Additional note 12

Lecture 32

1. **Revision questions of L31**

   - What is the main difference between the random and fixed effect one-way ANOVA model?

     Answer: For fixed effect model, treatment effect α_i are fixed parameters to be estimated. For random effect model, treatment effect A_i are random and follow $N(0, \sigma_A^2)$. The variance σ_A^2 is the model parameter to be estimated but A_i are not model parameters.

   - From a practical perspective, when should one use random effect model?

     Answer: If there are many treatment levels but there are only a few observations for each level.

   - From a theoretical perspective, when should one use random effect model?

     Answer:
     Random effect model: The factor levels comprise a small portion of the whole population and constitutes a random sample. The research interest is on the population.

     Fixed effect model: The factor levels form a major portion of the whole population and the research interest is on individual treatment factor.

2. The estimates of fixed effect $\boldsymbol{\beta}$ and random effect $\boldsymbol{u}$ are

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= (\mathbf{X}^\top \mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^\top \mathbf{V}^{-1}\boldsymbol{y} \\ \tilde{\boldsymbol{u}} &= \mathbf{G}\mathbf{Z}^\top \mathbf{V}^{-1}(\boldsymbol{y} - \mathbf{X}\hat{\boldsymbol{\beta}}).\end{aligned}$$

3. Comparing fixed effect model M1: $\boldsymbol{Y} = (\alpha\beta)_{ij} + \epsilon_{ijk},\ \epsilon_{ijk} \sim \text{NID}(0, \sigma^2)$ with
   random effect model M2: $\boldsymbol{Y} = \mu + (AB)_{ij} + \epsilon_{ijk},\ \epsilon_{ijk} \sim \text{NID}(0, \sigma^2),\ (AB)_{ij} \sim (0, \sigma_A^2)$,
   the estimates of treatment group effect is: $\widehat{(\alpha\beta)}_{ij} = \bar{Y}_{ij}$ under M1 and
   the estimates of treatment group effect is: $\hat{\mu} + \widehat{(AB)}_{ij}$ under M2.

4. Under M2, $\hat{\boldsymbol{\beta}} = \hat{\mu}$ and $\tilde{\boldsymbol{u}} = \widehat{(AB)}_{ij}$ are estimated condition on $\mathbf{V} = \boldsymbol{Z}\boldsymbol{G}\mathbf{Z}^\top + \boldsymbol{R}$ in which $\boldsymbol{G}$ depends on σ_A^2 and $\boldsymbol{R}$ depends on σ^2. Since σ_A^2 and σ^2 are variances of all random components and all observations respectively and they are estimated based on all random effects and all observations respectively.
   With *borrowing strength*, the estimates $\hat{\mu} + \widehat{(AB)}_{ij}$ of M2 are closer to the overall mean than $\bar{Y}_{ij}$ in M1.

5. The marginal variance is $\text{Var}(\boldsymbol{Y}) = \mathbf{V} = \boldsymbol{Z}\boldsymbol{G}\mathbf{Z}^\top + \boldsymbol{R}$ which is bold diagonal and each bold matrix accounts for the dependence between observations within block and independence between blocks (independence of observations from different blocks). Note that conditional on $\boldsymbol{u}$ (being observed and so no contributed variability) $\text{Var}(\boldsymbol{Y}|\boldsymbol{u}) = \boldsymbol{R}$.

Lecture 33

1. **Revision questions of L32**

   - What are the conditional and marginal distributions of Y?

     Answer:
     Condition on the unobserved $\boldsymbol{u}$, $\boldsymbol{Y} \sim \text{N}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{u}, \boldsymbol{R})$.
     Marginally integrate out the unobserved $\boldsymbol{u}$, $\boldsymbol{Y} \sim \text{N}(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{Z}\boldsymbol{G}\mathbf{Z}^\top + \boldsymbol{R})$.

- What kind of likelihood, conditional or marginal, we can make use of to provide the estimator of $\boldsymbol{\beta}$ and $\boldsymbol{u}$ for given σ_A^2 and σ^2?

  Answer: Conditional likelihood which contain both $\boldsymbol{\beta}$ and $\boldsymbol{u}$ to be estimated.

- How do we compare the estimate for treatment i effect between the fixed effect model and random effect model?

  Answer: In the random effect model, estimates of treatment effect A_i borrow strength (information) from other treatment groups to estimate the common variance σ_A^2 and so A_i are closer to the overall mean $\bar{Y}_{\bullet\bullet}$ than $\hat{\alpha}_i$ in the fixed effect model which are estimated by the treatment group mean $\bar{Y}_{i\bullet}$.

2. In the estimation of σ_A^2 and σ^2, there are three methods, MM, REML and ML.

   MM: $\hat{\sigma}^2 = \text{MSR}$ and $\hat{\sigma}_A^2 = \frac{1}{r}(\text{MS}_A - \text{MSR})$.

