

## Grado en Estadística

---

**Título:** Estudio de la variabilidad otorgada al resultado de un tratamiento no farmacológico por causa del intervencionista

**Autor:** Tommy Harold Herrera Santos

**Director:** Erik Cobo Valeri, Roser Rius Carrasco

**Departamento:** Estadística e Investigación Operativa



UNIVERSITAT DE  
BARCELONA



UNIVERSITAT POLITÈCNICA DE CATALUNYA  
BARCELONATECH

Facultat de Matemàtiques i Estadística

## **RESUMEN**

En una intervención no farmacológica, donde el tratamiento sobre el paciente lo realiza un intervencionista, aparece una fuente de variabilidad, como resultado de las diferencias en los antecedentes y/o formación de los profesionales.

Debido a este factor en la variabilidad, y que las instituciones buscan mantener el anonimato de los intervencionistas, se busca una manera de cuantificar el impacto de este efecto, mediante la variabilidad que añade al modelo.

Este estudio analiza cómo variarán los resultados de un modelo de efectos mixtos, al modificar los parámetros de origen de los datos, mediante una serie de simulaciones con datos generados.

## **PALABRAS CLAVE**

Intervención no farmacológica, Intervencionista, Modelo de efectos mixtos, Variable de efecto aleatorio, Simulaciones

## TITLE

Study of the variability given to the result of a non-pharmacological treatment through the interventionist

## SUMMARY

In a non-pharmacological intervention, in which the patient is treated by an interventionist, a source of variability is added, since there are differences in the background or expertise of each professional.

Since this additional variability exists, and the institutions want to keep their professionals anonymous, there's a reason to find a way to quantify said variability, without associating it with a group of individuals. Since said effects from each individual won't be of use in the future, as they won't be the ones carrying out the interventions, an estimation of each interventionist's effect won't help, but knowledge of the variance between these individuals will. In addition, if this variability is too high, adjusting the standardization of the professionals' formation or the treatment's itself should be taken into consideration.

This study adjusts these interventions into a linear mixed effects model, in which the treatment and the effect from the interventionists are handled as fixed and random effects respectively. The objective is studying what statistical analysis may be more appropriate if the interventionist adds an additional component that affects the variability of the result.

A series of simulations with generated data will be done to analyze how the results vary depending on the parameters from the source of the data. The generated data simulates the systolic blood pressure (SBP) of patients treated with a therapy whose effects might vary between interventionists.

First, I introduce the topic and the associated concepts, such as the description of a non-pharmacological intervention or the types of effect according to explanatory variables. Next, I justify the use of the mixed effects Analysis of Variance (ANOVA) model. I implement the simulations with the R software; and finally, I interpret them.

Programming a function for the case allows modifying the values of the variables associated with the random effect, to analyze how the results are modified. In particular, the evolution of the result is studied according to (1) the number of interventionists, (2) the number of patients per interventionist, and (3) the standard deviation of the distribution of the effects of the interventionists.

Spotting patterns in the behavior of the results after changes in its parameters will help determine how impactful the variability coming from the interventionists is in the results.

R software is used for the necessary computation throughout the study. Specifically, the lmer package has been used, which allows the formulation of mixed effects models. For the visual representation in graphs, or tables, the *ggplot2* and *kable* packages have also been used.

## **KEYWORDS**

Non-Pharmacological Interventions, Interventionist, Mixed effects model, Random effects, Simulations

## **Clasificación AMS**

62-XX : Statistics

62PXX : Applications

62P10 : Applications to biology and medical sciences

## ÍNDICE

INTRODUCCIÓN.....	3
METODOLOGÍA.....	4
I - CONTEXTO Y JUSTIFICACIÓN DEL ESTUDIO .....	5
I.1 - Intervenciones no farmacológicas .....	5
I.2 - Intervencionista.....	5
I.3 - Privacidad del intervencionista .....	5
II - GUÍAS DE PUBLICACIÓN.....	6
II.1 - Reproducibilidad en la investigación de salud.....	6
II.2 - Red EQUATOR.....	6
II.3 - CONSORT y su extensión CONSORT-NPT.....	6
III - VARIABILIDAD POR CAUSA DEL DISEÑO EXPERIMENTAL.....	8
IV - EFECTOS DE VARIABLES. FIJOS Y ALEATORIOS.....	9
V - MODELO DE EFECTOS MIXTOS .....	10
VI – SIMULACIÓN DE DATOS GENERADOS.....	11
VI.1 - Contexto de los datos.....	11
VI.2 - Generación de datos .....	11
VI.3 - Modelo teórico para los datos generados.....	14
VI.4 - Ajuste del modelo de efectos mixtos para los datos generados .....	14
VI.5 - Interpretación de los coeficientes de efectos aleatorios .....	15
VII - ESTIMACIÓN DE LA DISTRIBUCIÓN MEDIANTE SIMULACIÓN.....	17
VIII - EVOLUCIÓN DE LOS RESULTADOS SEGÚN CAMBIOS EN LOS PARÁMETROS ..	18
VIII.1 – Resultados según el número de intervencionistas.....	19
VIII.1.1 - <i>Efecto fijo</i> .....	19
VIII.1.2 - <i>Efecto aleatorio</i> .....	21
VIII.1.3 – <i>Otras simulaciones</i> .....	23
VIII.2 – Resultados según el número de pacientes por intervencionista.....	24
VIII.2.1 - <i>Efecto Fijo</i> .....	24
VIII.2.2 - <i>Efecto Aleatorio</i> .....	26
VIII.2.3 – <i>Otras simulaciones</i> .....	27
VIII.3 - Resultados según la desviación típica del efecto del intervencionista.....	28
VIII.3.1 - <i>Omisión del efecto aleatorio</i> .....	28
VIII.3.2 - <i>Efecto fijo</i> .....	29
VIII.3.3 - <i>Efecto aleatorio</i> .....	30
CONCLUSIÓN .....	33

BIBLIOGRAFÍA..... 36  
ANEXOS ..... 38

## **INTRODUCCIÓN**

Con la implementación de intervenciones no farmacológicas, donde el tratamiento que se efectúa sobre el paciente viene dado por un intervencionista, a diferencia de un tratamiento farmacológico donde este factor adicional no aparece, habrá un componente adicional que afectará a la variabilidad del resultado. El impacto de dicha variabilidad cambiará dependiendo de las características de cada caso. En el caso de que dicho impacto sea grande, sería apropiado integrar cambios en la estandarización de la formación de los intervencionistas, o en el proceso del tratamiento en cuestión, para obtener así una reducción en esta variabilidad.

Para investigar esta variabilidad, se realiza una serie de simulaciones con datos generados, que simulan un caso realista de pacientes con alta presión arterial sistólica a los cuales se les trata con un tratamiento experimental. Se busca analizar los resultados de las simulaciones y así determinar qué cambios en los parámetros del modelo aumentan o reducen la variabilidad proveniente del efecto del intervencionista que realiza el tratamiento.

Inicialmente se introduce el tema y los conceptos asociados, como la descripción de una intervención no farmacológica o los tipos de efecto en variables explicativas.

Seguidamente se justifica el uso del modelo de efectos mixtos para el caso de intervención no farmacológica.

Se realiza la implementación de las simulaciones en el software R y finalmente se realiza una interpretación de los resultados.



## **METODOLOGÍA**

En el estudio se ajusta el resultado de una intervención médica no farmacológica como resultado de un modelo de efectos mixtos, en el que se trata al efecto asociado al tratamiento como un efecto fijo, y al efecto que proviene del intervencionista que realiza el tratamiento sobre el paciente como efecto aleatorio, con una población de niveles de tamaño desconocido.

Para la investigación del comportamiento de un modelo de efectos mixtos se han realizado simulaciones de valores generados. En estas simulaciones se modifican los valores de las variables asociadas al efecto aleatorio, para analizar así como se modifican los resultados. En particular se estudia la evolución del comportamiento del resultado según variaciones en el número de intervencionistas en la muestra, el número de pacientes asociados a cada intervencionista de la muestra, y por último la desviación típica de la distribución de la cual proceden los efectos de los intervencionistas.

Se ha utilizado el software R para la computación necesaria a lo largo de todo el estudio. En concreto se ha utilizado el paquete *lmer* que permite formular modelos de efectos mixtos. Para la representación visual en gráficos, o tablas, se han utilizado también los paquetes *ggplot2* y *kable*.

## **I - CONTEXTO Y JUSTIFICACIÓN DEL ESTUDIO**

En la introducción se ha indicado que el objetivo es estudiar el impacto del efecto proveniente del intervencionista en una intervención no farmacológica. Una vez contextualizados estos conceptos, surge la justificación de estudiar su variabilidad intentando evitar problemas de privacidad.

### **I.1 - Intervenciones no farmacológicas**

En medicina, a la hora de prevenir, diagnosticar, tratar o aliviar los síntomas de una enfermedad, se puede realizar un tratamiento que involucre fármacos o no, lo que denominamos intervenciones no farmacológicas, que incluyen cualquier intervención no química o biológica realizada sobre el paciente por el cuidador con el objetivo de obtener un beneficio relevante. Como las farmacológicas, pueden presentar posibles efectos secundarios, deben tener sustento teórico y deben estar definidas de forma replicable. Estas intervenciones, generalmente más sofisticadas o “complejas”, no permiten descartar el efecto variable del profesional involucrado: I@s intervencionistas.

### **I.2 - Intervencionista**

Intervencionista es cualquier especialista que realice procedimientos en las unidades de la población objetivo para diagnosticar, prevenir y tratar síntomas de múltiples dolencias. Entre ellos, se puede encontrar por ejemplo una cirujana o un fisioterapeuta, quienes realizan la intervención.

Aunque los procesos de las intervenciones se encuentran estandarizados, su complejidad puede incluir variaciones, quizás relacionadas con los niveles de aprendizaje, experiencia o formación previos, lo que podría provocar un efecto diferente para distintos intervencionistas.

Incluso, podría imaginarse un efecto diferente en cada una de sus intervenciones, ya que es prácticamente imposible replicar el mismo proceso múltiples veces de la misma manera, aunque dada la imposibilidad de repetir la misma intervención en el mismo individuo en las mismas condiciones, esta variabilidad será indistinguible de (“estará confundida con”) la variabilidad natural entre las pacientes.

### **I.3 - Privacidad del intervencionista**

Por contraparte, mientras la legislación no lo permita explícitamente, consideramos prudente disociar los resultados de los intervencionistas, para respetar su privacidad, y trataremos a los intervencionistas como un efecto aleatorio, que podríamos rebautizar como “anónimo”.

## II - GUÍAS DE PUBLICACIÓN

A la hora de realizar un estudio de investigación en la salud, existen una serie de guías que buscan estandarizarlos. La guía CONSORT-NPT se centra en intervenciones no farmacológicas, pero no indica la necesidad de mostrar la variabilidad proveniente del efecto intervencionista.

### II.1 - Reproducibilidad en la investigación de salud

La reproducibilidad es el principio esencial del método científico, y se refiere a la capacidad que tenga una prueba o experimento de ser reproducido o replicado. A nivel práctico, si queremos aplicar una intervención en futuros pacientes, necesitamos que esos resultados se reproduzcan. La Ciencia siempre estudia todas las curaciones, pero aquellas que no logra reproducir (“milagros”), las abandona.

Centrado en la investigación de la salud, los organismos pueden mostrar mucha variación, lo que dificulta la replicación exacta de la investigación. Por ende, la reproducibilidad es un proceso de intentar hacer el experimento lo más reproducible posible, asegurando que el investigador pueda defender su posición.

### II.2 - Red EQUATOR

Con el objetivo de estandarizar y hacer reproducibles los informes de investigación de salud, la red EQUATOR (*Enhancing the QUALity and Transparency Of health Research*) formada por investigadores, editores de revistas científicas, metodólogos y representantes de todas las partes implicadas como la población atendida, ha desarrollado una serie de guías para diferentes áreas de investigación. Además de listar los puntos esenciales (“*checklist*”) de los artículos, busca también la precisión y transparencia de los mismos, para alcanzar la transparencia.

### II.3 - CONSORT y su extensión CONSORT-NPT

La guía CONSORT orienta los informes de ensayos aleatorizados, o *randomized controlled trials*, en concreto se centra en el diseño más habitual, de dos grupos o pruebas paralelas de los llamados estudios finales, confirmatorios o “pivote” usados para solicitar su autorización a las agencias de regulación del mercado farmacológico (FDA, EMEA, etc).

Para algunos casos más concretos, CONSORT va desarrollando extensiones.

La extensión para intervenciones no farmacológicas, o CONSORT-NPT es la guía dirigida a estudios no farmacológicos, como pueden ser los relacionados con cirugía, rehabilitación o psicoterapia.

