MATH 60604A
Statistical modelling
§ 6d -Random slope model

Léo Belzile

HEC Montréal
Department of Decision Sciences

Model formulation

We consider a linear mixed model with a random slope and a random intercept for the `revenge` data, of the form

$$\boldsymbol{Y}_i \mid \boldsymbol{\mathcal{B}}_i = \boldsymbol{b}_i \sim \mathsf{No}_5\left(\mathbf{X}_i\beta + \mathbf{Z}_i\boldsymbol{b}_i, \sigma^2\mathbf{I}_5\right)$$

$$\boldsymbol{\mathcal{B}}_i \sim \mathsf{No}_2(\mathbf{0}_2, \boldsymbol{\Omega})$$

where $\mathbf{Z}_i = [\mathbf{1}_5, \texttt{time}_i]$ is a 5×2 model matrix for the random effects and $\boldsymbol{\Omega} = \left(\begin{smallmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{smallmatrix}\right)$.
The columns of $\mathbf{Z}_i$ typically include as covariates

- time or
- indicators for categorical variables (group effect).

Random effects on the predictor variables

Suppose the matrix $\mathbf{Z}_i = [\mathbf{1}_{n_i}, \mathbf{X}_{1i}]$.

$$Y_{ij} = (\beta_0 + b_i) + (\beta_1 + b_{1i})\mathrm{X}_{ij1} + \beta_2\mathrm{X}_{ij2} + \cdots + \beta_p\mathrm{X}_{ijp} + \varepsilon_{ij}.$$

- The effect of the variable X_1 for group i is $\beta_1 + b_{1i}$
- The parameter β_1 is the “slope” of X_1 averaged over the entire population.
- $\beta_1 + b_{1i}$ is the effect of X_1 specific to group i.

Covariance of the response

- The covariance matrix of Y_{ij} depends on the predictors in $\mathbf{Z}_i$ which have random effects.
- For example, if $\mathbf{Z}_i = [\mathbf{1}_{n_i}, \mathbf{X}_{1i}]$, the marginal variance of Y_{ij} is

$$\mathsf{Var}\left(Y_{ij} \mid \mathbf{X}_i\right) = \omega_{11} + \mathrm{X}_{ij1}^2\omega_{22} + 2\mathrm{X}_{ij1}\omega_{12} + \sigma_{\varepsilon}^2.$$

- If we assume that the residuals are independent, the covariance between two observations in the same group $(j \neq k)$ is

$$\mathsf{Cov}\left(Y_{ij}, Y_{ik} \mid \mathbf{X}_i\right) = \omega_{11} + \mathrm{X}_{ij1}\mathrm{X}_{1ik}\omega_{22} + (\mathrm{X}_{ij1} + \mathrm{X}_{1ik})\omega_{12}.$$

- It may be difficult to estimate parameters if the errors has a complex covariance structure (not to mention computational costs).

SAS code for random intercept model

```
proc mixed data=statmod.revenge;
model revenge = sex age vc wom t
    / ddfm=kenwardroger solution;
random intercept t / subject=id type=un v=1 vcorr=1;
run;
```

The output includes information about the number of covariance parameters, the number of random effects, etc.

Dimensions	
Covariance Parameters	4
Columns in X	6
Columns in Z per Subject	2
Subjects	80
Max Obs per Subject	5

Covariance matrix of response

Covariance Parameter Estimates		
Cov Parm	**Subject**	**Estimate**
UN(1,1)	id	0.3064
UN(2,1)	id	-0.05268
UN(2,2)	id	0.01730
Residual	id	0.2055

Estimated V Matrix for Subject 1					
Row	**Col1**	**Col2**	**Col3**	**Col4**	**Col5**
1	0.4239	0.1830	0.1476	0.1122	0.07682
2	0.1830	0.3704	0.1468	0.1287	0.1106
3	0.1476	0.1468	0.3515	0.1452	0.1444
4	0.1122	0.1287	0.1452	0.3672	0.1782
5	0.07682	0.1106	0.1444	0.1782	0.4175

- The variance of the random intercept is $\omega_{11} = 0.3064$
- The variance of the random slope is $\omega_{22} = 0.01730$
- The correlation between the random effects is -0.72.

Model comparison

- We can test whether $\mathscr{H}_0 : \omega_{12} = 0$ versus $\mathscr{H}_a : \omega_{12} \neq 0$ by fitting the model with diagonal covariance and performing a likelihood ratio test (REML, since they have the same fixed effects)
  - the test statistic is $R = 8.98$
  - its null distribution is χ^2_1
  - the *p*-value is 0.002:
  - the correlation between the random effects is strongly significant.

We can do similar comparisons with the random intercept-only model, but the null distribution is $\frac{1}{2}\chi^2_1 + \frac{1}{2}\chi^2_2$ and the asymptotic approximation is poor …probably better to use information criteria.