October 7, 2020

Advanced topics in Biostatistics 2020/2021: Analysis of Variance

Silvia Calderazzo
Division of Biostatistics - C060
s.calderazzo@dkfz-heidelberg.de

Outline

Problem set-up: compare the means of k groups.

- Comparison when $k = 2$ (t-test)
- Comparison when $k \geq 3$ (**An**alyis **O**f **Va**riance - ANOVA)
  - Data are grouped according to the levels of **one** factor: one-way ANOVA
  - Data are grouped according to the levels of **two** factors: two-way ANOVA
- Post-hoc tests
- Assumptions: check & remedies
- Other applications of ANOVA
- Conclusions

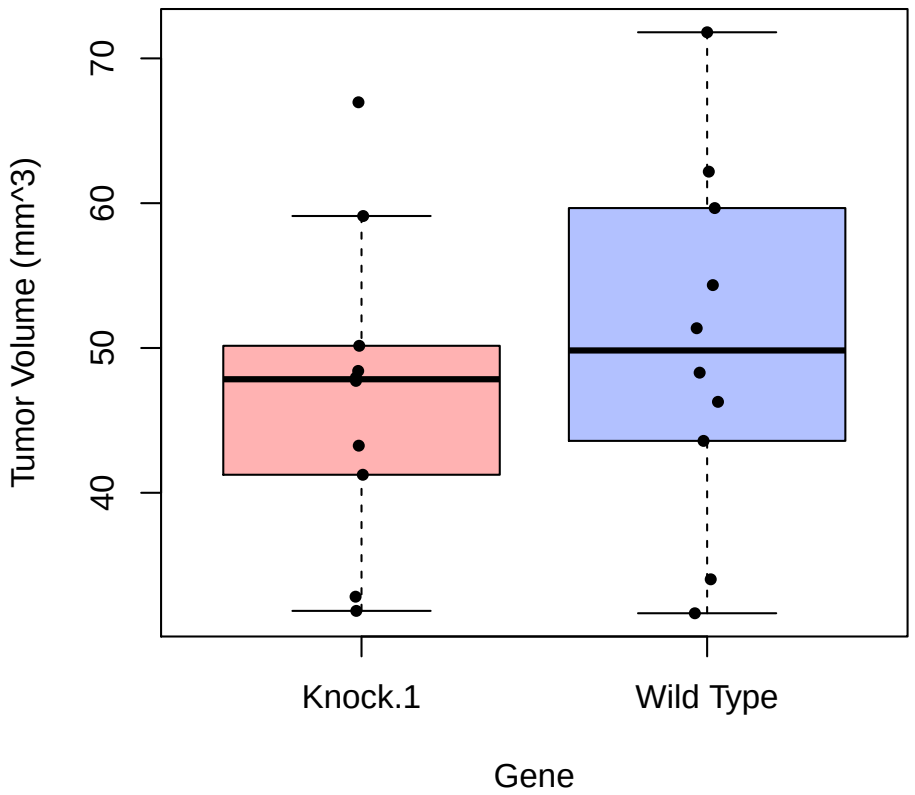
Compare the means of two groups: example

Suppose we measure tumor volume in two groups:

- Wild Type (Sample mean=50.32, Sample SD=12.4, $n_{WT} = 10$).
- Knock 1 (Sample mean=46.95, Sample SD=10.7, $n_{K1} = 10$);

Aim: we wish to know if belonging to wild type or knock 1 makes a difference on mean tumor volume.

Approach: a t-test to compare two means.



t-test: the formal set-up

Assumptions

- Let Y_{ij} be observation i, $i = 1, ..., n_j$ in group j, $j = 1, 2$ (e.g. tumor volume in group j={Knock 1,Wild Type});
- $Y_{ij} \sim N(\mu_j, \sigma^2)$, and are independent.

Set of hypotheses

- We wish to test at level α

$$H_0 : \mu_1 = \mu_2 \qquad \textbf{vs} \qquad H_1 : \mu_1 \neq \mu_2$$

- The question is whether the data provide enough evidence to reject H_0.

Test: t-test

$$T = \frac{\bar{Y}_{.1} - \bar{Y}_{.2}}{\hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \overset{H_0, assumpt.}{\sim} t_{n_1+n_2-2}$$

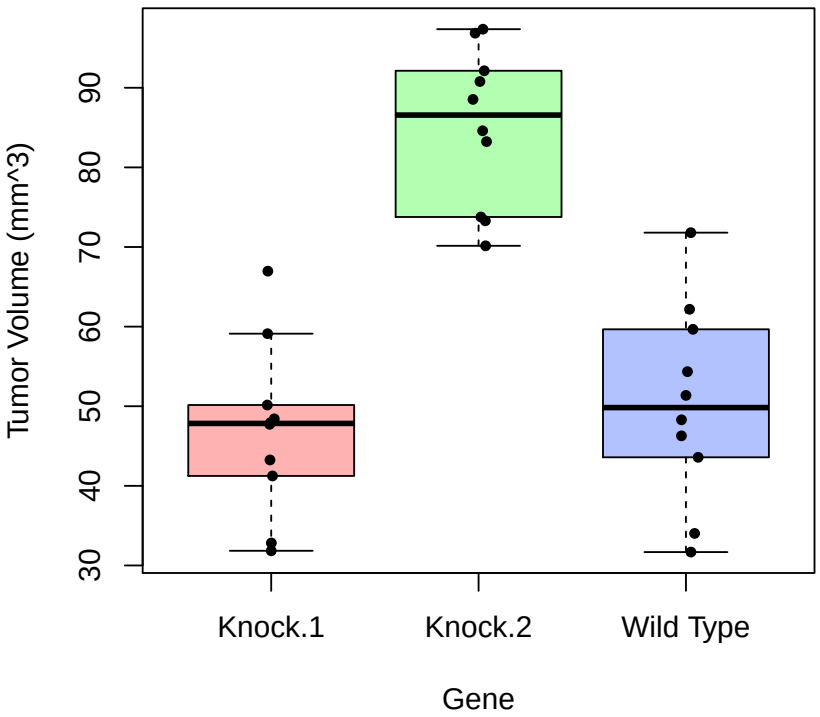
($\bar{Y}_{.j}$: mean of group j; $\hat{\sigma}$: pooled estimated standard deviation)

...and compare the observed $|T|$ with the $1 - \alpha/2$ quantile of a $t_{n_1+n_2-2}$ distribution (large values -> low p-values).