En las páginas web de Equator o CONSORT y otras guías se pueden encontrar diferentes herramientas para lograr un estudio útil. Entre éstas, se encuentra el *checklist* o lista de comprobación para verificar que se ha informado de los puntos esenciales. Entre los puntos añadidos en CONSORT-NPT relacionados con los intervencionistas, se encuentran los siguientes:

*13.a - "The number of care providers or centers performing the intervention in each group and the number of patients treated by each care provider or in each center."*

*15 - "When applicable, a description of care providers (case volume, qualification, expertise, etc.) and centers (volume) in each group."*

*20 - "In addition, take into account the choice of the comparator, lack of or partial blinding, and unequal expertise of care providers or centers in each group."*

Aunque se tiene en cuenta que hay un efecto que considerar, relacionado con las diferencias entre cada uno de los intervencionistas, no se solicita ningún parámetro concreto que determine o cuantifique el impacto de este efecto intrínseco de una intervención no farmacológica, donde un intervencionista se encuentra involucrado. Lo que abordaré en este trabajo.

### III - VARIABILIDAD POR CAUSA DEL DISEÑO EXPERIMENTAL

El problema fundamental de la inferencia causal reside en la imposibilidad de observar a la vez y en las mismas condiciones la respuesta a la intervención en estudio y a su referencia, comparador o control. Al posicionar al individuo únicamente en uno de los dos grupos, se estará perdiendo el efecto potencial que tendría al recibir el otro tratamiento. Si se asigna la intervención en estudio, la respuesta al control no podrá observarse, será un valor perdido o "*missing*". Si la decisión de que intervención asignar se tomó al azar, será un "*missing at random*", lo que permite usar la inferencia estadística para estimar el efecto de esa intervención.

Por supuesto, la asignación al azar se puede combinar con otras técnicas experimentales, como la estratificación, que consiste en colocar a individuos de características similares en estratos en los que se observan los resultados de los procesos en comparación. Esto se logra parcialmente con métodos de muestreo apropiados, que distribuyan los grupos homogéneamente respecto a estas características de los pacientes.

De la misma manera, en una intervención no farmacológica existe una variación que proviene del intervencionista, y el cambio potencial que tendría el efecto, en el caso que la intervención hubiese sido realizada por otro intervencionista. De igual modo al caso previo, un mismo individuo, no podrá recibir una misma intervención por diferentes intervencionistas.

#### **IV - EFECTOS DE VARIABLES. FIJOS Y ALEATORIOS**

En un modelo de regresión, el resultado esperado de cada individuo vendrá determinado por las características que lo describen, dichas características se almacenarán en una serie de variables explicativas.

Estas variables explicativas (véanse más detalles en Anexo 1) estarán formadas por diferentes niveles, y cada uno de ellos tendrá un efecto diferente en el resultado. En el caso de las variables de efectos fijos, dichos niveles pertenecerán a una población de opciones posibles conocidas, y todas ellas serán estudiadas y medidas o estimadas.

Para el caso de las intervenciones no farmacológicas, el resultado de una intervención también vendrá determinado por varias variables explicativas. Entre ellas, se encuentra la persona que realiza la intervención. Se podría decir que el efecto de cada uno de los intervencionistas provendrá de un conjunto de niveles teóricamente infinito (ya que irá aumentando con el tiempo), y que cada uno de ellos tendrá un efecto diferente (aunque similar), añadiendo así el factor "Intervencionista" una varianza mayor a la variable respuesta en la que observamos el efecto.

Teniendo en cuenta esta descripción de la variable que determina el intervencionista, no será adecuado tratarla como una variable de efecto fijo, sino como una variable de efecto aleatorio, cuya descripción se ajusta al caso en cuestión.

Debido a que nos encontramos con este número indefinido de niveles en el efecto aleatorio, el estudio del efecto de cada uno de ellos no nos será de mucha utilidad, teniendo en cuenta que estos intervencionistas del presente no serán quienes sigan haciéndolo dentro de x años. Lo que sí que puede resultarnos útil será el efecto que tendrá este factor en la varianza. Un buen indicador para determinar este efecto puede ser la proporción de dicho factor respecto a la varianza total.

A la hora de estudiar los efectos de esta población indefinida de niveles, será más indicado utilizar el análisis de varianza de un modelo que contenga tanto efectos aleatorios como efectos fijos, donde se tendrá en cuenta esta restricción de la cantidad y la selección aleatoria de niveles a estudiar respecto a la población de niveles del factor intervencionista, además de añadir el efecto fijo de tratamiento. Dicho modelo que incluye tanto factores fijos como aleatorios se llama modelo de efectos mixtos.

## V - MODELO DE EFECTOS MIXTOS

Un modelo de efectos mixtos se caracteriza por contar con variables explicativas de efectos fijos, y efectos aleatorios simultáneamente.

En un caso que cuenta con “N” individuos, “p” variables de efecto fijo, “q” variables de efecto aleatorio y “J” niveles que engloban tanto los niveles de los efectos aleatorios estudiados como los que no se han tenido en cuenta, el modelo se expresa de la siguiente manera:

$$y = \beta_0 + X\beta + Zu + \varepsilon$$

Siguiendo esta notación:

$y$  : Vector de longitud  $N*1$  con contiene la variable resultado.

$\beta_0$  : Valor intercept, escalar que representa el valor medio de la variable respuesta cuando todas las variables explicativas tienen valor 0.

$X$  : Matriz de tamaño  $N*p$  que contiene los valores de cada individuo para cada variable explicativa de efectos fijos “p”.

$\beta$  : Matriz de tamaño  $N*p$  que contiene los valores estimados del coeficiente asociado a cada variable explicativa de efectos fijos “p”.

$Z$  : Matriz de tamaño  $N*qJ$  que contiene los valores de cada individuo para cada variable explicativa de efectos aleatorios “q”, que pertenecerán al conjunto de niveles J.

$u$  : Matriz de tamaño  $N*qJ$  que contiene los valores estimados del coeficiente asociado a cada variable explicativa de efectos aleatorios “q”, que pertenecerán al conjunto de niveles J.

$\varepsilon$  : Matriz de tamaño  $N*1$  que contiene el residuo de la predicción del modelo para cada individuo, o diferencia entre el resultado real y la predicción  $X\beta + Zu$ .

Hay que tener en cuenta que, para las variables de efecto aleatorio, el número de niveles “J” puede ser un valor conocido o desconocido.

## **VI – SIMULACIÓN DE DATOS GENERADOS**

Para estudiar la variabilidad que produce el efecto de un intervencionista en una intervención no farmacológica, se realiza un estudio de unos datos creados bajo parámetros fijos de efectos de tratamiento e intervención.

Se realizará la computación en el software R.

### **VI.1 - Contexto de los datos**

Este caso intenta simular un estudio sobre un tratamiento sobre individuos que padecen de algún tipo de problema cardiaco y el valor de la variable respuesta será la presión arterial sistólica.

La presión arterial sistólica (PAS) indica cuánta presión ejerce la sangre contra las paredes de la arteria cuando el corazón late, es la cifra superior (y más alta) en una medición de la presión arterial. Para pacientes adultos, valores por debajo de 120 mmHg se consideran normales, y valores por encima de este valor se pueden considerar como una presión arterial alta. El tratamiento en cuestión buscará reducir este valor.

### **VI.2 - Generación de datos**

Para realizar una simulación inicial, los datos tendrán las siguientes características:

- Hay un total de 50 individuos de los cuales se extrae su valor de presión arterial sistólica al final del seguimiento, cuando se considera que ya se han manifestado los efectos de la intervención.
- La mitad de los individuos (25) pertenecerán al grupo de control, por lo que no tendrán un intervencionista asociado. Por otro lado, la otra mitad de individuos, habrán recibido un mismo tratamiento no farmacológico. Se dividen en 5 grupos de tamaño 5, donde en cada uno de ellos los pacientes recibirán el tratamiento por un intervencionista diferente.

Los valores de frecuencia cardiaca tendrán una distribución normal de media 135 y desviación típica 7, que permiten reproducir situaciones razonables en salud, con aproximadamente casi todos los valores entre 120 y 150.

El efecto de recibir el tratamiento se tratará como un efecto fijo que reduce la variable respuesta en 10 unidades, efecto que podríamos considerar entre mediano o grande.

Cada uno de los 5 intervencionistas tendrá un efecto aditivo asociado de distribución normal de media 0 y desviación típica de 5 (para la primera simulación). Esto representa que habrá intervencionistas que obtendrán diferentes efectos en los pacientes. Algunos de ellos serán



mejores (en este caso, valores negativos, que bajan la PAS), y otros no tanto. Tiene valor esperado de 0 ya que se espera que la intervención tenga el efecto de reducir en 10 unidades la variable respuesta, independientemente del intervencionista. La desviación típica del intervencionista, de 5, la he escogido para representar un efecto intervencionista que casi nunca alcance a anular el efecto de la intervención, ya que:

$$| \text{Normal}(0, 5) | \leq 10$$

en la mayoría (95%) de los casos.

Por lo tanto, a los valores generados con la Normal(135, 7), habrá que añadirle tanto el efecto del tratamiento, como el efecto del intervencionista que corresponda a cada individuo.

El *dataset* creado contará con 3 columnas:

- *bp* : Variable respuesta, numérica. Se trata de la presión arterial sistólica del individuo, en mmHg.
- *tract* : Variable explicativa, categórica y dicotómica. Indica si el paciente ha recibido o no el tratamiento. Tendrá valor 1 si ha recibido el tratamiento, y valor 0 en caso contrario. Se trata de un efecto fijo ya que tiene dos únicos niveles y se estudian ambos.
- *int* : Variable explicativa, categórica y politómica. Indica el intervencionista que ha hecho el tratamiento en el individuo en cuestión. Tiene 6 niveles, "none" en caso de que no haya recibido el tratamiento, ya que no hay intervencionista asociado, y valores "a, b, c, d, e" para indicar los 5 intervencionistas. Se trata de un efecto aleatorio ya que se estudian 5 niveles de una población indefinida.

```
#----- CREACIÓN DE DATOS -----
set.seed(1235)
#etiqueta de tratamiento
tract<- rep(c(0,1),each=25)
int <- rep(c("none", "a", "b", "c", "d", "e"), c(25, 5, 5, 5,5, 5))
#valores de una distribución normal con valores entre 120 y 150
dat <- rnorm(50, mean=135, sd=7)
# efecto de intervencionista
ef_int <- rnorm(5, mean=0, sd=5)
# añadir efecto del tratamiento e intervencionista
dat[26:50] <- dat[26:50]-10 + rep(ef_int, each=5)
tabla <- data.frame(bp=round(dat, 2) , tract=as.factor(tract), int=as.factor(int))
```

La tabla creada tiene la siguiente forma:

VI.2 – 1. Estructura de la tabla de datos generados

bp tract int		bp tract int	
130.11	0 none	122.71	1 d
126.01	0 none	133.87	1 e
141.93	0 none	126.18	1 e
135.78	0 none	135.66	1 e
135.80	0 none	138.36	1 e
146.89	0 none	156.48	1 e

Tabla VI.2 – 1. El valor de la presión arterial sistólica de cada individuo se encuentra en la variable “bp”, la variable “tract” es una variable dicotómica que indica si el paciente se encuentra en el grupo control o tratamiento, y la variable “int” indica el intervencionista que realiza la intervención.

Los valores generados tienen la siguiente distribución, estratificados según si pertenecen al grupo de control ( $tract=0$ ) o al grupo de tratamiento ( $tract=1$ ).

VI.2 –2. PAS según tratamiento

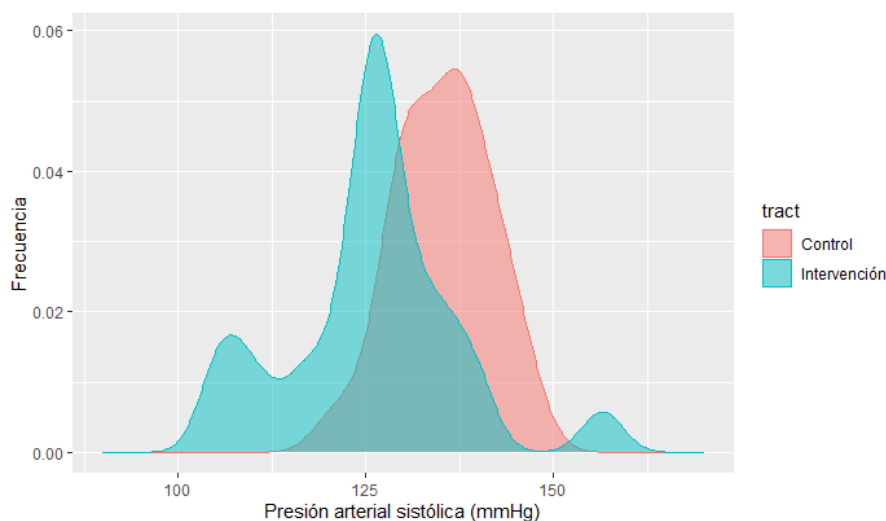


Figura VI.2 –2. En rojo, distribución de la PAS en el grupo control, con valores centrados en 135; y en azul está la del grupo tratado, centrado en 125 con mayor variabilidad.

Con esta generación aleatoria de datos, por ejemplo, se puede detectar como el grupo de pacientes del grupo control ( $tract=0$ ) tiene una distribución de valores con menor desviación (7), y centrada en valores más altos (135), mientras que entre los valores del grupo de pacientes que reciben el tratamiento se encuentran valores centrados en un valor menor (125), además que cuentan con una mayor desviación ( $7+5=12$ ).

