Design & Analysis of Experiments

Experimental Designs & AN(C)OVA & MAN(C)OVA

Nagiza F. Samatova, samatova@csc.ncsu.edu
Professor, Department of Computer Science
North Carolina State University

NC STATE UNIVERSITY
Department of Computer Science

Outline

- Designs
  - One-way Between-Groups ANOVA (Analysis of Variance)
  - One-way Within-Groups ANOVA
  - Two-way Factorial ANOVA
  - ANCOVA: Adjusting for Covariates
  - MANOVA: Multivariate Analysis of Variable (>1 Response Vars)
  - MANCOVA: Multivariate Analysis of Covariance
- Fitting ANOVA and ANCOVA Models in R
  - Formulas for different types of designs
  - Order matters: Type I, II, and III ordering of formula terms
  - Examples: ANOVA and ANCOVA
  - Assessing ANOVA and ANCOVA Assumptions
- Fitting MANOVA Models in R
  - Examples: One-way MANOVA
  - Assessing MANOVA Assumptions

Experimental Designs

AN(C)OVA & MAN(C)OVA

Focus: **Design of experiments ≡ ANalysis Of VAriance**

Design #1: One-way Between-Groups ANOVA

- Examplar Study: Goal: To study two treatments of anxiety
  - **Treatment (Independent Variable)**: Two treatments
    - CBT: Cognitive Behavior Therapy
    - EMDR: Eye Movement Desensitization & Reprocessing Therapy
  - **Response (Dependent Variable)**: (collected after 5 weeks of treatment)
    - STAI: State-Trait Anxiety Inventory; a self-report measure of anxiety
  - **Subjects**:
    - Randomly divided between two ***independent groups***: CBT & EMDR

		Subjects
Treatment	CBT	s_1
		s_2
		...
		$s_{n/2}$
	EMDR	$s_{n/2+1}$
		...
		...
		s_n

One-way Between-Groups Balanced ANOVA

- Treatment is a **between-groups** factor with two levels
- **Balanced design**: equal number of subjects in each treatment condition; otherwise, **unbalanced**
- **One-way**: because a single classification variable

F-tests to assess the effects in ANOVA designs

- If the F-test for Treatment is significant then reject the null hypothesis: H_0: the mean STAI scores are the same
- Conclude: the mean STAI scores changed over time for two therapies differed after five weeks of treatment

Design #2: One-way Within-Groups ANOVA

- Examplar Study: Goal: To study ***longitudinally*** one treatment of anxiety
  - **Treatment (Independent Variable)**: One treatment
    - CBT: Cognitive Behavior Therapy
  - **Response (Dependent Variable)**: (collected after 5 weeks of treatment)
    - STAI: State-Trait Anxiety Inventory; a self-report measure of anxiety
  - **Subjects**:
    - The same subjects over different time points (**dependent groups**)
  - **Time**: Two different time points: 5 weeks & 6 months

One-way Within-Groups Balanced ANOVA

	Time	
CBT Treatment	**5 weeks**	**6 months**
	s_1	s_1
	s_2	s_2
	…	…
	…	…
	s_n	s_n

- Time is a **within-groups** factor with two levels: each subject is measured under both levels
  - **Repeated measures** ANOVA
- **One-way**: because a single classification variable

Paired F-tests to assess effects in ANOVA designs

- If the F-test for Treatment is significant then reject the null hypothesis: H_0: the mean STAI scores are the same
- Conclude: the mean STAI scores change over time: between 5 weeks and 6 months

Design #3: Factorial (Mixed-Model) ANOVA

- Examplar Study: Goal: To study ***two treatments*** of anxiety ***longitudinally***
  - **Treatment (Independent Variable)**: Two treatments
    - CBT: Cognitive Behavior Therapy
    - EMDR: Eye Movement Desensitization & Reprocessing Therapy
  - **Response (Dependent Variable)**: (collected after 5 weeks of treatment)
    - STAI: State-Trait Anxiety Inventory; a self-report measure of anxiety
  - **Subjects**: Randomly assigned to two ***independent groups***: CBT & EMDR
  - **Time**: Two **dependent groups** over time: 5 weeks & 6 months

		5 wks	6 mo.
Treatment	CBT	s_1	s_1
		s_2	s_2
		...	...
		$s_{n/2}$	$s_{n/2}$
	EMDR	$s_{n/2+1}$	$s_{n/2+1}$
		...	...
		...	...
		s_n	s_n

Two-way Factorial ANOVA Design

- **Main effects**: Impact of Therapy (averaged across Time) and Time (averaged across Therapy type)
- **Interaction effect**: Interaction of Therapy & Time
- **Factorial ANOVA**: cross 2 factors (**two-way**) or more

Three F-tests to assess ANOVA design effects

- F-test for Therapy: Significant→ CBI & EMDR differ
- F-test for Time: Significant → change over time
- F-test for Therapy × Time interaction: Significant → two treatments had a differential impact over time: different change from 5 wks to 6 mo. for 2 therapies

Design #4: ANCOVA

- Goal: To study the treatments of anxiety with ***confounding factors***
- **Confounding Factor (Nuisance Variable)**: Other factors (than Treatment) that could explain the post-therapy differences on the dependent variable
  - Depression level:
    - A self-reported measure such as Beck Depression Inventory (BDI)
    - Although subjects were assigned randomly to treatment conditions, it is possible that two therapy groups differed in patient depression levels at the start of the study
    - → Any post-therapy differences might be due to the preexisting depression differences and not to experimental manipulations

ANCOVA Design

- Because depression could also explain the group difference on the dependent variable, it is a **confounding factor**
- Because the study is not interested in depression, this confounding factor is called a **nuisance variable**

Design #5-6: MANOVA & MANCOVA

- Goal: To study the treatments of anxiety with ***multiple dependent variables*** and/or ***confounding factors***
- **Multiple Dependent Variables:** To increase the validity of the study
  - STAI: One dependent variable
  - Family ratings: Another dependent variable
  - Therapy ratings: Yet another dependent variable
  - A measure assessing the impact of anxiety on the daily functioning

- **MANOVA**: **Multivariate Analysis of Variance**
  - There is more than one dependent variable

- **MANCOVA**: **Multivariate Analysis of Covariance**
  - Besides multiple dependent variables, there are covariates present

Experimental Designs

FITTING ANOVA & ANCOVA MODELS

ANOVA and ANCOVA in R: Formula

aov (***formula***, data = dataframe

Design	Formula
One-way ANOVA	**Response**.STAI **~** **Treatments**.CBT.EMDR
One-way ANCOVA with one covariate	**Response**.STAI **~** **Covariate**.Depression **+** **Treatments**.CBT.EMDR
Two-way Factorial ANOVA	**Response**.STAI **~** **Treatments**.CBT.EMDR ***** **Times.**5wks.6mo
Two-way Factorial ANCOVA with two covariates	**Response**.STAI **~** **Covariate**.Depression **+** **Covariate**.Gender **+** **Treatments**.CBT.EMDR ***** **Times.**5wks.6mo
One-way within groups ANOVA	**Response**.STAI **~** **Times.**5wks.6mo **+** **Error (Subject/Times.**5wks.6mo)
Repeated measures ANOVA with one within groups factor (W) and one between-groups factor (B)	**Response ~ B * W +** **Error (Subject/W)**

Formula Terms: Order Counts!

