Expandiendo el modelo

Eduardo Garcia-Garzon

IA Specialist

AEMCCO 2022 TERUEL

SHAKERS

Introducción

1. Añadiendo complejidad
2. El modelo de pendientes aleatorias
3. Añadiendo predictores
4. Interacciones inter-niveles

Key messages

1. Ajustar pendientes aleatorias.
2. Especificar predictores.
3. Los límites de Jamovi.

Expandiendo el modelo

- Hasta ahora, los modelos vistos son modelos de interceptos aleatorios.
- Únicamente valoramos la presencia de desviaciones de la media global de la variable.
- Podemos tener modelos de intercepto aleatorio con uno o más efectos aleatorios y con predictores en los diferentes niveles de interés.
- Podemos predecir variables categóricas utilizando el modelo lineal generalizado.

Añadiendo complejidad

Añadir más efectos aleatorios

- Nos permite descomponer la varianza en diferentes componentes.
- Cuidado con el tipo de efectos (cruzados vs jerárquicos).
- En el ejemplo: ¿qué tipo de efectos tenemos?

Añadir más efectos aleatorios

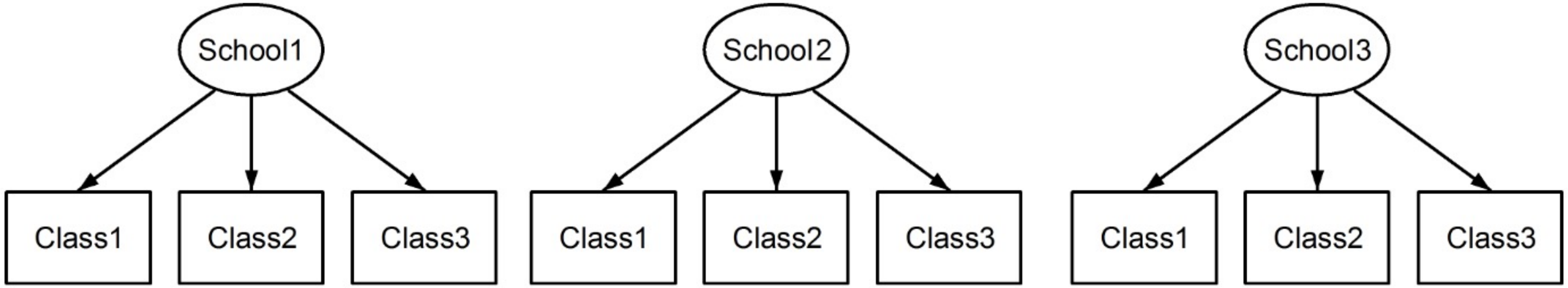
- Nos permite descomponer la varianza en diferentes componentes.
- Cuidado con el tipo de efectos (cruzados vs jerárquicos).
- En el ejemplo: ¿qué tipo de efectos tenemos?
  - **¡Cruzados!**
  - **Trial 1 para el sujeto 1 es diferente que trial 1 para sujeto 2**

Random Components

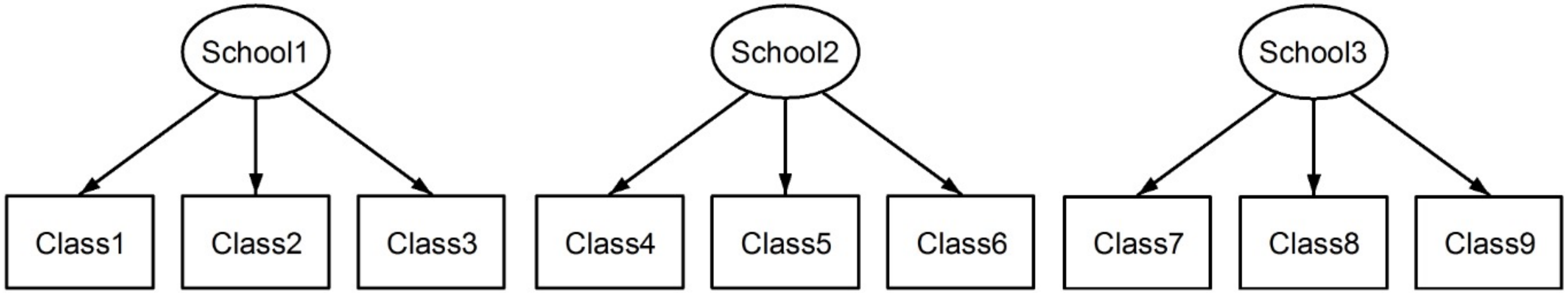
Groups	Name	SD	Variance	ICC
sujeto	(Intercept)	1.96	3.86	0.27
trial	(Intercept)	2.48	6.15	0.38
Residual		3.19	10.20	

