Grass Example: Random Effects One-Way

Six grass samples were randomly randomly selected from a large area. (We have the idea that each sample represents a genotype that exists in the area.) We then select 5 tillers from each grass sample and measure the chlorophyll content on each tiller.

For illustration, will first consider the fixed effects models. Note that lm() is used to fit the fixed effects (one-way ANOVA) model.

But since grasses were randomly selected, the random effects model is more appropriate. Note that lmer() from the lme4 package is used to fit the random effects model.

Also for illustration, REML and ML estimation are considered. In practice, default REML estimation is preferred.

```
library(emmeans)
Grasses <- read.csv("C:/hess/STAT512/RNotes/Random1/R1_Grasses_1way.csv")
str(Grasses)
```

```
## 'data.frame':    30 obs. of  2 variables:
##  $ grass: int  1 1 1 1 1 2 2 2 2 2 ...
##  $ y    : num  3.94 4.97 3.7 4.66 3.5 4.13 6.47 4.91 3.63 6.4 ...
```

```
#Important: Need to define grass as.factor!!!!
Grasses$grass <- as.factor(Grasses$grass)
```

FIXED Effects (one-way ANOVA) Model (for Comparison)

```
Model1 <- lm(y ~ grass, data = Grasses)
anova(Model1)
```

```
## Analysis of Variance Table
##
## Response: y
##           Df  Sum Sq Mean Sq F value Pr(>F)
## grass      5  4.9137 0.98275  1.3184 0.2898
## Residuals 24 17.8896 0.74540
```

```
summary(Model1)
```

```
##
## Call:
## lm(formula = y ~ grass, data = Grasses)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4860 -0.6195 -0.0390  0.5120  1.3620
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   4.1540     0.3861  10.759 1.16e-10 ***
## grass2        0.9540     0.5460   1.747   0.0934 .
## grass3        0.3040     0.5460   0.557   0.5829
## grass4        1.1620     0.5460   2.128   0.0438 *
```

```
## grass5        0.4420     0.5460   0.809   0.4262
## grass6        0.8680     0.5460   1.590   0.1250
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8634 on 24 degrees of freedom
## Multiple R-squared:  0.2155, Adjusted R-squared:  0.05204
## F-statistic: 1.318 on 5 and 24 DF,  p-value: 0.2898
```

```
emout <- emmeans(Model1, ~grass)
emout
```

```
##  grass emmean    SE df lower.CL upper.CL
##  1       4.15 0.386 24     3.36     4.95
##  2       5.11 0.386 24     4.31     5.90
##  3       4.46 0.386 24     3.66     5.25
##  4       5.32 0.386 24     4.52     6.11
##  5       4.60 0.386 24     3.80     5.39
##  6       5.02 0.386 24     4.23     5.82
##
## Confidence level used: 0.95
```

```
emmeans <- summary(emout)$emmean
```

RANDOM Effects Model (using lmer)

The notation (1|grass) specifies that grass be treated as a random effect. REML estimation is used by default. Because we have a random effects model (with no fixed effects), no ANOVA table is returned. The rand() function gives tests of variance components. The ranef() function can be used to calculate "blups".

```
library(lme4)
library(pbkrtest)
library(lmerTest)
Model2 <- lmer(y ~ (1|grass), data = Grasses)
summary(Model2)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: y ~ (1 | grass)
##    Data: Grasses
##
## REML criterion at convergence: 78.6
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.87901 -0.75344  0.05554  0.49831  1.86951
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  grass    (Intercept) 0.04747  0.2179
##  Residual             0.74540  0.8634
## Number of obs: 30, groups:  grass, 6
##
## Fixed effects:
```

```
##             Estimate Std. Error    df t value Pr(>|t|)
## (Intercept)    4.776      0.181 5.000   26.39 1.46e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(Model2)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##      Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
```

```
rand(Model2)
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## y ~ (1 | grass)
##             npar  logLik    AIC     LRT Df Pr(>Chisq)
## <none>         3 -39.280 84.561
## (1 | grass)    2 -39.364 82.728 0.16776  1     0.6821
```

```
ranef(Model2)
```

```
## $grass
##   (Intercept)
## 1 -0.15014125
## 2  0.08026318
## 3 -0.07672097
## 4  0.13049811
## 5 -0.04339203
## 6  0.05949297
##
## with conditional variances for "grass"
```

