Introducción a los modelos mixtos

Eduardo Garcia-Garzon

IA Specialist

AEMCCO 2022

TERUEL

SHAKERS

Introducción

1. Evaluar efectos aleatorios
2. Límites de los modelos mixtos
3. Elementos claves a la hora de ajustar un modelo mixto

Key messages

1. Qué es el ICC
2. Cuándo falla un MLM.
3. 5 aspectos a considerar

¿Necesitamos ese efecto aleatorio?

Justificación efecto aleatorio

- A menudo, tenemos que justificar la definición de un efecto aleatorio.
- Justificación teórica → no suele ser suficiente.
- Método más común:
  - **El Coeficiente de Correlación Intraclase (ICC)**

Justificación efecto aleatorio

- A menudo, tenemos que justificar la definición de un efecto aleatorio.
- Justificación teórica → no suele ser suficiente.
- Método más común:
  - **El Coeficiente de Correlación Intraclase (ICC)**

ICC

$$ICC = \frac{\boldsymbol{\tau}^2}{\boldsymbol{\tau}^2 + \sigma^2}$$

- Proporción de varianza explicada por el efecto aleatorio ****
- ICC condicional vs ICC marginal

Cómo utilizar el ICC

- El ICC se reporta para cada modelo que calculamos.
- Se suele reportar el ICC ajustado o marginal:

el % de varianza explicado por la variable de agrupación teniendo en cuenta los posibles efectos fijos

$$ICC_{adj} = \frac{\tau_{RE}^2}{\tau_{RE}^2 + \tau_{FE}^2 + \sigma^2}$$

Cómo utilizar el ICC

- El ICC se reporta para cada modelo que calculamos.
- Existe el ICC condicional :
  - el % de varianza explicado todos los efectos (fijos + aleatorios) del modelo

$$ICC_{cond} = \frac{\tau_{RE}^2 + \tau_{FE}^2}{\tau_{RE}^2 + \tau_{FE}^2 + \sigma^2}$$

Cálculo ICC

- El primer cálculo del ICC se realiza con el llamado **el modelo nulo.**
- El modelo nulo es un modelo de intercepto aleatorio que no incluye ningún efecto fijo.
- Se utiliza para:
    - **Calcular el ICC → justificar inclusion efectos aleatorios**
    - Servir de baseline para la comparación con modelos más complejos

Miniejercicios

- El primer cálculo del ICC se realiza con el llamado **el modelo nulo.**
- Calcular el ICC para los datos del ejemplo 2 utilizando sujeto y trial como efectos aleatorios.
- ¿En qué caso observamos una mayor influcencia de las variables de segundo nivel?

Random Components

Groups	Name	SD	Variance	ICC
id	(Intercept)	0.978	0.956	0.474
Residual		1.030	1.062	

Nota. Number of Obs: 700 , groups: id 100

La variabilidad entre participantes explica un 47% de la varianza total

Model Info

Info	
Estimate	Linear mixed model fit by REML
Call	relation ~ 1 +(1 | id)
AIC	2235.44
BIC	2251.76
LogLikel.	–1116.06
R-squared Marginal	0.00
R-squared Conditional	0.47
Converged	yes
Optimizer	bobyqa

[3]

La variabilidad entre participantes explica un 24% de la varianza total

Random Effect LRT

Test	N. par	AIC	LRT	df	p
(1 | id)	2.00	2483.60	247.49	1.00	$<.001$

- Los likelihood ratio test bajo REML son conservadores.
- En aplicación puede utilizarse una alternativa bootstrap.

Entonces, ¿todo se basa en el ICC?

Ajuste del modelo

- La evaluación de efectos fijos y aleatorios suele contemplarse desde dos perspectivas:
    - Efectos aleatorios: ICC / Ajuste modelo / LRT.
    - Efectos fijos: Ajuste modelo /LRT
- Clave cuando decidamos la complejidad del modelo

Model Info

Info	
Estimate	Linear mixed model fit by REML
Call	relation_neg ~ 1 + leaving+(1 | id)
~~AIC~~	~~1718.00~~
BIC	1744.82
LogLikel.	–859.31
R-squared Marginal	0.02
R-squared Conditional	0.50
Converged	yes
Optimizer	bobyqa

[3]

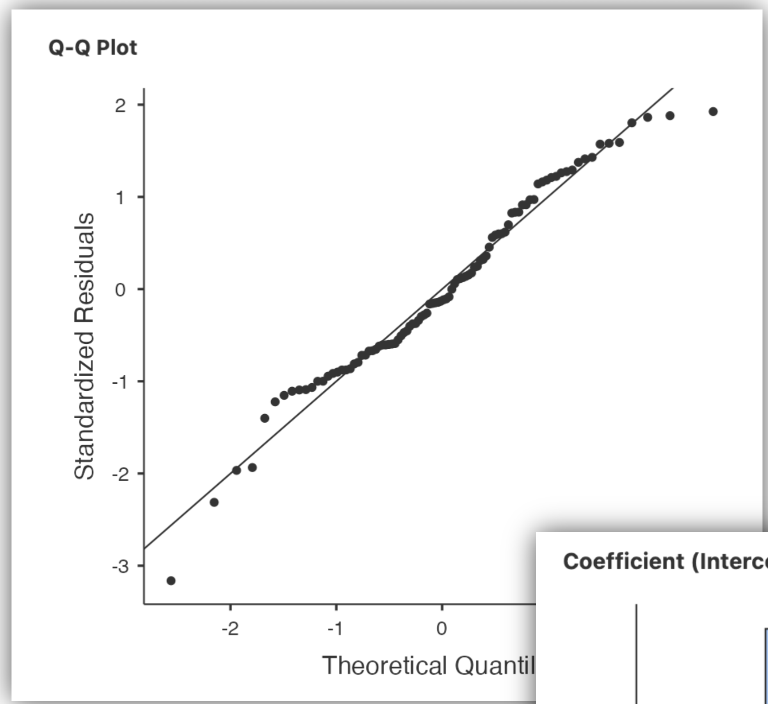
Supuestos estadísticos

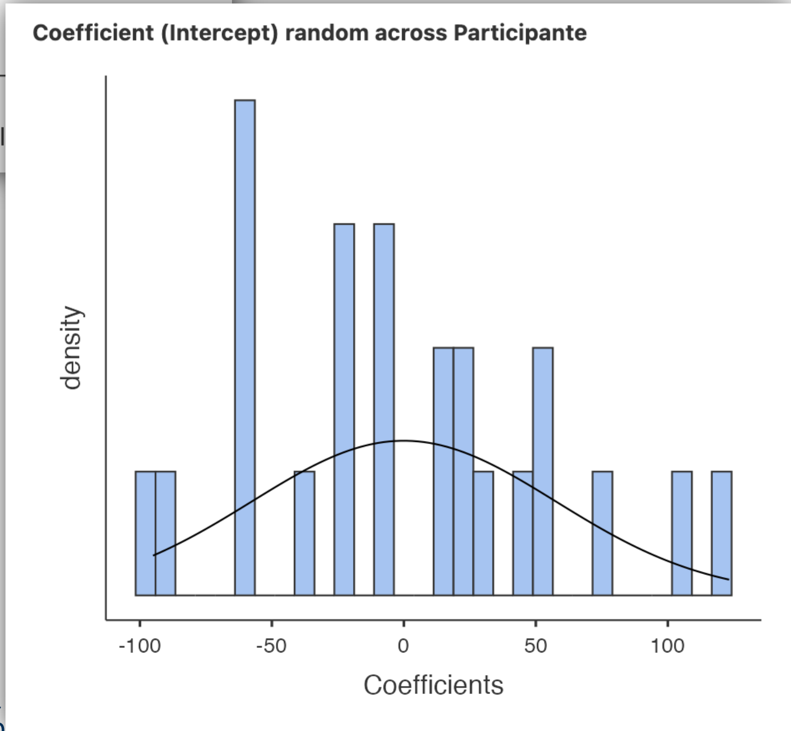
Supuestos

- Modelos mixtos conllevan una expansion de los supuestos tradicionales del modelo de regresión.
- Los más importantes no son rutinariamente evaluados.
- Afectar de manera importante a la estimación de los efectos fijos y aleatorios.
- **¡El modelo mixto no siempre es adecuado!**

Assumptions, properties, and advantages of HLM. When modeling clustered data with HLM, 10 assumptions are made:

1. All relevant predictors are included in the model.
2. All relevant random effects are included in the model.
3. The covariance structure of the within-cluster residuals, **R**, is properly specified (when the outcome is continuous).
4. The covariance structure of the random effects, **G**, is properly specified (for all outcomes scales).
5. The within-cluster residuals and the random effects do not covary [$Cov(\mathbf{u}, \boldsymbol{\varepsilon}) = \mathbf{0}$].
6. The within-cluster residuals follow a multivariate normal distribution (when the outcome is continuous).
7. The random effects follow a multivariate normal distribution (for all outcome scales).
8. The predictor variables do not covary with the residuals/random effects at any other level [$Cov(\mathbf{X}, \boldsymbol{\varepsilon}) = \mathbf{0}$, $Cov(\mathbf{X}, \mathbf{u}) = \mathbf{0}$].
9. Sample size is sufficiently large for asymptotic inference at each level.
10. With or without preprocessing, missing data are assumed to be missing completely at random (MCAR) or missing at random (MAR).