Nota. Number of Obs: 1800 , groups: sujeto 50, trial 36

With this naming convention, nested and fixed effects will be **different**:



With this naming convention, nested and fixed effects will be **equivalent**:



Añadir más efectos aleatorios

- "In other words, for a nested structure, there is an equivalent crossed structure. (The reverse is not true; Errikson, 2021)".
- **Jamovi (lme4) estima ambos modelos de manera similar –> clave estructura de los datos**

Random Components

Groups	Name	SD	Variance	ICC
trial	(Intercept)	2.31	5.33	0.27
condicion	(Intercept)	1.20	1.44	0.09
Residual		3.75	14.06	

Nota. Number of Obs: 1800 , groups: trial 36, condicion 2

Añadir más efectos aleatorios

- Podemos complicarlo con tres niveles.
- Cuidado, Jamovi entiende que todos los niveles son cruzados.
- Los resultados varían en función de nuestro diseño experimental.

Random Components

Groups	Name	SD	Variance	ICC
sujeto	(Intercept)	1.96	3.86	0.27
trial	(Intercept)	2.33	5.41	0.35
condicion	(Intercept)	1.20	1.44	0.12
Residual		3.19	10.20	

Nota. Number of Obs: 1800 , groups: sujeto 50, trial 36, condicion 2

Predictores

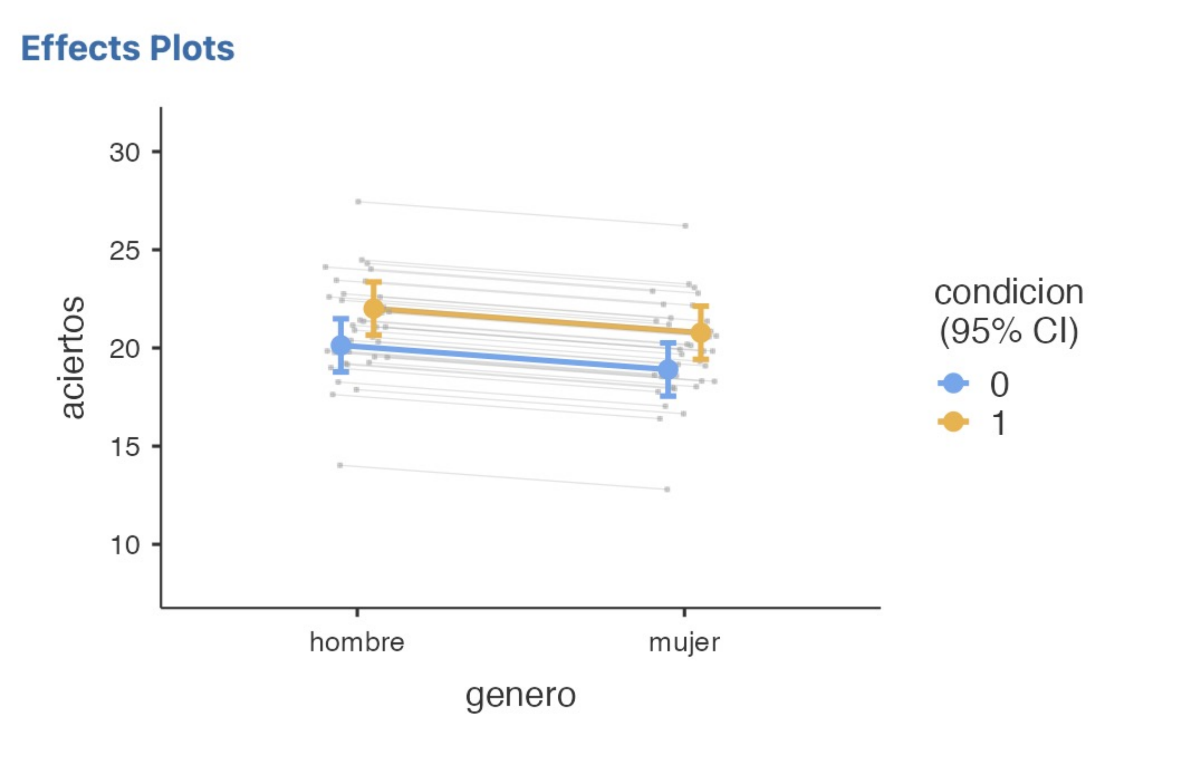
- Podemos tener predictores en todos los niveles de nuestro modelo
- Cuidado con la especificación en base de datos.
- Si son predictores continuos, es necesario considerer cómo centramos nuestros datos.

