STAT 510 Homework 8 Solutions Spring 2017

1. Please see the solution of problem 1 of exam 2 in 2016.
   Available at: https://dnett.github.io/S510/exam2sol2016.pdf.

2. (a) A linear-mixed effects model for the overall quality score is

$$y_{ijk} = \mu + \alpha_i + u_{ij} + \epsilon_{ijk},$$

where

- $\alpha_i$ is the fixed effect corresponding to temperature level $i = 1, 2, 3$,
- $u_{ij}$ is the random effect corresponding to cooler $j = 1, 2, 3, 4$ at temperature level $i$,
- $\epsilon_{ijk}$ is the random error for beef cut $k = 1, 2$ in cooler $j$ at temperature level $i$, and
- $u_{ij} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_u^2)$ independent of $\epsilon_{ijk} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$.

In matrix form, this model is

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{\epsilon},$$

where

- $\boldsymbol{y} = (y_{111}, y_{112}, y_{121}, \ldots, y_{142}, y_{211}, \ldots, y_{342})'$,
- $\boldsymbol{X} = (\mathbf{1}_{24\times 1}, \boldsymbol{I}_{3\times 3} \otimes \mathbf{1}_{8\times 1})$,
- $\boldsymbol{\beta} = (\mu, \alpha_1, \alpha_2, \alpha_3)'$,
- $\boldsymbol{Z} = (\boldsymbol{I}_{12\times 12} \otimes \mathbf{1}_{2\times 1})$,
- $\boldsymbol{u} = (u_{11}, u_{12}, \ldots, u_{34})'$,
- $\boldsymbol{\epsilon} = (\epsilon_{111}, \epsilon_{112}, \epsilon_{121}, \ldots, \epsilon_{142}, \epsilon_{211}, \ldots, \epsilon_{342})'$, and
- $\begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{\epsilon} \end{pmatrix} \sim \mathcal{N}\left( \begin{pmatrix} \mathbf{0}_{12\times 1} \\ \mathbf{0}_{24\times 1} \end{pmatrix}, \begin{pmatrix} \sigma_u^2 \boldsymbol{I}_{12\times 12} & \mathbf{0}_{12\times 24} \\ \mathbf{0}_{24\times 12} & \sigma_\epsilon^2 \boldsymbol{I}_{24\times 24} \end{pmatrix} \right)$.

(b) ANOVA table:

| Source | DF | Sums of Squares | Mean Squares | Expected Mean Squares |
|---|---|---|---|---|
| temperature | 3–1=2 | $\sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{2}(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot})^2$ | $\frac{8}{2}\sum_{i=1}^{3}(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot})^2$ | $\sigma_\epsilon^2 + 2\sigma_u^2 + 4\sum_{i=1}^{3}(\alpha_i - \bar{\alpha}_\cdot)^2$ |
| cooler(temp) | (4–1)(3)=9 | $\sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{2}(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot})^2$ | $\frac{2}{9}\sum_{i=1}^{3}\sum_{j=1}^{4}(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot})^2$ | $\sigma_\epsilon^2 + 2\sigma_u^2$ |
| cut(cooler,temp) | (2–1)(3)(4)=12 | $\sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{2}(y_{ijk} - \bar{y}_{ij\cdot})^2$ | $\frac{1}{12}\sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{2}(y_{ijk} - \bar{y}_{ij\cdot})^2$ | $\sigma_\epsilon^2$ |
| c. total | 24–1=23 | $\sum_{i=1}^{3}\sum_{j=1}^{4}\sum_{k=1}^{2}(y_{ijk} - \bar{y}_{\cdot\cdot\cdot})^2$ | | |

