STAT 512 Homework 9

Kathleen Wendt

04/21/2020

Part 1: Antihistamine data

For this problem use the data described in Ott & Longnecker problem 19.23 (p 1080 in the 7th Edition). We will verify that this is a BIBD in parts A and B. Note that Person acts as the blocking variable.

Question 1: BIBD components

Identify t, r, b and k. (4 pts)

t = 6 antihistamine treatments

r = 5 replicates of each treatment

b = 10 blocks (patients)

k = 3 natural block size (treatments per patient)

N = 30 observations

tr = bk

30 = 30

Question 2: BIBD lambda

Compute λ (and make sure it is an integer).

λ = r(k-1)/(t-1) = 5(3-1)/(6-1) = 10/5 = 2

Based on Q1 and Q2, this is a BIBD.

Question 3: BIBD with block as fixed

Treat Person (block) as fixed. Fit an appropriate model (using `lm()`) and include the Type 3 ANOVA table (using `Anova(, type = 3)`) in your assignment.

```
## Anova Table (Type III tests)
##
## Response: area_red
##              Sum Sq Df F value    Pr(>F)
## (Intercept) 2207.43  1 72.9953 3.796e-07 ***
## treatments  1747.06  5 11.5543 0.0001033 ***
## person       512.79  9  1.8841 0.1336294
## Residuals    453.61 15
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Question 4: Fixed BIBD CLD

Calculate `emmeans` and Tukey adjusted p-values for their differences. Summarize using a `cld` display and include it in your assignment. (4 pts)

```
##  treatments emmean  SE df lower.CL upper.CL .group
##  E            31.6 2.7 15     25.8     37.3  1
##  D            35.5 2.7 15     29.8     41.3  1
##  B            37.5 2.7 15     31.8     43.3  1
##  A            41.4 2.7 15     35.6     47.1  12
##  F            51.5 2.7 15     45.8     57.3   23
##  C            55.2 2.7 15     49.5     61.0    3
##
## Results are averaged over the levels of: person
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 6 estimates
## significance level used: alpha = 0.05
```

Question 5: Tukey's HSD

Calculate Tukey's HSD value by hand.

Tukey's HSD = 12.63

See code appendix for hand calculations.

Question 6: BIBD with block as random

Treat Person (block) as random. Fit an appropriate model (using `lmer()` from `lme4` package) and include the Type 3 ANOVA table (using `anova(, ddf="Kenward-Roger")`).

```
## Type III Analysis of Variance Table with Kenward-Roger's method
##            Sum Sq Mean Sq NumDF  DenDF F value    Pr(>F)
## treatments 1843.4  368.68     5 18.857  11.976 2.557e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Question 7: Random BIBD CLD

Calculate emmeans and Tukey adjusted p-values for their differences. Summarize using a cld display and include it in your assignment. (4 pts)

```
##  treatments emmean   SE df lower.CL upper.CL .group
##  E            33.1 2.85 24     27.2     38.9  1
##  D            34.9 2.85 24     29.0     40.8  1
##  B            36.1 2.85 24     30.2     42.0  1
##  A            41.8 2.85 24     36.0     47.7  12
##  F            50.8 2.85 24     44.9     56.7   23
##  C            56.1 2.85 24     50.2     62.0    3
##
```

```
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 6 estimates
## significance level used: alpha = 0.05
```

Question 8: Fixed vs. random

Comparing your results from #4 and #7, you should find that the emmeans are not the same. Is the ranking of the treatments the same? Are the conclusions about significant differences the same?

The estimated marginals means differ between the fixed effects and random effects models, but the ranking of the treatments and the conclusions about differences between treatments are the same.

Part 2: Wheat variety data

In a variety trial, a total of 7 varieties (Var) of wheat are considered. For each variety, n = 3 reps are randomly assigned to positions in a single field. The response variable is the Yield at the end of the trial. The data is available from Canvas as "Varieties.csv".