Condición – predictor nivel trial
Genero – predictor nivel sujeto

Fixed Effects Parameter Estimates

				95% Confidence Interval				
Names	Effect	Estimate	SE	Lower	Upper	df	t	p
(Intercept)	(Intercept)	20.45	0.48	19.52	21.39	61.50	42.94	<.001
genero1	mujer - hombre	–1.23	0.55	–2.32	–0.14	48.00	–2.22	0.031
condicion1	1 - 0	1.87	0.79	0.32	3.42	34.00	2.37	0.024

Condición – predictor nivel trial
Genero – predictor nivel sujeto



Interacción inter-nivel
Dos variables de diferente nivel

Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval Lower	95% Confidence Interval Upper	df	t	p
(Intercept)	(Intercept)	20.45	0.48	19.52	21.39	61.50	42.94	<.001
genero1	mujer - hombre	–1.23	0.55	–2.32	–0.14	48.00	–2.22	0.031
condicion1	1 - 0	1.87	0.79	0.32	3.42	34.00	2.37	0.024
genero1 ✻ condicion1	mujer - hombre ✻ 1 - 0	–0.35	0.30	–0.94	0.24	1714.00	–1.17	0.241

Why You Should *Always* Include a Random Slope for the Lower-Level Variable Involved in a Cross-Level Interaction*

Jan Paul Heisig
WZB Berlin Social Science Center

Merlin Schaeffer
University of Copenhagen

dictor and the outcome. We argue that multilevel models involving cross-level interactions should always include random slopes on the lower-level components of those interactions. Failure to do so will usually result in severely anti-conservative statistical inference.

Introducing a random slope term on the lower-level variable involved in a cross-level interaction, reduces the absolute *t*-ratio by 31% or more in three quarters of cases, with an average reduction of 42%.

Ejemplo 3

- Vamos a intentar predecir el afecto positivo de nuestros teniendo en cuenta:
    - Datos anidados en sujetos (id) y practicas (diary)
    - Variables nivel 1: performance_type & audience
    - Variables nivel 2: years_study & absorption
- ¿A qué modelo final llegamos?
- ¿Qué conclusions extraemos?

Añadiendo complejidad

Que sería de un taller de metodología sin ecuaciones….

Tradicionalmente, condensamos la notación a estas dos ecuaciones principales..

$$Nota = (b_{intercepto\_estudiante}) + b_{tutoría} * tutoría + \varepsilon$$

$$b_{intercepto\_estudiante} \sim N\ (b_{intercepto}, \tau)$$

$$\varepsilon \sim N\ (0, \sigma)$$

Que sería de un taller de metodología sin ecuaciones….

Tradicionalmente, condensamos la notación a estas dos ecuaciones principales..

$$Nota = (b_{intercepto} + effestud_0) + b_{tutoría} * tutoría + \varepsilon$$

$$effestud_0 \sim N\ (0, \tau_0)$$

$$\varepsilon \sim N\ (0, \sigma)$$

Que sería de un taller de metodología sin ecuaciones….

¿Cómo se amplia ahora la estructura de var-covar de los efectos?

$$Nota = (b_{intercepto} + effestud_0) + (b_{tutoría} + effestud_1) * tutoría + \varepsilon$$

$$\varepsilon \sim N\ (0, \sigma)$$

$$?$$

Que sería de un taller de metodología sin ecuaciones….