Por lo tanto, un valor de presión arterial sistólica de un individuo en el grupo de control seguirá la siguiente distribución:

$$X_{tract=0} = x_0, x_0 \sim \text{Normal}(135, 7)$$

Y un paciente del grupo que recibe el tratamiento, tendrá un valor generado por la siguiente distribución:

$$X_{\text{tract}=1} = x_0 - 10 + z, \quad x_0 \sim \text{Normal}(135, 7), \quad z \sim \text{Normal}(0, 5)$$

### VI.3 - Modelo teórico para los datos generados

Bajo las características explicados, el modelo teórico será el siguiente:

$$y = \beta_0 + X\beta + Zu + \varepsilon$$

Donde  $\beta$  será el coeficiente de la variable tratamiento (efecto fijo) que contará con 2 niveles,  $u$  serán los coeficientes de la variable categórica del efecto aleatorio, que contará con 5 niveles observados, de un conjunto desconocido de "J" niveles. Generaremos N=50 resultados.

### VI.4 - Ajuste del modelo de efectos mixtos para los datos generados

Se utiliza el paquete de R "lme4", ya que contiene la función "lmer" que permitirá indicar si los efectos que afectan al resultado son fijos o aleatorios.

Una vez generados los datos simulados, se ajusta un modelo mixto con la función *lmer*, con la variable *bp* como variable respuesta, y las variables *tract* y *int* como variables explicativas de efectos fijos y aleatorios respectivamente.

```
mod <- lmer(bp ~ tract + (1|int), data=tabla)
```

El resumen del modelo obtenido por R es el siguiente:

VI.4-1. Salida de R del modelo generado

```
Linear mixed model fit by REML ['lmerMod']
Formula: bp ~ tract + (1 | int)
Data: tabla

REML criterion at convergence: 346.9

Scaled residuals:
  Min       1Q   Median       3Q      Max
-2.38042 -0.65526  0.03282  0.48569  2.64415

Random effects:
 Groups   Name      Variance Std.Dev.
 int      (Intercept) 52.44    7.241
 Residual                    61.30    7.830
Number of obs: 50, groups: int, 6

Fixed effects:
              Estimate Std. Error t value
(Intercept) 135.298     7.409  18.262
tract1      -9.498     8.236  -1.153
```

Figura VI.4–1. El bloque del efecto aleatorio indica las varianzas asociadas al efecto tratamiento y la varianza residual. Y el de efectos fijos, las estimaciones de los coeficientes de la constante y del efecto tratamiento, junto con sus errores estándar.

La salida de R presenta los resultados en dos partes: efectos fijos (*fixed effects*) y efectos aleatorios (*random effects*). La parte de efectos fijos se asemeja a la de un modelo lineal, con el valor estimado del efecto del tratamiento, -9.498, próximo al teórico usado para generar los datos, -10.

El valor de la constante, 135.298, también se aproxima al valor teórico, 135, con un resultado aproximado de. También coincide su desviación típica 7.409 con el valor fijado 7.

Por otro lado, en la parte de efectos aleatorios, se puede ver la variancia del efecto intervencionista, “*int*” es un 46.1% ( $52.44 / (61.30 + 52.44)$ ) de la variancia total ( $61.30 + 52.44 = 113.74$ ).

En algunas simulaciones, la función considera que el efecto aleatorio es suficientemente pequeño como para que se pueda retirar, por lo que pasa a eliminar el efecto aleatorio de intervencionista. Aparece el siguiente *warning*:

*boundary (singular) fit: see help('isSingular')*

Dependiendo de los parámetros del modelo, variará la frecuencia con la que ocurre este suceso.

## VI.5 - Interpretación de los coeficientes de efectos aleatorios

Mediante la función *ranef()* del paquete *lme4*, se pueden extraer las estimaciones de los efectos de cada uno de los niveles de las variables de efecto aleatorio. En este caso, estos son las estimaciones de cada efecto:

VI.5– 1. Estimación de coeficientes de intervencionistas

Intervencionista	Estimación	ValorReal
a	-1.487	1.573
b	-7.938	-10.076
c	1.318	0.943
d	-1.869	0.258
e	9.977	8.056
none	0.000	0.000

Tabla VI.5– 1. La tabla cuenta con las estimaciones de los coeficientes junto con los valores reales asociados a cada uno de los intervencionistas, además del nivel “none” que indica el efecto nulo de los pacientes del grupo control (*tract* = 0).

Las estimaciones del efecto de cada uno de los intervencionistas son valores cuya importancia puede ser vital para esos pacientes y quizás podrían interesar a un gestor de la calidad de la

atención sanitaria, pero desde el punto de vista científico su importancia es relativamente nula. Como se trata de un efecto aleatorio, cada uno de los niveles observados, no tienen por qué ser los niveles que se observarán en el futuro. Por ello, el efecto que tiene cada uno de ellos no resulta tan importante. Por otro lado, las propiedades de la varianza de la variable que se pueden extraer del modelo sí es una información que se podrá trasladar a un estudio futuro, con niveles diferentes de la variable.

## VII - ESTIMACIÓN DE LA DISTRIBUCIÓN MEDIANTE SIMULACIÓN

Para lograr una estimación más ágil, he creado una función que genera múltiples(n) simulaciones, para observar cómo fluctúan los parámetros estimados.

En ella se añade un bucle en el que se crea un conjunto aleatorio de datos con los mismos parámetros para cada uno de las iteraciones. De cada uno de los modelos creados con cada conjunto de valores, se estiman los parámetros de interés, como la variancia del efecto aleatorio, o las estimaciones del valor del efecto fijo de tratamiento.

Con ello, analizaré su distribución. Por ejemplo, a continuación, un gráfico de la distribución que corresponde a la proporción de varianza que pertenece al efecto aleatorio de intervencionista en un bucle de 300 iteraciones, manteniendo los valores del punto anterior.

VII – 1. Distribución de la proporción de varianza debida al efecto aleatorio

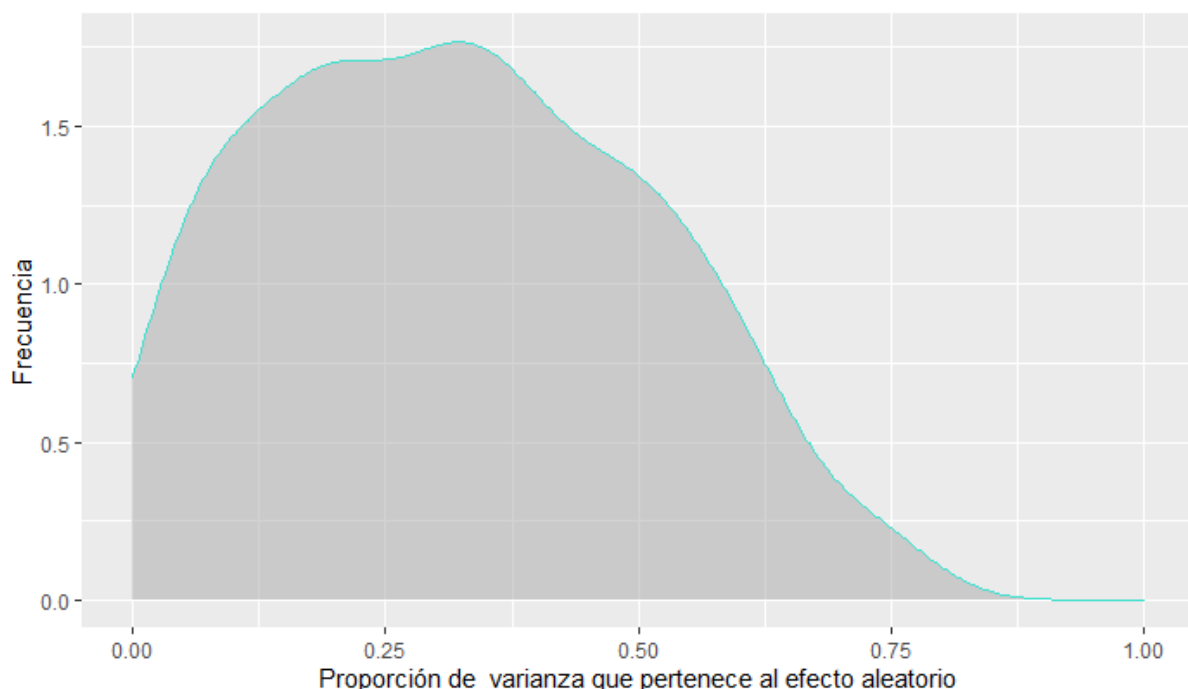


Figura VII – 1. Gráfico de densidad de la distribución de la proporción de varianza asociada al efecto intervencionista. Aunque todos los valores provienen de una misma distribución, con una proporción teórica de  $(5/12)$  0.416, el rango de valores varía entre 0 y 0.8.

Al realizar el análisis repetidas veces, hay una proporción de casos en los que se retira el efecto aleatorio en el modelo. En este caso en concreto, un 11% del total (33 casos). En el gráfico no se muestran; por lo que la proporción de 0 podría ser mayor.

## VIII - EVOLUCIÓN DE LOS RESULTADOS SEGÚN CAMBIOS EN LOS PARÁMETROS

Analizaré ahora si las variaciones en los resultados dependen de los parámetros usados en las simulaciones.

No todos los parámetros del modelo influirán en las estimaciones del efecto aleatorio. Estudiamos 3 parámetros:

- Número de pacientes por intervencionista.
- Número de intervencionistas.
- Desviación típica del efecto del intervencionista.

Por otro lado, los siguientes parámetros del modelo son intrínsecos de la naturaleza de los datos generados y ajenos a los intervencionistas:

- Media de la distribución normal de la que provienen los datos.
- Desviación típica de la distribución normal de la que provienen los datos.
- Valor del efecto fijo del tratamiento.
- Media de la distribución normal de la que provienen los efectos de intervencionista.

Para analizar los resultados, se modificará la función creada anteriormente para almacenar los resultados para diferentes valores de los parámetros de interés. Una vez parametrizada permitirá al usuario elegir los valores para los parámetros del modelo.

Por lo tanto, la función podrá calcular las distribuciones de los parámetros para diferentes valores de cada una de las variables de interés, para así observar la evolución de los resultados al cambiar dichos parámetros.

Variaremos (1) el número de pacientes por intervencionista, (2) el número de intervencionistas y (3) la desviación típica del efecto del intervencionista. En cada ejecución mantendré fijos 2 de ellos e iré variando el tercero. Las dos constantes los fijaré a 10, 10 y 5 respectivamente.

En el Anexo 3 se encuentra el código de R utilizado para recrear las simulaciones, y en el Anexo 4 está el código de R para la creación de los gráficos.

## VIII.1 – Resultados según el número de intervencionistas

El primer caso a estudiar será el que varíe el número de intervencionistas.

A diferencia de los parámetros fijos del modelo, el número de intervencionistas no modifica el contexto de los datos, sino que nutre la información, ya que al recibirla de más o menos intervencionistas, variará la calidad de estimación de los parámetros.

Inicio la simulación para los valores  $value=5*i$ , donde  $i \in \{5, 6, \dots, 15\}$ . Cada uno de los intervencionistas tendrá 10 pacientes asignados; y el efecto del intervencionista se distribuye según una distribución Normal(0,5). Para mantener el balanceo de la proporción de pacientes con y sin tratamiento, aumentará el número de pacientes sin tratamiento.

En esta simulación no se encuentra ningún caso en que la función elimine el efecto aleatorio.

### VIII.1.1 - Efecto fijo

En cuanto a la parte de efectos fijos del modelo, estudio la constante y el efecto de tratamiento, junto con sus errores estándar.

VIII.1.1 –1. Evolución del estimador de la constante del modelo

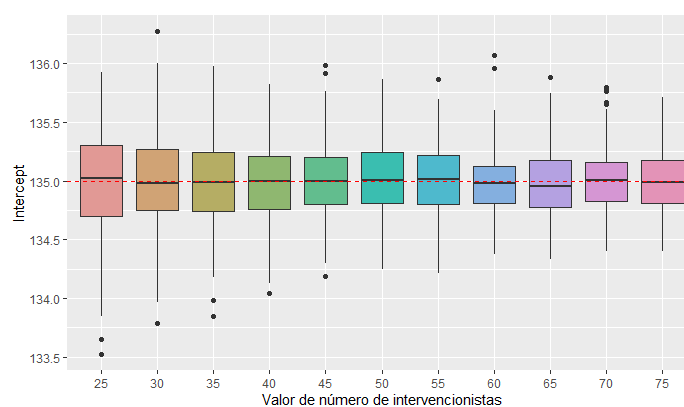


Figura VIII.1.1 –1. La línea discontinua horizontal roja representa el valor teórico (135). Los valores estimados se encuentran centrados a su alrededor; y la precisión de la estimación aumenta cuando se aumenta el número de intervencionistas.