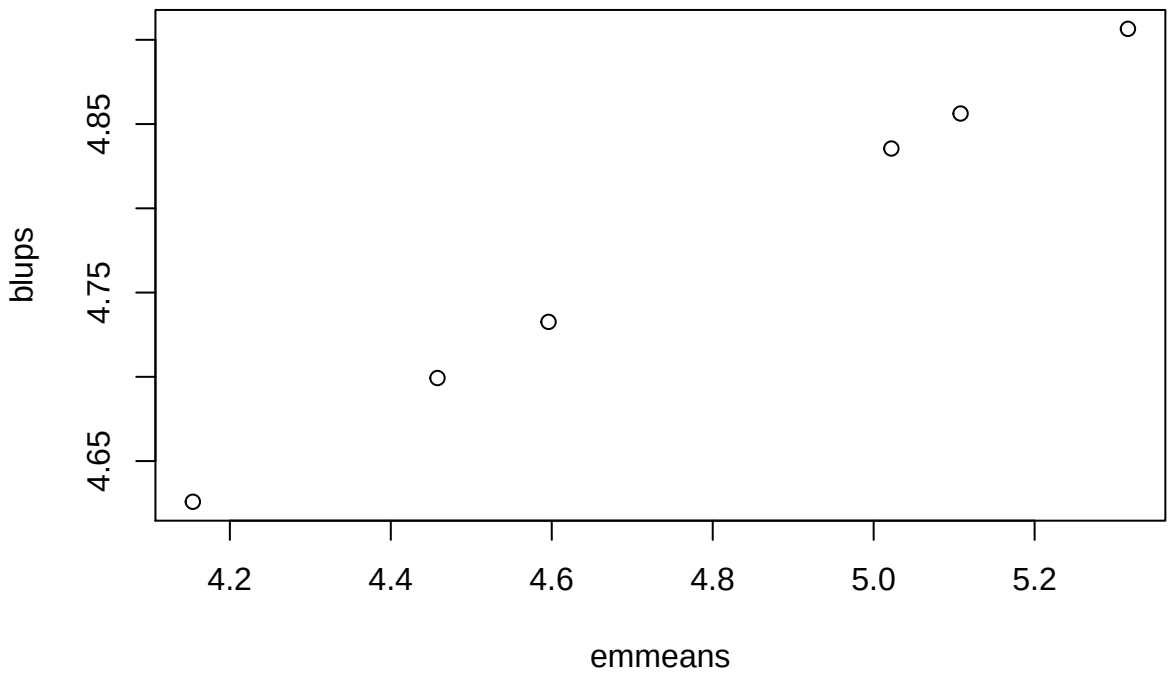
```
blups <- ranef(Model2)$grass + 4.776
```

Compare emmeans and blups

```
Temp <- data.frame(emmeans, blups)
colnames(Temp) <- c("emmeans","blups")
Temp
```

```
##   emmeans    blups
## 1   4.154 4.625859
## 2   5.108 4.856263
## 3   4.458 4.699279
## 4   5.316 4.906498
## 5   4.596 4.732608
## 6   5.022 4.835493
```

```r
plot(blups ~ emmeans, data = Temp)
```



RANDOM Effects Model (using lmer with ML estimation)

Primarily for Illustration. Default REML estimation is standard. Note that the estimated variance component associated with grass is different detpending on whether REML or ML estimation is used.

```r
Model3 <- lmer(y ~ (1|grass), data = Grasses, REML = FALSE)
summary(Model3)
```

```
## Linear mixed model fit by maximum likelihood . t-tests use
##   Satterthwaite's method [lmerModLmerTest]
## Formula: y ~ (1 | grass)
##    Data: Grasses
##
##      AIC      BIC   logLik deviance df.resid
##     82.9     87.1    -38.4     76.9       27
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.91058 -0.76284  0.07823  0.46725  1.92789
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  grass    (Intercept) 0.01472  0.1213
```

```
##  Residual             0.74539  0.8634
## Number of obs: 30, groups:  grass, 6
##
## Fixed effects:
##             Estimate Std. Error     df t value Pr(>|t|)
## (Intercept)   4.7757     0.1652 6.0010    28.9 1.13e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(Model3)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##      Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
```