



# Taller: Realización de Modelos Lineales Mixtos como alternativa para el análisis de datos

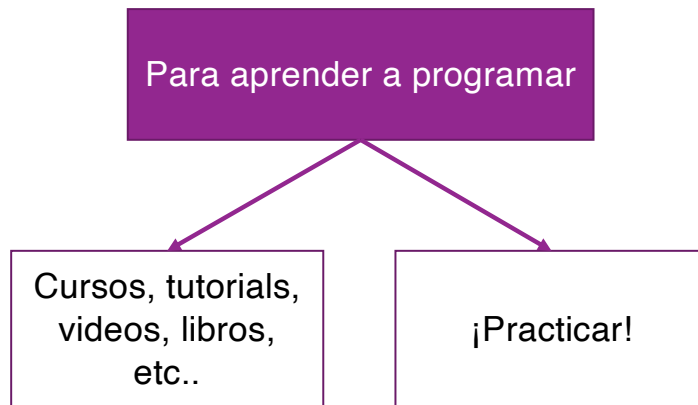
Gabriela E. López-Tolsa

Programa de Doctorado en Psicología  
Escuela Internacional de Doctorado UNED

21, 24 y 28 de noviembre, y 1 de diciembre de 2022

# Estructura del taller

Disclaimer: no es un curso de programación, ni de R.



Para aprender R/RStudio



# Estructura del taller

- Objetivo:

Analizar datos mediante la ejecución e interpretación de Modelos Lineales Mixtos utilizando R.

- Temario:

1. Instalación y breve introducción a R y RStudio
2. Introducción teórica al uso de Modelos Lineales de Efectos Mixtos
3. Preparación de los datos para su análisis
4. Modelos Lineales Mixtos frecuentistas
5. Modelos Lineales Mixtos Bayesianos
6. Reporte de resultados conforme a los criterios de ciencia abierta

# Certificados

- Para obtener el certificado es necesario cumplir con los siguientes requisitos:
  - Entregar las evidencias (scripts y bases de datos) de los ejercicios guiados hechos durante el taller.
  - Entregar evidencias de un análisis de datos no guiado.
    - Un análisis con MLM de sus propios datos (si no cuentan con datos se los podría facilitar).
    - Los requisitos se enviarán por correo electrónico y estarán disponibles en GitHub.
- El certificado será expedido por la EIDUNED, pero no cuenta para créditos.

# ¿Qué es programar?

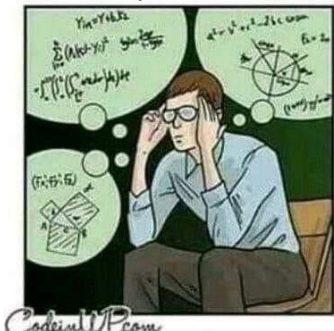
- Programar es redactar un conjunto de **instrucciones** para que las lleve a cabo un ordenador.
- Las instrucciones deben estar redactadas en un **lenguaje que el ordenador pueda interpretar**.
- Y ese lenguaje (igual que cualquier idioma) debe llevar una **cierta estructura**, en sus “oraciones” (líneas), sangrías, espacios, caracteres, etc.
- Si aprendes a programar en un lenguaje, será más fácil aprender otros.
- También a veces requieren o se pueden usar **programas especiales** para **redactar/leer el código**, a las que se les llama Entorno de Desarrollo Integrado (IDE).
- Una de las habilidades más importante a la hora de programar es **leer**. En muchas ocasiones podemos cometer errores “simples” y fáciles de arreglar, que pueden arreglarse tomándolo con calma y leyendo cuál es el problema.

# ¿Qué es programar?

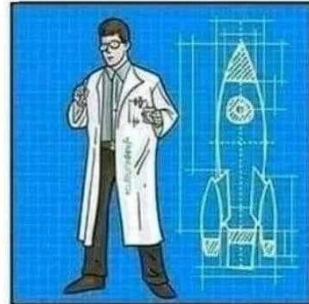
Lo que la gente  
cree que es



Lo que yo creo  
que es



Lo que los padres  
creen que es



Lo que realmente  
es



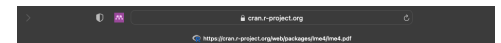
En programación, muchas veces es más importante saber buscar información que saber hacer algo.

¿Dónde buscar?

Google

stackoverflow

CRAN  
R  
The Comprehensive R  
Archive Network



Package 'lme4'

November 1, 2022

Version 1.1-31

Title Linear Mixed-Effects Models using 'Eigen' and S4

Description Fit linear and generalized linear mixed-effects models.

The models and their components are represented using S4 classes and methods. The core computational algorithms are implemented using the 'Eigen' C++ library for numerical linear algebra and 'RcppEigen' " glue".

Depends R (>= 3.5.0), Matrix (>= 1.2-1), methods, stats

# ¿Qué son R y Rstudio?



Es un lenguaje de programación diseñado para hacer análisis estadísticos y gráficas.

Es software libre, lo que quiere decir que es gratuito, y que su código base puede modificarse.

Además de su código base se le pueden añadir paquetes o librerías que cumplan funciones específicas. Estas también se comparten como software libre.



Es un Entorno de Desarrollo integrado (IDE, por sus siglas en inglés), que permiten programar de forma más amigable.

Dentro de RStudio vamos a poder compilar paquetes, correr el código, ver gráficos, etc.

