Anova

Mark Andrews

July 5, 2018

```
library(dplyr)
library(pander)
library(ggplot2)
library(readr)
library(lme4)
```

One-way Anova

```
data(PlantGrowth)
M <- aov(weight ~ group, data=PlantGrowth)
pander(M)
```

Table 1: Analysis of Variance Model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	2	3.766	1.883	4.846	0.01591
Residuals	27	10.49	0.3886	NA	NA

We can do Tukey's range test to perform multiple comparisons:

```
TukeyHSD(M)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = weight ~ group, data = PlantGrowth)
##
## $group
##             diff        lwr       upr     p adj
## trt1-ctrl -0.371 -1.0622161 0.3202161 0.3908711
## trt2-ctrl  0.494 -0.1972161 1.1852161 0.1979960
## trt2-trt1  0.865  0.1737839 1.5562161 0.0120064
```

Note that we can also we can do Anova using `lm()`:

```
M <- lm(weight ~ group, data=PlantGrowth)
anova(M)
```

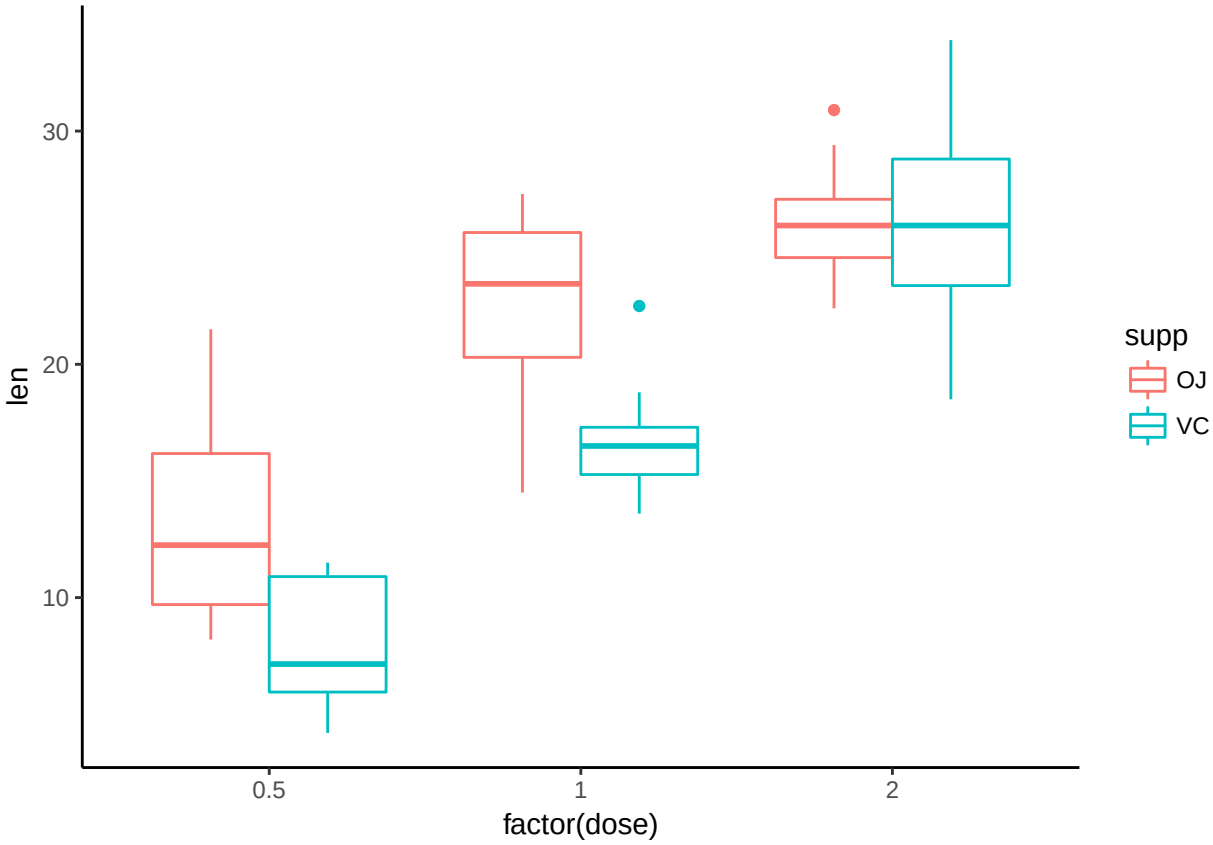
```
## Analysis of Variance Table
##
## Response: weight
##           Df  Sum Sq Mean Sq F value  Pr(>F)
## group      2  3.7663  1.8832  4.8461 0.01591 *
## Residuals 27 10.4921  0.3886
## ---
```

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

Two-way anova

```
data("ToothGrowth")

ggplot(ToothGrowth,
       aes(x = factor(dose), y = len, col = supp)) +
  geom_boxplot() +
  theme_classic()
```




```
M <- aov(len ~ supp*dose, data=ToothGrowth)
```

Repeated measures Anova

Oneway

```
Df <- read_table('../data/recall_data.txt')
```

```
## Parsed with column specification:
## cols(
##   Observation = col_integer(),
##   Subject = col_character(),
```

```
##   Valence = col_character(),
##   Recall = col_integer()
## )
```

```
M <- aov(Recall ~ Valence + Error(Subject/Valence), data=Df)
pander(M)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Residuals	4	105.1	26.27	NA	NA
Valence	2	2030	1015	189.1	1.841e-07
Residuals1	8	42.93	5.367	NA	NA

Multiple comparisons, with Bonferroni correction

```
with(Df,
     pairwise.t.test(x=Recall, g=Valence),
     p.adjust.methods='bonferroni',
     paired=T)
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  Recall and Valence
##
##     Neg     Neu
## Neu 1.9e-05 -
## Pos 0.00014 7.1e-08
##
## P value adjustment method: holm
```

Twoway

```
Df <- read_table('../data/recall_data2.txt')
```

```
## Parsed with column specification:
## cols(
##   Observation = col_integer(),
##   Subject = col_character(),
##   Task = col_character(),
##   Valence = col_character(),
##   Recall = col_integer()
## )
```

```
M <- aov(Recall ~ Valence*Task + Error(Subject/(Task*Valence)), data=Df)
pander(M)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Residuals	4	349.1	87.28	NA	NA
Task	1	30	30	7.347	0.05351
Residuals1	4	16.33	4.083	NA	NA
Valence	2	9.8	4.9	1.459	0.2883
Residuals2	8	26.87	3.358	NA	NA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Valence:Task	2	1.4	0.7	0.2907	0.7553
Residuals	8	19.27	2.408	NA	NA

Multilevel models

The above can be done using multilevel models too.

```
M <- lmer(Recall ~ Valence*Task + (1|Subject),
          data=Df)
```