Two Way ANOVA

Lecture 20

Alexandra Chronopoulou

ILLINOIS

COLLEGE OF LIBERAL ARTS & SCIENCES

Department of Statistics
101 Illini Hall, MC-374
725 S. Wright St.
Champaign, IL 61820-5710

Two factor experiments: Examples

- Investigate the effects of selling price (factor A) and type of promotional campaign (factor B) on sales of a product.
- Study the relationship between the response to three blood pressure lowering drug types (factor A) for hypertensive males and females (factor B).
- Investigate the type of fertilizer (factor A) and variety of corn (factor B) on the crop yield.

One factor at a time vs. Multi-factor designs

- Single-factor
  1. Do not explore the entire space of treatment combinations.
  2. Interactions cannot be estimated.
  3. Full randomization is not possible.
  4. Multiple stages increase complexity of the analysis.

- MultiFactor
  1. Efficient replication.
  2. Assessment of Interactions.
  3. Validity of Findings.

Rats Example

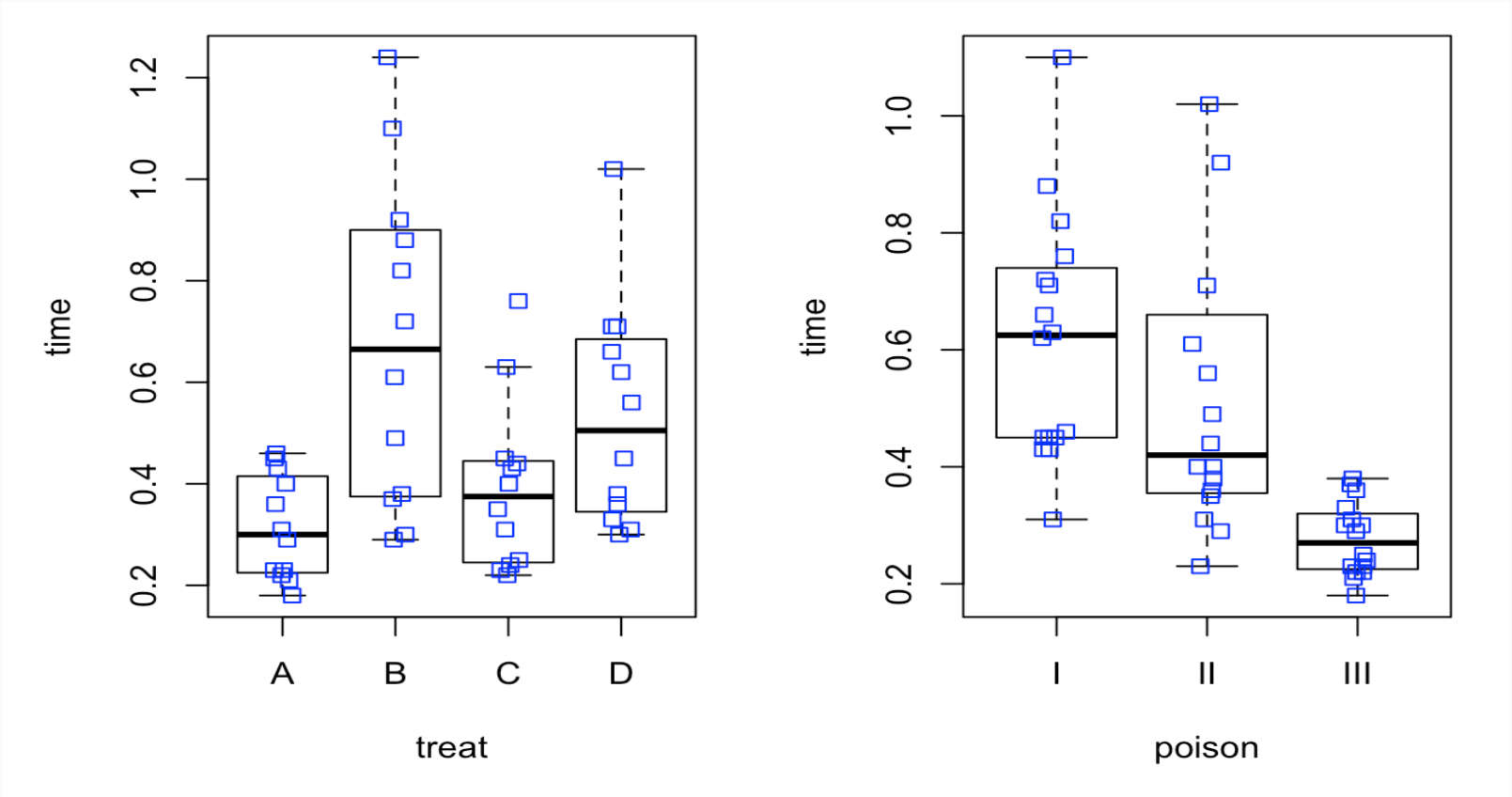
- The `rats` data set consists of the *survival times* (the response) of rats that are randomly allocated to 3 poisons (I, II, III) and four treatments (A, B, C, D).
- This is an experiment with 2 factors, poisons and treatments, each having a certain number of levels, 3 and 4 respectively.
- This is a *factorial structure* since each of the 3 poisons appears together with each of the 4 treatments (*crossed effects*).

