Lab 11 - Randomized Complete Block Designs

April 28, 2017

In this lab we will cover

1) Desigining a Randomized Complete Block Factorial Experiment
2) Running the RCB analysis

Note, in this lab we will be fitting a mixed effects model for the first time. That is to say, our model will have both fixed and random effect.

Part I - Designing a Randomized Complete Block Factorial Experiment

Suppose a researcher wants to design a factorial experiment consisting of a factor, `Poison` (3 levels: P1, P2, and P3), and a factor, `Medication` (4 levels: M1, M2, M3, and M4). Say 48 volunteers were selected from 4 different hospitals (12 volunteers from each hospital).
Note that these 48 volunteers were not randomly sampled from the entire population (why?), but they were randomly sampled within Hospital. Therefore we treat each hospital as a random block effect.

In this part of the lab we design the experiment with 4 blocks for a 3x4 factorial experiment, assuming each treatment is replicated once per block. How would you make a randomized plan for this design?

We can use the `fac.layout()` function (similar to how it was used in lab 9) to randomize the treatments within each block with slightly different arguments. The steps are as follows:

NOTE: At the end of the lab, the solution to this question goes through randomly assigning block as well as as treatment. Since each volunteer was selected from a hospital, we cannot randomly assign Hospital/Block in our case

1. Define the number of blocks, `b`, the number of treatments, `t`, and the total number of observations, `n`.

```
# Number of blocks
b=4

# Number of unique treatments
t=12

# Total number of observations
n=b*t
```

2. Create lists of the elements that will not be randomized. In addition to the subjects, for a RCBD, the blocks will also not be randomized. Note the number of subjects here (`t`) corresponds to the number of subjects *per block*.

```
# Creates list of factors that should not be randomized
# we have 4 different hospitals, so 4 blocks, with 12 volunteers per hospital
unrandomized.unit = list(Blocks = c("M1", "M2", "M3", "M4"), Units = 12)
```

3. Again we will create a data frame of the 48 treatment combinations in standard order. For this study, for each subject we want to randomly assign each subject a `Block`, `Poison` and `Medication` level where each combination is used exactly once.

```r
library(dae)
```

```
## Loading required package: ggplot2
```

```r
# Creates data frame of the 48 treatments in the standard order. times = 4 because there are 4 blocks
treatments2 = fac.gen(list(Poison=c("P1","P2","P3"), Medication = c("M1","M2","M3","M4")),times = 4)
```

4. Now, we can use the `fac.layout()` function to randomly assign treatments to the subjects in each block where each of the 12 treatments appears exactly once in each block.

```r
# Randomizes the 12 treatments within each block
treatments2.randomized = fac.layout(unrandomized=unrandomized.unit,randomized=treatments2)
```

Here is the result of the randomization. The last three columns indicate the `Block`, `Poison` and `Medication` level randomly assigned to the subjects listed in column 3.

```r
treatments2.randomized
```

```
##    .Units .Permutation Blocks Units Poison Medication
## 1       1            2     M1     1     P1         M3
## 2       2            9     M1     2     P1         M1
## 3       3            1     M1     3     P2         M2
## 4       4            8     M1     4     P2         M1
## 5       5            4     M1     5     P3         M4
## 6       6            3     M1     6     P2         M4
## 7       7           11     M1     7     P3         M3
## 8       8            6     M1     8     P1         M4
## 9       9           10     M1     9     P1         M2
## 10     10           12     M1    10     P3         M1
## 11     11            7     M1    11     P2         M3
## 12     12            5     M1    12     P3         M2
## 13     13           14     M2     1     P1         M3
## 14     14           21     M2     2     P1         M1
## 15     15           13     M2     3     P2         M2
## 16     16           20     M2     4     P2         M1
## 17     17           16     M2     5     P3         M4
## 18     18           15     M2     6     P2         M4
## 19     19           23     M2     7     P3         M3
## 20     20           18     M2     8     P1         M4
## 21     21           22     M2     9     P1         M2
## 22     22           24     M2    10     P3         M1
## 23     23           19     M2    11     P2         M3
## 24     24           17     M2    12     P3         M2
## 25     25           38     M3     1     P1         M3
## 26     26           45     M3     2     P1         M1
## 27     27           37     M3     3     P2         M2
## 28     28           44     M3     4     P2         M1
## 29     29           40     M3     5     P3         M4
## 30     30           39     M3     6     P2         M4
## 31     31           47     M3     7     P3         M3
## 32     32           42     M3     8     P1         M4
## 33     33           46     M3     9     P1         M2
## 34     34           48     M3    10     P3         M1
## 35     35           43     M3    11     P2         M3
## 36     36           41     M3    12     P3         M2
## 37     37           26     M4     1     P1         M3
```

```
## 38      38             33     M4     2     P1          M1
## 39      39             25     M4     3     P2          M2
## 40      40             32     M4     4     P2          M1
## 41      41             28     M4     5     P3          M4
## 42      42             27     M4     6     P2          M4
## 43      43             35     M4     7     P3          M3
## 44      44             30     M4     8     P1          M4
## 45      45             34     M4     9     P1          M2
## 46      46             36     M4    10     P3          M1
## 47      47             31     M4    11     P2          M3
## 48      48             29     M4    12     P3          M2
```

Part II - Doing an RCB analysis

Question 2 (50 points)

The security department at a university campus is concerned about the level of bike theft and the possibility of this being due to organized criminal activity. To try and determine whether there were patterns of theft over the week, they monitored traffic on campus to determine how many non-affiliated cars (without campus parking stickers) were entering campus but not using campus facilities (ie, stopping for very short periods). These were broken down by day of the week and measured over 12 randomly selected weeks in the year.