- **The order in which the effects appear in a formula matters when**
  - More than one factor
  - Design is unbalanced: the greater the imbalance in sample sizes, the greater the impact the order of the terms will have on the results
  - Covariates are present
  - $y \sim A * B$ and $y \sim B * A$ will produce different results
- **Type I (sequential) Ordering**: $y \sim A + B + A * B$
  - The impact of A on y
  - The impact of B on y, controlling for A
  - The interaction of A and B, controlling for the A and B main effects
  - Summary: the effects are adjusted for those factors that appear earlier in the formula. A is unadjusted. B is adjusted for the A. The $A{:}B$ interaction is adjusted for A and B.
- **Best Practices**:
  - Strive for balanced designs
  - Place more fundamental effects earlier in the formula
  - List covariates first in the formula
  - Followed by main factors, followed by two-way interactions, followed three-way interactions, and so on

Other Formula Term Ordering Schemes

- **Type I (sequential) Ordering**: $y \sim A + B + A * B$
  - Effects are adjusted for those factors that appear earlier in the formula. A is unadjusted. B is adjusted for A. The $A{:}B$ interaction is adjusted for A and B.
  - R's aov() function deploys Type I approach by default

- **Type II (hierarchical) Ordering**: :
  - Effects are adjusted for other effects at the same or lower level. A is adjusted for B and B is adjusted for A. The $A{:}B$ interaction is adjusted for both A and B.
  - help(Anova, package="car"): allows to choose Type I, II, or III

- **Type III (hierarchical) Ordering**: :
  - Each effect is adjusted for every other effect in the model. A is adjusted for B and $A{:}B$. B is adjusted for A and $A{:}B$. The $A{:}B$ interaction is adjusted for both A and B.
  - SAS and SPSS employ Type III approach by default

Fitting ANOVA Models

ONE-WAY BETWEEN-GROUPS

Ex: One-way Between-Groups ANOVA

```
# install.packages("multcomp")
require (multcomp)
attach (cholesterol)
head (cholesterol)
```

```
Treatments (trt: Dependent Var):
  * Same drug but administered differently
      - 20 mg once per day (1time)
      - 10 mg twice per day (2times)
      - 5 mg four times per day (4times)
  * drugD
  * drugE
Response Var: Cholesterol Reduction
```

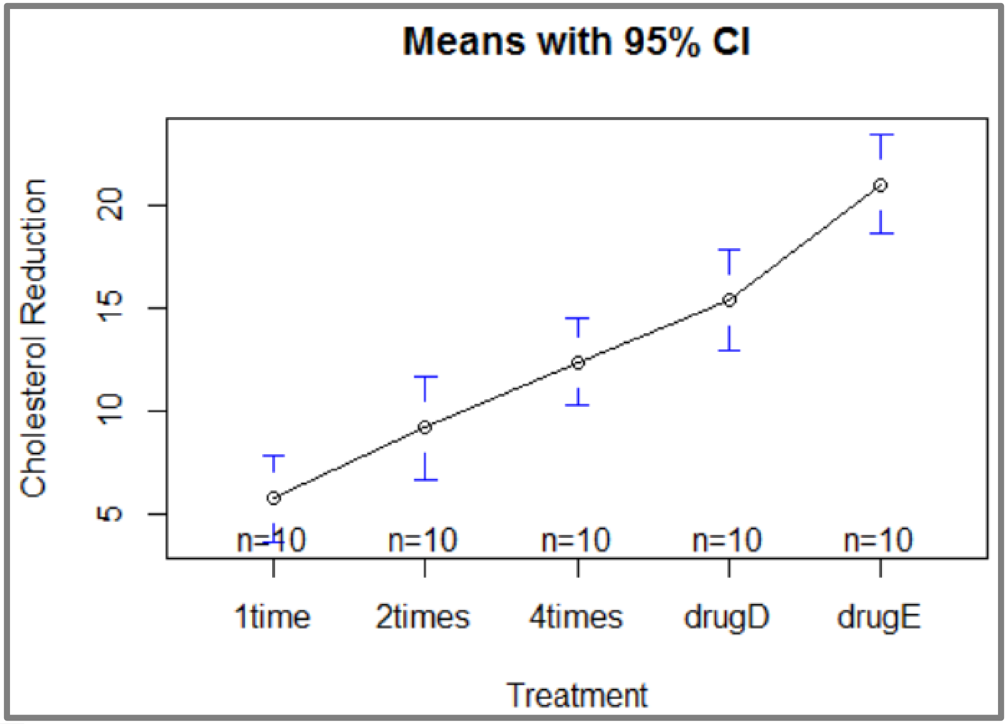
```
> table(trt)
trt
 1time 2times 4times  drugD  drugE 
    10     10     10     10     10 
> aggregate (response, by=list(trt), FUN=mean)
  Group.1        x
1   1time  5.78197
2  2times  9.22497
3  4times 12.37478
4   drugD 15.36117
5   drugE 20.94752
> aggregate (response, by=list(trt), FUN=sd)
  Group.1        x
1   1time 2.878113
2  2times 3.483054
3  4times 2.923119
4   drugD 3.454636
5   drugE 3.345003
```

File: Design_Experiments.ANOVA.MANOVA.R

Ex: One-way Between-Groups ANOVA (cont.)

```
> fit <- aov (response ~ trt)
> summary(fit)
            Df Sum Sq Mean Sq F value   Pr(>F)
trt          4 1351.4   337.8   32.43 9.82e-13 ***
Residuals   45  468.8    10.4
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
```

- F-test is significant:
  - five treatments are not equally effective
  - it does not tell which treatments differ from one another
- drugE appears to produce the greatest cholesterol reduction based on the plotmeans()




```
library (gplots)
plotmeans (response ~trt, xlab="Treatment",
           ylab = "Cholesterol Reduction",
           main="Means with 95% CI")
```