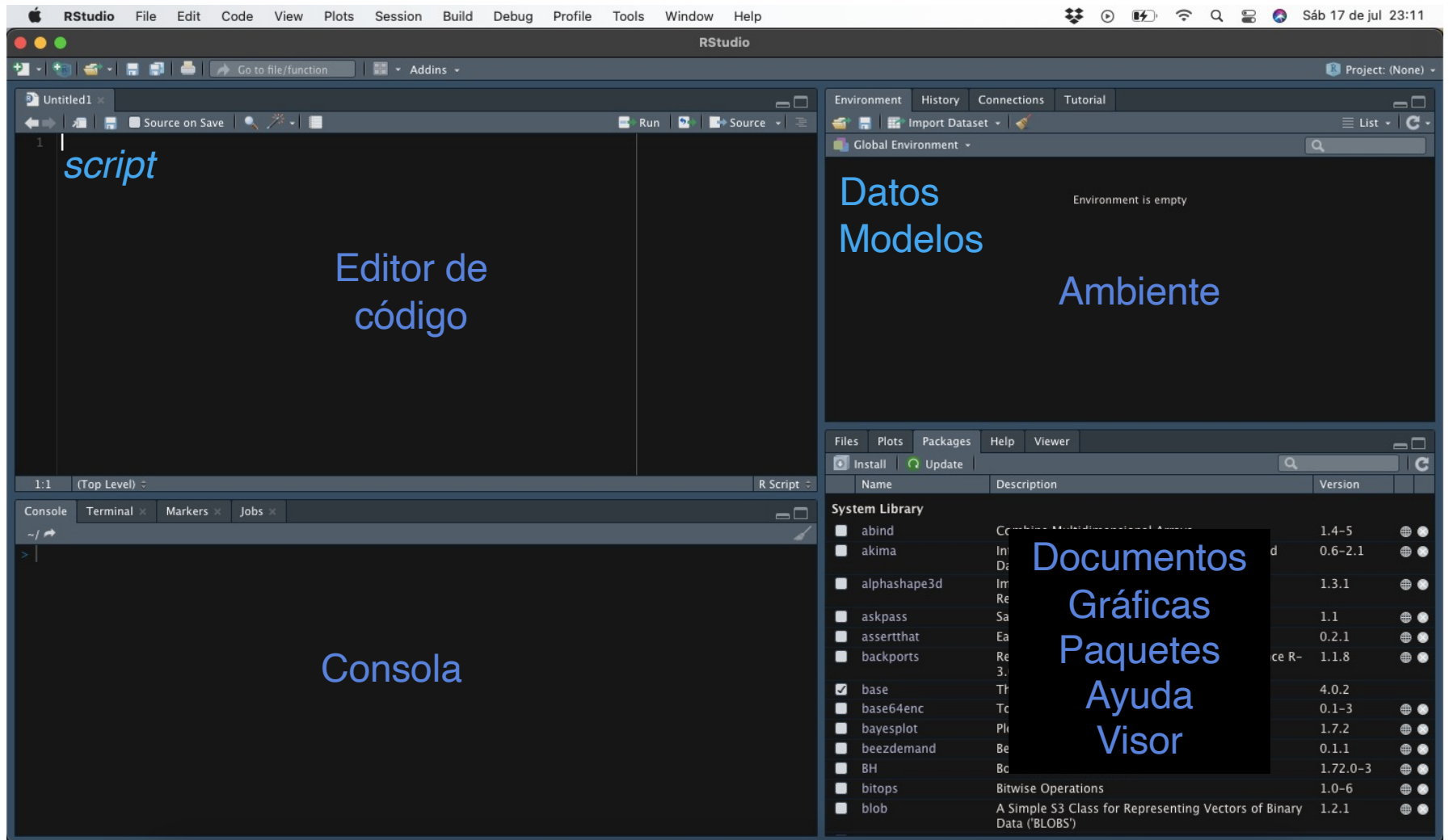
Tiene una interfaz gráfica, con elementos como “botones” y “menús” que nos permiten trabajar de forma más cómoda.

# Instalación de R y RStudio

- Si no han instalado R y RStudio, este es el momento...



# Familiarizarnos con RStudio



**RStudio** File Edit Code View Plots Session Build Debug Profile Tools Window Help

Sáb 17 de jul 23:11

**RStudio**

Project: (None)

Environment History Connections Tutorial

Global Environment

Environment is empty

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
<b>System Library</b>		
<input type="checkbox"/> abind	Combine Multidimensional Arrays	1.4-5
<input type="checkbox"/> akima	Interpolation of Irregularly and Regularly Spaced Data	0.6-2.1
<input type="checkbox"/> alphashape3d	Implementation of the 3D Alpha-Shape for the Reconstruction of 3D Sets from a Point Cloud	1.3.1
<input type="checkbox"/> askpass	Safe Password Entry for R, Git, and SSH	1.1
<input type="checkbox"/> assertthat	Easy Pre and Post Assertions	0.2.1
<input type="checkbox"/> backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.8
<input checked="" type="checkbox"/> base	The R Base Package	4.0.2
<input type="checkbox"/> base64enc	Tools for base64 encoding	0.1-3
<input type="checkbox"/> bayesplot	Plotting for Bayesian Models	1.7.2
<input type="checkbox"/> beezdemand	Behavioral Economic Easy Demand	0.1.1
<input type="checkbox"/> BH	Boost C++ Header Files	1.72.0-3
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> blob	A Simple S3 Class for Representing Vectors of Binary Data ('BLOBS')	1.2.1

1:1 (Top Level) R Script

Console Terminal Markers Jobs

> |

## shortcuts

Ctrl + 1: lleva al editor.  
Ctrl + 2: lleva a la consola.  
Ctrl + I: borra consola.  
# en el script para poner comentarios, es decir texto que no es parte del código.  
Ctrl + enter: ejecuta línea seleccionada.

### Símbolos necesarios:

Virgulilla (~):

- Windows: altGr + +
- Mac: alt + ñ

Barra vertical/Pleca (|):

- Windows: alt + 124
- Mac: alt + 1

# Familiarizarnos con RStudio

¡Vamos a hacer nuestro primer programa!

1. En el Editor escriban la siguiente línea:

```
1 print("¡Hola Mundo")
```

2. Pongan el cursor sobre cualquier parte de la línea de código y den ctrl + click.
3. Observen la Consola:

```
> print("¡Hola Mundo")  
[1] "¡Hola Mundo"  
> |
```

Cuando por fin te compila  
tu "Hola Mundo"...



Ningún Don Ramón, ¡señor Programador!

# Hablemos de análisis de datos...

Supongamos que tenemos un experimento, ya recogimos los datos, y estamos listos para analizarlos...

Tenemos:

- 2 grupos.
- 8 sujetos/grupo.
- 5 últimas sesiones.

¿Qué prueba escogerían ustedes?

**Yo escogería un Modelo Lineal  
de Efectos Mixtos**

# ¿Por qué Modelos Lineales Mixtos?

## ANOVAs

## M. Lineales Mixtos



Lidian con datos “anidados” (nested).



No tienen problemas con datos desbalanceados



Toman en cuenta la variabilidad entre sujetos y entre ítems *al mismo tiempo*.



Indican si un efecto es significativo...



... pero también nos dicen en qué dirección se da el efecto y de que tamaño es.

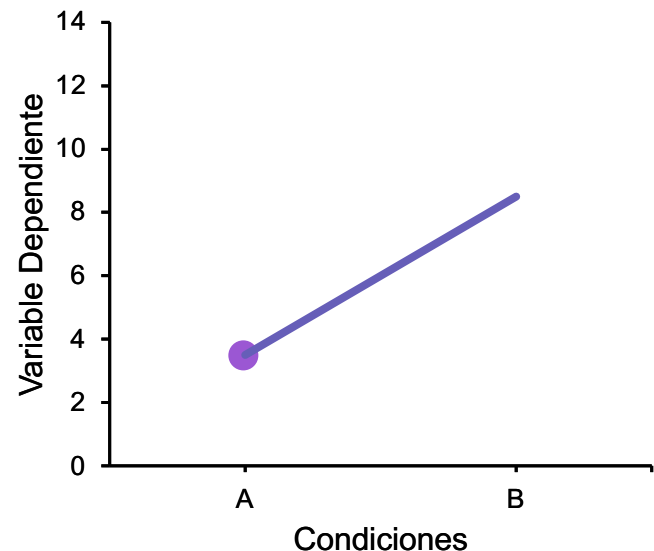


Va más allá de la mentalidad de sig. vs no sig., y provee un enfoque de “mejor modelo”.



# ¿Qué son los Modelos lineales de efectos mixtos?

- Son una extensión de los modelos lineales en los que se incluye una combinación de efectos fijos y aleatorios como variables predictoras.
- Es decir, nos van a generar una línea recta entre dos puntos (por ejemplo, condición A y condición B).
- Esta línea recta va a tener:
  - un **intercepto** (punto de partida) y
  - una **pendiente** (cambio del intercepto al segundo punto).
- La pendiente de esa línea nos va a indicar la *diferencia* entre una y otra.