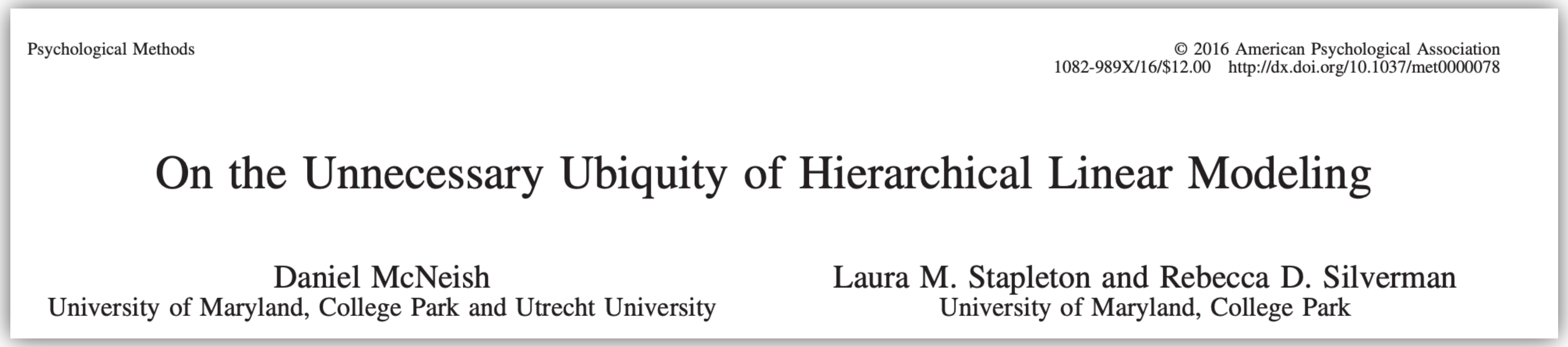
https://psycnet.apa.org/doiLanding?doi=10.1037%

Assumptions, properties, and advantages of HLM. When modeling clustered data with HLM, 10 assumptions are made:

1. All relevant predictors are included in the model.
2. All relevant random effects are included in the model.
3. The covariance structure of the within-cluster residuals, **R**, is properly specified (when the outcome is continuous).
4. The covariance structure of the random effects, **G**, is properly specified (for all outcomes scales).
5. The within-cluster residuals and the random effects do not covary [$Cov(\mathbf{u}, \boldsymbol{\varepsilon}) = \mathbf{0}$].
6. The within-cluster residuals follow a multivariate normal distribution (when the outcome is continuous).
7. The random effects follow a multivariate normal distribution (for all outcome scales).
8. The predictor variables do not covary with the residuals/random effects at any other level [$Cov(\mathbf{X}, \boldsymbol{\varepsilon}) = \mathbf{0}$, $Cov(\mathbf{X}, \mathbf{u}) = \mathbf{0}$].
9. Sample size is sufficiently large for asymptotic inference at each level.
10. With or without preprocessing, missing data are assumed to be missing completely at random (MCAR) or missing at random (MAR).

- Riesgo de un efecto de sesgo de heterogeneidad.
- No es fácil de detectar.
- Solución: moverse a un modelo efectos fijos.

