

STAT 425 Assignment 8

Due Wednesday, May 12, 11:59 pm. Submit through Moodle.

Name: (insert your name here)

Netid: (insert)

Submit your computational work both as an R markdown (*.Rmd) document and as a pdf, along with any files needed to run the code. Embed your answers to each problem in the document below after the question statement. If you have hand-written work, please scan or take pictures of it and include in a pdf file, ideally combined with your pdf output file from R Markdown.

Most relevant class notes: 11.1.ExperimentalDesign1, 11.2.ExperimentalDesign2, 12.RandomEffects

Problem 1 (15 points, 3 points for each part)

The one-way random factor ANOVA model with n_i observations for each level of the random factor has the form

$$y_{ij} = \mu + \alpha_i + e_{ij}, \quad i = 1, 2, \dots, k; \quad j = 1, 2, \dots, n_i$$

where μ is a constant, **but the α_i and e_{ij} are independent random variables** distributed as

$$\alpha_i \sim N(0, \sigma_\alpha^2) \text{ and } e_{ij} \sim N(0, \sigma^2).$$

The total number of observations is $n = n_1 + n_2 + \dots + n_k$.

a) Show that the model assumptions stated above imply $E(y_{ij}) = \mu$ for all i and j .

Answer:

b) In general, if U_1 and U_2 are independent (and therefore uncorrelated) random variables, then we know that $Var(U_1 + U_2) = Var(U_1) + Var(U_2)$. Using this fact, obtain a simplified expression for $Var(y_{ij})$.

Answer:

c) For any 4 random variables U_1, U_2, U_3, U_4 , we know that

$$Cov(U_1 + U_2, U_3 + U_4) = Cov(U_1, U_3) + Cov(U_1, U_4) + Cov(U_2, U_3) + Cov(U_2, U_4).$$

Using these facts, show that if $i \neq i'$, then $Cov(y_{ij}, y_{i'j'}) = 0$, i.e., observations from different groups are uncorrelated.

Answer:

d) Now, we show that observations from the same group have nonzero covariance. Show that if $j \neq j'$, then $Cov(y_{ij}, y_{ij'}) = \sigma_\alpha^2$. (Useful fact: for any random variable U , $Cov(U, U) = Var(U)$)

Answer:

e) The final step is to obtain the intra-class correlation coefficient for two observations from the same group. Remembering the definition of correlation,

$$Corr(U, V) = \frac{Cov(U, V)}{\sqrt{Var(U)Var(V)}},$$

show that if $j \neq j'$ then

$$Corr(y_{ij}, y_{ij'}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma^2}.$$

Answer:

Problem 2 (15 points, 3 points for each part)

In assignment 7 we analyzed the `ctsib` data from `faraway` using `Subject` as a blocking variable as we studied the primary fixed factors `Surface` and `Vision`. In this problem we reanalyze the data using a mixed model analysis. We'll continue to treat `Surface`, `Vision` and their interactions as fixed effects, but we'll treat `Subject` as a random effect. Recall that each of the 40 subjects contributed 12 observations to the study, 2 attempts for each of the 2×3 combinations of levels of `Surface` and `Vision`. We expect the observations from the same subject to be correlated. Therefore we'll consider a linear mixed model of the form

$$y_{ijk} = \mu + \alpha_i^{Surf} + \beta_j^{Vis} + (\alpha\beta)_{ij}^{Surf*Vis} + \gamma_k^{Subj} + e_{ijk}$$

where the first four terms on the right side of the equation are fixed effects that satisfy reference cell constraints, γ_k^{Subj} are independent $N(0, \sigma_\gamma^2)$ random subject effects, and e_{ijk} are random independent errors.

a) We'll need the following libraries:

```
library(faraway)
library(lme4)
```

```
## Loading required package: Matrix
```

Then we can fit the model described above using the REML method as follows:

```
mod1.reml = lmer(CTSIB ~ Surface*Vision + (1 | Subject), data=ctsib)
summary(mod1.reml, corr=FALSE)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: CTSIB ~ Surface * Vision + (1 | Subject)
## Data: ctsib
##
## REML criterion at convergence: 463.1
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.47966 -0.69276 -0.01271  0.70899  2.91572
##
## Random effects:
## Groups Name Variance Std.Dev.
## Subject (Intercept) 0.07246 0.2692
## Residual 0.12400 0.3521
## Number of obs: 480, groups: Subject, 40
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)      2.53750    0.05798  43.766
## Surfacenorm      -0.75000    0.05568 -13.470
## Visiondome       -0.15000    0.05568  -2.694
## Visionopen       -0.66250    0.05568 -11.899
## Surfacenorm:Visiondome 0.10000    0.07874   1.270
## Surfacenorm:Visionopen 0.06250    0.07874   0.794
```

Based on the results above, calculate the REML estimate of the intraclass correlation coefficient for two observations from the same subject.

Answer:

b) Use the `confint` function on the fitted model to obtain bootstrap 95% confidence intervals for the estimated parameters (as demonstrated on the very last lecture slide of the entire course). Does the confidence interval for the **Subject** random effect variance exclude zero? What does this mean?

Answer:

c) Use the maximum likelihood (ML) method to refit the above model and show the model summary. Compare the subject variance component estimates from ML and REML. Which estimate is smaller and by how much?

Answer:

d) Use ML to fit the reduced model without the fixed effect interactions between **Surface** and **Vision** and show the model summary. The model should still include the **Subject**

random effect. How much did the subject random effect variance change compared with part c).

Answer:

e) By comparing the models in c) and d), compute the log-likelihood ratio test statistic for $H_0 : (\alpha\beta)_{ij}^{Surf*Vis} = 0$ (no fixed effect interactions between **Surface** and **Vision**). Also compute the chi-square p-value. Which model, full or reduced, should we use based on the result?

Answer: