en las estimaciones bayesiana y con matriz de dispersión son las más estables.

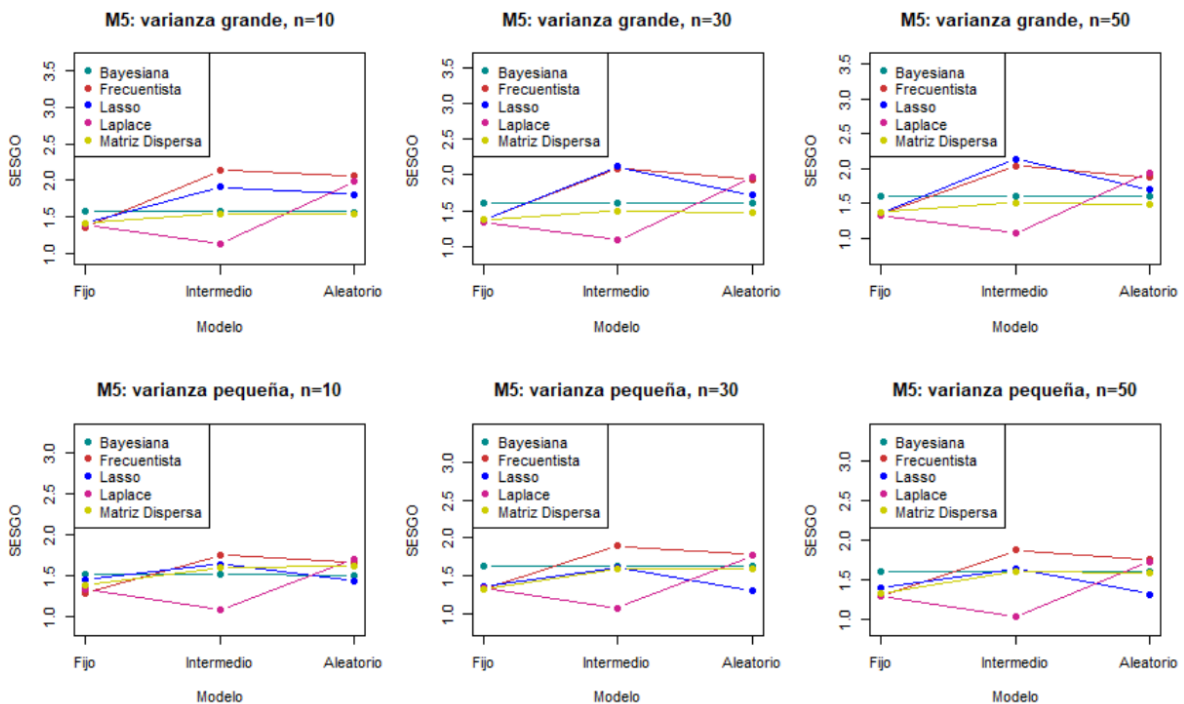
Cuando la varianza es grande se tiene que Laplace muestra un decrecimiento en el sesgo cuando se estima el modelo de incorporar algunos coeficientes aleatorios,

caso contrario LASSO y frecuentista muestran un aumento en el valor del sesgo en el modelo intermedio. Cuando la varianza es pequeña, la estimación con el método Laplace muestra nuevamente disminución en el modelo intermedio, la frecuentista un leve aumento.

En relación con el tamaño de muestra utilizado, se observa comportamientos muy similares entre los tamaños de muestra propuestos.

Imagen 5

SESGO del coeficiente M5 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

En la imagen 6 se presenta las estimaciones de los sesgos para la variable M10 de los distintos escenarios propuestos. Esta es de las variables que en la mayoría de los modelos completos de efectos aleatorios para las alternativas de estimación no presentó significancia estadística, además de que mide la cantidad de vocalizaciones en los últimos minutos para que concluya el experimento, es el minuto 4 del día tres en el cual se suministra al sujeto la anfetamina o solución

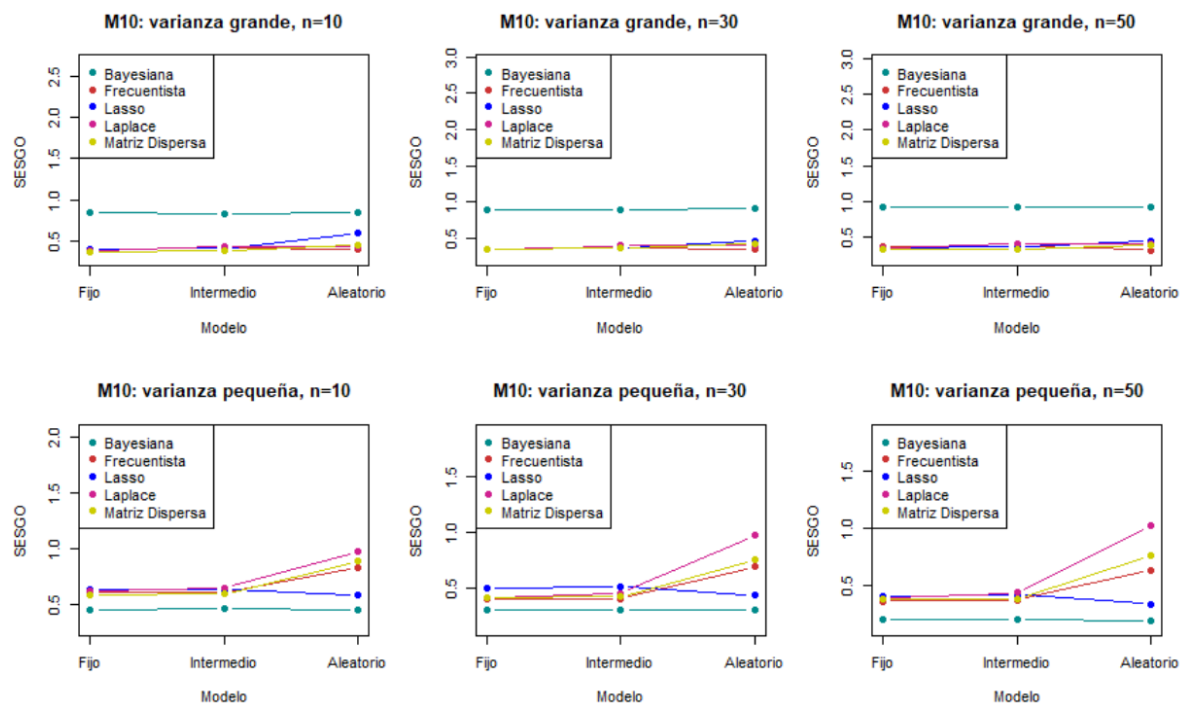
salina, es importante para el investigador conocer el comportamiento de esta variable.

Al considerar la varianza grande, todas las alternativas presentadas son muy estables, sin embargo, a pesar de que las estimaciones bayesianas también son estables, son las que presentan un valor mayor en el sesgo. Por otra parte, al considerar una varianza pequeña se tiene un aumento en el valor del sesgo en todas las alternativas, siendo más notorio en el modelo de efectos aleatorios completo, con excepción de la estimación bayesiana, la cual se mantiene muy estable en los escenarios planteados. En ambos casos la estimación bayesiana muestra una mayor estabilidad en los valores obtenidos en las estimaciones.

El tamaño de muestra presenta comportamientos similares en las estimaciones.

Imagen 6

SESGO del coeficiente M10 en los modelos de coeficientes fijos y efectos aleatorios, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

A continuación, una breve descripción del comportamiento de las demás variables incorporadas en los modelos, los gráficos se presentan en el anexo 6.1:

Para el intercepto (Imagen A5), cuando la varianza es grande, se tiene un leve aumento en el modelo con todos los coeficientes aleatorios, sin embargo, todas las estimaciones presentan un comportamiento similar, por otra parte, cuando la varianza es pequeña, las estimaciones se comportan de manera muy similar y se observa que los valores del sesgo son levemente superiores en la estimación bayesiana manteniéndose constante, pero sin alejarse de las demás. Gráficamente se observa que es la estimación LASSO con varianza grande la que presenta mayor cambio entre el modelo de efectos fijos y el completo de efectos aleatorios.

Para DIA2 (Imagen A8): presenta un comportamiento similar en los escenarios que involucran distintos tamaños de muestra y variabilidad en los datos. En general se tiene un cambio en los valores del sesgo cuando se tiene el modelo de coeficientes aleatorios completo, siendo la estimación bayesiana la más estable.

Específicamente, cuando la varianza es grande, se tiene un leve cambio en el modelo con todos los coeficientes aleatorios, por lo que el tamaño de muestra pareciera no mostrar diferencias. Cuando se tiene varianza las estimaciones también presentan similitudes, con algunos leves descensos en el sesgo, sobresale el caso de la estimación con matriz de dispersión que la cual únicamente cuando considera varianza pequeña y  $n=50$  se aleja del comportamiento de las demás estimaciones.

La variable M4 (Imagen A9) presenta mucha estabilidad a nivel general, el tamaño de muestra no parece presentar diferencias en el comportamiento de los valores sesgo, únicamente sobresale la estimación LASSO, misma que presenta una disminución en el valor del sesgo para el modelo completo, y parece ser mayor cuando se considera varianza grande.

La estimación bayesiana presenta mucha estabilidad y se aleja de las demás estimaciones cuando se considera varianza grande.

En el caso de M6 (Imagen A10), en general se tiene que a medida que se incorporan coeficientes de efectos aleatorios en el modelo, se presentan cambios en el valor del sesgo, lo que puede considerarse presencia de *shrinkage*. Son la estimación bayesiana y la estimación con matriz de dispersión las que presentan mayor estabilidad en todos los escenarios considerados. Además, en la estimación LASSO muestra a nivel gráfico un aumento en el valor del sesgo cuando se considera varianza grande y una disminución cuando se considera varianza pequeña. El tamaño de la muestra no pareciera mostrar cambios en el comportamiento; y son los escenarios de varianza pequeña los que muestran mayor estabilidad al incorporar coeficientes aleatorios en los modelos.

La variable M9 (Imagen A11) en todos los escenarios hay un aumento en el valor del sesgo para los modelos que incorporan todos los coeficientes aleatorios. Es la estimación bayesiana la que presenta una mayor estabilidad en los modelos considerados, curiosamente la estimación LASSO en esta variable es de las más estables en los modelos y escenarios considerados, el tamaño de muestra parece no mostrar diferencia a nivel gráfico.

Finalmente, al considerar varianza grande se tiene que los cambios en el sesgo entre modelos de únicamente efectos fijos y al modelo completo de efectos aleatorios son menos notorios que los cambios que se presentan al considerar una varianza pequeña.

En el caso de la variable M11 (Imagen A12) se tiene un comportamiento similar en los escenarios planteados. A medida que se incorporan coeficientes de efectos aleatorios en el modelo completo se tiene un aumento en los valores del sesgo. A nivel gráfico se puede apreciar que los escenarios en los cuales se consideró una varianza grande los aumentos en el valor del sesgo son levemente menores que los observados en los modelos para los cuales se consideró varianza pequeña.

El tamaño de muestra parece no mostrar diferencias entre los modelos, y es en las estimaciones bayesiana, LASSO y con matriz de dispersión donde se muestra mayor estabilidad en los sesgos de esta variable.

Los gráficos con las estimaciones de los errores cuadráticos medios (CME) se presentan en el anexo 6.2 de la imagen A13 a la A22, ambos inclusive.

A partir de los resultados anteriores que muestran el comportamiento de las estimaciones de los coeficientes fijos en los distintos escenarios propuestos, los cambios sobresalen cuando se considera el aumento en la cantidad de los efectos aleatorios en los modelos, así como cambios en la varianza, el tamaño de muestra no parece gráficamente mostrar diferencias entre los valores de los coeficientes, así como los del SESGO. Lo anterior permite interpretar al investigador que se tiene *shrinkage* en los modelos analizados.

A pesar de que se presentaron tres alternativas para la estimación de los modelos multinivel además de las frecuentista y bayesiana inicialmente consideradas, a partir de los resultados anteriores se concluye que es la alternativa bayesiana la que muestra una mayor estabilidad en las estimaciones a partir de la base de datos madre y de los resultados obtenidos en el estudio de simulaciones; sin embargo, es importante indicar que en este caso el modelo completo de efectos aleatorios dos de las variables más importantes para el investigador no resultaron significativas, pero son interpretados los coeficientes obtenidos a partir de la estimación bayesiana.

Se debe recordar que la función de enlace utilizada el logaritmo, por lo que para la interpretación los coeficientes deben exponenciarse; también, debe considerarse el tipo de codificación utilizada en las variables, pues todas son dicotómicas.

Se tiene que:

DIA2:  $\exp(-0.7479) = 0.47$  Si el día en que se realizan las mediciones es el día 2, en la cantidad de vocalizaciones se tendrá un aumento promedio de 0.47 en relación con el día 1.

DIA3:  $\exp(1.9253) = 6.86$  Si el día en que se realizan las mediciones es el día 3, en la cantidad de vocalizaciones se tendrá un aumento promedio de 6.86 en relación con el día 1.

M1:  $\exp(0.9063) = 2.48$  Si las mediciones se realizan en el minuto 1 en el día 1, la cantidad de vocalizaciones aumenta en 2.48.

M4:  $\exp(-1.1401) = 0.32$  Si las mediciones se realizan en el minuto 10 en el día 1, la cantidad de vocalizaciones aumenta en 0.32.

M5:  $\exp(1.6020) = 4.96$  Si las mediciones se realizan en el minuto 1 en el día 2, la cantidad de vocalizaciones aumenta en 4.96.

M6:  $\exp(0.6833) = 1.98$  Si las mediciones se realizan en el minuto 4 en el día 2, la cantidad de vocalizaciones aumenta en 1.98.

M9:  $\exp(0.5688) = 1.77$  Si las mediciones se realizan en el minuto 1 en el día 3, la cantidad de vocalizaciones aumenta en 1.77.

M10:  $\exp(0.0569) = 1.06$  Si las mediciones se realizan en el minuto 4 en el día 3, la cantidad de vocalizaciones aumenta en 1.06.

M11:  $\exp(-0.1558) = 0.86$  Si las mediciones se realizan en el minuto 7 en el día 3, la cantidad de vocalizaciones aumenta en 0.86.

En los casos de las variables medición que corresponden a los minutos de los días en los cuales se contabilizaron las vocalizaciones, si se deseara conocer la diferencia entre la media de vocaciones de la primera medición y la onceava, se debe calcular la resta de dichos valores medios.

Ahora bien, en relación con la estimación de los efectos aleatorios y para su respectiva interpretación se consideraron los valores obtenidos en las estimaciones del paquete STATA 15 a pesar de que no se obtuvo la convergencia de los tres modelos de interés; sin embargo, se considera relevante dar interpretación a estos, pues mostraron mayor estabilidad y similitud a las estimaciones frecuentistas que los obtenidos con la plataforma de programación R. Es importante destacar que el objetivo de la interpretación de los valores es dar mayor provecho al análisis.

En el modelo de efectos aleatorios, la estimación de la varianza del coeficiente de regresión para la variable M1 según sujeto, es 1.4091, así, la desviación estándar es 1.1870. Considerando el coeficiente (también obtenido por STATA 15) de la variable  $2.2178 = e^{0.7965}$ , entonces se puede tomar valores entre 1.0307 y 3.4048 – donde el intervalo no contiene el valor cero –, eso nos indica cuál podría ser la relación entre la cantidad de vocalizaciones emitidas por los sujetos y la primer medición del día 1 del experimento, además de cuánto podría cambiar de sujeto a sujeto. Para la toma de decisiones se debe decidir si son suficientes los resultados obtenidos o se necesita implementar otros mecanismos de análisis.

En el modelo de efectos aleatorios, la estimación de la varianza del coeficiente de regresión para la variable M5 según sujeto, es 1.2568, así, la desviación estándar es 1.1211. Considerando el coeficiente (también obtenido por STATA 15) de la variable  $4.1803 = e^{1.4304}$ , entonces puede tomar valores entre 3.0593 y 5.3014 – donde el intervalo no contiene el valor cero –, eso nos indica cuál podría ser la relación entre la cantidad de vocalizaciones emitidas por los sujetos y la primer medición del día dos del experimento, además de cuánto podría cambiar de sujeto a sujeto. Para la toma de decisiones se debe considerar si son suficientes los resultados obtenidos o se necesita implementar otros mecanismos de análisis.