Multiple Comparisons: TukeyHSD(): pairwise

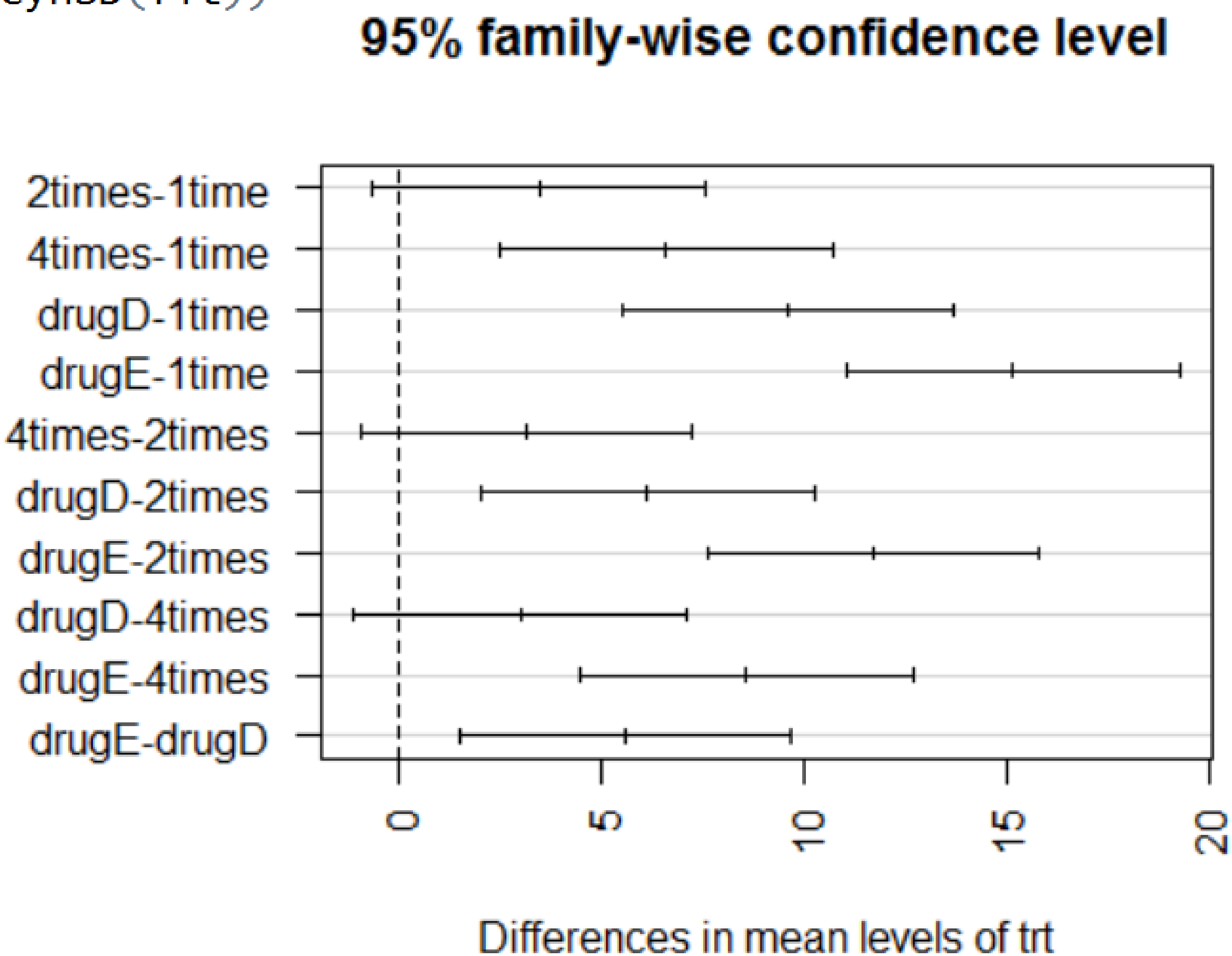
- F-test is significant:
  - Conclusion: five treatments are not equally effective
  - But: it does not tell which treatments differ from one another

TukeyHSD() provides a test of all **pairwise differences** between group means

```
> TukeyHSD(fit)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = response ~ trt)

$trt
                  diff        lwr       upr     p adj
2times-1time   3.44300 -0.6582817  7.544282 0.1380949
4times-1time   6.59281  2.4915283 10.694092 0.0003542
drugD-1time    9.57920  5.4779183 13.680482 0.0000003
drugE-1time   15.16555 11.0642683 19.266832 0.0000000
4times-2times  3.14981 -0.9514717  7.251092 0.2050382
drugD-2times   6.13620  2.0349183 10.237482 0.0009611
drugE-2times  11.72255  7.6212683 15.823832 0.0000000
drugD-4times   2.98639 -1.1148917  7.087672 0.2512446
drugE-4times   8.57274  4.4714583 12.674022 0.0000037
drugE-drugD    5.58635  1.4850683  9.687632 0.0030633
```

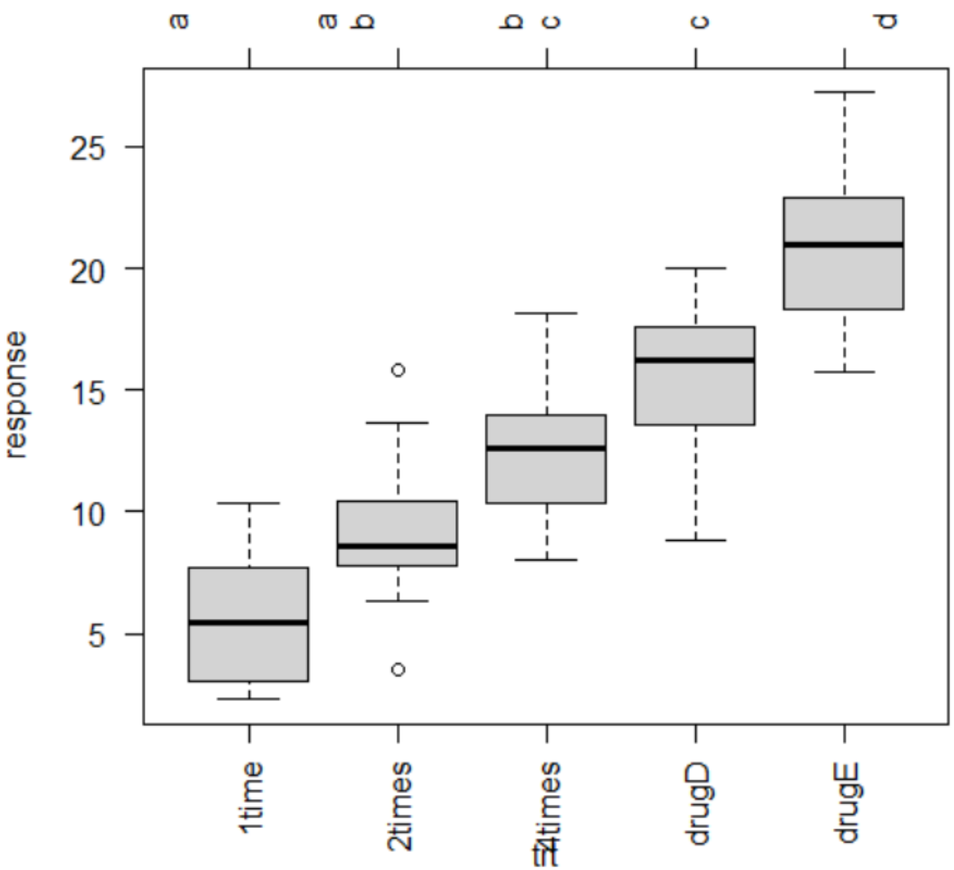
Plot: TukeyHSD(): pairwise comparisons

```
par(las=2)
par (mar=c(5,8,4,2))
plot (TukeyHSD(fit))
```



Multiple Comparisons: glht()

ghlt() provides a test of **comparing multiple** group means



- 1time & 2times are not significantly different (share letter a)
- 2times & 4times are not significantly different (share letter b)
- 1time & 4times are different (they do not share a letter)
- drugE is superior and is different from all the competitors

```
library(multicomp)
par (mar = c(5,4,6,2))
comp <- glht(fit, linfct=mcp(trt="Tukey"))
plot (cld(comp, level=0.05), col="lightgrey")
```

