

corre

correlations with random effects

Marco Cattaneo

Department of Clinical Research
University of Basel

22 September 2022

background

- ▶ we often encounter correlated data points, e.g. in ophthalmology we may measure both eyes of a patient, and maybe several times

background

- ▶ we often encounter correlated data points, e.g. in ophthalmology we may measure both eyes of a patient, and maybe several times
- ▶ if we are interested in regression, we can use random effects to deal with the correlation among data points, e.g. `variable ~ (1 | patient / eye)`

background

- ▶ we often encounter correlated data points, e.g. in ophthalmology we may measure both eyes of a patient, and maybe several times
- ▶ if we are interested in regression, we can use random effects to deal with the correlation among data points, e.g. `variable ~ (1 | patient / eye)`
- ▶ disregarding the correlation among data points, e.g. `variable ~ 1`, usually leads to:

background

- ▶ we often encounter correlated data points, e.g. in ophthalmology we may measure both eyes of a patient, and maybe several times
- ▶ if we are interested in regression, we can use random effects to deal with the correlation among data points, e.g. `variable ~ (1 | patient / eye)`
- ▶ disregarding the correlation among data points, e.g. `variable ~ 1`, usually leads to:
 - ▶ slightly less efficient estimates

background

- ▶ we often encounter correlated data points, e.g. in ophthalmology we may measure both eyes of a patient, and maybe several times
- ▶ if we are interested in regression, we can use random effects to deal with the correlation among data points, e.g. `variable ~ (1 | patient / eye)`
- ▶ disregarding the correlation among data points, e.g. `variable ~ 1`, usually leads to:
 - ▶ slightly less efficient estimates
 - ▶ much too short confidence intervals

background

- ▶ we often encounter correlated data points, e.g. in ophthalmology we may measure both eyes of a patient, and maybe several times
- ▶ if we are interested in regression, we can use random effects to deal with the correlation among data points, e.g. `variable ~ (1 | patient / eye)`
- ▶ disregarding the correlation among data points, e.g. `variable ~ 1`, usually leads to:
 - ▶ slightly less efficient estimates
 - ▶ much too short confidence intervals
- ▶ if we are interested in correlation instead of regression, no simple solution seem to exist, e.g. something like `cor(c(variable1, variable2) ~ (1 | patient / eye))`

background

- ▶ we often encounter correlated data points, e.g. in ophthalmology we may measure both eyes of a patient, and maybe several times
- ▶ if we are interested in regression, we can use random effects to deal with the correlation among data points, e.g. `variable ~ (1 | patient / eye)`
- ▶ disregarding the correlation among data points, e.g. `variable ~ 1`, usually leads to:
 - ▶ slightly less efficient estimates
 - ▶ much too short confidence intervals
- ▶ if we are interested in correlation instead of regression, no simple solution seem to exist, e.g. something like `cor(c(variable1, variable2) ~ (1 | patient / eye))`
- ▶ this is very surprising, considering how often correlation estimates taking care of random effects would be useful

goal

- ▶ R package `corre` on CRAN providing a user-friendly implementation of correlation analysis with random effects: e.g.
`corre(c(variable1, variable2) ~ (1 | patient / eye), data)`

goal

- ▶ R package `corre` on CRAN providing a user-friendly implementation of correlation analysis with random effects: e.g.
`corre(c(variable1, variable2) ~ (1 | patient / eye), data)`
- ▶ the main target is the global correlation coefficient, with random effects considered as a nuisance to be taken care of

goal

- ▶ R package `corre` on CRAN providing a user-friendly implementation of correlation analysis with random effects: e.g.
`corre(c(variable1, variable2) ~ (1 | patient / eye), data)`
- ▶ the main target is the global correlation coefficient, with random effects considered as a nuisance to be taken care of
- ▶ point estimates are relatively easy, although several reasonable estimators can be constructed (as in the related problem of variance components estimation, e.g. ML vs REML)

goal

- ▶ R package `corre` on CRAN providing a user-friendly implementation of correlation analysis with random effects: e.g.
`corre(c(variable1, variable2) ~ (1 | patient / eye), data)`
- ▶ the main target is the global correlation coefficient, with random effects considered as a nuisance to be taken care of
- ▶ point estimates are relatively easy, although several reasonable estimators can be constructed (as in the related problem of variance components estimation, e.g. ML vs REML)
- ▶ a central goal of the package is producing confidence intervals (and p-values) via multilevel bootstrapping

goal

- ▶ R package `corre` on CRAN providing a user-friendly implementation of correlation analysis with random effects: e.g.
`corre(c(variable1, variable2) ~ (1 | patient / eye), data)`
- ▶ the main target is the global correlation coefficient, with random effects considered as a nuisance to be taken care of
- ▶ point estimates are relatively easy, although several reasonable estimators can be constructed (as in the related problem of variance components estimation, e.g. ML vs REML)
- ▶ a central goal of the package is producing confidence intervals (and p-values) via multilevel bootstrapping
- ▶ code is available (in part based on existing packages, such as `lme4` and `lmeresampler`), but should be further developed and documented