https://psycnet.apa.org/doiLanding?doi=10.1037%2Fmet0000078

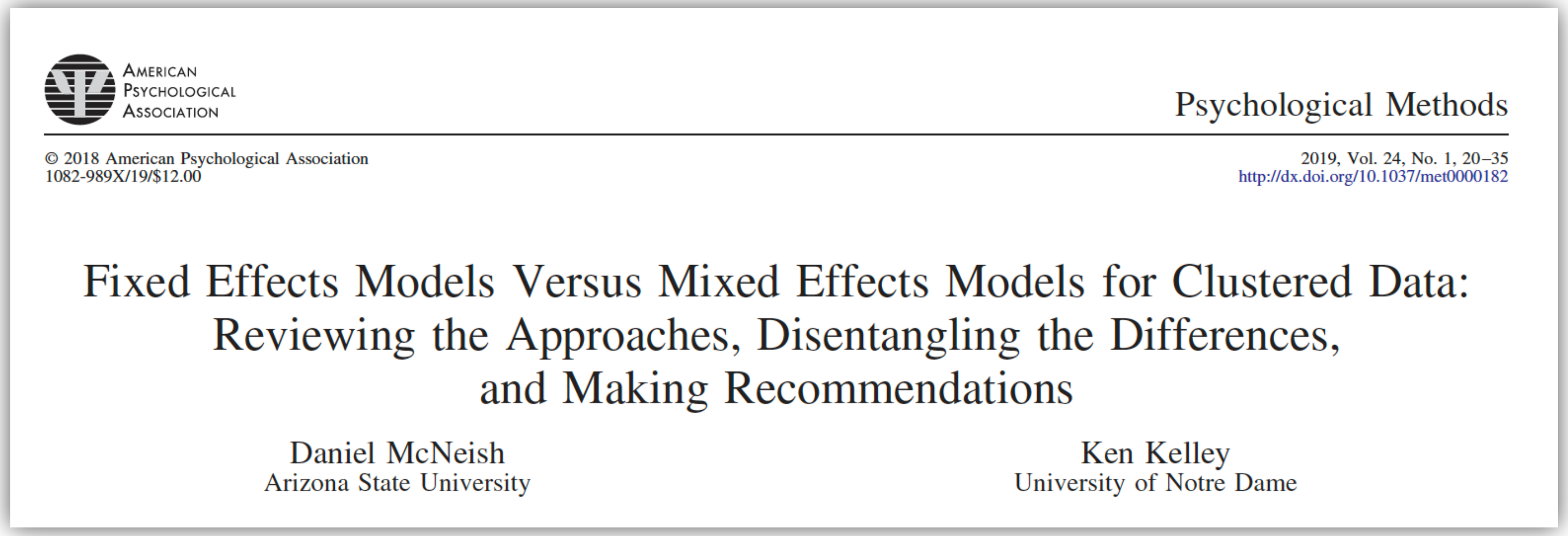

Psychological Methods

© 2016 American Psychological Association
1082-989X/16/$12.00 http://dx.doi.org/10.1037/met0000078

On the Unnecessary Ubiquity of Hierarchical Linear Modeling

Daniel McNeish
University of Maryland, College Park and Utrecht University

Laura M. Stapleton and Rebecca D. Silverman
University of Maryland, College Park


American Psychological Association

Psychological Methods

© 2018 American Psychological Association
1082-989X/19/$12.00

2019, Vol. 24, No. 1, 20–35
http://dx.doi.org/10.1037/met0000182

Fixed Effects Models Versus Mixed Effects Models for Clustered Data: Reviewing the Approaches, Disentangling the Differences, and Making Recommendations

Daniel McNeish
Arizona State University

Ken Kelley
University of Notre Dame

Table 2

Comparison of the Types of Modeling Questions That Can Be Assessed With MEMs and FEMs

Modeling problem	MEM	FEM
Accommodation of clustering	Random effects must be explicitly modeled by the user. The covariance matrix of the random effects also must be explicitly modeled.	Cluster affiliation dummy variables are included directly in the model.
Common estimation method	(Restricted) Maximum likelihood	Ordinary least squares
Predictors at Level 2	Allowed and coefficients are directly estimated. Proper specification is required, meaning that no relevant variables are omitted.	Generally not estimable (although there are proposed methods that claim to be able to provide estimates under particular circumstances). Omitted Level 2 variable bias is not a concern.
Omitted variable bias	A concern at all levels	Only a concern at Level 1
Accommodation of variability at Level 2	Predictor variables and random effects	Cluster affiliation dummy variables
Coefficient interpretation	Coefficients at either level are interpreted conditional on the variables explicitly included in the model.	Level 1 coefficients are conditional on all Level 2 variables (measured and unmeasured) being accounted for.
Level 2 sample size requirement	30 is the general recommendation; can be reduced (to about 10) if corrective	Viable with very small Level 2 sample sizes

Assumptions, properties, and advantages of HLM. When modeling clustered data with HLM, 10 assumptions are made:

1. All relevant predictors are included in the model.
2. All relevant random effects are included in the model.
3. The covariance structure of the within-cluster residuals, $\mathbf{R}$, is properly specified (when the outcome is continuous).
4. The covariance structure of the random effects, $\mathbf{G}$, is properly specified (for all outcomes scales).
5. The within-cluster residuals and the random effects do not covary [$Cov(\mathbf{u}, \boldsymbol{\varepsilon}) = \mathbf{0}$].
6. The within-cluster residuals follow a multivariate normal distribution (when the outcome is continuous).
7. The random effects follow a multivariate normal distribution (for all outcome scales).
8. The predictor variables do not covary with the residuals/random effects at any other level [$Cov(\mathbf{X}, \boldsymbol{\varepsilon}) = \mathbf{0}$, $Cov(\mathbf{X}, \mathbf{u}) = \mathbf{0}$].
9. Sample size is sufficiently large for asymptotic inference at each level.
10. With or without preprocessing, missing data are assumed to be missing completely at random (MCAR) or missing at random (MAR).

