Outlier detection in multicenter trials: a simulation study

Natalie Dennis Emmanuel Curis Lucie Biard Sylvie Chevret

October 6th, 2017

Centre de Recherche Epidémiologie et Statistique Sorbonne Paris Cité
Inserm UMR 1153

Introduction

- In a multicenter trial, in which a clinical trial is conducted at more than one research or medical centers, there exists a natural heterogeneity among observations
  - ⇒ This is due to the differences between the centers at which the clinical trial takes place
- How do we detect if there is an outlying center in a multicenter trial?

Objectives

- Determine if it is possible to detect an outlying center in a linear mixed-effects model
- If so, analyze the behavior of each method, including the Type I error and power

Linear Mixed-Effects Model

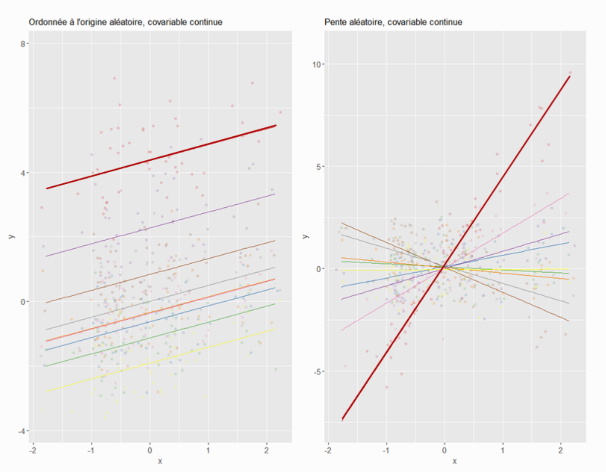
y_{ij} = response of i-th member of center j, $i = 1, ..., n_j$, $j = 1, ..., m$
m = number of centers
n_j = size of center j
x_{ij} = covariate vector of i-th member of center j for fixed effects
β = fixed effects parameter
u_{ij} = covariate vector of i-th member of center j for random effects
γ_j = random effect parameter

$$y_{ij} = x_{ij}^t \beta + u_{ij}^t \gamma_j + \epsilon_{ij}, \qquad i = 1, ..., n_j;\ j = 1, ..., m$$

Source: Czado, Claudia. “Linear Mixed Models.” *Technical University of Munich.*

https://www.statistics.ma.tum.de/en/people/professors/claudia-czado/research/.

Linear Mixed Model with Random Effect Outlier



Outlier Detection Methods

We will consider the following outlier detection methods to determine if there is a random-effect outlier:

- Boxplot extremities
  - Values below Q1−1.5× IQR or above Q3 +1.5×IQR
- Dixon's Q Test [1]
  - $\text{Q} = \frac{x_{(2)} - x_{(1)}}{x_{(n)} - x_{(1)}}$ and $\text{Q} = \frac{x_{(n)} - x_{(n-1)}}{x_{(n)} - x_{(1)}}$
- Grubbs' Test [2]
  - $G = \frac{\max |x_i - \bar{x}|}{s}$
- Likelihood Ratio Test, introduced by Langford & Lewis (1998) [4]

Method by Langford & Lewis

Consider the following linear mixed model with a random intercept and fixed slope:

$$y_{ij} = x_{ij}^t \beta + \gamma_j + \epsilon_{ij}$$

Langford & Lewis propose using the likelihood ratio test to compare this model to the following model:

$$y_{ij} = x_{ij}^t \beta + (1 - h)\gamma_j + h\alpha + \epsilon_{ij}$$

where $h = 1$ for the center we wish to test as an outlier and $h = 0$ otherwise.

Simulation Method

2200 simulations of the following models with 10, 30, 50 centers and 10, 20, 50 individuals per center (9 combinations):

1. One random intercept with a continuous covariate
2. One random intercept with a discrete covariate
3. One random slope with a continuous covariate
4. One random slope with a discrete covariate

Incorporating an outlying center

For each combination, we studied the following cases:

1. No outlying random effect
   - $\gamma_1, ..., \gamma_m \sim \mathcal{N}(0, 1)$
2. One center with an outlying fixed effect
   - $\gamma_1 = 3$, $\gamma_1 = 4$ and $\gamma_1 = 6$
3. One center with an outlying random effect
   - $\gamma_1 \sim \mathcal{N}(3, 1)$, $\gamma_1 \sim \mathcal{N}(4, 1)$, and $\gamma_1 \sim \mathcal{N}(6, 1)$

Bias of the fixed effect, $\hat{\beta}$

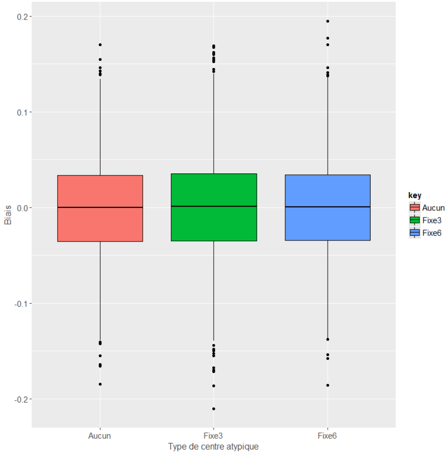


Figure: Bias of $\hat{\beta}$ for (1) no outlying random effect, (2) $\gamma_1 = 3$, and (3) $\gamma_1 = 6$