# Efectos fijos vs. efectos aleatorios

## Efectos fijos

- Son aquellos que nosotros manipulamos (VI).

## Efectos aleatorios

- Son las posibles fuentes de variabilidad que *no podemos controlar* (VE), así que solo las *incorporamos* en el modelo.

~~Selección~~  
Selección



# Entonces, ¿cuándo usarlos?

- Cuando tenemos al menos **dos grupos**.
- Si tenemos **grupos desbalanceados** (con diferentes  $n$ )
- Si tenemos **datos anidados** (varias mediciones de los mismos sujetos).
- Si tenemos acceso a los **datos individuales** (que nos permiten poner los efectos aleatorios).

# ¿Qué elementos tienen los Modelos lineales de efectos mixtos?

Su forma básica es:

Variable dependiente

Variable(s) independiente(s)

error de fuente desconocida,  
que intentaremos minimizar

$$VD \sim VI + \text{efecto(s) aleatorio(s)} + \varepsilon$$

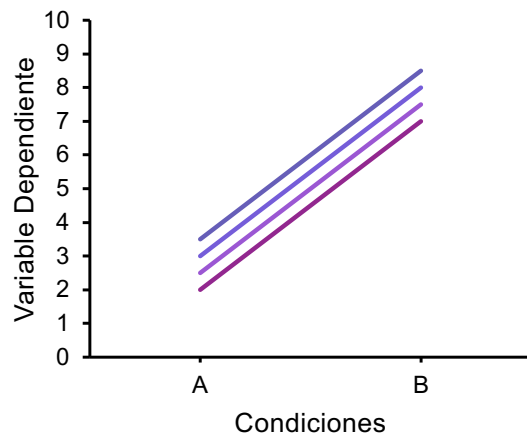
es función

posibles variables extrañas  
que podemos identificar

The diagram shows the equation  $VD \sim VI + \text{efecto(s) aleatorio(s)} + \varepsilon$ . Annotations with arrows point to each part: 'Variable dependiente' points to 'VD'; 'Variable(s) independiente(s)' points to 'VI'; 'error de fuente desconocida, que intentaremos minimizar' points to ' $\varepsilon$ '; 'es función' points to the tilde symbol '~'; and 'posibles variables extrañas que podemos identificar' points to 'efecto(s) aleatorio(s)'.

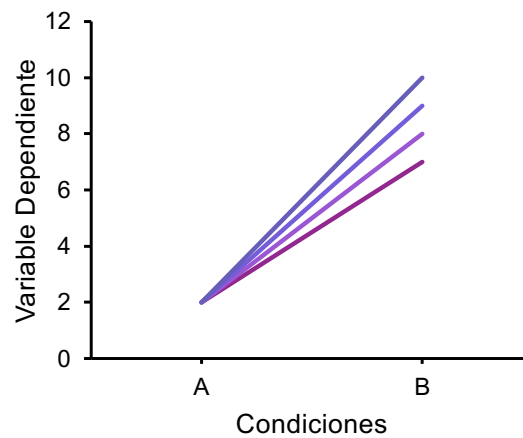
# Tipos de efectos aleatorios

Interceptos aleatorios



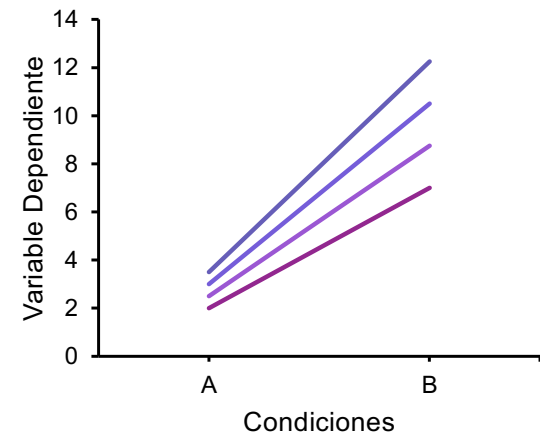
$$VD \sim VI + (1|sujeto) + \varepsilon$$

Pendientes aleatorias



$$VD \sim VI + (1 + VI|sujeto) + \varepsilon$$

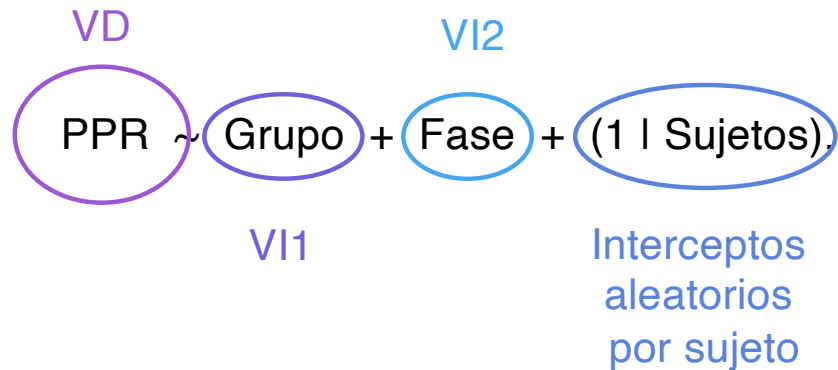
Interceptos y pendientes aleatorios



$$VD \sim VI + (1|sujeto) + (1 + VI|sujeto) + \varepsilon$$

# ¿Cómo se hace el análisis?

1. Construimos un modelo con base en nuestras variables principales.

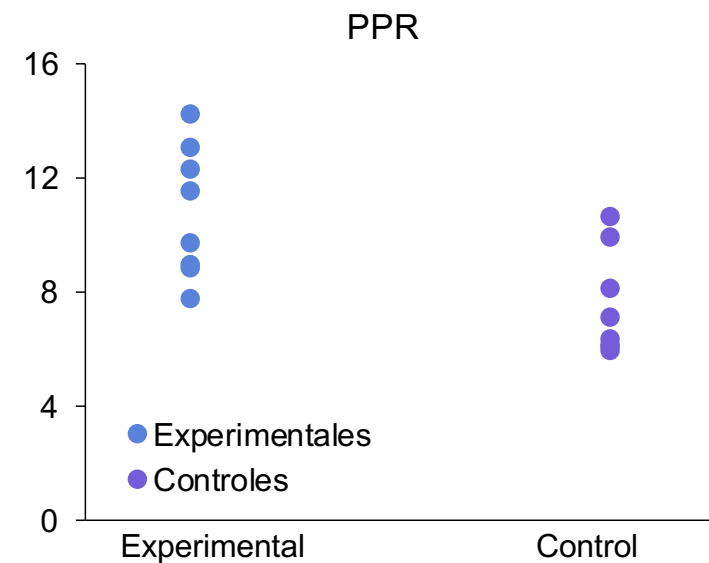


2. Comparamos el modelo con el más sencillo posible (modelo nulo) y, de ser oportuno, con algunos más complejos (por ejemplo, intercepto y pendientes aleatorias; tres variables en lugar de 2, etc.).