When $k \geq 3$: example

Suppose we measure a variable (e.g. tumor volume) in **multiple** groups (e.g. three) representing the levels of **one factor** (e.g. gene type):

- Wild Type (Sample mean=50.32, Sample SD=12.4, $n_{WT} = 10$).
- Knock 1 (Sample mean=46.95, Sample SD=10.7, $n_{K1} = 10$);
- Knock 2 (Sample mean=85.07, Sample SD=9.9, $n_{K1} = 10$);

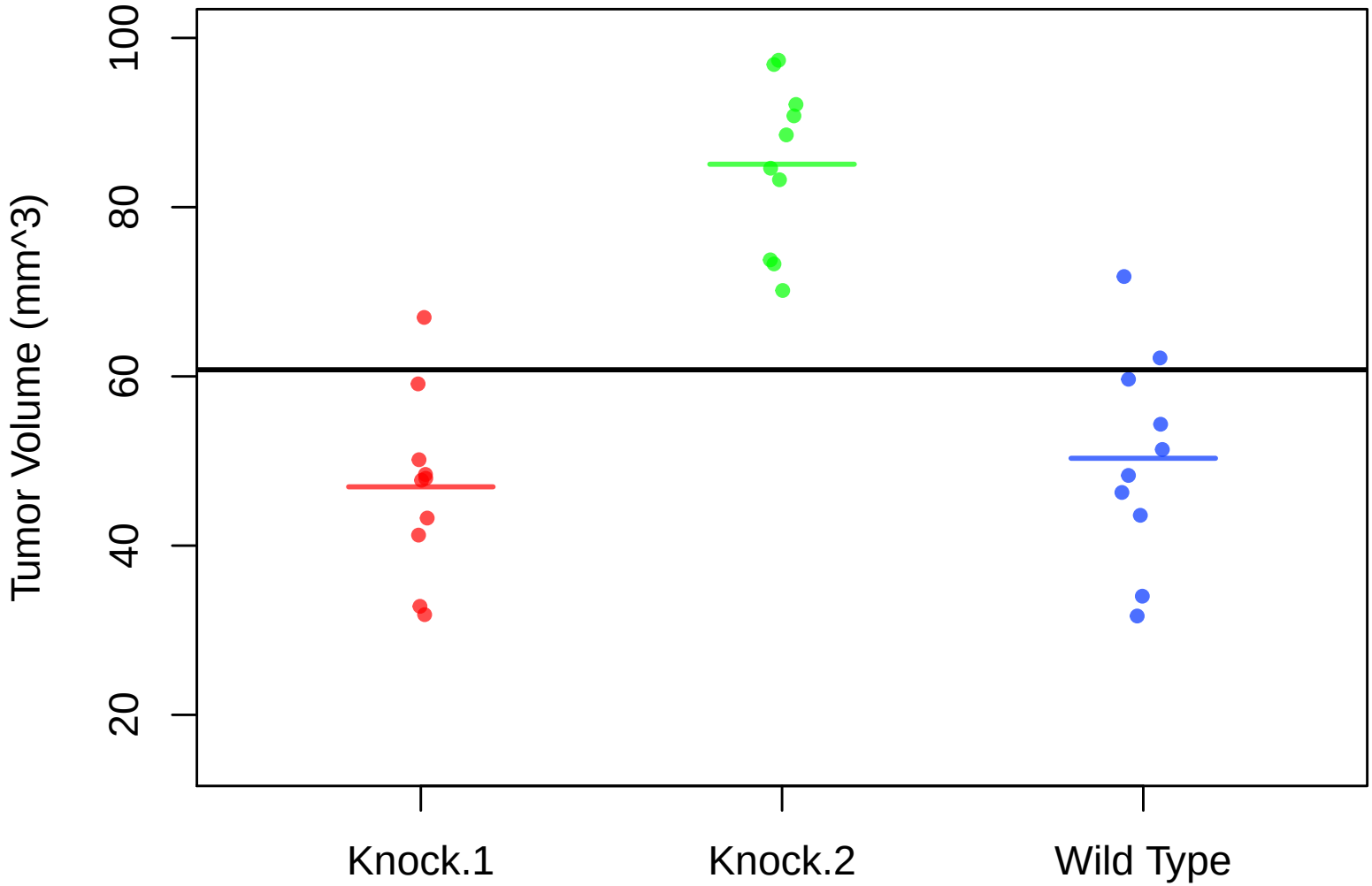


Aim: we wish to know if the gene has an effect on tumor volume, and in particular if:

1. At least one of the means is different from the others;
2. Which means are different.

Approach: One-way ANOVA answers question (1). Post-hoc analyses answer question (2).

Variance decomposition: intuition



Variance decomposition: 'the maths'

Let Y_{ij} be observation $i = 1, ..., n_j$ in group j, $j = 1, ..., k$ (e.g. tumor volume in group j={Knock 1, Knock 2, Wild Type});

Then,

$$\underbrace{\sum_{j=1}^{k}\sum_{i=1}^{n_j}(Y_{ij} - \bar{Y}_{..})^2}_{Total\ Deviance\ (SST)} = \underbrace{\sum_{j=1}^{k} n_j(\bar{Y}_{.j} - \bar{Y}_{..})^2}_{Between\ Deviance\ (SSB)} + \underbrace{\sum_{j=1}^{k}\sum_{i=1}^{n_j}(Y_{ij} - \bar{Y}_{.j})^2}_{Residual\ Deviance\ (SSR)},$$

where $\bar{Y}_{.j}$ is the mean of the observations in group j, and $\bar{Y}_{..}$ the overall mean.

ANOVA test

Assumptions

- $Y_{ij} \sim N(\mu_j, \sigma^2)$, where $\mu_j = \eta + \gamma_j$ and η is the global mean, and are independent.
- *Equivalently:* $Y_{ij} = \eta + \gamma_j + \epsilon_{ij}$, *where* $\epsilon_{ij} \sim N(0, \sigma^2)$ *and are independent.*

Set of hypotheses

- We wish to test at level α

$H_0 : \gamma_1 = \gamma_2 = \cdots = \gamma_k (= 0)$ **vs** H_1: at least one γ_j is different (from zero)

Test: *F*-test

$$T = \frac{\sum_{j=1}^{k} n_j (\bar{Y}_{.j} - \bar{Y}_{..})^2 / (k-1)}{\sum_{j=1}^{k} \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2 / (n-k)} \overset{H_0, assumpt.}{\sim} F_{k-1, n-k},$$

where $n = \sum_{j=1}^{k} n_j$.

