

Exam 1 Redo Final Questions

Dominic Adducci

Exam 1 Question 9

Given the initial pictures of these data, you want to also consider a random intercept and slope model. Write out the complete data model formulation for this model including dimensions for all matrices.

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{b} + \mathbf{E}$$

Dimensions:

- $\mathbf{Y} = (\Sigma n_i \times 1)$: where Σn_i is the total number of observations.
- $\mathbf{X} = (\Sigma n_i \times p)$: where p is the number of fixed effects, which in this case is 2. The fixed effects are the intercept and the evaluation number.
- $\beta = (p \times 1)$
- $\mathbf{Z}(\Sigma n_i \times q)$: Where q is the number of random effects, which in this case is 2. The random effects are a random intercept and a random slope for each subject ID.
- $\mathbf{b} = (q \times 1)$
- $\mathbf{E} = (\Sigma n_i \times 1)$

$$Y \sim N(\mathbf{X}\beta, ZGZ^T + R)$$

The distribution of the random effects is a multivariate normal with mean 0, and variance is a matrix \mathbf{G} .

$$\mathbf{b} \sim N(0, \mathbf{G})$$

$$G = \begin{bmatrix} G_1 & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & G_n \end{bmatrix}$$

$$G_i = \begin{bmatrix} \tau_0^2 & \tau_{01} \\ \tau_{01} & \tau_1^2 \end{bmatrix}$$

The distribution of the error term is a multivariate normal with mean 0, and the variance is a matrix \mathbf{R} .

$$\mathbf{E} \sim N(0, \mathbf{R})$$

$$R = \begin{bmatrix} R_1 & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & R_n \end{bmatrix}$$

$$R_i = \sigma^2 I_n$$

Exam 1 Question 10

The output below is for the random intercept and random intercept and slope models. Both models used a Satterthwaite adjustment for the denominator DF. Provide an intuition/explanation as to why the DF decreases for the random intercept and slope model compared to the random intercept model even though the adjustment method is the same.

In the random intercept and slope model the DF are lower because you are estimating both random slope and random intercept versus only estimating a random intercept. More degrees of freedom are used to estimate the random intercept and slope model than the random slope only model.

Exam 1 Question 11

$$\begin{aligned}
 U &= A^T Y \\
 E[U] &= E[A^T Y] = A^T \mu I_n = 0 \\
 V[U] &= V[A^T Y] = A^T V[Y] A = A^T \sigma^2 I_n A = \sigma^2 A^T A \\
 L &= \prod_{i=1}^{n-1} \frac{|A^T A|^{1/2}}{\sqrt{2\pi\sigma^2}} \exp \left[\frac{-1}{2\sigma^2} U^T (A^T A)^{-1} U \right] \\
 \log L(U) &= \frac{n-1}{2} \log(2\pi) - \frac{1}{2} \log((\sigma^2)^{n-1} |A^T A|) - \frac{1}{2\sigma^2} U^T (A^T A)^{-1} U \\
 \frac{\delta}{\delta \sigma^2} \log L(U) &= -\frac{n-1}{2\sigma^2} + \frac{1}{2\sigma^4} U^T (A^T A)^{-1} U = 0 \\
 \hat{\sigma}^2 &= \frac{U^T (A^T A)^{-1} U}{n-1} = \frac{\sum_i (Y_i - \bar{Y})^2}{n-1}
 \end{aligned}$$

Where:

$$\sum_i (Y_i - \bar{Y})^2 = \sum_i X_i^2 - n\bar{X}^2$$

Taking the expected value of this to show that $\hat{\sigma}$ is unbiased:

$$\begin{aligned}
 E[\sum_i X_i^2] &= nE[X_i^2] = n(\sigma^2 + \mu^2) \\
 E[n\bar{X}^2] &= nE[\bar{X}^2] = n(Var[\bar{X}] + E[\bar{X}]^2) = n \left(Var \left[\frac{1}{n} \sum X_i \right] + \mu^2 \right) \\
 &= n \left(\frac{n}{n^2} \sigma^2 + \mu^2 \right) = \sigma^2 + n\mu^2
 \end{aligned}$$

Putting these two expected values together:

$$E[\sum_i (Y_i - \bar{Y})^2] = E[\sum_i X_i^2 - n\bar{X}^2] = n(\sigma^2 + \mu^2) - \sigma^2 - n\mu^2 = n\sigma^2 - \sigma^2$$

Thus showing that the estimator is unbiased:

$$E[\hat{\sigma}^2] = E \left[\frac{\sum_i (Y_i - \bar{Y})^2}{n-1} \right] = \frac{\sigma^2(n-1)}{n-1} = \sigma^2$$

The likelihood is referred to as restricted because you can manipulate the mean to 0.