The data for this question is in CampusCars.csv.

a) (5 points) Should we regard week as being a random effect? Argue one way or the other.

Weeks are chosen randomly through the year, and as such could reasonably be treated as a random effect. However, if researcher is specifically interested in the weeks that were selected (e.g. weeks were selected associated with different holidays), then a fixed effect analysis would be more appropriate.

b) (10 points) Write down an appropriate statistical model for these data. What is the design of this model?

This is a randomized blocks design. An appropriate model is

$$y_{ij} = \mu + d_i + w_j + \epsilon_{ij}$$

*where d_i represents the effect of the ith day and w_j is the random effect of the jth week.

c) (5 points) Conduct an analysis of variance model to determine whether there are differences in days of the week. Report your null hypotheses, statistics and conclusions.

First, loading in the data and fit the model with and without the Day variable.

```r
cars = read.csv('CampusCars.csv',head=TRUE)
cars$Week = as.factor(cars$Week)

library('lme4')
```

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

```r
mod = lmer(Cars ~ Day + (1|Week),data=cars, REML = F)
modR = lmer(Cars ~ (1|Week),data=cars, REML = F)
```

Then we can test between the two models with the `anova` function.

```r
anova(mod, modR)
```

```
## Data: cars
## Models:
## modR: Cars ~ (1 | Week)
## mod: Cars ~ Day + (1 | Week)
##      Df    AIC    BIC   logLik deviance  Chisq Chi Df Pr(>Chisq)
## modR  3 252.40 258.68 -123.197   246.40
## mod   7 193.55 208.21  -89.775   179.55 66.845      4  1.051e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Note that, in order to ensure we are appropriately accounting for the random effects we need to set `REML = F` when fitting the models, and we directly compare to a reduced model. This completes a *likelihood ratio test*

Here our null hypothesis is that there are no differences between days of the week: $d_i = 0$ for all i. The Chisq statistic for this test is 133.11 and we reject the null hypothesis and conclude that the number of suspicious vehicles on campus changes with week day.

d) (10 points) Conduct a Tukey HSD proceedure to test for differences between the effects of different week days. Summarize in just a few words which days are different from eachother.

To do TukeyHSD with a mixed model, we used the `glht` function from the `multcomp` package. `glht` stands for "General Linear Hypothesis Test" and `multcomp` stands for "Multiple Comparison".

The use of the function is: `summary(glht(YOUR MODEL, linfct=mcp(YOUR FIXED FACTOR="Tukey")))`

```r
library(multcomp)
```

```
## Loading required package: mvtnorm

##
## Attaching package: 'mvtnorm'

## The following object is masked from 'package:dae':
##
##     rmvnorm

## Loading required package: survival

## Loading required package: TH.data

## Loading required package: MASS

##
## Attaching package: 'TH.data'

## The following object is masked from 'package:MASS':
##
##     geyser
```

```r
summary(glht(mod, linfct = mcp(Day = 'Tukey')))
```

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lmer(formula = Cars ~ Day + (1 | Week), data = cars, REML = F)
##
## Linear Hypotheses:
##                  Estimate Std. Error z value Pr(>|z|)
```

```
## Mon - Fri == 0  3.583e+00  3.024e-01  11.850  < 1e-04 ***
## Thu - Fri == 0  1.333e+00  3.024e-01   4.409 0.000121 ***
## Tue - Fri == 0  1.333e+00  3.024e-01   4.409 0.000100 ***
## Wed - Fri == 0  1.667e+00  3.024e-01   5.512  < 1e-04 ***
## Thu - Mon == 0 -2.250e+00  3.024e-01  -7.441  < 1e-04 ***
## Tue - Mon == 0 -2.250e+00  3.024e-01  -7.441  < 1e-04 ***
## Wed - Mon == 0 -1.917e+00  3.024e-01  -6.339  < 1e-04 ***
## Tue - Thu == 0  2.220e-16  3.024e-01   0.000 1.000000
## Wed - Thu == 0  3.333e-01  3.024e-01   1.102 0.805395
## Wed - Tue == 0  3.333e-01  3.024e-01   1.102 0.805384
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

Monday and Friday have different means from all other days. We do not have evidence that Tuesday, Wednesday and Thursdya have different means from eachother, however they are different from Monday and Friday.

e) (10 points) How much variation is due to differences between weeks as opposed to differences between days within a week? What is the estimate of week-to-week variation? Provide an interpretation in terms of what you would expect to observe if you repeated this experiment.