(c) A test of $H_0 : \alpha_1 - \alpha_2 = 0$ can be based on

$$t = \frac{\bar{y}_{1\cdot\cdot} - \bar{y}_{2\cdot\cdot} - 0}{\sqrt{\frac{2MS_{\text{cooler(temp)}}}{4\cdot 2}}} = \frac{\bar{y}_{1\cdot\cdot} - \bar{y}_{2\cdot\cdot}}{\sqrt{\frac{1}{4}\left(\frac{2}{9}\sum_{i=1}^{3}\sum_{j=1}^{4}(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot})^2\right)}} = \frac{\bar{y}_{1\cdot\cdot} - \bar{y}_{2\cdot\cdot}}{\sqrt{\frac{1}{18}\sum_{i=1}^{3}\sum_{j=1}^{4}(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot})^2}}.$$

The numerator should be obvious, but why use $\frac{2MS_{\text{cooler(temp)}}}{4\cdot 2}$ in the denominator? Notice that since the $u_{ij}$ and $e_{ijk}$ are all independent,

$$\begin{aligned}
\text{Var}(\bar{y}_{1\cdot\cdot} - \bar{y}_{2\cdot\cdot}) &= \text{Var}(\bar{y}_{1\cdot\cdot}) + \text{Var}(\bar{y}_{2\cdot\cdot}) - 2\,\text{Cov}(\bar{y}_{1\cdot\cdot}, \bar{y}_{2\cdot\cdot}) \\
&= \text{Var}\left(\mu + \alpha_1 + \bar{u}_{1\cdot} + \bar{\epsilon}_{1\cdot\cdot}\right) + \text{Var}\left(\mu + \alpha_2 + \bar{u}_{2\cdot} + \bar{\epsilon}_{2\cdot\cdot}\right) \\
&\qquad\qquad - 2\,\text{Cov}\left(\mu + \alpha_1 + \bar{u}_{1\cdot} + \bar{\epsilon}_{1\cdot\cdot}, \mu + \alpha_2 + \bar{u}_{2\cdot} + \bar{\epsilon}_{2\cdot\cdot}\right) \\
&= \text{Var}\left(\bar{u}_{1\cdot} + \bar{\epsilon}_{1\cdot\cdot}\right) + \text{Var}\left(\bar{u}_{2\cdot} + \bar{\epsilon}_{2\cdot\cdot}\right) - 2\,\text{Cov}\left(\bar{u}_{1\cdot} + \bar{\epsilon}_{1\cdot\cdot}, \bar{u}_{2\cdot} + \bar{\epsilon}_{2\cdot\cdot}\right) \\
&= \text{Var}\left(\bar{u}_{1\cdot}\right) + \text{Var}\left(\bar{\epsilon}_{1\cdot\cdot}\right) + \text{Var}\left(\bar{u}_{2\cdot}\right) + \text{Var}\left(\bar{\epsilon}_{2\cdot\cdot}\right) \\
&= \frac{\sigma_u^2}{4} + \frac{\sigma_\epsilon^2}{2\cdot 4} + \frac{\sigma_u^2}{4} + \frac{\sigma_\epsilon^2}{2\cdot 4} \\
&= \frac{2\left(\sigma_\epsilon^2 + 2\sigma_u^2\right)}{4\cdot 2} \\
&= \frac{2EMS_{\text{cooler(temp)}}}{4\cdot 2}.
\end{aligned}$$

(d) The degrees of freedom are 9, since the denominator is based on $MS_{\text{cooler(temp)}}$.

(e) The noncentrality parameter is

$$\frac{\alpha_1 - \alpha_2 - 0}{\sqrt{\frac{2(\sigma_\epsilon^2 + 2\sigma_u^2)}{4\cdot 2}}} = \frac{2(\alpha_1 - \alpha_2)}{\sqrt{\sigma_\epsilon^2 + 2\sigma_u^2}}.$$

3. Please see the solution of problem 2 of exam 2 in 2016.
   Available at: https://dnett.github.io/S510/exam2sol2016.pdf.