### VIII.1.1 – 2. Evolución de la estimación del efecto del tratamiento

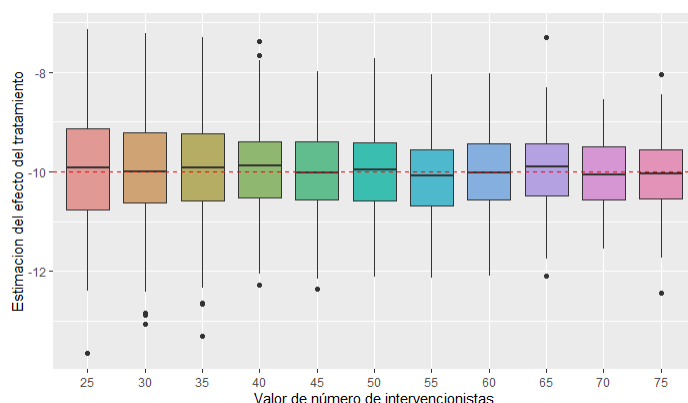


Figura VIII.1.1 – 2. Los valores estimados se centran alrededor del valor teórico (-10, línea discontinua horizontal roja); y su precisión aumenta al aumentar el número de intervencionistas.

En estos primeros *boxplots*, se ve como ambos estimadores se centran alrededor de sus valores teóricos (135 y -10), independientemente del número de intervencionistas. Por otro lado, el rango de posibles valores se reduce, por lo que aumenta su precisión. Más adelante realizo una simulación con un número aún mayor de intervencionistas, para observar si estos límites de los intervalos se siguen reduciendo.

En estas simulaciones el nº de pacientes por intervencionista sigue fijo, por lo que, al aumentar el número de intervencionistas, aumenta el número total de individuos N:

$$N = 10 * value * 2, \quad value = 5 * i, \quad i \in \{5, 6, \dots, 15\}$$

Ya que ni la constante  $\beta_0$  ni el efecto intervencionista  $\beta$  varían con la evolución del número de intervencionistas, sus valores esperados permanecen constantes. Pero al aumentar N disminuyen sus oscilaciones, con intervalos más reducidos, como anticipa el TLC.

A continuación, el error estándar de ambos estimadores.

### VIII.1.1 – 3. Evolución del error estándar de la estimación de la constante del modelo

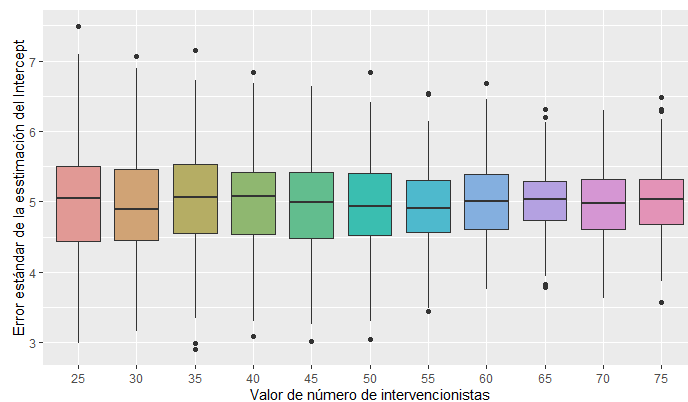


Figura VIII.1.1 – 3. Error estándar de la estimación de la constante: la precisión aumenta con el número de observaciones.

#### VIII.1.1 – 4. Evolución del error estándar de la estimación del efecto del tratamiento

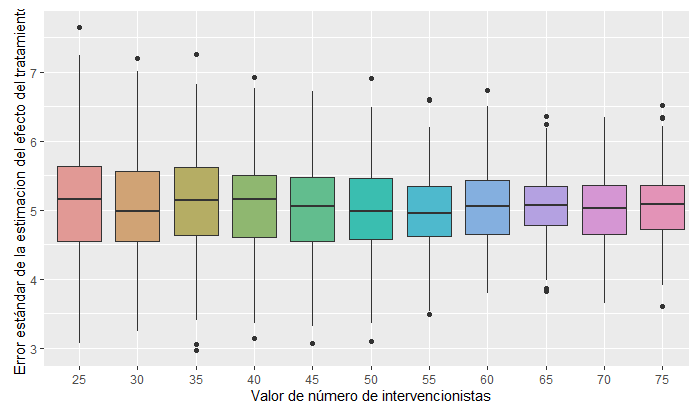


Figura VIII.1.1 – 4. Error estándar de la estimación del del efecto del tratamiento: la precisión mejora al aumentar el nº de observaciones

Los dos estimadores tienen un comportamiento prácticamente idéntico en cuanto a error estándar, con valores alrededor de 5, y cuyos intervalos se reducen cuando la N aumenta.

#### VIII.1.2 - Efecto aleatorio

En un modelo que contiene tanto efectos fijos como aleatorios, cuando se habla de varianza, habrá una parte de ella que vendrá explicada por parte del efecto aleatorio del modelo, y la restante será considerada como residual.

#### VIII.1.2 – 1. Evolución la desviación típica residual

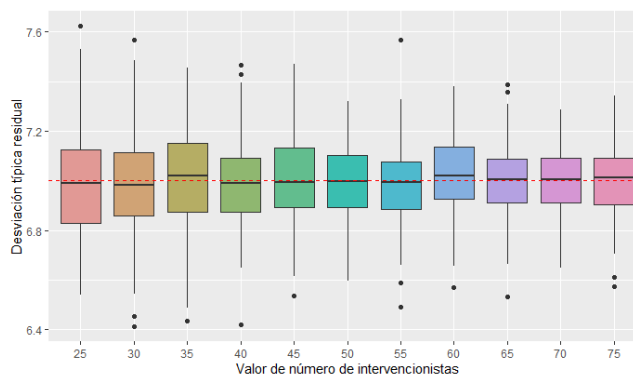


Figura VIII.1.2 – 1. Las estimaciones de la desviación típica residual del modelo se centran en el valor teórico 7, ya que proviene de una  $N(135,7)$ . La precisión de las estimaciones aumenta al aumentar el número de observaciones.

### VIII.1.2 – 2. Evolución de la desviación típica del efecto Intervencionista

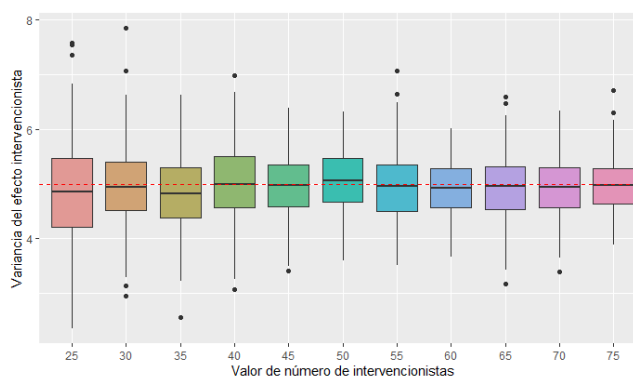


Figura VIII.1.2 – 2. Las estimaciones de la desviación típica de los efectos de intervencionistas se centran en el valor teórico 5, ya que proviene de una  $N(0,5)$ . La precisión aumenta a medida que crece el número de intervencionistas y, por tanto, de observaciones.

Esta simulación no varía la desviación típica del efecto intervencionista ni la varianza residual del modelo, por lo que el valor central de sus estimaciones coincide con los valores teóricos, 7 y 5. Aun así, sus intervalos son cada vez más estrechos, indicando mayor precisión.

La proporción de la desviación que vendrá explicada por el efecto aleatorio, será:

$$\text{Proporción } S_{EfAleatorio} = \frac{S_{EfAleatorio}}{S_{Residual} + S_{EfAleatorio}} = \frac{S_{EfAleatorio}}{S_{Total}}$$

El valor teórico de esta proporción es:

$$\frac{S_{EfAleatorio}}{S_{Residual} + S_{EfAleatorio}} = \frac{5}{7 + 5} = 0.41\hat{6}$$

### VIII.1.1 – 3. Evolución de la proporción de desviación del efecto Intervencionista

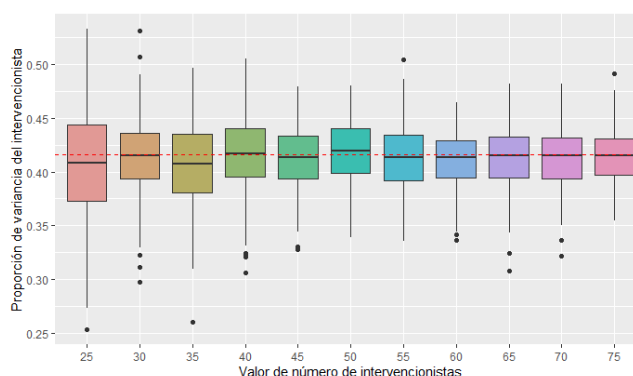


Figura VIII.1.1 – 3. La proporción de variabilidad asociada al efecto aleatorio se centra en el valor teórico (0.416) con un rango menor al aumentar el número de pacientes.

Los valores simulados se encuentran centrados en su valor teórico: 41.7% de la desviación total del modelo vendrá explicada por el efecto aleatorio del intervencionista.

### VIII.1.3 – Otras simulaciones

Con el objetivo de investigar si se mantiene el mismo comportamiento cuando aumenta el número de intervencionistas, he realizado una simulación posterior, ahora con valores:

$$value=20*i, \text{ donde } i \in \{5, 6, \dots, 15\}$$

Los gráficos del Anexo 2.1 muestran cómo se mantienen los mismos patrones, que resumo en la siguiente Tabla.

VIII.1.3 – 1. Tabla resumen de los parámetros según la evolución de la variable número de intervencionistas

	Parámetro de evolución	
	Intervencionistas	
	Valor medio	Precisión
Constante del modelo $\beta_0$	✓	✓
Tratamiento $\beta$	✓	✓
Std. Intercept	✓	✓
Std. Tratamiento	✓	✓
S Residual	✓	✓
S Efecto Intervencionista	✓	✓
Proporción de desviación del efecto intervencionista	✓	✓

Tabla VIII.1.3 – 1. Los ticks en las celdas muestran que los resultados se ajustan a lo esperado, o bien se mejora la precisión de las estimaciones. En el caso de aumentar el número de intervención no hay ninguna anomalía.

## VIII.2 – Resultados según el número de pacientes por intervencionista

A continuación, muestro los resultados del mismo estudio para la variable número de pacientes por intervencionista.

De igual manera que con el número de intervencionistas, al aumentar el valor de esta variable se estará aumentando el número total de individuos observados, lo cual conllevará la mejor estimación de algunos parámetros. A diferencia del caso anterior, la mejora en la precisión será menor.

Este caso también tomará el conjunto de valores  $value=5*i$ , donde  $i \in \{5, 6, \dots, 15\}$ , por lo que habrá un total de 10 intervencionistas en cada simulación, con  $value$  número de pacientes. Se sigue manteniendo la proporción de pacientes con y sin tratamiento, por lo que también aumentará el número de pacientes sin tratamiento.

En esta simulación en ninguna ocasión la función elimina el efecto aleatorio.

### VIII.2.1 - Efecto Fijo

De manera similar a la simulación anterior, en ésta las estimaciones de la constante y del efecto tratamiento se vuelven más precisas al aumentar el  $n^o$  de pacientes por intervencionista. El tamaño  $N$  total de la muestra también aumenta:

$$N = value * 10 * 2, \quad value = 5 * i, \quad i \in \{5, 6, \dots, 15\}$$

VIII.2.1 – 1. Evolución del estimador de la constante

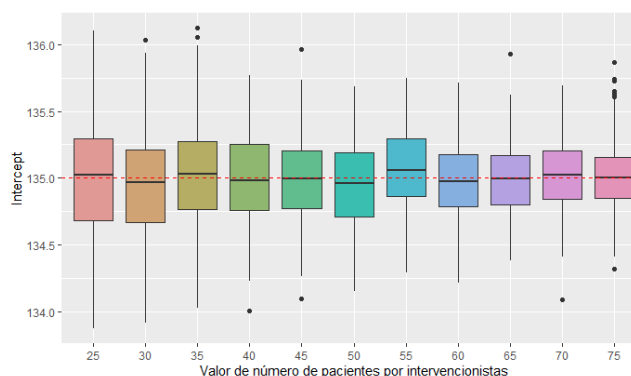


Figura VIII.2.1 – 1. La línea discontinua horizontal roja representa el valor teórico (135). Las estimaciones se centran en el valor teórico y se reducen los intervalos al aumentar el  $n^o$  de pacientes por intervencionista.

### VIII.2.1 – 2. Evolución de la estimación del efecto del tratamiento

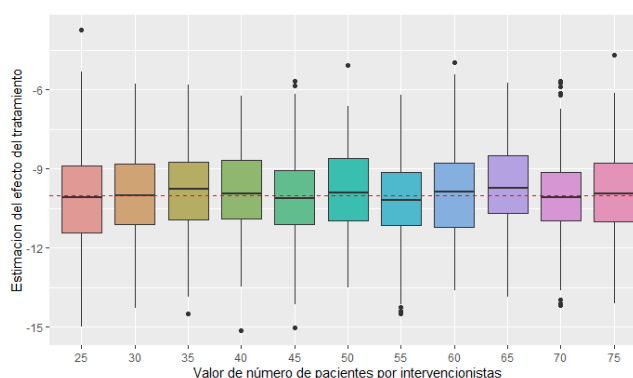


Figura VIII.2.1 – 2. Los resultados se centran alrededor de la línea discontinua horizontal rojo, que representa el valor teórico (-10).