Tradicionalmente, condensamos la notación a estas ecuaciones principales..

$$Nota = (b_{intercepto} + effestud_0) + (b_{tutoría} + effestud_1) * tutoría + \varepsilon$$

$$\begin{matrix} effestud_0 \\ effestud_1 \end{matrix} = MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_0 & \rho_{01} \\ \rho_{01} & \tau_1 \end{bmatrix}\right)$$

$$\varepsilon \sim N\,(0, \sigma)$$

Que sería de un taller de metodología sin ecuaciones….

Tradicionalmente, condensamos la notación a estas ecuaciones principales..

$$Nota = b_{intercepto\_estud} + b_{tutoría\_estud} * tutoría + \varepsilon$$

$$\begin{matrix} b_{intercepto\_estud} \\ b_{tutoría\_estud} \end{matrix} = MVN\left(\begin{bmatrix} b_{intercepto} \\ b_{tutoría} \end{bmatrix}, \begin{bmatrix} \tau_0 & \rho_{01} \\ \rho_{01} & \tau_1 \end{bmatrix}\right)$$

$$\varepsilon \sim N\,(0, \sigma)$$

Que sería de un taller de metodología sin ecuaciones….

Tradicionalmente, condensamos la notación a estas dos ecuaciones principales..

$$Nota = b_{intercepto\_estudiante} + b_{tutoría} * tutoría + \varepsilon$$

$$b_{intercepto\_estudiante} \sim N\ (b_{intercepto}\ , \tau)$$

http://mfviz.com/hierarchical-models/?fbclid=IwAR074-goLDlzwMwWmq86qINQRGIEclsm_SAfXin4JulC_RSGZU6MJlDBXTc

Que sería de un taller de metodología sin ecuaciones….

Tradicionalmente, condensamos la notación a estas ecuaciones principales..

$$Nota = b_{intercepto\_estud} + b_{tutoría\_estud} * tutoría + \varepsilon$$

$$\begin{matrix} b_{intercepto\_estud} \\ b_{tutoría\_estud} \end{matrix} = MVN\left(\begin{bmatrix} b_{intercepto} \\ b_{tutoría} \end{bmatrix}, \begin{bmatrix} \tau_0 & \rho_{01} \\ \rho_{01} & \tau_1 \end{bmatrix}\right)$$

$$\varepsilon \sim N\,(0, \sigma)$$

ICC con pendientes aleatorias

El uso del ICC si hay pendientes aleatorias no es sencillo…

Caution: For models with random slopes and random intercepts, the ICC would differ at each unit of the predictors. Hence, the ICC for these kind of models cannot be understood simply as proportion of variance (see *Goldstein et al. 2010*). For convenience reasons, as the `` `icc()` `` function also extracts the different random effects variances, the ICC for random-slope-intercept-models is reported nonetheless, but it is usually no meaningful summary of the proportion of variances.

To get a meaningful ICC also for models with random slopes, use `` `adjusted = TRUE` ``. The adjusted ICC uses the mean random effect variance, which is based on the random effect variances for each value of the random slope (see *Johnson et al. 2014*).

Ejemplo de una mala idea
(sin interceptos aleatorios)

Random Components

Groups	Name	SD	Variance	ICC
sujeto	condicion0	1.70	2.90	
	condicion1	2.16	4.67	
Residual		3.92	15.39	

Nota. Number of Obs: 1800 , groups: sujeto 50

Random Parameters correlations

Groups	Param.1	Param.2	Corr.
sujeto	condicion0	condicion1	0.99

The Importance of Random Slopes in Mixed Models for Bayesian Hypothesis Testing

Psychological Science
2022, Vol. 33(4) 648–665
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/09567976211046884
www.psychologicalscience.org/PS
SAGE

