SCIENCE MEETS LIFE

One-way anova

VIB-UGENT CENTER FOR PLANT SYSTEMS BIOLOGY

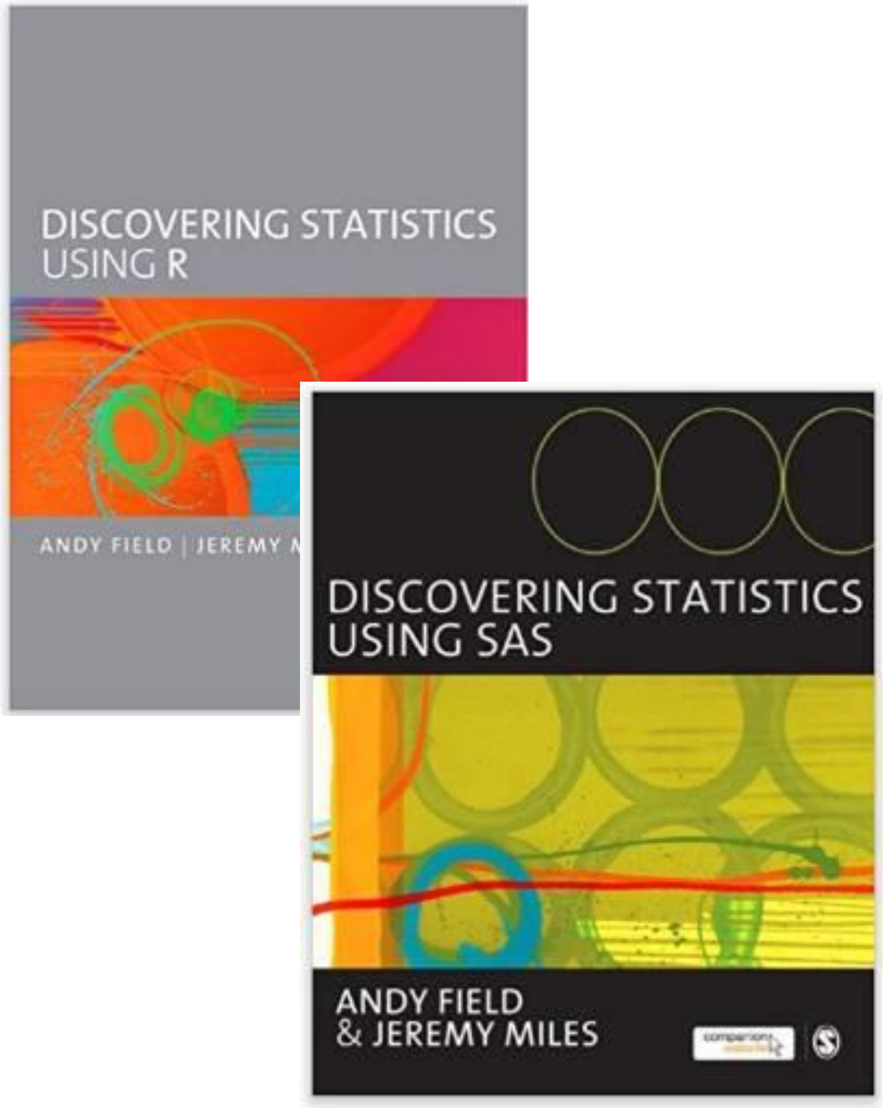
UNIVERSITEIT GENT

23/05/2019

Veronique Storme

References

- Discovering statistics using R” by Andy Field
  http://www.uk.sagepub.com/dsur/main.htm
- SAS course notes



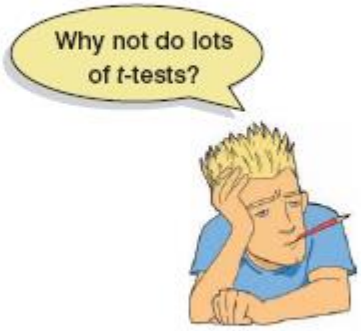
Aims

- Understand the basic principles of ANOVA
  - Why it is done?
  - What it tells us?
- Theory of one-way independent ANOVA
- Following up an ANOVA:
  - *Post hoc* tests

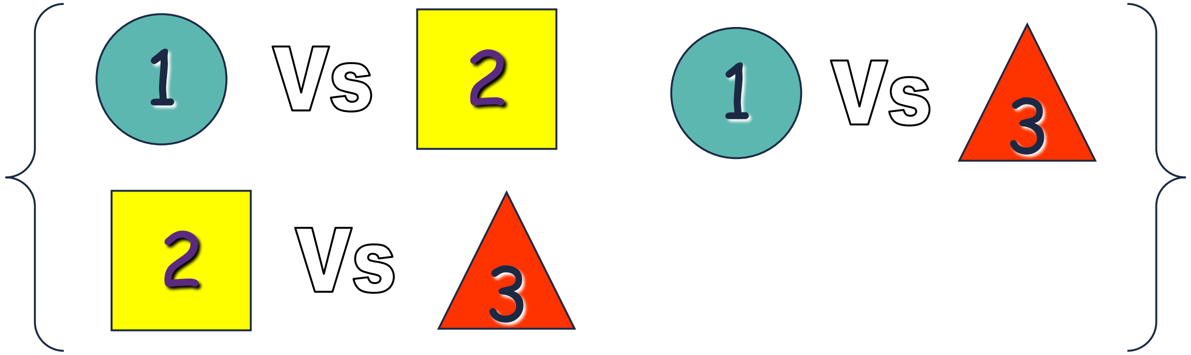
When and Why

- When we want to compare means we can use a *T*-test. This test has limitations:
  - You can compare only 2 means: often we would like to compare means from 3 or more groups.
  - It can be used only with one predictor/independent variable.
- ANOVA
  - Compares several means.
  - Can be used when you have manipulated more than one independent variable.
  - It is an extension of regression (the general linear model).

Why Not Use Lots of *T*-Tests?



- If we want to compare several means why don't we compare pairs of means with *T*-tests?
  - Can't look at several independent variables.
  - Inflates the Type I error rate.



$$\text{familywise error} = 1 - 0.95^n$$

Multiple Comparisons

Comparisonwise Error Rate (α=0.05)	Number of Comparisons	Experimentwise Error Rate (α=0.05)
.05	1	.05
.05	3	.14
.05	6	.26
.05	10	.40

What Does ANOVA Tell Us?