3. Seleccionamos el modelo que se ajuste mejor a los datos con base en una prueba de razón de verosimilitud.

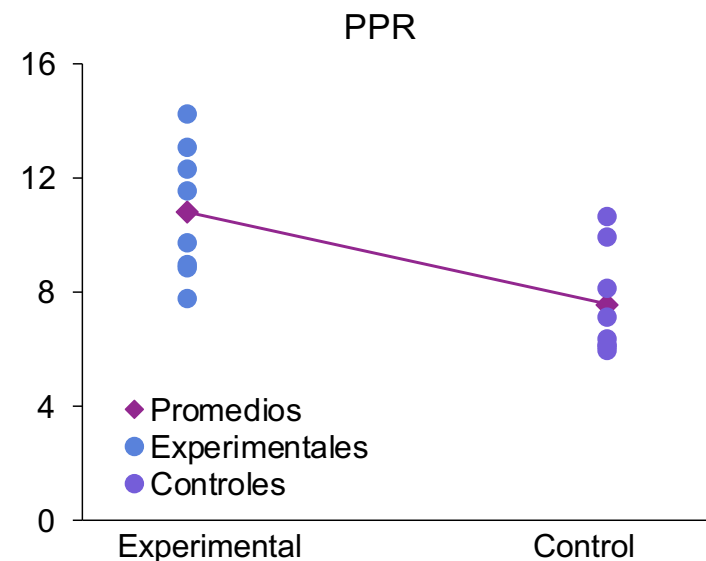
# Ejercicio guiado 1:

- Experimento de 30 sesiones, se analizan las últimas 5.
- Dos grupos: experimental (n = 8) y control (n = 8).
- Objetivo del análisis: comparar la pausa post-reforzamiento (tiempo que ocurre entre el inicio del ensayo y la primera sesión) entre grupos.



# Ejercicio guiado 1:

- Experimento de 30 sesiones, se analizan las últimas 5.
- Dos grupos: experimental (n = 8) y control (n = 8).
- Objetivo del análisis: comparar la pausa post-reforzamiento (tiempo que ocurre entre el inicio del ensayo y la primera sesión) entre grupos.



# Preparación de datos para el análisis

- De forma general, R funciona mejor con archivos .csv que .xlsx, pero hay paqueterías que nos permiten leer archivos de excel.
- Los archivos deben acomodarse de forma vertical.
- En variables de tipo *caracter*, R lee por default por orden alfabético, por lo que yo recomiendo utilizar, por ejemplo, la A para el grupo experimental y la B para el grupo control.

	A	B	C	D	E
1	IF	Sesion	Grupo	Sujeto	PPR
2	15	1 B		1	7.27213114
3	15	1 B		3	5.43934426
4	15	1 B		5	5.46557375
5	15	1 B		7	5.49016391
6	15	1 B		9	10.3967213
7	15	1 B		11	9.88688527
8	15	1 B		13	8.30983604
9	15	1 B		15	5.37049182
10	15	2 B		1	7.68032789
11	15	2 B		3	5.79344258
12	15	2 B		5	6.70000001

# Preparación del script

- Para los MLM frequentistas vamos a usar dos paquetes:
  - lme4 → hace los modelos lineales
  - lmerTest → calcula el nivel de significancia de los parámetros estimados.
  - Opcional: readxl

```
10 install.packages("lme4") #nótese que va entre comillas
11 install.packages("lmerTest")
12 install.packages("readxl") #solo si se va a trabajar con excel
13
```

- Esto se puede poner en el Editor, o directamente en la Consola.



¡Vamos a R!

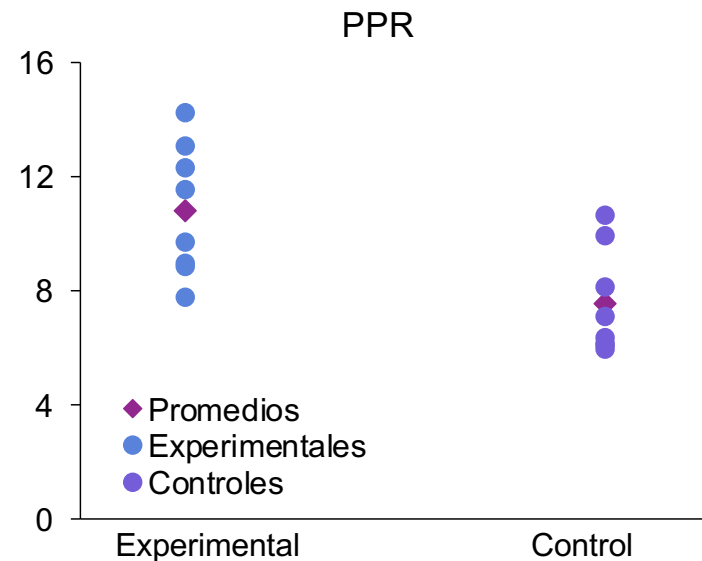
# Repaso Sesión 1:

- ¿Cuál es la diferencia entre R y RStudio?
- ¿Qué es un MLM?
- Un ejemplo de efecto fijo
- Un ejemplo de efecto aleatorio
- ¿Cuál es el primer modelo que construimos?
- ¿Cómo sabemos qué modelo elegir?
- ¿Cómo reportamos los resultados?

# Sesión 2

# Ejercicio guiado 1:

- Redacción de resultados.
  - Primero redactamos la sección de análisis de datos
  - Después se describen los resultados.
  - Es recomendable también poner disponibles tanto la base de datos como los scripts.



Resaltar sus ventajas, ya que aún no es un método muy utilizado.

Citar los artículos/tutorials que nos ayudaron a llevar a cabo el análisis.

“El uso de Modelos Lineales de Efectos Mixtos ha ido incrementando en psicología y otras ciencias de la conducta, ya que **presenta varias ventajas** sobre otros métodos estadísticos, como que toman en cuenta la varianza intra- y entre-sujetos, lidian bien con datos desbalanceados y/o anidados y nos indican la dirección y magnitud del efecto (**Boisgontier & Cheval, 2016; Brown, 2021; Harrison et al., 2018; Young, 2018; Winter, 2013**). Se utilizaron Modelos Lineales Mixtos para evaluar las diferencias en la pausa post-reforzamiento (PPR) de los grupos A y B. Primero se **construyó un modelo nulo**, y luego **se usó un test de razón de verosimilitud** para determinar **cuál es el modelo que mejor se ajustaba** a los datos. El modelo que mejor se ajustaba a los datos será descrito en la sección de resultados. Todos los análisis de efectos mixtos se llevaron a cabo usando los paquetes **lme4 (Bates, Mächler, Bolker & Walker, 2015) y lmerTest (Kuznetsova, Brockhoff & Christensen, 2017)** en R (R Team, 2012), usando el ambiente **Rstudio (Rstudio Team, 2020)**. Los códigos utilizados y bases de datos se pueden encontrar en <https://osf.io/ejemplo/>.”

Describir brevemente cómo se hizo el análisis.

Citar los paquetes y el software que utilizamos.

Fomentando la ciencia abierta, poner a disposición códigos y bases de datos.

Incluir la formula. Si son varios modelos pueden incluirse en una tabla para facilitar la lectura del texto.

Describir de qué fue función nuestra VD, con base en el mayor modelo.

Describir el modelo.