Question 9: Fixed effects model

Treat Variety as fixed and fit a one-way ANOVA model using `lm()`. Fit an appropriate model and construct the ANOVA table (using `Anova(, type = 3)`).

```
## Anova Table (Type III tests)
##
## Response: yield
##              Sum Sq Df F value    Pr(>F)
## (Intercept) 1008.33  1 165.430 3.821e-09 ***
## var          667.62  6  18.255 7.017e-06 ***
## Residuals     85.33 14
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Question 10: Null hypothesis

In the ANOVA table, you should find a test corresponding to Variety. Specifically, F = 18.26 and p-value <.0001. State the null hypothesis for this test.

No difference in yield by wheat variety.

Question 11: Estimated marginal means

Give the emmeans for each variety.

var	emmean	SE	df	lower.CL	upper.CL
A	18.33333	1.425393	14	15.27617	21.39050
B	31.00000	1.425393	14	27.94284	34.05716
C	22.66667	1.425393	14	19.60950	25.72383
D	25.00000	1.425393	14	21.94284	28.05716
E	33.66667	1.425393	14	30.60950	36.72383
F	17.33333	1.425393	14	14.27617	20.39050
G	22.66667	1.425393	14	19.60950	25.72383

Question 12: Random effects model

Treat Variety as random and fit a one-way random effects model using `lmer()`. Fit an appropriate model and include "Random Effects" table (giving the variance component estimates) in your assignment.

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: yield ~ (1 | var)
##    Data: wheat_data
##
## REML criterion at convergence: 113.4
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.52307 -0.39126  0.01374  0.51853  1.36196
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  var      (Intercept) 35.039   5.919
##  Residual              6.097   2.469
## Number of obs: 21, groups:  var, 7
##
## Fixed effects:
##             Estimate Std. Error     df t value Pr(>|t|)
## (Intercept)   24.381      2.301  6.005    10.6 4.14e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Question 13: Variation from variety

Use the `rand()` function to get a test corresponding to Variety. Specifically, Chi.sq = 19 and p-value $<$ 0.0001. State the null hypothesis for this test.

Null hypothesis: No additional variation from wheat type; σ^2 for variety = 0.

Question 14: Best Linear Unbiased Predictors

Give the BLUPs for each variety. Remember to include the intercept in the BLUP estimation.

```
## $var
##   (Intercept)
## A   -5.716094
## B    6.256198
## C   -1.620310
```

```
## D    0.585112
## E    8.776680
## F   -6.661275
## G   -1.620310
##
## with conditional variances for "var"
```

Appendix

```
# load packages
library(readxl)
library(tidyverse)
library(janitor)
library(car)
library(emmeans)
# set global options
knitr::opts_chunk$set(echo = FALSE,
                      warning = FALSE,
                      message = FALSE)
# read and prepare allergy/antihistamine data
allergy_data <- readxl::read_xlsx("data/ex19-23.xlsx") %>%
  janitor::clean_names() %>%
  dplyr::mutate(person = as.factor(person),
                treatments = as.factor(treatments))
# 3. treat person as fixed in bibd lm
med_bibd_fixed <- lm(area_red ~ treatments + person, data = allergy_data)
# 3. call anova type 3 table on fixed bibd lm
car::Anova(med_bibd_fixed, type = 3)
# 4. emmeans for fixed bibd lm
med_bibd_fixed_em <- emmeans::emmeans(med_bibd_fixed, pairwise ~ treatments)
# 4. cld for emmeans
emmeans::cld(med_bibd_fixed_em$emmeans)
# 5. calculate tukey critical value in R
qtukey(0.95, 6, 15)
# 5. calculate se diff in means "by hand"
sqrt(2*3*(453.61/15)/(6*2))
# 5. calculate tukey hsd "by hand"
4.595*(sqrt((453.61/15)/4))
# 5. short-cut with se from contrasts output
4.595*3.89/sqrt(2)
# detach emmeans
detach(package:emmeans)
# load random effects testing packages; lmer masked: lme4 -> lmerTest
library(lme4)
library(lmerTest)
# reload emmeans to work correctly with lmer
library(emmeans)
# 6. treat person as random in allergy bibd
med_bibd_rand <- lmer(area_red ~ treatments + (1|person), data = allergy_data)
# 6. call anova type 3 table on bibd random
anova(med_bibd_rand, ddf = "Kenward-Roger")
# 7. emmeans for random effects bibd lm
med_bibd_rand_em <- emmeans::emmeans(med_bibd_rand, pairwise ~ treatments)
# 7. cld for emmeans
emmeans::cld(med_bibd_rand_em$emmeans)
# detach all additional packages in session thus far
pacman::p_unload(pacman::p_loaded(), character.only = TRUE)
# load packages (again)
library(tidyverse)
library(janitor)
```

```r
library(car)
library(emmeans)
library(kableExtra)
# read and prepare wheat variety data
wheat_data <- readr::read_csv("data/Varieties.csv") %>%
  janitor::clean_names() %>%
  dplyr::mutate(var = as.factor(var))
# 9. treat variety as fixed in wheat lm
wheat_fixed_lm <- lm(yield ~ var, data = wheat_data)
# 9. call anova type 3 table on fixed wheat lm
car::Anova(wheat_fixed_lm, type = 3)
# 11. compute emmeans
wheat_lm_em <- emmeans::emmeans(wheat_fixed_lm, pairwise ~ var)
# 11. display emmeans
kableExtra::kable(wheat_lm_em$emmeans)
# detach emmeans
detach(package:emmeans)
# load random effects testing packages; lmer masked: lme4 -> lmerTest
library(lme4)
library(lmerTest)
# reload emmeans to work correctly with lmer
library(emmeans)
# 12. treat variety as random in wheat lm
wheat_rand_lm <- lmer(yield ~ (1|var), data = wheat_data)
# 12. call summary on random wheat lm; type 3 anova table empty because
# intercept is only fixed effect
summary(wheat_rand_lm)
# 13. test potency/value of random-effect term (variety) in lm
lmerTest::rand(wheat_rand_lm)
# 14. compute blups for wheat variety
lme4::ranef(wheat_rand_lm)
```