- Null hyothesis:
  - Like a *T*-test, ANOVA tests the null hypothesis that the means are the same.
- Experimental hypothesis:
  - The means differ.
- ANOVA is an omnibus test
  - It test for an overall difference between groups.
  - It tells us that the group means are different.
  - It doesn’t tell us exactly which means differ.

Assumptions ANOVA

1. The measure is interval-level continuous data.
2. The measure is normally distributed (within each group), *i.e.*, Y|X is normally distributed.
3. The variance of the measure is the same in each group.
4. The observations are independent

Note that 2,3 and 4 are the same as saying

$$\epsilon_i \sim i.i.d.N(0, \sigma^2)$$

Theory of ANOVA

- If the experiment is successful, then the model will explain more variance than it can't
  - SS_M will be greater than SS_E

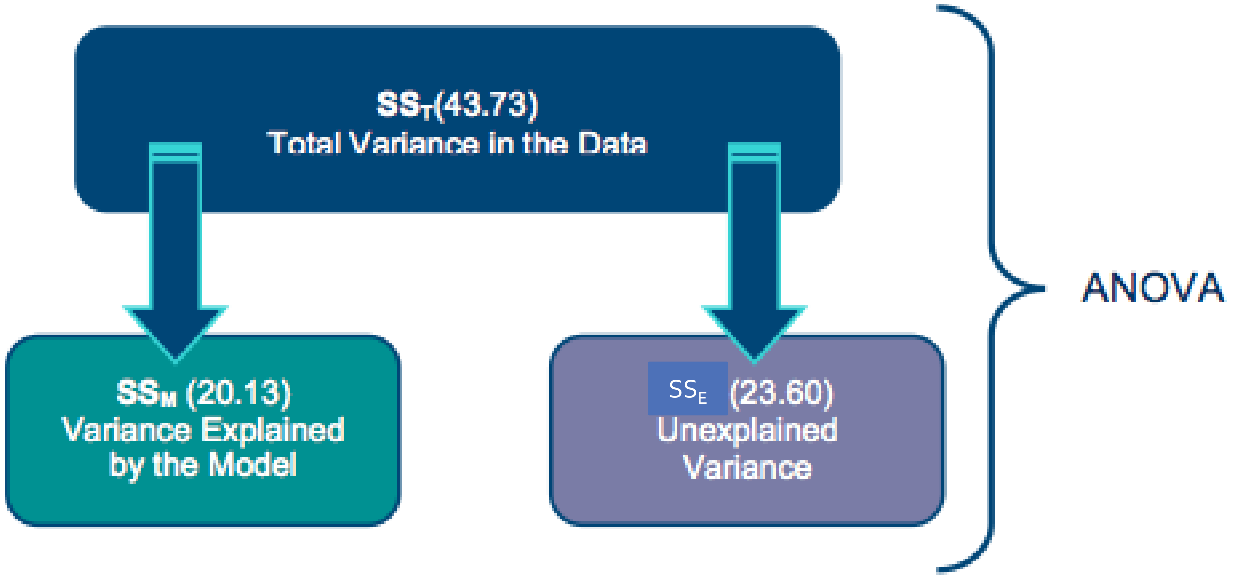


Figure 10.4: Partitioning variance for ANOVA

ANOVA by Hand

- Testing the effects of Viagra on libido using three groups:
  - Placebo (sugar pill)
  - Low dose viagra
  - High dose viagra
- The outcome/dependent variable (DV) was an objective measure of libido.