En el modelo de efectos aleatorios, la estimación de la varianza del coeficiente de regresión para la variable M11 según sujeto, es 0.8178, así, la desviación estándar es 0.9043. Considerando el coeficiente (también obtenido por STATA 15) de la variable  $1.0351 = e^{0.0345}$ , entonces puede tomar valores entre 0.1308 y 1.9394– donde el intervalo no contiene el valor cero –, eso nos indica cuál podría ser la relación entre la cantidad de vocalizaciones emitidas por los sujetos y la séptima medición del día 3 del experimento, además de cuánto podría cambiar de sujeto a sujeto. Para la toma de decisiones se debe decidir si son suficientes los resultados obtenidos o se necesita implementar otros mecanismos de análisis.

En ambos modelos se calcula el valor de la correlación intraclase, para el modelo de efectos fijos el *rho* es:  $[0.3248/(0.3248 + 0.2242)] = 0.5917$ , valor que se considera alto y muestra la homogeneidad a lo interno de los grupos.

Para el modelo de efectos aleatorios completo, el *rho* es:  $[(0.2562 + 1.4091 + 1.2568 + 0.8178) / (0.2562 + 1.4091 + 1.2568 + 0.8178 + 0.3169)] = 0.9219$ , un valor muy cercano a 1, lo cual es indicio de que hay una alta homogeneidad a lo interno de los grupos.

## 6. Conclusiones

### 6.1 Conclusiones relevantes del caso de estudio

Los resultados iniciales obtenidos en el análisis descriptivo de los datos mostraron que en el día tres del experimento, la cantidad de vocalizaciones era mayor a los días anteriores especialmente al suministrarle la anfetamina a los sujetos, sin embargo, esto se esperaba debido a que en la teoría ya se indicaba que la anfetamina es un potenciador de cierto comportamiento eufórico.

Los análisis gráficos arrojan información de que la distribución de los datos podría tener un comportamiento asimétrico y que podría ser modelado con una distribución Poisson.

Para este análisis resultó de gran relevancia contar con el experto de datos quien por su *expertise* facilitó la comprensión de la información con la cual se contaba así como los lineamientos para en la generación de todos los posibles modelos, la selección del modelo con el cual se trabajó, y las variables a las cuales se les estimó los coeficientes aleatorios.

Es importante mencionar que el modelo con el cual se trabajó estaba conformado únicamente por variables dicotómicas, además que la cantidad de observaciones con las que se contaba hizo compleja la estimación de los modelos pues se tenían pocos registros.

Las variables de interés para el investigador eran el día tres y las mediciones de ese día pues corresponden al momento del experimento cuando se suministró la anfetamina en los sujetos, las variables consideradas en el modelo de estudio son M9, M10 y M11, las cuales corresponden a los minutos 1,4 y 7 del día tres.

Como parte de los resultados obtenidos, particularmente son las variables de mayor interés las que pierden significancia en los modelos al incluir una mayor cantidad de efectos aleatorios. La variable M9 pierde significancia en tres de las cinco alternativas, M10 pierde significancia en tres de las cinco alternativas y M11 pierde

significancia en dos de las cinco alternativas, DIA3 mantienen su significancia en todas las alternativas consideradas. En este punto es importante resaltar que es en la alternativa de matriz de dispersión en la cual las variables de interés del investigador se mantuvieron significativas.

Al comparar visualmente los coeficientes de efectos fijos de los modelos sin efectos aleatorios *versus* los coeficientes de efectos fijos en los modelos de coeficientes aleatorios completos, se visualiza un cambio que podría reflejar la presencia de *shrinkage*.

Al exponenciarse únicamente los coeficientes para su posterior interpretación se tiene que la variable DIA3 y M5 tiene los coeficientes más altos, y M4 el más bajo, lo demás se mantienen similares.

En relación con los valores del *rho*, se esperaba que estos fuesen altos, pues la misma estructura en la cual fueron conformados los grupos, buscó que fuesen lo más similar posibles entre sí.

En el caso de los efectos aleatorios, los intervalos de los coeficientes no incluyen el cero, lo que nos permite poder entender la posible relación entre la cantidad de vocalizaciones y los momentos en los cuales se realizaron las mediciones. Los intervalos de M1, M5 y M11 corresponde a (1.0307, 3.4048), (3.0593, 5.3014) y (0.1308, 1.9394) respectivamente. El intervalo de la variable M11 es más pequeño de los tres estimados y es de las variables de mayor interés para el investigador.

En el caso de la variable M1, como se ha indicado corresponde a la primera reacción del sujeto en el campo abierto en el día uno, pareciera que las vocalizaciones son más variables. En el caso de la variable M5, es el primer acercamiento al campo abierto con una inyección de placebo para acostumbrar al sujeto a la inyección, corresponde a las vocalizaciones del minuto uno del día dos, es decir las vocalizaciones apenas se coloca en el área del experimento. La variable M11 presenta un intervalo más pequeño, corresponde al minuto siete del día tres, casi al

finalizar el experimento en su totalidad y cuando ya se había suministrado la sustancia de interés y se tenían varios minutos en el campo abierto.

En relación con las alternativas consideradas, es importante considerar que es la estimación bayesiana la que presenta mayor estabilidad en las estimaciones con los datos de la base madre, así como en el estudio de simulaciones; sin embargo, también es la alternativa que presenta problemas es la estimación de los efectos aleatorios en la plataforma de programación R y problemas de convergencia en el programa STATA 15. Es importante que el investigador en este punto, considere estos aspectos encontrados para análisis posteriores.

A partir de lo encontrado en la literatura, y que la estimación bayesiana había tenido cierta ventaja en relación con la estimación frecuentista en el análisis preliminar realizado, y considerando que la estimación Laplace tiene un enfoque bayesiano e incluye parámetros para mejorar las estimaciones al trabajar con modelos multinivel; a criterio de experto se esperaba que la alternativa Laplace tuviese mayor estabilidad en las estimaciones y que no se perdiera la significancia en las variables de interés del investigador al incorporar los efectos aleatorios, pero no fue lo que ocurrió.

En el caso de la alternativa LASSO se considera que para la estimación de los modelos específicos utilizando la base de datos madre es muy estable, sin embargo, en el estudio de simulación arrojó casos en los cuales el ponderador de ciertos coeficientes fijos fue cero, lo que en este caso particular no se buscaba. Por ende, es recomendable que se hagan estudios de simulaciones más amplios o se consideren valores de *Lambda* más bajos, siempre que sea seleccionado con criterios razonables. Así, el estudio de simulaciones arrojó un resultado importante que debe ser considerado: un mismo valor *Lambda* no debe ser considerado como el óptimo en distintos modelos, más si se consideran distintos escenarios.

En el caso de la alternativa de matriz de dispersión, en términos de especificación del modelo es más compleja que las demás alternativas, por lo que se debe tener cierta pericia en el fenómeno de estudio, para “aprovechar este modelo”; sin

embargo – en este caso se hizo la especificación de los parámetros mínimos necesarios – es la única en la que las variables de interés del experto resultaron significativas, por lo que puede ser una buena opción para realizar estimaciones, siempre que se realice el análisis previo de este paquete para obtener un mejor ajuste.

Se recomienda en casos similares al expuesto, se utilice modelos de enfoque bayesiano con algún ajuste en la estimación así como la alternativa de matriz dispersa. Por otra parte, es siempre importante también considerar varios programas, o paquetes de un mismo programa o *software*, para analizar la consistencia entre los resultados obtenidos, y determinar qué tan similares o distintos son los resultados, así como posibles falencias.

Para el estudio de simulaciones y a partir de la teoría revisada, se esperaba que mostrara de manera clara cuáles eran los escenarios en los que se tenía el fenómeno de *shrinkage*, al menos en este caso de estudio, que permitiera identificar cuáles de las alternativas presentadas mostraban un buen ajuste, así como una mayor rapidez de las estimaciones.

El estudio de simulaciones arrojó información adicional en la alternativa LASSO que fue mencionada anteriormente y permitió analizar visualmente el comportamiento de cada alternativa en los escenarios planteados. También, sugiere que el tamaño de muestra no es un escenario que condicione la presencia del efecto *shrinkage*, pues el comportamiento en las estimaciones del SESGO y CME es muy similar con las muestras de 10, 30 y 50 en todas las alternativas consideradas.

Así mismo, el tipo de varianza (grande o pequeña) si es un factor que permite identificar de manera más clara en los gráficos la presencia de *shrinkage*, y en términos generales se tiene que, en las estimaciones con varianza pequeña, son en la mayoría de los coeficientes más estables, es decir, el SESGO se mantiene constante, caso contrario, las estimaciones con varianza grande muestran cambios en los gráficos del SESGO.

Por otra parte, se tiene que la estimación bayesiana presenta una mayor estabilidad en las estimaciones de los coeficientes, ante la estimación frecuentista. En el caso del SESGO se mantienen constantes en los modelos a pesar de la inclusión de más coeficientes aleatorios.

A partir de lo anterior, se concluye que las alternativas propuestas son un mecanismo que permite solventar la problemática del *shrinkage* en el presente análisis; de las cuales sobresalen las bayesiana y el ajuste con Laplace pues son muy estables, presentan valores de SESGO y errores cuadráticos más bajos, además de la estimación que incorpora matriz de dispersión responde a los intereses del experto en neurociencia.

En este caso puntual, se tiene que todas las alternativas mostraron resultados favorables, y a partir de las referencias bibliográficas son buenas alternativas en casos de análisis similar a este, a pesar de que algunas de las variables importantes para el experto en neurociencias no mostraron significancia estadística, los valores del SESGO y CME mostraron cierta estabilidad en las variables analizadas.

Es importante indicar también, que en las referencias bibliográficas se mencionaron otras alternativas como el modelo de localización general (GLOM), el método de descomposición de Cholesky o "Stochastic Search Variable Selection" (SSVS) que no fueron analizadas en este caso, por lo que no se deben considerar como únicas las alternativas a utilizar.

La alternativa que se escoja dependerá de las características que tenga el conjunto de datos, por lo que no se puede concluir de manera certera con una recomendación puntual si no mencionar los hallazgos positivos y los resultados particulares a los que se les debe prestar atención en cada caso.

En este caso se utilizó únicamente el SESGO y el error cuadrático medio para determinar la mejor alternativa, sin embargo, se recomienda que en estudios posteriores se utilicen otros estadísticos en la toma de decisiones.

## 6.2 Hallazgos adicionales

A partir de consultas a expertos, indican que la cantidad de registros por grupo y en los niveles considerados es baja, lo que podría provocar problemas en la convergencia de las estimaciones, por ello es muy importante que los investigadores consideren este aspecto, pues, aunque cuando se trabaja con datos reales esto es lo que ocurre, se deben buscar estructuras simples que logren representar lo que se desea estudiar.

Para la selección del valor *Lambda*, el proceso consistió en un vector con valores de 0 a 500 a partir de referencias consultadas, pero se podría considerar también intervalos más cercanos entre sí, es decir, no usar valores con intervalos de cinco, sino más pequeños como intervalos de tres o dos, y un valor máximo diferente, en valor de 500 se utilizó porque era la recomendación en el caso consultado, y en este caso el intervalo contenía el valor óptimo, pero no se debe considerar como una regla única.

En la selección del valor de *Lambda* se realizó con los datos de la base madre, por lo que puede ocurrir que, en el mismo modelo, pero bajo otro conjunto de datos el *Lambda* óptimo sea otro. En la regresión LASSO, a pesar de que el valor de *Lambda* seleccionado (75) permitió la estimación de todos los coeficientes de interés en el modelo, en algunos de los casos en el estudio de simulaciones, las estimaciones generadas de los coeficientes fueron cero, es decir, en valor de *Lambda* seleccionado realizó una ponderación y selección generando una posible conclusión de que dicha variable no aportaba información al modelo, desde la perspectiva matemática. Únicamente en los casos de las estimaciones para el modelo de efectos aleatorio completo con varianza grande y  $n=30$ , y el modelo de efectos fijos con varianza grande y  $n=30$ , las estimaciones son distintas a cero para todos los coeficientes. Por lo anterior y a pesar de que se presentaron errores bajos, se recomienda considerar otro valor de *Lambda* más pequeño y hacer las respectivas estimaciones, o realizar un estudio de simulaciones con el valor de *Lambda* de manera más rigurosa.

En el caso de la estimación bayesiana, debido a que se carece de informativas previas, se utilizan las previas del paquete `MCMCglmm`, por ende, estimaciones que cuenten con información, podrán generar estimaciones más precisas, sin embargo, esta información debe ser generada por expertos en el tema.

Es importante indicar que algunos de los paquetes utilizados cuentan con componentes "*For advanced users*" que no fueron utilizados, ya que se decidió utilizar los parámetros requeridos mínimos en el análisis pues se está en una fase inicial y en el caso de incluir o considerar información se podrían generar resultados erróneos si no se tiene certeza de la misma.

La plataforma de programación R mostró una ventaja analítica en relación con el paquete STATA 15, pues este último no logró convergencia en modelos bayesianos con mayor cantidad de efectos aleatorios utilizando los datos de la base madre en las estimaciones preliminares, es importante rescatar que las estimaciones en los modelos que sí se lograron estimar con STATA 15 eran las mismas que en la plataforma de programación R, con diferencias en las estimaciones hasta el cuarto decimal, en el caso de la plataforma de programación R también se presentaron casos en los cuales los paquetes tampoco lograron la convergencia del modelo; sin embargo, es necesario que se conozca el método de estimación considerado por cada programa, así como sus ventajas y desventajas.

Los paquetes utilizados presentaron rapidez en la convergencia, los tiempos de estimación en el estudio de simulaciones también, específicamente en `glmmr` y `hglm2`.

Para la generación de bases de datos en los distintos escenarios se programó una selección con muestreo de reemplazo condicionada a la base de datos madre, los casos muestreados y graficados muestran un comportamiento similar. Sin embargo, en el presente caso únicamente se seleccionó aleatoriamente la variable de respuesta cantidad de vocalizaciones emitidas y se asignó de manera aleatoria también el tratamiento de anfetamina a la mitad de los sujetos. Por ello en casos de

estudio más complejos se debería utilizar algún mecanismo para que el proceso de la generación de las bases de datos más robusto.

Los resultados obtenidos en el análisis permiten comprender el fenómeno *shrinkage*, se debe tener claro que como una primera aproximación permite mejorar la comprensión del problema y así abrir las puertas a posteriores análisis, además de posibles mejoras al mismo.

## Bibliografía

- Ahmed, E. S., Fallahpour, S. (2012). Shrinkage estimation strategy in quasi-likelihood models. *Statistics and Probability Letters*. 82, 2170-2179.
- Alam, M., Ronnegard, and L., Shen, X. (2014). *Fitting Conditional and Simultaneous Autoregressive Spatial Models in hglm*. Submitted, 7(2): 5-18.URL: <https://journal.r-project.org/archive/2015/RJ-2015-017/RJ-2015-017.pdf>
- Aparicio, A., Morera, M. (2007). La convivencia del análisis multinivel para la investigación en salud: Una aplicación para Costa Rica. *Población y Salud en Mesoamérica*. 4(2), 1-23.
- Barnard, J., McCulloch, R. and Meng, X. (2000). MODELING COVARIANCE MATRICES IN TERMS OF STANDARD DEVIATIONS AND CORRELATIONS, WITH APPLICATION TO SHRINKAGE. *Statistica Sinica*. 10, 1281-1311.
- Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*: 67(1), 1-48. doi:10.18637/jss.v067.i01.
- Bolker, B. (2019). Getting started with the glmmTMB package.
- Bondell, H., Krishna, A. and Ghosh S. (2010). Joint Variable Selection for Fixed and Random Effects in Linear Mixed-Effects Models. *Biometrics*. 66 (4), 1069-1077.
- Bottai, M. and Orsini, N. (2004). Confidence intervals for the variance components of random-effects linear models. *The Stata Journal*. 4 (4), 429-435.
- Box, G.E.P and Tiao, G.C. (1973) Bayesian Inference in Statistical Inference. Addison-Wesley, P.C.

