HW#11 Solution (week13 HW)

Due 12/1/2019

HW 25.5 Filling machines: CH25PR5-1factor-REM.txt

```
Machines = read.table(url(
  "https://raw.githubusercontent.com/npmldabook/Stat3119/master/Week-13/CH25PR5-1factor-REM.txt"))
names(Machines) = c("Response", "machine", "units")
Machines$machine = as.factor(Machines$machine);
str(Machines)
```

```
## 'data.frame':    120 obs. of  3 variables:
##  $ Response: num  -0.14 0.2 0.07 0.18 0.38 0.1 -0.04 -0.27 0.27 -0.21 ...
##  $ machine : Factor w/ 6 levels "1","2","3","4",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ units   : int  1 2 3 4 5 6 7 8 9 10 ...
```

(a) Follows the defintion from one-factor random effects model

R immplementation b) test $H_0 : \sigma^2_\mu = 0$ vs. $H_a : \sigma^2_\mu > 0$

```
summary(aov(Response~machine , data=Machines ))
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## machine       5  2.289  0.4579   14.78 3.64e-11 ***
## Residuals   114  3.531  0.0310
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results: Same F-test as one-way ANOVA for fixed model. $F - stat = 14.78 \sim F(5, 114)$, P<0.0001, we reject null.

R immplementation c)

```
library(lme4)
```

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

```
# We want to have a random effect per Officer
fit1 <- lmer(Response ~ (1 | machine), data = Machines)
fixef(fit1)
```

```
## (Intercept)
##   0.2276667
```

```r
confint(fit1, oldNames = FALSE)
```

```
## Computing profile confidence intervals ...
```

```
##                              2.5 %    97.5 %
## sd_(Intercept)|machine 0.07612263 0.2741633
## sigma                  0.15539471 0.2015542
## (Intercept)            0.09685478 0.3584785
```

Results: The estimate of overall mean is 0.2277 and 95% CI is (0.097, 0.36).

HW 25.6 Filling machines: CH25PR5-1factor-REM.txt

a) total variablity

```r
summary(fit1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response ~ (1 | machine)
##    Data: Machines
##
## REML criterion at convergence: -57.5
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.01114 -0.71863  0.04014  0.76492  1.91487
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  machine  (Intercept) 0.02134  0.1461
##  Residual             0.03097  0.1760
## Number of obs: 120, groups:  machine, 6
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  0.22767    0.06177   3.686
```

```r
ICC = 0.02134/(0.02134+ 0.03097)

MSTR= 0.4579
MSE= 0.0310
F.upp= qf(.975, 5,114)
F.low= qf(.025, 5,114)
n=20
L = (MSTR/MSE*(1/F.upp)-1)/20
U = (MSTR/MSE*(1/F.low)-1)/20
Lower.limit =  L/(1+L)
upper.limit =  U/(1+U)

paste("The estimate of proportion of variance explained is", round(ICC,4) )
```

```
## [1] "The estimate of proportion of variance explained is 0.408"
```

```
paste("The 95% CI is (",  round(Lower.limit, 4), ",",  round(upper.limit,4), ")." )
```

```
## [1] "The 95% CI is ( 0.1841 , 0.816 )."
```

b) we square of the CI for σ^2 from confint output.

From the output, for σ^2 , estimate=0.03097, 95% CI
$(0.15539471^2, 0.2015542^2) = (0.0241, 0.041)$.

c)

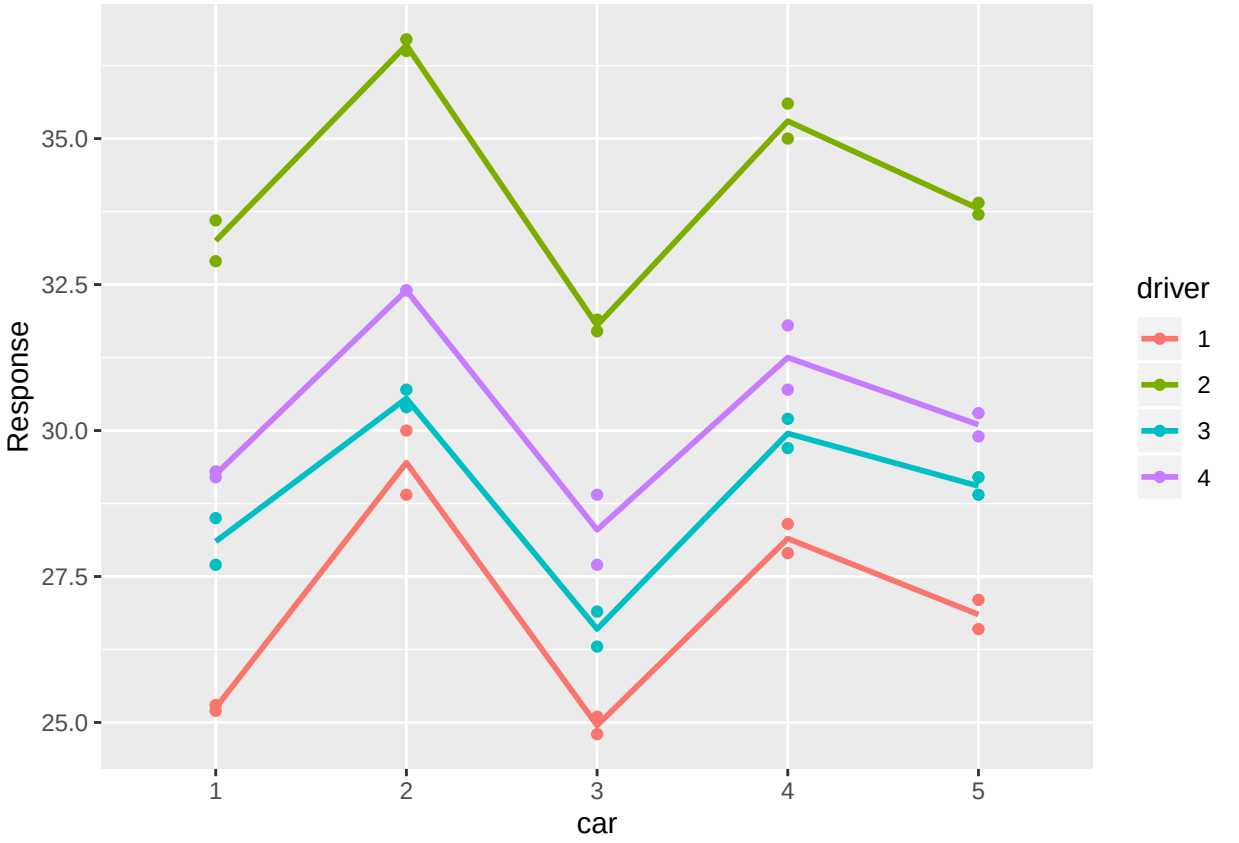
From the output, for σ^2_μ , estimate=0.02134, 95% CI
$(0.07612263^2, 0.2741633^2) = (0.00579, 0.0752)$.

25.15 - Miles per gallon 2-factor random effects model:

```
Miles = read.table(url(
  "https://raw.githubusercontent.com/npmldabook/Stat3119/master/Week-13/CH25PR15-2factorREM.txt"))
names(Miles) = c("Response", "driver", "car", "units")
Miles$driver = as.factor(Miles$driver)
Miles$car = as.factor(Miles$car)

str(Miles)
```

```
## 'data.frame':    40 obs. of  4 variables:
##  $ Response: num  25.3 25.2 28.9 30 24.8 25.1 28.4 27.9 27.1 26.6 ...
##  $ driver  : Factor w/ 4 levels "1","2","3","4": 1 1 1 1 1 1 1 1 1 1 ...
##  $ car     : Factor w/ 5 levels "1","2","3","4",..: 1 1 2 2 3 3 4 4 5 5 ...
##  $ units   : int  1 2 1 2 1 2 1 2 1 2 ...
```

```
# use ggplot and group by machine
library(ggplot2)
ggplot(Miles, aes(x =car , y = Response , group = driver , col = driver )) +
  geom_point() + stat_summary(fun.y = mean, geom = "line", size=1)
```



(a-b) ANOVA test

