

SELECTION OF TREATMENT EFFECTS IN A GENERALIZED LINEAR MIXED MODELS USING A BOOTSTRAP PROCEDURES

Fatoretto, M. B. ; Moral, R. A. ; Demétrio, C. G. B.

"Luiz de Queiroz" College of Agriculture
University of São Paulo - Brazil

VIII Encontro dos Alunos
Pós-graduação em Estatística e Experimentação Agronômica

Overview

1 Introduction

2 Experiment

3 Methodology

4 Results

5 Conclusions

6 Future works

7 References

Introduction

- Dose-response experiments are common in entomology and yield data;
- Usually overdispersed;

Statistical Objectives

- ① Compare different models to overdispersion;
- ② Propose a bootstrap interval to compare treatments

Practical Objectives

- ① Compare treatments in a Biological control dataset;

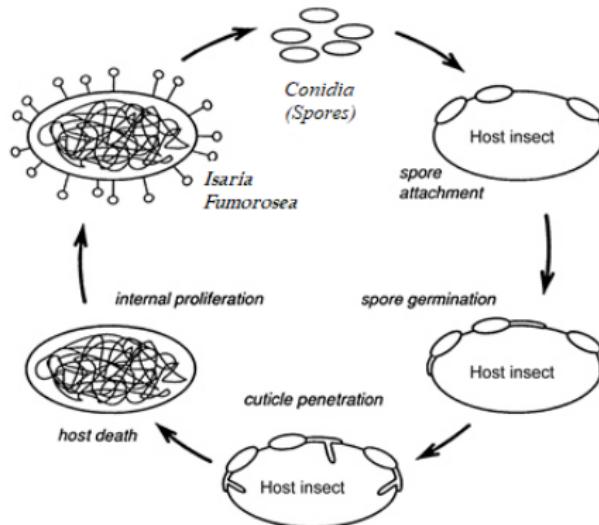
Motivation

- The fungus *Isaria Fumosorosea* is commonly found in soil and infecting several species of arthropods.
- Isolates may be collected from different places or insects.
- Fungi-based biopesticides have the capacity to infect a large number of pests and to remain in the environment.



Motivation

Biological mode of action

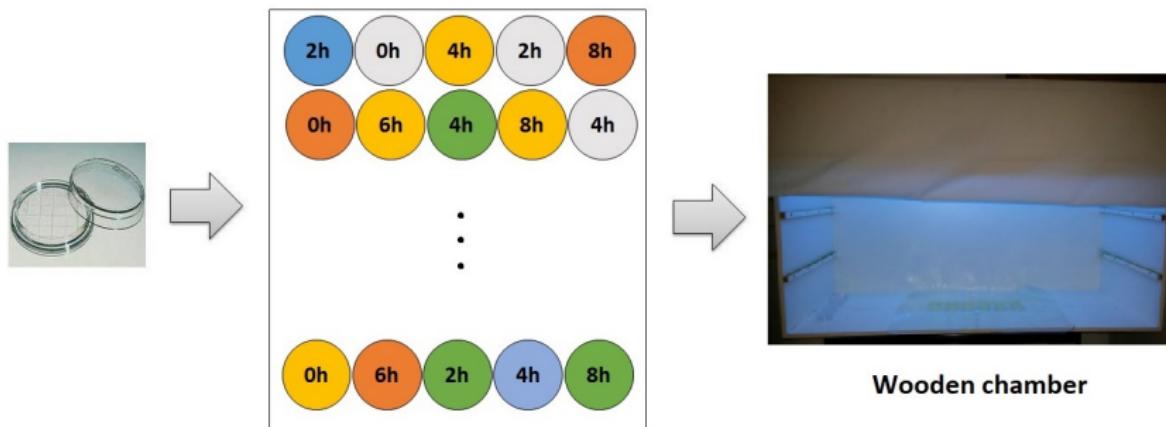


Problem: The ultraviolet radiation (UV-B) influences the efficacy and conidia germination of fungi in the field.

