

# Assignment8 KEY

40 points total, 2 points per problem part unless otherwise noted.

```
library(dplyr)
library(ggplot2)
library(car)
library(emmeans)
InData <- read.csv("~/Dropbox/STAT512/Assigns/Assign8/Biomass.csv")
str(InData)

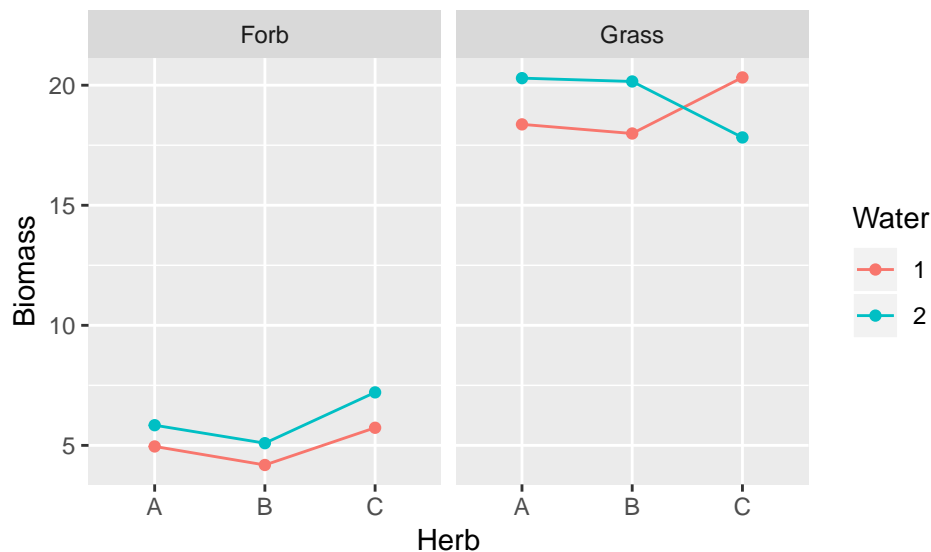
## 'data.frame': 36 obs. of 4 variables:
## $ Type : Factor w/ 2 levels "Forb","Grass": 2 2 2 2 2 2 2 2 2 2 ...
## $ Herb : Factor w/ 3 levels "A","B","C": 1 1 1 1 1 1 2 2 2 2 ...
## $ Water : int 1 1 1 2 2 2 1 1 1 2 ...
## $ Biomass: num 16.4 19.9 18.9 18.2 21.2 ...

InData$Water <- as.factor(InData$Water)
```

## 1. Summary Graph (4pts)

```
SumStats <- summarize(group_by(InData, Type, Herb, Water),
                        n = n(),
                        Biomass = mean(Biomass),
                        sd = sd(Biomass))

#SumStats
qplot(x = Herb, y = Biomass, colour = Water, group = Water, data = SumStats) +
  geom_line() +
  facet_grid(. ~ Type)
```