4. (a) The covariance between the heights of two plants (i.e., genotypes $k = 1, 2$) on the same table (i.e., watering level $j$ and greenhouse $i$) is

$$\begin{aligned}
\text{Cov}(y_{ij1}, y_{ij2}) &= \text{Cov}(\mu + g_i + \omega_j + t_{ij} + \gamma_1 + \phi_{j1} + e_{ij1}, \mu + g_i + \omega_j + t_{ij} + \gamma_2 + \phi_{j2} + e_{ij2}) \\
&= \text{Cov}(g_i + t_{ij} + e_{ij1}, g_i + t_{ij} + e_{ij2}) \qquad \text{dropping fixed effects} \\
&= \text{Cov}(g_i, g_i) + \text{Cov}(t_{ij}, t_{ij}) \qquad \text{since } g_i, t_{ij}, e_{ijk} \text{ are all independent} \\
&= \sigma_g^2 + \sigma_t^2.
\end{aligned}$$

The variance of any single observation is

$$\begin{aligned}
\text{Var}(y_{ijk}) &= \text{Var}(\mu + g_i + \omega_j + t_{ij} + \gamma_k + \phi_{jk} + e_{ijk}) \\
&= \text{Cov}(g_i + t_{ij} + e_{ijk}, g_i + t_{ij} + e_{ijk}) \qquad \text{dropping fixed effects} \\
&= \text{Cov}(g_i, g_i) + \text{Cov}(t_{ij}, t_{ij}) + \text{Cov}(e_{ijk}, e_{ijk}) \qquad \text{since } g_i, t_{ij}, e_{ijk} \text{ are all independent} \\
&= \sigma_g^2 + \sigma_t^2 + \sigma_e^2.
\end{aligned}$$

Hence, the correlation is

$$\begin{aligned}
\operatorname{Corr}(y_{ij1}, y_{ij2}) &= \frac{\operatorname{Cov}(y_{ij1}, y_{ij2})}{\sqrt{\operatorname{Var}(y_{ij1})\operatorname{Var}(y_{ij2})}} \\
&= \frac{\sigma_g^2 + \sigma_t^2}{\sigma_g^2 + \sigma_t^2 + \sigma_e^2}.
\end{aligned}$$

(b) If there are no watering level main effects, the fixed effects will be the same for each watering level $j$ when averaged across the other factors (i.e., averaged over $i$ and $k$). Written in terms of the model parameters, $\mu + \omega_j + \bar{\gamma}_\cdot + \bar{\phi}_{j\cdot}$ would be equal for all $j$. This happens if and only if $\omega_j + \bar{\phi}_{j\cdot}$ is equal for all $j$, so the null hypothesis of no watering level main effects is

$$H_0 : \omega_1 + \bar{\phi}_{1\cdot} = \omega_2 + \bar{\phi}_{2\cdot} = \omega_3 + \bar{\phi}_{3\cdot}$$

*Comments:* Note that $H_0 : \omega_1 = \omega_2 = \omega_3$ is *not* the null hypothesis of no watering level main effects. Even if $\omega_1 = \omega_2 = \omega_3$, there could still be main effects from the interaction terms.

(c) Let

- $\boldsymbol{\beta} = (\mu, \omega_1, \omega_2, \omega_3, \gamma_1, \gamma_2, \phi_{11}, \phi_{12}, \phi_{21}, \phi_{22}, \phi_{31}, \phi_{32})'$,
- $\boldsymbol{X} = (\mathbf{1}_{24\times1}, \mathbf{1}_{4\times1} \otimes \boldsymbol{I}_{3\times3} \otimes \mathbf{1}_{2\times1}, \mathbf{1}_{12\times1} \otimes \boldsymbol{I}_{2\times2}, \mathbf{1}_{4\times1} \otimes \boldsymbol{I}_{6\times6})$,
- $\boldsymbol{u} = (g_1, g_2, g_3, g_4, t_{11}, t_{12}, t_{13}, t_{21}, \ldots, t_{43})'$,
- $\boldsymbol{Z} = (\boldsymbol{I}_{4\times4} \otimes \mathbf{1}_{6\times1}, \boldsymbol{I}_{12\times12} \otimes \mathbf{1}_{2\times1})$.