Experiment

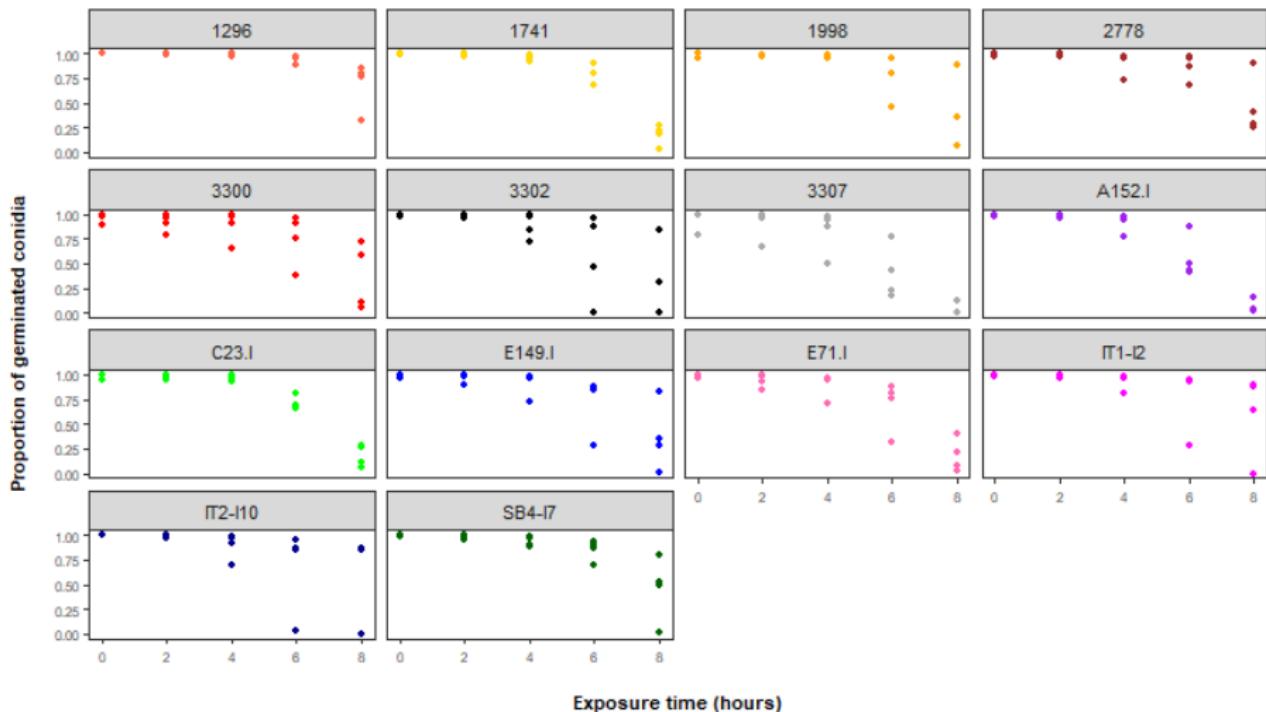
GOAL: To select isolates of *I. fumosorosea* with high UV-B radiation tolerance for developing of a new mycopesticide.

- Randomized block complete design with 4 blocks
- 70 Petri dishes: 14 isolates \times 5 exposure time (0, 2, 4, 6, 8 hours)

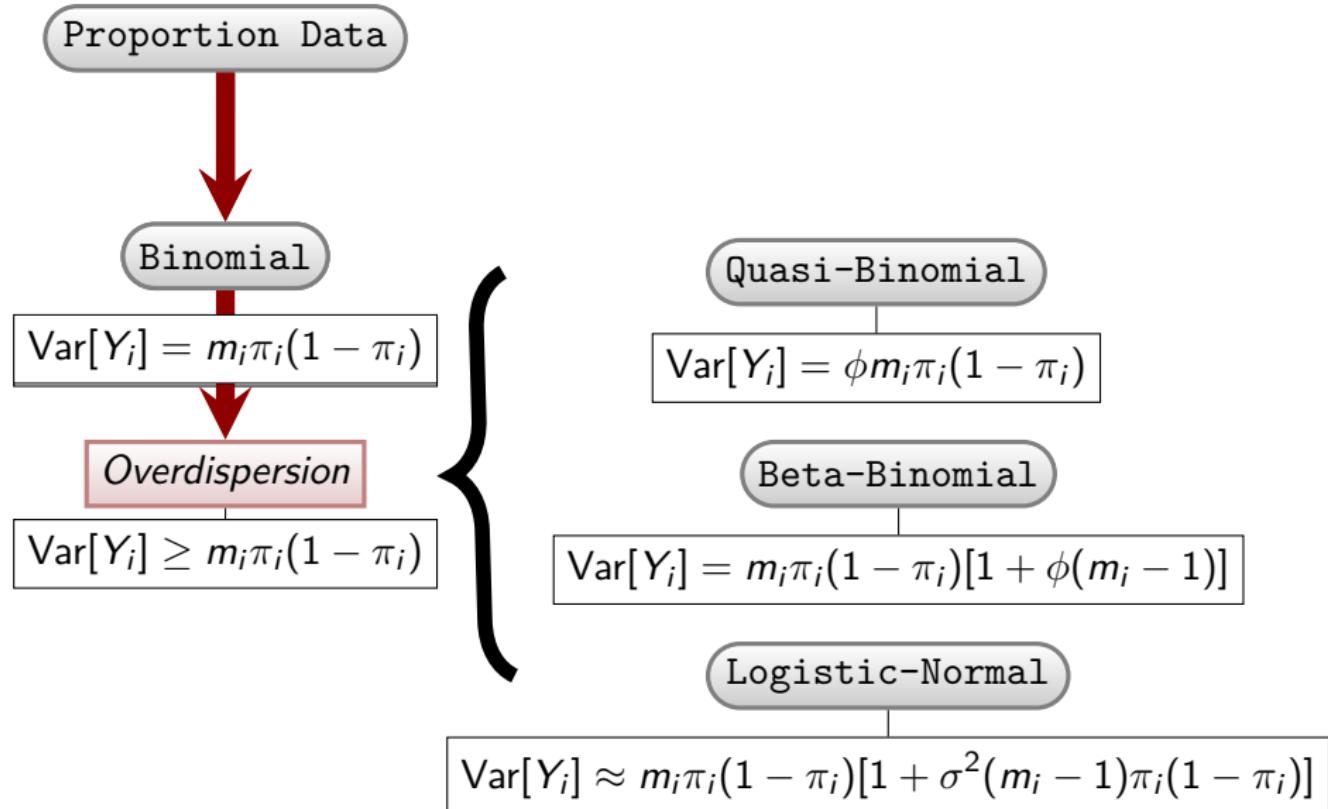


Outcome: proportion of germinated conidia in each Petri dish.

Observed data



Methodology



Model fitting

- **Logistic-normal:**

block ($i = 1, \dots, 4$) Isolate($j = 1, \dots, 14$) exposure time($k = 1, \dots, 5$)