## 2. Three-way ANOVA table

```
options(contrasts=c("contr.sum","contr.poly"))
Model1 <- lm(Biomass ~ Type*Herb*Water, data = InData)
Anova(Model1, type = 3)
```

```
## Anova Table (Type III tests)
##
```

```
## Response: Biomass
##              Sum Sq Df    F value    Pr(>F)
## (Intercept)  5473.8  1 5539.1323 < 2.2e-16 ***
## Type        1678.7  1 1698.7721 < 2.2e-16 ***
## Herb         5.1    2   2.5644  0.097872 .
## Water        5.9    1   5.9711  0.022271 *
## Type:Herb     5.4    2   2.7185  0.086259 .
## Type:Water    0.7    1   0.7092  0.408021
## Herb:Water    7.9    2   3.9973  0.031742 *
## Type:Herb:Water 13.1  2   6.6440  0.005055 **
## Residuals    23.7  24
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

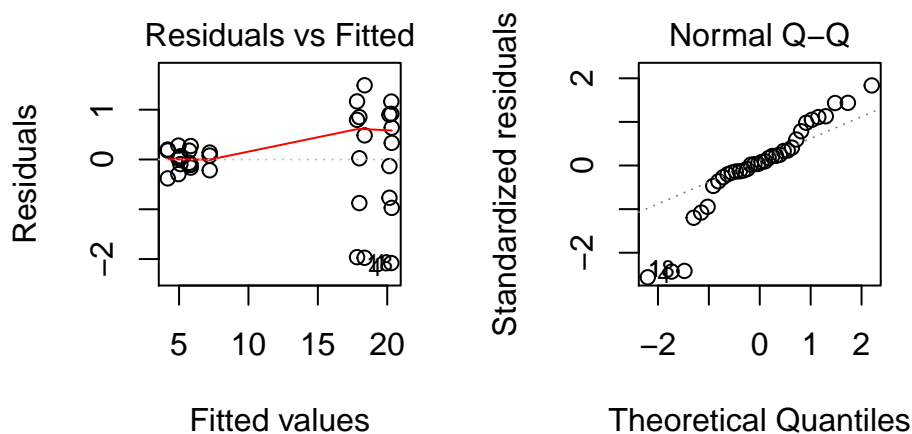
### 3. Three-way Diagnostic plots

**Note:** Diagnostic plots not required for full credit, but shown here for completeness.

Based on the plot of Resids vs Fitted, there is strong evidence of unequal variance.

(Based on the QQplot, there is also some evidence against normality. But this is actually driven by the unequal variance.)

```
par(mfrow=c(1,2))
plot(Model1, which = c(1:2))
```



### 4. Three-way emmeans

```
emout1 <- emmeans(Model1, pairwise ~ Water|Herb*Type)
emout1$contrasts
```

```
## Herb = A, Type = Forb:
## contrast estimate      SE df t.ratio p.value
## 1 - 2    -0.8848361 0.8116673 24  -1.090  0.2865
##
## Herb = B, Type = Forb:
## contrast estimate      SE df t.ratio p.value
## 1 - 2    -0.9084060 0.8116673 24  -1.119  0.2741
##
## Herb = C, Type = Forb:
## contrast estimate      SE df t.ratio p.value
## 1 - 2    -1.4730603 0.8116673 24  -1.815  0.0821
##
## Herb = A, Type = Grass:
## contrast estimate      SE df t.ratio p.value
```

```
## 1 - 2    -1.9244287 0.8116673 24    -2.371    0.0261
##
## Herb = B, Type = Grass:
## contrast      estimate          SE df t.ratio p.value
## 1 - 2         -2.1667163 0.8116673 24    -2.669    0.0134
##
## Herb = C, Type = Grass:
## contrast      estimate          SE df t.ratio p.value
## 1 - 2          2.4991714 0.8116673 24     3.079    0.0051
```

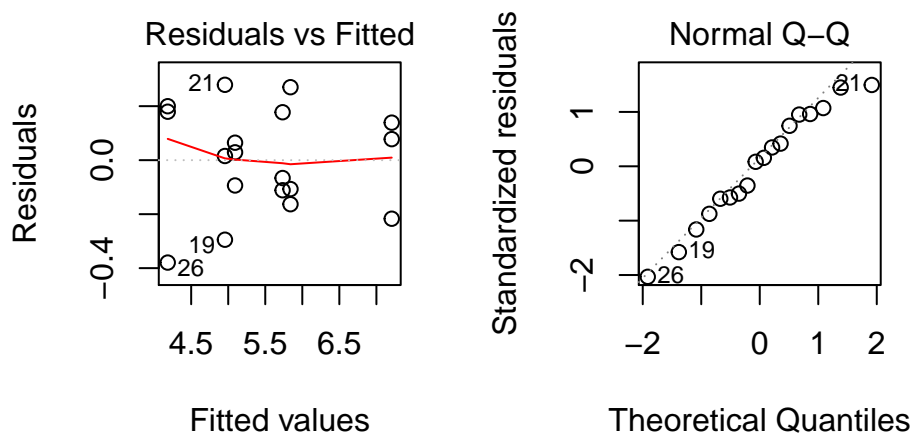
#### 5. FORB Two-way ANOVA table