- We call *treatment* each combination of the factor level means.
- In this experiment, the factors are (fully) crossed, since each combination of their levels is represented, which results in $3 \times 4 = 12$ treatments.
- The experiment is replicated 4 times and since all treatment groups have an equal number of rats, the design is called *balanced*.
- Denote by y_{ijk} the survival time for the kth animal that receives the ith poison and the jth treatment.

Survival times (unit: 10 hours) in 3×4 factorial structure

	"Treatment"			
Poison	A	B	C	D
I	.31	.82	.43	.45
	.45	1.10	.45	.71
	.46	.88	.63	.66
	.43	.72	.76	.62
II	.36	.92	.44	.56
	.29	.61	.35	1.02
	.40	.49	.31	.71
	.23	1.24	.40	.38
III	.22	.30	.23	.30
	.21	.37	.25	.36
	.18	.38	.24	.31
	.23	.29	.22	.33

Side-by-side Boxplots



Cell Means Model for Two Factors

Equal Sample Sizes for all Treatment Means

$$y_{ijk} = \mu_{ij} + \varepsilon_{ijk}$$

- μ_{ij} are the mean of the ith level of factor A and the jth level of factor B
- ε_{ijk} are independent $\mathcal{N}(0, \sigma^2)$
- $i = 1, \ldots, a$
  $j = 1, \ldots, b$
  $k = 1, \ldots, n$, where $n > 1$.
  $n_T = nab$: total sample size

Usually, the means μ_{ij} are decomposed as follows:

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

where

- μ is the overall mean
- α_i are the factor A, `poison`, effects (fixed)
- β_j are the factor B, `"treatment"`, effects (fixed)
- $(\alpha\beta)_{ij}$ are the interaction effects (fixed)

Factor Effects Model for Two Factors

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

- e_{ijk} are independent $\mathcal{N}(0, \sigma^2)$
- This model is equivalent to the cell means model, but it is more useful when we want to understand the significance of each term in a software like R.
- Here again we need to impose constraints to ensure that the estimators for the effects are unique.

Notation

	Sum	Average
Cell (i,j)	$y_{ij\cdot} = \sum_{k=1}^{n} y_{ijk}$	$\bar{y}_{ij\cdot} = \frac{y_{ij\cdot}}{n}$
Row i	$y_{i\cdot\cdot} = \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk}$	$\bar{y}_{i\cdot\cdot} = \frac{y_{i\cdot\cdot}}{bn}$
Column j	$y_{\cdot j\cdot} = \sum_{i=1}^{a} \sum_{k=1}^{n} y_{ijk}$	$\bar{y}_{\cdot j\cdot} = \frac{y_{\cdot j\cdot}}{an}$
Overall	$y_{\cdots} = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk}$	$\bar{y}_{\cdots} = \frac{y_{\cdots}}{nab}$

Fitting of ANOVA

- Using least squares method, the estimated treatment means are:

$$\hat{\mu}_{ij} = \bar{y}_{ij\cdot}$$

- The factor effects estimators depend on the constraints that we impose. For example, under the *sum-constraints* we have

$$\hat{\alpha}_i = \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot}, \quad \hat{\beta}_j = \bar{y}_{\cdot j\cdot} - \bar{y}_{\cdot\cdot\cdot}$$

$$\widehat{(\alpha\beta)}_{ij} = \bar{y}_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot}$$

- The fitted values and residuals compute as usual as

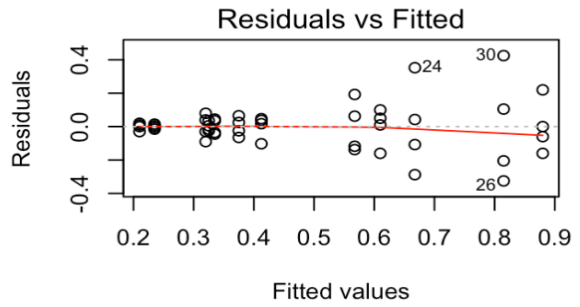
$$\hat{y}_{ijk} = \bar{y}_{ij\cdot}, \quad r_{ij} = y_{ijk} - \hat{y}_{ijk}$$

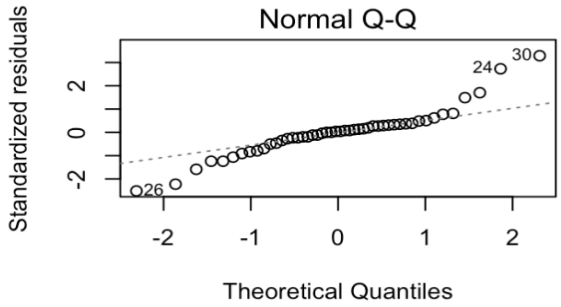
```
g=lm(time ~ poison*treat, rats)
summary(g)
```

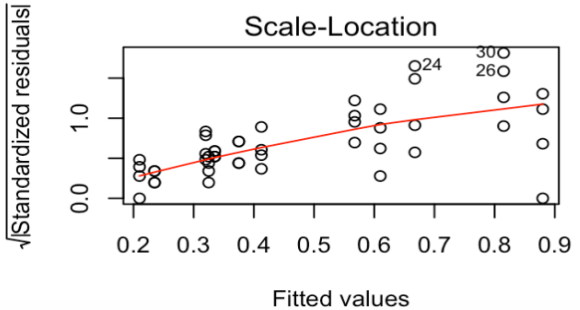
```
##
## Call:
## lm(formula = time ~ poison * treat, data = rats)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.32500 -0.04875  0.00500  0.04312  0.42500
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       0.41250    0.07457   5.532 2.94e-06 ***
## poisonII         -0.09250    0.10546  -0.877   0.3862
## poisonIII        -0.20250    0.10546  -1.920   0.0628 .
## treatB            0.46750    0.10546   4.433 8.37e-05 ***
## treatC            0.15500    0.10546   1.470   0.1503
## treatD            0.19750    0.10546   1.873   0.0692 .
## poisonII:treatB   0.02750    0.14914   0.184   0.8547
## poisonIII:treatB -0.34250    0.14914  -2.297   0.0276 *
## poisonII:treatC  -0.10000    0.14914  -0.671   0.5068
## poisonIII:treatC -0.13000    0.14914  -0.872   0.3892
## poisonII:treatD   0.15000    0.14914   1.006   0.3212
## poisonIII:treatD -0.08250    0.14914  -0.553   0.5836
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1491 on 36 degrees of freedom
## Multiple R-squared:  0.7335, Adjusted R-squared:  0.6521
## F-statistic:  9.01 on 11 and 36 DF,  p-value: 1.986e-07
```

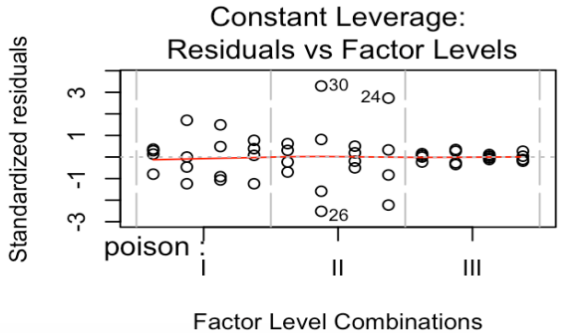
Checking Model Assumptions

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



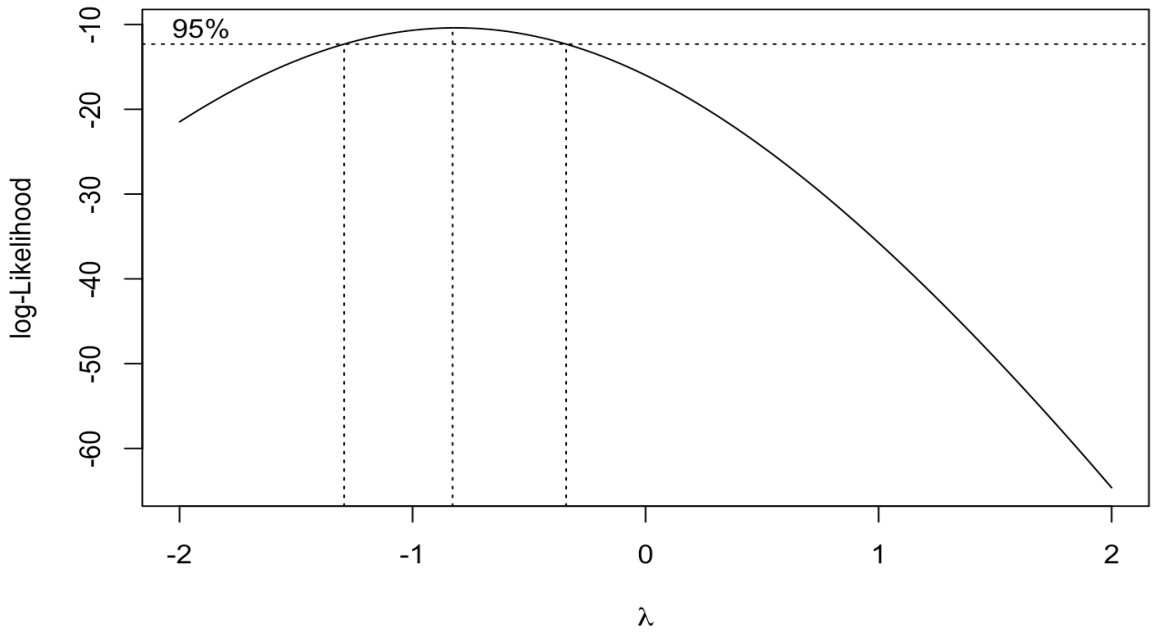






- The plots show that the constant variance assumption is not satisfied, so we check whether a Box-Cox (power) transformation can fix this issue.

```
library(MASS)
par(mfrow=c(1,1))
boxcox(g)
```



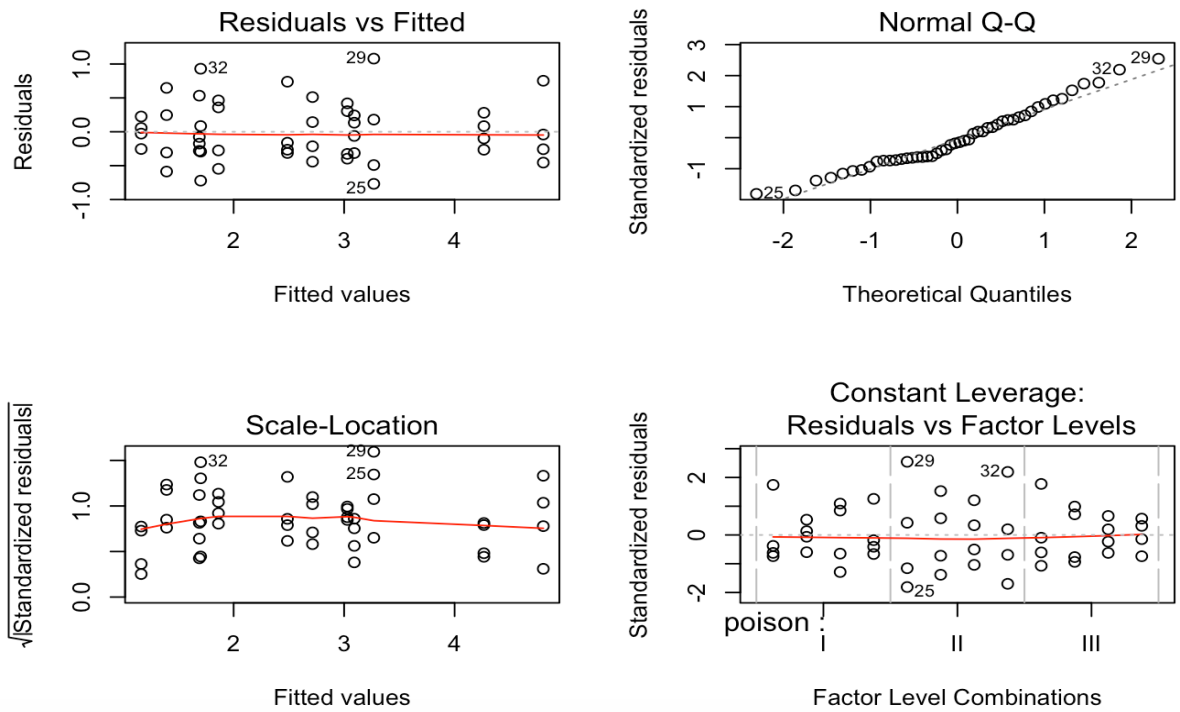
- We take the inverse (reciprocal) transformation of the survival time (response), i.e. the death rate, and we re-fit the model.

```
ginv <- lm(time^-1 ~ poison*treat, data=rats)
summary(ginv)
```

```
##
## Call:
## lm(formula = time^-1 ~ poison * treat, data = rats)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.76847 -0.29642 -0.06914  0.25458  1.07936
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       2.48688    0.24499  10.151 4.16e-12 ***
## poisonII          0.78159    0.34647   2.256 0.030252 *
## poisonIII         2.31580    0.34647   6.684 8.56e-08 ***
## treatB           -1.32342    0.34647  -3.820 0.000508 ***
## treatC           -0.62416    0.34647  -1.801 0.080010 .
## treatD           -0.79720    0.34647  -2.301 0.027297 *
## poisonII:treatB  -0.55166    0.48999  -1.126 0.267669
## poisonIII:treatB -0.45030    0.48999  -0.919 0.364213
## poisonII:treatC   0.06961    0.48999   0.142 0.887826
## poisonIII:treatC  0.08646    0.48999   0.176 0.860928
## poisonII:treatD  -0.76974    0.48999  -1.571 0.124946
## poisonIII:treatD -0.91368    0.48999  -1.865 0.070391 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.49 on 36 degrees of freedom
## Multiple R-squared:  0.8681, Adjusted R-squared:  0.8277
## F-statistic: 21.53 on 11 and 36 DF,  p-value: 1.289e-12
```

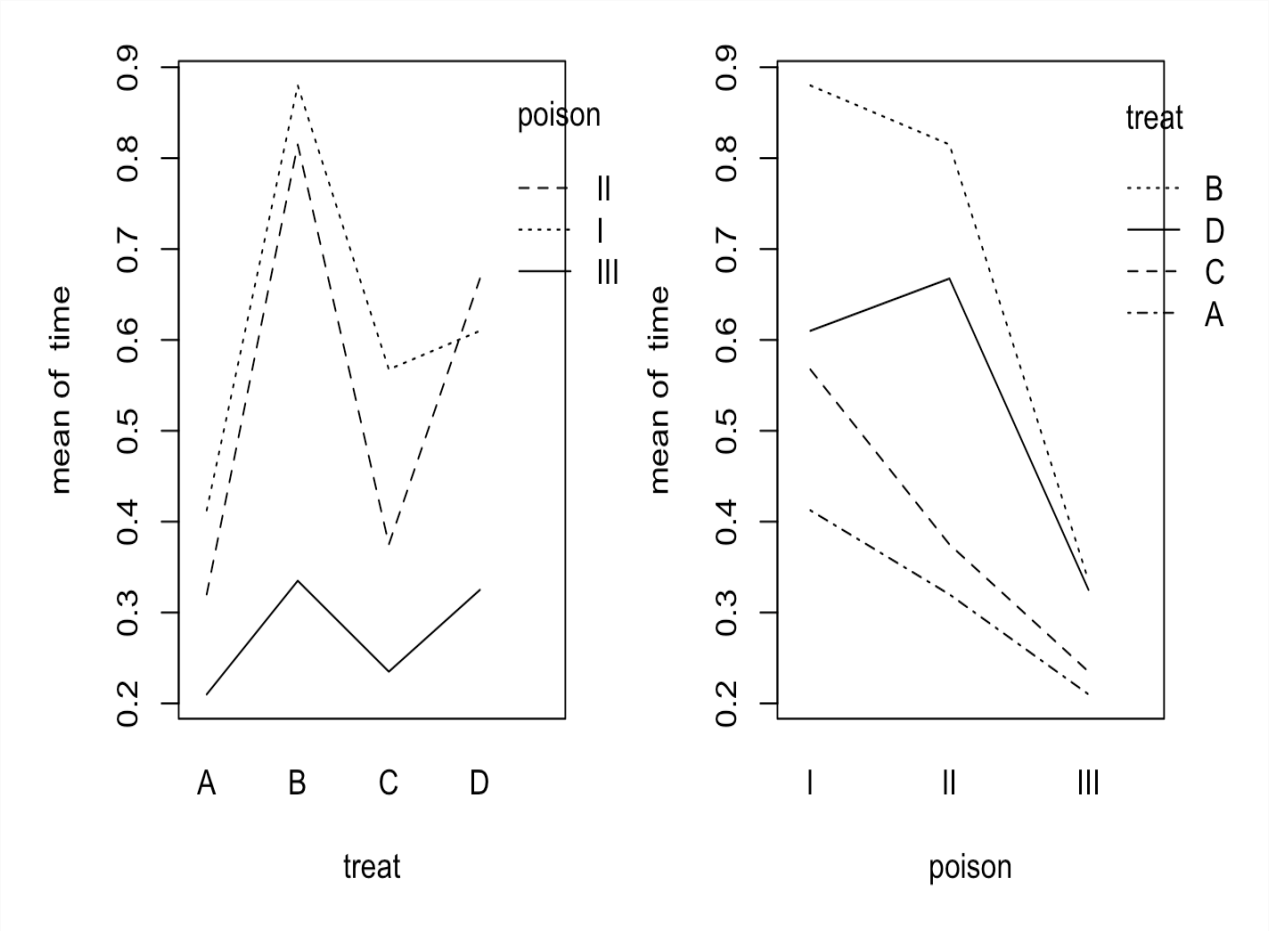
We check the model assumptions for the transformed model:

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



- The diagnostics for the transformed model look good.
- The next goal is to understand the effects (main and interaction).
- We start by creating an *interaction plot*, that is the plot of the response against the factor level combinations.
- This plot will help us to visually inspect the presence (or not) of interactions and the nature of the main effects.