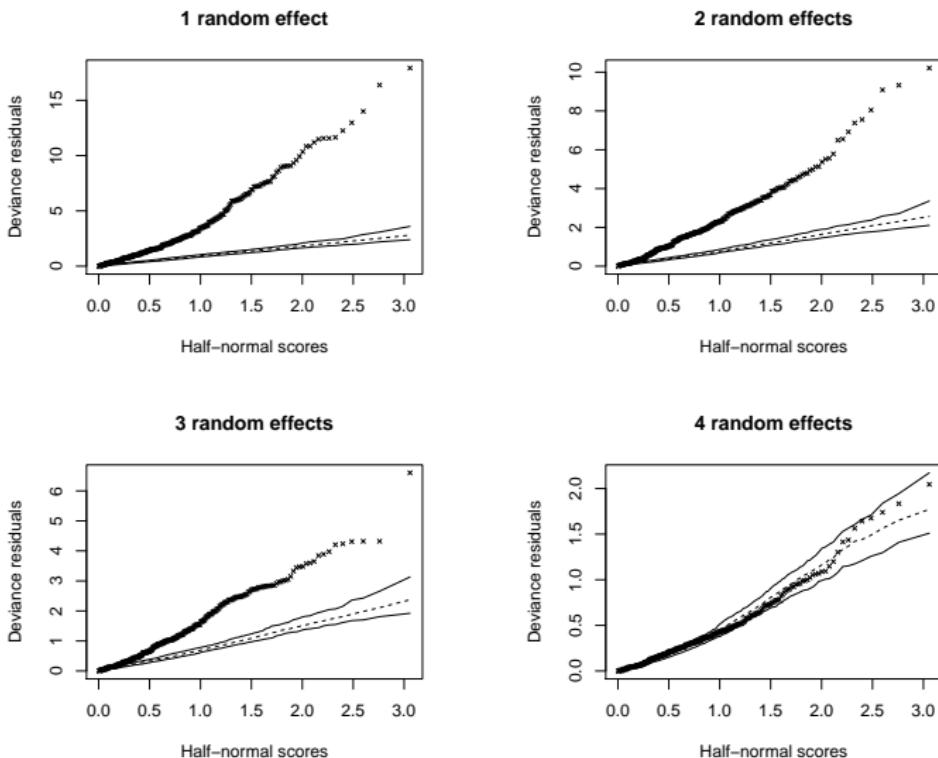
$$\eta_{ijk} = \beta_{0j} + \beta_{1j} t_k + r_i + b_{0ij} + b_{1ij} t_k + z_{ijk}$$

$$r_i \sim N(0, \sigma_B^2) \quad z_{ijk} \sim N(0, \sigma_O^2)$$

$$\begin{bmatrix} b_{0ij} \\ b_{1ij} \end{bmatrix} \sim N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma \right] \text{ with } \Sigma = \begin{bmatrix} \sigma_I^2 & \sigma_{IS} \\ \sigma_{IS} & \sigma_S^2 \end{bmatrix}$$

R packages

- `lme4` \Rightarrow `glmer` (bobyqa method);
- `hnp` \Rightarrow half-normal plots.



Model selection

1st step: Random effects

Likelihood ratio tests for the comparison of a series of generalized linear mixed models

Distribution of random effects	Covariance matrix	LR test ($D_q - D_p$)
$\begin{bmatrix} r_i \\ b_{0ij} \end{bmatrix} \sim N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma \right]$	$\Sigma = \begin{bmatrix} \sigma_B^2 & 0 \\ 0 & \sigma_I^2 \end{bmatrix}$	
$\begin{bmatrix} r_i \\ b_{0ij} \\ b_{1ij} \end{bmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Sigma \right]$	$\Sigma = \begin{bmatrix} \sigma_B^2 & 0 & 0 \\ 0 & \sigma_I^2 & \sigma_{IS} \\ 0 & \sigma_{IS} & \sigma_S^2 \end{bmatrix}$	925.70(<0.0001) Reference distribution: $\frac{1}{2}\chi_1^2 + \frac{1}{2}\chi_2^2$
$\begin{bmatrix} r_j \\ b_{0ij} \\ b_{1ij} \\ z_{ijk} \end{bmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \Sigma \right]$	$\Sigma = \begin{bmatrix} \sigma_B^2 & 0 & 0 & 0 \\ 0 & \sigma_I^2 & \sigma_{IS} & 0 \\ 0 & \sigma_{IS} & \sigma_S^2 & 0 \\ 0 & 0 & 0 & \sigma_O^2 \end{bmatrix}$	430.19(<0.0001) Reference distribution: $\frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2$

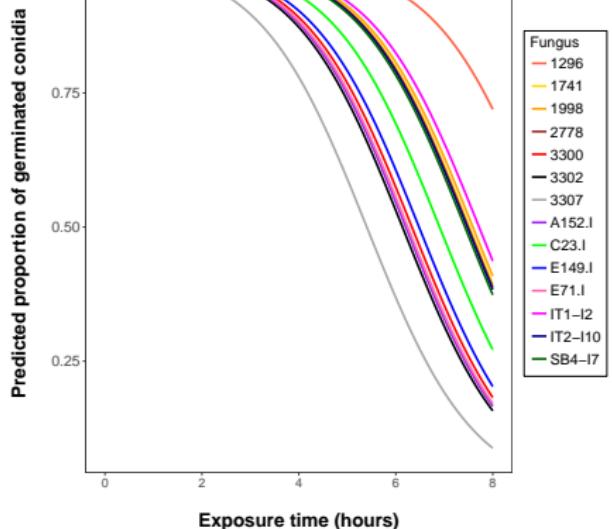
2st step: Fixed effects

Likelihood-ratio tests for the logistic-normal models with separate, parallel and coincident regression lines.