```
Model2 <- lm(Biomass ~ Herb*Water, data = InData[InData$Type == "Forb",])
Anova(Model2, type = 3)
```

```
## Anova Table (Type III tests)
##
## Response: Biomass
##              Sum Sq Df    F value    Pr(>F)
## (Intercept) 544.92   1 10432.1562 < 2.2e-16 ***
## Herb         10.17   2    97.3895 3.820e-08 ***
## Water         5.33   1   102.1229 3.197e-07 ***
## Herb:Water    0.33   2     3.1846 0.07772 .
## Residuals    0.63  12
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

#### 6. FORB diagnostic plots look much better! Some evidence of unequal variance, but not severe.

```
par(mfrow = c(1,2))
plot(Model2, which = c(1:2))
```



#### 7. FORB emmeans #1 (interaction comparisons)

```
emout2 <- emmeans(Model2, pairwise ~ Water|Herb)
emout2
```

```
## $emmeans
## Herb = A:
## Water    emmean      SE df lower.CL upper.CL
## 1        4.955550 0.131953 12 4.668049 5.243051
## 2        5.840386 0.131953 12 5.552885 6.127887
##
```

```
## Herb = B:
##   Water    emmean      SE df lower.CL upper.CL
##   1      4.183644 0.131953 12 3.896143 4.471145
##   2      5.092050 0.131953 12 4.804549 5.379551
##
## Herb = C:
##   Water    emmean      SE df lower.CL upper.CL
##   1      5.734037 0.131953 12 5.446536 6.021538
##   2      7.207097 0.131953 12 6.919597 7.494598
##
## Confidence level used: 0.95
##
## $contrasts
## Herb = A:
##   contrast    estimate      SE df t.ratio p.value
##   1 - 2      -0.8848361 0.1866097 12  -4.742  0.0005
##
## Herb = B:
##   contrast    estimate      SE df t.ratio p.value
##   1 - 2      -0.9084060 0.1866097 12  -4.868  0.0004
##
## Herb = C:
##   contrast    estimate      SE df t.ratio p.value
##   1 - 2      -1.4730603 0.1866097 12  -7.894  <.0001
```

#### 8. FORB LSD #1 (interaction comparisons)

```
n = 3 #for each treatment combination mean
qt(0.975, df = 12)*sqrt(2*(0.63/12)/3)
```

```
## [1] 0.4076186
```

```
qt(0.975, df = 12)*0.187
```

```
## [1] 0.407438
```

```
#or peeling MSResid out of ANOVA table with R
#numerator is SSResid, denominator is df resid from ANOVA table
MSResid=(Anova(Model2, type = 3)[[1]][5])/(Anova(Model2, type = 3)[[2]][5])
MSResid
```

```
## [1] 0.05223477
```

```
qt(.975,12)*sqrt(2*MSResid/n)
```

```
## [1] 0.4065876
```

#### 9. FORB emmeans #2 (main effect comparison)

```
emout3 <- emmeans(Model2, pairwise ~ Water)
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```
emout3$contrasts
```

```
##   contrast    estimate      SE df t.ratio p.value
##   1 - 2      -1.088768 0.1077392 12 -10.106  <.0001
##
## Results are averaged over the levels of: Herb
```

10. FORB LSD #2 (main effect comparison)

```
n=9 #because we are comparing Water across 3 levels of Herb (each with 3)
qt(0.975, df = 12)*sqrt(2*(0.63/12)/9)
```

```
## [1] 0.2353387
```

```
qt(0.975, df = 12)*0.108
```

```
## [1] 0.2353118
```

```
#or this method using same MSResid from ANOVA
qt(.975,12)*sqrt(2*MSResid/n)
```

```
## [1] 0.2347435
```