“**El mejor Modelo Lineal Mixto** para dar cuenta de los resultados del Experimento 1 **incluyó al grupo** (Experimental vs. Control) **como efecto fijo e interceptos aleatorios por sujeto** (**PPR ~ Grupo + (1|Sujeto)**). La **PPR fue función del grupo** ( $\chi^2_{(1)} = 8.45, p = .004$ ). La **PPR fue** **3.26 ± 1.04 s** (Pendiente + EEM) **mayor para el grupo A** que para el grupo B (**t = -3.122, p <.01**).”

Incluir el resultado de la prueba de verosimilitud.

Poner el tamaño y dirección de la(s) pendiente(s)/efecto.

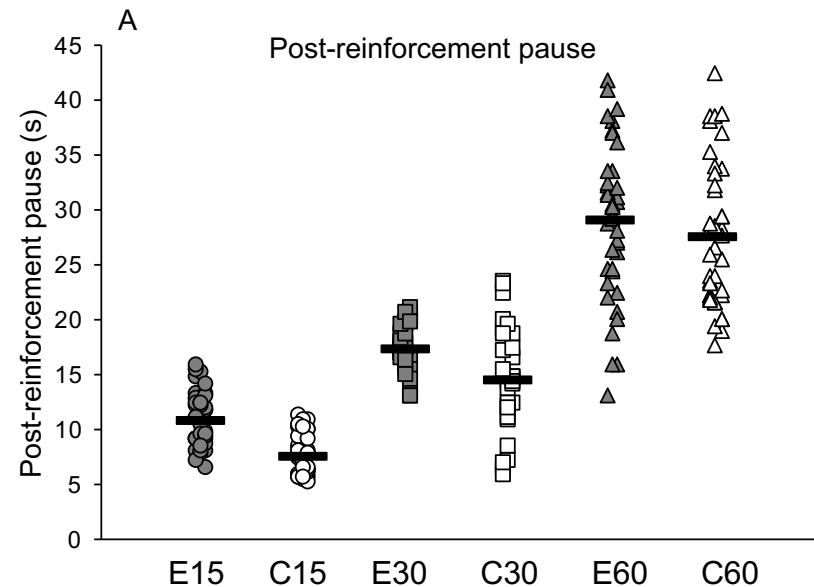
Si lo consideraran pertinente, poner la significancia de la(s) pendiente(s).

## Ejercicio guiado 2a:

- Es la base de datos completa del Ejercicio 1.
- 30 sesiones, analizaremos las últimas 5.
- 6 grupos:

	IF 15	IF 30	IF 60
Experimental	n = 8	n = 6	n = 8
Control	n = 8	n = 6	n = 8

- Objetivo del análisis: comparar la pausa post-reforzamiento de los grupos experimentales y control, para los 3 valores de IF.



## Ejercicio guiado 2a:

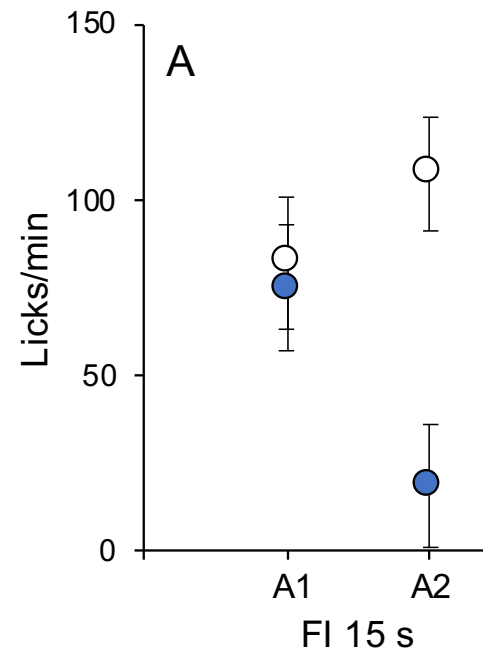
- Redacción de resultados.

“El mejor Modelo Lineal Mixto para dar cuenta de los resultados de este estudio incluyó al intervalo fijo (IF) (15, 30 o 60 s) como efecto fijo e interceptos aleatorios por sujeto ( $PPR \sim IF + (1|Sujeto)$ ). La pausa post-reforzamiento fue función del intervalo fijo, ( $\chi^2_{(2)} = 65.45, p = .000$ ). La pausa post-reforzamiento incremento conforme aumento el IF, siendo  $6.71 \pm 1.76$  s mayor durante el IF 30, que durante el IF 15 ( $t = 3.81, p = .000$ ) y  $19.09 \pm 1.63$  s mayor durante el IF 60 que durante el IF 15 ( $t = 11.713, p = .000$ ).”



## Ejercicio guiado 2b:

- Dos grupos, experimental y control (n = 4).
- Se evaluó la cantidad de bebida (licks) en ratas, comparando dos condiciones.
- Se incluyen las últimas 3 sesiones de la fase 1 y las primeras 3 de la fase 2.
- Objetivo del análisis: comparar la tasa de licks en ambas fases, y en función del grupo.



## Ejercicio guiado 2b:

- Redacción de resultados.

“**El mejor Modelo Lineal Mixto** para dar cuenta de los resultados de este estudio **incluyó al grupo** (Experimental vs. Control) **y la fase** (1 vs. 2) **como efectos fijos y pendientes aleatorios entre fases para cada sujeto** (**Licks ~ Grupo + Fase + (1 + Fase|Sujeto)**). La **tasa de licks fue función del grupo y las fases** ( $\chi^2_{(2)} = 7.65, p = .021$ ). La **tasa de licks del grupo control fue mayor por  $59.456 \pm 22.86$  (Pendiente + EEM) licks que la del grupo experimental** ( $t = 2.60, p = .041$ ). Por otro lado, **la tasa de licks fue menor por  $13.943 \pm 11.549$  licks durante la fase 2**, aunque la diferencia entre fases no fue significativa ( $t = -1.207, p = .267$ ). Sin embargo, **las pendientes entre fases tendieron a ser más grandes para los sujetos del grupo Experimental (1-4), que para los del grupo control (5-8), como se puede observar en la tabla X.**”

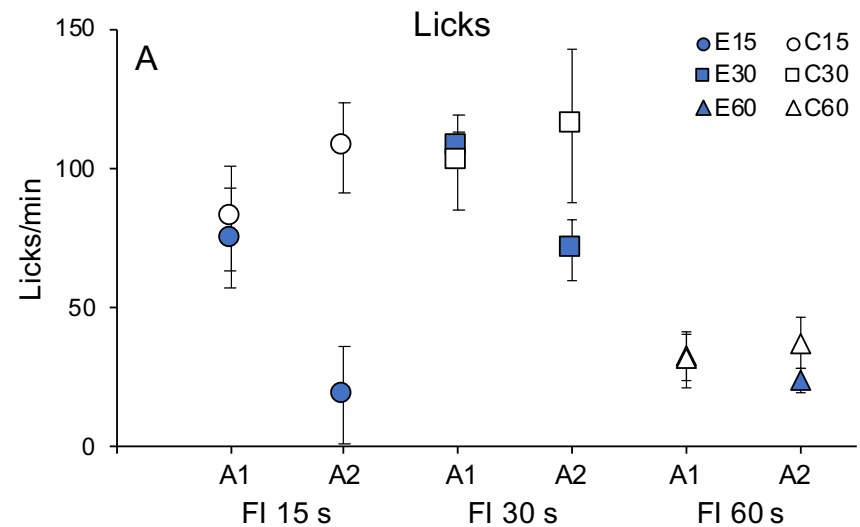
# Repaso sesiones 1 y 2:

- ¿Cuál es la diferencia entre R y RStudio?
- ¿Qué es un MLM?
- Un ejemplo de efecto fijo y uno de aleatorio
- ¿Cuál es el primer modelo que construimos?
- ¿Cómo sabemos qué modelo elegir?
- Dos cosas que se deban incluir en la redacción del análisis de datos.
- Tres datos que se deban incluir en la redacción de la sección de resultados.