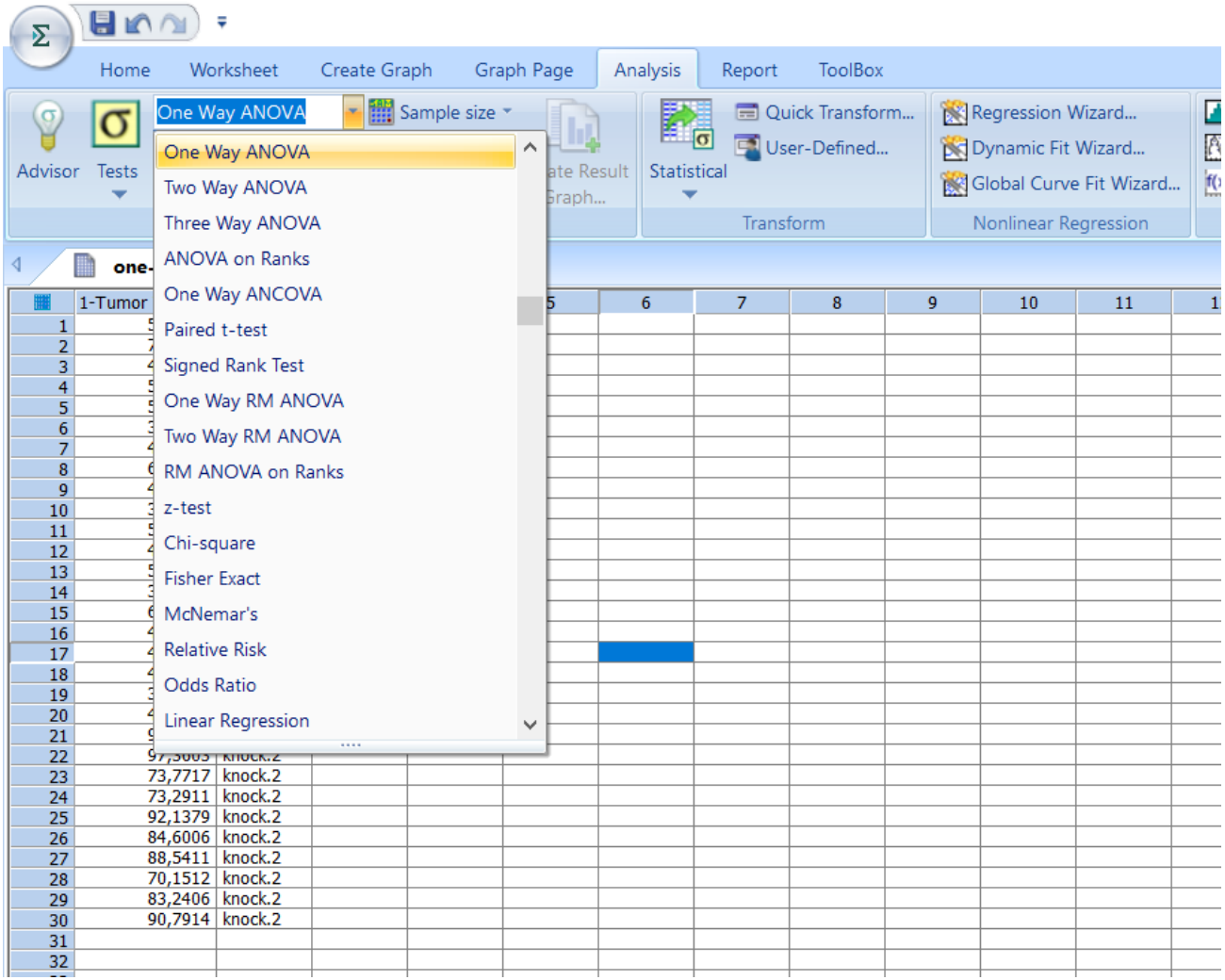
Compare the observed T with the $1 - \alpha$ quantile of a $F_{k-1,n-k}$ distribution (large values -> low *p*-values).

One-way ANOVA: output

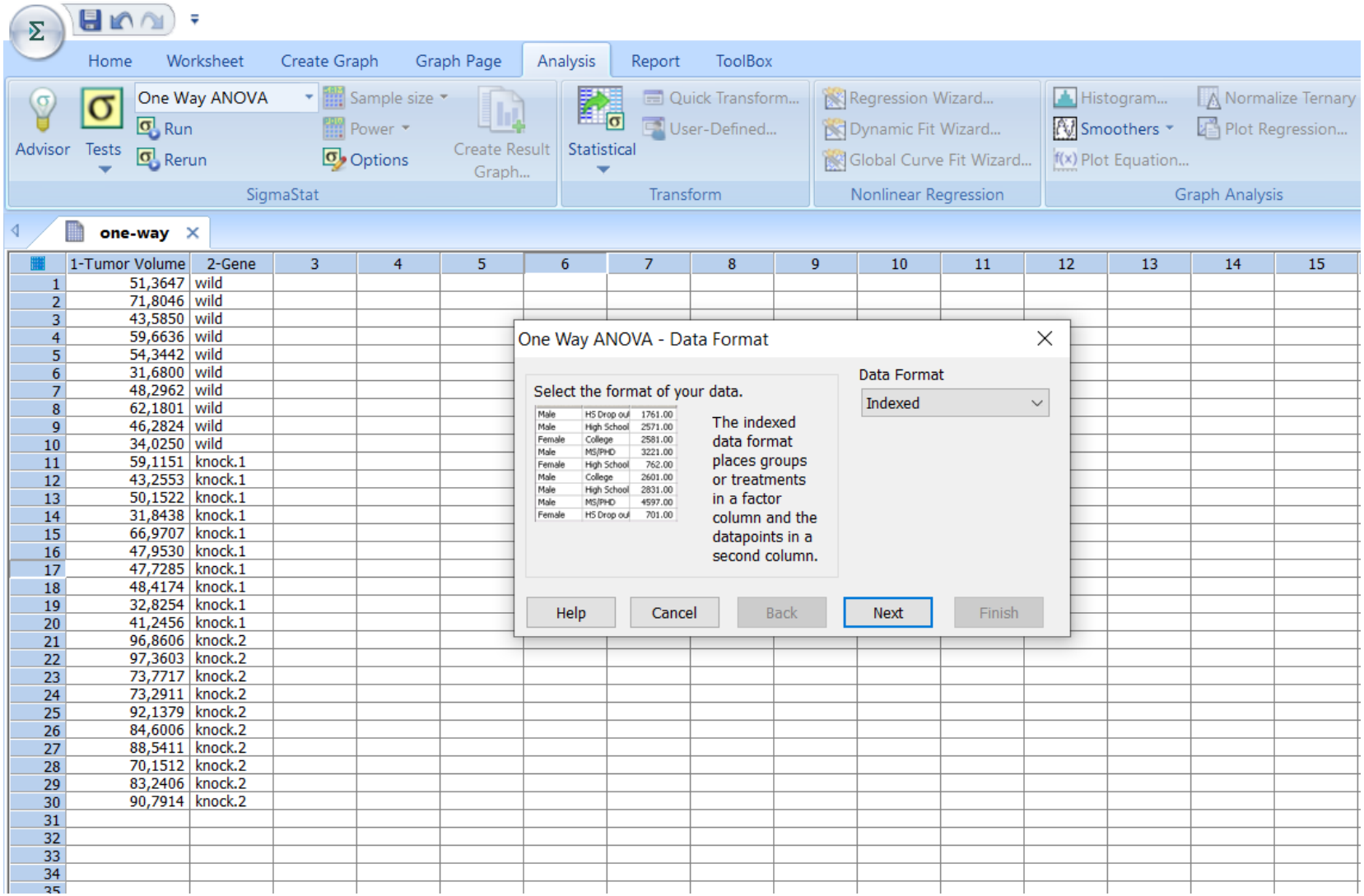
...In practice, your favorite software will give you an output of this type:

