Media Example: Unbalanced RCB

An experiment to compare the growth of snapdragons in various growing media. Response variable is stem length. There are seven growing media (t = 7) randomly assigned in three blocks (b = 3). But with 2 missing observations for a total of n=19 observations.

Notes:

1. Analysis is the same as the balanced case, just remember to use Anova(, type =3) for tests and emmeans() for estimates and pairwise comparisons.
2. The analysis would be the same if the observations with missing values were deleted. We just include them here to easily calculate predicted values.

```
library(car)
library(emmeans)
Media <- read.csv("C:/hess/STAT512/RNotes/ExpDesign1/ED1_MediaRCB.csv")
str(Media)
```

```
## 'data.frame':    21 obs. of  3 variables:
##  $ blk   : int  1 1 1 1 1 1 1 2 2 2 ...
##  $ media : Factor w/ 7 levels "CLARION","CLINTON",..: 1 2 4 5 3 6 7 1 2 4 ...
##  $ stemln: num  NA 32.1 35.7 36 31.8 38.2 32.5 32.3 NA 35.9 ...
```

```
#Important: Need to define block as.factor!
Media$blk <- as.factor(Media$blk)
Media
```

```
##    blk   media stemln
## 1    1 CLARION     NA
## 2    1 CLINTON   32.1
## 3    1    KNOX   35.7
## 4    1  ONEILL   36.0
## 5    1 COMPOST   31.8
## 6    1  WABASH   38.2
## 7    1 WEBSTER   32.5
## 8    2 CLARION   32.3
## 9    2 CLINTON     NA
## 10   2    KNOX   35.9
## 11   2  ONEILL   34.2
## 12   2 COMPOST   28.0
## 13   2  WABASH   37.8
## 14   2 WEBSTER   31.1
## 15   3 CLARION   31.5
## 16   3 CLINTON   29.1
## 17   3    KNOX   33.1
## 18   3  ONEILL   31.2
## 19   3 COMPOST   29.2
## 20   3  WABASH   31.9
## 21   3 WEBSTER   29.7
```

```
aggregate(stemln ~ media, FUN = mean, data = Media)
```

```
##     media   stemln
## 1 CLARION 31.90000
## 2 CLINTON 30.60000
```

```
## 3 COMPOST 29.66667
## 4    KNOX 34.90000
## 5  ONEILL 33.80000
## 6  WABASH 35.96667
## 7 WEBSTER 31.10000
```

```
Model1 <- lm(stemln ~ blk + media, data = Media)
Anova(Model1, type = 3)
```