# Sesión 3

## Ejercicio guiado 2c:

- Es la base de datos completa del Ejercicio 2b.
- Analizamos la última sesión de la fase 1, y la primera de la fase 2, en función del grupo, en función del IF.
- Objetivo del análisis: comparar la diferencia en la tasa de licks que hay entre ambos grupos, y ver cómo interactúa la diferencia con el IF.



## Ejercicio guiado 2c:

- Redacción de resultados.

“**Los mejores Modelos Lineales Mixtos** para dar cuenta de los resultados de este estudio **incluyeron la interacción entre Grupo** (Experimental y Control) **y Fase** (1 y 2) **como efecto fijo e interceptos aleatorios por sujeto** (**Licks ~ Grupo\*Fase + (1|Sujeto)**), para los grupos de IF 15 e IF 30, mientras que para el grupo de IF 60, ninguno de los modelos fue mejor que el nulo. Los **licks fueron función de la interacción durante el IF 15** ( $\chi^2_{(1)} = 8.87, p = .003$ ), **y durante el IF 30** ( $\chi^2_{(1)} = 9.11, p = .003$ ), mientras que la tasa de Licks durante el IF 60 no fue función de las variables independientes manipuladas en este estudio. **El efecto de la interacción fue mayor durante el IF 15** (82.26;  $t = 3.49, p = .012$ ), **que durante el IF 30** (49.95;  $t = 3.56, p = .011$ ).

# Fin de MLM-frecuentistas

¿Preguntas? ¿Dudas?

# Requisitos para la obtención de certificado

## Ejercicios guiados (50%)

1. Ejercicio 1: Modelo lineal mixto con una variable independiente.
2. Ejercicio 2a: Modelo lineal mixto con dos variables independientes. Interceptos aleatorios.
3. Ejercicio 2b: Modelo lineal mixto con dos variables independientes. Pendientes aleatorias.
4. Ejercicio 2c: Modelo lineal mixto con dos variables independientes. Interacción.
5. Ejercicio 3: Modelo lineal mixto Bayesiano.

- **Puntajes:**

- Scripts: **7% c/u (total: 35%)**
- Interpretación: **3% c/u (total 15%)**

## Ejercicio no guiado (50%)

	Sección	Porcentaje
1.	Base de datos en formato (largo), puede ser extensión .csv o .xlsx.	<b>5%</b>
2.	Breve descripción de las variables a analizar (definición y/o niveles) (50-100 palabras).	<b>3%</b>
3.	Justificación del tipo de análisis a utilizar: frecuentista o Bayesiano (50-100 palabras).	<b>2%</b>
4.	Script en formato R.	<b>30%</b>
5.	Interpretación de resultados.	<b>10%</b>
	TOTAL:	<b>50%</b>

Fecha de entrega: 08-diciembre-2022  
Envío a: glopez@psi.uned.es



# Frecuentista vs Bayesiana

## Frecuentista

- Vamos a averiguar la probabilidad de obtener esos resultados si la hipótesis *nula* es verdadera.
  - Mi hipótesis nula es que no hay diferencias entre grupo A y B.
  - Mi hipótesis alternativa es que hay diferencias entre A y B.
  - Si  $p = 0.04$ , eso significa que la probabilidad de que la hipótesis nula sea cierta es del 4% (es decir, muy baja).

## Bayesiana

- Vamos a calcular la probabilidad de que cada una de las hipótesis sea cierta, dado:
  - Información previa que tengamos
  - La evidencia de los datos obtenidos.

Hipótesis nula: todos los estudiantes del taller tienen (por lo menos) un título de master.

Primer ensayo:

A favor de la hipótesis *Nula*:

6

A favor de la hipótesis *Alternativa*:

4

Segundo ensayo (con info adicional):

A favor de la hipótesis *Nula*:

0

A favor de la hipótesis *Alternativa*:

10

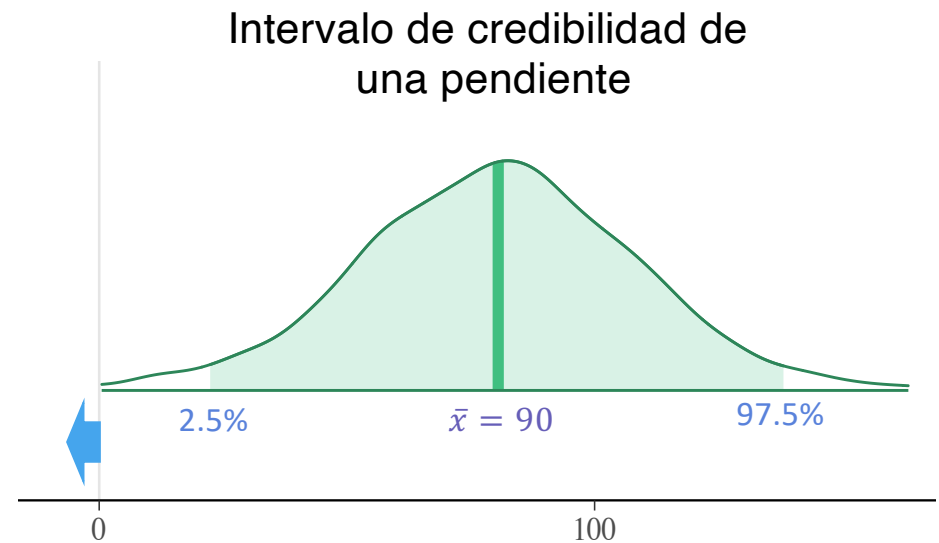
# Ventajas de la estadística Bayesiana

- De forma general, la estadística Bayesiana enfatisa la evidencia sobre las decisiones.
- El valor  $p$  nos dice la probabilidad de obtener ese valor (o uno así de extremo), si la hipótesis nula es cierta.
- Con Bayes no decidimos si un  $p < .05$  (o a  $.01$ ) es significativo de forma 'arbitraria' (lo cuál puede ser la causa de la crisis de replicabilidad), sino que se determina qué tan probable es que la hipótesis nula o alternativa sea cierta con base en los datos analizados.
- El teorema de Bayes indica la probabilidad de que una hipótesis sea cierta, basado en una distribución anterior (prior), y dado cierta evidencia.
- Puede aceptarse la hipótesis nula.
- Permite la construcción de conocimiento de forma más sencilla (por ejemplo, usando los datos de estudios anteriores para determinar nuestra prior y contrastarlo directamente con nuestros datos).