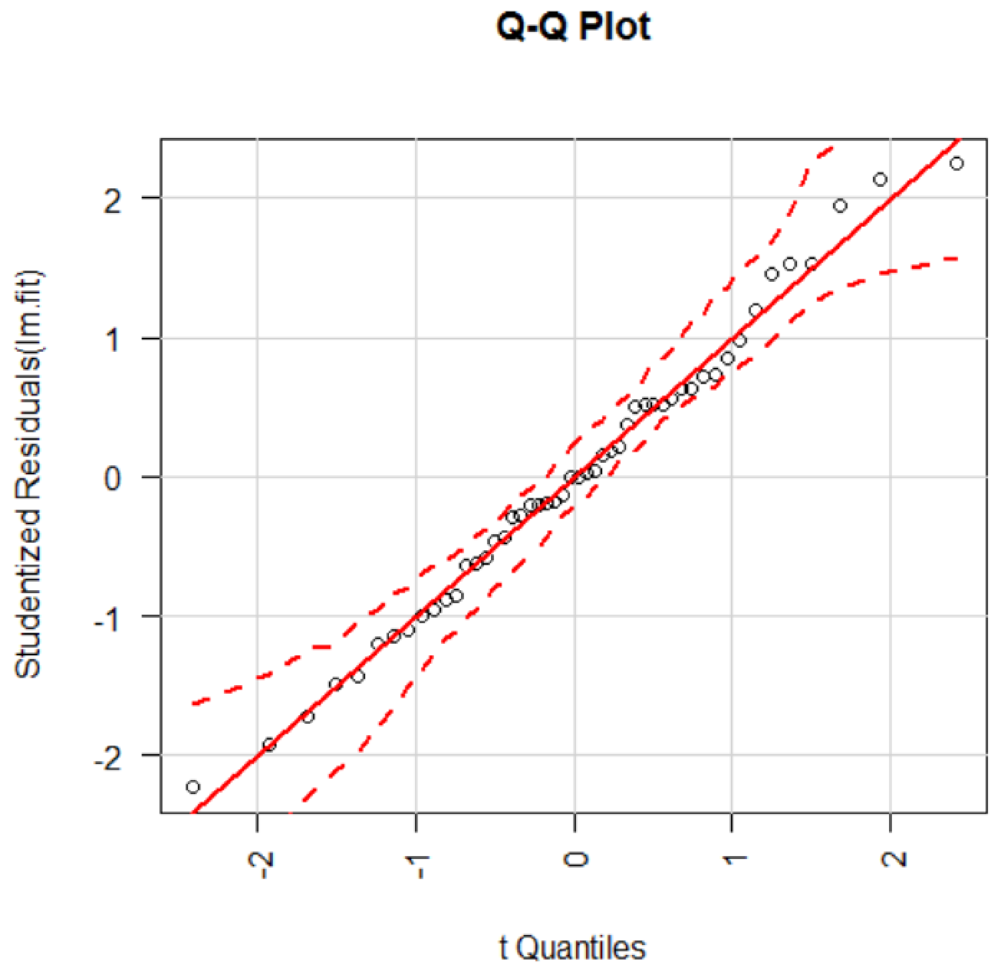
Assumptions: One-way ANOVA

- **Assumptions underlying the statistical tests**:
  - Dependent variable is normally distributed
  - Dependent variable has equal variance in each trt group
  - ANOVA results could be sensitive to outliers

Normality Assumption: One-way ANOVA

Dependent variable must be normally distributed

Data falls within 95% CI → normality assumption is met




```r
library (car)
lm.fit <- lm(response ~ trt, data=cholesterol)
qqPlot (lm.fit, simulate=TRUE,
        main="Q-Q Plot", labels=FALSE)
```

Equal Variance Assumption: One-way ANOVA

Dependent variable must have equal variance for each group

```
> bartlett.test (response ~ trt, data=cholesterol)

 Bartlett test of homogeneity of variances

data:  response by trt
Bartlett's K-squared = 0.5797, df = 4, p-value = 0.9653
```

- **Fail to reject the null hypothesis (p—value=0.96)**
- **H_0: Equality / homogeneity of variances**

Lack of Outliers Assumption: One-way ANOVA

ANOVA can be sensitive to the presence of outliers

```
> library(car)
> outlierTest(fit)

No Studentized residuals with Bonferonni p < 0.05
Largest |rstudent|:
   rstudent unadjusted p-value Bonferonni p
19 2.251149           0.029422           NA
```

- **Fail to reject the null hypothesis (NA: p-value>1)**
- **H_0: There is no outliers in the data**

Fitting ANCOVA Models

ONE-WAY BETWEEN-GROUPS

Ex: One-way Between-Groups ANCOVA

- **Treatments**: four doses of drugs administered to pregnant mice
- **Response**: Post-birth weight
- **Covariate**: Gestation time

```r
data (litter, package="multcomp")
attach(litter)
head(litter)
table(dose)
fit <- aov(weight ~ gesttime + dose)
summary(fit)
```

```
            Df Sum Sq Mean Sq F value  Pr(>F)
gesttime     1  134.3  134.30   8.049 0.00597 **
dose         3  137.1   45.71   2.739 0.04988 *
Residuals   69 1151.3   16.69
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
0.1 ' ' 1
```

- Gestation time was related to birth weight
- Drug dosage was related to birth weight after controlling for gestation time
- Mean birth weight isn't the same for each of the drug dosages, after controlling for gestation time:
  - F-test does not say which treatments do not have the same mean birth weight, i.e. which means differ from one another
  - → use glht() for multiple comparisons

ANCOVA: Adjusting Group Means

- Original grop means before adjusting for the covariate

```
> aggregate(weight, by=list(dose), mean)
  Group.1        x
1       0 32.30850
2       5 29.30842
3      50 29.86611
4     500 29.64647
```

- **Adjusted group means** after partialing out the effects of the covariate

```
install.packages("effects")
library (effects)
effect ("dose", fit)
```

```
 dose effect
dose
       0        5       50      500
32.35367 28.87672 30.56614 29.33460
```

ANCOVA: Multiple Comparisons

- Mean birth weight isn't the same for each of the drug dosages, after controlling for gestation time:
  - F-test does not say which treatments do not have the same mean birth weight, i.e. which means differ from one another
  - → use glht() for multiple comparisons

```r
# c(3, -1, -1, -1): comparison of the no-drug group
#  with the average of the other three drug groups
library(multicomp)
contrast <- rbind("no drug vs drug" = c(3, -1,-1, -1))
summary(glht(fit, linfct=mcp(dose=contrast)))
```

```
	 Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: User-defined Contrasts


Fit: aov(formula = weight ~ gesttime + dose)

Linear Hypotheses:
                     Estimate Std. Error t value Pr(>|t|)  
no drug vs drug == 0    8.284      3.209   2.581    0.012 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
(Adjusted p values reported -- single-step method)
```

- **Significant t-statistic (p-value<0.05) →**
- **Fail to reject H_0**
- **No-drug group has higher birth weight than drug conditions**

Assumptions: One-way ANCOVA

- **Assumptions underlying the statistical tests**:
  - Dependent/response variable is ***normally distributed***
  - Response variable has ***equal variance*** in each group
  - ANOVA results could be sensitive to ***outliers***
  - ***Homogeneity / equality of regression slopes*** between response variable and covariate for each group