Klaus Oberauer
Department of Psychology, University of Zurich

For classical frequentist statistics, Barr et al. (2013) demonstrated through analysis and simulations that neglecting random effects can lead to a serious inflation of Type I errors, or false positives (i.e., obtaining a significant result for an effect that is zero). In particular, if there are true individual differences in the effects of predictors in the data, but the model does not specify them as random slopes, false-alarm rates increase substantially above the nominal alpha level. Conversely, including random effects that are not warranted by the data does not jeopardize the validity of statistical inferences. Barr et al. (2013) therefore recommended keeping mixed models "maximal" by default, that is, to include all random effects and their correlations. Matuschek

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Balancing Type I error and power in linear mixed models

Hannes Matuschek [a,*], Reinhold Kliegl [a], Shravan Vasishth [a], Harald Baayen [b], Douglas Bates [c]

[a] University of Potsdam, Germany
[b] University of Tübingen, Germany
[c] University of Wisconsin-Madison, USA

[HTML] Balancing Type I error and power in linear mixed models

C[x] CODE

H Matuschek, R Kliegl, S Vasishth, H Baayen… - Journal of memory and …, 2017 - Elsevier

Linear mixed-effects models have increasingly replaced mixed-model analyses of variance for statistical inference in factorial psycholinguistic experiments. Although LMMs have many advantages over ANOVA, like ANOVAs, setting them up for data analysis also requires some care. One simple option, when numerically possible, is to fit the full variance-covariance structure of random effects (the maximal model; Barr, Levy, Scheepers & Tily, 2013), presumably to keep Type I error down to the nominal α in the presence of random effects …

☆ Guardar Citar Citado por 1060 Artículos relacionados Las 13 versiones

Ejemplo de una idea más acertada

Random Components

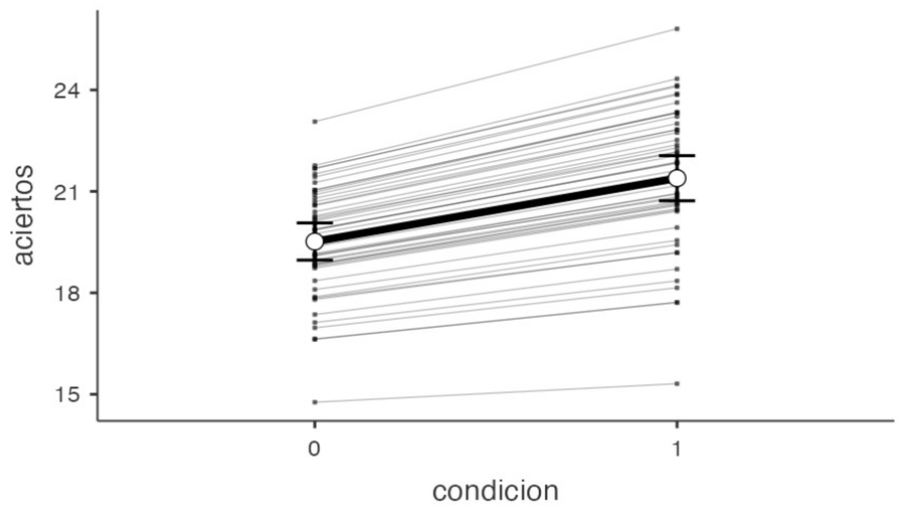
Groups	Name	SD	Variance	ICC
sujeto	(Intercept)	1.93	3.71	0.19
	condicion1	0.52	0.27	
Residual		3.92	15.39	

Nota. Number of Obs: 1800 , groups: sujeto 50

Random Parameters correlations

Groups	Param.1	Param.2	Corr.
sujeto	(Intercept)	condicion1	0.88

Effects Plots



Trabajando con pendientes aleatorias

- Siempre testear la necesidad de incluir pendientes aleatorias.
- Si introducimos las mismas, añadir el intercepto aleatorio.
- Comparaciones basadas en ajuste.

Workflow de modelos mixtos

- 1. Modelo nulo de intercepto aleatorio → ICC
- 2. Modelo intercepto aleatorio con covariables nivel 1.
- 3. Modelo intercepto aleatorio covariables nivel 2.
- 4. Modelo intercepto y pendientes aleatorias con covariables selecionadas.
- 5. Añadir interacciones inter-nivel.