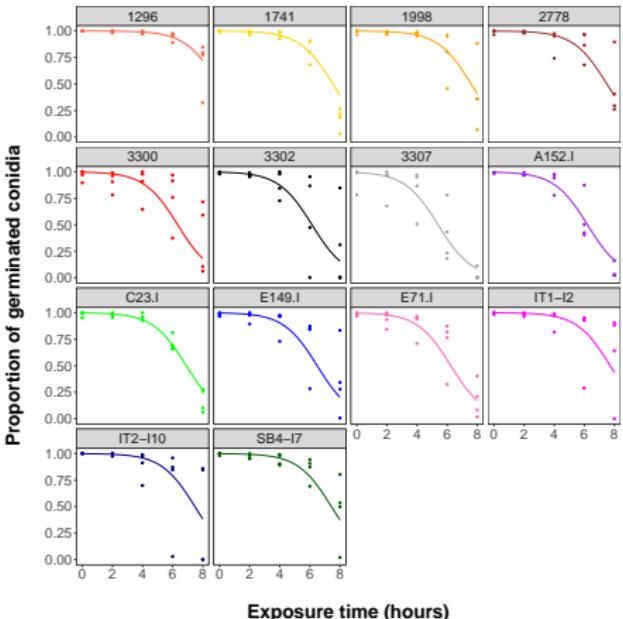
Test	χ^2	df	p-value
Separate \times parallel	12.57	13	0.4838
Parallel \times coincident	25.86	13	0.0177

Selected model: parallel lines

$$\eta_{ijk} = \beta_0 j + \beta_1 t_k + r_i + b_{0ij} + b_{1ij} t_k + z_{ijk}, \quad \begin{cases} i = 1, \dots, 4 \\ j = 1, \dots, 14 \\ k = 1, \dots, 5 \end{cases}$$



Fitted proportions using a parallel lines logistic-normal model.



Observed data and curves of fitted values

Next step: Confidence Intervals.

Nonparametric bootstrap

Resampling from \hat{F} (empirical distribution function). That is, sample from the data (with replacement):

$$(\mathbf{X}_i, \mathbf{Z}_i, \mathbf{y}_i)$$

for $b = 1, \dots, \beta$ (bootstrap sampling times)
Resampling ijk plots with equal probability
Fit the logistic normal model
Estimate parameters $\beta_{0i}^{(b)}, \beta_1^{(b)}, \sigma_B^{(b)}, \sigma_O^{(b)}, \sigma_I^{(b)}, \sigma_S^{(b)}$ and $\sigma_{JS}^{(b)}$

Multilevel nonparametric bootstrapping

for $b = 1, \dots, \beta$ (bootstrap sampling times)
Randomly i block with equal probability
Within each block sampled, sample jk plots with equal probability
Fit the logistic normal model
Estimate parameters $\beta_{0i}^{(b)}, \beta_1^{(b)}, \sigma_B^{(b)}, \sigma_O^{(b)}, \sigma_I^{(b)}, \sigma_S^{(b)}$ and $\sigma_{JS}^{(b)}$