The Data

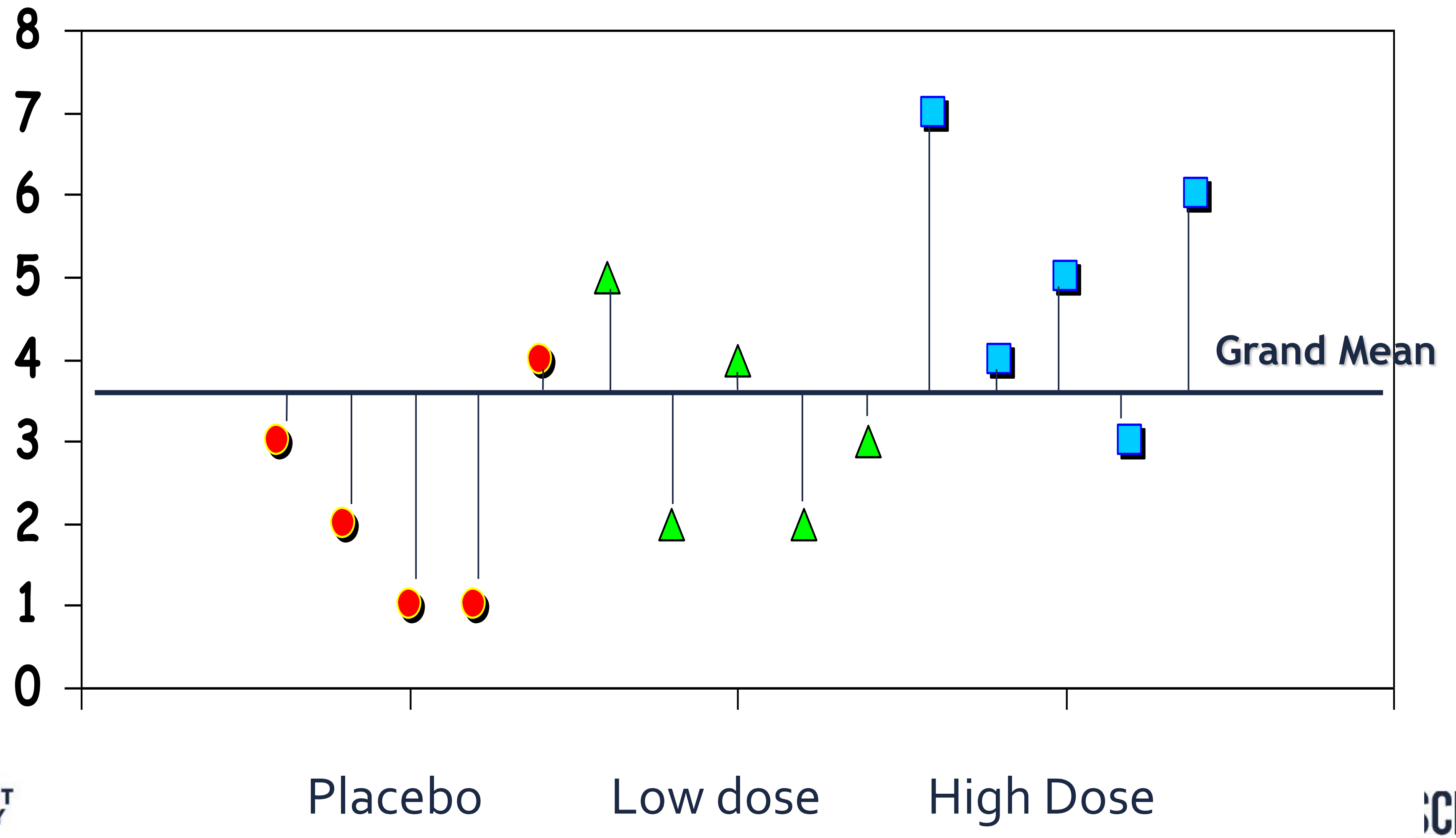
Table 10.1: Data in **Viagra.dat**

	Placebo	Low Dose	High Dose
x_{ij}	3	5	7
	2	2	4
	1	4	5
	1	2	3
	4	3	6
$\bar{x}_{.j}$	**2.20**	**3.20**	**5.00**
s_j	**1.30**	**1.30**	**1.58**
s_j^2	**1.70**	**1.70**	**2.50**
Grand Mean = **3.467** Grand SD = **1.767** $\bar{x}_{..}$ Grand Variance = **3.124** s^2			

VIB-UGENT CENTER FOR SYSTEMS BI

SCIENCE MEETS LIFE

Total Sum of Squares (SS_T):



Step 1: Calculate SS_T

$$SS_T = \sum_{i=1}^{N} (x_i - \bar{x}_{..})^2$$

$$s^2 = \frac{SS}{N-1}$$

$$SS = s^2(N-1)$$

$$SS_T = s^2(N-1)$$

$$SS_T = 3.124(15-1)$$

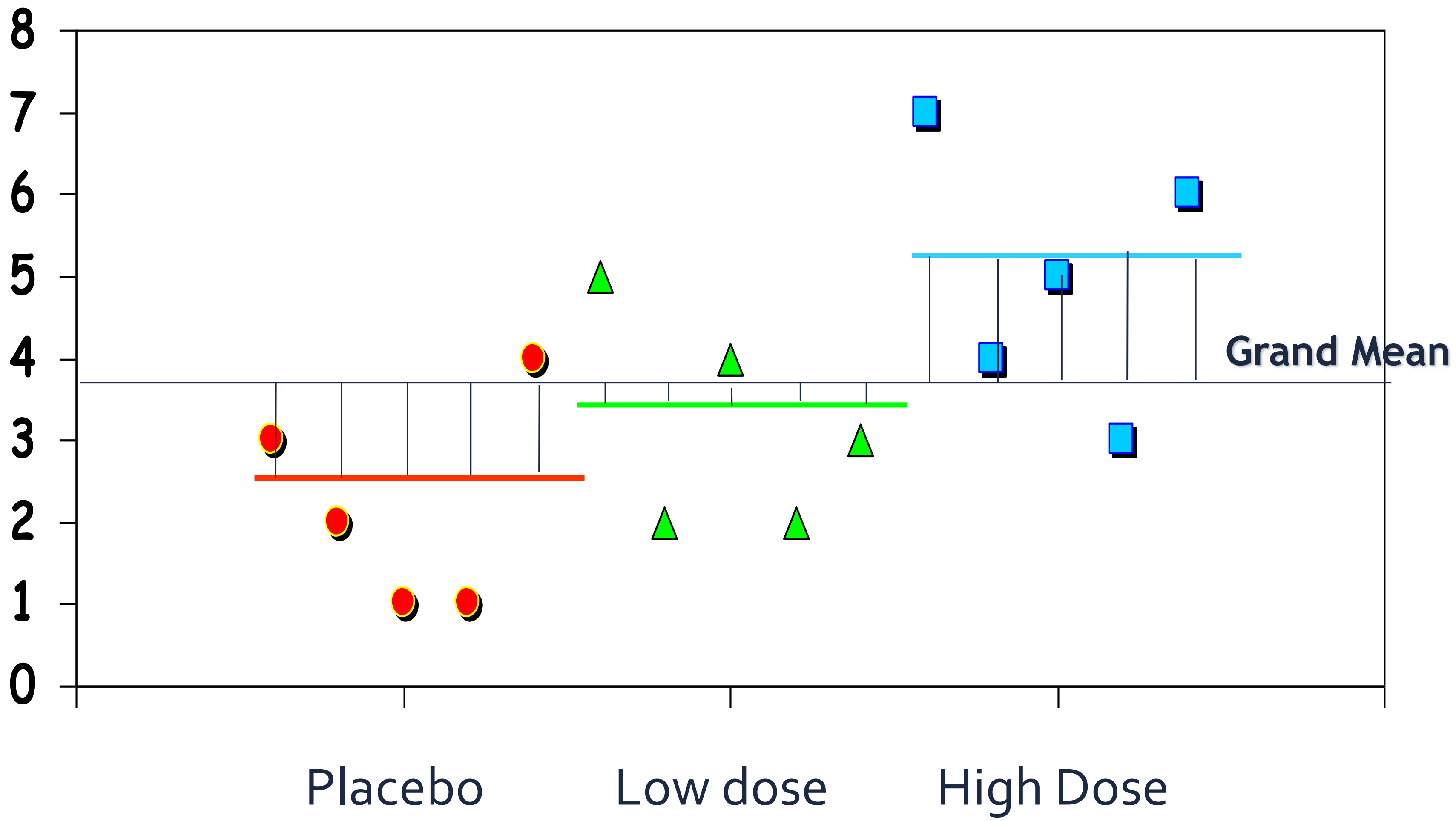
$$SS_T = 43.74$$

Degrees of Freedom

- Degrees of freedom (*df*) are the number of values that are free to vary.
- In general, the *df* are one less than the number of values used to calculate the SS.