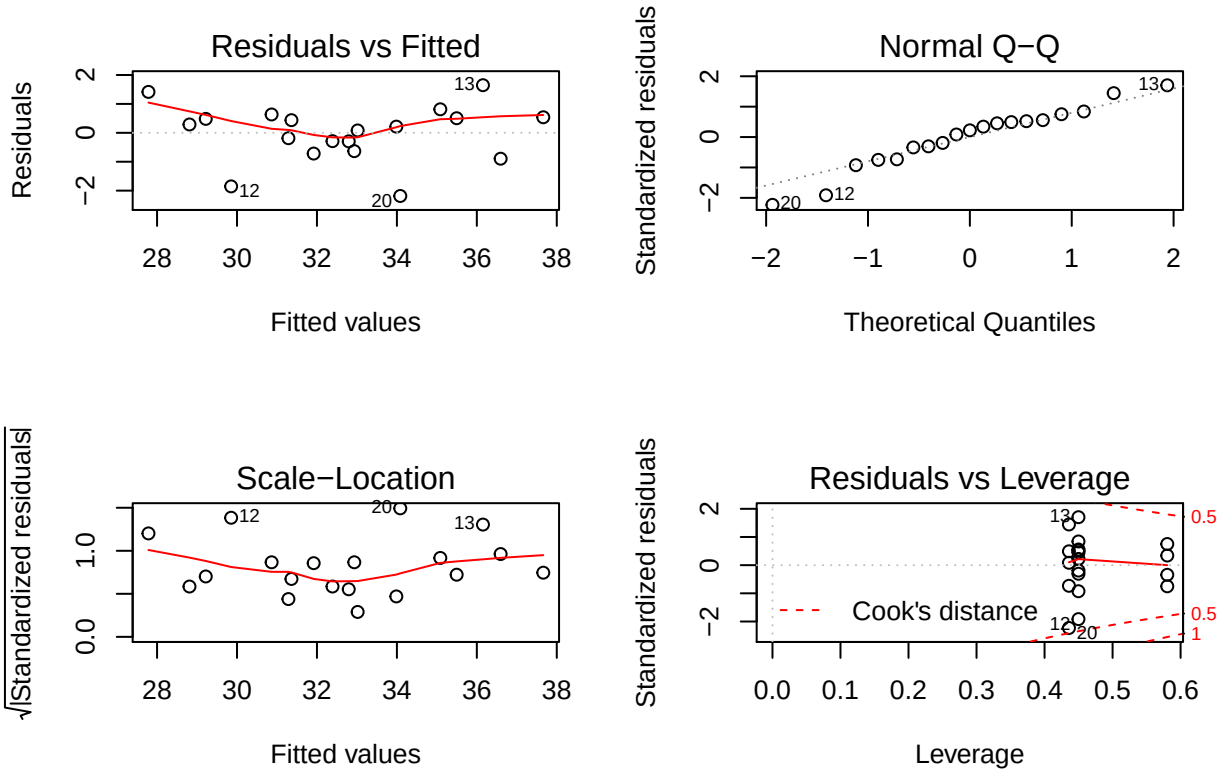
```
## Anova Table (Type III tests)
##
## Response: stemln
##              Sum Sq Df F value    Pr(>F)
## (Intercept) 1556.40  1 914.723 3.656e-11 ***
## blk           40.80  2  11.989  0.002208 **
## media         92.85  6   9.095  0.001420 **
## Residuals     17.01 10
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
emmeans(Model1, pairwise ~ media)
```

```
## $emmeans
##  media     emmean        SE df lower.CL upper.CL
##  CLARION 32.74779 0.9488549 10 30.63361 34.86197
##  CLINTON 30.69324 0.9488549 10 28.57906 32.80742
##  COMPOST 29.66667 0.7531039 10 27.98865 31.34469
##  KNOX    34.90000 0.7531039 10 33.22198 36.57802
##  ONEILL  33.80000 0.7531039 10 32.12198 35.47802
##  WABASH  35.96667 0.7531039 10 34.28865 37.64469
##  WEBSTER 31.10000 0.7531039 10 29.42198 32.77802
##
## Results are averaged over the levels of: blk
## Confidence level used: 0.95
##
## $contrasts
##  contrast            estimate       SE df t.ratio p.value
##  CLARION - CLINTON  2.0545455 1.362416 10   1.508  0.7357
##  CLARION - COMPOST  3.0811189 1.211400 10   2.543  0.2393
##  CLARION - KNOX    -2.1522145 1.211400 10  -1.777  0.5881
##  CLARION - ONEILL  -1.0522145 1.211400 10  -0.869  0.9698
##  CLARION - WABASH  -3.2188811 1.211400 10  -2.657  0.2049
##  CLARION - WEBSTER  1.6477855 1.211400 10   1.360  0.8102
##  CLINTON - COMPOST  1.0265734 1.211400 10   0.847  0.9731
##  CLINTON - KNOX    -4.2067599 1.211400 10  -3.473  0.0626
##  CLINTON - ONEILL  -3.1067599 1.211400 10  -2.565  0.2325
##  CLINTON - WABASH  -5.2734266 1.211400 10  -4.353  0.0169
##  CLINTON - WEBSTER -0.4067599 1.211400 10  -0.336  0.9998
##  COMPOST - KNOX    -5.2333333 1.065050 10  -4.914  0.0075
##  COMPOST - ONEILL  -4.1333333 1.065050 10  -3.881  0.0340
##  COMPOST - WABASH  -6.3000000 1.065050 10  -5.915  0.0019
##  COMPOST - WEBSTER -1.4333333 1.065050 10  -1.346  0.8170
##  KNOX - ONEILL      1.1000000 1.065050 10   1.033  0.9342
##  KNOX - WABASH     -1.0666667 1.065050 10  -1.002  0.9424
##  KNOX - WEBSTER     3.8000000 1.065050 10   3.568  0.0543
```

```
##  ONEILL - WABASH   -2.1666667 1.065050 10  -2.034  0.4507
##  ONEILL - WEBSTER   2.7000000 1.065050 10   2.535  0.2420
##  WABASH - WEBSTER   4.8666667 1.065050 10   4.569  0.0123
##
## Results are averaged over the levels of: blk
## P value adjustment: tukey method for comparing a family of 7 estimates
```

```
par(mfrow=c(2,2))
plot(Model1)
```