Recomiendo que lean Wagenmakers et al., 2018; y Young, 2019 para ampliar información

# Estadística Bayesiana

- Al análisis bayesiano nos da una distribución (la distribución posterior) de los parámetros plausibles y su probabilidad relativa.
- La **media de la distribución** representa el valor con mayor probabilidad de ocurrir.
- En el caso de los MLMB, nos da la distribución de posibles resultados del intercepto, la pendiente, etc.
- **Intervalo de credibilidad** (2.5%-97.5% de las muestras) nos puede aportar información adicional, por ejemplo, ¿cuál es la probabilidad de que mi pendiente sea negativa ( $< 0$ )?



## BRMS- algunos términos nuevos.

- En Bayes vamos a estar hablando de “converger”.
- Que un modelo converja quiere decir que se encontró una **solución estable**.
- Es como armar un puzzle, y que todas las piezas encajen.
- Para encontrar esa solución se llevarán a cabo **iteraciones**, que significa probar sustituyendo cada elemento del modelo con valores distintos, hasta encontrar una solución.

$$A + B = 5$$

$$1 + 1 = \cancel{5}$$

$$2 + 2 = \cancel{5}$$

$$3 + 3 = \cancel{5}$$

$$3 + 2 = 5$$

Se encontró  
**una** solución.

Cadenas-conjunto de iteraciones.

$$A + B = 5$$

$$1 + 1 = \cancel{5}$$

$$1 + 2 = \cancel{5}$$

$$1 + 3 = \cancel{5}$$

$$1 + 4 = 5$$

$$A + B = 5$$

$$1 + 1 = \cancel{5}$$

$$2 + 1 = \cancel{5}$$

$$3 + 1 = \cancel{5}$$

$$4 + 1 = 5$$

$$A + B = 5$$

$$1 + 1 = \cancel{5}$$

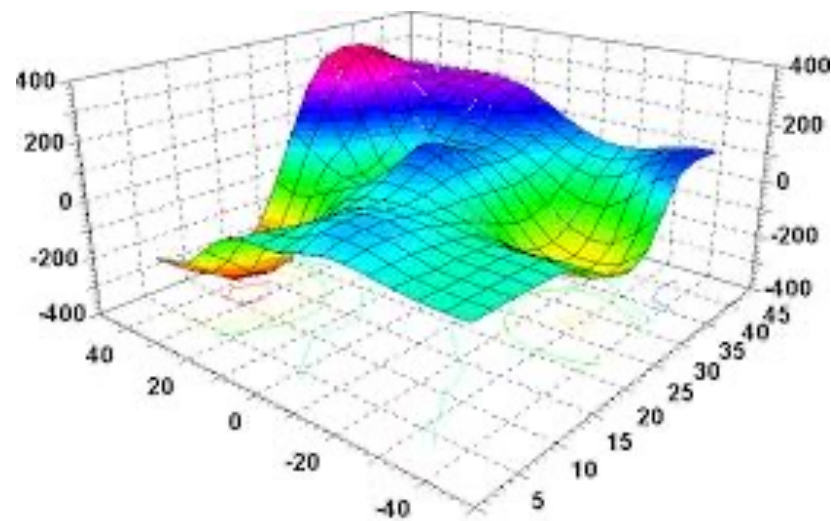
$$2 + 2 = \cancel{5}$$

$$3 + 3 = \cancel{5}$$

$$3 + 2 = 5$$

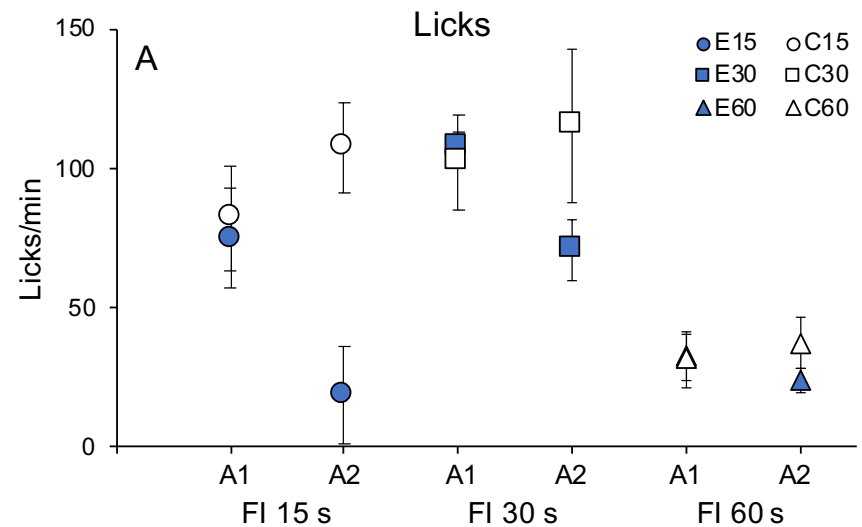
Iteraciones

*Seed*- indica el punto de partida de las interacciones



## Ejercicio 3

- Vamos a re-analizar los datos del ejercicio 2c.
- Analizamos la última sesión de la fase 1, y la primera de la fase 2, en función del grupo, en función del IF.
- Objetivo del análisis: comparar la diferencia en la tasa de licks que hay entre ambos grupos, y ver cómo interactúa la diferencia con el IF.





# Repaso sesiones 1-3:

- ¿Cuál es la diferencia entre R y RStudio?
- ¿Qué es un MLM?
- Un ejemplo de efecto fijo y uno de aleatorio
- ¿Cómo sabemos qué modelo elegir?
- Menciona alguna diferencia entre estadística Bayesiana y Frecuentista.
- ¿Qué son las iteraciones?

# Requisitos para la obtención de certificado

## Ejercicios guiados (50%)

1. Ejercicio 1: Modelo lineal mixto con una variable independiente.
2. Ejercicio 2a: Modelo lineal mixto con dos variables independientes. Interceptos aleatorios.
3. Ejercicio 2b: Modelo lineal mixto con dos variables independientes. Pendientes aleatorias.
4. Ejercicio 2c: Modelo lineal mixto con dos variables independientes. Interacción.
5. Ejercicio 3: Modelo lineal mixto Bayesiano.

- **Puntajes:**

- Scripts: **7% c/u (total: 35%)**
- Interpretación: **3% c/u (total 15%)**

## Ejercicio no guiado (50%)

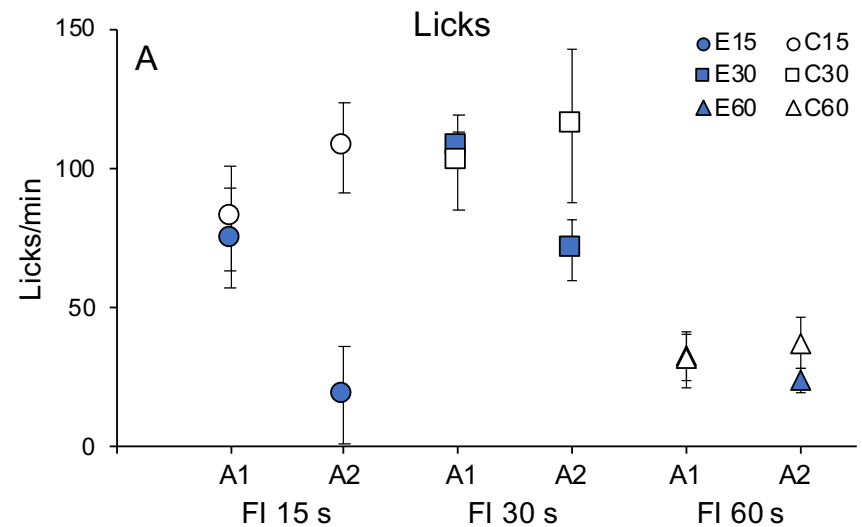
	Sección	Porcentaje
1.	Base de datos en formato (largo), puede ser extensión .csv o .xlsx.	<b>5%</b>
2.	Breve descripción de las variables a analizar (definición y/o niveles) (50-100 palabras).	<b>3%</b>
3.	Justificación del tipo de análisis a utilizar: frecuentista o Bayesiano (50-100 palabras).	<b>2%</b>
4.	Script en formato R.	<b>30%</b>
5.	Interpretación de resultados.	<b>10%</b>
	TOTAL:	<b>50%</b>