Homogeneity of Regression Slopes: ANCOVA

```
library (multcomp)
fit2 <- aov (weight ~ gesttime*dose, data=litter)
summary(fit2)
```

```
              Df Sum Sq Mean Sq F value  Pr(>F)
gesttime       1  134.3  134.30   8.289 0.00537 **
dose           3  137.1   45.71   2.821 0.04556 *
gesttime:dose  3   81.9   27.29   1.684 0.17889
Residuals     66 1069.4   16.20
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

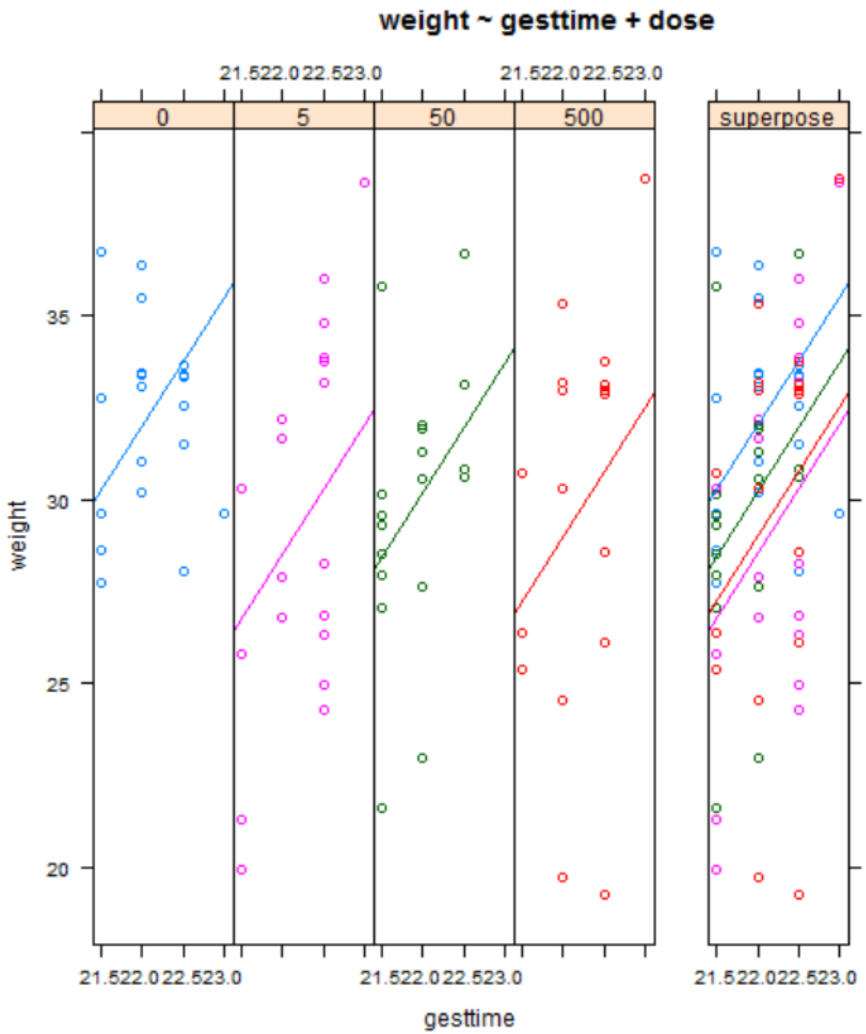
- **Fail to reject H_0**
- **The gestation*dose interaction is nonsignificant supporting the assumption of equality of slopes**

If H_0 were rejected then other tests could be tried out:

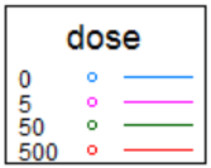
- nonparametric ANCOVA: help(sm.ancova, package="sm")

ANCOVA: Visualization

```r
library (HH)
ancova (weight ~ gesttime + dose, data=litter)
```



Relationship between the dependent variable (birth weight), the covariate (gestation time), and the factor (dose)



- Regression lines for predicting birth weight from gestation time are parallel in each group but have different intercepts

Fitting ANOVA Models

TWO-WAY FACTORIAL

Ex: Two-Way Between-Groups Factorial ANOVA

Subjects are assigned to groups from the cross-classification of factors

- Cross-classification of factors (Independent variable):
  - dose: Three levels of ascorbic acid (0.5, 1, and 2)
  - supp: Two delivery methods (Orange Juice (OJ) and Vitamic C (VC))
- Dependent Variable: Response
  - Tooth length (len) for pigs

```
attach (ToothGrowth)
table (supp, dose)
```

```
    dose
supp 0.5  1  2
  OJ  10 10 10
  VC  10 10 10
```

balanced design

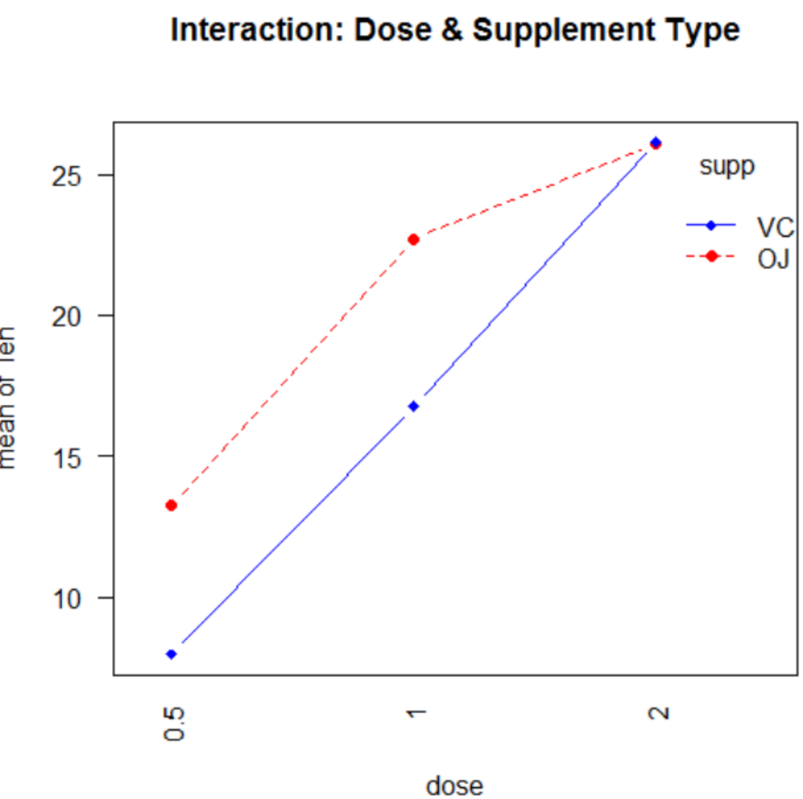
```
> aggregate (len,
+             by=list(supp,dose),
+             FUN=mean)
  Group.1 Group.2     x
1      OJ     0.5 13.23
2      VC     0.5  7.98
3      OJ     1.0 22.70
4      VC     1.0 16.77
5      OJ     2.0 26.06
6      VC     2.0 26.14
```

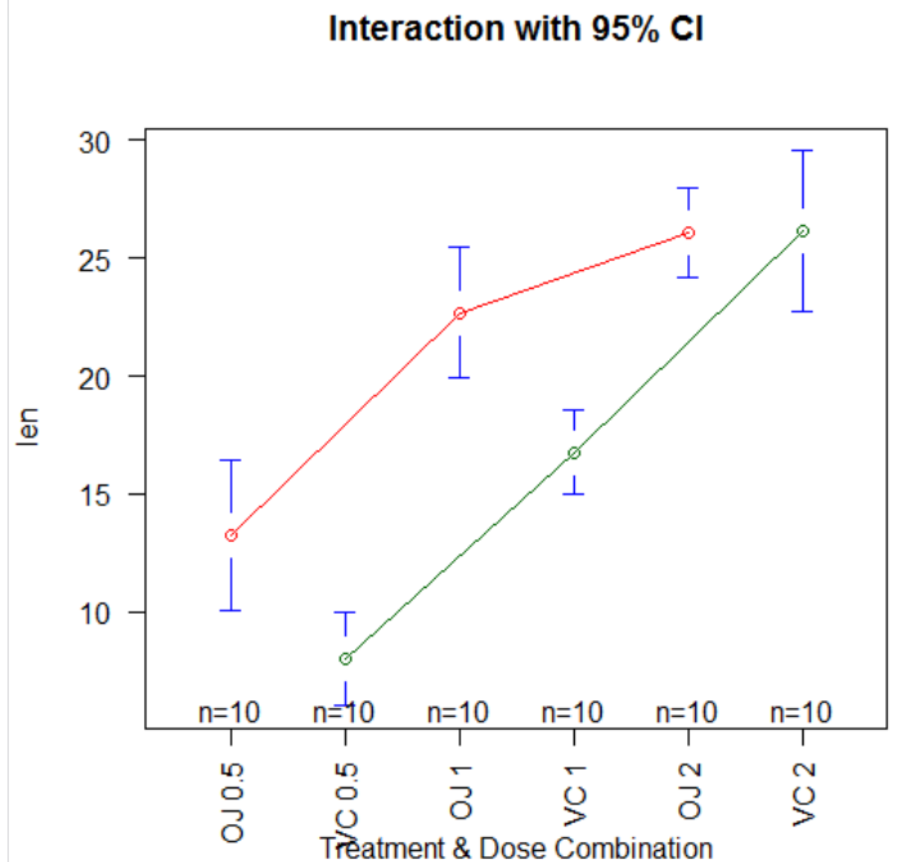
```
> fit <- aov(len ~ supp*dose)
> summary(fit)
            Df Sum Sq Mean Sq F value   Pr(>F)    
supp         1  205.4   205.4  12.317 0.000894 ***
dose         1 2224.3  2224.3 133.415  < 2e-16 ***
supp:dose    1   88.9    88.9   5.333 0.024631 *  
Residuals   56  933.6    16.7                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- Main effects: sup and dose
  - significant (p-value < 0.05)
- Interaction between factors: sup*dose
  - significant (p-value < 0.05)
- Reject the null hypothesis of equal means

