

STAT 8320 Spring 2015 Assignment 4

March 17, 2015

► 1. **Solution.** (a). Define

$$\begin{aligned}\mathbf{Y}_i &= (Y_{i1}, Y_{i2})' \\ \boldsymbol{\beta} &= (\beta_0, \beta_0)' \\ \mathbf{b}_i &= (b_{0i}, b_{1i})' \\ \mathbf{W}_i &= \begin{pmatrix} 1 & W_{i1} \\ 1 & W_{i2} \end{pmatrix} \\ \boldsymbol{\epsilon}_i &= (\epsilon_{i1}, \epsilon_{i2})' \\ \mathbf{D} &= \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} \\ \boldsymbol{\Sigma} &= \begin{pmatrix} \sigma^2 & \\ & \sigma^2 \end{pmatrix}\end{aligned}$$

Then the model can be written as

$$\mathbf{Y}_i = \boldsymbol{\beta} + \mathbf{W}_i \mathbf{b}_i + \boldsymbol{\epsilon}_i$$

The marginal variance/covariance matrix of \mathbf{Y}_i is

$$\begin{aligned}Var(\mathbf{Y}_i) &= Var(\boldsymbol{\beta} + \mathbf{W}_i \mathbf{b}_i + \boldsymbol{\epsilon}_i) = Var(\mathbf{W}_i \mathbf{b}_i) + Var(\boldsymbol{\epsilon}_i) \\ &= \mathbf{W}_i Var(\mathbf{b}_i) \mathbf{W}_i' + \boldsymbol{\Sigma} \\ &= \mathbf{W}_i \mathbf{D} \mathbf{W}_i' + \boldsymbol{\Sigma} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 2 & \\ & 2 \end{pmatrix} \\ &= \begin{pmatrix} 10 & 11 \\ 11 & 18 \end{pmatrix}\end{aligned}$$

(b) The conditional variance/covariance matrix of \mathbf{Y}_i is

$$Var(\mathbf{Y}_i|\mathbf{b}_i) = Var(\boldsymbol{\epsilon}_i) = \boldsymbol{\Sigma} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

(c). The hypotheses are

$$H_0 : cov(b_{0i}, b_{1i}) = 0 \quad \text{v.s.} \quad H_a : cov(b_{0i}, b_{1i}) \neq 0$$

Then statistic is

$$\Lambda = -2(\ell(\text{Reduced Model}) - \ell(\text{Full Model})) = 426 - 420 = 6 \sim \chi^2(1)$$

Then the value of statistic is greater than $\chi^2_{0.95}(1) = 3.84$ with the P-value=0.014. Thus, we reject the null hypothesis, that is, we should favor the model for which the random effects parameters are dependent.

► **2. Solution.** (a). We have the form of model

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

where $\mathbf{Y} = (y_1, y_2)'$, $\mathbf{X} = (X_1, X_2)'$, $\boldsymbol{\beta} = \beta$, $\mathbf{b} = (b_1, b_2)'$, $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2)'$, and

$$\begin{aligned} \mathbf{Z} &= \begin{pmatrix} 1 & \\ & 2 \end{pmatrix} \\ \mathbf{D} &= \begin{pmatrix} \tau^2 & \frac{\phi\tau^2}{1+\phi^2} \\ \frac{\phi\tau^2}{1+\phi^2} & \tau^2 \end{pmatrix} \\ \boldsymbol{\Sigma} &= \begin{pmatrix} \sigma^2 & \\ & \sigma^2 \end{pmatrix} \end{aligned}$$

(b). The marginal variance/covariance matrix of \mathbf{Y} is that

$$Var(\mathbf{Y}) = \mathbf{ZDZ}^T + \boldsymbol{\Sigma} = \begin{pmatrix} \tau^2 + \sigma^2 & \frac{2\phi\tau^2}{1+\phi^2} \\ \frac{2\phi\tau^2}{1+\phi^2} & 4\tau^2 + \sigma^2 \end{pmatrix}$$

Then the marginal variance of Y_2 is $4\tau^2 + \sigma^2$ and the marginal covariance between Y_1 and Y_2 is

$$cov(Y_1, Y_2) = \frac{2\phi\tau^2}{1+\phi^2}$$

(c). Nothing. Because in the restricted likelihood function there is no parameters other than those from variance and covariance matrix, we can only test the variances and covariances, but not the parameters of fixed and random effects based on REML.