Por otro lado, y a diferencia del caso anterior, ahora no se aprecia mejora en la precisión de la estimación.

### VIII.2.1 – 3. Evolución del error estándar de la estimación del Intercept del modelo

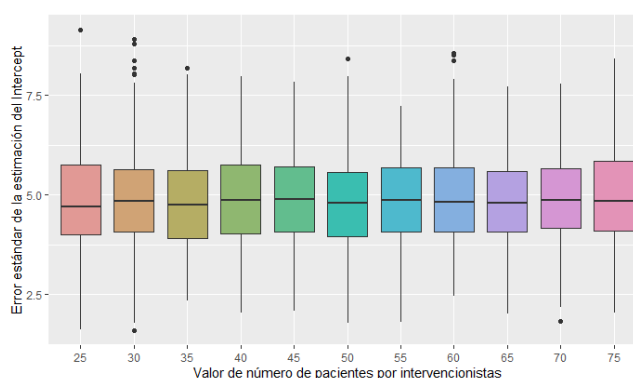


Figura VIII.2.1 – 3. Las distribuciones se encuentran centradas en un mismo valor (5), y mantienen un rango similar.

### VIII.2.1 – 4. Evolución del error estándar de la estimación del efecto del tratamiento

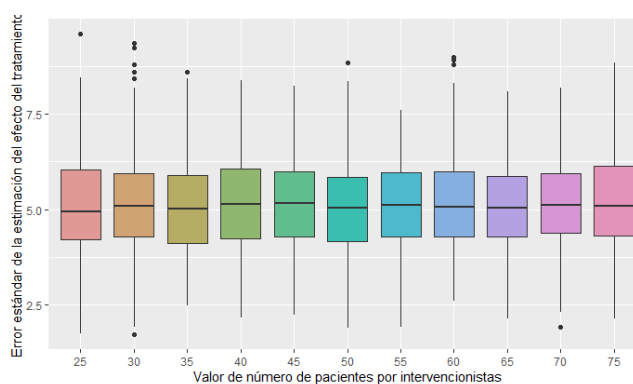


Figura VIII.2.1 – 4. Los resultados se centran en el valor de la constante (5) y con una precisión similar.

Los valores de los errores estándar se encuentran también centrados en 5, pero sus rangos son constantes independientemente del número de pacientes que tiene cada intervencionista.

### VIII.2.2 - Efecto Aleatorio

La distribución de la estimación de la desviación residual del modelo, se centra en el valor teórico 7, y su rango se reduce a medida que aumenta el número de pacientes por intervencionista.

VIII.2.2 – 1. Evolución la desviación típica residual

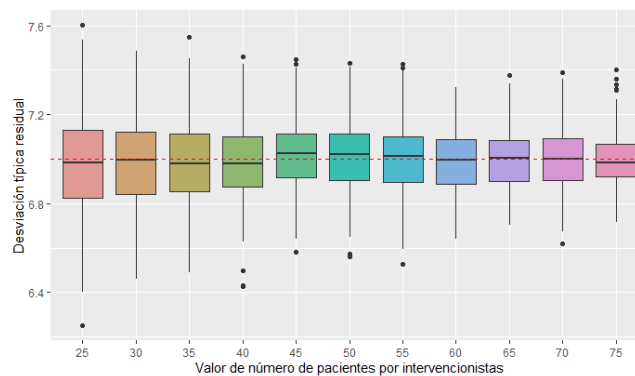


Figura VIII.2.2 – 1. La precisión de las estimaciones mejora al aumentar el número de pacientes por intervencionista, y se centran en el valor teórico (7) marcado con la línea discontinua horizontal roja.

En el caso de la desviación asociada al efecto intervencionista, los valores se encuentran centrados ligeramente bajo el valor teórico 5, y sus rangos no se reducen al aumentar el número de pacientes por intervencionista.

VIII.2.2 – 2. Evolución de la desviación típica del efecto Intervencionista

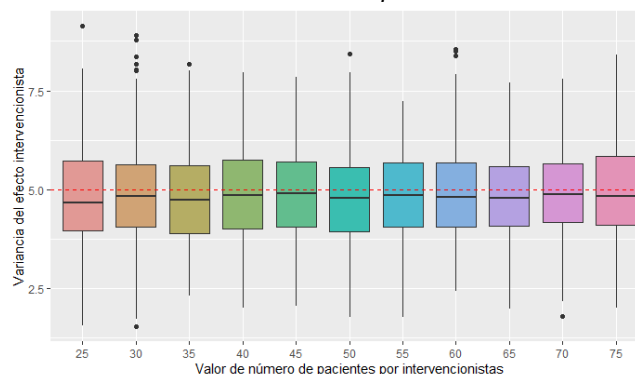


Figura VIII.2.2 – 2. Los valores encuentran centrados ligeramente por debajo del valor teórico (5) marcado con la línea discontinua horizontal y su precisión no aumenta junto con la variable de evolución.

De forma parecida, para la proporción de la desviación asociada al efecto intervencionista, en el gráfico VIII.2.2 – Figura 3, se puede ver como dicha proporción se encuentra ligeramente por debajo del valor teórico 0.416, y su precisión no aumenta.

VIII.2.2 – 3 Proporción de desviación del efecto Intervencionista según nº paciente/intervencionista

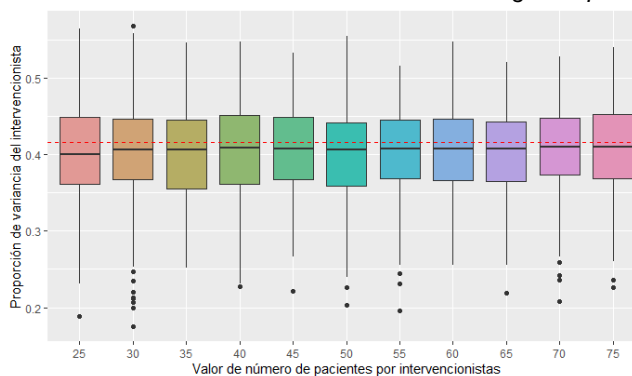


Figura VIII.2.2 – 3. Las estimaciones de la proporción de la variabilidad asociada al efecto aleatorio se centran ligeramente por debajo del valor teórico (0.416) y la precisión de las estimaciones no aumenta al aumentar el número de pacientes asociados a cada intervencionista.

VIII.2.3 – Otras simulaciones

De igual manera que con el caso anterior, se realiza una simulación con valores más grandes del nº de pacientes por intervencionista:

$$value=20*i, \text{ donde } i \in \{5, 6, \dots, 15\}$$

En los gráficos del Anexo 2.2, se puede ver como se mantienen los mismos patrones.

VIII.2.3 – 1. Tabla resumen de los parámetros según la evolución de la variable número de pacientes por intervencionista

	Parámetro de evolución	
	Pacientes por Intervencionista	
	Valor medio	Precisión
Constante del modelo $\beta_0$	✓	✓
Tratamiento $\beta$	✓	✓
Std. Intercept	✓	Constante
Std. Tratamiento	✓	Constante
S Residual	✓	✓
S Efecto Intervencionista	Bajo el valor teórico	Constante
Proporción de desviación del efecto intervencionista	Bajo el valor teórico	Constante

Tabla VIII.2.3 – 1. Los ticks en las celdas muestran que los resultados se ajustan a lo esperado, o bien se mejora la precisión de las estimaciones. En este caso surgen parámetros que no se ajustan al resultado esperado o bien no se aumenta la precisión de sus estimaciones al aumentar el número de pacientes por intervencionista.



### VIII.3 - Resultados según la desviación típica del efecto del intervencionista

Finalmente, se realizará la simulación para la desviación típica del efecto intervencionista. A diferencia de los casos anteriores, el número de individuos en cada iteración será constante:

$$N = 10 * 10 * 2 = 200$$

De los cuales 100 pertenecerán al grupo control, y los 100 restantes recibirán el tratamiento por 10 intervencionistas, es decir, 10 pacientes asociados a cada intervencionista, este efecto vendrá distribuido según una  $N(0, value)$ , donde:

$$value = 2 + (i * 0.5), i \in \{0, 1, \dots, 10\}$$

Se escoge este rango de valores con el objetivo de que los efectos del intervencionista no sean mayores al efecto fijo del tratamiento. Con un efecto aleatorio de -10 (resultante de la distribución Normal(0,  $value$ )), la influencia del intervencionista estaría anulando el efecto del tratamiento.

En el otro extremo, una desviación de 2 causará un conjunto de efectos más cercanos a 0, por ello, en ocasiones la función *lmer* considerará que no es necesario añadir el efecto aleatorio, ya que la desviación simulada es muy pequeña.

#### VIII.3.1 - Omisión del efecto aleatorio

El siguiente diagrama de barras indica la proporción de casos en los que ocurre esta omisión del efecto aleatorio y, por lo tanto, fija los efectos de cada uno de los intervencionistas a 0, que provocará también una desviación de 0 en dicho efecto.

VIII.3.1 – 1. Proporción de casos en los que se omite el efecto aleatorio

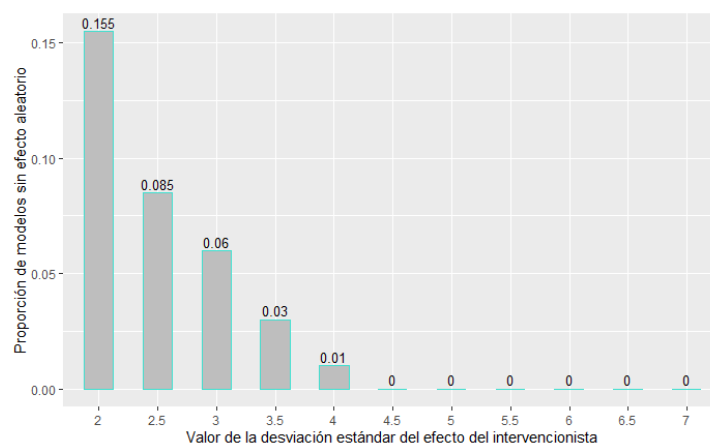


Figura VIII.3.1 – 1. La proporción de casos en los que se omite el efecto aleatorio se reduce a medida que aumenta la desviación típica de los efectos de los intervencionistas.

Como se ve en la gráfica, con una desviación de 2, hay un 0.155 de casos en los que se retira el efecto aleatorio. Este porcentaje se va reduciendo hasta que con una desviación de 4.5 no hay ningún caso en el que ocurra dicha omisión.

**VIII.3.2 - Efecto fijo**

En cuanto al efecto fijo, los valores de estimación de la constante se encuentran centrados en el valor teórico y con rango de valores similar. El error estándar aumenta en valor y rango a medida que aumenta la desviación típica de los efectos de los intervencionistas, por lo que se vuelve una estimación menos precisa. Los valores en los que se centran cada una de las distribuciones de errores estándar coincidirá con la desviación del efecto aleatorio.

VIII.3.2 – 1. Evolución del estimador de la constante del modelo

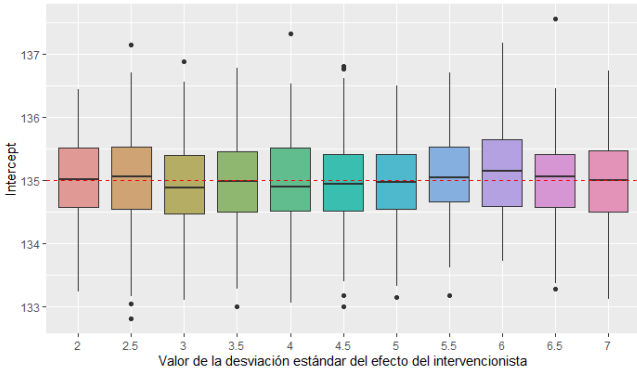


Figura VIII.3.2 – 1. Se centra en el valor teórico (135) y la precisión de su estimación se mantiene constante.

VIII.3.2 – 2. Error estándar de la estimación de la constante del modelo según la variación del efecto del intervencionista

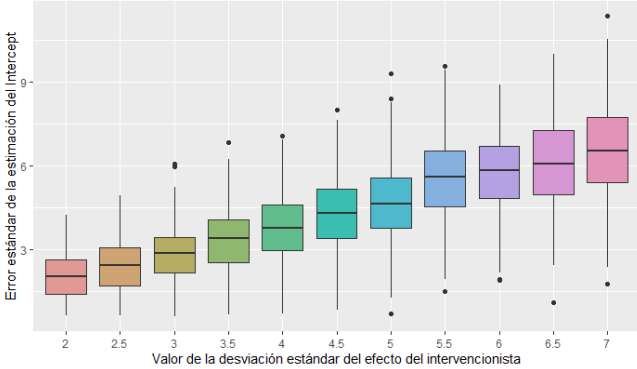


Figura VIII.3.2 – 2. El valor central se aproxima al teórico, que aparece en el eje x. La dispersión también aumenta.