- Modelo maximalista → no dejarse fuera ningún efecto fijo.
- Importante explorar todos los efectos posibles de interceptos/pendientes aleatorias.
- Cuidado: mayor complejdidad, mayor sobreajuste, peor convergencia

https://psycnet.apa.org/doiLanding?doi=10.1037%2Fmet0000078

Assumptions, properties, and advantages of HLM. When modeling clustered data with HLM, 10 assumptions are made:

1. All relevant predictors are included in the model.
2. All relevant random effects are included in the model.
3. The covariance structure of the within-cluster residuals, **R**, is properly specified (when the outcome is continuous).
4. The covariance structure of the random effects, **G**, is properly specified (for all outcomes scales).
5. The within-cluster residuals and the random effects do not covary [$Cov(\mathbf{u}, \boldsymbol{\varepsilon}) = \mathbf{0}$].
6. The within-cluster residuals follow a multivariate normal distribution (when the outcome is continuous).
7. The random effects follow a multivariate normal distribution (for all outcome scales).
8. The predictor variables do not covary with the residuals/random effects at any other level [$Cov(\mathbf{X}, \boldsymbol{\varepsilon}) = \mathbf{0}$, $Cov(\mathbf{X}, \mathbf{u}) = \mathbf{0}$].
9. Sample size is sufficiently large for asymptotic inference at each level.
10. With or without preprocessing, missing data are assumed to be missing completely at random (MCAR) or missing at random (MAR).

Tamaño muestral

- Tenemos tantos tamaños muestrales relevantes como niveles.
- La N total suele ser de menor importancia que la N de los componentes.
- Cada tamaño muestral afecta de manera diferente a la estimación de sus componentes y la corr. de estos.

https://psycnet.apa.org/doiLanding?doi=10.1037%2Fmet0000078

Tamaño muestral

Multilevel Modelling of Country Effects: A Cautionary Tale

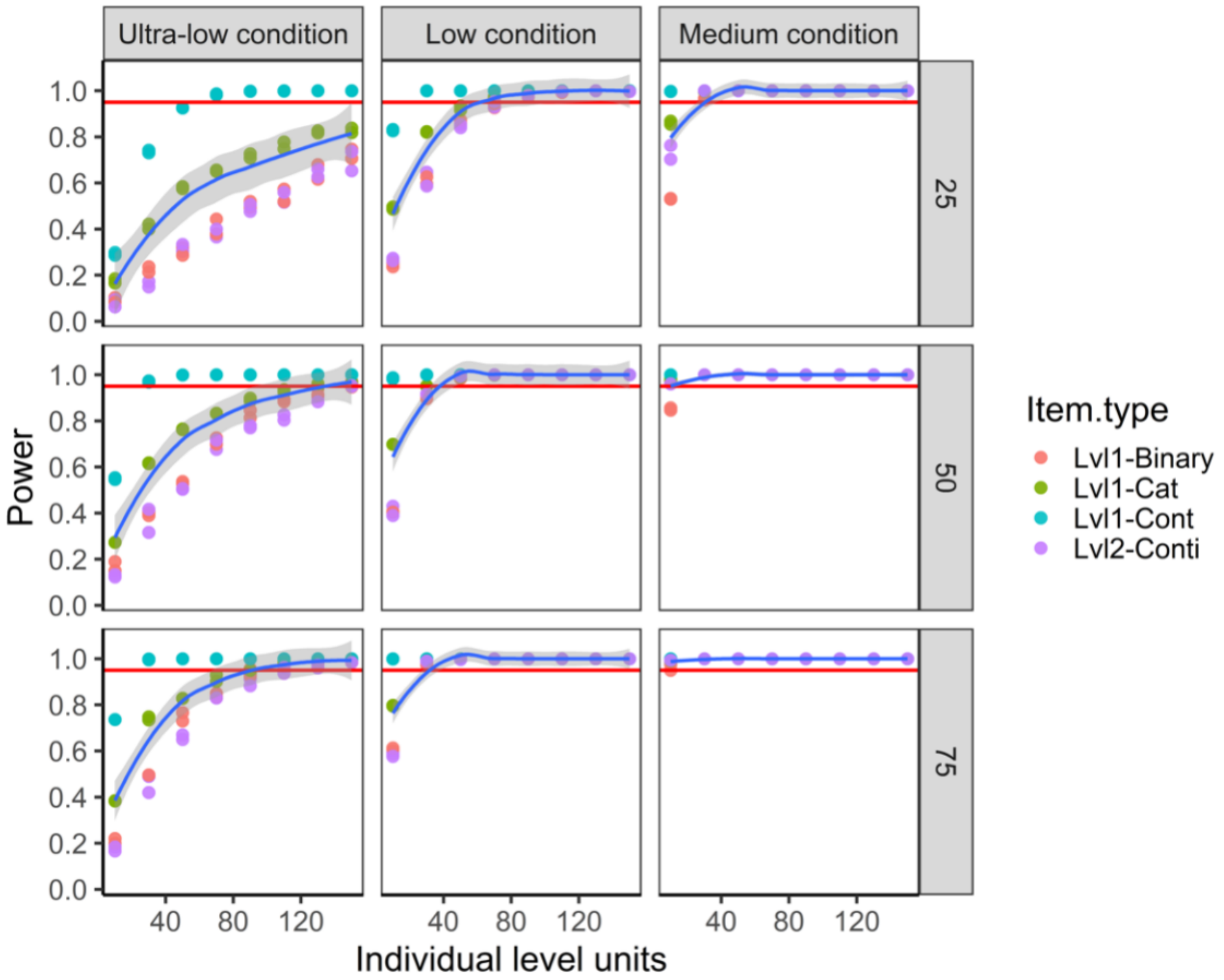
Mark L. Bryan[1,*] **and Stephen P. Jenkins**[1,2,3,*]