Source of variation	Degrees of freedom	Deviance (SS)	Variance (MS)	F
Between	$k-1$	$SSB = \sum_{j=1}^{k} n_j(\bar{Y}_{.j} - \bar{Y}_{..})^2$	$SSB/(k-1)$	$\frac{SSB/(k-1)}{SSR/(n-k)}$
Residual	$n-k$	$SSR = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(Y_{ij} - \bar{Y}_{.j})^2$	$SSR/(n-k)$	
Total	$n-1$	$SST = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(Y_{ij} - \bar{Y}_{..})^2$		

One-way ANOVA with Sigma-plot

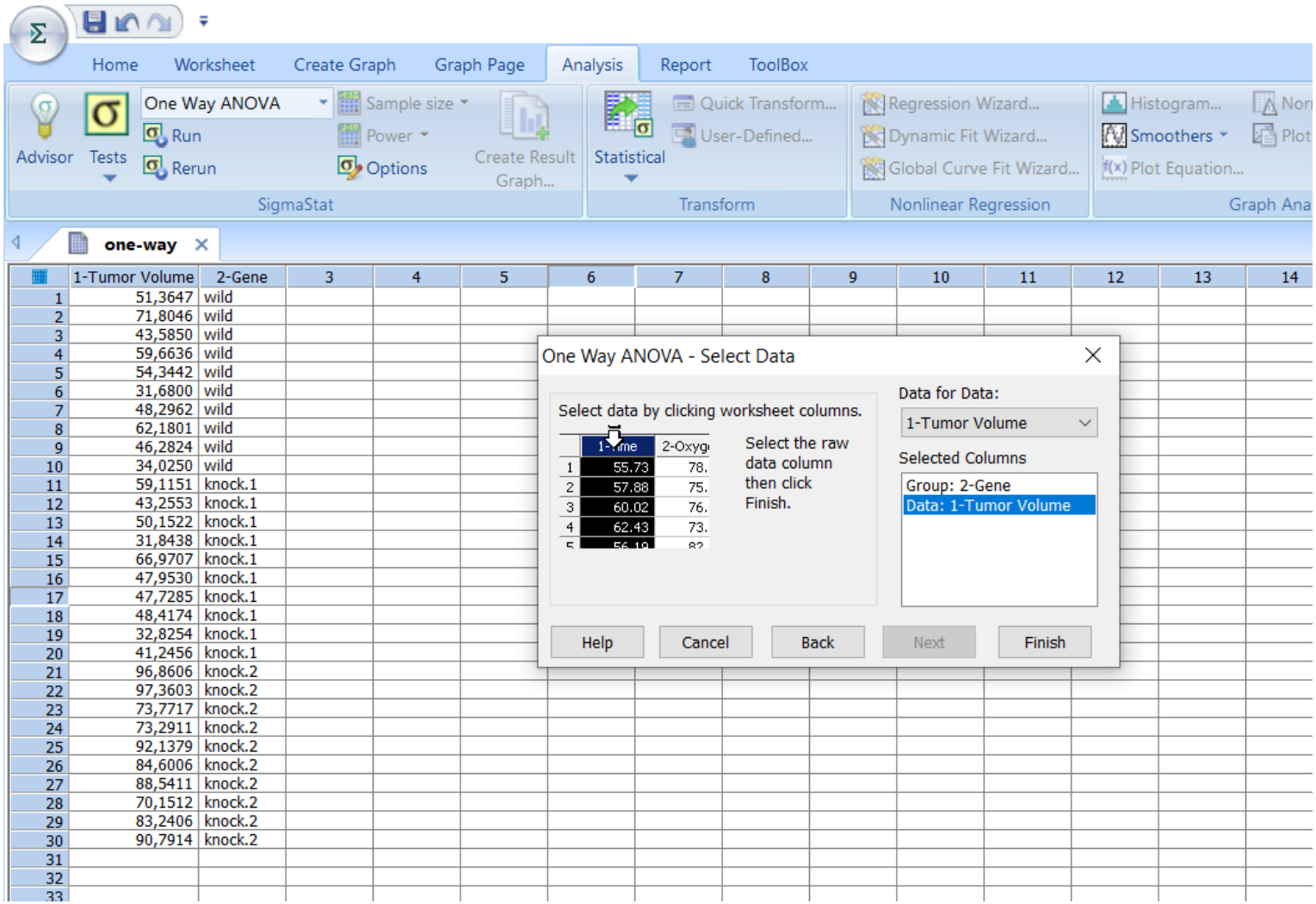


One-way ANOVA with Sigma-plot



	1-Tumor Volume	2-Gene
1	51,3647	wild
2	71,8046	wild
3	43,5850	wild
4	59,6636	wild
5	54,3442	wild
6	31,6800	wild
7	48,2962	wild
8	62,1801	wild
9	46,2824	wild
10	34,0250	wild
11	59,1151	knock.1
12	43,2553	knock.1
13	50,1522	knock.1
14	31,8438	knock.1
15	66,9707	knock.1
16	47,9530	knock.1
17	47,7285	knock.1
18	48,4174	knock.1
19	32,8254	knock.1
20	41,2456	knock.1
21	96,8606	knock.2
22	97,3603	knock.2
23	73,7717	knock.2
24	73,2911	knock.2
25	92,1379	knock.2
26	84,6006	knock.2
27	88,5411	knock.2
28	70,1512	knock.2
29	83,2406	knock.2
30	90,7914	knock.2

One-way ANOVA with Sigma-plot



One-way ANOVA with Sigma-plot

One Way Analysis of Variance

Data source: one-way in one-way

Dependent Variable: Tumor Volume

Normality Test (Shapiro-Wilk): Passed (P = 0,581)

Equal Variance Test (Brown-Forsythe): Passed (P = 0,789)

Group Name	N	Missing	Mean	Std Dev	SEM
wild	10	0	50,323	12,408	3,924
knock.1	10	0	46,951	10,733	3,394
knock.2	10	0	85,075	9,873	3,122

Source of Variation	DF	SS	MS	F	P
Between Groups	2	8908,375	4454,188	36,447	<0,001
Residual	27	3299,642	122,209		
Total	29	12208,017			

The differences in the mean values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P = <0,001).

One-way ANOVA: poll

Two-way ANOVA: example (1)

- Suppose now that the data are grouped according to the levels of **two factors**, A and B, e.g. gene and treatment.
- Data can be summarized in a two-way table.