Visualization: Two-Way Factorial ANOVA

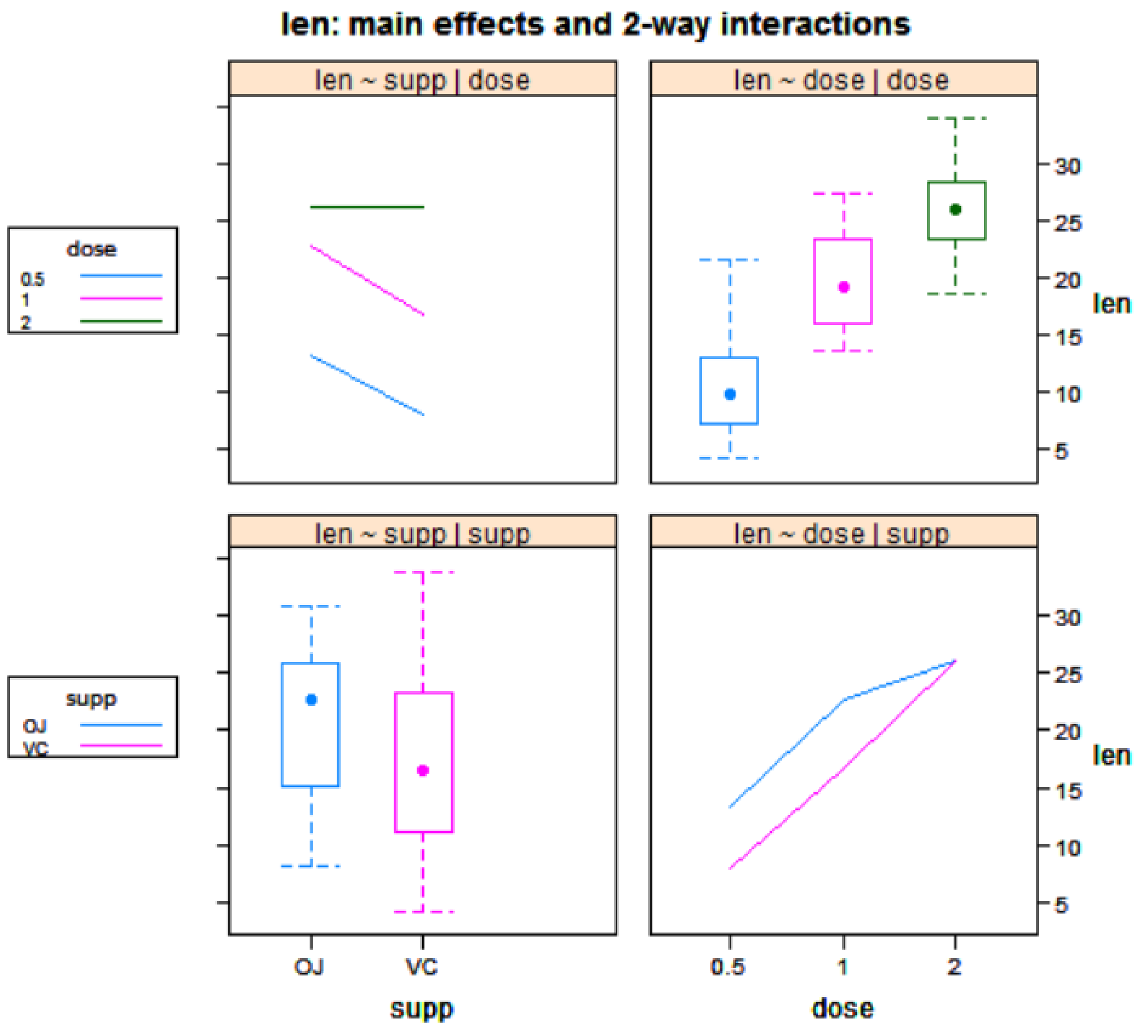
```r
interaction.plot (dose, supp, len, type="b",
         col=c("red","blue"), pch=c(16,18),
         main="Interaction: Dose & Supplement Type")
```






```r
library(gplots)
plotmeans(len~interaction(supp,dose, sep=" "),
         connect=list(c(1,3,5), c(2,4,6)),
         col=c("red","darkgreen"),
         main="Interaction with 95% CI",
         xlab="Treatment & Dose Combination")
```

Visualization: Two-Way Factorial ANOVA




```
library (HH)
interaction2wt (len~supp*dose)
```

Fitting ANOVA Models

REPEATED MEASURES

- ***The same subjects are measured more than once***

Ex: Repeated Measures ANOVA

Subjects are measured more than once

- Factors (Independent variables):
  - Type: Quebec vs Mississippi (Between-Group)
  - Conc: CO_2 concentration (Within-Group)
- Dependent Variable / Response:
  - Uptake: CO_2 uptake by plants