[1]Institute for Social and Economic Research, University of Essex, Colchester CO4 3SQ, UK, [2]Department of Social Policy, London School of Economics, London WC2A 2AE, UK and [3]IZA, Bonn 53113, Germany

*Corresponding author. Email: s.jenkins@lse.ac.uk

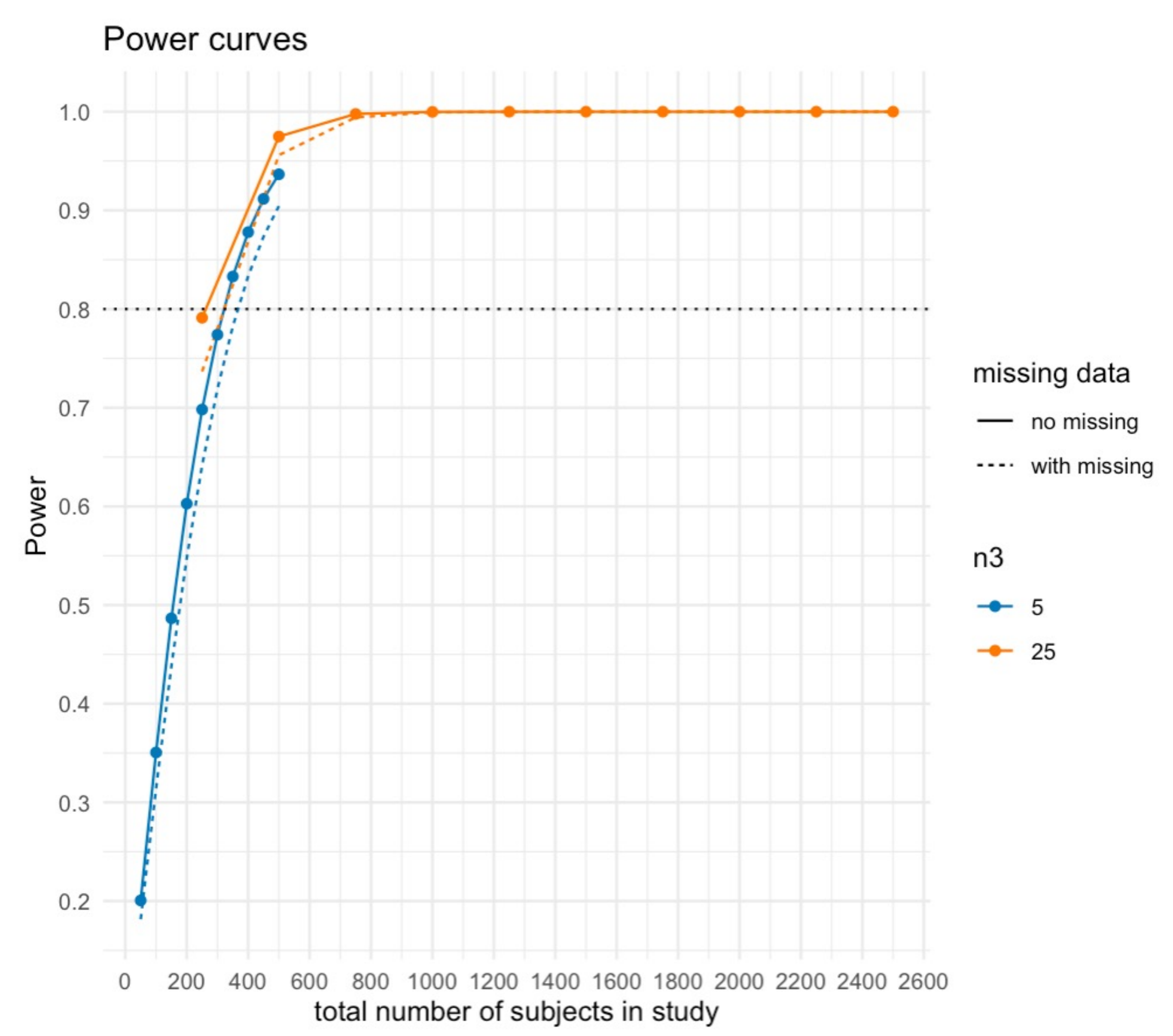
els. With large sample sizes of individuals within each country but only a small number of countries, analysts can reliably estimate individual-level effects but estimates of parameters summarizing country effects are likely to be unreliable. Multilevel modelling methods are no panacea.