   ML: $\text{argmax}_{\sigma_A^2,\sigma^2} \; -\frac{rt}{2}\log(2\pi) - \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}(\boldsymbol{y}-\mathbf{X}\hat{\boldsymbol{\beta}})^\top \mathbf{V}^{-1}(\boldsymbol{y}-\mathbf{X}\hat{\boldsymbol{\beta}})$.

   REML: $\text{argmax}_{\sigma_A^2,\sigma^2} \; -\frac{rt}{2}\log(2\pi) - \frac{1}{2}\log|\mathbf{V}| - \frac{1}{2}(\boldsymbol{y}-\mathbf{X}\hat{\boldsymbol{\beta}})^\top \mathbf{V}^{-1}(\boldsymbol{y}-\mathbf{X}\hat{\boldsymbol{\beta}}) - \frac{1}{2}\log|\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{X}|$.

3. REML is less biased than ML estiamtes and the difference between ML and REML is that REML accounts for the lost of df.

Lecture 34

1. **Revision questions of L33**

   - In the estimation of σ_A^2 and σ^2, what likelihood function, marginal or conditional, do we consider for ML and REML methods?

     Answer: marginal using $\boldsymbol{Y} \sim \text{N}(\mathbf{X}, \boldsymbol{Z}\boldsymbol{G}\mathbf{Z}^\top + \boldsymbol{R})$ as it does not involve $\boldsymbol{u}$ but σ_A^2 and σ^2 together in $\boldsymbol{V} = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}^\top + \boldsymbol{R}$ rather than $\boldsymbol{R}$ alone in the conditional likelihood.

   - The likelihood involve $\boldsymbol{\beta}$. How to deal with it? Estimate $\boldsymbol{\beta}$ together with σ_A^2 and σ^2 or substitute it by its estimate?

     Answer: substitute by its estimate and so the marginal likelihood is called profile likelihood.

   - Which method, ML or REML, give unbiased estimator of σ_A^2?

     Answer: REML method making use of residual likelihood. It corrects for the lost of df as it subtract the term $\frac{1}{2}\log|\mathbf{X}^\top\boldsymbol{V}^{-1}\mathbf{X}|$ from the profile likelihood and so often gives larger estimate of σ_A^2.

2. *Longitudinal* data favors random effect model as each time series naturally forms a random effect component. There are different models, fixed effect model (M1) for each time series and random effects model (M2 & M3). There are different assumptions for the random intercepts B_{i0} and random slopes B_{i1},

$$\text{Dependence M2:} \quad \boldsymbol{G} = \boldsymbol{I}_{18} \otimes \begin{bmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{bmatrix}. \tag{1}$$

$$\text{Independence M3:} \quad \boldsymbol{G} = \boldsymbol{I}_{18} \otimes \begin{bmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_1^2 \end{bmatrix}. \tag{2}$$

   Note the R codes to formulate these two models. The general model $\boldsymbol{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{u} + \boldsymbol{\epsilon}$, $\boldsymbol{u} \sim \text{N}(\mathbf{0}, \boldsymbol{G})$, $\boldsymbol{\epsilon} \sim \text{N}(\mathbf{0}, \boldsymbol{R})$. Note the different matrices and vectors $\boldsymbol{Y}$, $\mathbf{X}$, $\boldsymbol{\beta}$, $\mathbf{Z}$, $\boldsymbol{u}$, $\boldsymbol{G}$ and $\boldsymbol{R}$ and their dimensions.

Outcomes $\boldsymbol{Y}$: 180×1 vector $(y_{11}, y_{12}, \ldots, y_{18,10})$ of Reaction in P.4.

Fixed effect design matrix $\mathbf{X}$:

M1: 180×36 matrix in LHS of P.7;

M2, M3: 180×2 matrix in RHS of P.7.

Fixed effect parameters $\boldsymbol{\beta}$:

M1: 36×1 vector of subject specific intercept and slope $\beta_{10}, \beta_{11}, \beta_{20}, \beta_{21}, \ldots, \beta_{18,0}, \beta_{18,1}$;

M2, M3: 2×1 vector of common intercept and slope (β_0, β_1).

Random effect design matrix $\mathbf{Z}$: NA for fixed effect model M1;

M2, M3: 180×36 matrix in LHS of P.7.

Random effect $\mathbf{u}$: NA for fixed effect model M1;

M2, M3: 36×1 vector of subject specific relative intercept and slope effect $B_{10}, B_{11}, B_{20}, B_{21}, \ldots, B_{18,0}, B_{18,1}$ with respect to the common intercept and slope (β_0, β_1).

Random effect variance $\boldsymbol{G}$: NA for fixed effect model M1;

M2: 36×36 matrix defined in equation (1) and in P.10 as an example.

M3: 36×36 matrix defined in equation (2).

Residual variance $\boldsymbol{R}$: All M1,M2,M3: 180×180 diagonal matrix of σ^2, ie $\sigma^2 \mathbf{I}_{180}$.

---

last adjustments: May 28, 2021 by JC