Mixed Effect Models in R

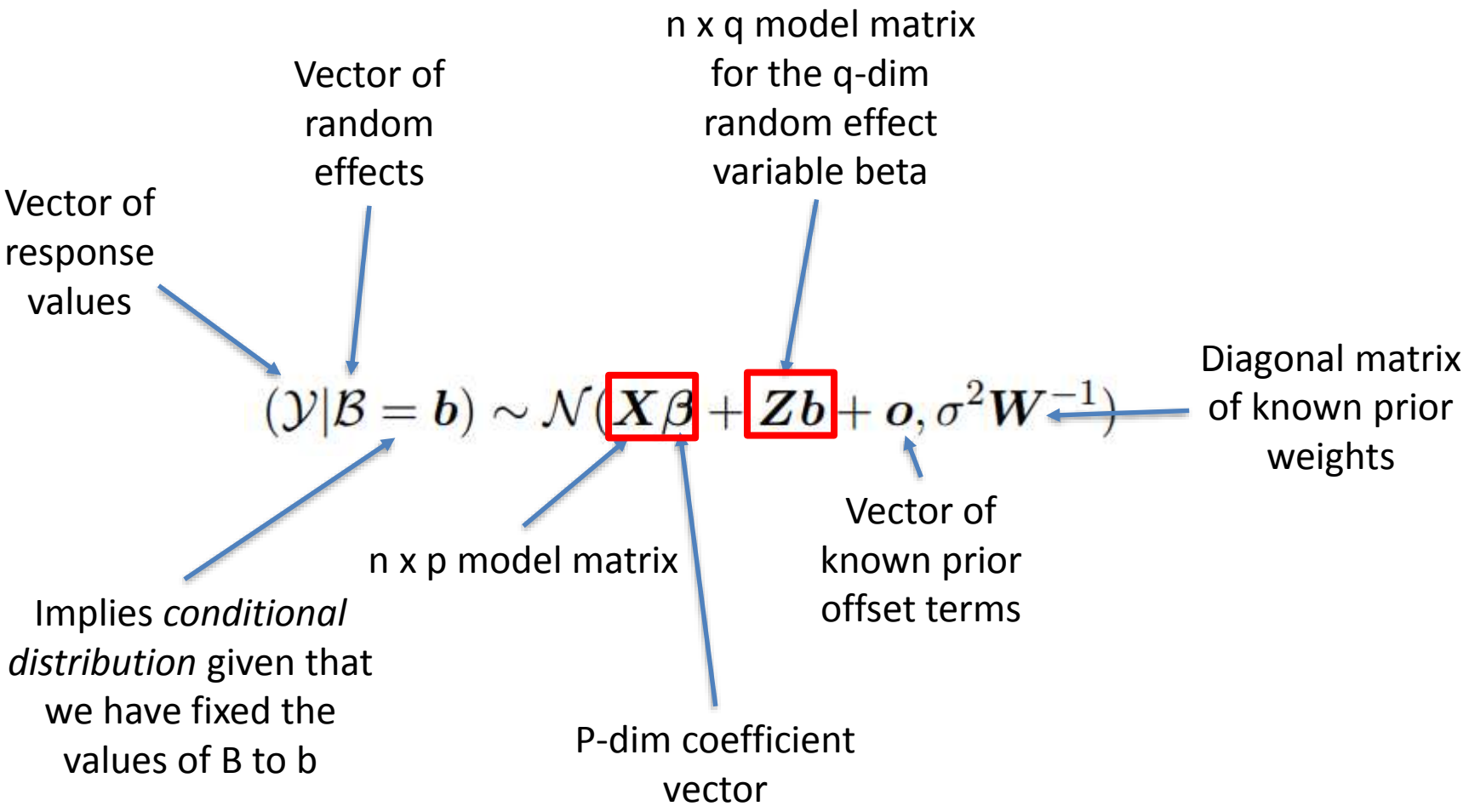
Hannah Tavalire, PhD

What is a mixed effect model?

