MA50260 Statistical Modelling

Lecture 15: (Linear) Mixed Effects Models - Estimation and Inference

Ilaria Bussoli

March 26, 2024

Linear Mixed Effect Models

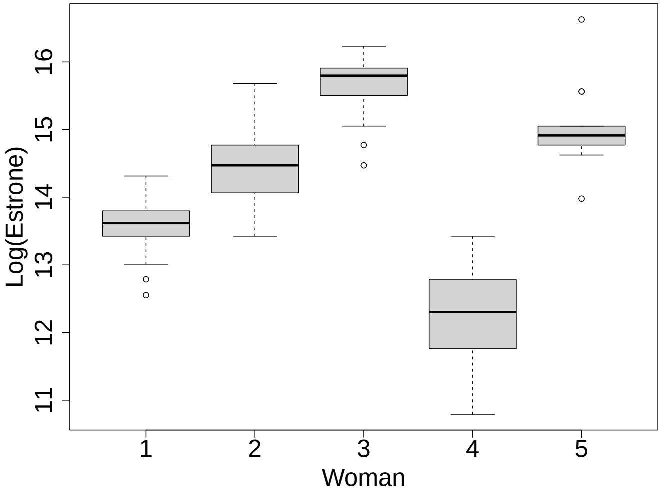
A mixed effects (normal) linear model is written as

$$\mathbf{Y} = \mathbf{X}\underline{\beta} + \mathbf{Zb} + \underline{\epsilon}$$

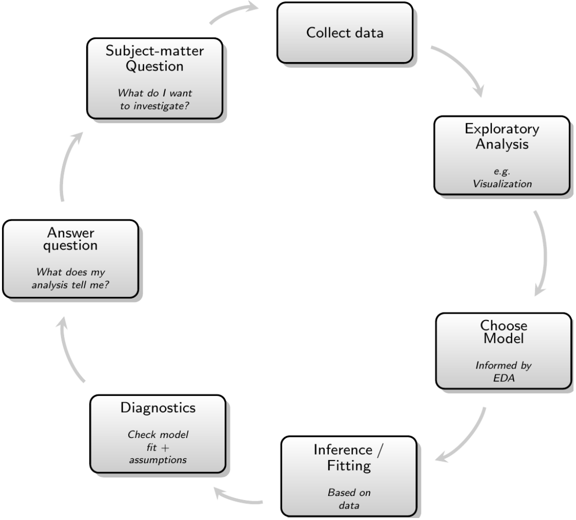
where

- $\mathbf{Y}$ is the vector of responses;
- $\mathbf{X}$ is the design matrix ;
- $\underline{\beta}$ are the regression coefficients;
- $\underline{\epsilon}$ are the vector of residual errors;
- $\mathbf{Z}$ is the matrix of covariates associated to q (unknown) random effects.