Por otro lado, en el caso del efecto fijo del tratamiento, las estimaciones se centran en el valor teórico -10, pero sus rangos aumentan cuando aumenta la variación del efecto del intervencionista; es decir, aumenta el error estándar del estimador.

VIII.3.2 – 3. Estimación del efecto del tratamiento según aumenta su dispersión.

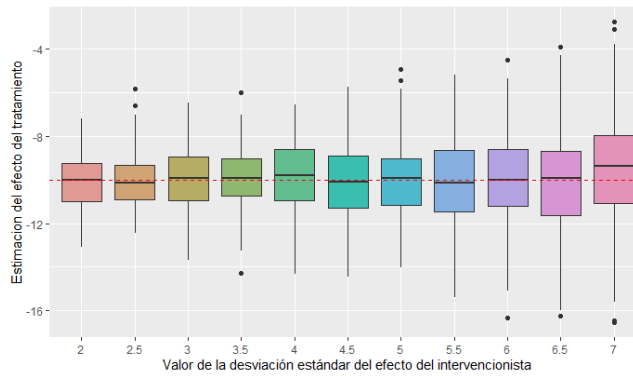


Figura VIII.3.2 – 3. Estimación del efecto del tratamiento. Según aumenta la dispersión del efecto, aumenta el error estándar del estimador.

VIII.3.2 – 4. Evolución del error estándar de la estimación del efecto del tratamiento según aumenta la dispersión del efecto.

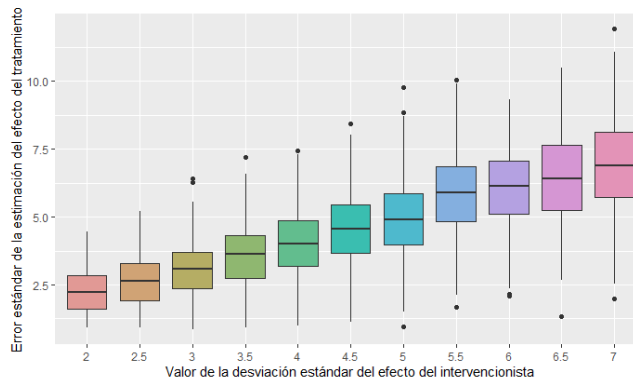


Figura VIII.3.2 – 4. Error estándar de la estimación del efecto tratamiento. El valor central evoluciona junto con su dispersión, mostrada en el eje x. La precisión de las estimaciones también se reduce.

**VIII.3.3 - Efecto aleatorio**

En cuanto a los componentes de la varianza, la estimación de la desviación típica residual está centrada en el valor teórico, con una dispersión similar.

VIII.3.3– 1. *Estimación de la desviación típica residual según aumenta la dispersión del efecto del intervencionista.*

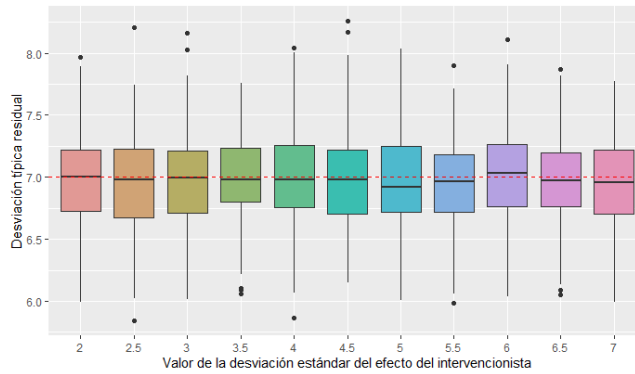


Figura VIII.3.3– 1. *Estimación de la desviación típica residual del modelo. Los valores se encuentran centrados en el valor teórico (7) y su precisión se mantiene muy similar independientemente de la desviación del efecto intervencionista.*

VIII.3.3– 2. *Evolución de la desviación típica del efecto Intervencionista según aumenta su dispersión teórica.*

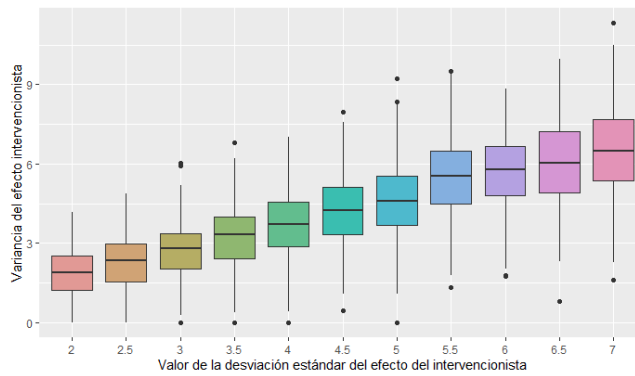


Figura VIII.3.3– 2. *Estimación de la desviación típica del efecto aleatorio. Los valores centrales de cada distribución coinciden con el valor teórico, indicado en el eje x. La precisión de las estimaciones va disminuyendo.*

La desviación típica del efecto aleatorio se incrementa al aumentar su valor teórico. Y también se reduce la precisión de las estimaciones, lo que se observa por el aumento de los rangos.

VIII.3.3– 3. *Evolución de la proporción de desviación del efecto Intervencionista al aumentar su valor teórico.*

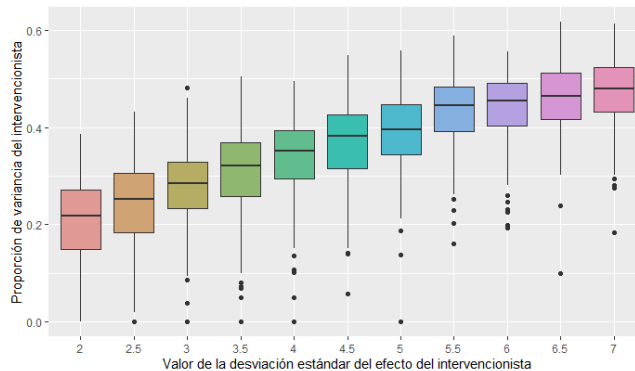


Figura VIII.3.3– 3. Estimación de la proporción de variabilidad asociada al efecto aleatorio. El valor central coincide con el valor teórico. La precisión no varía considerablemente.

La estimación de la proporción de variabilidad asociada al efecto aleatorio del intervencionista aumenta con su valor teórico:

$$\text{Proporción}_{S_{EfAleatorio}}(value) = \frac{value}{7 + value}$$

$$value = 2 + (i * 0.5), i \in \{0, 1, \dots, 10\}$$

si bien la precisión de las estimaciones permanece similar.

VIII.3.3– 1. Tabla resumen de los parámetros según la evolución de la variable desviación típica de los efectos de intervencionistas

	Parámetro de evolución	
	Desviación efectos Intervencionistas	
	Valor medio	Precisión
Constante del modelo $\beta_0$	✓	Constante
Tratamiento $\beta$	✓	Se reduce
Std. Intercept	✓	Se reduce
Std. Tratamiento	✓	Se reduce
S Residual	✓	Constante
S Efecto Intervencionista	✓	Se reduce
Proporción de desviación del efecto intervencionista	✓	Constante

Tabla VIII.3.3– 1. Los ticks en las celdas muestran que los resultados se ajustan a lo esperado, o bien se mejora la precisión de las estimaciones. En el caso de aumentar la desviación típica del efecto intervencionista surgen cambios en la precisión: o bien se mantiene constante o bien se reduce.

## CONCLUSIÓN

### Interpretación

En el caso de las intervenciones no farmacológicas, es apropiado referirse al efecto que proviene de la variabilidad entre intervencionistas como un efecto aleatorio, y añadiendo el efecto del tratamiento que se estará implantando en el paciente como un efecto fijo, quedaría resultante un modelo de efectos mixtos, que contendrá un efecto aleatorio y un fijo.

Si bien debido a la naturaleza de este efecto aleatorio la estimación de los coeficientes de cada nivel de la variable no es útil, ya que se tratará de efectos asociados a individuos que no se repetirán en una iteración futura, existen maneras de mostrar el efecto que tiene la variable sobre el comportamiento en el resultado. Dicho efecto puede ser representado por la variabilidad que otorga al modelo, mediante una proporción, o de manera visual.

Tanto la estimación como la variabilidad del efecto mencionado dependerá de diferentes factores. A través del estudio, se han detectado 3 parámetros de un modelo que influyan al efecto aleatorio: El número de intervencionistas, el número de pacientes por intervencionista y la variabilidad entre los efectos de cada intervencionista. Aumentar o reducir la escala de estos 3 parámetros provocará diferentes modificaciones en el comportamiento del modelo, y por lo tanto en la variabilidad del efecto aleatorio. Dichos resultados se encuentran resumidos en la siguiente tabla.

*Tabla resumen de los parámetros según la evolución según las variables de evolución estudiadas*

	Parámetro de evolución					
	Intervencionistas		Pacientes por Intervencionista		Desviación efectos Intervencionistas	
	Valor medio	Precisión	Valor medio	Precisión	Valor medio	Precisión
Constante del modelo	✓	✓	✓	✓	✓	Constante
Tratamiento	✓	✓	✓	✓	✓	Se reduce
Std. Intercept	✓	✓	✓	Constante	✓	Se reduce
Std. Tratamiento	✓	✓	✓	Constante	✓	Se reduce
S Residual	✓	✓	✓	✓	✓	Constante
S Efecto Intervencionista	✓	✓	Bajo el valor teórico	Constante	✓	Se reduce
Proporción de desviación del efecto intervencionista	✓	✓	Bajo el valor teórico	Constante	✓	Constante

*Los ticks en las celdas muestran que los resultados se ajustan a lo esperado, o bien se mejora la precisión de las estimaciones.*

En el caso del número de intervencionistas, los resultados se encuentran siempre centrados en el valor teórico, además de que, al aumentar la escala, las estimaciones son cada vez más precisas y cercanas a estos valores teóricos.

En cuanto al número de pacientes por intervencionista, al aumentar el parámetro, la precisión de la estimación de los efectos fijos mejora, pero se mantiene constante para la de sus errores estándar. La estimación de la desviación residual mejorará, y en el caso de la desviación del efecto aleatorio, ésta se encontrará ligeramente por debajo del valor teórico, además que su precisión se mantendrá en un rango consistente de valores.

En el caso de aumentar la desviación de la distribución de la que provienen los efectos de cada intervencionista, la precisión de parámetros como los errores estándar de los parámetros fijos, o de la desviación del efecto aleatorio se reduce, aumentando así el rango de valores que aparecen. En cambio, en el caso de la estimación del intercept del modelo, o de la desviación residual, la precisión de sus estimaciones se mantiene constante. Las distribuciones de las estimaciones de todos los parámetros se encuentran centradas en sus valores teóricos.

### **Limitaciones**

Pese a que las simulaciones generadas cuentan con parámetros que simulan un caso realista, siguen sin ser datos de un experimento real. Aun así, el procedimiento del análisis se podría realizar de la misma manera en cualquier modelo que cuente con un efecto fijo (tratamiento) y un efecto aleatorio (intervencionista). En caso de añadir variables explicativas adicionales que influyan en el resultado, este procedimiento no será el apropiado, por otro lado, puede ser de utilidad si se observa una alta variabilidad proveniente del intervencionista en un modelo con únicamente las dos variables presentadas.

Debido a que como se ha comentado, se preserva el anonimato del intervencionista, no ha sido posible encontrar una base de datos que se ajusten a las características que se buscaban, por ello se ha decidido realizar el estudio mediante datos generados.

En la interpretación de la proporción de desviación asociada al efecto intervencionista, no hay un lindar de valores que indique cuándo la variabilidad asociada al intervencionista se considera significativamente alta.