Tamaño muestral



Tamaño muestral

- Pasamos de 5 a 50 sujetos en en incrementos de 5.
- Calculamos potencia para 5 y 25 unidades segundo nivel.
- Sin y 20% datos perdidos
- Paquete: powerlmm



Perils and pitfalls of mixed-effects regression models in biology

Matthew J. Silk[1,2], Xavier A. Harrison[1] and David J. Hodgson[1]

[1] Centre for Ecology and Conservation, University of Exeter, Penryn, Cornwall, UK
[2] Environment and Sustainability Institute, University of Exeter, Penryn, Cornwall, UK

Table 1 The perils of mixed modelling highlighted in this paper together with their potential consequences and solutions to avoid them.

Peril	Example	Consequences	Potential solutions
#1. Anticonservative significance tests at low sample size	Comparing crop yields in split-plot experiments with few replicates	*p*-value of Wald-like Chi-square test of significance is too low, causing high rates of Type I error	If better replication is not possible, use corrections for small sample size and accept that answers are approximations Move to statements of credibility based on Bayesian analyses
#2. Pseudoreplicated with group-level predictors	Infer the effect of maternal traits on the performance of several offspring per mother	Increased Type I error if pseudoreplication is not recognised Increased Type I error if true replication is small even if pseudoreplication is recognised	Have a firm grasp of the design level at which true replication occurs, and of the correct mixed-model specification If better replication of experimental units is not possible, use corrections for small sample size Move to statements of credibility based on Bayesian analyses with correct specification of design hierarchy
#3. Too few levels of a random effect	Fitting sex as a random effect	Model degeneracy Biased estimation of random effect variance Inaccurate estimation of random effect variance High error potential for questions related to random effects (less substantial for questions related to fixed effects)	Fit the variable as a fixed effect rather than random effect Use a Bayesian model with strong priors for the size of the random effect variance component. Such an approach requires caution and an ability to justify the inclusion of prior knowledge

#5a. Confounding by cluster	Multiple observation of foraging behaviour are made at a succession of sites to test the hypothesis that foraging rates are associated with disturbance levels. However, sites have different mean levels of disturbance. Both disturbance (fixed effect) and site (random effect) are used as explanatory variables	Biased estimates of fixed and random effect parameters	Use within-group mean centring of variables alongside a group-level covariate
#5b. Informative cluster sizes	A model with offspring weight as a response and maternal pathogen load as an explanatory variable and maternal ID as a random effect, if high pathogen loads also cause reduced litter sizes	Possibility of biased estimates for fixed effect parameters if they are correlated with the random effect especially if the model only includes random intercepts	Fit cluster size as a covariate where appropriate (see main text) Joint modelling of the response variable of interest and cluster size in a multivariate model
#6. Group means are not normally distributed	An unmeasured variable causes differences between sub-populations Skewed differences in a trait between sub-populations that are unexplained by fixed effects	Fixed effect estimates robust unless (a) mis-specification of random effects is extreme or (b) fixed effects are correlated with random effects Random effect estimates can become less accurate and systematically biased	Use non-Gaussian random effect distributions (challenging, but made available by use of Bayesian models) Fit the variable as a fixed effect instead. Take extra care to check for other violations of mixed model assumptions
#7. Use of categorical random effects for autocorrelated data	Region is used as a random effect to control for spatial autocorrelation when modelling abundance of a species in response to a range of habitat variables	Poorly fitting model (with inaccurate predictions) and potential of statistical errors unless scale of random effect is correct	Use correlograms to check for covariance in the residuals of the model to ensure that categorical random effect is effective

Claves estadísticas

Key 1: La estimación del modelo

- Librerías estándar utilizan algún algoritmo de optimización basado en REML.
- REML > ML; aunque puede no converger en *bastantes ocasiones.*
- Problemas de optmización con pocos niveles, falta de información, estructuras complejas, modelos no-lineales, etc.
- Querido Bayes…

Key 2: Grados de libertad

- Existen dos alternativas para estimar los grados de libertad:
    - **Satterhwite**
    - **Kenward-Rodger**
- Ambas pueden aplicarse con REML.
- K-R es un poco más conservadora, aunque no tiene ventajas evidentes sobre Satterhwite, y tarda más en calcularse.
- Bajo muestras amplias, ambas aproximaciones deberían dar resultados similares.

Key 3: Valores p

- En Jamovi, podemos obtener una *de las posibles* aproximaciones a los valores p existentes.
- Positivo: nos "facilita" la interpretación de estos modelos.
- Cons: el cálculo sea incorrecto o inadecuado.
- Como se reflejó Douglas Bates (autor de lme4)…

that group. If those people feel that I am a heretic for even suggesting that a p-value provided by SAS could be other than absolute truth and that I should be made to suffer a slow, painful death by being burned at the stake for my heresy, then I suppose that we will be able to look forward to an exciting finale to the conference dinner at UseR!2006 next month. (Well, I won't be looking forward to such a finale but the rest of you can.)