Example: Analysis of Estrone Levels



Today's Lecture



Motivation

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: logestrone ~ (1 | woman)
##    Data: estroneaov
##
## REML criterion at convergence: 157.8
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.56578 -0.41058  0.08074  0.57329  2.84352
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  woman    (Intercept) 1.7494   1.3227
##  Residual             0.3254   0.5705
## Number of obs: 80, groups:  woman, 5
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  14.1751     0.5949   23.83
```

Notation

We assume

$$\mathbf{b} \sim \mathrm{MVN}_q(0, \mathcal{D}),$$

with

$$\mathcal{D} = \sigma_b^2 \mathbf{I}_q.$$

The linear mixed effects model is then

$$\mathbf{Y} \sim \mathrm{MVN}_n\left(\mathbf{X}\underline{\beta}, \Sigma + \mathbf{Z}\mathcal{D}\mathbf{Z}^T\right),$$

where

$$\mathbf{\Sigma} = \sigma_\epsilon^2 \mathbf{I}_n.$$

We are interested in estimating $\underline{\beta}$ and $\underline{\theta} = (\sigma_b^2, \sigma_\epsilon^2)$.

Maximum-Likelihood Estimation

By writing

$$\mathbf{V}(\underline{\theta}) = \Sigma + \mathbf{Z}\mathcal{D}\mathbf{Z}^T,$$

we get

$$\ell(\underline{\beta}, \underline{\theta}) = -\frac{1}{2}\left\{\log|\mathbf{V}(\underline{\theta})| + \left(\mathbf{y} - \mathbf{X}\underline{\beta}\right)^T \mathbf{V}(\underline{\theta})^{-1}\left(\mathbf{y} - \mathbf{X}\underline{\beta}\right)\right\} \qquad (1)$$

where $|V(\underline{\theta})|$ is the determinant of $V(\underline{\theta})$.

Maximum-Likelihood Estimation

By writing

$$\mathbf{V}(\underline{\theta}) = \Sigma + \mathbf{Z}\mathcal{D}\mathbf{Z}^T,$$

we get

$$\ell(\underline{\beta}, \underline{\theta}) = -\frac{1}{2}\left\{\log|\mathbf{V}(\underline{\theta})| + \left(\mathbf{y} - \mathbf{X}\underline{\beta}\right)^T \mathbf{V}(\underline{\theta})^{-1}\left(\mathbf{y} - \mathbf{X}\underline{\beta}\right)\right\} \qquad (1)$$

where $|V(\underline{\theta})|$ is the determinant of $V(\underline{\theta})$.

For fixed $\underline{\theta}$, the maximum likelihood estimate for $\underline{\beta}$ is

$$\hat{\underline{\beta}}_\theta = \left(\mathbf{X}^T\mathbf{V}(\underline{\theta})^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{V}(\underline{\theta})^{-1}\mathbf{y}.$$

The variance parameters can be estimated by maximizing the **profile likelihood**, i.e. the likelihood with $\underline{\beta}$ replaced with $\hat{\underline{\beta}}_\theta$ in (1).

Estimation of the Random Effects

We can write

$$f(\mathbf{y}, \mathbf{b}) = f(\mathbf{y} \mid \mathbf{b}) f(\mathbf{b}).$$

Then

$$\frac{\partial \ell(\underline{\beta}, \underline{\theta}, \mathbf{b})}{\partial \mathbf{b}} = \mathbf{Z}^T \Sigma^{-1} \left(\mathbf{y} - \mathbf{X}\underline{\beta} - \mathbf{Z}\mathbf{b} \right) - \mathcal{D}^{-1}\mathbf{b},$$

i.e. $\hat{\mathbf{b}}$ solves

$$\left(\mathbf{Z}^T \Sigma^{-1} \mathbf{Z} + \mathcal{D}^{-1} \right) \hat{\mathbf{b}} = \mathbf{Z}^T \Sigma^{-1} \left(\mathbf{y} - \mathbf{X}\underline{\hat{\beta}} \right).$$

The solution

$$\hat{\mathbf{b}} = \mathcal{D}\mathbf{Z}^T \mathbf{V}(\underline{\theta})^{-1} \left(\mathbf{y} - \mathbf{X}\underline{\hat{\beta}} \right),$$

is known as the **best linear unbiased predictor (BLUP)**.

Restricted Maximum Likelihood Estimation (REML)

The ML estimates for σ_b^2 and σ_ϵ^2 are biased.

Restricted Maximum Likelihood Estimation (REML)

The ML estimates for σ_b^2 and σ_ϵ^2 are biased.

We instead use **Restricted Maximum Likelihood Estimation**:

1. Find matrix $\mathbf{K}$ such that $k^T\mathbf{X} = 0$ for all columns of $\mathbf{K}$.
2. Consider the likelihood function for the model

$$\mathbf{K}^T\mathbf{Y} \sim \mathrm{MVN}_n(0\,,\,\mathbf{K}^T\mathbf{V}(\underline{\theta})\mathbf{K})$$

We could thus estimate the variance components first, and then the regression coefficients.

Model Comparison for Fixed Effects

Suppose that $\mathcal{M}_1$ and $\mathcal{M}_2$ only differ in their fixed effects, with $\mathcal{M}_1$ nested in $\mathcal{M}_2$.

We can compare $\mathcal{M}_1$ and $\mathcal{M}_2$ using the likelihood ratio test statistic

$$2\left[\ell\left(\underline{\hat{\beta}}^{(2)}, \hat{\mathbf{b}}^{(2)}, \underline{\hat{\theta}}^{(2)}\right) - \ell\left(\underline{\hat{\beta}}^{(1)}, \hat{\mathbf{b}}^{(1)}, \underline{\hat{\theta}}^{(1)}\right)\right].$$

We compare to the chi-square distribution with degrees of freedom equal to the difference in the dimensions of the parameter spaces.

Note, we cannot use the REML estimates.

We can also use the AIC to compare models differing in terms of their fixed effects only.

Model Comparison for Random Effects

Models differing in their random effects structure can be compared using the likelihood ratio test statistic

$$2(\ell_2 - \ell_1) \sim \chi^2_{p_2 - p_1}.$$

Model Comparison for Random Effects

Models differing in their random effects structure can be compared using the likelihood ratio test statistic

$$2(\ell_2 - \ell_1) \sim \chi^2_{p_2 - p_1}.$$

- An important case we wish to test is $H_0 : \sigma_b^2 = 0$
- However, $2(\ell_2 - \ell_1)$ is no longer chi-squared distributed
- If you use the chi-square distribution, p-values will tend to be larger than they should be
- So, take care when using this approach for borderline results.

Prediction

Suppose we aim to predict the next observation for one of the $I = 5$ women in the estrone example.

The estimate is

$$\hat{\mu}_i = \hat{\mu} + \hat{b}_i,$$

where $\hat{\mu}$ is the estimated fixed effect and $\hat{b}_i$ is the best linear unbiased predictor.

Prediction

Suppose we aim to predict the next observation for one of the $I = 5$ women in the estrone example.

The estimate is

$$\hat{\mu}_i = \hat{\mu} + \hat{b}_i,$$

where $\hat{\mu}$ is the estimated fixed effect and $\hat{b}_i$ is the best linear unbiased predictor.

Suppose another woman is added to the study. Then, our estimate is

$$\hat{\mu}_{I+1} = \hat{\mu}.$$

Diagnostics

The observed residuals are defined as

$$\underline{\hat{\epsilon}} = \mathbf{y} - \mathbf{X}\underline{\hat{\beta}} - \mathbf{Z}\hat{\mathbf{b}}.$$

As for the normal linear regression model, we assess model fit using

- PP and QQ plots
- Plots of the predictor against the residuals
- Plots of the explanatory variables against the residuals

If the number of groups is not small, we should check for

- Normality of the group effects
- Constant variance