Fecha de entrega: 08-diciembre-2022  
Envío a: [glopez@psi.uned.es](mailto:glopez@psi.uned.es)

# Sesión 4

## Ejercicio 3

- Vamos a re-analizar los datos del ejercicio 2c.
- Analizamos la última sesión de la fase 1, y la primera de la fase 2, en función del grupo, en función del IF.
- Objetivo del análisis: comparar la diferencia en la tasa de licks que hay entre ambos grupos, y ver cómo interactúa la diferencia con el IF.



## Ejercicio guiado 3:

- Análisis de datos

... se comparó el mejor modelo, con el nulo, y se calculó el factor Bayes de la comparación. Los parámetros estimados se obtuvieron corriendo simulaciones de 8000 iteraciones (4 cadenas de 2000 iteraciones, con 1000 de calentamiento), de forma que se obtuvo una distribución posterior de 4000 muestras para cada parámetro. Se utilizó el estadístico  $\hat{R}$  para verificar la convergencia, y todos los estimados cayeron en el rango de  $\hat{R} = 1 \pm 0.$ , indicando que el modelo convergió correctamente. Para evaluar la fuerza de la evidencia de que la interacción fue mayor (o menor, en función de la dirección de la pendiente) que cero, se utilizó la probabilidad posterior ( $P(\delta)$ ). Una probabilidad posterior mayor (o menor) a 95% ( $P(\delta) > 0 = .95$ ) se consideró evidencia convincente.

## Ejercicio guiado 3:

- Redacción de resultados.

“El mejor Modelo Lineal Mixto Bayesiano ( $FB_{10} = 64434824.20$ ). para dar cuenta de los resultados de este experimento **incluyó la interacción entre Grupo** (Experimental y Control) **y Fase** (1 y 2) **como efecto fijo e interceptos aleatorios por sujeto** (**Licks ~ Grupo\*Fase + (1|Sujeto)**). Las **interacciones positivas indicaron** que **hubo evidencia convincente** de que la **tasa de licks fue mayor en la Fase A que en la B para el grupo experimental, pero menor en la fase A que en la fase B para el grupo control** ( $E(\mu) 81.70 \pm 30.56$  (95% I.C. = 20.00 – 143.25,  $P(\delta) > 0 = .995$ ).”

# Reporte de Resultados conforme a la Ciencia Abierta

# ¿Qué es la ciencia abierta?

- Un movimiento que busca que el conocimiento científico sea **accesible para todas las personas**.
- Es importante porque permite un **mayor avance** en la ciencia.
- **Características:**
  - Transparencia en sus prácticas
  - Conocimiento accesible
    - Para el público en general (fácil de entender)
    - Para el público especialista (fácil de obtener).
  - Es fácilmente replicable y reproducible
  - Favorece la colaboración



# Ciencia Abierta

Genera “conocimiento que se puede **acceder, usar, modificar y compartir libremente** para cualquier propósito (si acaso con requerimientos para preservar su procedencia y apertura)”  
(Open Knowledge Foundation, s/f).

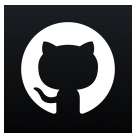
Esta apertura/libertad aplicaría para:

- Bases de datos
- Software
- Tecnología (hardware)
- Contenido
- Técnicas o métodos

# ¿Cómo hacer ciencia abierta?

- Cuidar los estándares de las citas/referencias, particularmente de los métodos y datos.
  - Transparencia en los datos, haciéndolos públicos.
  - Transparencia en los métodos de análisis, poniéndolos a disposición de otros científicos.
- Transparencia (explicación adecuada) en los métodos y protocolos utilizados.
  - Transparencia en el diseño de los estudios.
  - Pre-registro de los estudios.
  - Pre-registro de los planes de análisis.
  - Fomento de la replicación.

# ¿Dónde poner los datos y scripts?



**GitHub**

- Es un servicio que permite almacenar y seguir los cambios en códigos.
- Está enfocado en programación.
- No es muy intuitivo.
- Permite cambiar la privacidad de los repositorios.

**OSF**



- Es un repositorio del *Center for Open Science*, que permite crear proyectos para compartir datos, código, y otros elementos.
- Su uso es bastante intuitivo aunque nos e tenga experiencia.
- Permite añadir colaboradores de forma fácil.
- Permite crear DOIs para compartir los proyectos.
- Permite cambiar la privacidad de los repositorios.
- Está conectado con otros recursos de ciencia abierta, como los preprints.



32.4MB

Private

Make Public

0



# Behavioral history and SID

Project

Analysis

Communication

Data

Hypothesis

Instrumentation

Methods and Measures

Procedure

Project

Software

Other

Uncategorized

lón

2022-04-15 02:04 AM

Name

Modified

Behavioral history and SID

- OSF Storage (United States)

Abstract.pdf

2022-04-15 01:57 AM

+ Additional Material

Data.xlsx


2022-04-15 01:46 AM

Script\_BLMM\_Licks\_feb2022.R

2022-04-15 01:46 AM




Script\_BLMM\_LP\_feb2022.R


2022-04-15 02:01 AM





Search or jump to...


[Pull requests](#) [Issues](#) [Codespaces](#) [Marketplace](#) [Explore](#)


  

 **GElopezTolsa / Timing\_and\_SIB** Public


 Unpin


 Unwatch **1**


 Fork **0**

 Star **0**


[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#) [Settings](#)


 main


 1 branch


 0 tags


Go to file


 **GElopezTolsa** Merge branch 'main' of <https://github.com/GElopezTolsa/Ti...> be56736


 Posterior\_distributions\_and\_PPche... updated PPchecks files

 R\_codes\_and\_data\_sets Added codes and datasets

 Supplemental\_material Added codes and datasets


 .DS\_Store updated PPchecks files

 .gitattributes Initial commit


 README.md Update README.md

README.md


# Timing\_and\_SIB


 main


**Timing\_and\_SIB / R\_codes\_and\_data\_sets /**


 **GElopezTolsa** Added codes and datasets


..


 BLMM\_LP\_Licks.R Added cod


 BLMM\_Last\_Lick.R Added cod


 BLMM\_PRP\_BP.R Added cod

 BLMM\_Peak\_Licks.R Added cod

 Data\_LP\_Licks.csv Added cod

 Data\_PRP\_BP.csv Added cod

 data\_licks.csv Added cod

 data\_peak.csv Added cod