```
summary(aov(Response~driver*car  , data=Miles ))
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## driver       3 280.28   93.43  531.60  < 2e-16 ***
## car          4  94.71   23.68  134.73 3.66e-14 ***
## driver:car  12   2.45    0.20    1.16    0.371
## Residuals   20   3.52    0.18
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results: (a) For Interaction, F-test is the same. p=0.37 and we don't reject H_0: no interaction.

```
  MSA= 93.43
  MSB= 23.68
  MSAB= 0.20

# For factor A
 (F.statA = MSA/MSAB)
```

```
## [1] 467.15
```

```r
# p-value for testing factor A
 (pf.A= 1- pf(F.statA, 3, 12))
```

```
## [1] 1.094125e-12
```

```r
# For factor B
 (F.statB = MSB/MSAB)
```

```
## [1] 118.4
```

```r
# p-value for testing factor B
 (pf.B= 1- pf(F.statB, 4, 12))
```

```
## [1] 1.560327e-09
```

Results: (b) For factor A and B, F-test using MSAB as denominator, and both factors had significant effects with p<0.0001.

(c-e) Estimation and CI

```r
library(lme4)
fit.Miles <- lmer(Response ~ (1 | driver  ) + (1 | car) + (1 | driver:car), data = Miles )
summary(fit.Miles)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response ~ (1 | driver) + (1 | car) + (1 | driver:car)
##    Data: Miles
##
## REML criterion at convergence: 86.8
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.54828 -0.61796 -0.08964  0.61863  1.99002
##
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  driver:car (Intercept) 0.01399  0.1183
##  car        (Intercept) 2.93576  1.7134
##  driver     (Intercept) 9.31843  3.0526
##  Residual               0.17579  0.4193
## Number of obs: 40, groups:  driver:car, 20; car, 5; driver, 4
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)   30.048      1.709   17.58
```

```r
confint(fit.Miles, oldNames = FALSE)
```

```
## Computing profile confidence intervals ...
```

```
##                                 2.5 %     97.5 %
## sd_(Intercept)|driver:car  0.0000000  0.4164478
## sd_(Intercept)|car         0.9554788  3.8124705
## sd_(Intercept)|driver      1.5358586  6.7100121
## sigma                      0.3165500  0.5618089
## (Intercept)               26.3640965 33.7308423
```

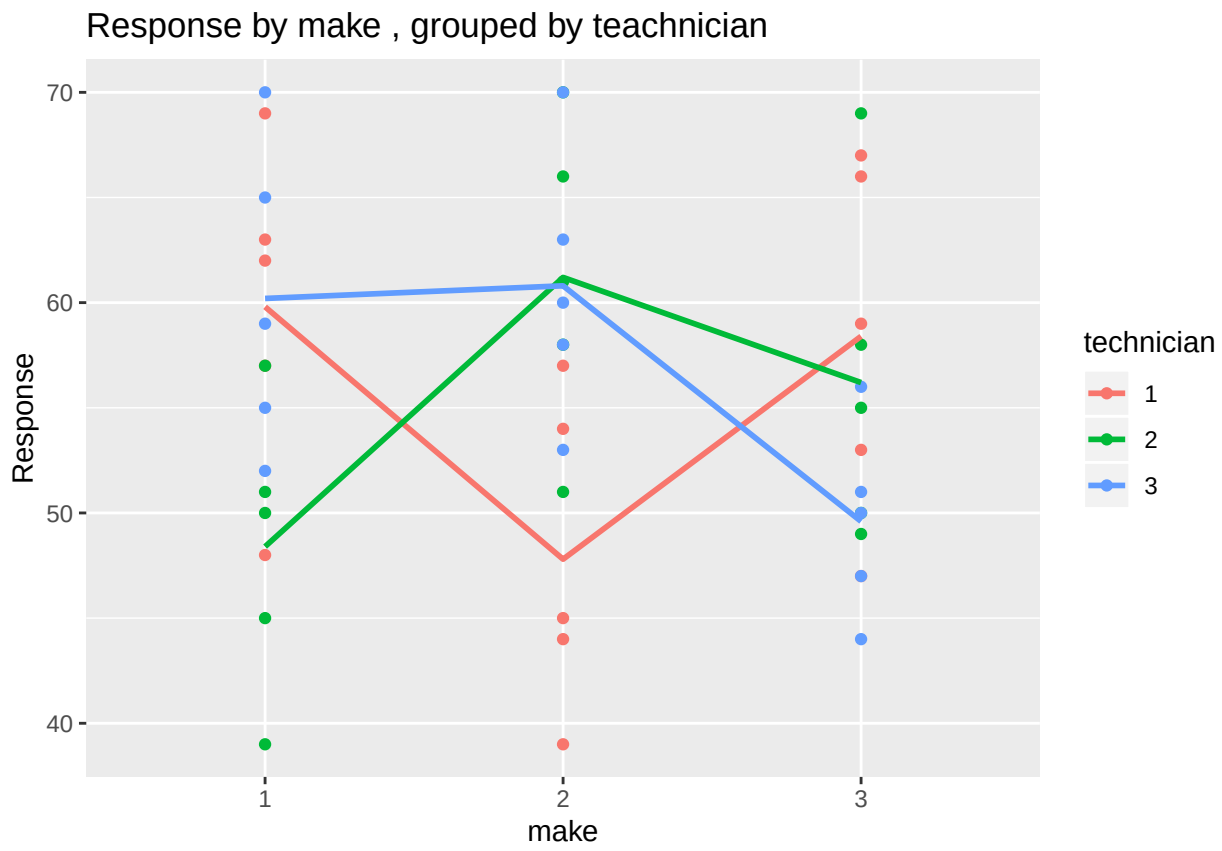
Results: (c)-(e) For driver: σ_a^2 estimate = 9.31843 with 95% CI for σ_a as (1.5358586 , 6.7100121). For car: σ_b^2 estimate= 2.93576 , with 95% CI for σ_b as (0.9554788 , 3.8124705). Driver had more effects than car on the miles.