- Brenes, J.C., Sequeira, A. & Rojas, M. (2017-2019). *Motivación y plasticidad neuroconductual: Efectos cognitivos, emocionales y sociales del enriquecimiento ambiental aleatorio e impredecible en ratas*. Instituto de Investigaciones Psicológicas, Universidad de Costa Rica.
- Cai, B. and Dunson, D. (2006). Bayesian Covariance Selection in Generalized Linear Mixed Models. *Biometrics*. 62 (2), 446-457.
- Carlin B.P. and Louis, T.A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall/CRC.
- Carter, D. (1979). Comparison of different shrinkage formulas in estimating population multiple correlation coefficients. *Educational and Psychological Measurement*. 39.
- Chen, Z. (2003). Random Effects Selection in Linear Mixed Models. *Biometrics*. 59, 762 - 769.
- Cheng, Z., Liado, Z. & Schorfheide, F. (2015). Shrinkage Estimation of High-Dimensional Factor Models with Structural Instabilities. *Economic Studies Advance Access*. 21.
- Chipman, H., Kolaczyk, E. and McCulloch, R. E. (1997). Adaptive Bayesian wavelet shrinkage. *Journal of the American Statistical Association*. 92, 1413-1421.
- Cnaan, A., Laird, N. and Slasor, P. (1997). Tutorial in biostatistics: Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Statistics in medicine*. 16, 2349-2380.
- Cohen, M. (1998). Determining Sample Sizes for Survey with Data Analyzed by Hierarchical Linear Models. *Journal of Official Statistics*. 14 (3), 267-275.
- Congdon, P. (2003). *Applied Bayesian Modelling*. Wiley.

- Daniels, M. (2001). Shrinkage Estimators for Covariance Matrices. *Biometrics*. 57, 1173 - 1184.
- De la Cruz, Francisco. (2008). Modelos Multinivel. *Revista Peruana de Epidemiología*. 12 (3), 1-8.
- Bates , D., Maechler , M., Bolker, B. Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01
- Ejaz Ahmed, S., Hossain, S., Doksum, K. and You, J. (2007). SHRINKAGE, PRETEST AND ABSOLUTE PENALTY ESTIMATORS IN PARTIALLY LINEAR MODELS. *Australian & New Zealand Journal of Statistics*. 49(4), 435-454
- Fahrmeir, L., Kneib, T. and Konrath, S. (2010). Bayesian regularisation in structured additive regression: a unifying perspective on shrinkage, smoothing and predictor selection. *Statistics and Computing*. 20 (2), 203-219.
- Finch, W. et ál (2014). *Multilevel Models using R*. Series: Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences.
- Gaviria, J.L., y Castro M. (2005). *Modelos Jerárquicos Lineales*. (Serie Cuadernos de Estadística número 29). Madrid: La Muralla, S.A.
- Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2004). *Texts in Statistical Science. Bayesian Data Analysis*. Chapman & Hall/CRC.
- Gelman, A. and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Groll, A. (2017). glmmLasso: *Variable Selection for Generalized Linear Mixed Models by L1-Penalized Estimation*. R package version 1.5.1. <https://CRAN.R-project.org/package=glmmLasso>

- Gujarati, D.N. and Porter, D.C. (2010). *Econometría*. (5.<sup>a</sup> ed.). Mc Graw Hill. México D.F.
- Hadfield, J. D. (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software*: 33(2), 1-22. URL: <http://www.jstatsoft.org/v33/i02/>.
- Hamilton, L.C. (2004). *Statistic with STATA*. Thamson, Brooks/Cole.
- Hossain, S., Ejaz Ahmed, S., Doksum, K.A. (2015). Shrinkage, pretest, and penalty estimators in generalized linear models. *Statistical Methodology*. 24, 52-68
- Huang, Y., Liu, D. and Wu, H. (2006). Hierarchical Bayesian Methods for Estimation of Parameters in a Longitudinal HIV Dynamic System. *Biometrics*. 62, 413-423.
- Jarrod D Hadfield (2010). MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software*, 33(2), 1-22. URL <http://www.jstatsoft.org/v33/i02/>
- Kinney, S. and Dunson, D. (2006). Fixed and random effects selection in linear and logistic models. *Biometrics*. 63 (3), 690-698.
- Kreft, I.G.G., de Leeuw, J. and Aiken, L. (1995). The Effect of Different Forms of Centering in Hierarchical Linear Models. *Multivariate Behavioral Research*, 30 (1), 1-22.
- Kutner, M., Nachtsheim, C., Neter, J. and Li, W. (2005). *Applied Linear Statistical Models*. McGraw-Hill.
- Lee, Y., Nelder, J.A. (2006). Double hierarchical generalized linear models. *Appl. Statist.* 55 (2), 1-29.
- Longford, N.T. (1993). *Random Coefficient Models*. Oxford: Clarendon.

- López-González E. (1998). TRATAMIENTO DE LA COLINEALIDAD EN REGRESIÓN MÚLTIPLE. *Psicothema*, 10 (2), 491-507
- López, L., Franco, M., Barreto, P. (2007). Sobre la construcción del mejor predictor lineal insesgado (BLUP) y restricciones asociadas. *Revista Colombiana de Estadística*. 30 (1), 13-36.
- Lui, Q. (1993). *Laplace Approximations to Likelihood Functions for Generalized Linear Mixed Models*. (Tesis de Doctorado en Philosophy) Universidad Estatal de Oregon.
- Lykou, A., Ntzoufras, I. (2013). On Bayesian lasso variable selection and the specification of the shrinkage parameter. *Statistics and Computing*. 23, 361-390.
- Milliken, G., Johnson, D. (2002). *Analysis of Messy Data. Volumen III: Analysis of covariance*. Chapman & Hall/CRC.
- Mora-Escobar, H. (2011). *Factorizaciones de Cholesky, matrices definidas y semidefinidas positivas*. Recuperado de la dirección electrónica [http://hectormora.info/fact\\_chol.pdf](http://hectormora.info/fact_chol.pdf) [Consulta 1 may. 2017].
- Morris, J., Baggerly, K. and Coombes, K. (2003). Bayesian Shrinkage Estimation of the Relative Abundance of mRNA Transcripts Using SAGE. *Biometrics*. 59, 476-486.
- Ntzoufras, I. (2009). *Bayesian Modeling Using WinBUGS*. Wiley.
- Ogden, H. (2018). *glmmsr: Fit a Generalized Linear Mixed Model*. R package version 0.2.2. <https://CRAN.R-project.org/package=glmmsr>
- Paccapelo, M. (2015). *Modelos de Selección Genómica para Caracteres Cuantitativos basados en Marcadores Moleculares aplicados al mejoramiento de maíz*. Tesis para optar por el título de Magíster en Estadística Aplicada. Universidad de Córdoba Colombia.

- Pan, J. and Huang, Ch. (2014). Random effects selection in generalized linear mixed models via shrinkage penalty function. *Statistics and Computing*. 24 (5), 725-738.
- Pereira, M., Andreatini, R., Schwarting, R. and Brenes, J. (2013). Amphetamine-induced appetitive 50-kHz calls in rats: a marker of affect in mania? *Psychopharmacology*. Springer-Verlag Berlin Heidelberg.
- Press, S.J. (2003). *Subjective and Objective Bayesian Statistic Principles, Models and Applications*. Wiley.
- R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Revilla, M. (2010). Quality in Unimode and Mixed-Mode designs: A Multitrait-Multimethod approach. *Survey Research Methods*. 4 (3), 151-164.
- Royle, A. and Link, W. (2002). Random effects and shrinkage estimation in capture-recapture models. *Journal of Applied Statistics* 29 (1-4), 329-351
- Rüsch, H., Jungwirth, D. and Hilsdorf, H. (1983). *Creep and Shrinkage. Their Effect on the Behavior of Concrete Structures*. New York Inc.: Springer-Verlag.
- Sanders, W. (1994). The Tennessee Value-Added Assessment System (TVAAS): Mixed-Model Methodology in Educational Assessment. *Journal of Personnel Evaluation in Education*. 8, 299-311.
- Sequeira, A., Fornaguera, J. & Brenes, J.C. (2016-2019). Pry01-163-2017-Cambios en la plasticidad neuronal relacionados con la incubación de la sensibilidad por la anfetamina: evaluación molecular y estructural. Instituto de Investigaciones en Salud. Universidad de Costa Rica.
- Soto-Liria, J., García-Pérez, J., Andújar-Rodríguez, A. (2000). ACERCA DE LA CORRECCIÓN DE LA MULTICOLINEALIDAD EN MODELOS DE

REGRESIÓN LINEAL de Anales de Economía Aplicada. XIV Reunión ASEPELT-España. Oviedo, 22 y 23 de junio de 2000.

Stata14. (2015). Estados Unidos, Texas: StataCorp.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*. 58 (1), 267-288.

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society*. 73 (3), 273–282.

Uhl, N., Eisenberg, T. (1970). PREDICTING SHRINKAGE IN THE MULTIPLE CORRELATION COEFFICIENT. *Educational and Psychological Measurement*. 30, 487-489.

Wackerly, D., Mendenhall III, W. and Scheaffer, R. (2010). *Estadística matemática con aplicaciones*. Cengage Learning, Inc. México.

Walker, D. (2007). A Comparison of Eight Shrinkage Formulas under Extreme Conditions. *Journal of Modern Applied Statistical Methods*. 6 (1), 162-172.

West, B.T., Welch, K.B. and Galecki, A.T. (2007). *Linear Mixed Models: A Practical Guide Using Statistical Software*. Chapman & Hall/CRC.

Xie, X., Kou, S.C., and Brown L. (2016). OPTIMAL SHRINKAGE ESTIMATION OF MEAN PARAMETERS IN FAMILY OF DISTRIBUTIONS WITH QUADRATIC VARIANCE. *Annals of statistics*. 44 (2), 564-597

Xu, S., Yuan, M., Karlsson, M., Dunne, A., Nandy, P. and Vermeulen, A. (2012). Shrinkage in Nonlinear Mixed-Effects Population Models: Quantification, Influencing Factors, and Impact. *The AAPS Journal*. 4 (4). 927-936.

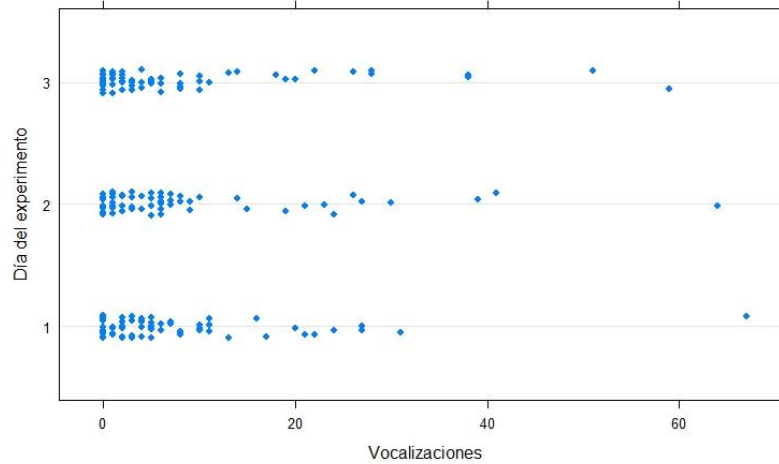
Yi, N. and Ma, S. (2012). Hierarchical Shrinkage Priors and Model Fitting for High-dimensional Generalized Linear Models. *Statistical Applications in Genetics and Molecular Biology*. 11 (6). Art 3.

## Anexo:

### Anexo 1: Estadísticas descriptivas

Gráfico A1

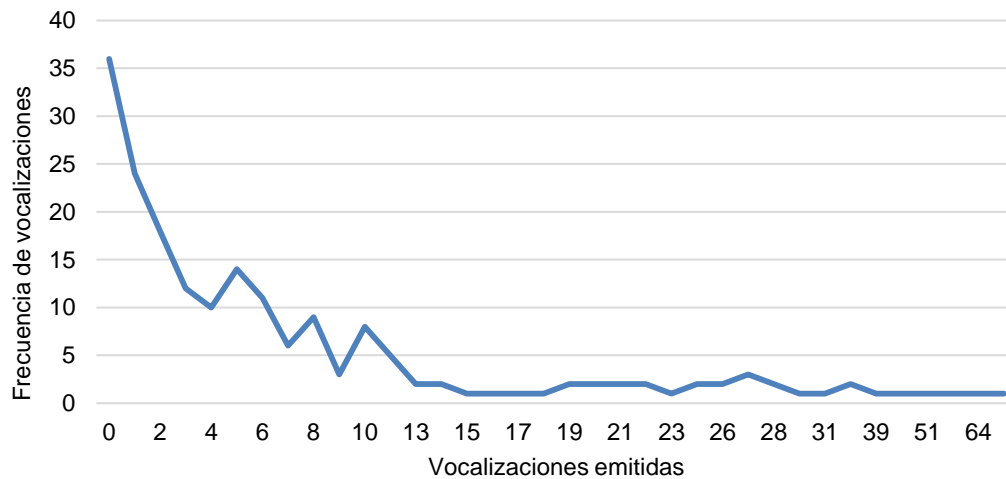
Cantidad de vocalizaciones emitidas por los sujetos del grupo control según día



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Gráfico A2

Conteo de vocalizaciones emitidas por las ratas del grupo control durante el experimento



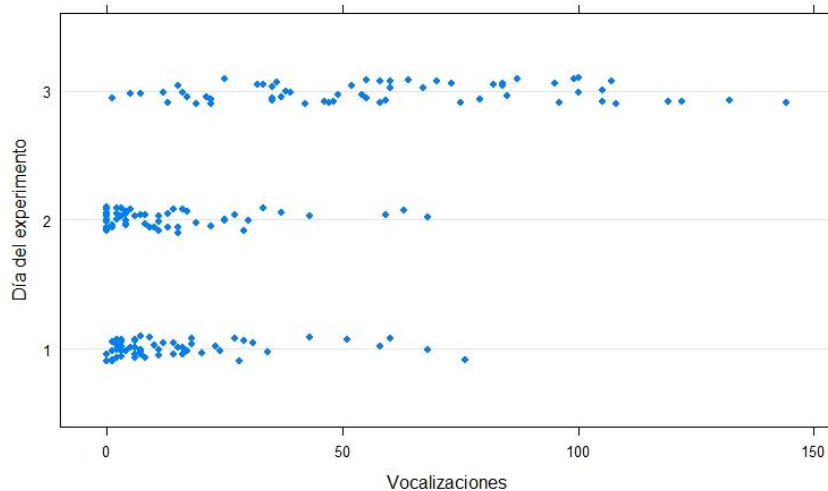
Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

**Cuadro A1**  
**Estadísticas generales de las vocalizaciones emitidas según sujetos del grupo control**

Sujeto	Total de vocalizaciones emitidas	Total de registros	Promedio de vocalizaciones emitidas	Varianza de vocalizaciones emitidas
Fe09	38	12	3,17	13,97
Fe14	55	12	4,58	16,45
Fe20	162	12	13,50	153,18
Fe28	52	12	4,33	9,33
Ju06	92	12	7,67	65,88
Ju13	74	12	6,17	69,79
Ju17	263	12	21,92	636,81
Ju24	63	12	5,25	49,84
Mr05	62	12	5,17	78,33
Mr07	112	12	9,33	151,52
Mr20	37	12	3,08	12,08
Mr22	120	12	10,00	106,91
My10	89	12	7,42	135,36
My12	88	12	7,33	53,15
My17	182	12	15,17	339,42
My24	61	12	5,08	29,17
<b>Total</b>	<b>1550</b>	<b>192</b>	<b>8,07</b>	<b>134,64</b>