$$df_T = N - 1 = 15 - 1 = 14$$

Model Sum of Squares (SS_M):



Step 2: Calculate SS_M

$$SS_M = \sum_{j=1}^{k} n_j(\bar{x}_{.j} - \bar{x}_{..})^2$$

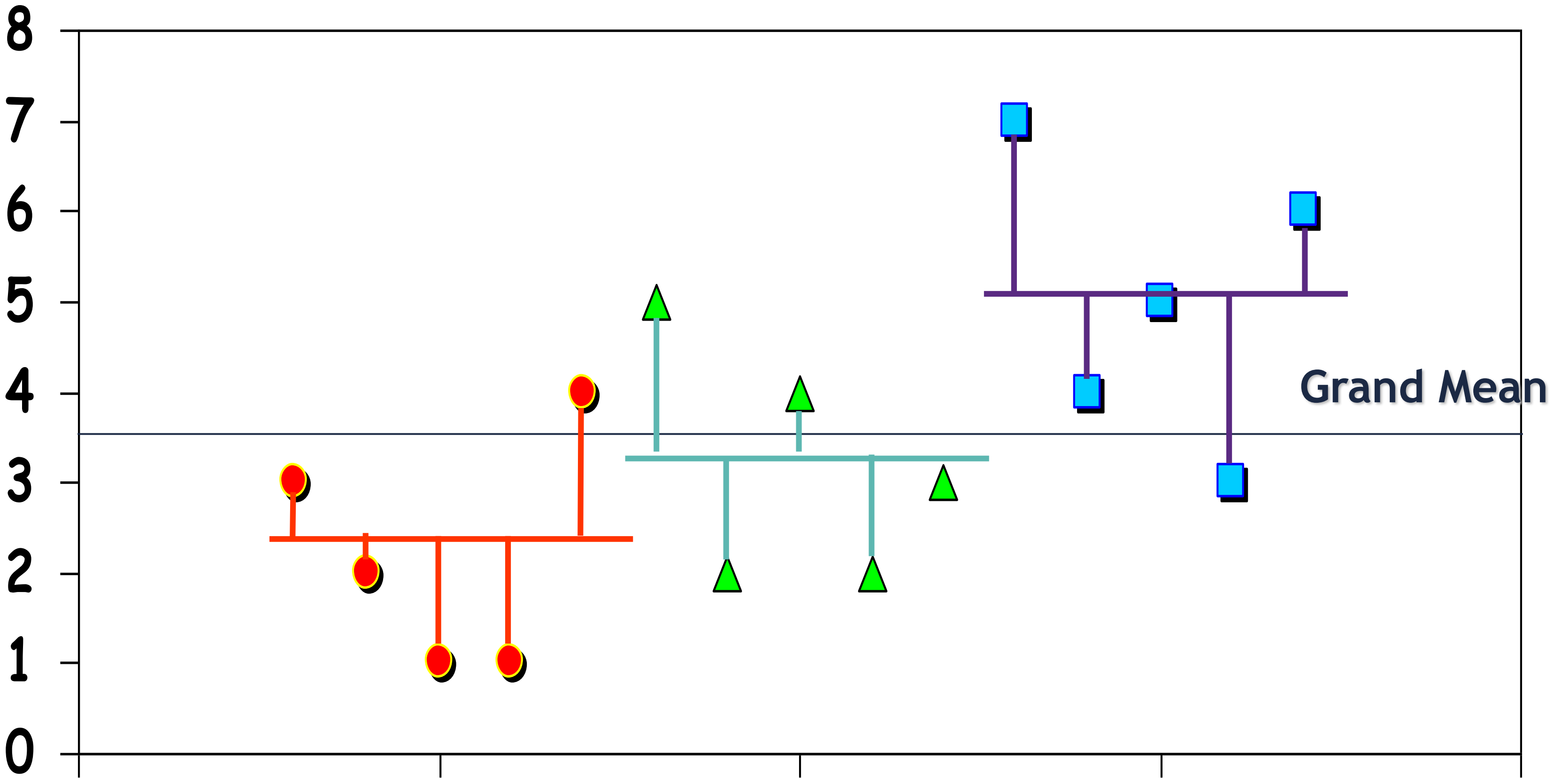
$$SS_M = 5(2.2 - 3.467)^2 + 5(3.2 - 3.467)^2 + 5(5.0 - 3.467)^2$$
$$SS_M = 20.135$$

Model Degrees of Freedom

- How many values did we use to calculate SS_M?
  - We used the 3 means.

$$df_M = k - 1 = 3 - 1 = 2$$

Residual Sum of Squares (SS_E):



Placebo Df = 4

Low dose Df = 4

High Dose Df = 4

Step 3: Calculate SS_E

$$SS_E = \sum_{j=1}^{k} \sum_{i=1}^{N} (x_{ij} - \bar{x}_{.j})^2$$

$$s^2 = \frac{SS}{N-1}$$

$$SS = s^2(N-1)$$

$$SS_E = \sum_{j=1}^{k} s_j^2(n_j - 1)$$

$$SS_E = s_1^2(n_1 - 1) + s_2^2(n_2 - 1) + s_3^2(n_3 - 1)$$

Step 3: Calculate SS_E

$$
\begin{aligned}
SS_E &= s_1^2(n_1 - 1) + s_2^2(n_1 - 1) + s_3^2(n_1 - 1) \\
&= 1.70(5 - 1) + 1.70(5 - 1) + 2.50(5 - 1) \\
&= 23.60
\end{aligned}
$$

Residual Degrees of Freedom

- How many values did we use to calculate SS_E?
  - We used the 5 scores for each of the SS for each group.

$$df_E = df_1 + df_2 + df_3$$
$$df_E = (n_1 - 1) + (n_2 - 1) + (n_3 - 1)$$
$$df_E = N - k = 12$$

Double Check

$$SS_T = SS_M + SS_E$$
$$43.74 = 20.14 + 23.60$$
$$43.74 = 43.74$$

$$df_T = df_M + df_E$$
$$14 = 2 + 12$$

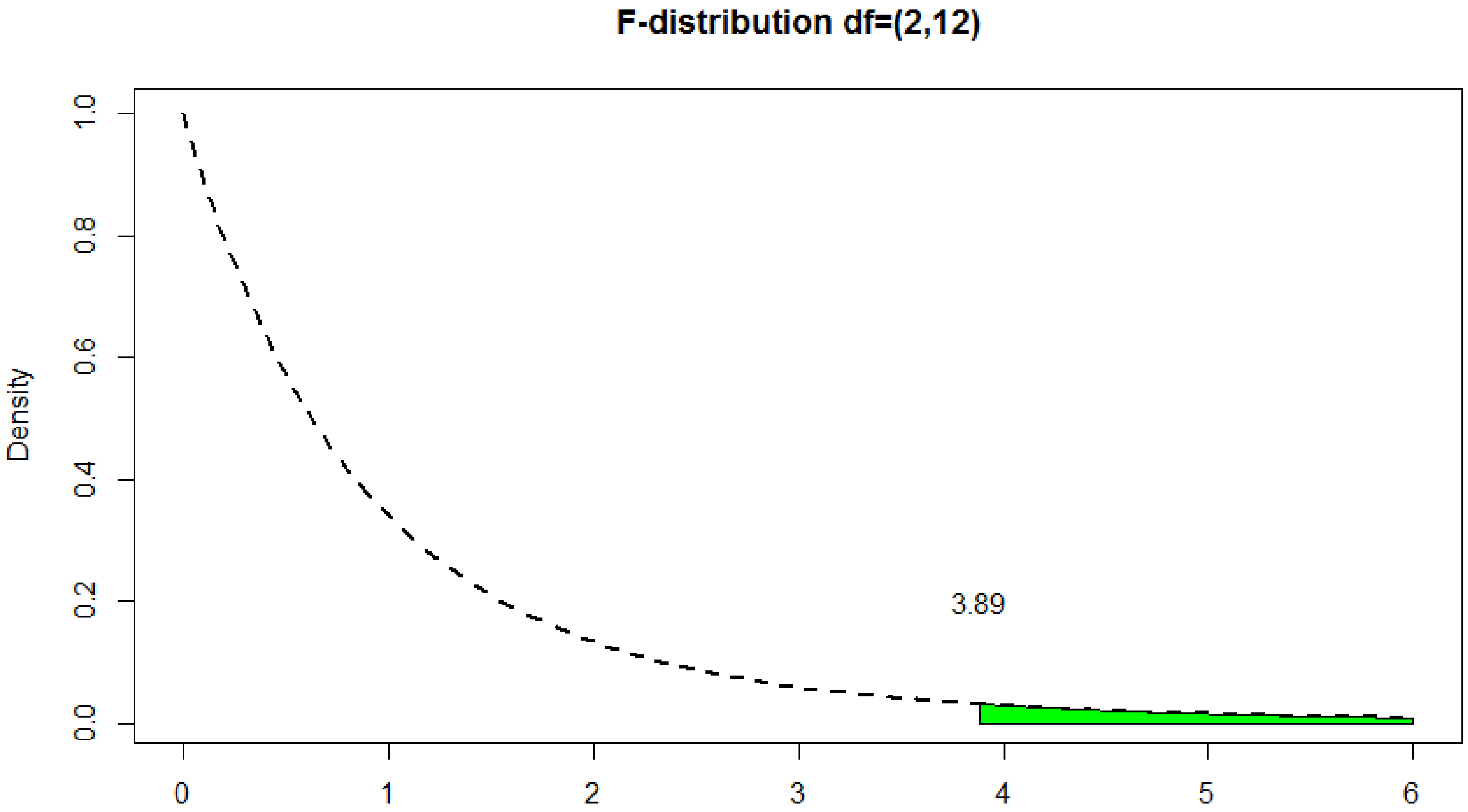
Step 4: Calculate the Mean Squares

$$MS_M = \frac{SS_M}{df_M} = \frac{20.135}{2} = 10.067$$
$$MS_E = \frac{SS_E}{df_E} = \frac{23.60}{12} = 1.967$$

Step 5: Calculate the *F*-Ratio

$$F = \frac{MS_M}{MS_E} = \frac{10.067}{1.967} = 5.12$$

F Statistic and Critical Values at α=0.05



$$F(\text{Model df, Error df}) = MS_M \,/\, MS_E$$

Step 6: Construct a Summary Table

Source	SS	*df*	MS	*F*
Model	20.14	2	10.067	5.12*
Residual	23.60	12	1.967	
Total	43.74	14		

Workflow

- Summary statistics
- Graphs (boxplots, barcharts)
- Statistical model
- Test the assumptions
- Post-hoc tests