Interaction Plots



Partitioning of Total Sum of Squares

$$\underbrace{y_{ijk} - \bar{y}_{\cdots}}_{\text{Total Deviation}} = \underbrace{\bar{y}_{ij\cdot} - \bar{y}_{\cdots}}_{\substack{\text{Deviation of estimated} \\ \text{treatment mean around} \\ \text{overall mean}}} + \underbrace{y_{ijk} - \bar{y}_{ij\cdot}}_{\substack{\text{Deviation} \\ \text{around estimated} \\ \text{treatment mean}}}$$

$$TSS = FSS + RSS$$

where

$$TSS = \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{\cdots})^2$$

$$FSS = n \sum_i \sum_j (\bar{y}_{ij\cdot} - \bar{y}_{\cdots})^2$$

$$RSS = \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{ij\cdot})^2 = \sum_i \sum_j \sum_k e_{ijk}^2$$

Partitioning of Treatment Sum of Squares

$$\underbrace{\bar{y}_{ij\cdot} - \bar{y}_{\cdots}}_{\substack{\text{Deviation of estimated} \\ \text{treatment mean around} \\ \text{overall mean}}} = \underbrace{\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots}}_{\substack{\text{A main} \\ \text{effect}}} + \underbrace{\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots}}_{\substack{\text{B main} \\ \text{effect}}} + \underbrace{\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdots}}_{\substack{\text{A B interaction} \\ \text{effect}}}$$

$$FSS = SSA + SSB + SSAB \quad \text{(Orthogonal Decomposition)}$$

where

$$SSA = nb \sum_i (\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots})^2$$

$$SSB = na \sum_j (\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots})^2$$

$$SSAB = n \sum_i \sum_j (\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\cdots})^2$$

ANOVA Table

Source of Variation	SS	df	MS
Factor A	SSA	$a-1$	$MSA = \frac{SSA}{a-1}$
Factor B	SSB	$b-1$	$MSB = \frac{SSB}{b-1}$
AB Interactions	$SSAB$	$(a-1)(b-1)$	$MSAB = \frac{SSAB}{(a-1)(b-1)}$
Error	RSS	$ab(n-1)$	$MSE = \frac{RSS}{ab(n-1)}$
Total	TSS	$nab-1$	

- In order to test for the statistical significance of the interaction terms, we use partial F-tests. So, we fit a main effects model (i.e. no interactions)

$$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}$$

- Then, we compare the two nested models:

$$\begin{cases} H_0: \text{ smaller model with } p_0 \text{ coefficients} \\ H_\alpha: \text{ larger model with } p_\alpha \text{ coefficients} \end{cases}$$

The F-test is formulated as

$$F = \frac{(RSS_0 - RSS_\alpha)/(p_\alpha - p_0)}{MSE_\alpha} \sim F_{p_\alpha - p_0, n - p_\alpha} \text{ under the } H_0$$

- We can also perform F-tests directly using the ANOVA table, where for the interaction term we have:

$$F_{AB} = \frac{MSAB}{MSE} \sim F_{(a-1)(b-1),nab-1}$$

```
anova(g)
```

```
## Analysis of Variance Table
##
## Response: time
##              Df  Sum Sq Mean Sq F value    Pr(>F)
## poison        2 1.03301 0.51651 23.2217 3.331e-07 ***
## treat         3 0.92121 0.30707 13.8056 3.777e-06 ***
## poison:treat  6 0.25014 0.04169  1.8743    0.1123
## Residuals    36 0.80073 0.02224
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Hierarchy principle: we test for main effects *only if* the interaction term is not statistically significant.

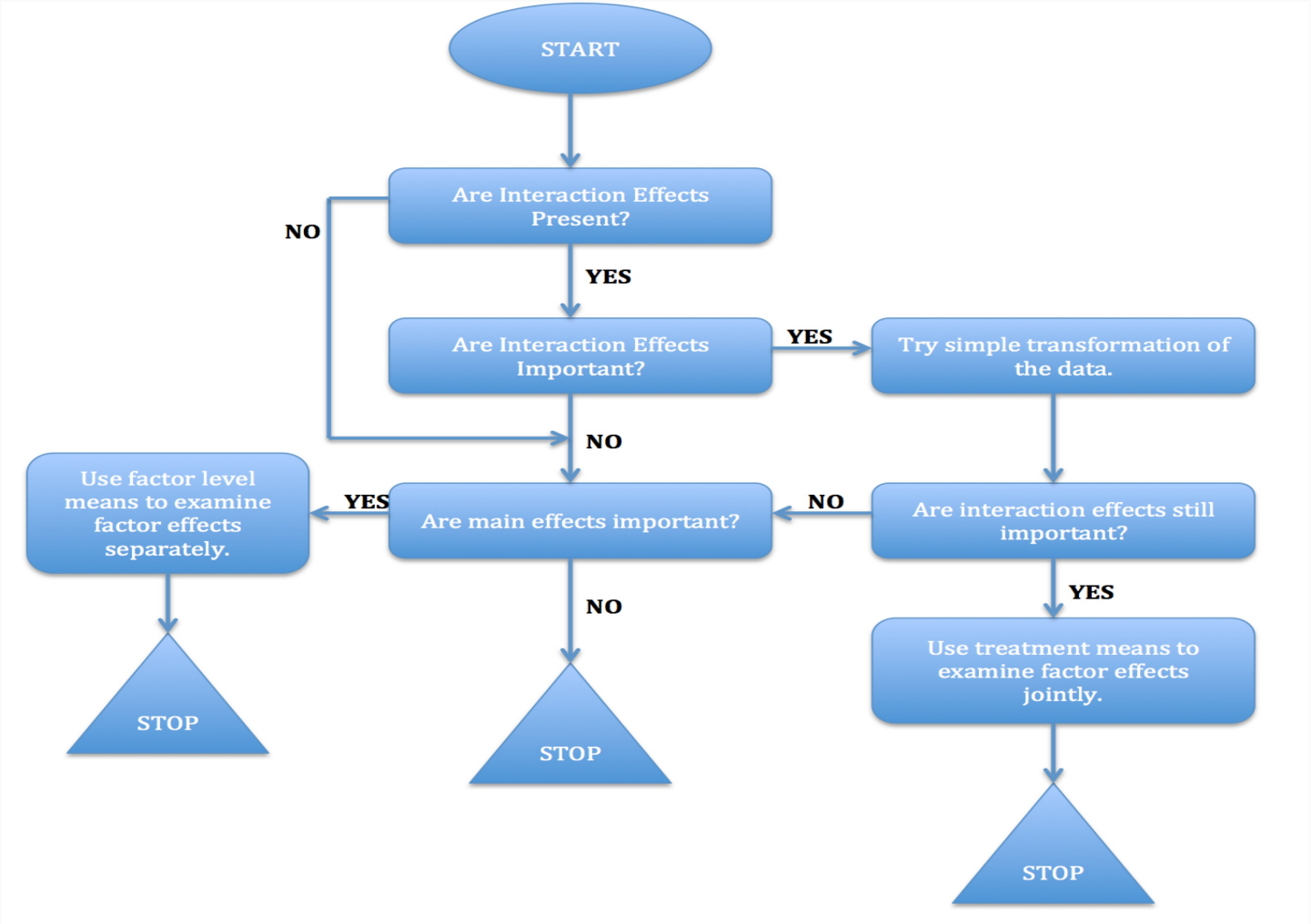
```
maineffectg <- lm(time^-1 ~ poison + treat, data=rats)
summary(maineffectg)
```

```
##
## Call:
## lm(formula = time^-1 ~ poison + treat, data = rats)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.82757 -0.37619  0.02116  0.27568  1.18153
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.6977     0.1744  15.473  < 2e-16 ***
## poisonII      0.4686     0.1744   2.688  0.01026 *
## poisonIII     1.9964     0.1744  11.451 1.69e-14 ***
## treatB       -1.6574     0.2013  -8.233 2.66e-10 ***
## treatC       -0.5721     0.2013  -2.842  0.00689 **
## treatD       -1.3583     0.2013  -6.747 3.35e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4931 on 42 degrees of freedom
## Multiple R-squared:  0.8441, Adjusted R-squared:  0.8255
## F-statistic: 45.47 on 5 and 42 DF,  p-value: 6.974e-16
```