25.16 Disk drive service (a-d)- 2-factor mixed effects model

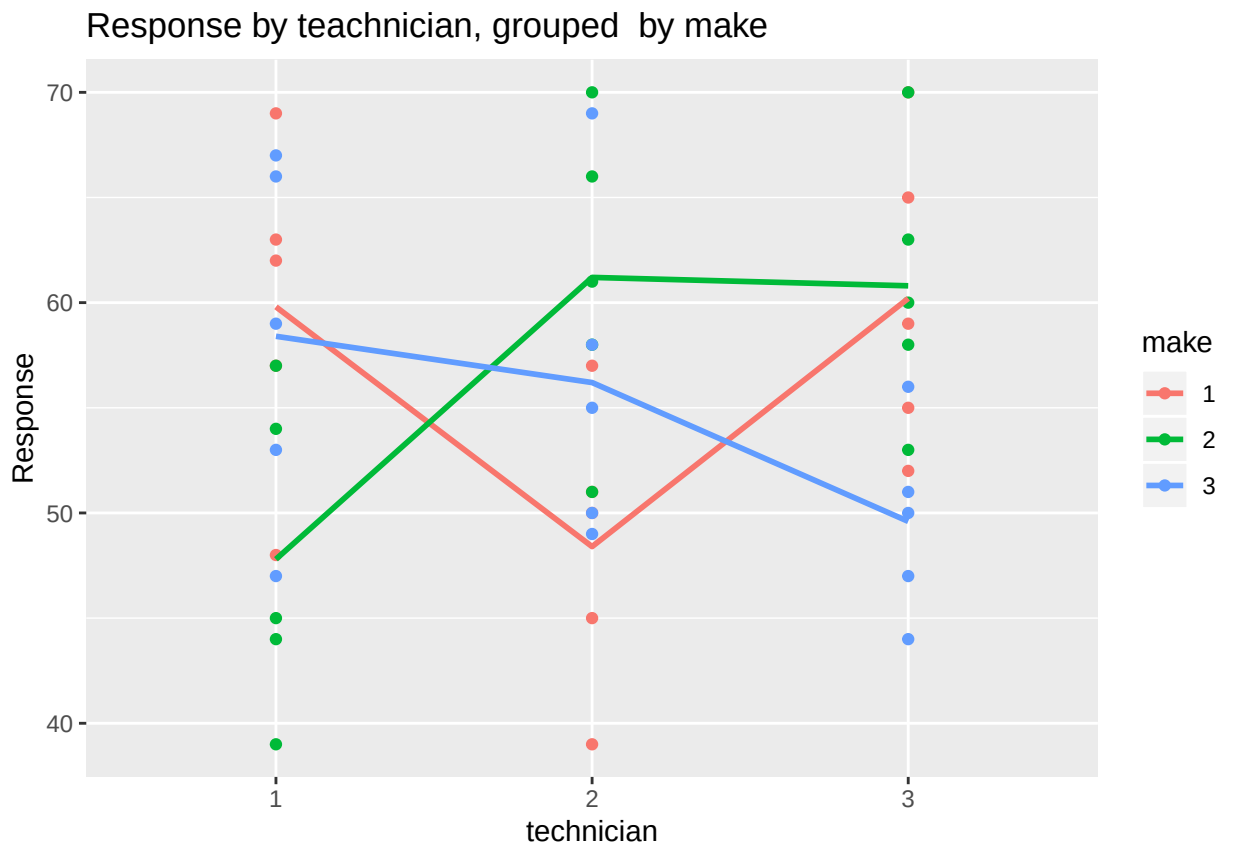
```
Disk = read.table(url(
  "https://raw.githubusercontent.com/npmldabook/Stat3119/master/Week-13/CH25PR16-2factorMEM.txt"))
names(Disk) = c("Response", "technician", "make", "units")
Disk$make = as.factor(Disk$make)
Disk$technician = as.factor(Disk$technician)

str(Disk)
```

```
## 'data.frame':    45 obs. of  4 variables:
##  $ Response  : num  62 48 63 57 69 57 45 39 54 44 ...
##  $ technician: Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1 1 1 1 ...
##  $ make      : Factor w/ 3 levels "1","2","3": 1 1 1 1 1 2 2 2 2 2 ...
##  $ units     : int  1 2 3 4 5 1 2 3 4 5 ...
```

```
# use ggplot and group by machine
ggplot(Disk, aes(x =make , y = Response , group = technician , col = technician )) +
  geom_point() + stat_summary(fun.y = mean, geom = "line", size=1)+
   labs(title="Response by make , grouped by teachnician")
```




```
ggplot(Disk, aes(x =technician , y = Response , group = make , col =make  )) +
  geom_point() + stat_summary(fun.y = mean, geom = "line", size=1)+
    labs(title="Response by teachnician, grouped  by make")
```