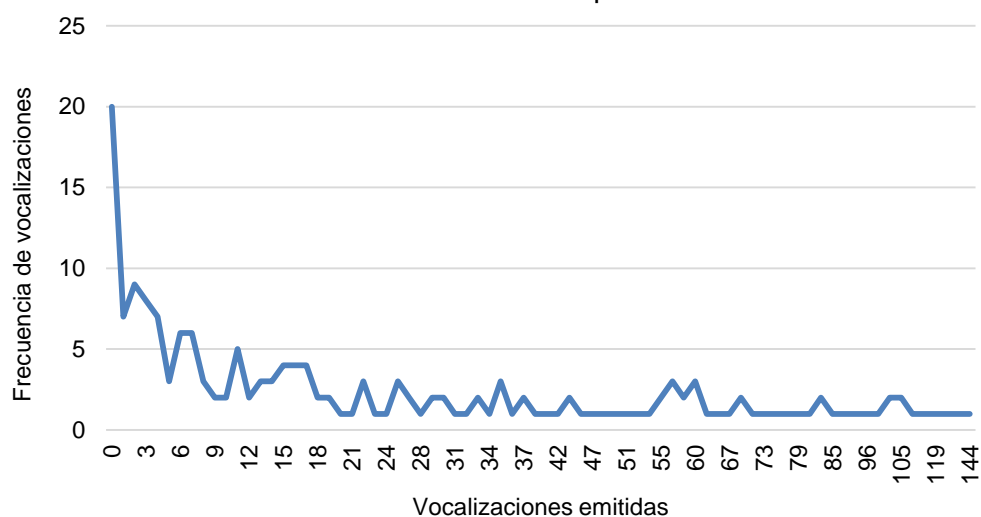
Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

**Gráfico A3**  
**Cantidad de vocalizaciones emitidas por los sujetos del grupo tratamiento según día**



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Gráfico A4  
 Conteo de vocalizaciones emitidas por las ratas del grupo  
 tratamiento durante el experimento



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Cuadro A2  
 Estadísticas generales de las vocalizaciones emitidas según sujetos del grupo tratamiento

Sujeto	Total de vocalizaciones emitidas	Total de registros	Promedio de vocalizaciones emitidas	Varianza de vocalizaciones emitidas
Fe12	191	12	15,92	396,27
Fe18	470	12	39,17	1348,88
Fe30	487	12	40,58	732,99
Ju03	629	12	52,42	2353,90
Ju08	240	12	20,00	199,09
Ju15	453	12	37,75	1115,84
Ju16	523	12	43,58	2437,36
Mr03	302	12	25,17	1057,97
Mr06	206	12	17,17	599,06
Mr08	266	12	22,17	862,52
Mr17	145	12	12,08	160,45
My01	183	12	15,25	91,84
My05	312	12	26,00	1045,27
My15	161	12	13,42	293,90
My21	588	12	49,00	1630,73
<b>Total</b>	<b>5156</b>	<b>180</b>	<b>28,64</b>	<b>1058,29</b>

Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

## **Anexo 2: Selección del modelo**

Para la selección del modelo se realizan sesiones de trabajo en conjunto con el experto en neurociencias.

El objetivo en esta fase es identificar un modelo parsimonioso que permita predecir el número de vocalizaciones emitidas por los roedores que fueron expuestos a la sustancia de anfetamina.

Se parte de los siguientes requisitos deseados por la investigadora y el experto en neurociencias:

- Se deben incluir al menos dos variables dentro de cada nivel para respetar la estructura de anidamiento indicada inicialmente y que constituye la base de datos.
- Se busca un modelo parsimonioso.

Para lo anterior, se decide analizar todos los posibles modelos que estén conformados por dos variables en el primer nivel (medición) para los días uno y dos, y al menos dos mediciones del día tres<sup>31</sup>; además de las variables del segundo nivel (día). La cantidad de modelos se estima a partir de las combinaciones posibles de las mediciones – 2 en el día 1, 2 en el día 2 y al menos 2 en el día 3 –, con un total de 396 modelos posibles. Ello utilizando el enfoque frecuentista.

Para cada caso se analiza la significancia de las variables “medición” incorporadas y se realiza la primera depuración en la que los modelos son clasificados en tres categorías: 1. No pasa a segunda fase, 2. Pasa a segunda fase con reserva, 3. Pasa a segunda fase.

En la primera categoría se incluyen los modelos en los cuales al menos una de las variables de medición del día uno o del día dos no resultan significativas y en los casos en los cuales para el día tres se tiene solo una medición. En la segunda

---

<sup>31</sup> En el día tres se suministra la sustancia de interés anfetamina, y se desea mantener la mayor información posible en el modelo.

categoría se incluyen los modelos en los cuales al menos una de las variables tiene valores de significancia superiores o iguales a 0.040, los cuales son considerados para este caso valores cercanos al límite definido (0.050). Y, en la tercera categoría se incluyen los modelos en los cuales todas las variables medición del día uno y dos resultaron significativas, y en el día tres al menos dos de las variables de medición resultaron significativas (119) y son los que pasan a la segunda fase de depuración.

La segunda fase de análisis corresponde a ordenar y ranquear los 119 modelos de acuerdo al valor del estadístico AIC.

En la tercera fase de depuración son analizados uno a uno los 119 modelos, de manera conjunta por la investigadora y el experto en neurociencias, y son seleccionados cuatro modelos a partir de la expertis de ambos. Dichos modelos son presentados a continuación:

Cuadro A3

Ranking de modelos seleccionados preliminarmente para el análisis

Modelo	VARIABLES INCORPORADAS	AIC
1	M1 M4 M5 M6 M9(N.S.) M10 M11 M12 <sup>a/</sup>	4 664,8
2	M1 M4 M5 M6 M9 M10 M11	4 668,9
3	M3 M4 M5 M6 M9 M10 M11 M12	4 670,2
4	M1 M4 M5 M8 M9(N.S.) M10 M11 M12 <sup>a/</sup>	4 677,5

<sup>a/</sup> En estos casos se incluye en el modelo la variable M9 a pesar de que ésta no es significativa (N.S.), ello porque corresponden a modelos que incorporan mayor cantidad de variables en el día 3 del experimento.

En la cuarta fase de análisis para determinar el modelo a utilizar en la presente investigación, se realizó la estimación frecuentista para los cuatros modelos antes indicados, considerando tres escenarios para cada caso:

- Modelo de efectos fijos
- Modelo con algunos efectos aleatorios (seleccionados a criterio de experto)
- Modelo de efectos aleatorios completo

Finalmente, se selecciona el segundo modelo pues además de presentar el segundo AIC más bajo, tuvo todas las variables resultan significativas.

## Anexo 3: Paquetes utilizados

### 3.1 glmer

Este paquete permite ajustar modelos lineales generalizados mixtos (GLMM), en los cuales se puede incorporar la estimación de los parámetros para los efectos fijos y aleatorios en la especificación del predictor lineal, vía máxima verosimilitud. Dicho predictor lineal se relaciona con la media condicional de la variable de respuesta a través de la función de enlace indicada.

La aproximación de verosimilitud para los GLMM es la Gauss adaptativa y el argumento `nAGQ` controla el número de nodos en el modelo.

Para la especificación de los modelos lineales generalizados se utiliza la siguiente fórmula:

```
glmer(formula, data=NULL, family=gaussian, control=glmerControl(),
      start=NULL, verbose=0L, nAGQ=1L, subset, weights, na.action,
      offset, contrasts = NULL, mustart, etastart, devFunOnly = FALSE)
```

Argumento	Detalle
<b>Formula</b>	Para los efectos fijos del modelo la variable de respuesta se coloca a la izquierda del operador ~ y las covariables de la derecha separadas por el operador +, por otra parte los efectos aleatorios se diferencian colocando el símbolo   el cual separa a los factores de agrupación siendo estos los que están a la derecha.
<b>Data</b>	Hace referencia a la base de datos en la cual se encuentran las variables nombradas en la formula.
<b>Family</b>	Permite especificar la descripción de la variable de respuesta así como la función de enlace.
<b>Control</b>	Corresponde a estructuras de control para el ajuste de los modelos mixtos. Los parámetros de control más generales: optimizador, parámetros de verificación, así como el número máximo de iteraciones.
<b>Start</b>	En la lista de valores iniciales para los parámetros de un modelo, o un vector. Por defecto se tiene que para un vector de datos las diagonales tienen valor 1 y fuera de la diagonal se tiene valores 0.
<b>Verbose</b>	Valor entero: Si > 0 se genera una salida detallada durante la optimización de las estimaciones del parámetro. Si se genera > 1 salida detallada

Argumento	Detalle
	durante la iterativa penalización individual de los mínimos cuadrados ponderados.
<b>nAGQ</b>	Valor entero: indica el número de puntos por eje para evaluar la aproximación de verosimilitud "Gauss-Hermite". Por defecto el valor es 1.
<b>Subset</b>	Indica el subconjunto de las filas de datos que deben utilizarse en el ajuste del modelo. Por defecto son incluidas todas las observaciones.
<b>Weights</b>	Hace referencia a un vector opcional de "pesos previos" que se utilizará en el proceso de ajuste. Debe ser NULL o un vector numérico.
<b>na.action</b>	Este argumento indica que hacer cuando se tienen valores faltantes (NA) en la base de datos, de manera automática son eliminados.
<b>Offset</b>	Éste se puede usar para especificar un componente conocido a priori que se incluirá en el predictor lineal durante el ajuste. Debe ser NULL o un vector numérico de longitud igual al número de casos.
<b>Contrasts</b>	Corresponde a una lista opcional para crear una matriz de diseño o modelo para incorporar factores de un grupo de variables dummies o interacciones.
<b>Mustart</b>	Valores iniciales para el vector de promedios.
<b>Etastart</b>	Valores iniciales para el predictor lineal.
<b>devFunOnly</b>	Parámetro lógico, devuelve la función de deviancia.

### 3.2 MCMCglmm:

Este paquete permite ajustar modelos lineales generalizados mixtos (GLMM) utilizando los métodos de Cadenas de Markov vía Monte Carlo, en la fórmula son especificados los efectos fijos y aleatorios. Dicha fórmula se presenta a continuación:

```
MCMCglmm(fixed, random=NULL, rcov=~units, family="gaussian",
         mev=NULL, data,start=NULL, prior=NULL, tune=NULL,
         pedigree=NULL, nodes="ALL", scale=TRUE, nitt=13000, thin=10,
         burnin=3000, pr=FALSE, pl=FALSE, verbose=TRUE, DIC=TRUE,
         singular.ok=FALSE, saveX=TRUE, saveZ=TRUE, saveXL=TRUE,
         slice=FALSE, ginverse=NULL, trunc=FALSE)
```

Argumento	Detalle
<b>Fixed</b>	Se utiliza para indicar los que serán efectos fijos en el modelo, la variable de respuesta se coloca a la izquierda del operador ~ y las covariables a la derecha, separadas por el operador +.
<b>Random</b>	<p>Para modelos más simples, la sintaxis del modelo tiene <code>landom1 + random2 + ... + randomk</code>. Pero se puede utilizar la función: <code>formvariance.function(formula)</code>: <code>linking.function(random.terms)</code>. Las funciones de variación disponibles son: <code>idv</code>, <code>idh</code>, <code>us</code>, <code>cor[]</code> y <code>ante[]</code>. <code>idv</code> ajusta una variación constante en todos los componentes de varianza en la formula. <code>idh</code> y <code>us</code> ajustan los componentes de varianza y covarianza en cada componente. <code>corg</code> fija los valores de la diagonal de la matriz de varianzas en 1, <code>corgh</code> fija los valores de la diagonal de las varianzas especificados en la priori, y <code>cors</code> permite submatrices de correlación. <code>ante[]</code> ajusta estructuras de dependencia de orden diferente y el número puede ser prefijado por una <code>c</code> para mantener todos los coeficientes de regresión del mismo orden iguales.</p> <p>Para <code>linking.function</code> están únicamente disponibles <code>mm</code> y <code>str</code>. <code>mm</code> se ajusta a un modelo de múltiples miembros donde varios términos aleatorios están separados por el operador +. <code>str</code> permite que existan covarianzas entre múltiples términos aleatorios que también están separados por el operador +. En ambos casos, los niveles de todos los términos aleatorios múltiples deben ser los mismos.</p>
<b>Rcov</b>	Fórmula para la estructura de covarianza residual. Esto se debe configurar de modo que cada punto de datos se asocie con un residuo único.
<b>Family</b>	Permite especificar la descripción de la variable de respuesta así como la función de enlace.
<b>Mev</b>	En un vector de información opcional para las mediciones de error de la varianzas en cada dato para en el análisis de los efectos aleatorios.
<b>Data</b>	Hace referencia a la base de datos en la cual se encuentran las variables nombradas en la formula.
<b>Start</b>	Lista opcional de valores iniciales: se tienen cuatro elementos posibles: <code>R</code> (estructura <code>R</code> ), <code>G</code> (estructura <code>G</code> ), <code>liab</code> (variables latentes) debe contener los valores iniciales donde <code>G</code> en sí es también una lista con tantos elementos como componentes de efectos aleatorios y <code>QUASI</code> que es un valor lógico, si es verdadero los valores iniciales se obtienen de manera heurística, si es falso se general mediante una distribución <code>Z</code> .
<b>Prior</b>	<p>Listado de especificaciones a priori, tiene tres elementos: <code>R</code> (estructura <code>R</code>), <code>G</code> (estructura <code>G</code>) y <code>B</code> (efectos fijos).</p> <p>Para las matrices <code>G</code> y <code>R</code> las estructuras de varianza son listas con las varianzas y covarianzas esperadas (<code>v</code>) y el grado de creencia (<code>nu</code>) para el Wishart inverso, así como el vector medio (<code>alpha.mu</code>) y la matriz de</p>

Argumento	Detalle
	covarianza ( $\alpha.V$ ). Los valores predeterminados son $\mu=0$ , $V=1$ , $\alpha.\mu=0$ y $\alpha.V=0$ . En el caso de $B$ se tiene el valor esperado ( $\mu$ ) y la matriz de varianza y covarianza ( $V$ ), los valores predeterminados son $\mu=0$ y $V=I*1e+10$ , donde está $I$ una matriz de identidad de dimensión apropiada.
<b>Tune</b>	Matriz opcional que define la propuesta de distribución para las variables latentes. Si es NULL, se utiliza un logaritmo adaptativo.
<b>Pedigree</b>	Ordena los árboles genealógicos con tres columnas identificadoras: <code>id</code> , <code>dam</code> and <code>sire</code> o <code>phylo</code> .
<b>Nodes</b>	Puede definirse como <code>ALL</code> o <code>TIPS</code> . En el primer caso estima los efectos para todos los individuos, en el segundo, las estimaciones de hacen para los subconjuntos de interés.
<b>Scale</b>	Valor lógico: ¿debe ser escalar a la longitud de la unidad?
<b>Nitt</b>	Número de iteraciones MCMC.
<b>Thin</b>	Intervalo de “adelgazamiento”.
<b>Burnin</b>	Estimaciones de “calentamiento” o iniciales.
<b>Pr</b>	Valor lógico: ¿debería guardarse la distribución posterior de los efectos aleatorios?
<b>P1</b>	Valor lógico: ¿debería guardarse la distribución posterior de variables latentes?
<b>Verbose</b>	Valor lógico: si los diagnósticos de TRUE MH se imprimen en la pantalla.
<b>DIC</b>	Valor lógico: calcula el criterio de información de deviancia.
<b>singular.ok</b>	Valor lógico: si el falso las dependencias lineales en lo efectos fijos son eliminadas, si es verdadero, son estimadas a través de información priori.
<b>saveX</b>	Valor lógico: guardar matriz de diseño de efectos fijos.
<b>saveZ</b>	Valor lógico: guardar matriz de diseño de efectos aleatorios.
<b>saveXL</b>	Valor lógico: guardar matriz de diseño de parámetros estructurales.
<b>Slice</b>	Valor lógico: ¿se debe utilizar el muestreo de corte? (se aplica únicamente a casos binarios con residuos independientes).
<b>Ginverse</b>	Lista de matrices inversas ( $\text{solve}(A)$ ) que son proporcionales a la estructura de covarianza de los efectos aleatorios. Los nombres de las matrices deben corresponder a columnas en los datos que están asociados con los términos aleatorios.
<b>Trunc</b>	Valor lógico: ¿deberían truncarse las variables latentes en los modelos binarios?