Type I errors

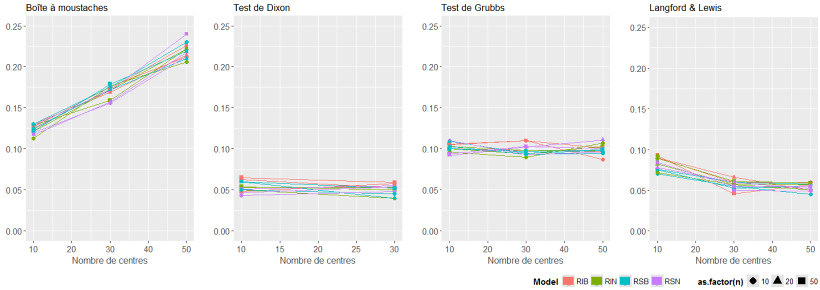


Figure: Type I errors for (1) Boxplot extremities, (2) Dixon's Q test, (3) Grubbs' test and (4) the Langford & Lewis method

Power of tests with $\gamma_1 = 6$ et $\gamma_1 \sim \mathcal{N}(6,1)$

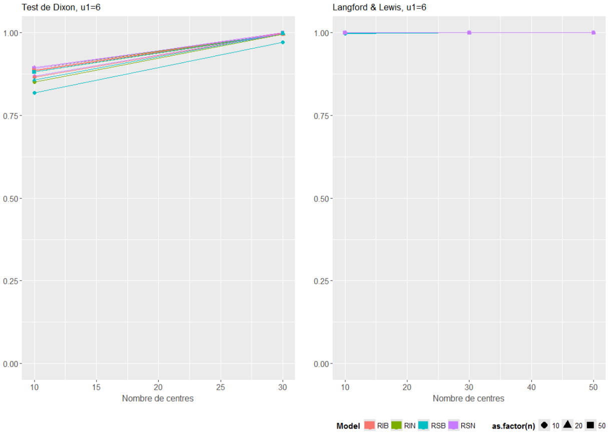


Figure: Power of (1) Dixon's Q test and (2) the Langford & Lewis method

Power of tests with $\gamma_1 = 3$ et $\gamma_1 \sim \mathcal{N}(3,1)$

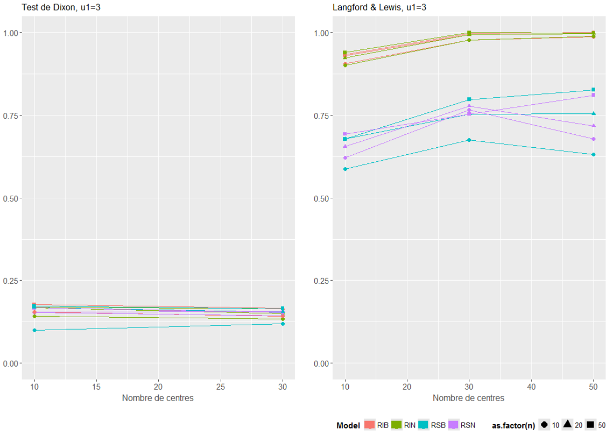


Figure: Power of (1) Dixon's Q test and (2) the Langford & Lewis method

Dixon's Q Test: Detecting the correct center

	$n = 10$	$n = 20$	$n = 50$
$\gamma_1 = 3$	0.875	0.921	0.929
$\gamma_1 = 4$	0.993	0.996	0.998
$\gamma_1 = 6$	1.000	1.000	1.000
$\gamma_1 \sim \mathcal{N}(3, 1)$	0.956	0.964	0.966
$\gamma_1 \sim \mathcal{N}(4, 1)$	0.991	0.994	0.987
$\gamma_1 \sim \mathcal{N}(6, 1)$	1.000	1.000	1.000

Table: Proportion of the correct center detected by Dixon's Q test when $p < 0.05$ and $m = 30$ centers

Conclusion

- The Grubbs' test, while more powerful than Dixon's Q Test, cannot be considered because of its elevated Type I error rate.
- There is further work to be done in developing Langford & Lewis's method:
  - Recoding to test the maximum and minimum random effects instead of the center of choice
  - Adjusting p-value to compensate for testing each random effect
- Testing these methods on more complex models

Thank you for your attention!

References

W.J Dixon.
Analysis of extreme values.
Ann. Math. Statist., 21(4):488–506, 1950.

Frank E. Grubbs.
Procedures for detecting outlying observations in samples.
Technometrics, 11(1):1–21, 1969.

Nan M. Laird and James H. Ware.
Random-effects models for longitudinal data.
Biometrics, 38(4):963–974, 1982.

Ian H. Langford and Toby Lewis.
Outliers in multilevel data.
Journal of the Royal Statistical Society: Series A, 161(2):121–160, 1998.