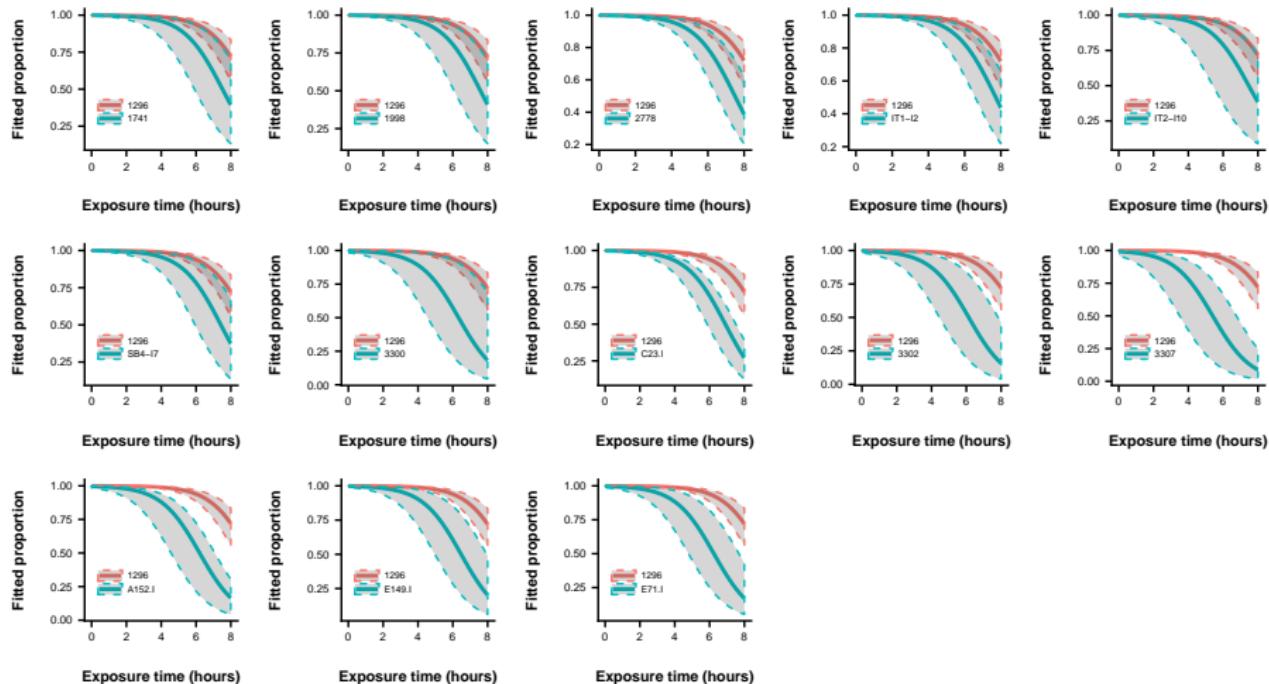
percentile confidence intervals

Compare nonparametric bootstrap intervals

Coverage rate of the 95% CI and HDI of parameters obtained by the Hierarchical and Paired Nonparametric bootstrap methods

Method	β_{01}	β_{02}	β_{03}	β_{04}	β_{05}	β_{06}	β_{07}	β_{08}	β_{09}	β_{010}	β_{011}	β_{012}
Hier. (HDI)	0.90	0.93	0.88	0.95	0.87	0.92	0.96	0.88	0.88	0.9	0.9	0.88
Hier. (CI)	0.94	0.91	0.88	0.94	0.89	0.87	0.91	0.89	0.87	0.89	0.84	0.87
Paired (HDI)	0.60	0.65	0.70	0.78	0.70	0.73	0.73	0.80	0.76	0.77	0.81	0.81
Paired (CI)	0.62	0.61	0.72	0.77	0.75	0.76	0.73	0.81	0.76	0.77	0.84	0.83
	β_{013}	β_{014}	β_1	σ_B	σ_O	σ_I	σ_S	σ_{IS}				
Hier. (HDI)	0.91	0.93	0.90	0.67	0.65	0.56	0.59	0.54				
Hier. (CI)	0.87	0.93	0.90	0.68	0.63	0.63	0.61	0.53				
Paired (HDI)	0.73	0.72	0.73	0.62	0.61	0.57	0.6	0.43				
Paired (CI)	0.74	0.73	0.76	0.6	0.65	0.56	0.59	0.4				

Similar isolates



HDI interval for the proportion of germinated conidia for each isolate

Conclusions

- The logistic-normal model with four random effects fits the data well.
- The use of a GLMM allowed us to model the correlation between observations of each block per isolate;
- Due to the inclusion of four random effects, some convergence problems may arise due to numerical integration problems, so it is important to take into consideration different approaches;
- The interval obtained by nonparametric bootstrap allowed us to compare the proportion of germinated conidia between isolate 1296 and the others.