### 3.3 glmmLasso

Este paquete proporciona un enfoque de selección de variables para modelos mixtos lineales generalizados por estimación penalizada por L1. El modelo final incluye sólo las variables correspondientes a los efectos fijos distintos a cero, los cuales se ajustan mediante la puntuación de Fisher. Tanto para el algoritmo principal como para la reestimación final, Fisher puede elegir dos métodos para el cálculo de los parámetros de varianza y covarianza de efectos aleatorios, una estimación de tipo EM y una clase REML.

Para la especificación de los modelos lineales generalizados se utiliza la siguiente fórmula:

```
glmmLasso(fix=formula, rnd=formula, data, lambda, family =
  gaussian(link="identity"), switch.NR=FALSE, final.re=FALSE,
  control = list())
```

Argumento	Detalle
<b>Fix</b>	Se utiliza para indicar los efectos fijos en el modelo, la variable de respuesta se coloca a la izquierda del operador ~ y las covariables a la derecha, separadas por el operador +. Para variables categóricas se debe utilizar <code>as.factor(.)</code>
<b>Rnd</b>	Corresponde a la parte de efectos aleatorios del modelo, con el factor de agrupación a la izquierda del operador ~ y, a la derecha los términos aleatorios separados por los operadores +. Se utiliza el término "list()". Si se establece en NULL, no se incluyen efectos aleatorios.
<b>Data</b>	Hace referencia a la base de datos en la cual se encuentran las variables nombradas en la formula.
<b>Lambda</b>	Hace referencia al parámetro de penalización que controla la contracción de los términos fijos y controla la selección de variables. El valor de dicho parámetro debe determinarse mediante el uso de criterios de información o validación cruzada.
<b>Family</b>	Permite especificar la descripción de la variable de respuesta así como la función de enlace.  Este paquete permite ajustar modelos de respuesta ordinal: <code>family=acat()</code> y <code>family=cumulative()</code> para el ajuste de una categoría adyacente o modelo acumulativo, respectivamente. Si no se

Argumento	Detalle
	indica el elemento <code>family</code> en la fórmula se ajusta un modelo lineal mixto; de lo contrario se ajusta un modelo mixto lineal generalizado.
<code>switch.NR</code>	Valor lógico. ¿Debería el algoritmo cambiar al de Newton-Raphson, cuando sea razonable? El valor predeterminado es FALSE.
<code>final.re</code>	Valor lógico. ¿Debería realizarse la reestimación final de la puntuación de Fisher? El valor predeterminado es FALSE.
<b>Control</b>	Corresponde a una lista de valores de control para el algoritmo de estimación, para especificar dichos valores utilice la función <code>glmmLassoControl</code> . El valor predeterminado es una lista vacía.

### 3.4 glmsr

El paquete `glmsr` proporciona funciones para realizar estimaciones en modelos mixtos lineales generalizados, por ende, el usuario puede determinar qué método usar para la estimación. Entre ellas se tienen: *Laplace* y Gauss Adaptativa del paquete (`lme4`), y se incluyen las aproximaciones por reducción secuencial y muestreo de importancia (ambos proporcionan una estimación precisa a la probabilidad en algunas situaciones en las que no es posible utilizar Gauss Adaptativa).

La fórmula para especificar el modelo lineal generalizado mixto corresponde a:

```
glmm(formula, subformula = NULL, data = NULL, family = gaussian,
      method = NULL, control = list(), weights = NULL, offset = NULL,
      prev_fit = NULL, verbose = 1L, lme4_control =
      set_lme4_control())
```

Argumento	Detalle
<b>Formula</b>	Para los efectos fijos del modelo la variable de respuesta se coloca a la izquierda del operador <code>~</code> y las covariables de la derecha separadas por el operador <code>+</code> , por otra parte los efectos aleatorios se diferencian colocando el símbolo <code> </code> el cual separa a los factores de agrupación siendo estos los que están a la derecha.
<b>Subformula</b>	Subformula que indica la relación entre la variable dependiente y las covariables.

Argumento	Detalle
<b>Data</b>	Hace referencia a la base de datos en la cual se encuentran las variables nombradas en la formula.
<b>Family</b>	Permite especificar la descripción de la variable de respuesta así como la función de enlace.
<b>Method</b>	Mecanismos para aproximar la verosimilitud del modelo: "Laplace", "AGQ" (Cuadratura Gaussiana Adaptativa de lme4), "SR" (Reducción secuencial) y "IS" (Muestreo por importancia).
<b>Control</b>	Listado de parámetros extra que permite controlar la aproximación de la verosimilitud.
<b>Weights</b>	Vector opcional de "pesos previos" que son utilizados en el proceso de ajuste. Debe ser NULL o un vector numérico.
<b>Offset</b>	Se utiliza para especificar el componente informativo previo conocido a ser incorporado en la estimación. Debe ser NULL o un vector numérico del mismo tamaño el número de casos.
<b>prev_fit</b>	Objeto <code>glmmFit</code> , muestra el resultado de un ajuste previo.
<b>Verbose</b>	Corresponde a detalles que se imprimen con la estimación del modelo. Para <code>verbose = 0</code> , no se imprime información. Para <code>verbose = 1</code> (definida previamente), muestra un resumen cada segundo durante el ajuste del modelo. Para <code>verbose = 2</code> , muestra el resultado del valor del parámetro, así como la verosimilitud en cada estado de la optimización.
<b>lme4_ctrl</b>	Parámetro de control.

### 3.5 hglm2

`hglm2` permite ajustar modelos lineales jerárquicos generalizados. Y, extiende la función `hglm` al permitir varios efectos aleatorios, bajo la especificación de `lme4`, e implementando técnicas de matriz dispersa utilizando la biblioteca `Matrix`.

La fórmula para especificar el modelo lineal generalizado mixto corresponde a:

```
hglm2(meanmodel, data = NULL, family = gaussian(link = identity),
      rand.family = gaussian(link = identity), method = "EQL", conv =
      1e-6, maxit = 50, startval = NULL, X.disp = NULL, disp = NULL,
      link.disp = "log", weights = NULL, fix.disp = NULL, offset =
```

NULL, sparse = TRUE, vcovmat = FALSE, calc.like = FALSE, RandC = NULL, bigRR = FALSE, verbose = FALSE

Argumento	Detalle
<b>Meanmodel</b>	Los efectos fijos se especifican colocando la variable de respuesta a la izquierda del operador ~ y las covariables del lado derecho separadas por el operador +, para los efectos aleatorios, le utiliza el símbolo  , en el cual los niveles se indican a la derecha.
<b>Data</b>	Hace referencia a la base de datos en la cual se encuentran las variables nombradas en la formula.
<b>Family</b>	Permite especificar la descripción de la variable de respuesta así como la función de enlace.
<b>rand.family</b>	Describe la función de distribución y enlace que se utilizará para los efectos aleatorios.
<b>Method</b>	Método de estimación, de manera predeterminada se tiene EQL, también se tiene la opción EQL1, que mejora la estimación de los GLMM (especialmente para los modelos de Poisson con un gran número de niveles en los efectos aleatorios).
<b>Conv</b>	Los criterios de convergencia (cambio en el predictor lineal entre iteraciones).
<b>Maxit</b>	Número máximo de iteraciones.
<b>Startval</b>	Un vector de valores iniciales en el siguiente orden: efectos fijos, efecto aleatorio, varianza de los efectos aleatorios, varianza de los residuos.
<b>x.disp</b>	La matriz de diseño para los efectos fijos en el modelo.
<b>disp.</b>	Parte de la fórmula en la cual se especifican los efectos fijos en el modelo.
<b>link.disp</b>	La función de enlace para la parte de dispersión del modelo.
<b>Weights</b>	Ponderaciones previas especificadas.
<b>fix.disp</b>	Un valor numérico si el parámetro de dispersión del modelo es conocido.
<b>Offset</b>	Un desplazamiento para el predictor lineal del modelo medio.
<b>Sparse</b>	Valor lógico. Si es verdadero, el cálculo se realizará utilizando una técnica de matriz dispersa.
<b>Vcovmat</b>	Valor lógico. Si es verdadero, se exporta la matriz de varianza y covarianza.
<b>calc.like</b>	Valor lógico. Si es verdadero, las probabilidades se calcularán en la convergencia y se mostrarán a través de los métodos de impresión o resumen en el objeto de salida.
<b>RandC</b>	Vector que indica el número de columnas de Z que se utilizarán para cada uno de los términos de efectos aleatorios.

Argumento	Detalle
<b>bigRR</b>	Valor lógico. Si es verdadero, y solo para el modelo Gaussiano con un término de efecto aleatorio, se utilizará un algoritmo específico para ajuste rápido cuando se tienen problemas de alta dimensión de variables ( $p \gg n$ ).
<b>Verbose</b>	Valor lógico. Si es VERDADERO, se imprime más información durante el proceso de ajuste del modelo.

### 3.6 glmmTMB

`glmmTMB` permite ajustar los modelos mixtos lineales y lineales generalizados con varias extensiones, incluida la inflación cero. Los modelos se ajustan utilizando la estimación de máxima verosimilitud a través de “TMB” (*Template Model Builder*). Supone que los efectos aleatorios son gaussianos en la escala del predictor lineal y se integran utilizando la aproximación de *Laplace*. Los gradientes se calculan mediante diferenciación automática.

La fórmula para especificar el modelo lineal generalizado mixto corresponde a:

```
glmmTMB(formula, data = NULL, family = gaussian(), ziformula = ~0,
         dispformula = ~1, weights = NULL, offset = NULL, contrasts =
         NULL, na.action = na.fail, se = TRUE, verbose = FALSE, doFit =
         TRUE, control = glmmTMBControl(), REML = FALSE)
```

Argumento	Detalle
<b>Formula</b>	Fórmula combinada de efectos fijos y aleatorios, siguiendo la sintaxis <code>lme4</code> .
<b>Data</b>	Base de datos.
<b>Family</b>	Permite especificar la descripción de la variable de respuesta así como la función de enlace.
<b>Ziformula</b>	Una fórmula unilateral (es decir, sin variable de respuesta) para la inflación cero que combina efectos fijos y aleatorios: el valor predeterminado <code>~ 0</code> especifica que no hay inflación cero. Especificando <code>~</code> establece la fórmula de inflación cero idéntica al lado derecho de la fórmula (es decir, la fórmula de efectos condicionales); los términos también se pueden sumar o restar. Cuando se utiliza <code>~</code> . Como la fórmula de inflación cero en modelos donde la fórmula de efectos condicionales contiene un término de compensación, el término de compensación se eliminará automáticamente. El modelo de inflación cero utiliza un enlace <i>logit</i> .

Argumento	Detalle
<b>dispformula</b>	Una fórmula unilateral para dispersión que contiene solo efectos fijos: el valor predeterminado $\sim 1$ especifica la dispersión estándar dada a cualquier familia. El argumento se ignora para las familias que no tienen un parámetro de dispersión. El modelo de dispersión utiliza un enlace de registro. En los modelos mixtos gaussianos, $\text{dispformula} = \sim 0$ fija el parámetro para que sea 0, forzando la variación en los efectos aleatorios.
<b>Weights</b>	Hace referencia a un vector opcional de "pesos" que se utilizará en el proceso de ajuste. No se escala automáticamente para tener la suma 1.
<b>Offset</b>	Parámetro de compensación o ponderación únicamente para el modelo condicional.
<b>Contrasts</b>	Una lista opcional de parámetros para especificar los valores de matrices numéricas o funciones de denominación de cadenas de caracteres, que se utilizarán como valores de reemplazo para la función de reemplazo de contrastes y cuyos nombres son los nombres de columnas de datos que contienen factores.
<b>na.action</b>	Este parámetro permite indicar como se van a manejar los valores perdidos. El valor predeterminado es "na.omit".
<b>Se</b>	Indicación para pedir si quiero que devolver errores estándar.
<b>Verbose</b>	Lógica que indica si se debe imprimir alguna indicación de progreso en la consola.
<b>DoFit</b>	Si se ajusta al modelo completo o (si es FALSO) devuelve los datos preprocesados y los objetos de parámetros, sin ajustar el modelo.
<b>Control</b>	Parámetros de control.
<b>REML</b>	Lógico; Utilice la estimación REML en lugar de la probabilidad máxima.

## Anexo 4: Estimación adicional de modelos a partir del enfoque frecuentista y bayesiano

### Anexo 4.1: Estimación frecuentista con R

Cuadro A4

Estimación frecuentista de coeficientes para los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, utilizando el paquete `glmmTMB`

Variable	Modelo de efectos fijos		Modelo con efectos aleatorios		Modelo completo de efectos aleatorios	
	Coeficiente	Error estándar	Coeficiente	Error estándar	Coeficiente	Error estándar
Intercepto	2.0120	0.1252	2.0581	0.1283	1.7111	0.1575
DIA2	-0.6971	0.1289	-0.8287	0.1402	-0.4923	0.1893
DIA3	1.4892	0.1621	1.4000	0.1783	1.7248	0.2476
M1	0.7268	0.0550	0.4535	0.2154	0.5630	0.2213
M4	-1.2377	0.1113	-1.2501	0.1118	-1.5226	0.3012
M5	1.3306	0.0685	1.2252	0.2132	1.2060	0.2518
M6	0.6649	0.0789	0.6649	0.0793	0.2286 <sup>a/</sup>	0.2716
M9	0.2829	0.0464	0.2935	0.0467	0.7868	0.2897
M10	0.3852	0.0454	0.3958	0.0457	0.3258 <sup>a/</sup>	0.1807
M11	0.3204	0.0461	-0.3024 <sup>a/</sup>	0.2063	0.2135 <sup>a/</sup>	0.1297
Efectos aleatorios	Varianza	Desviación estándar	Varianza	Desviación estándar	Varianza	Desviación estándar
ID_SUJ	0.2971	0.5451	0.2882	0.5368	0.3105	0.5572
ID_DIA	0.2132	0.4618	0.2493	0.4993	0.4915	0.7011
M1			1.2266	1.1078	1.2910	1.1362
M5			1.0890	1.0436	1.5800	1.2570
M11			0.8755	0.9357	0.1615	0.4019
M4					1.6492	1.2842
M6					1.5290	1.2365
M9					2.3643	1.5376
M10					0.6514	0.8071

<sup>a/</sup> Coeficientes que no resultaron significativos al modelo, con  $\alpha=0.05$ .









**Imagen A5**  
**Estimación bayesiana de coeficientes para el modelo de efectos fijos utilizando un modelo multinivel con regresión Poisson**

Multilevel structure

---

id\_suj  
 {U0}: random intercepts

id\_suj>id\_dia  
 {UU0}: random intercepts

---

Model summary

---

Likelihood:

y ~ mepoisson(xb\_y)

Priors:

{y:m1 m4 m5 m6 m9 m10 m11 dia2 dia3 \_cons} ~ normal(0,10000) (1)  
 {U0} ~ normal(0,{U0:sigma2}) (1)  
 {UU0} ~ normal(0,{UU0:sigma2}) (1)

Hyperpriors:

{U0:sigma2} ~ igamma(.01,.01)  
 {UU0:sigma2} ~ igamma(.01,.01)

---

(1) Parameters are elements of the linear form xb\_y.

Bayesian multilevel Poisson regression	MCMC iterations =	40,000
Random-walk Metropolis-Hastings sampling	Burn-in =	10,000
	MCMC sample size =	30,000

Group Variable	No. of Groups	Observations per Group		
		Minimum	Average	Maximum
id_suj	31	12	12.0	12
id_dia	93	4	4.0	4

Family : Poisson  
Link : log

Number of obs = 372  
Acceptance rate = .3058  
Efficiency: min = .001388  
avg = .004684  
max = .01479

Log marginal likelihood

		Mean	Std. Dev.	MCSE	Median	Equal-tailed [95% Cred. Interval]	
Y	m1	.7618429	.0491654	.006481	.7616149	.6669862	.8562709
	m4	-1.209656	.1113084	.011476	-1.208699	-1.4313	-.9945216
	m5	1.30617	.0474114	.005746	1.305705	1.215758	1.397126
	m6	.6406255	.0617382	.006616	.6378869	.527365	.7689328
	m9	.2901904	.0362458	.003538	.2903359	.2180221	.3597824
	m10	.3815858	.0467809	.004513	.380962	.2951576	.4754425
	m11	.3225404	.0368962	.003268	.3216146	.2495912	.3934896
	dia2	-.5377845	.0326963	.005067	-.5417106	-.5921774	-.4706812
	dia3	1.580345	.1396562	.021385	1.586352	1.306684	1.848858
	_cons	1.936049	.0614936	.007224	1.937159	1.811507	2.057083
id_suj	U0:sigma2	.3248467	.1060246	.005033	.3082677	.1712044	.5762625
id_suj>id_dia	UU0:sigma2	.2241669	.0452368	.002159	.2190612	.1499192	.3254848

Note: Default priors are used for model parameters.

