ANOVA as Linear Models

Prof. Lauren Perry

ANOVA

What is ANOVA, really?

The one-way ANOVA we've seen is just a linear model with one categorical predictor:

$$y = \beta_0 + \beta_1 x + \epsilon$$

which can be written as

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

- μ is the baseline mean
- α_i is the effect of category i

Example

```r
anova(aov(weight ~ feed, chickwts))
```

Is exactly the same as

```r
anova(lm(weight ~ feed, chickwts))
```

```
## Analysis of Variance Table
##
## Response: weight
##           Df Sum Sq Mean Sq F value    Pr(>F)
## feed       5 231129   46226  15.365 5.936e-10 ***
## Residuals 65 195556    3009
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Two-Way ANOVA

We can extend the one-way ANOVA to other settings, such as the two-way ANOVA.

- This is a linear model with two categorical predictors, plus their interaction

$$y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ij}$$

- ANOVA adds predictors into the model using *sequential hypothesis testing*.

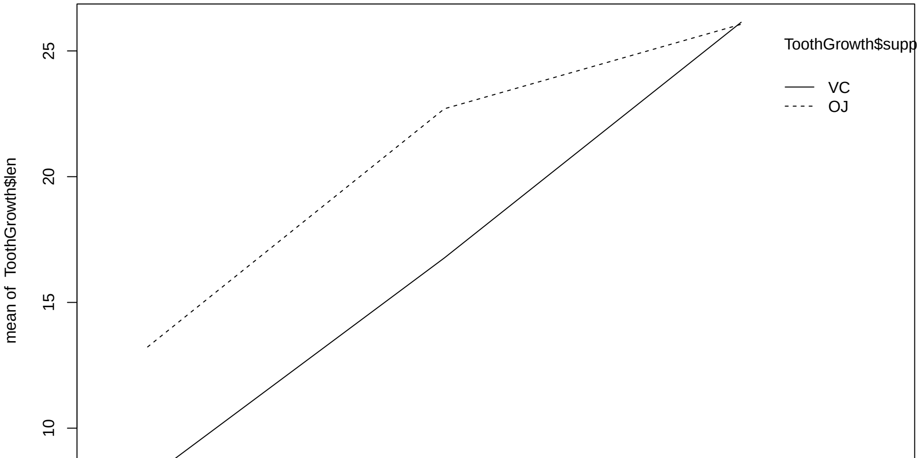
Two-Way ANOVA

```
anova(lm(len ~ supp*as.factor(dose), ToothGrowth))
```

```
## Analysis of Variance Table
##
## Response: len
##                      Df  Sum Sq Mean Sq F value    Pr(>F)
## supp                  1  205.35  205.35  15.572 0.0002312 ***
## as.factor(dose)       2 2426.43 1213.22  92.000 < 2.2e-16 ***
## supp:as.factor(dose)  2  108.32   54.16   4.107 0.0218603 *
## Residuals            54  712.11   13.19
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interaction Plot

```
interaction.plot(as.factor(ToothGrowth$dose), ToothGrowth$supp, ToothGrowtl
```



Sequential Hypothesis Testing

	Null Model	Alternative
1.	$y \sim 1$	$y \sim \text{supp}$
2.	$y \sim \text{supp}$	$y \sim \text{supp} + \text{dose}$
3.	$y \sim \text{supp} + \text{dose}$	$y \sim \text{supp} + \text{dose} + \text{interaction}$