- Assumptions (LME)
  - The explanatory variables are related **linearly** to the response…
  - Errors have constant variance.
  - Errors are independent.
  - Errors are normally distributed.
  - How would you check this…?
- ML versus REML? Depends on the questions…
  - **ML estimator** attempts to find the parameter values that maximize the likelihood function, given your observations.
  - **ML**: principle of ML applied to the full model
  - **REML**: ML applied to the least-squared residuals- subtract out the fixed effects then the distribution of the residuals depends on the variance components
  - REML better for random effects, ML better for fixed effects

Linear Mixed Effect Models

Distribution assumed multivariate normal with mean zero



(Bates et al., 2015)

What is a mixed effect model?

- Involves both *fixed* and *random* effects
  - Multiple definitions out there, but for the sake of today…
    - Fixed effect= Constant across individuals… we are interested in comparing levels
    - Random effect= Predictor variables where you are interested in making inferences about the distribution of values not necessarily the differences between particular levels
      - Allowed to vary across individuals
      - Can be uneven or within group replication can be small

Main uses of Mixed Effect Models?

- Experiments (blocks/plots)
- Variance component models
- Longitudinal data

Popular packages

- lme4
- nlme

- Some I have found useful…
  - MCMCglmm
  - coxme
  - VarComp

Lme4 versus nlme

- **PRO LME4**:
  - (1) more efficient linear algebra tools, giving improved performance on large problems;
  - (2) simpler syntax and more efficient implementation for fitting models with crossed random effects;
  - (3) the implementation of profile likelihood confidence intervals on random-effects parameters; and
  - (4) the ability to fit generalized linear mixed models
- **PRO NLME**: The main advantage of nlme relative to lme4 is a user interface for fitting models with structure in the residuals (various forms of heteroscedasticity and autocorrelation) and in the random-effects covariance matrices (e.g., compound symmetric models). With some extra effort, the computational machinery of lme4 can be used to fit structured models that the basic lmer function cannot handle. Actually gives significance values…

(Bates et al., 2015)

lme4

Formula	Alternative	Meaning
`(1 | g)`	`1 + (1 | g)`	Random intercept with fixed mean.
`0 + offset(o) + (1 | g)`	`-1 + offset(o) + (1 | g)`	Random intercept with *a priori* means.
`(1 | g1/g2)`	`(1 | g1) + (1 | g1:g2)`	Intercept varying among `g1` and `g2` within `g1`.
`(1 | g1) + (1 | g2)`	`1 + (1 | g1) + (1 | g2)`	Intercept varying among `g1` and `g2`.
`x + (x | g)`	`1 + x + (1 + x | g)`	Correlated random intercept and slope.
`x + (x || g)`	`1 + x + (1 | g) + (0 + x | g)`	Uncorrelated random intercept and slope.

Table 2: Examples of the right-hand-sides of mixed-effects model formulas. The names of grouping factors are denoted `g`, `g1`, and `g2`, and covariates and *a priori* known offsets as `x` and `o`.

Coding is similar in nlme… as you will see in the tutorial

(Bates et al., 2015)

Offset versus Weight

- Offset= A ‘structural’ predictor, not estimated by the model but assumed to have a value of 1. Differing levels of exposure to an event; common when dealing with count data (age of each forest plot could differ)
- Weight= Scales the variance of the response. Known values varying from observation to observation (area of each forest plot or number of trees measured per plot could differ)

lme4 & nlme Tutorial

- lme4
  - lmer (linear)
  - glmer (generalized- specify error distribution and link function)
    - binomial(link = "logit")
    - gaussian(link = "identity")
    - Gamma(link = "inverse")
    - inverse.gaussian(link = "1/mu^2")
    - poisson(link = "log")
    - quasi(link = "identity", variance = "constant")
    - quasibinomial(link = "logit")
    - quasipoisson(link = "log")
  - Also nlmer, glmer.nb, etc
- nlme
  - lme
  - gnls
  - nlme

References

Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., et al. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution, 24*(3), 127-135.

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics, 53*(3), 983-997.

Laird, N. M., & Ware, J. H. (1982). RANDOM-EFFECTS MODELS FOR LONGITUDINAL DATA. *Biometrics, 38*(4), 963-974.

Lindstrom, M. J., & Bates, D. M. (1990). NONLINEAR MIXED EFFECTS MODELS FOR REPEATED MEASURES DATA. *Biometrics, 46*(3), 673-687.