One-Way ANOVA using SAS: output

analysis of variance with Levene's test for equality of variances

The GLM Procedure

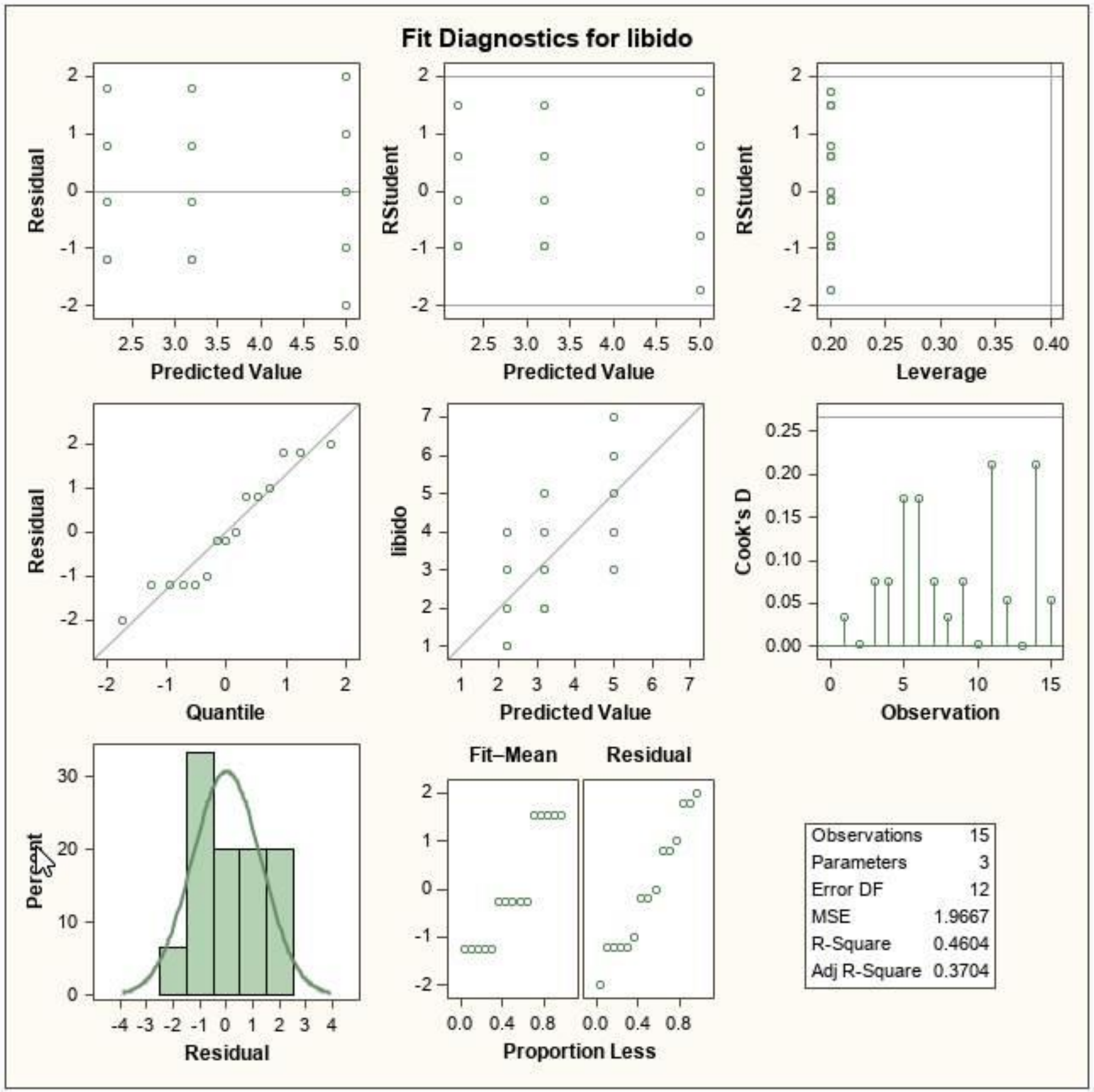
Dependent Variable: libido

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	20.13333333	10.06666667	5.12	0.0247
Error	12	23.60000000	1.96666667		
Corrected Total	14	43.73333333			

R-Square	Coeff Var	Root MSE	libido Mean
0.460366	40.45324	1.402379	3.466667

Source	DF	Type III SS	Mean Square	F Value	Pr > F
dose	2	20.13333333	10.06666667	5.12	0.0247

Verifying the assumptions



Why Use Follow-Up Tests?

- The *F*-ratio tells us only that the experiment was successful
  - i.e. group means were different
- It does not tell us specifically which group means differ from which.
- We need additional tests to find out where the group differences lie.

Multiple Comparison Methods

Control Comparisonwise Error Rate → Pairwise t-tests

Control Experimentwise Error Rate → Compare All Pairs: Tukey; Compare to Control: Dunnett

VIB-UGENT CENTER FOR PLANT SYSTEMS BIOLOGY

UNIVERSITEIT GENT

SCIENCE MEETS LIFE

Post Hoc Tests

- Compare each mean against all others.
- In general terms they use a stricter criterion to accept an effect as significant.
  - Hence, control the familywise error rate.
  - Simplest example is the Bonferroni method:

$$\boxed{\text{Bonferroni } \alpha = \frac{\alpha}{\text{number of tests}}}$$

Tukey's Multiple Comparison Method

- This method is appropriate when you consider pairwise comparisons only.
- The experimentwise error rate is
  - equal to alpha when ***all*** pairwise comparisons are considered
  - less than alpha when ***fewer*** than all pairwise comparisons are considered.

Special Case of Comparing to a Control

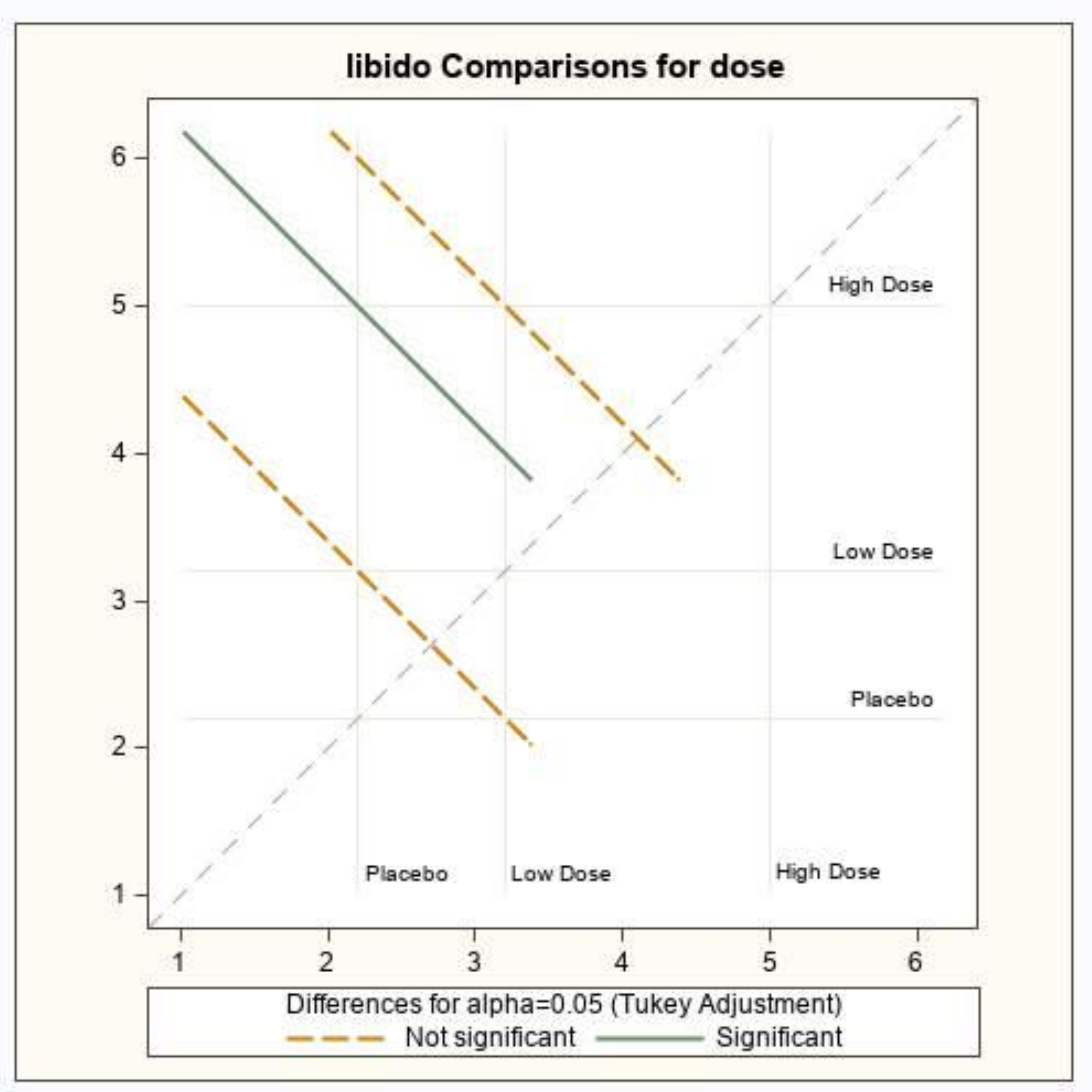
- Comparing to a control is appropriate when there is a natural reference group, such as a placebo group in a drug trial.
  - Control comparison computes and tests k-1 groupwise differences, where k is the number of levels of the classification variable.
  - An example is the *Dunnett* method.