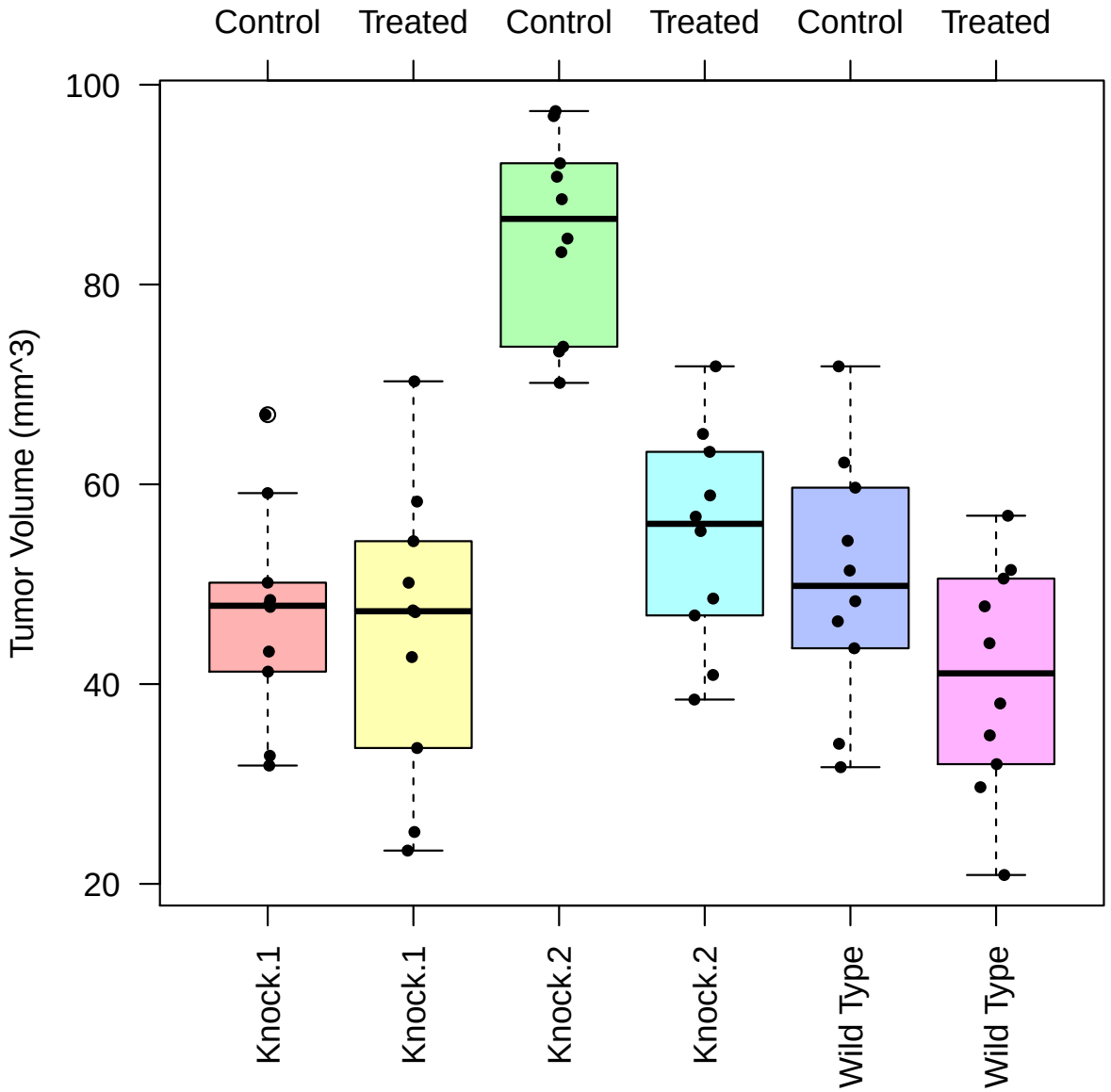
		Treatment		
		Control	Treated	Total
Gene	Wild Type	10	10	20
	Knock 1	10	10	20
	Knock 2	10	10	20
	Total	30	30	60

Aim: we wish to know if the gene and treatment have an effect on tumor volume, and in particular if:

1. There is a difference in the mean of tumor volume according to the gene;
2. There is a difference in the mean of tumor volume according to the treatment;
3. There is an interaction effect between gene and treatment on tumor volume;
4. Which specific comparisons are significant?

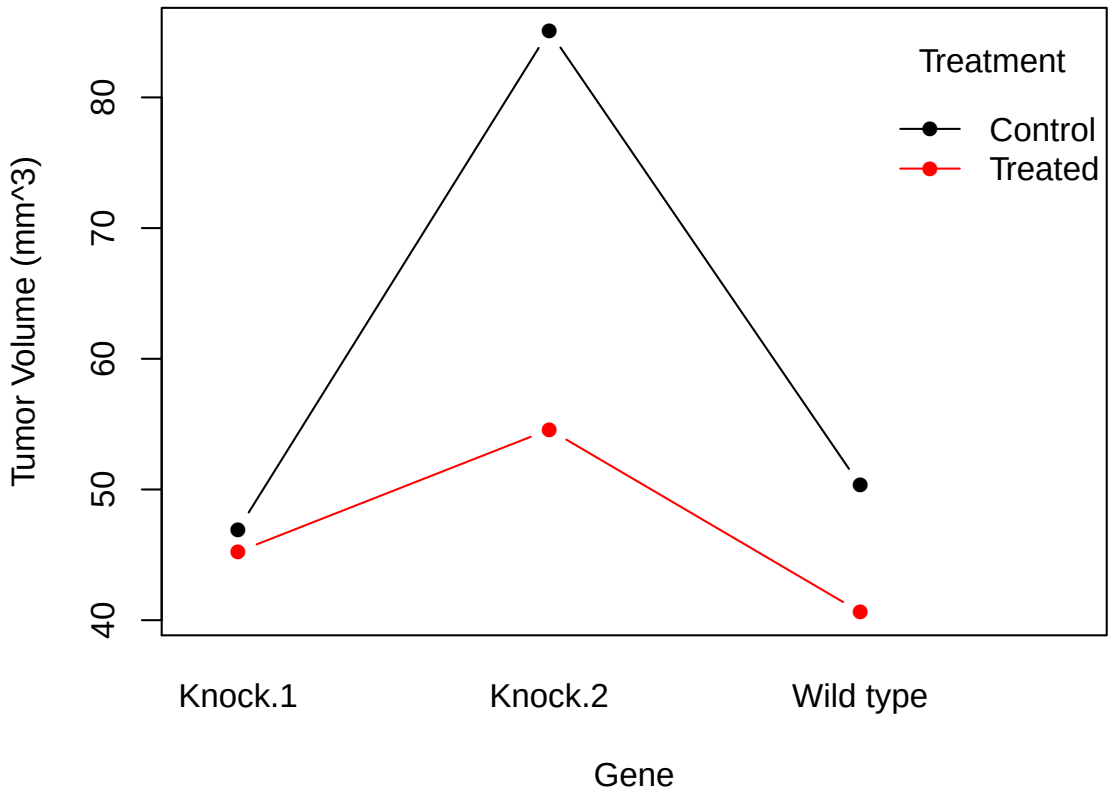
Approach: Two-way ANOVA answers questions (1)-(3). Post-hoc analyses answer question (4).