## BIBLIOGRAFÍA

- Stimulus. *Las Terapias No Farmacológicas* [En línea]. [consulta: Marzo de 2022] Disponible en:

<https://stimuluspro.com/blog/las-terapias-no-farmacologicas/>

- Explorable. Reproducibilidad [En línea]. [consulta: Marzo de 2022] Disponible en:

<https://explorable.com/es/reproducibilidad>

- Evidencias en Pediatría. *Ensayo clínico (IV). Resultados. Variables y medidas de efecto e impacto* [En línea]. [consulta: Marzo de 2022] Disponible en:

<https://evidenciasenpediatria.es/articulo/6736/ensayo-clinico-iv-resultados-variables-y-medidas-de-efecto-e-impacto>

- National Library of Medicine, MedlinePlus. *Presión arterial alta* [En línea]. [consulta: Abril de 2022] Disponible en:

<https://medlineplus.gov/spanish/highbloodpressure.html>

- OMRON. *Gráficos de presión arterial* [En línea]. [consulta: Abril de 2022] Disponible en:

<https://www.omron-healthcare.es/es/salud-y-estilo-de-vida/salud-cardiovascular/control-de-la-presion-arterial/graficos-de-presion-arterial.html>

- EQUATOR Network [En línea]. [consulta: Abril de 2022] Disponible en:

<https://www.equator-network.org/>

- CONSORT. *Transparent reporting of trials* [En línea]. [consulta: Abril de 2022] Disponible en:

<http://www.consort-statement.org/>

- CONSORT. *Non-Pharmacologic Treatment Interventions* [En línea]. [consulta: Abril de 2022] Disponible en:

<http://www.consort-statement.org/extensions?ContentWidgetId=558>

- UCLA Statistical Methods and Data Analytics. *Introducción a modelos mixtos lineales* [En línea]. [consulta: Abril de 2022] Disponible en:

<https://stats.oarc.ucla.edu/other/mult-pkg/introduction-to-linear-mixed-models/>

-

- Francesc Carmona, MODELOS LINEALES [En línea]. Universitat de Barcelona. 27 de noviembre de 2001. Disponible en:

<https://docplayer.es/36072897-Modelos-lineales-francesc-carmona.html>

- Github. Coding Club. *Introduccion to linear mixed models* [En línea]. [consulta: Abril de 2022] Disponible en:

<https://ourcodingclub.github.io/tutorials/mixed-models/>

- Github. Modelos mixtos con R [En línea]. [consulta: Mayo de 2022] Disponible en:

[https://fhernanb.github.io/libro\\_modelos\\_mixtos/pac-lme4.html](https://fhernanb.github.io/libro_modelos_mixtos/pac-lme4.html)

## **ANEXOS**

### **Anexo 1 - Clasificación de las variables según su naturaleza**

Inicialmente las variables pueden dividirse según cualitativas y cuantitativas. Las variables cualitativas son aquellas en las que cada valor corresponde a una característica o cualidad del sujeto al que pertenece.

Las variables cualitativas pueden clasificarse en nominales y ordinales. Las nominales son aquellas en las que la escala de medida se compone de palabras que permiten únicamente clasificar a los sujetos, sin posibilidad de ordenar o jerarquizar sus valores. Por otro lado, las variables cualitativas ordinales sí permiten que sus valores estén jerarquizados.

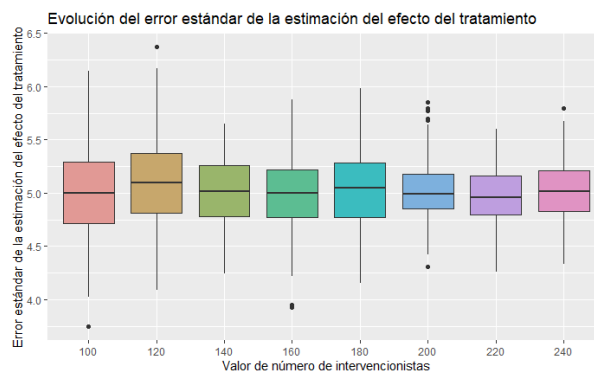
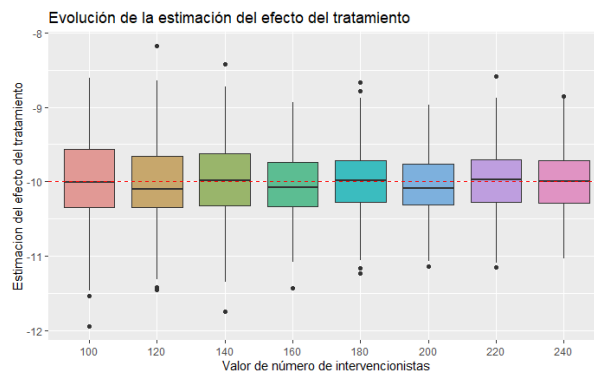
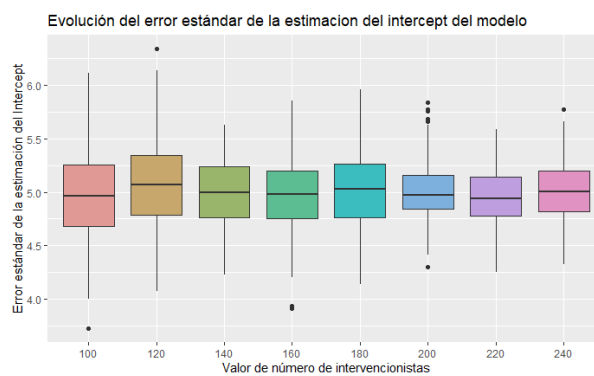
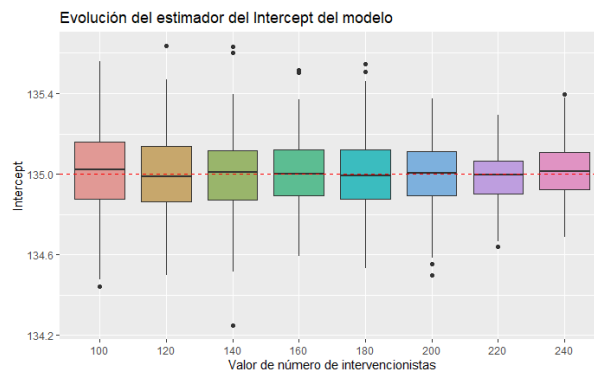
Por último, las variables cuantitativas son aquellas en las que los datos se expresan realmente con números por lo que, además de operaciones de orden, permiten operaciones algebraicas. Este tipo de variables se dividen, a su vez, en discretas y continuas.

Una variable cuantitativa discreta es aquella que solo puede tomar cierto número de valores finitos, representados por números enteros, mientras que una variable continua puede tomar cualquier valor intermedio decimal entre los números enteros.

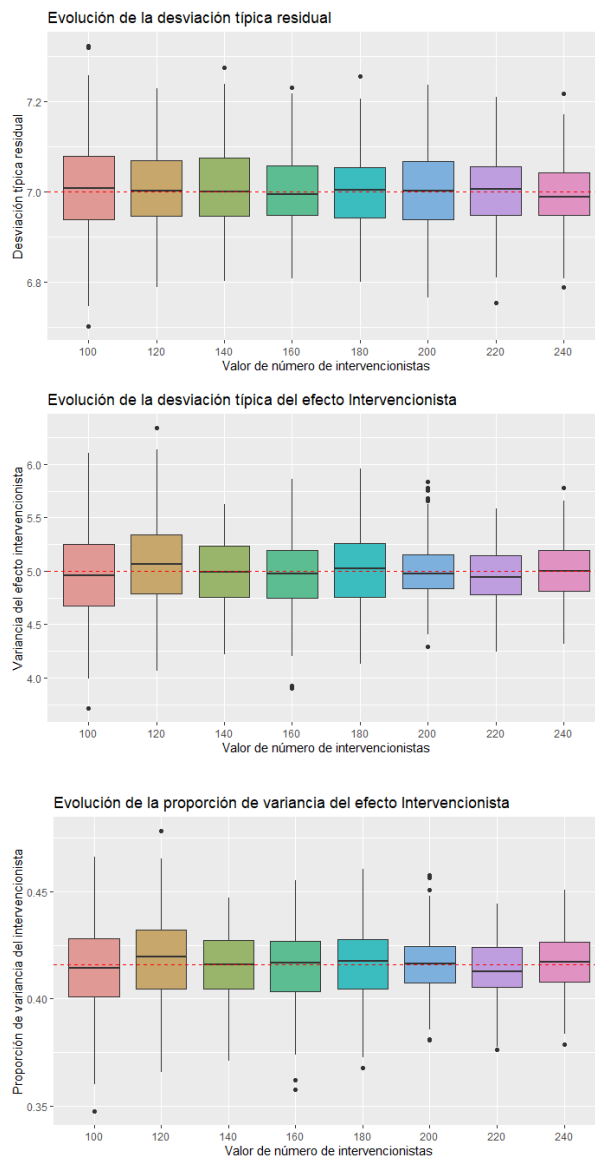
## Anexo 2 - Gráficos de simulaciones de mayor magnitud.

### Anexo 2.1 - Caso número de interaccionistas

#### Efecto fijo

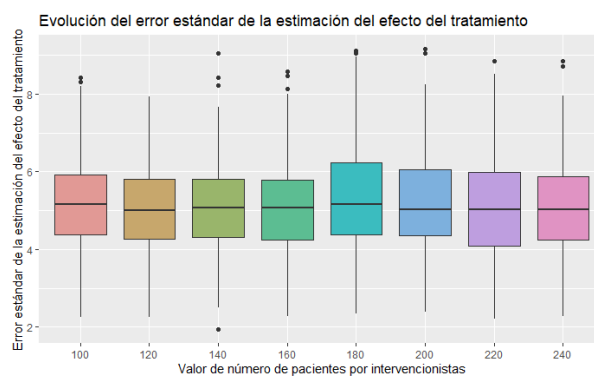
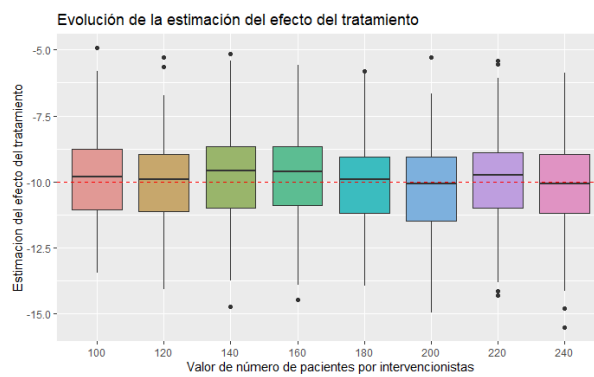
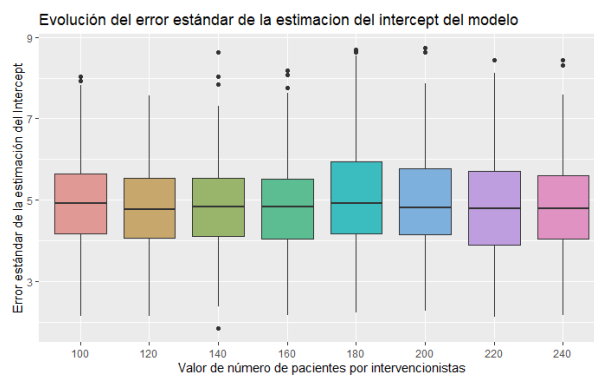
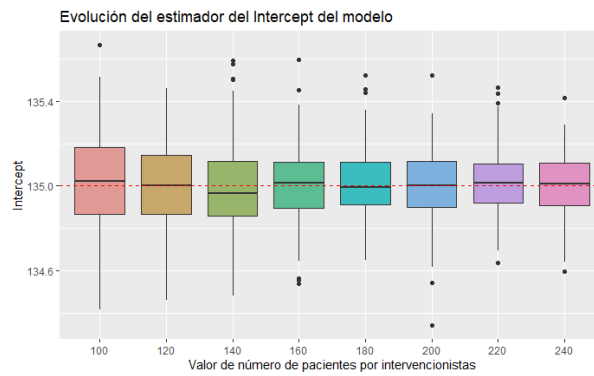


## Efecto aleatorio

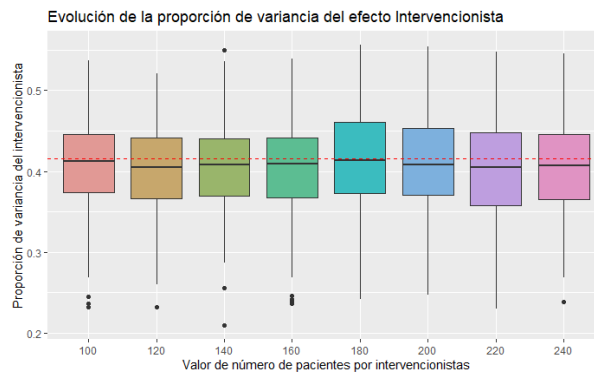
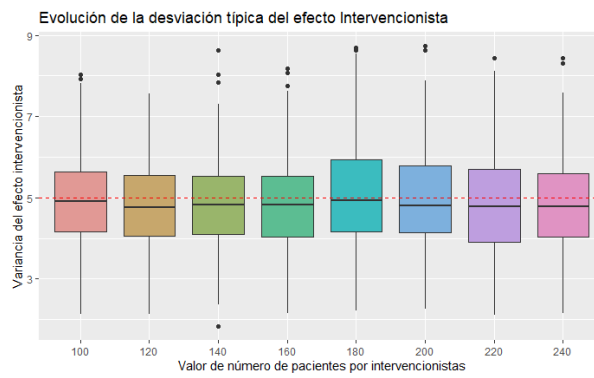
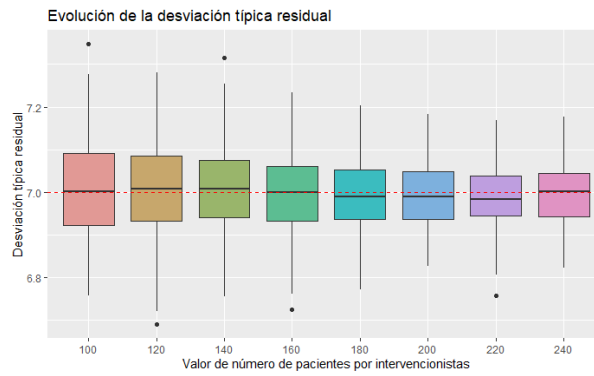


## Anexo 2.2 - Caso número de pacientes por interveccionista

### Efecto fijo



## Efecto aleatorio



### Anexo 3 – Código de R. Función de simulaciones.

En este caso se fijan los parámetros a los utilizados en el punto VIII.1 del trabajo. Con la variable número de intervencionistas como variable de evolución, que tomará los valores

$$value=5*i, i \in \{5, 6, \dots, 15\}$$

Al cambiar las variables *evo*, que indicará cuál es la variable de evolución, y el vector *values*, que contiene los valores que tomará dicha variable, se pueden simular distintos casos.