The parameter estimates calculated by lmer are the maximum likelihood or the REML (residual maximum likelihood) estimates and they are not based on observed and expected mean squares or on error strata. And that's a good thing because lmer can handle unbalanced designs with multiple nested or fully crossed or partially crossed grouping factors for the random effects. This is important for analyzing data from large observational studies such as occur in psychometrics.

Most of the research on tests for the fixed-effects specification in a mixed model begin with the assumption that these statistics will have an F distribution with a known numerator degrees of freedom and the only purpose of the research is to decide how to obtain an approximate denominator degrees of freedom. I don't agree.

https://stat.ethz.ch/pipermail/r-help/2006-May/094765.html

Behav Res (2017) 49:1494–1502
DOI 10.3758/s13428-016-0809-y

CrossMark

Evaluating significance in linear mixed-effects models in R

Steven G. Luke[1]

of freedom, were also evaluated. The results of these simulations suggest that Type 1 error rates are closest to .05 when models are fitted using REML and *p*-values are derived using the Kenward-Roger or Satterthwaite approximations, as these approximations both produced acceptable Type 1 error rates even for smaller samples.

Key 4: Centrar variables

A RECURRING PLOT LINE: CENTERING IN MULTILEVEL MODELS

We now introduce one of the most salient plot lines in the history of MLMs—the use of centering for lower-level variables. In single-level models, centering refers only to the rescaling of a

Annual Review of Psychology

Catching Up on Multilevel Modeling

Lesa Hoffman[1] and Ryan W. Walters[2]

Key 4: Centrar variables

- En un modelo mixto, cobra especial importancia cómo **centramos** los predictores.
- **Clave: cambia la interpretación de nuestros resultados**.
- Centrar no es sencillo en modelos con efectos cruzados, con interacciones multinivel, y otros monstruos parecidos.
- Tema muy discutido y con directrices poco claras.

Media global

- Variables niveles superiores / variables nivel 1
- Diferencias reflejan variabilidad entre-grupos.
- P.ej., diferencias a nivel rasgo (diferencia media de los sujetos)
- No cambia el ajuste del modelo

Media del grupo

- Variables nivel 1.
- Diferencias reflejan variabilidad intra-grupos.
- P.ej., diferencias a nivel estado (diferencia de media del sujeto)
- Cambia el ajuste del modelo

¿Cómo elegimos el método de centrado?

- Depende de la pregunta de investigación.
- En nivel 1: Si centramos por la media global → analizamos variabilidad total (var intra + var entre).
- Si centramos por la media del clúster → analizamos variabilidad inter-cluster → **recomendado**
- Podemos añadir ambas variables, conocido como efecto individual + efecto contextual

¿Cómo elegimos el método de centrado?

- Centrar por cluster mejora la estimación de interacciones inter-nivel y estimación de pendientes aleatorias.
- Importante: podemos centrar variables categóricas (Enders & Tofighi, 2007): poco aplicado.
- De nuevo, cuidado con las interpretaciones de cada uno de los tipos de centrado.

https://doi.org/10.1037/1082-989X.12.2.121

Predecimos notas a través de números de ejercicios por clase

Media global

- El intercepto es la nota media para una clase con un un número promedio de ejercicios.
- El coef. regres. representa cómo cambia las notas si aumenta un ejercicio (respecto al promedio de ejercicios de todas las clases).

Media grupo

- El intercepto es la nota media para un número de ejercicios promedio.
- El coef. regres. representa cómo cambia las notas si aumenta un ejercicio con respecto al promedio de ejercicios de tu clase.

https://doi.org/10.1037/1082-989X.12.2.121

Centering categorical predictors in multilevel models: Best practices and interpretation

Haley E. Yaremych[1], Kristopher J. Preacher[1], & Donald Hedeker[2]

[1]Department of Psychology & Human Development, Vanderbilt University

[2]Department of Public Health Sciences, University of Chicago

This is the authors' accepted version of the article which is currently in press at *Psychological Methods* (acceptance date: 8/10/2021).

Key 5: Estandarizar variables

- Recomedable estandarizar las variables para facilitar la interpretación de los modelos.
- Mejora la convergencia
- **Gelman & Hill (2008): Escalar los predictores dividiéndolos por 2 desviaciones estándar →** Comparar el efecto variables continuas y variables en los datos.

¿Y aquí acaba todo?

Solo acabamos de empezar…

Thank you for your attention

eduardo.garciag@uam.es

@edu_gargar

edugargar

Design by Dr. Ruggeri (Columbia University)