```r
summary(mod)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: Cars ~ Day + (1 | Week)
##    Data: cars
##
##      AIC      BIC   logLik deviance df.resid
##    193.5    208.2    -89.8    179.5       53
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6040 -0.3755  0.0892  0.4928  2.3590
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Week     (Intercept) 4.6736   2.1619
##  Residual             0.5486   0.7407
## Number of obs: 60, groups:  Week, 12
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  48.9167     0.6597   74.15
## DayMon        3.5833     0.3024   11.85
## DayThu        1.3333     0.3024    4.41
## DayTue        1.3333     0.3024    4.41
## DayWed        1.6667     0.3024    5.51
##
## Correlation of Fixed Effects:
##        (Intr) DayMon DayThu DayTue
## DayMon -0.229
## DayThu -0.229  0.500
## DayTue -0.229  0.500  0.500
## DayWed -0.229  0.500  0.500  0.500
```

About

```
4.6736/(4.6737+0.5486)
```

```
## [1] 0.8949314
```

percent of variability is due to differences between weeks. This means that in a new analysis we could expect that we would expect to see average observations over a week differ from each other by a standard deviation of about 2.16 cars, but the variability within a week (after accounting for day effects) would have have standard deviation only 0.74.

Part 1 example with randomly assigned blocks

1. Define the number of blocks, `b`, the number of treatments, `t`, and the total number of observations, `n`.

```
# Number of blocks
b=4

# Number of unique treatments
t=12

# Total number of observations
n=b*t
```

2. Create lists of the elements that will not be randomized. In addition to the subjects, for a RCBD, the blocks will also not be randomized. Note the number of subjects here (`t`) corresponds to the number of subjects *per block*.

```
# Creates list of factors that should not be randomized
unrandomized.unit = list(Subject = 48)
```

3. Again we will create a data frame of the 48 treatment combinations in standard order. For this study, for each subject we want to randomly assign each subject a `Block`, `Poison` and `Medication` level where each combination is used exactly once.

```
library(dae)

# Creates data frame of the 48 treatments in the standard order
treatments2 = fac.gen(list(Block = c("B1","B2","B3","B4"),Poison=c("P1","P2","P3"), Medication = c("M1"
```

4. Now, we can use the `fac.layout()` function to randomly assign treatments to the subjects in each block where each of the 12 treatments appears exactly once in each block.

```
# Randomizes the 12 treatments within each block
treatments2.randomized = fac.layout(unrandomized=unrandomized.unit,randomized=treatments2)
```

Here is the result of the randomization. The last three columns indicate the `Block`, `Poison` and `Medication` level randomly assigned to the subjects listed in column 3.

```
treatments2.randomized
```

```
##    .Units .Permutation Subject Block Poison Medication
## 1       1           22       1    B1     P1         M3
## 2       2           21       2    B3     P2         M4
## 3       3            1       3    B4     P3         M4
## 4       4           33       4    B3     P3         M4
## 5       5           36       5    B2     P3         M3
## 6       6           24       6    B4     P1         M4
## 7       7           32       7    B2     P1         M3
```

```
## 8        8           43       8  B2    P2          M3
## 9        9           12       9  B2    P1          M4
## 10      10           18      10  B4    P3          M3
## 11      11           19      11  B2    P1          M2
## 12      12           14      12  B1    P3          M1
## 13      13           46      13  B4    P2          M2
## 14      14           11      14  B1    P3          M4
## 15      15            7      15  B2    P2          M2
## 16      16            9      16  B4    P2          M3
## 17      17           40      17  B4    P1          M1
## 18      18           15      18  B1    P3          M2
## 19      19            8      19  B1    P3          M3
## 20      20           47      20  B3    P2          M3
## 21      21           23      21  B1    P1          M2
## 22      22           48      22  B1    P1          M1
## 23      23            5      23  B2    P3          M1
## 24      24           30      24  B1    P2          M2
## 25      25           31      25  B3    P2          M2
## 26      26           35      26  B3    P3          M1
## 27      27           45      27  B3    P3          M2
## 28      28           38      28  B3    P2          M1
## 29      29           28      29  B4    P3          M1
## 30      30           25      30  B2    P3          M4
## 31      31           20      31  B3    P1          M1
## 32      32            2      32  B1    P2          M3
## 33      33           26      33  B1    P1          M4
## 34      34           27      34  B4    P1          M2
## 35      35           41      35  B3    P1          M2
## 36      36            4      36  B1    P2          M1
## 37      37           17      37  B4    P2          M4
## 38      38           34      38  B3    P1          M4
## 39      39           39      39  B4    P1          M3
## 40      40            6      40  B2    P2          M1
## 41      41           44      41  B3    P3          M3
## 42      42           13      42  B4    P3          M2
## 43      43           16      43  B1    P2          M4
## 44      44           37      44  B4    P2          M1
## 45      45           29      45  B3    P1          M3
## 46      46           42      46  B2    P1          M1
## 47      47           10      47  B2    P2          M4
## 48      48            3      48  B2    P3          M2
```