Note: There is a high autocorrelation after 500 lags.

## Imagen A6

## Estimación bayesiana de coeficientes para el modelo de efectos aleatorios utilizando un modelo multinivel con regresión Poisson

Multilevel structure

---

```
id_suj
  {U0}: random intercepts
  {U1}: random coefficients for m1
  {U2}: random coefficients for m5
  {U3}: random coefficients for m11
```

```
id_suj>id_dia
  {UU0}: random intercepts
```

---

Model summary

---

Likelihood:

```
y ~ mepoisson(xb_y)
```

Priors:

```
{y:m1 m4 m5 m6 m9 m10 m11 dia2 dia3 _cons} ~ normal(0,10000)      (1)
                                           {U0} ~ normal(0,{U0:sigma2})  (1)
                                           {U1} ~ normal(0,{U1:sigma2})  (1)
                                           {U2} ~ normal(0,{U2:sigma2})  (1)
                                           {U3} ~ normal(0,{U3:sigma2})  (1)
                                           {UU0} ~ normal(0,{UU0:sigma2}) (1)
```

Hyperpriors:

```
{U0:sigma2} ~ igamma(.01,.01)
{U1:sigma2} ~ igamma(.01,.01)
{U2:sigma2} ~ igamma(.01,.01)
{U3:sigma2} ~ igamma(.01,.01)
{UU0:sigma2} ~ igamma(.01,.01)
```

---

(1) Parameters are elements of the linear form `xb_y`.



## Anexo 5: Selección del valor para *Lambda* en la regresión LASSO

### 5.1 Uso de Validación cruzada

```

N<-dim(base)[1]
ind<-sample(N,N)
lambda <- seq(500,0,by=-5)
kk<-5
nk <- floor(N/kk)
Devianz_ma<-matrix(Inf,ncol=kk,nrow=length(lambda))

for(j in 1:length(lambda))
{
  print(paste("Iteration ", j,sep=""))
  for (i in 1:kk)
  {
    if (i < kk)
    {
      indi <- ind[(i-1)*nk+(1:nk)]
    }else{
      indi <- ind[((i-1)*nk+1):N]
    }
    base.train<-base[-indi,]
    base.test<-base[indi,]

    glm2 <- glmLasso(fix = y ~ 1+DIA2+DIA3+M1+M4+M5+M6+M9+M10+M1,
rnd=list(Sujeto=~ID_DIA),
family = poisson(link="log"), data = base.train,
lambda=lambda[j],
switch.NR=F,final.re=TRUE)

    y.hat<-predict(glm2,base.test)
    Devianz_ma[j,i]<-sum(family$dev.resids(base.test$points,y.hat,
wt=rep(1,length(y.hat))))
  }
  print(sum(Devianz_ma[j,]))
}
Devianz_vec<-apply(Devianz_ma,1,sum)
opt2<-which.min(Devianz_vec)

opt2
lambda[opt2]

```

## 5.2 Uso de BIC y AIC

```

lambda <- seq(500, 0, by=-5)
BIC_vec <- rep(Inf, length(lambda))
AIC_vec <- rep(Inf, length(lambda))
for(j in 1:length(lambda))
{
  glm1 <- glmLasso(fix = y ~ DIA2+DIA3+M1+M4+M5+M6+M9+M10+M11, rn
d = list(Sujeto=~1), family = poisson(link="log"), data = base, la
mbda = lambda[j], switch.NR = T, final.re=TRUE)
  BIC_vec[j] <- glm1$bic
  AIC_vec[j] <- glm1$aic
}
opt.bic<-which.min(BIC_vec)
opt.aic<-which.min(AIC_vec)

opt.bic
opt.aic
lambda[opt.bic]
lambda[opt.aic]

```

## 5.3 Resumen de estimaciones

Cuadro A5

Estimación de los coeficientes fijos para los modelos de efectos fijos, según valores de *Lambda*

Variable	Valores de <i>Lambda</i>					
	$\lambda=0$	$\lambda=50$	$\lambda=75$	$\lambda=80$	$\lambda=100$	$\lambda=370$
Intercepto	1.9245	1.9246	1.9248	1.9246	1.9248	1.9253
DIA2	-0.4738	-0.4738	-0.4739	-0.4738	-0.4739	-0.2004
DIA3	1.6404	1.6404	1.6401	1.6403	1.6402	1.6395
M1	0.7283	0.7283	0.7283	0.7283	0.7283	0.7283
M4	-1.2363	-1.2363	-1.2363	-1.2363	-1.2363	-1.2362
M5	1.3306	1.3306	1.3306	1.3306	1.3306	1.0569
M6	0.6649	0.6649	0.6649	0.6649	0.6649	0.0000 <sup>a/</sup>
M9	0.2818	0.2818	0.2818	0.2818	0.2818	0.2817
M10	0.3841	0.3841	0.3841	0.3841	0.3841	0.3840
M11	0.3192	0.3192	0.3192	0.3192	0.3192	0.3192

<sup>a/</sup> Coeficientes que según la estimación de la regresión LASSO no deben ser seleccionados en el modelo.

**Cuadro A6**  
Estimación de los errores estándar de los coeficientes fijos para los modelos de efectos fijos, según valores de *Lambda*

Variables	Valores de <i>Lambda</i>					
	$\lambda=0$	$\lambda=50$	$\lambda=75$	$\lambda=80$	$\lambda=100$	$\lambda=370$
Intercepto	0.1051	0.1051	0.1051	0.1051	0.1051	0.1051
DIA2	0.0711	0.0711	0.0711	0.0711	0.0711	0.0594
DIA3	0.1062	0.1062	0.1062	0.1062	0.1062	0.1061
M1	0.0551	0.0551	0.0551	0.0551	0.0551	0.0551
M4	0.1117	0.1117	0.1117	0.1117	0.1117	0.1117
M5	0.0688	0.0688	0.0688	0.0688	0.0688	0.0566
M6	0.0793	0.0793	0.0793	0.0793	0.0793	NA <sup>a/</sup>
M9	0.0464	0.0464	0.0464	0.0464	0.0464	0.0464
M10	0.0454	0.0454	0.0454	0.0454	0.0454	0.0454
M11	0.0460	0.0460	0.0460	0.0460	0.0460	0.0460

<sup>a/</sup> Coeficientes que según la estimación de la regresión LASSO no deben ser seleccionados en el modelo.

**Cuadro A7**  
Estimación de coeficientes fijos para los modelos de efectos aleatorios completos, según valores de *Lambda*

Variable	Valores de <i>Lambda</i>					
	$\lambda=0$	$\lambda=50$	$\lambda=75$	$\lambda=80$	$\lambda=100$	$\lambda=370$
Intercepto	2.0924	2.0896	2.0869	2.0861	2.0621	2.0535
DIA2	-0.7030	-0.7011	-0.6921	-0.6915	-0.7224	-0.5164
DIA3	1.2413	1.1631	1.1652	1.1633	1.5276	1.2340
M1	0.5350	0.5387	0.5371	0.5376	0.4062 <sup>a/</sup>	0.5947
M4	-1.3507	-1.3410	-1.3471	-1.3438	-1.4186	-1.1504
M5	1.1867	1.1837	1.1719	1.1710	1.2249	0.9239
M6	0.4498 <sup>a/</sup>	0.4261 <sup>a/</sup>	0.4257 <sup>a/</sup>	0.4238 <sup>a/</sup>	0.5309	0.0000 <sup>b/</sup>
M9	0.5966	0.6102	0.6049	0.6055	0.4449 <sup>a/</sup>	0.7017
M10	0.3252	0.3336	0.3291	0.3295	0.0000 <sup>b/</sup>	0.3467
M11	0.1947 <sup>a/</sup>	0.2078 <sup>a/</sup>	0.2027 <sup>a/</sup>	0.2039 <sup>a/</sup>	0.1106 <sup>a/</sup>	0.1811 <sup>a/</sup>

<sup>a/</sup> Coeficientes que no son significativos al modelo con  $\alpha=0.05$ .

<sup>b/</sup> Coeficientes que según la estimación de la regresión LASSO no deben ser seleccionados en el modelo.

Cuadro A8  
Estimación de los errores estándar para los coeficientes fijos en los modelos  
con todos los efectos aleatorios, según valores de *Lambda*

Variables	Valores de <i>Lambda</i>					
	$\lambda=0$	$\lambda=50$	$\lambda=75$	$\lambda=80$	$\lambda=100$	$\lambda=370$
Intercepto	0.1173	0.1180	0.1186	0.1187	0.1143	0.1176
DIA2	0.1938	0.1940	0.1909	0.1912	0.1853	0.1481
DIA3	0.3054	0.3128	0.3127	0.3127	0.2714	0.3203
M1	0.2180	0.2182	0.2179	0.2180	0.2054 <sup>a/</sup>	0.2174
M4	0.2858	0.2843	0.2843	0.2840	0.2744	0.2525
M5	0.2585	0.2580	0.2554	0.2556	0.2502	0.1874
M6	0.2512 <sup>a/</sup>	0.2558 <sup>a/</sup>	0.2511 <sup>a/</sup>	0.2520 <sup>a/</sup>	0.2448	NA <sup>b/</sup>
M9	0.2625	0.2634	0.2624	0.2625	0.2497 <sup>a/</sup>	0.2562
M10	0.1611	0.1620	0.1607	0.1608	NA <sup>b/</sup>	0.1533
M11	0.1108 <sup>a/</sup>	0.1083 <sup>a/</sup>	0.1086 <sup>a/</sup>	0.1083 <sup>a/</sup>	0.0901 <sup>a/</sup>	0.1127 <sup>a/</sup>

<sup>a/</sup> Coeficientes que no son significativos al modelo con  $\alpha=0.05$ .

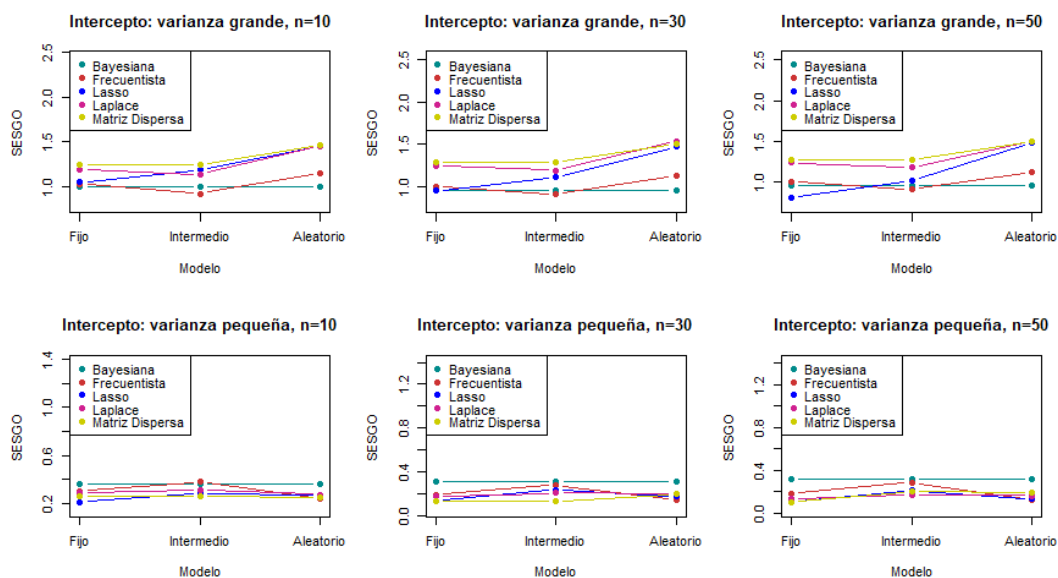
<sup>b/</sup> Coeficientes que según la estimación de la regresión LASSO no deben ser seleccionados en el modelo.

## Anexo 6: Estimación de coeficientes a partir del estudio de simulaciones

### 6.1 Estimación del SESGO

Imagen A7

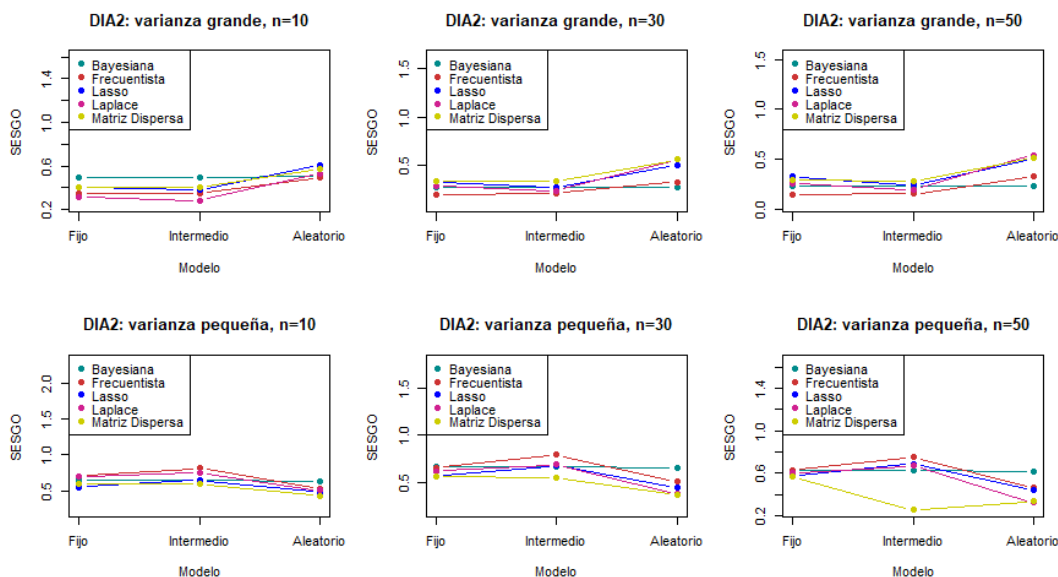
SESGO del coeficiente intercepto en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A8

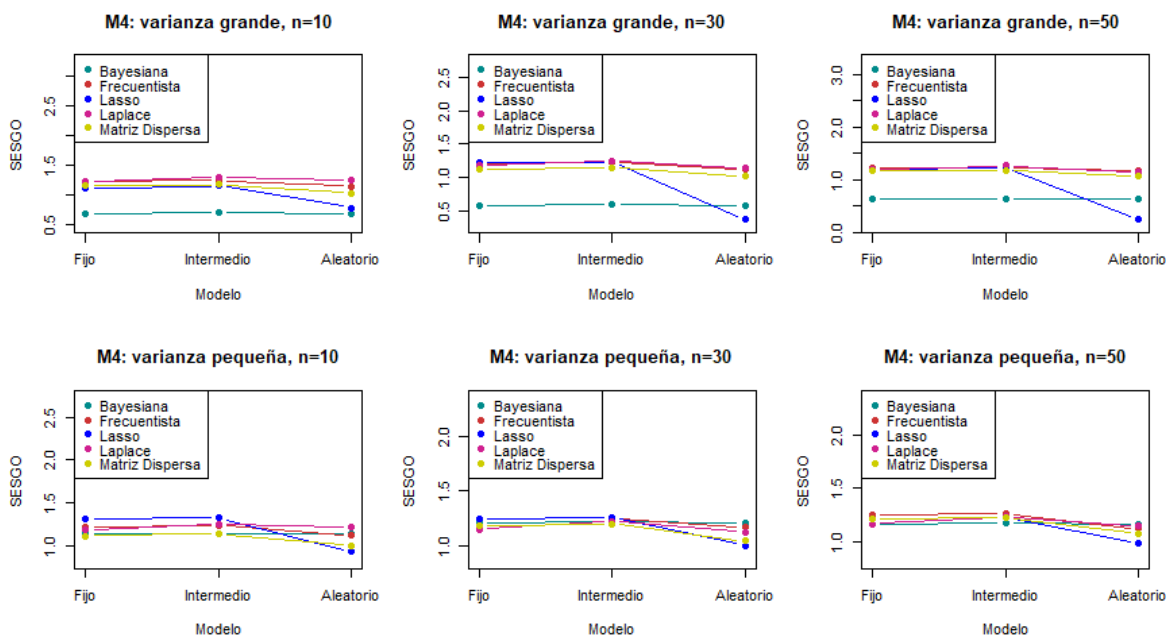
SESGO del coeficiente DIA2 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A9