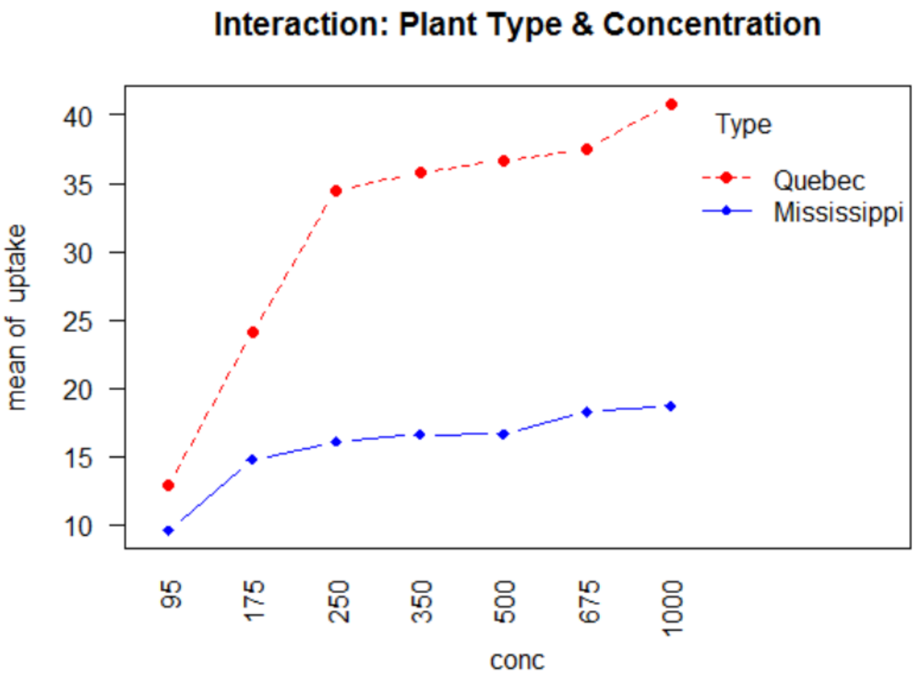
```r
data (CO2)
head(CO2)
w1b1 <- subset(CO2, Treatment=='chilled')
attach(w1b1)
fit <- aov(uptake ~ conc*Type +
             Error(Plant/(conc), w1b1))
summary(fit)
```

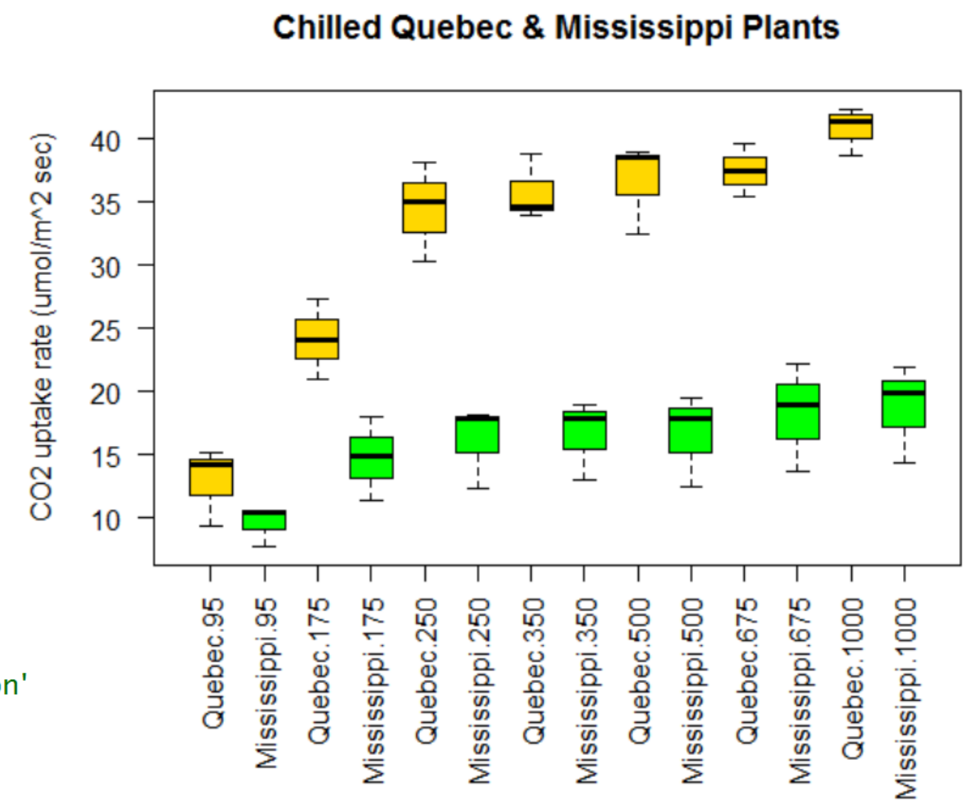
```
Error: Plant
          Df Sum Sq Mean Sq F value  Pr(>F)
Type       1 2667.2  2667.2   60.41 0.00148 **
Residuals  4  176.6    44.1
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05

Error: Plant:conc
          Df Sum Sq Mean Sq F value   Pr(>F)
conc       1  888.6   888.6  215.46 0.000125 ***
conc:Type  1  239.2   239.2   58.01 0.001595 **
Residuals  4   16.5     4.1
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05

Error: Within
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals 30    869   28.97
```

- The Type and concentration main effects are significant
- Type × conc interaction is significant

Visualization: Repeated Measures ANOVA

```
boxplot(uptake ~ Type*conc, data=w1b1,
        col=c("gold","green"),
        main="Chilled Quebec & Mississippi Plants",
        ylab="CO2 uptake rate (umol/m^2 sec)")
```






```
par(las=2)
par(mar=c(10,4,4,2))
with(w1b1,
     interaction.plot (conc, Type, uptake,
      type="b", col=c("red","blue"), pch=c(16,18),
      main="Interaction: Plant Type & Concentration'
```