► **3. Solution.** (a). The intercepts should be same among the different graphs, but the increments of different graphs should be different, and the time points of the maximum weights should be different.

(b). We have the form of model

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \mathbf{e}_i$$

where

$$\begin{aligned}\mathbf{Y}_i &= (y_{i1}, \dots, y_{in_i})', \\ \mathbf{X}_i &= \begin{pmatrix} 1 & t_{i1} & t_{i1}^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{in_i} & t_{in_i}^2 \end{pmatrix}, \\ \mathbf{Z}_i &= (t_{i1}, \dots, t_{in_i})', \\ \boldsymbol{\beta} &= (\beta_0, \beta_1, \beta_2)', \\ \mathbf{b}_i &= b_{1i}, \\ \mathbf{e} &= (e_1, \dots, e_{in_i})', \\ \text{var}(\mathbf{e}_i) = \boldsymbol{\Sigma} &= \begin{pmatrix} \sigma^2 & & \\ & \ddots & \\ & & \sigma^2 \end{pmatrix} = \sigma^2 \mathbf{I}_{n_i \times n_i}, \\ \text{var}(\mathbf{b}_i) = \mathbf{D} &= \begin{pmatrix} \sigma_b^2 & & \\ & \ddots & \\ & & \sigma_b^2 \end{pmatrix} = \sigma_b^2 \mathbf{I}_{n_i \times n_i},\end{aligned}$$

(c). The marginal variance/covariance matrix of \mathbf{Y} is that

$$\begin{aligned} \text{Var}(\mathbf{Y}_i) &= \mathbf{Z}_i \mathbf{D} \mathbf{Z}_i^T + \Sigma \\ &= \begin{pmatrix} t_{i1} \\ \vdots \\ t_{in_i} \end{pmatrix} \begin{pmatrix} \sigma_b^2 & & \\ & \ddots & \\ & & \sigma_b^2 \end{pmatrix} (t_{i1}, \dots, t_{in_i}) + \begin{pmatrix} \sigma^2 & & \\ & \ddots & \\ & & \sigma^2 \end{pmatrix} \\ &= \begin{pmatrix} t_{i1}^2 \sigma_b^2 + \sigma^2 & t_{i1} t_{i2} \sigma_b^2 & \cdots & t_{i1} t_{in_i} \sigma_b^2 \\ t_{i2} t_{i1} \sigma_b^2 & t_{i2}^2 \sigma_b^2 + \sigma^2 & \cdots & t_{i2} t_{in_i} \sigma_b^2 \\ \vdots & \vdots & \ddots & \vdots \\ t_{in_i} t_{i1} \sigma_b^2 & t_{in_i} t_{i2} \sigma_b^2 & \cdots & t_{in_i}^2 \sigma_b^2 + \sigma^2 \end{pmatrix} \end{aligned}$$

(d). Because the marginal covariance \mathbf{Y} is

$$\text{cov}(Y_{ij}, Y_{ik}) = t_{ij} t_{ik} \sigma_b^2$$

then the correlation of Y_i is that

$$\begin{aligned} \text{cor}(Y_{ij}, Y_{ik}) &= \frac{\text{cov}(Y_{ij}, Y_{ik})}{\sqrt{\text{var}(Y_{ij})} \sqrt{\text{var}(Y_{ik})}} \\ &= \frac{t_{ij} t_{ik} \sigma_b^2}{\sqrt{t_{ij}^2 \sigma_b^2 + \sigma^2} \sqrt{t_{ik}^2 \sigma_b^2 + \sigma^2}} \\ &= \frac{jk}{\sqrt{j^2 + 1} \sqrt{k^2 + 1}} \\ &= \frac{1}{\sqrt{1/j^2 + 1} \sqrt{1/k^2 + 1}} \end{aligned}$$

The correlations will increase with the increase of time, j and k , but no trend just with temporal separation. This is not so realistic. In common sense, we usually may think that the correlations may be smaller with large temporal separation than the correlations with small temporal separation, because status of one time point is more likely to affect or to be affected by the status of the near time point. The reason causing this unrealistic result may be we simply assume the conditional independence while the data may not have this property.

(e). There are two advantages of the marginal covariance derived hierarchically. First, compared to the unstructured covariance structure, the hierarchical marginal

covariance have less unknown parameters to estimate, so it can reduce the computation, and avoid suffering overfitting problem. Secondly, it easily to understand and interpret the variance components, we can know that which parts of variation come from random effect and which parts come from the violation of conditional independence.