For Illustration: We calculate predicted values and look at the parameter estimate information. Not typically of interest.

```
summary(Model1)
```

```
##
## Call:
## lm(formula = stemln ~ blk + media, data = Media)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.1846 -0.4649  0.2135  0.5211  1.6468
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  34.4434     1.1388  30.244 3.66e-11 ***
## blk2         -1.5091     0.7866  -1.919 0.084014 .
## blk3         -3.5776     0.7398  -4.836 0.000686 ***
```

```
## mediaCLINTON  -2.0545     1.3624  -1.508 0.162475
## mediaCOMPOST  -3.0811     1.2114  -2.543 0.029191 *
## mediaKNOX      2.1522     1.2114   1.777 0.106005
## mediaONEILL    1.0522     1.2114   0.869 0.405433
## mediaWABASH    3.2189     1.2114   2.657 0.024016 *
## mediaWEBSTER  -1.6478     1.2114  -1.360 0.203629
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.304 on 10 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.8891, Adjusted R-squared:  0.8003
## F-statistic: 10.02 on 8 and 10 DF,  p-value: 0.0007043
```

```
Out <- data.frame(Media, Yhat = predict(Model1, newdata = Media))
Out
```

```
##    blk   media stemln     Yhat
## 1    1 CLARION     NA 34.44336
## 2    1 CLINTON   32.1 32.38881
## 3    1    KNOX   35.7 36.59557
## 4    1  ONEILL   36.0 35.49557
## 5    1 COMPOST   31.8 31.36224
## 6    1  WABASH   38.2 37.66224
## 7    1 WEBSTER   32.5 32.79557
## 8    2 CLARION   32.3 32.93427
## 9    2 CLINTON     NA 30.87972
## 10   2    KNOX   35.9 35.08648
## 11   2  ONEILL   34.2 33.98648
## 12   2 COMPOST   28.0 29.85315
## 13   2  WABASH   37.8 36.15315
## 14   2 WEBSTER   31.1 31.28648
## 15   3 CLARION   31.5 30.86573
## 16   3 CLINTON   29.1 28.81119
## 17   3    KNOX   33.1 33.01795
## 18   3  ONEILL   31.2 31.91795
## 19   3 COMPOST   29.2 27.78462
## 20   3  WABASH   31.9 34.08462
## 21   3 WEBSTER   29.7 29.21795
```

For Illustration: The only difference between Model1 and Model2 is the order of the terms. We use these models to examine difference between Type1 anova() and Type3 Anova() tests. Due to imblance (missing data), these tests do NOT match. We get different Type 1 tests depending on ordering of terms. Type 3 tests do not depend on ordering of terms. Type 3 tests preferred!

```
Model2 <- lm(stemln ~ media + blk, data = Media)
anova(Model1)
```

```
## Analysis of Variance Table
##
## Response: stemln
##           Df Sum Sq Mean Sq F value   Pr(>F)
## blk        2 43.495 21.7474  12.781 0.001758 **
## media      6 92.850 15.4750   9.095 0.001420 **
## Residuals 10 17.015  1.7015
## ---
```

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

```
anova(Model2)
```

```
## Analysis of Variance Table
##
## Response: stemln
##           Df Sum Sq Mean Sq F value   Pr(>F)
## media      6 95.547 15.9244  9.3591 0.001267 **
## blk        2 40.798 20.3992 11.9890 0.002208 **
## Residuals 10 17.015  1.7015
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Anova(Model2, type = 3)
```

```
## Anova Table (Type III tests)
##
## Response: stemln
##              Sum Sq Df F value    Pr(>F)
## (Intercept) 1556.40  1 914.723 3.656e-11 ***
## media         92.85  6   9.095  0.001420 **
## blk           40.80  2  11.989  0.002208 **
## Residuals     17.01 10
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```