Both interactin plots show lack of parallelism, suggesting possible interaction.

(a, c, d) ANOVA test

```
summary(aov(Response~technician*make  , data=Disk ))
```

```
##                Df Sum Sq Mean Sq F value   Pr(>F)
## technician      2   24.6   12.29   0.236 0.790779
## make            2   28.3   14.16   0.272 0.763283
## technician:make  4 1215.3  303.82   5.841 0.000994 ***
## Residuals       36 1872.4   52.01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results: (a-c) For Interaction, F-test statistic is the same= 5.84, p=0.00099, can't reject H0, significant interaction. For random factor -technician, F-test statistic is the same =0.24, p=0.79 , can't reject H0: $\sigma_a^2 = 0$.

```
  MSB= 14.16
  MSAB= 303.82

# For fixed factor B
 (F.statB = MSB/MSAB)
```

```
## [1] 0.04660654
```

```r
# p-value for testing factor B
 (pf.B= 1- pf(F.statB, 2,4 ))
```

```
## [1] 0.9549734
```

Results: b) For fixed factor- make, F*=0.047, p=0.96, can't reject H0: all β_j are the same.

(b) estimate σ^2_{ab}

```r
fit.Disk <- lmer(Response ~ make + (1 | technician) + (1 | make:technician), data = Disk)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.00208956
## (tol = 0.002, component 1)
```

```r
summary(fit.Disk)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Response ~ make + (1 | technician) + (1 | make:technician)
##    Data: Disk
##
## REML criterion at convergence: 301.6
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.57344 -0.74154 -0.02433  0.70806  1.82596
##
## Random effects:
##  Groups          Name        Variance  Std.Dev.
##  make:technician (Intercept) 3.089e+01 5.558117
##  technician      (Intercept) 1.046e-06 0.001023
##  Residual                    5.202e+01 7.212348
## Number of obs: 45, groups:  make:technician, 9; technician, 3
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  56.1333     3.7102  15.130
## make2         0.4667     5.2470   0.089
## make3        -1.4000     5.2470  -0.267
##
## Correlation of Fixed Effects:
##       (Intr) make2
## make2 -0.707
## make3 -0.707  0.500
## convergence code: 0
## Model failed to converge with max|grad| = 0.00208956 (tol = 0.002, component 1)
```

Results: Estimated σ^2_{ab}= 30.89, Residual σ^2= 52, σ^2_{ab} is slightly smaller.