Practical Issues - var-cov complexity depends on sample sizes

- number of repeated observations n_i
  - simpler choices if n is small
    - random-int with CS assumption more reasonable ($q = 2$, number of var-covar parameters)
    - $\boldsymbol{R} = \texttt{UN}$ doesn't require estimation of too many parameters (*i.e.*, $q = 6$ if $n = 3$)
    - need to check degree of within-subject variation of outcome variable (esp. for categorical outcomes)
  - more choices as n gets large
- number of subjects N
  - more limitied choices as N is small
    - not much information for estimation of var-covar matrix
  - more possibilities as N is large

Practical Issues - estimation problems usually occur for variance-covariance parameters, not fixed effects

- near-zero variance components
  - rescale time to larger units (*e.g.*, months instead of days)
- extrapolation of intercept variance
  - for models with multiple random effects, ensure that 0 value (of variables with random effects) is meaningful; centering variables helps
- near-unity correlation of variance terms
  - orthogonal polynomials instead of raw metric for time effects
  - may suggest simpler model - not enough information from data to simultaneously estimate all model parameters

Practical Issues - ML or REML estimation

- ML estimates of variance parameters biased downward when N is small and p (number of covariates) is large
  - *e.g.*, ML estimate of error variance = SSE / N in ordinary regression, instead of SSE / $(N - p - 1)$

- use of likelihood-ratio tests for comparison of nested models is tricky under REML
  - ok only if the covariates in the two models are identical; *i.e.*, ok only for comparisons of different models of $V(\boldsymbol{y})$

Practical Issues - Model selection

- tests of fixed effects depends on structure fitted for $V(\boldsymbol{y})$
- many possible models for $V(\boldsymbol{y})$: random intercepts, random intercepts and trends, random intercepts trends and AR(1) errors ...
- 2-step procedure for model selection
  1. fit model with all fixed effects of potential interest and perform model selection of $V(\boldsymbol{y})$ structure
     - Likelihood ratio tests for nested models
       simulation studies suggest halved p-values for null hypothesis tests of var-covar parameters (see Snijders & Bosker (1999) Multilevel Analysis, page 90)
     - AIC and BIC for non-nested models (penalized versions of log-likelihood value for number of parameters)
  2. test fixed effects using model selected for $V(\boldsymbol{y})$

Practical Issues - Statistical tests of fixed effects

- likelihood-ratio tests for nested models
  - *e.g.*, model with group and time versus model with group, time, and group by time
  - sample size must be identical (careful with inclusion of covariates)
  - only valid under ML, not REML

- Wald tests (z-statistics) for specific model parameters
  - estimate / standard error $\sim$ N(0,1) under null hypothesis
  - approximate $t-$ and $F-$ statistics are better (*e.g.*, using Kenward and Roger adjustment); differences emerge as N is small

Practical Issues - Statistical tests of var-covar parameters

- Wald tests generally not recommended
  - normality for the sampling distribution of variance parameters is not reasonable

- LR tests under REML or ML
  - REML tests slightly better, but both accept null hypothesis (of variance parameters = 0) too often
  - simulation studies suggest halved p-values for null hypothesis tests of var-covar parameters (see Snijders & Bosker (1999) Multilevel Analysis, page 90)