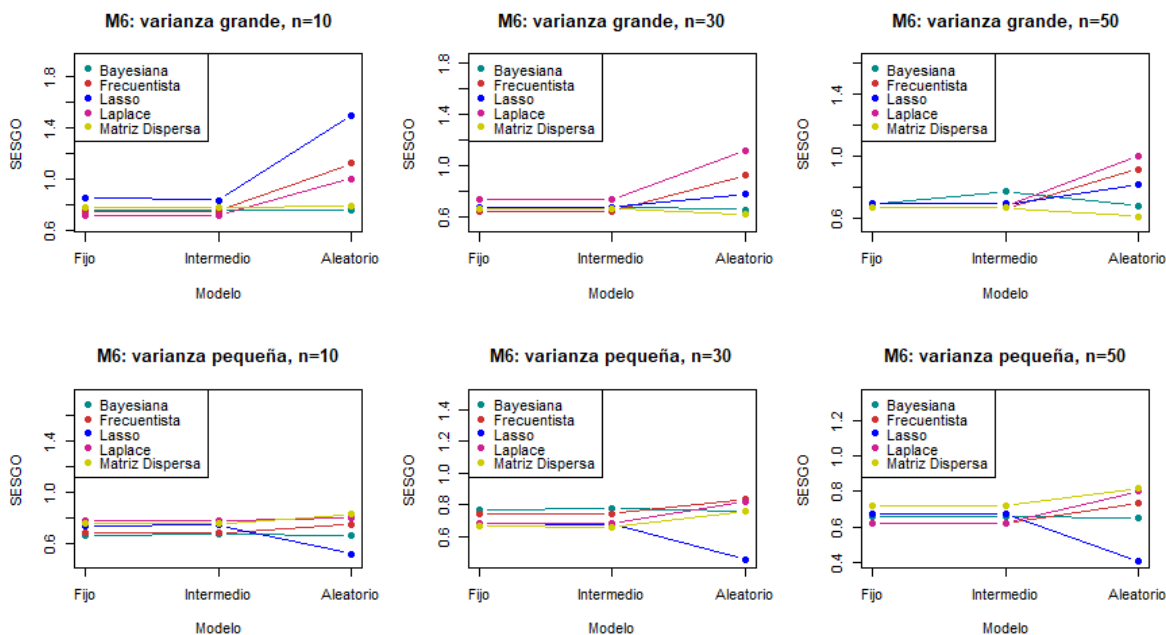
SESGO del coeficiente M4 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A10

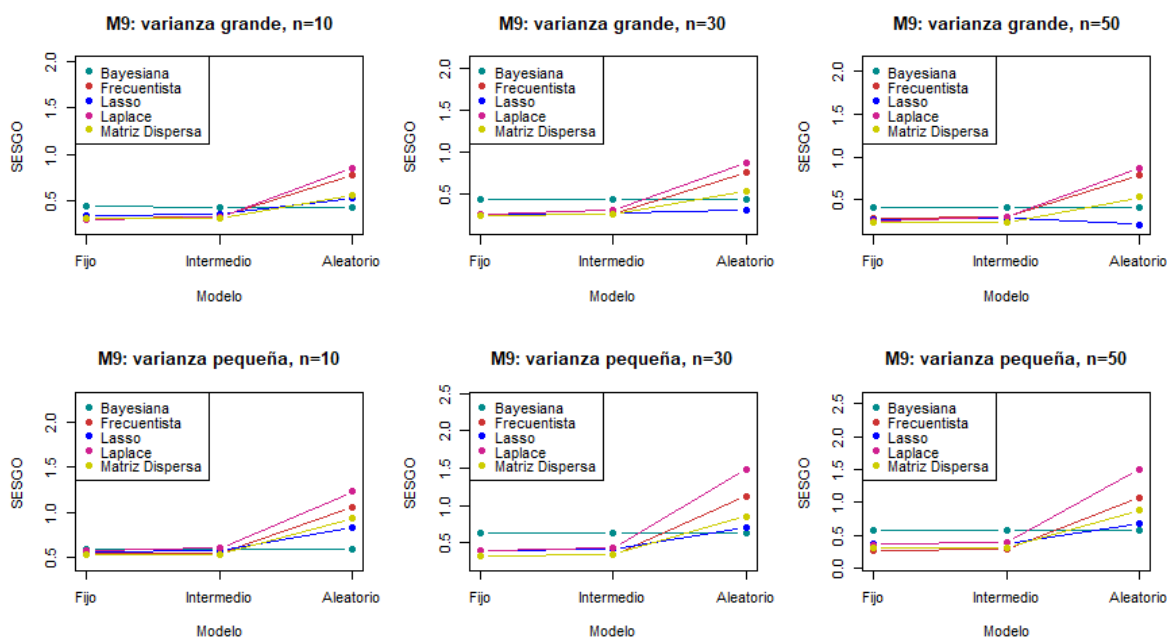
SESGO del coeficiente M6 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A11

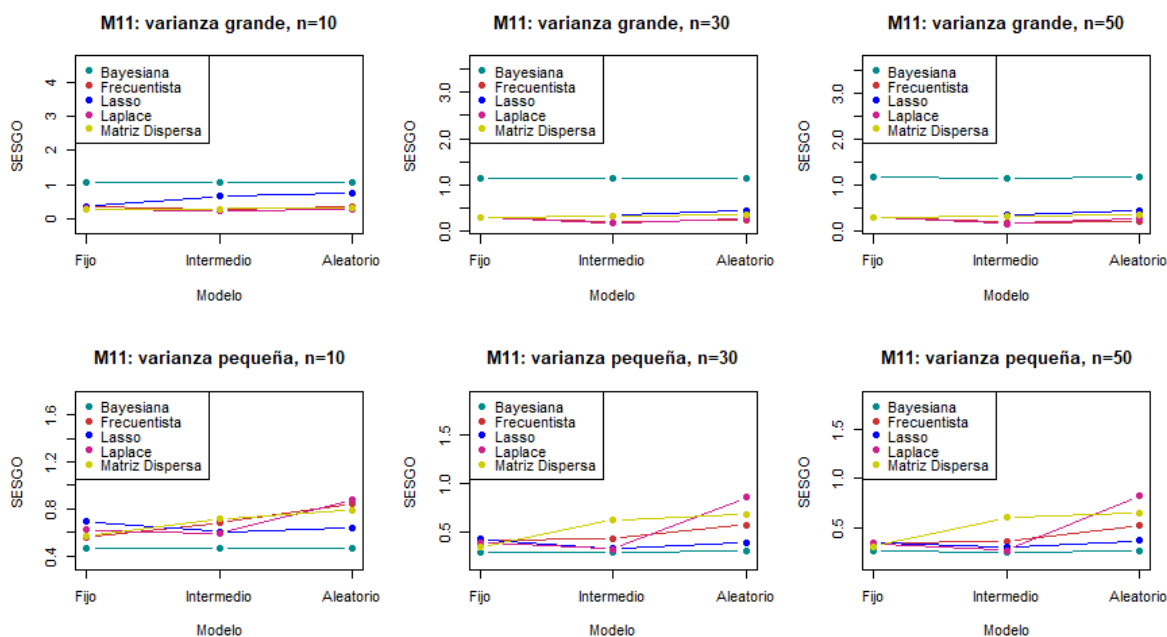
SESGO del coeficiente M9 en los modelos de coeficientes fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A12

SESGO del coeficiente M11 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza

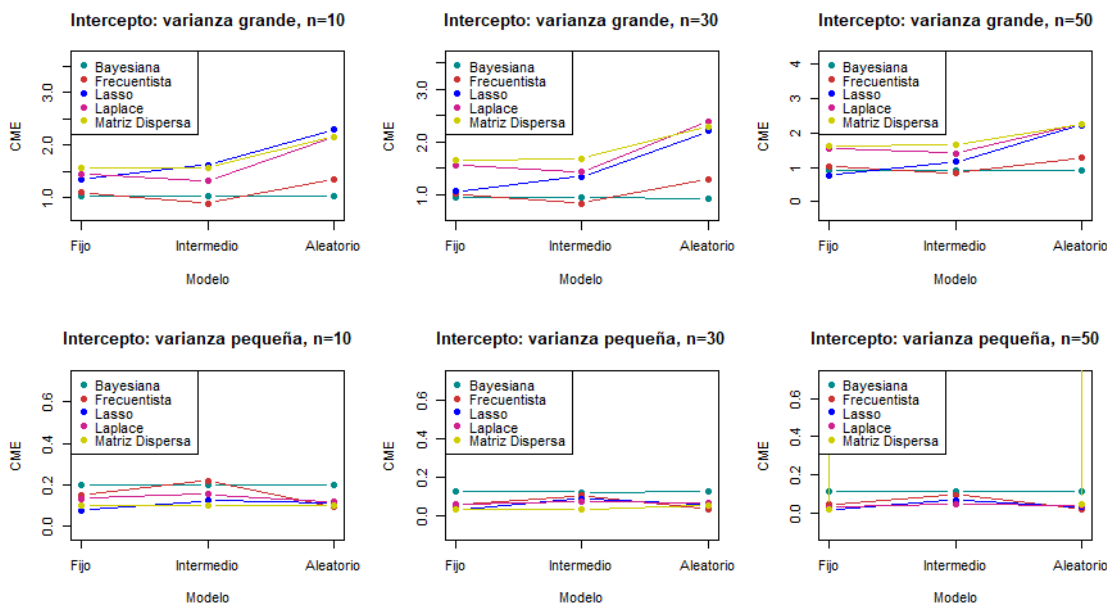


Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

## Anexo 6.2 Estimación del CME

Imagen A13

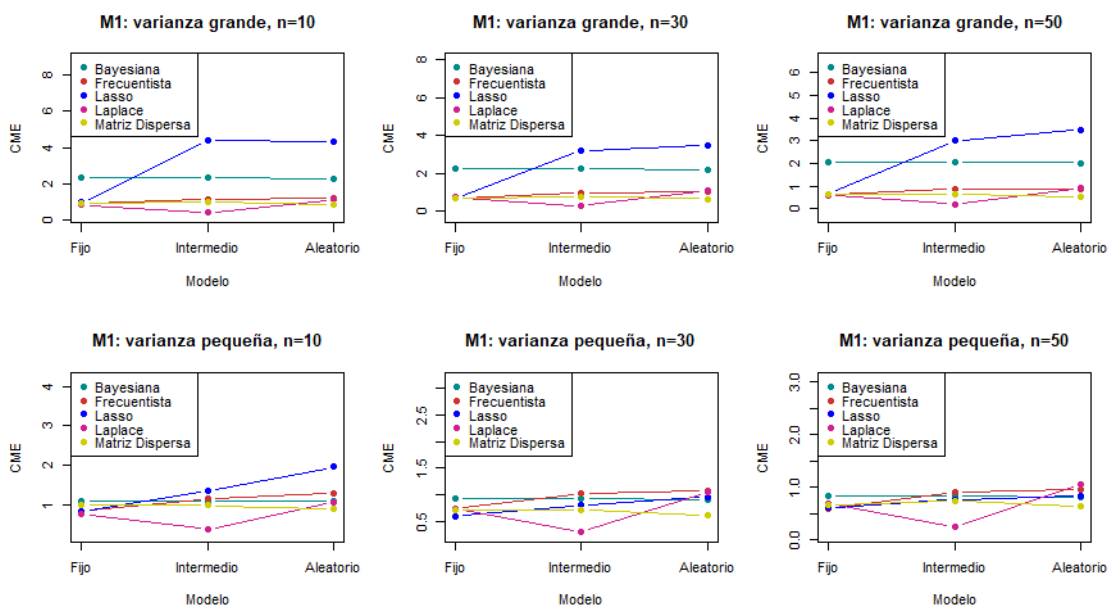
CME del coeficiente intercepto en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A14

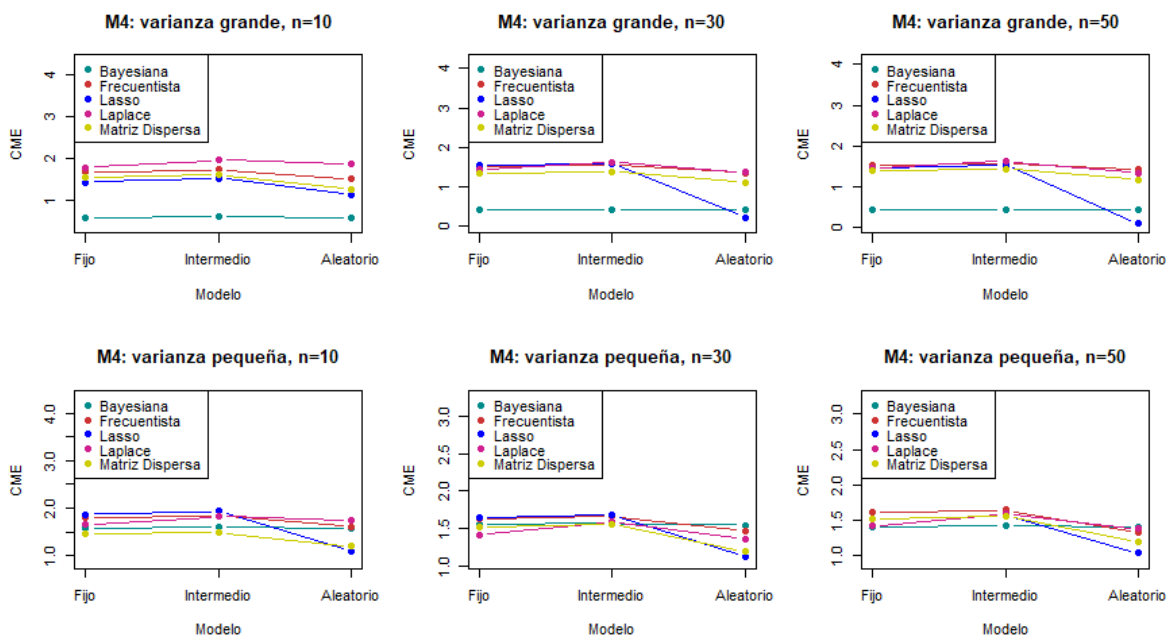
CME del coeficiente M1 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A15