```
anova(maineffectg, ginv)
```

```
## Analysis of Variance Table
##
## Model 1: time^-1 ~ poison + treat
## Model 2: time^-1 ~ poison * treat
##   Res.Df     RSS Df Sum of Sq      F Pr(>F)
## 1     42 10.2139
## 2     36  8.6431  6    1.5708 1.0904 0.3867
```

Strategy for analysis: 2 or more factor studies



Estimation of Factor Level Means

When interactions are *not* statistically significant, we analyze the factor level means: [1]

- Factor Level Means: $\hat{\mu}_{i\cdot} = \bar{y}_{i\cdot\cdot}$, $s^2_{\hat{\mu}_{i\cdot}} = \frac{MSE}{bn}$
- Differences of Factor Level Means:

$$\hat{\mu}_{i\cdot} - \hat{\mu}_{i'\cdot} = \bar{y}_{i\cdot\cdot} - \bar{y}_{i'\cdot\cdot}, \; s^2_{\hat{D}} = \frac{2MSE}{bn}$$

- Contrasts of Factor Level Means:

$$\hat{L} = \sum c_i \hat{\mu}_{i\cdot} = \sum c_i \bar{y}_{i\cdot\cdot}, \; s^2_{\hat{L}_{(i)}} = \frac{MSE}{bn} \sum c_i^2$$

where $\sum c_i = 0$

[1] Formulas are given for factor A

- For individual hypothesis test and CIs, the multiplier is $T_{(n-1)ab}(\alpha/2)$.
- For family hypothesis tests/intervals, we select the desired family multiplier:
  - Tukey Multiplier: $\frac{1}{\sqrt{2}} q_{a,(n-1)ab}(1-\alpha)$
  - Bonferroni Multiplier: $B = T_{(n-1)ab}(1-\alpha/2m)$, where m refers to the number of multiple comparisons.
  - Scheffé Multiplier:
    - $S^2 = (b-1)F_{b-1,(n-1)ab}(1-\alpha)$, if the contrasts involve the $\mu_{i\cdot}$ and
    - $S^2 = (a-1)F_{a-1,(n-1)ab}(1-\alpha)$, if the contrasts involve the $\mu_{\cdot j}$.

Rats Example Pairwise Differences

We use Tukey's method to look at the pairwise differences in the `rats` model:

```
TukeyHSD(aov(time^-1 ~ poison + treat, data=rats))
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = time^-1 ~ poison + treat, data = rats)
##
## $poison
##              diff        lwr       upr     p adj
## II-I   0.4686413 0.04505584 0.8922267 0.0271587
## III-I  1.9964249 1.57283950 2.4200103 0.0000000
## III-II 1.5277837 1.10419824 1.9513691 0.0000000
##
## $treat
##           diff        lwr         upr     p adj
## B-A -1.6574024 -2.1959343 -1.11887050 0.0000000
## C-A -0.5721354 -1.1106673 -0.03360355 0.0335202
## D-A -1.3583383 -1.8968702 -0.81980640 0.0000002
## C-B  1.0852669  0.5467351  1.62379883 0.0000172
## D-B  0.2990641 -0.2394678  0.83759598 0.4550931
## D-C -0.7862029 -1.3247347 -0.24767096 0.0018399
```

Estimation of Treatment Means

When interactions are *statistically significant*, we analyze the treatment means:

- Treatment Means: $\hat{\mu}_{ij} = \bar{y}_{ij\cdot}$, $s^2_{\hat{\mu}_{ij}} = \frac{MSE}{n}$
- Differences of Treatment Means:

$$\hat{D} = \hat{\mu}_{ij} - \hat{\mu}_{i'j'} = \bar{y}_{ij\cdot} - \bar{y}_{i'j'\cdot}, \quad i, j \neq i', j' \text{ and } s^2_{\hat{D}} = \frac{2MSE}{n}$$

- Contrasts of Treatment Means:

$$\hat{L} = \sum\sum c_{ij}\hat{\mu}_{ij} = \sum\sum c_{ij}\bar{y}_{ij\cdot}, \text{ where } \sum\sum c_{ij} = 0$$

with variance $s^2_{\hat{L}} = \frac{MSE}{n}\sum c^2_{ij}$.

Multipliers for Treatment Means

- For individual hypothesis test and CIs, the multiplier is $T_{(n-1)ab}(\alpha/2)$.
- For family hypothesis tests/intervals, we choose the desired family multiplier:
  - Tukey Multiplier: $\frac{1}{\sqrt{2}} q_{ab,(n-1)ab}(1-\alpha)$
  - Bonferroni Multiplier: $B = T_{(n-1)ab}(1-\alpha/2m)$, where m refers to the number of multiple comparisons.
  - Scheffé Multiplier: $S^2 = (ab-1)F_{ab-1,(n-1)ab}(1-\alpha)$, if the contrasts involve the $\mu_{i\cdot\cdot}$.