Experimental Designs

FITTING MANOVA & MANCOVA

There is more than one dependent (outcome) variable

Ex: MANOVA: Multivariate Analysis of Variance

There is more than one dependent (outcome) variable

- Factors (Independent variables):
  - shelf: Store shelf (1 is bottom; 2 is middle, 3 is top) (Between-Group)
- Dependent Variables / Outcome:
  - calories
  - fat
  - sugars

```
library(MASS)
attach(UScereal)
y <- cbind(calories, fat, sugars)
aggregate (y, by=list(shelf), FUN=mean)
```

```
  Group.1 calories       fat    sugars
1       1 119.4774 0.6621338  6.295493
2       2 129.8162 1.3413488 12.507670
3       3 180.1466 1.9449071 10.856821
```

```
> fit <- manova (y ~ shelf)
> summary(fit)
          Df  Pillai approx F num Df den Df  Pr(>F)
shelf      1 0.19594    4.955      3     61 0.00383 **
Residuals 63
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Significant F-value indicates that the three groups differ on the set of nutritional measures
  - Use summary.aov(fit) for pairwise comparison
  - Use TukeyHSD() to determine which shelves differ from each other

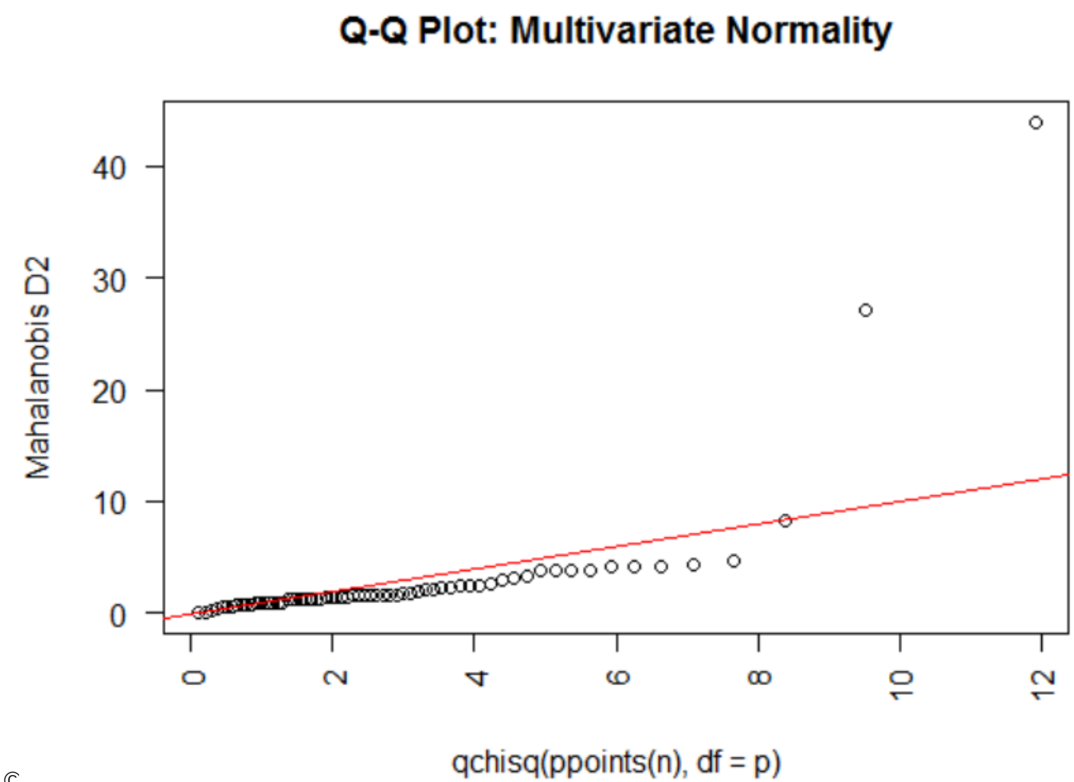
Assumptions: One-way MANOVA

- **Assumptions underlying the statistical tests**:
  - ***Multivariate normality*** of response
  - ***Homogeneity / equality of variance-covariance matrices:***
    - Use Box M test: sensitive to normality assumption
  - ***Multivariate outliers***

- **If any of the assumptions are not met**:
  - Consider robust nonparametric MANOVA provided by Wilks.test() in the "rrcov" package
  - Also, Adonis() function in the "vegan" package

Multivariate Normality Assumption: MANOVA

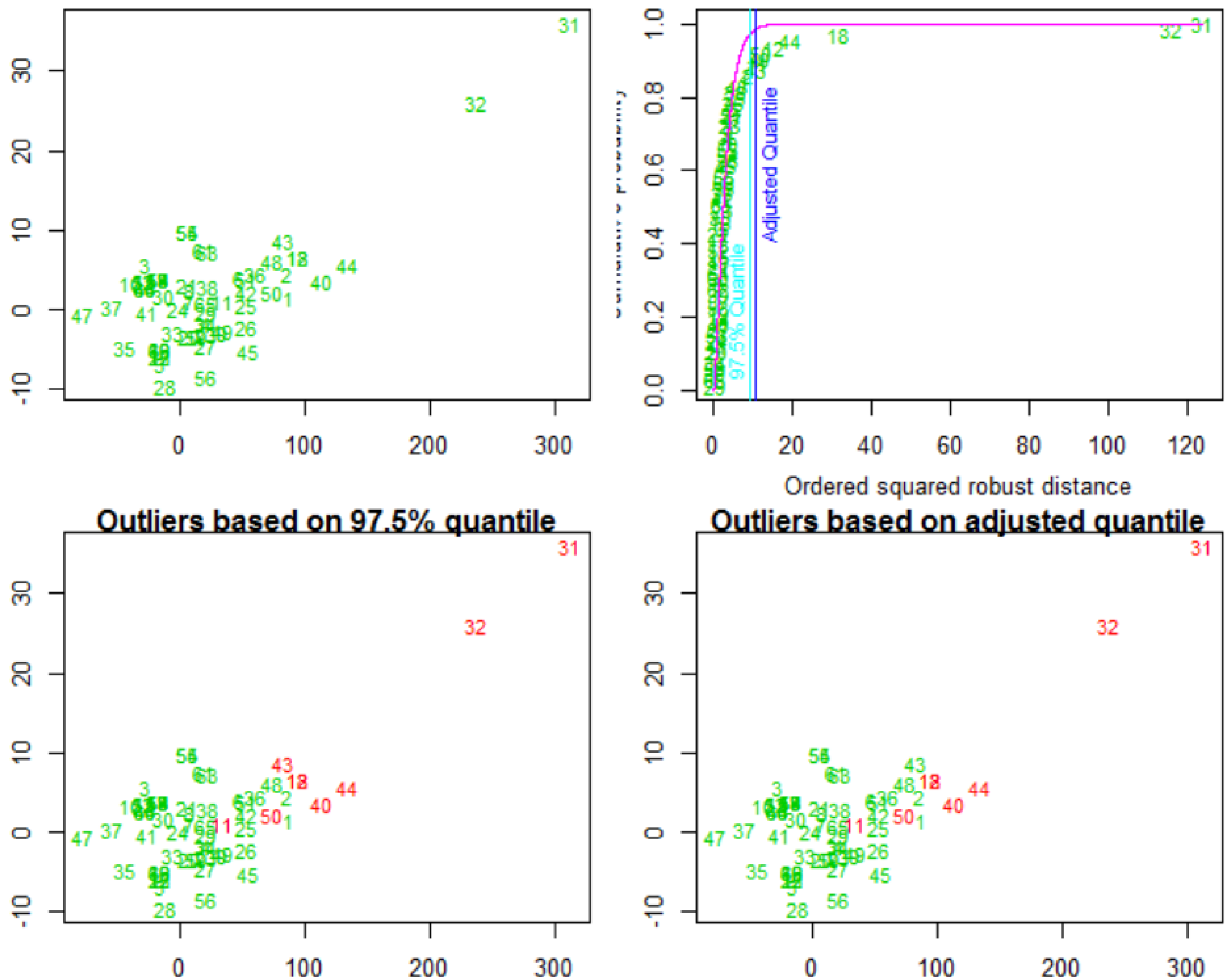
```r
center <- colMeans(y)
n <- nrow(y)
p <- ncol(y)
cov <- cov(y)
d <- mahalanobis (y, center, cov)
coord <- qqplot (qchisq(ppoints(n), df=p),
        d, main="Q-Q Plot: Multivariate Normality",
        ylab="Mahalanobis D2")
abline(a=0,b=1, col=c("red"))
```



- The data appears to violate multivariate normality (due to a couple of outliers):
  - Wheaties Honey Gold
  - Wheaties
- Delete them and re-run

Outliers: One-way MANOVA

```
library(mvoutlier)
outliers <- aq.plot(y)
outliers
```



Robust MANOVA

```
library (rrcov)
Wilks.test(y, shelf, method="mcd")
```

```
 Robust One-way MANOVA (Bartlett Chi2)

data:  x
Wilks' Lambda = 0.5107, Chi2-Value = 23.311, DF =
4.845, p-value = 0.0002549
sample estimates:
  calories       fat    sugars
1 119.8210 0.7010828  5.663143
2 128.0407 1.1849576 12.537533
3 160.8604 1.6524559 10.352646
```

Summary

- Designs
  - One-way Between-Groups ANOVA (Analysis of Variance)
  - One-way Within-Groups ANOVA
  - Two-way Factorial ANOVA
  - ANCOVA: Adjusting for Covariates
  - MANOVA: Multivariate Analysis of Variable (>1 Response Vars)
  - MANCOVA: Multivariate Analysis of Covariance
- Fitting ANOVA and ANCOVA Models in R
  - Formulas for different types of designs
  - Order matters: Type I, II, and III ordering of formula terms
  - Examples: ANOVA and ANCOVA
  - Assessing ANOVA and ANCOVA Assumptions
- Fitting MANOVA Models in R
  - Examples: One-way MANOVA
  - Assessing MANOVA Assumptions