Work in progress

$$\eta_{ijk} = \beta_{0j} + \beta_1 t_k + \gamma_i + b_{0ij} + b_{1ij} t_k + z_{ijk}$$

Method	γ_1	γ_2	γ_3	γ_4	β_{02}	β_{03}	β_{04}	β_{05}	β_{06}	β_{07}	β_{08}	β_{09}	β_{010}	β_{011}	β_{012}	β_{013}	β_{014}	β_1
$ML_{1000}(\text{CI})$	0.98	0.82	0.98	0.98	0.87	0.88	0.87	0.9	0.92	0.91	0.94	0.87	0.87	0.91	0.9	0.88	0.89	0.87
$ML_{1000}(\text{HDI})$	0.98	0.81	0.97	0.97	0.91	0.88	0.89	0.9	0.9	0.89	0.95	0.86	0.88	0.91	0.91	0.89	0.92	0.88
$ML_{2000}(\text{CI})$	0.98	0.74	0.98	0.98	0.95	0.94	0.92	0.91	0.87	0.9	0.89	0.94	0.91	0.87	0.94	0.93	0.92	0.85
$ML_{2000}(\text{HDI})$	0.98	0.8	0.98	0.98	0.93	0.91	0.92	0.92	0.87	0.9	0.89	0.95	0.93	0.88	0.94	0.9	0.92	0.87
$REML_{1000}(\text{CI})$	0.94	0.87	0.93	0.93	0.92	0.89	0.9	0.92	0.94	0.91	0.88	0.92	0.91	0.92	0.92	0.93	0.9	0.89
$REML_{1000}(\text{HDI})$	0.94	0.91	0.93	0.94	0.92	0.88	0.94	0.92	0.96	0.9	0.91	0.93	0.91	0.92	0.93	0.93	0.9	0.89
	σ_I	σ_S	σ_{IS}	σ_O														
$ML_{1000}(\text{CI})$	1	0.97	1	0.95														
$ML_{1000}(\text{HDI})$	0.98	0.96	0.99	0.95														
$ML_{2000}(\text{CI})$	0.98	0.94	0.99	0.98														
$ML_{2000}(\text{HDI})$	0.98	0.95	0.98	0.98														
$REML_{1000-}(\text{CI})$	0.95	0.95	0.98	0.93														
$REML_{1000-}(\text{HDI})$	0.95	0.95	0.99	0.95														

References

-  Davison, A. C. and Hinkley, D. V. (1997).
Bootstrap methods and their application, volume 1.
Cambridge university press.
-  Demétrio, C. G. B., Hinde, J., and de Andrade Moral, R. (2014).
Models for Overdispersed Data in Entomology, chapter 9, pages 219–259.
Springer, first edition.
-  Moral, R. A., Hinde, J., and Demétrio, C. G. (2017).
Half-normal plots and overdispersed models in r: The hnp package.
Journal of Statistical Software, 81(10):1–23.
-  Nelder, J. A. and Wedderburn, R. W. M. (1972).
Generalized linear models.
Journal of the Royal Statistical Society. Series A (General), 135(3):370–384.
-  Ren, S., Lai, H., Tong, W., Aminzadeh, M., Hou, X., and Lai, S. (2010).
Nonparametric bootstrapping for hierarchical data.
Journal of Applied Statistics, 37(9):1487–1498.
-  Ren, S., Yang, S., and Lai, S. (2006).
Intraclass correlation coefficients and bootstrap methods of hierarchical binary outcomes.
Statistics in medicine, 25(20):3576–3588.
-  Thai, H.-T., Mentré, F., Holford, N. H., Veyrat-Follet, C., and Comets, E. (2013).
A comparison of bootstrap approaches for estimating uncertainty of parameters in linear mixed-effects models.
Pharmaceutical statistics, 12(3):129–140.

Thank You!

mairafatoretto@gmail.com