En el caso de querer realizar una simulación con diferentes parámetros en los datos de origen, éstos se pueden modificar en el bloque “Parámetros fijos de generación de datos”

```
#library(lme4)
lista <- list()
lista_dats <- list()
lista_ef <- list()
no_random <- c()
j <- 0

#----- Parámetros que varían en cada bucle -----
# Indica los valores que tendrá la variable que cambia en cada iteración.
values <- 5*c(5:15)
# indica cuál de las 3 variables a cambiar es la que tomará los valores del vector "values"
# 1: Número de intervencionistas
# 2: Número de pacientes por intervencionista
# 3: Desviación típica del efecto del intervencionista
evo <- 1

#----- Parámetros fijos de generación de datos -----
# Iteraciones por bucle
m <- 200
#número de intervencionistas
n.int <- 10
#número de pacientes por intervencionista
n.x.int <- 10
# media de los datos
mean.dat <- 135
# desviación típica de los datos
sd.dat <- 7
# media del efecto del intervencionista (fijo a 0)
mean.ef <- 0
# desviación típica del efecto del intervencionista
sd.ef <- 5
#valor fijo del tratamiento
ef.trat <- -10

#-----          Bucle          -----
for(w in values){
  j <- j+1

  # determinar qué variable cambia en cada iteración
  if(evo==1){
    n.int <- w
  }else if(evo==2){
```



```

n.x.int <- w
}else{
sd.ef <- w
}
# tamaño total de la muestra
N <- n.int*n.x.int*2

tab <- data.frame()
dats<- data.frame()
ef_ints<- data.frame()
for(i in 1:m){
tract<- rep(c(0,1),each=N/2)
int <- rep(c("0", 1:n.int), c(N/2, rep(n.x.int, n.int)))

#valores de distribución normal entre 120 y 150
dat <- rnorm(N, mean=mean.dat, sd=sd.dat)
dats<- rbind(dats, dat)

# efecto de intervencionista
ef_int <- rnorm(n.int, mean=mean.ef, sd=sd.ef)
ef_ints<- rbind(ef_ints, ef_int)

# añadir efecto del tratamiento e intervencionista
dat[((N/2)+1):N] <- dat[((N/2)+1):N]+ef.trat + rep(ef_int, each=n.x.int)

tabla <- data.frame(bp=round(dat, 2) , tract=as.factor(tract), int=as.factor(int))
mod <- lmer(bp~ tract + (1|int), data=tabla)
summ <- summary(mod)

# extracción de los resultados del modelo
tab2 <- data.frame(i=i,
int.var=as.numeric(summ$varcor$int),res.var=summ$sigma^2,intercept=summ$coefficients[
1],intercept.std=
summ$coefficients[3],eff.tract=summ$coefficients[2],eff.tract.std=summ$coefficients[4],prop.
var=sqrt(as.numeric(summ$varcor$int))/(summ$sigma+
sqrt(as.numeric(summ$varcor$int))))

tab <- rbind(tab, tab2)
}
no_random <- c(no_random ,sum(tab$int.var==0))
lista[[j]] <- tab
lista_dats[[j]] <- dats
lista_ef[[j]] <- ef_ints
}
# no_random contiene el número de casos en los que se igualan los efectos aleatorios a 0,
para cada iteración del ciclo
# lista contiene los dataframes de información extraída en cada iteración del ciclo
# lista_dats contiene los dataframes de datos utilizados en cada iteración del ciclo
# lista_ef contiene los dataframes de efectos aleatorios (de cada intervencionista) en cada
iteración del ciclo

```

#### Anexo 4 – Código de R. Visualización gráfica de los resultados.

Primero se ejecuta el siguiente bloque de código para ajustar la base de datos al formato apropiado para el paquete *ggplot2*

```
# Bucle para cambiar el formato a uno que pueda leer el ggplot2

#library(ggplot2)
dat <- data.frame()
for(ciclo in 1:j){
  dat0 <- data.frame("prop.var"=lista[[ciclo]][["prop.var"],
                    "int.var"=lista[[ciclo]][["int.var"],
                    "res.var"=lista[[ciclo]][["res.var"],
                    "intercept"=lista[[ciclo]][["intercept"],
                    "intercept.std"=lista[[ciclo]][["intercept.std"],
                    "eff.tract"=lista[[ciclo]][["eff.tract"],
                    "eff.tract.std"=lista[[ciclo]][["eff.tract.std"]
                    ,"value"= as.factor(values[ciclo]))
  dat <- rbind(dat, dat0)
}

if(evo==1){
  evo.name <- "número de intervencionistas"
}else if(evo==2){
  evo.name <- "número de pacientes por intervencionistas"
}else{
  evo.name <- "la desviación estándar del efecto del intervencionista"
}
```

Gráficos para las simulaciones de número de intervencionistas y número de pacientes por intervencionista

```
# ----- GRÁFICOS SIMULACIÓN NÚMERO DE INTERVENCIONISTAS / NÚMERO DE
# PACIENTES POR INTERVENCIONISTA -----
# incluye rectas horizontales con el valor teórico de cada parámetro
#----- EFECTO ALEATORIO -----
#prop var
p1 <- ggplot(data = dat, aes(x=value, y=prop.var, fill=value)) +
  geom_boxplot() + labs(title="Evolución de la proporción de variancia del efecto
Intervencionista",x=paste(c("Valor de"), evo.name), y = "Proporción de variancia del
intervencionista") +
  geom_hline(yintercept=as.numeric(sd.ef/(sd.ef+sd.dat)),linetype="dashed", color = "red")
p1+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#int var
p2 <- ggplot(data = dat, aes(x=value, y=int.var, fill=value)) +
  geom_boxplot() + labs(title="Evolución de la variancia del efecto
Intervencionista",x=paste(c("Valor de"), evo.name), y = "Variancia del efecto
intervencionista") +
  geom_hline(yintercept=sd.ef^2,linetype="dashed", color = "red")
```

```

p2+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

p2.1 <- ggplot(data = dat, aes(x=value, y=sqrt(int.var), fill=value)) +
  geom_boxplot() + labs(title="Evolución de la desviación típica del efecto
Intervencionista",x=paste(c("Valor de"), evo.name), y = "Variancia del efecto
intervencionista") +
  geom_hline(yintercept=sd.ef,linetype="dashed", color = "red")
p2.1+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

p2.2 <- ggplot(data = dat, aes(x=value, y=(sqrt(int.var)-5), fill=value)) +
  geom_boxplot() + labs(title="",x=paste(c("Valor de"), evo.name), y = "Variancia del efecto
intervencionista") +
  geom_hline(yintercept=0,linetype="dashed", color = "red")
p2.2+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#res var
p3 <- ggplot(data = dat, aes(x=value, y=sqrt(res.var), fill=value)) +
  geom_boxplot() + labs(title="Evolución de la desviación típica residual",x=paste(c("Valor
de"), evo.name), y = "Desviación típica residual") +
  geom_hline(yintercept=sd.dat,linetype="dashed", color = "red")
p3+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#----- EFECTO FIJO -----
#intercept
p4 <- ggplot(data = dat, aes(x=value, y=intercept, fill=value)) +
  geom_boxplot() + labs(title="Evolución del estimador del Intercept del
modelo",x=paste(c("Valor de"), evo.name), y = "Intercept") +
  geom_hline(yintercept=mean.dat,linetype="dashed", color = "red")
p4+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#intercept std
p5 <- ggplot(data = dat, aes(x=value, y=intercept.std, fill=value)) +
  geom_boxplot() + labs(title="Evolución del error estándar de la estimacion del intercept del
modelo",x=paste(c("Valor de"), evo.name), y = "Error estándar de la estimación del
Intercept")
p5+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#eff tract
p6 <- ggplot(data = dat, aes(x=value, y=eff.tract, fill=value)) +
  geom_boxplot() + labs(title="Evolución de la estimación del efecto del
tratamiento",x=paste(c("Valor de"), evo.name), y = "Estimacion del efecto del tratamiento") +
  geom_hline(yintercept=ef.trat,linetype="dashed", color = "red")
p6+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#std error eff tract
p7 <- ggplot(data = dat, aes(x=value, y=eff.tract.std, fill=value)) +
  geom_boxplot() + labs(title="Evolución del error estándar de la estimación del efecto del
tratamiento",x=paste(c("Valor de"), evo.name), y = "Error estándar de la estimación del
efecto del tratamiento")
p7+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

```

## Gráficos para el caso de la desviación típica del efecto intervencionista

```
# ----- GRÁFICOS SIMULACIÓN DESVIACIÓN TÍPICA EFECTO
INTERVENCIONISTA -----
# incluye rectas horizontales con el valor teórico de cada parámetro, además del diagrama
de barras con los casos en los que se omite el efecto aleatorio
#----- EFECTO ALEATORIO -----
#prop var
p1 <- ggplot(data = dat, aes(x=value, y=prop.var, fill=value)) +
  geom_boxplot() + labs(title="Evolución de la proporción de variancia del efecto
Intervencionista",x=paste(c("Valor de"), evo.name), y = "Proporción de variancia del
intervencionista")
#+ stat_summary(fun.y = mean, geom="point",colour="black", size=2, shape=3)
p1+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#int var
p2 <- ggplot(data = dat, aes(x=value, y=int.var, fill=value)) +
  geom_boxplot() + labs(title="Evolución de la variancia del efecto
Intervencionista",x=paste(c("Valor de"), evo.name), y = "Variancia del efecto
intervencionista")
#+ geom_point(aes(x=factor(values^2), y=values),color = 'red')

p2.1 <- ggplot(data = dat, aes(x=value, y=sqrt(int.var), fill=value)) +
  geom_boxplot() + labs(title="Evolución de la desviación típica del efecto
Intervencionista",x=paste(c("Valor de"), evo.name), y = "Variancia del efecto
intervencionista")
#+ stat_summary(fun.y = mean, geom="point",colour="black", size=2, shape=3)
#+ geom_point(aes(x=factor(values), y=values),color = 'red')

p2.1+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#res var
p3 <- ggplot(data = dat, aes(x=value, y=sqrt(res.var), fill=value)) +
  geom_boxplot() + labs(title="Evolución de la desviación típica residual",x=paste(c("Valor
de"), evo.name), y = "Desviación típica residual")+
  geom_hline(yintercept=sd.dat,linetype="dashed", color = "red")
p3+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#----- EFECTO FIJO -----
#intercept
p4 <- ggplot(data = dat, aes(x=value, y=intercept, fill=value)) +
  geom_boxplot() + labs(title="Evolución del estimador del Intercept del
modelo",x=paste(c("Valor de"), evo.name), y = "Intercept") +
  geom_hline(yintercept=mean.dat,linetype="dashed", color = "red")
p4+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#intercept std
p5 <- ggplot(data = dat, aes(x=value, y=intercept.std, fill=value)) +
  geom_boxplot() + labs(title="Evolución del error estándar de la estimacion del intercept del
modelo",x=paste(c("Valor de"), evo.name), y = "Error estándar de la estimación del
Intercept")
```

```

p5+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#eff tract
p6 <- ggplot(data = dat, aes(x=value, y=eff.tract, fill=value)) +
  geom_boxplot() + labs(title="Evolución de la estimación del efecto del
tratamiento",x=paste(c("Valor de"), evo.name), y = "Estimacion del efecto del tratamiento") +
  geom_hline(yintercept=ef.trat,linetype="dashed", color = "red")
p6+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#std error eff tract
p7 <- ggplot(data = dat, aes(x=value, y=eff.tract.std, fill=value)) +
  geom_boxplot() + labs(title="Evolución del error estándar de la estimación del efecto del
tratamiento",x=paste(c("Valor de"), evo.name), y = "Error estándar de la estimación del
efecto del tratamiento")
p7+scale_fill_hue(l=70, c=50)+theme(legend.position="none")

#----- Diagrama de barras adicional -----
# casos de omisión de efecto aleatorio
ggplot(data=data.frame(val=as.factor(values), nran=no_random), aes(x=val,
y=nran/m))+geom_bar(stat="identity", width=0.5,color="turquoise",
fill="grey")+geom_text(aes(label=nran/m), vjust=-0.3, size=3.5) + labs(x=paste(c("Valor de"),
evo.name), y="Proporción de modelos sin efecto aleatorio")

```