Output all-pairwise comparisons

dose Least Squares Means								
dose	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
High Dose	5.0000	0.6272	12	7.97	<.0001	0.05	3.6335	6.3665
Low Dose	3.2000	0.6272	12	5.10	0.0003	0.05	1.8335	4.5665
Placebo	2.2000	0.6272	12	3.51	0.0043	0.05	0.8335	3.5665

Differences of dose Least Squares Means Adjustment for Multiple Comparisons: Tukey												
dose	_dose	Estimate	Standard Error	DF	t Value	Pr > |t|	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
High Dose	Low Dose	1.8000	0.8869	12	2.03	0.0652	0.1475	0.05	-0.1325	3.7325	-0.5662	4.1662
High Dose	Placebo	2.8000	0.8869	12	3.16	0.0083	0.0209	0.05	0.8675	4.7325	0.4338	5.1662
Low Dose	Placebo	1.0000	0.8869	12	1.13	0.2816	0.5163	0.05	-0.9325	2.9325	-1.3662	3.3662

diffogram



VIB-UGENT CENTER FOR PLANT SYSTEMS BIOLOGY

UNIVERSITEIT GENT

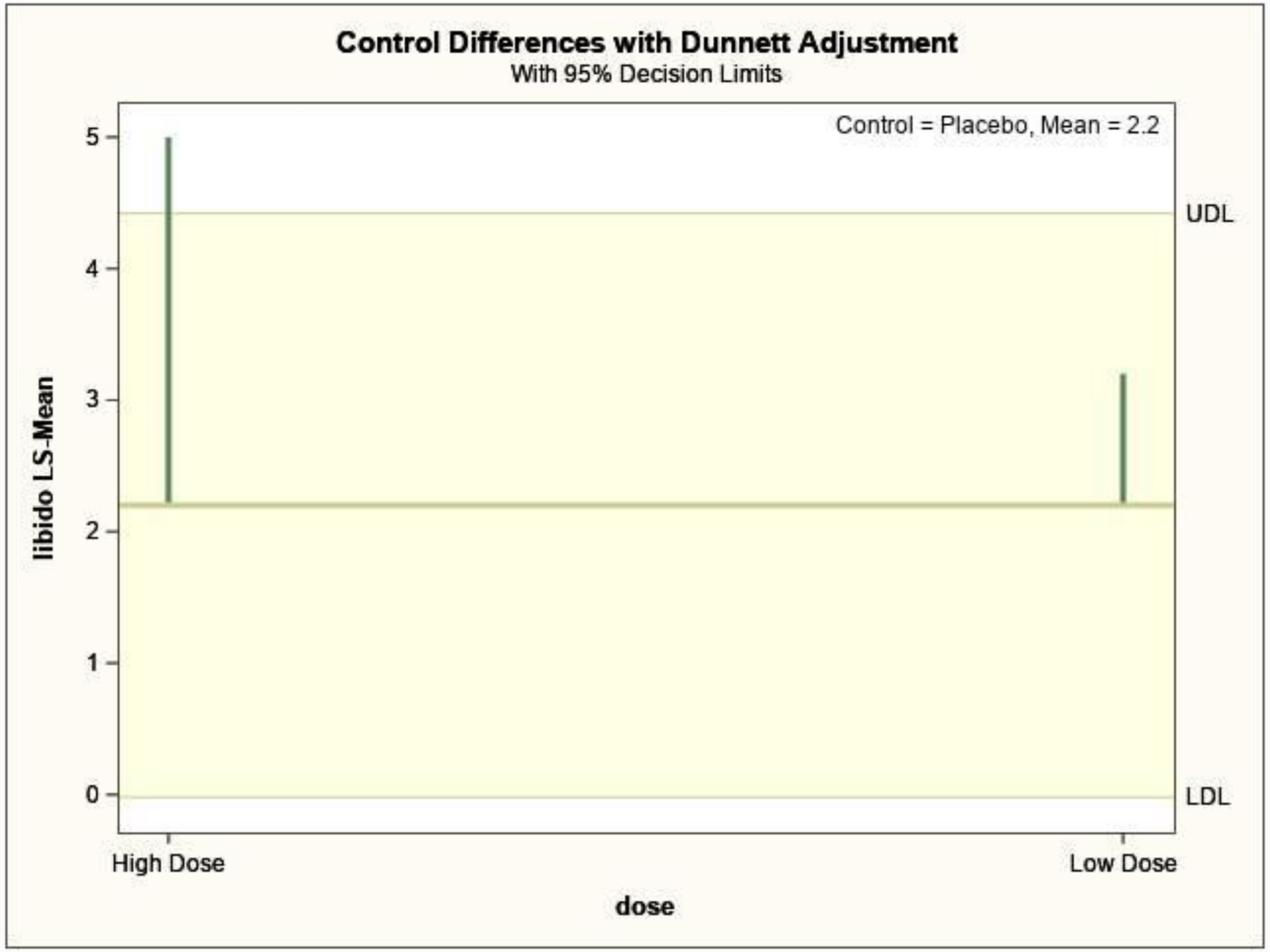
SCIENCE MEETS LIFE

Output comparisons vs a control

dose Least Squares Means								
dose	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
High Dose	5.0000	0.6272	12	7.97	<.0001	0.05	3.6335	6.3665
Low Dose	3.2000	0.6272	12	5.10	0.0003	0.05	1.8335	4.5665
Placebo	2.2000	0.6272	12	3.51	0.0043	0.05	0.8335	3.5665

Differences of dose Least Squares Means Adjustment for Multiple Comparisons: Dunnett												
dose	_dose	Estimate	Standard Error	DF	t Value	Pr > |t|	Adj P	Alpha	Lower	Upper	Adj Lower	Adj Upper
High Dose	Placebo	2.8000	0.8869	12	3.16	0.0083	0.0152	0.05	0.8675	4.7325	0.5805	5.0195
Low Dose	Placebo	1.0000	0.8869	12	1.13	0.2816	0.4459	0.05	-0.9325	2.9325	-1.2195	3.2195

Output comparisons vs a control



VIB-UGENT CENTER FOR PLANT SYSTEMS BIOLOGY

UNIVERSITEIT GENT

SCIENCE MEETS LIFE

When variances are not equal across groups

- If Levene’s test is significant then it is reasonable to assume that population variances are different across groups.
- Welch anova (variance weighted)
- Mixed models allow estimation of heterogeneous variances
- Unless the group variances are extremely different or the number of groups is large, the usual ANOVA test is relatively robust when the groups are all about the same size.

Quiz: anova as regression

- In the viagra dataset we compare 3 groups. How many dummy variables do we have in the regression model?
  a) 1
  b) 2
  c) 3
  d) 4

Quiz: anova as regression

- In the viagra dataset we compare 3 groups. How many dummy variables doe we have in the regression model
  a) 1
  b) 2
  c) 3
  d) 4

GLM Coding of CLASS Variables (default)

CLASS	Value	X_1	X_2	X_3
dose	High Dose	1	0	0
	Low Dose	0	1	0
	Placebo	0	0	1

Model is **overparameterized**

SAS ignores the last level, making it the ref level

	Parameter	Estimate		Standard Error	t Value	Pr > |t|
b0	Intercept	2.200000000	B	0.62716292	3.51	0.0043