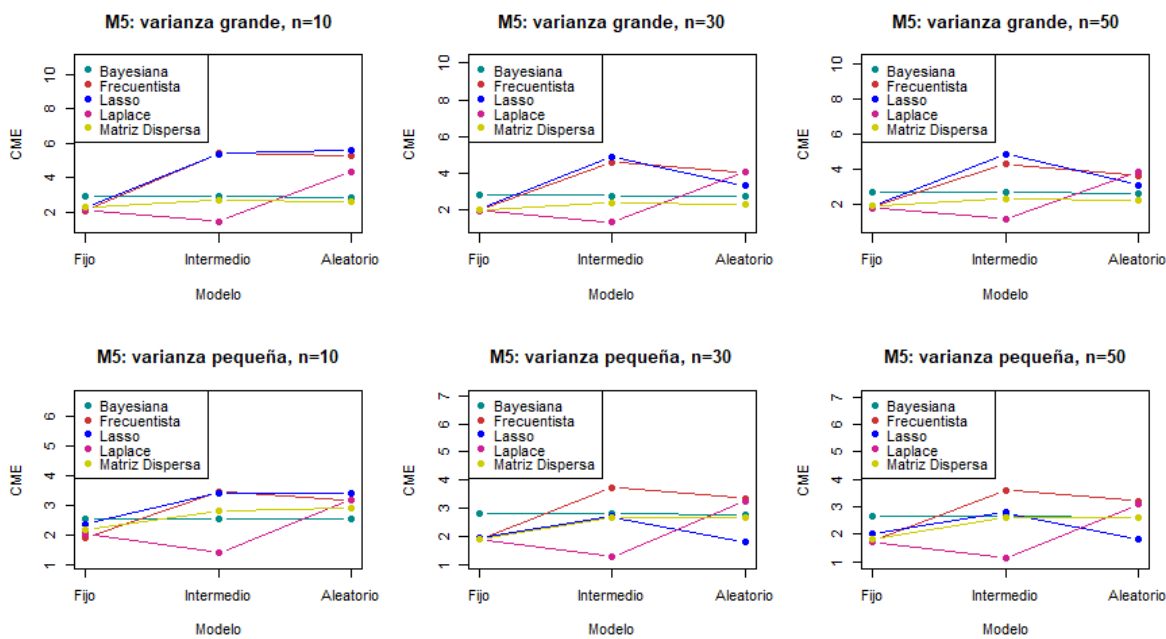
CME del coeficiente M4 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A16

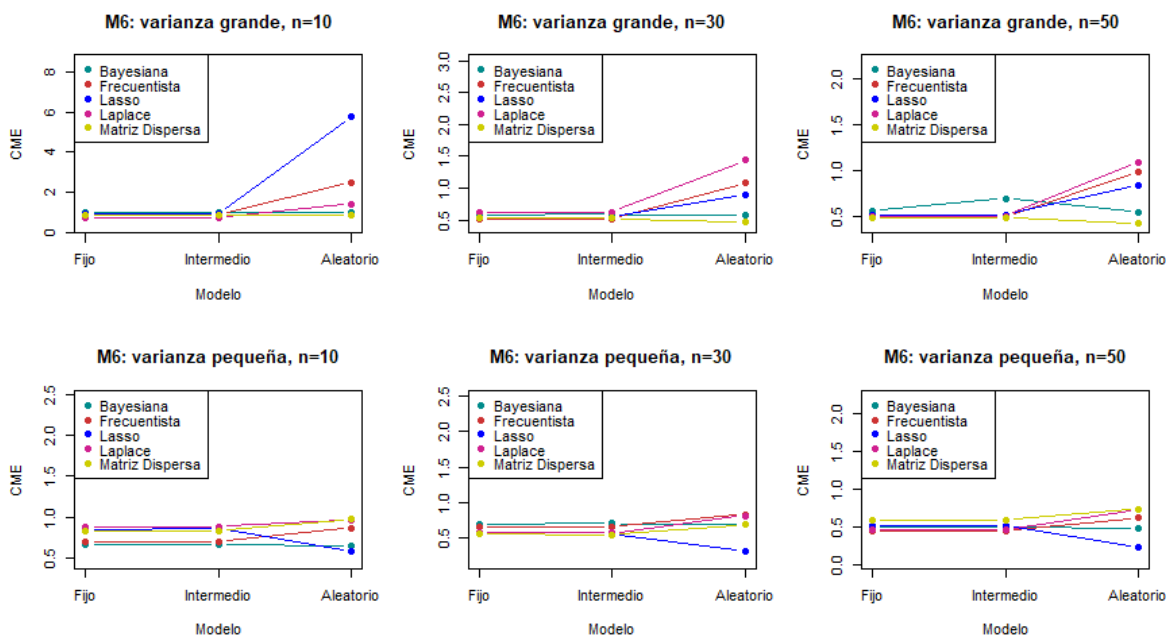
CME del coeficiente M5 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A17

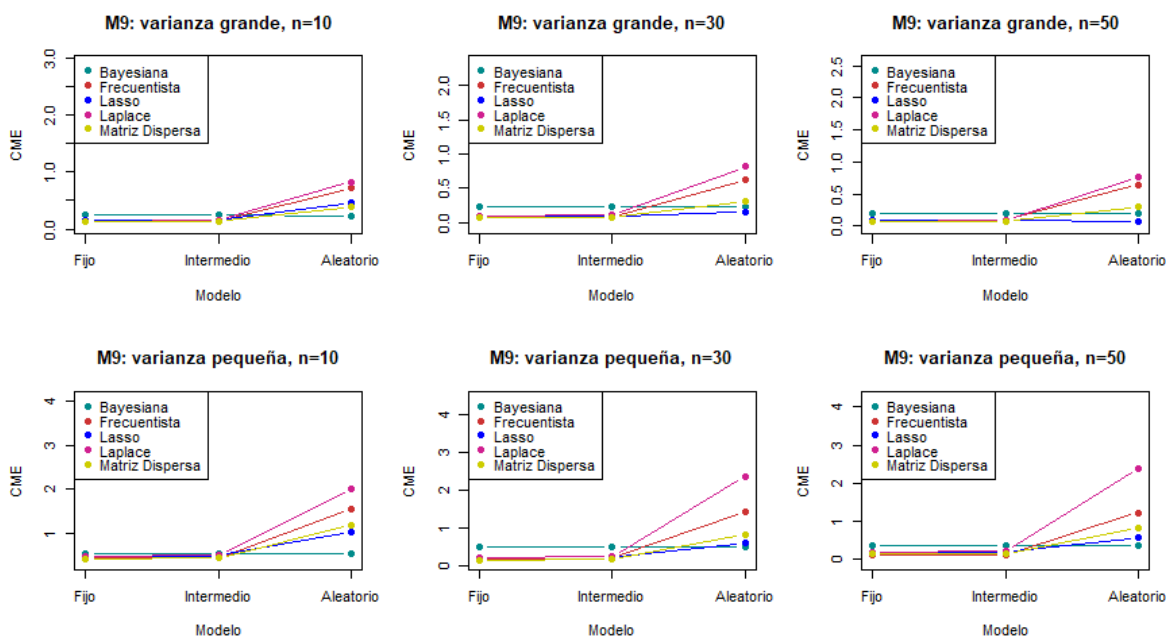
CME del coeficiente M6 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A18

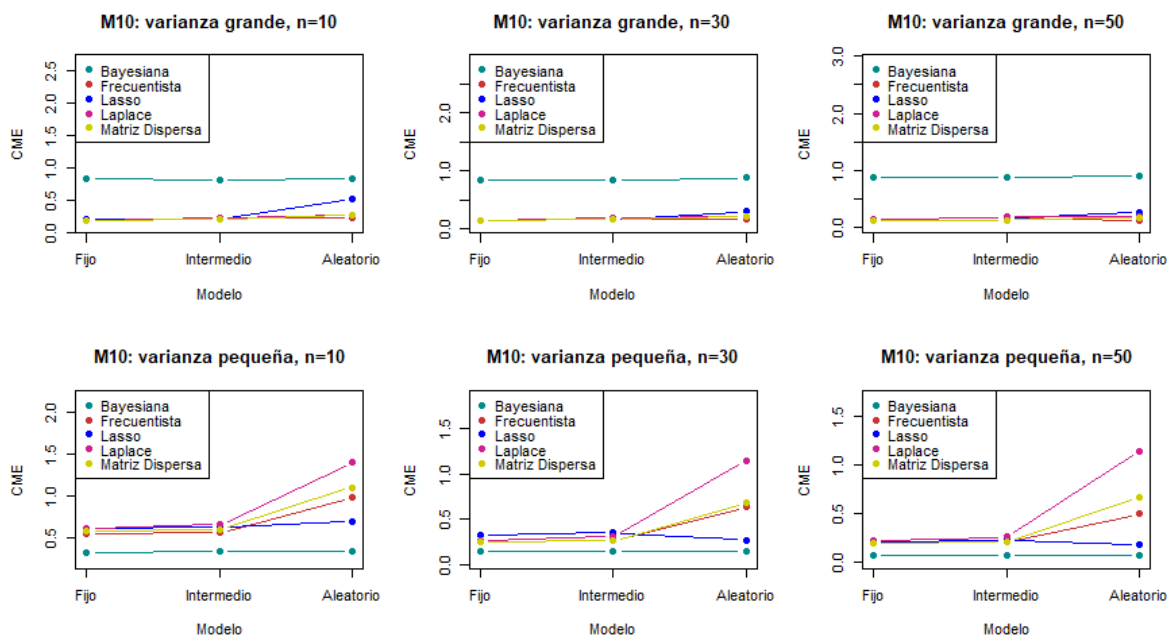
CME del coeficiente M9 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A19

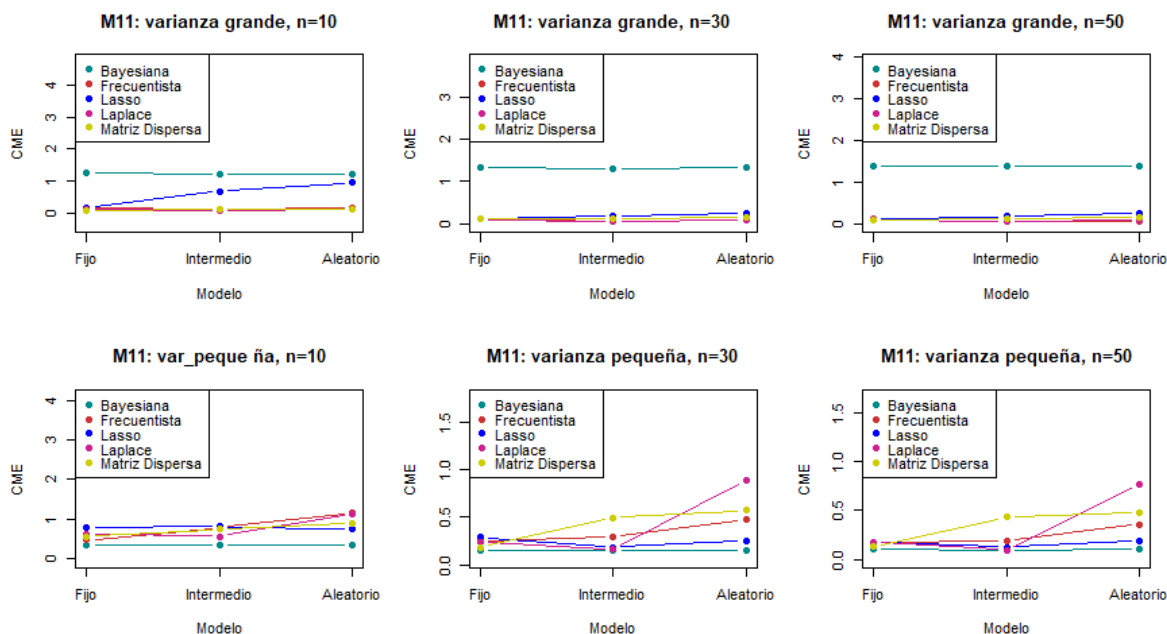
CME del coeficiente M10 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A20

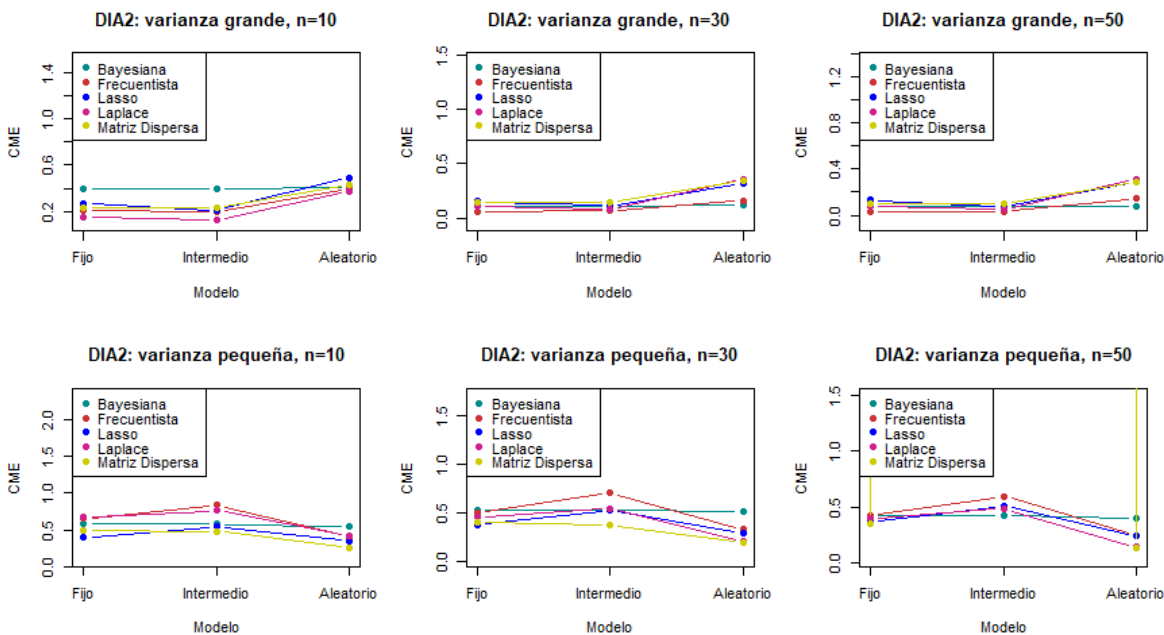
CME del coeficiente M11 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A21

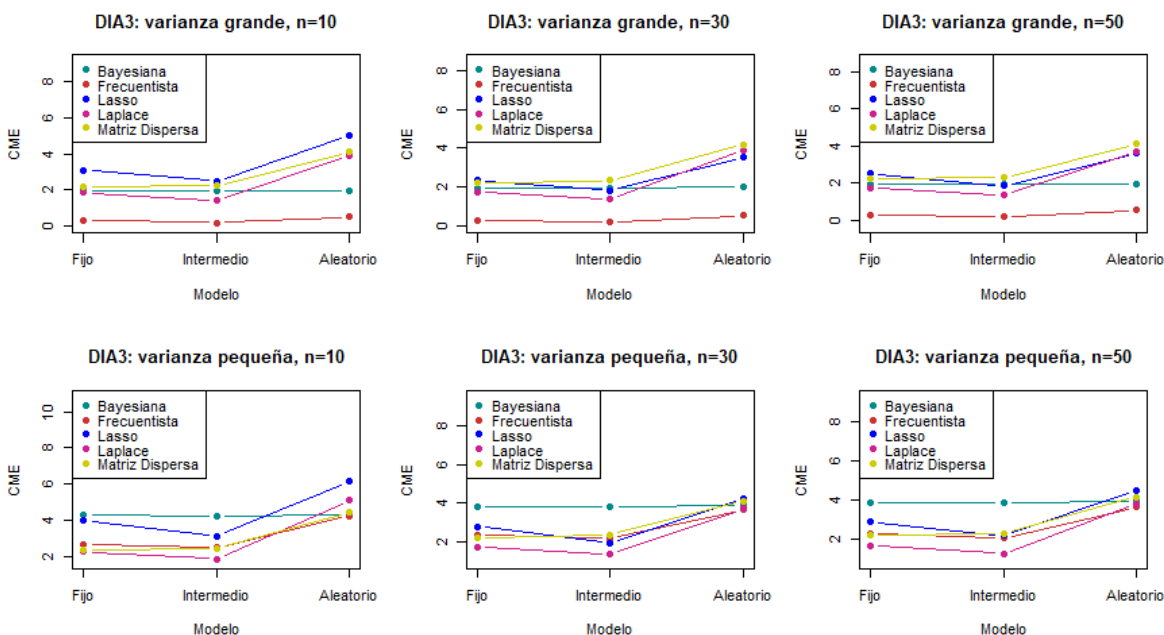
CME del coeficiente DIA2 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.

Imagen A22

CME del coeficiente DIA3 en los modelos de efectos fijos, efectos aleatorios y efectos aleatorios completos, según tipo de estimación, tamaño de muestra y varianza



Fuente: Elaboración propia a partir de datos del Instituto de Neurociencias de la Universidad de Costa Rica.