Two-way ANOVA: example (2)



Interaction effect

The best way to visualize an interaction effect is by plotting the means for each level combination of the two factors.



If an interaction effect is present lines are **not parallel**.

Two-way ANOVA: the formal set-up

Let Y_{ijz} be observation i having level j ($j = 1, ..., a$) of factor A and level z ($z = 1, ..., b$) of factor B (e.g. tumor volume for gene j={Knock 1, Knock 2, Wild Type} and treatment z={control, treated}); let each jz combination have r observations, thus $n = abr$.

The variance decomposition is, in this case,

$$\underbrace{\sum_{j=1}^{a}\sum_{z=1}^{b}\sum_{i=1}^{r}(Y_{ijz} - \bar{Y}_{...})^2}_{Total\ Deviance\ (SST)} = \underbrace{\sum_{j=1}^{a}\sum_{z=1}^{b}\sum_{i=1}^{r}(Y_{ijz} - \bar{Y}_{.jz})^2}_{Residual\ Deviance\ (SSR)} + \underbrace{br\sum_{j=1}^{a}(\bar{Y}_{.j.} - \bar{Y}_{...})^2}_{Factor\ A\ Deviance\ (SSA)}$$

$$+ \underbrace{ar\sum_{z=1}^{b}(\bar{Y}_{..z} - \bar{Y}_{...})^2}_{Factor\ B\ Deviance\ (SSB)} + \underbrace{r\sum_{j=1}^{a}\sum_{z=1}^{b}(\bar{Y}_{.jz} - \bar{Y}_{.j.} - \bar{Y}_{..z} + \bar{Y}_{...})^2}_{Interaction\ Deviance\ (SSI)}.$$

Two-way ANOVA: test

Assumptions

- $Y_{ijz} \sim N(\mu_{jz}, \sigma^2)$, where $\mu_{jz} = \eta + \gamma_j + \beta_z + (\gamma\beta)_{jz}$, and are independent.
- *Equivalently:* $Y_{ijz} = \eta + \gamma_j + \beta_z + (\gamma\beta)_{jz} + \epsilon_{ijz}$, *where* $\epsilon_{ijz} \sim N(0, \sigma^2)$ *and are independent.*

Set of hypotheses

We wish to test at level α:

$$H_0 : \gamma_1 = \gamma_2 = \cdots = \gamma_a \text{ (=0)} \quad \textbf{vs} \quad H_1 : \text{at least one } \gamma_j \text{ is different (from 0)}$$
$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_b \text{ (=0)} \quad H_1 : \text{at least one } \beta_z \text{ is different (from 0)}$$
$$H_0 : \gamma\beta_{11} = \gamma\beta_{12} = \gamma\beta_{21} = \cdots = \gamma\beta_{ab} \text{ (=0)} \quad H_1 : \text{at least one } \gamma\beta_{jz} \text{ is different (from 0)}$$

Test: *F*-test

$$T_A = \frac{SSA/(a-1)}{SSR/(n-ab)} \overset{H_0, assumpt.}{\sim} F_{a-1,n-ab}$$

$$T_B = \frac{SSB/(b-1)}{SSR/(n-ab)} \overset{H_0, assumpt.}{\sim} F_{b-1,n-ab}$$

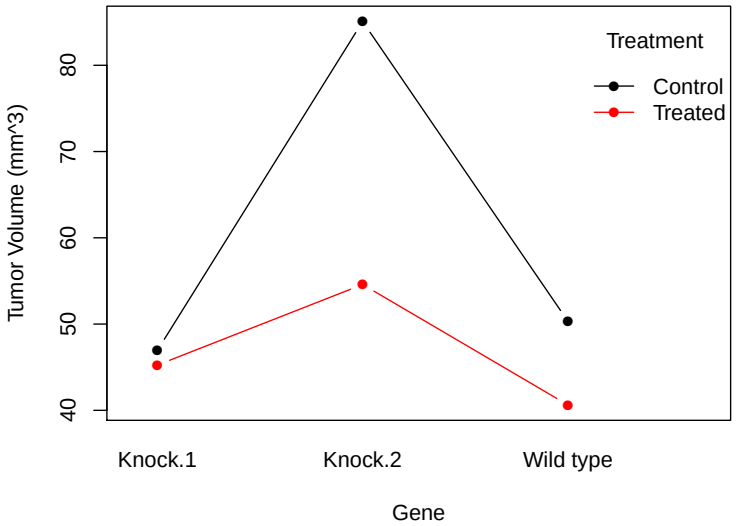
$$T_{AB} = \frac{SSI/[(a-1)(b-1)]}{SSR/(n-ab)} \overset{H_0, assumpt.}{\sim} F_{(a-1)(b-1),n-ab}.$$

Two-way ANOVA: output

Source of variation	Degrees of freedom	Deviance (SS)	Variance (MS)	F
Factor A	$a-1$	SSA	$SSA/(a-1)$	$\frac{SSA/(a-1)}{SSR/(n-ab)}$
Factor B	$b-1$	SSB	$SSB/(b-1)$	$\frac{SSB/(b-1)}{SSR/(n-ab)}$
Interaction AB	$(a-1)(b-1)$	SSI	$SSI/[(a-1)(b-1)]$	$\frac{SSI/[(a-1)(b-1)]}{SSR/(n-ab)}$
Residual	$n-ab$	SSR	$SSR/(n-ab)$	
Total	$n-1$	SST		