Si no detectas un efecto aleatorio relevante:

sesgos y problemas error tipo I

Workflow de modelos mixtos

- 1. Modelo nulo de intercepto aleatorio → ICC
- 2. Modelo más complejo posible, incluyendo interacciones inter-nivel y efectos aleatorios.
- 3. Googlear los errores de convergencia.
- 4. Eliminar efectos no relevantes.

Si la estructura es muy compleja:

Problemas de potencia y convergencia

Workflow de modelos mixtos

- 1. Modelo nulo de intercepto aleatorio → ICC
- 2. Modelo más complejo posible, incluyendo interacciones inter-nivel y efectos aleatorios.
- 3. Googlear los errores de convergencia.
- 4. Eliminar efectos no relevantes.

Es

MUY

PERO QUE MUY

POCO PRACTICO

Y cuando acabe este workshop,

¿qué hago en la vida real?

Tener teoría previa para seleccionar efectos fijos e interacciones de interés

Mantener la parsimonia e interpretabilidad.

Menos es más

Evaluar el modelo desde una perspectiva maximalista

¡Tiempo de trabajar!

Ejemplo 3

- ¿Cómo afecta al modelo si probamos añadir pendientes aleatorias?
- ¿Y si temenos efectos cruzados entre niveles?
- ¿Qué interpretación realizamos del ICC?
- ¿Y si hubiéramos seguido una estrategia maximalista…llegaríamos al mismo lugar?

¿Y si tengo una variable dicotómica?

Extensión del modelo multinivel

- La extensión a modelos lineales generalizados parece tan simple como en otro modelo de regresión
- Sin embargo, existen varios problemas en el uso de estos modelos: ICC, estimación de varianzas, potencia estadística, convergencia…
- Clave: estudiar cada tipo de modelo generalizado por separado.

Interpretación de los efectos

- Estimación original se realiza sobre los log-odds de la variable dicotómica.
- Intepretar **odds ratio [exp(log-odds)]:** ¿cuánto aumenta la probabilidad condicional de que ocurra X si aumenta una unidad del predictor?

Ejemplo

- In our example, $OR = exp(B_1) = exp(1.50) \approx 4.5$, indicates that the odds of owing Justin's album (instead of not owning it) are 4.5:1, that is, *multiplied* by 4.5/1 = 4.5 when GPA increases by one unit. Simply put, pupils are 4.5 times *more* likely to own the album when GPA increases by one unit (a 350% *increase*).
- Now imagine that the sign of B_1 is negative, that is, $B_1 = -1.50$. In such a case, $OR = exp(B_1) = exp(-1.50) \approx 0.22$ indicates that the odds of owning Justin's album (instead of not owning it) are 1:0.22, that is, *divided* by $1/0.22 \approx 4.5$ when GPA increases by one unit. Simply put, pupils are 4.5 times *less* likely to own the album when GPA increases by one unit (a 350% *decrease*).

https://www.rips-irsp.com/articles/10.5334/irsp.90/

¿Cómo calculamos el ICC?

- $Logit(odds) = (b_{intercepto} + effestud_0) + (b_{tutoría} + effestud_1) * tutoría$

- No necesitamos residuos individuales → intentamos predecir una probabilidad (e.g., 1/0)

- La varianza en el primer nivel se aproxima comúnmente como $\pi^2/3$= 3,29.

¿Cómo calculamos el ICC?

$$ICC_{adj} = \frac{\boldsymbol{\tau}_{RE}^2}{\boldsymbol{\tau}_{RE}^2 + \boldsymbol{\tau}_{FE}^2 + 3.29}$$

- 3.29 = varianza variable dicotómica en el primer nivel.
- Este ICC representa una aproximación → comúnmente se interpreta como el de un modelo lineal general.

Otros dos puntos claves

- Un modelo multinivel generalizado tiene una alta probabilidad de no converger.
- Normalmente, nos tocará ir probando estimadores y puntos de cuadratura.
- Fuerza examiner asuntos como la sobredispersión, adecuación de la función de enlace, etc..
- Paciencia y ánimo.

Es el momento de conocer a vuestro nuevo mejor amigo.

Thank you for your attention

eduardo.garciag@uam.es

@edu_gargar

edugargar

Design by Dr. Ruggeri (Columbia University)