► **4. Solution.** (a) We have the split-plot design model

$$Y_{ijk} = \mu + \rho_i + \alpha_j + e_{ij} + \beta_k + (\alpha\beta)_{jk} + \epsilon_{ijk}$$

where ρ is plot effect, α is pasture effect and β is mineral effect. The random effects are ρ , e and ϵ . From the ANOVA table, we have that

$$\sigma_{\text{plot}}^2 = 12.74, \quad \sigma_e^2 = 1.05, \quad \sigma_{\epsilon}^2 = 2.25$$

In addition, the test for the significance for interaction of the pasture and mineral effects yields a P-value of 0.4981, the Factor pasture effect yields a P-value of 0.0377, and the Factor mineral effect yields a P-value of 0.0932. So only pasture effect are significant at $\alpha = 0.05$.

(b).

(c). From the output "Differences of Least Squares Means", we can see that the only significance difference is the difference between pasture 1 and pasture 4 based on the Tukey-Kramer adjustment.

► **5. Solution.** (a). From the two graphs, we can see that there are a lot of distinct line, which indicate the difference of different subject, the random effects. The mean intercepts of the lines from different groups almost same, and they should be because the experimental subjects should be randomly assigned to one group and they all come from a same population. Furthermore, the trends of different groups are different show that there may be some interaction between group effect and visits. Finally, the score trends with respect to time seems to be linear, and two lack of fit tests show that it is reasonable assumption with P-values are 0.9127 and 0.9131.

To sum up, we should have 3 assumptions from the graphs: i) same intercept; 2) different slope; 3) linearity.

(b). Firstly, we will define some notations. We denote i as the index of group with $i = 0, 1$; j as the index of visit with $j = 1, 2, 3, 4, 5$ and k as the index of subject with $k = 1, 2, \dots, 47$. $\mathbf{Y}_k = (y_{i1k}, y_{i2k}, y_{i3k}, y_{i4k}, y_{i5k})'$ denotes the repeated responses of a single subject. X_1 is a indicator variable of lecithin group, where $X_{1k} = 1$ when

the subject in lecithin group. X_2 is a indicator variable of placebo group, where $X_{2k} = 1$ when the subject in placebo group. t is treated as a continuous variable of visit. X_1t and X_2t denote the interaction of group and visit, which will show the different slopes of different groups. β_0 is the identical intercept, β_1 and β_2 are the coefficients of X_1t and X_2t . b_{0k} is the random component of the intercept, where $b_{0k} \sim N(0, \sigma_b^2)$. ϵ_{ijk} is error term, and $\epsilon_{ijk} \sim N(0, \sigma^2)$. Now we can construct the model based on the general structure,

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

where the $\mathbf{X}\boldsymbol{\beta}$ are fixed effects, and $\mathbf{Z}\mathbf{b}$ are random effects, and

- $\mathbf{Y} = (\mathbf{Y}'_1, \mathbf{Y}'_2, \dots, \mathbf{Y}'_{47})'$;
- $\mathbf{X} = (\mathbf{1}, \mathbf{X}_1\mathbf{t}, \mathbf{X}_2\mathbf{t}) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ X_{11}t & X_{12}t & \cdots & X_{1,47}t \\ X_{21}t & X_{22}t & \cdots & X_{2,47}t \end{pmatrix}'$;
- $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)'$;
- $\mathbf{Z} = \mathbf{1} = (1, 1, \dots, 1)'$;
- $\mathbf{b} = b_0$;
- $\boldsymbol{\epsilon} = (\epsilon_{1,1,1}, \epsilon_{1,2,1}, \dots, \epsilon_{1,5,25}, \epsilon_{2,1,26}, \dots, \epsilon_{2,5,47})'$.

We can also know that,

$$\text{var}(\mathbf{b}) = \mathbf{D} = \sigma_b^2, \quad \text{var}(\boldsymbol{\epsilon}) = \boldsymbol{\Sigma} = \left\{ \{d\sigma^2\}_{j=1}^5 \right\}_{k=1}^{25}$$

where $\{d\cdot\}$ denotes a diagonal matrix.

In this model, we assume that

- different groups have same intercept, that is, there exists significant;
- intercept has random component;
- $\mathbf{b} \sim N(\mathbf{0}, \mathbf{D})$, $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, that is, variance-covariance structure is VC(variance component), and residual structure is normal and independent;
- the relation between score and visit is linear.