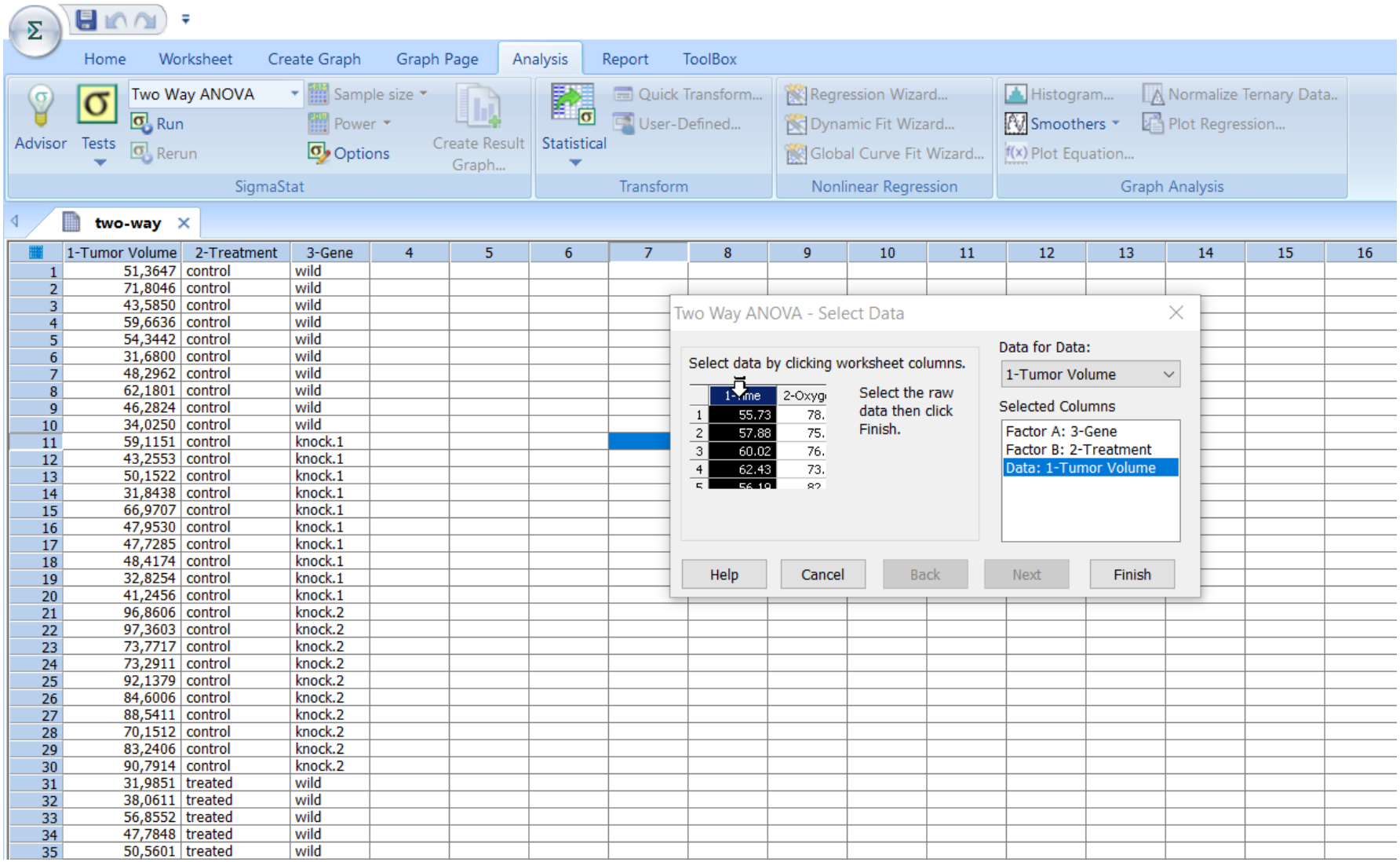
Interaction effect

- If an interaction effect is significant, care should be taken in interpreting the main effects.
- It can happen that an interaction effect is significant, but not the main effects: this means that some differences among the cell means are present, but they disappear when we look at the levels of one factor after averaging over the levels of the other factor.



  - If, on the other hand, the interaction not significant, you can remove it from the model (it is always a good idea to try to keep the model as simple as possible!).

Two-way ANOVA with Sigma-plot



	1-Tumor Volume	2-Treatment	3-Gene
1	51,3647	control	wild
2	71,8046	control	wild
3	43,5850	control	wild
4	59,6636	control	wild
5	54,3442	control	wild
6	31,6800	control	wild
7	48,2962	control	wild
8	62,1801	control	wild
9	46,2824	control	wild
10	34,0250	control	wild
11	59,1151	control	knock.1
12	43,2553	control	knock.1
13	50,1522	control	knock.1
14	31,8438	control	knock.1
15	66,9707	control	knock.1
16	47,9530	control	knock.1
17	47,7285	control	knock.1
18	48,4174	control	knock.1
19	32,8254	control	knock.1
20	41,2456	control	knock.1
21	96,8606	control	knock.2
22	97,3603	control	knock.2
23	73,7717	control	knock.2
24	73,2911	control	knock.2
25	92,1379	control	knock.2
26	84,6006	control	knock.2
27	88,5411	control	knock.2
28	70,1512	control	knock.2
29	83,2406	control	knock.2
30	90,7914	control	knock.2
31	31,9851	treated	wild
32	38,0611	treated	wild
33	56,8552	treated	wild
34	47,7848	treated	wild
35	50,5601	treated	wild

Two-way ANOVA with Sigma-plot

Two Way Analysis of Variance

Data source: two-way in twoway

Balanced Design

Dependent Variable: Tumor Volume

Normality Test (Shapiro-Wilk): Passed (P = 0,590)

Equal Variance Test (Brown-Forsythe): Passed (P = 0,915)

Source of Variation	DF	SS	MS	F	P
Gene	2	7713,622	3856,811	27,956	<0,001
Treatment	1	2925,687	2925,687	21,207	<0,001
Gene x Treatment	2	2207,243	1103,621	8,000	<0,001
Residual	54	7449,889	137,961		
Total	59	20296,441	344,007		

Main effects cannot be properly interpreted if significant interaction is determined. This is because the size of a factor's effect depends upon the level of the other factor.

The effect of different levels of Gene depends on what level of Treatment is present. There is a statistically significant interaction between Gene and Treatment. (P = <0,001)

Post-hoc tests : idea

- Suppose that you have run your one-way or two-way ANOVA.
- You additionally wish to test e.g. individual group differences.
- Post-hoc tests are designed for this purpose.

Post-hoc tests : idea

- Suppose that you have run your one-way or two-way ANOVA.
- You additionally wish to test e.g. individual group differences.
- Post-hoc tests are designed for this purpose.
- Essentially, all you need to do is to perform a set of pair-wise *t*-tests...

Post-hoc tests : idea

- Suppose that you have run your one-way or two-way ANOVA.
- You additionally wish to test e.g. individual group differences.
- Post-hoc tests are designed for this purpose.
- Essentially, all you need to do is to perform a set of pair-wise *t*-tests...
- ...But you run into a situation of **multiple testing** (this is the topic of Lecture 11, 16th December).

Post-hoc tests : idea

- Suppose that you have run your one-way or two-way ANOVA.
- You additionally wish to test e.g. individual group differences.
- Post-hoc tests are designed for this purpose.
- Essentially, all you need to do is to perform a set of pair-wise *t*-tests...
- ...But you run into a situation of **multiple testing** (this is the topic of Lecture 11, 16th December).
- The main problem in a multiple testing framework is that, unless your tests are perfectly dependent (which means that you essentially perform one test), your overall α (here we focus on the family-wise error rate, the probability of rejecting the null hypothesis, when true, in at least one comparison) level is higher than that of a single test.