11. The power is higher for the main effect comparison (#10) because the LSD (ME) is smaller.

12. GRASS Two-way ANOVA table

```
Model3 <- lm(Biomass ~ Herb*Water, data = InData[InData$Type == "Grass",])
Anova(Model3, type =3)
```

```
## Anova Table (Type III tests)
```

```
##
```

```
## Response: Biomass
```

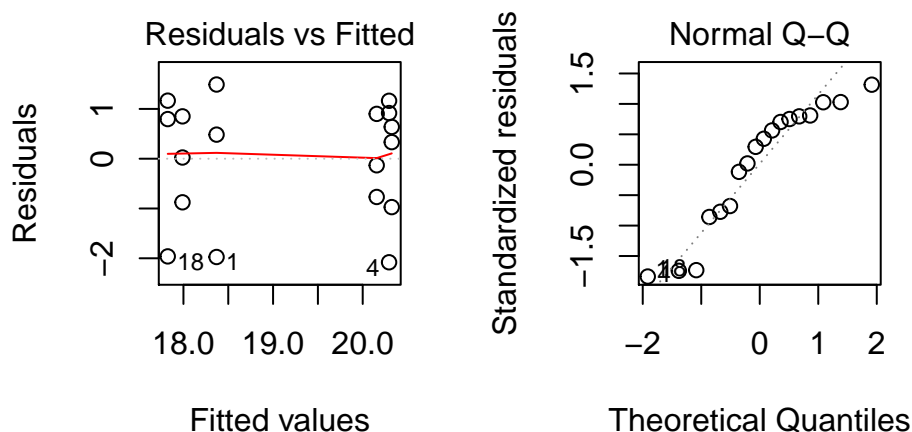
```
##          Sum Sq Df    F value    Pr(>F)
## (Intercept) 6607.6  1 3433.9972 4.029e-16 ***
## Herb          0.3   2   0.0694   0.9333
## Water        1.3   1   0.6586   0.4329
## Herb:Water    20.7  2   5.3786   0.0215 *
## Residuals    23.1 12
```

```
## ---
```

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

13. GRASS diagnostic plots look much better! Some evidence of skew (non-normality), but not severe.

```
par(mfrow = c(1,2))
plot(Model3, which = c(1:2))
```



14. GRASS emmeans (interaction comparisons)

```
emout4 <- emmeans(Model3, pairwise ~ Water|Herb)
emout4$contrasts
```

```
## Herb = A:
```

```
## contrast estimate      SE df t.ratio p.value
## 1 - 2      -1.924429 1.132601 12  -1.699  0.1150
##
## Herb = B:
## contrast estimate      SE df t.ratio p.value
## 1 - 2      -2.166716 1.132601 12  -1.913  0.0799
##
## Herb = C:
## contrast estimate      SE df t.ratio p.value
## 1 - 2       2.499171 1.132601 12   2.207  0.0476
```

15. No, because there is evidence of an interaction between Water and Herb. In particular, the estimated difference in mean response for Water 1 vs 2 changes sign when comparing across Herb. When there is evidence of an interaction, it does not make sense to look at main effects.
16. (**4pts**) The estimated differences are the SAME.  
The SEs are DIFFERENT (Combined SE = 0.812, Forb SE = 0.187, Grass SE = 1.13).  
The df is DIFFERENT (combined df = 24, split df = 12).
17. (1) The combined analysis showed strong evidence of unequal variance. The model assumptions are better satisfied using the split analyses. (2) The combined analysis showed evidence of a three-way interaction.
18. (1) When we split the analysis we reduce the df (hence reducing power). (2) When we split the analysis we cannot get direct comparisons of Forb versus Grass.