b1	dose High Dose	2.800000000	B	0.88694231	3.16	0.0083
b2	dose Low Dose	1.000000000	B	0.88694231	1.13	0.2816
b3	dose Placebo	0.000000000	B	.	.	.

ANOVA as Regression

$$libido_i = b_0 + b_1 X_1 + b_2 X_2 + e_i$$

X_1 = 1 if observation belongs to the High Dose group and 0 otw

X_2 = 1 if observation belongs to the Low Dose group and 0 otw

$$E(libido|Placebo) = b_0 = \bar{x}_{placebo}$$

$$E(libido|High) = b_0 + b_1 = \bar{x}_{high}$$

$$E(libido|Low) = b_0 + b_2 = \bar{x}_{low}$$

Parameter	Estimate		Standard Error	t Value	Pr > |t|
Intercept	2.200000000	B	0.62716292	3.51	0.0043
dose High Dose	2.800000000	B	0.88694231	3.16	0.0083
dose Low Dose	1.000000000	B	0.88694231	1.13	0.2816
dose Placebo	0.000000000	B	.	.	.

lsmeans

- Least-squares means are predictions from a linear model
  - Balanced data: lsmeans = arithmetic group means
    - All SE are equal (due to homogeneity of variance assumption)
  - Unbalanced data: lsmeans ≠ arithmetic group means, adjusted for imbalance

Multiple Choice Poll

- If you have 20 observations in your ANOVA and you calculate the residuals, to which of the following would they sum?

a. -20
b. 0
c. 20
d. 400
e. Unable to tell from the information given

Multiple Choice Poll – Correct Answer

- If you have 20 observations in your ANOVA and you calculate the residuals, to which of the following would they sum?

a. -20
b. 0
c. 20
d. 400
e. Unable to tell from the information given

Multiple Choice Poll

- If you have 20 observations in your ANOVA and you calculate the squared residuals, to which of the following would they sum?

a. -20
b. 0
c. 20
d. 400
e. Unable to tell from the information given

Multiple Choice Poll – Correct Answer

- If you have 20 observations in your ANOVA and you calculate the squared residuals, to which of the following would they sum?

a. -20
b. 0
c. 20
d. 400
e. Unable to tell from the information given

Multiple Choice Poll

- Which part of the ANOVA tables contains the variation due to nuisance factors?

a. Sum of Squares Model
b. Sum of Squares Error
c. Degrees of Freedom

Multiple Choice Poll – Correct Answer

- Which part of the ANOVA tables contains the variation due to nuisance factors?

a. Sum of Squares Model
b. Sum of Squares Error
c. Degrees of Freedom

Multiple Answer Poll

- A study is conducted to compare the average monthly credit card spending for males versus females. Which statistical method might be used?

    a. One-sample *t*-test
    b. Two-sample *t*-test
    c. One-way ANOVA
    d. Two-way ANOVA

Multiple Answer Poll – Correct Answers

- A study is conducted to compare the leaf area of leaf 4 of B73 versus Mo17. Which statistical method might be used?

  a. One-sample *t*-test
  b. Two-sample *t*-test
  c. One-way ANOVA
  d. Two-way ANOVA