Post-hoc tests : options

There are several solutions to this problem (some more and some less specific to ANOVA):

- Adjustment of the *p*-values (these are generally applicable to any test):

Post-hoc tests : options

There are several solutions to this problem (some more and some less specific to ANOVA):

- Adjustment of the *p*-values (these are generally applicable to any test):
  - **Bonferroni**, **Šidák**, **Holm-Bonferroni**, or **Holm-Šidák**: Šidák tends to have more power (but assumes independence between the tests), and the Holm variant is more powerful in both cases.
  - ...
- More specific to the *t*-test/ANOVA situation:

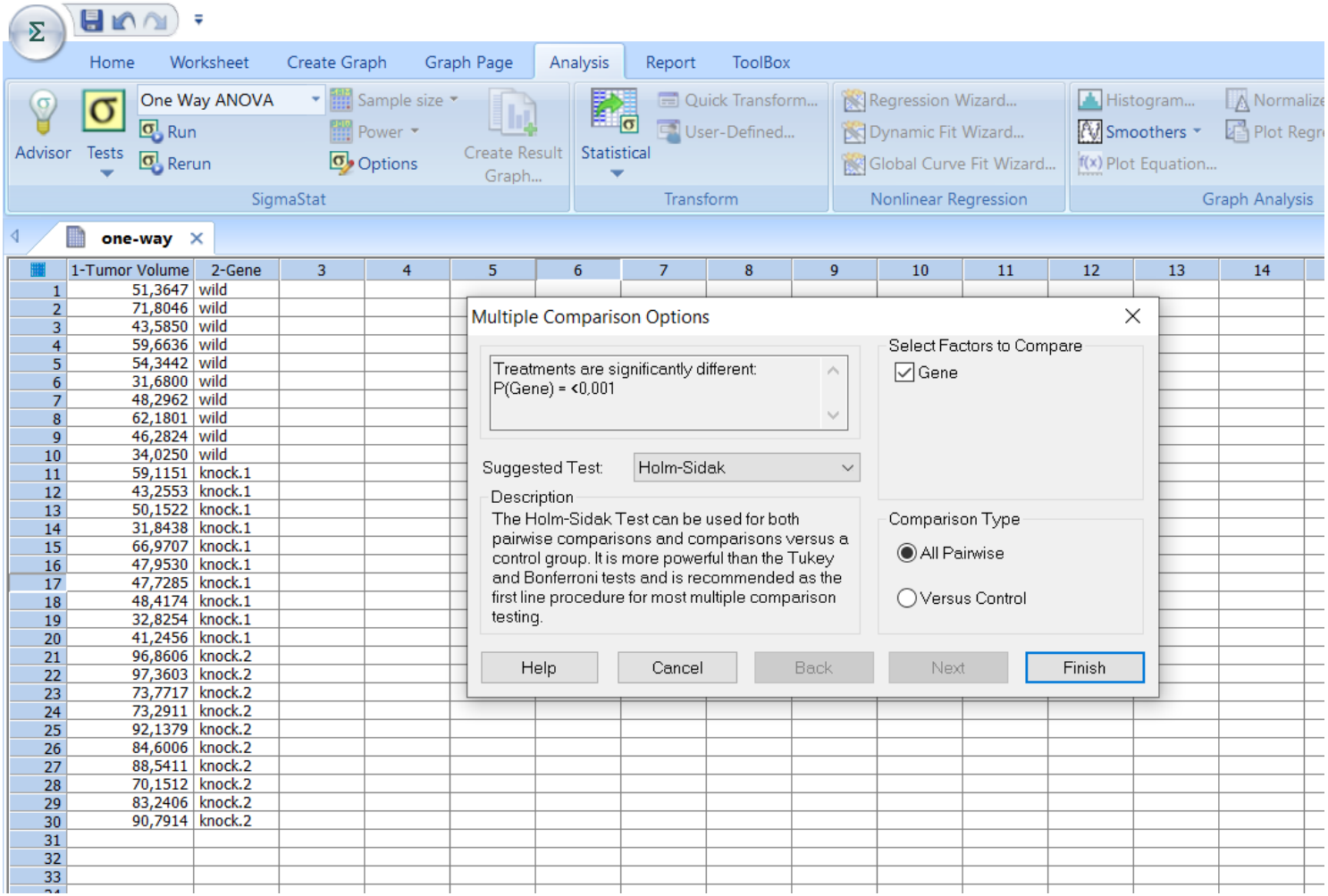
Post-hoc tests : options

There are several solutions to this problem (some more and some less specific to ANOVA):

- Adjustment of the *p*-values (these are generally applicable to any test):
  - **Bonferroni**, **Šidák**, **Holm-Bonferroni**, or **Holm-Šidák**: Šidák tends to have more power (but assumes independence between the tests), and the Holm variant is more powerful in both cases.
  - ...
- More specific to the *t*-test/ANOVA situation:
  - **Tukey** (requires a correction for unbalanced designs), **Student-Newman-Keuls**, **Duncan** (decreasingly conservative, increasingly powerful: no 'free lunch'): adjust the quantiles of the test statistics & pool the variances.
  - **Fisher LSD**: pools the variances, but does not perform a multiple testing correction.
  - **Dunnett**: appropriate to compare each group to a control.
  - ...

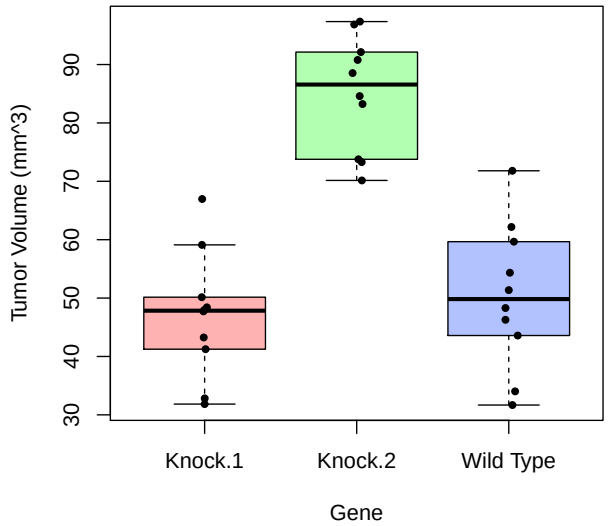
Note: Post-hoc tests, with the exception of Fisher LSD, can be used independently of the ANOVA result.

Post-hoc tests with sigma-plot: one-way ANOVA



	1-Tumor Volume	2-Gene
1	51,3647	wild
2	71,8046	wild
3	43,5850	wild
4	59,6636	wild
5	54,3442	wild
6	31,6800	wild
7	48,2962	wild
8	62,1801	wild
9	46,2824	wild
10	34,0250	wild
11	59,1151	knock.1
12	43,2553	knock.1
13	50,1522	knock.1
14	31,8438	knock.1
15	66,9707	knock.1
16	47,9530	knock.1
17	47,7285	knock.1
18	48,4174	knock.1
19	32,8254	knock.1
20	41,2456	knock.1
21	96,8606	knock.2
22	97,3603	knock.2
23	73,7717	knock.2
24	73,2911	knock.2
25	92,1379	knock.2
26	84,6006	knock.2
27	88,5411	knock.2
28	70,1512	knock.2
29	83,2406	knock.2
30	90,7914	knock.2

Post-hoc tests with sigma-plot: one-way ANOVA

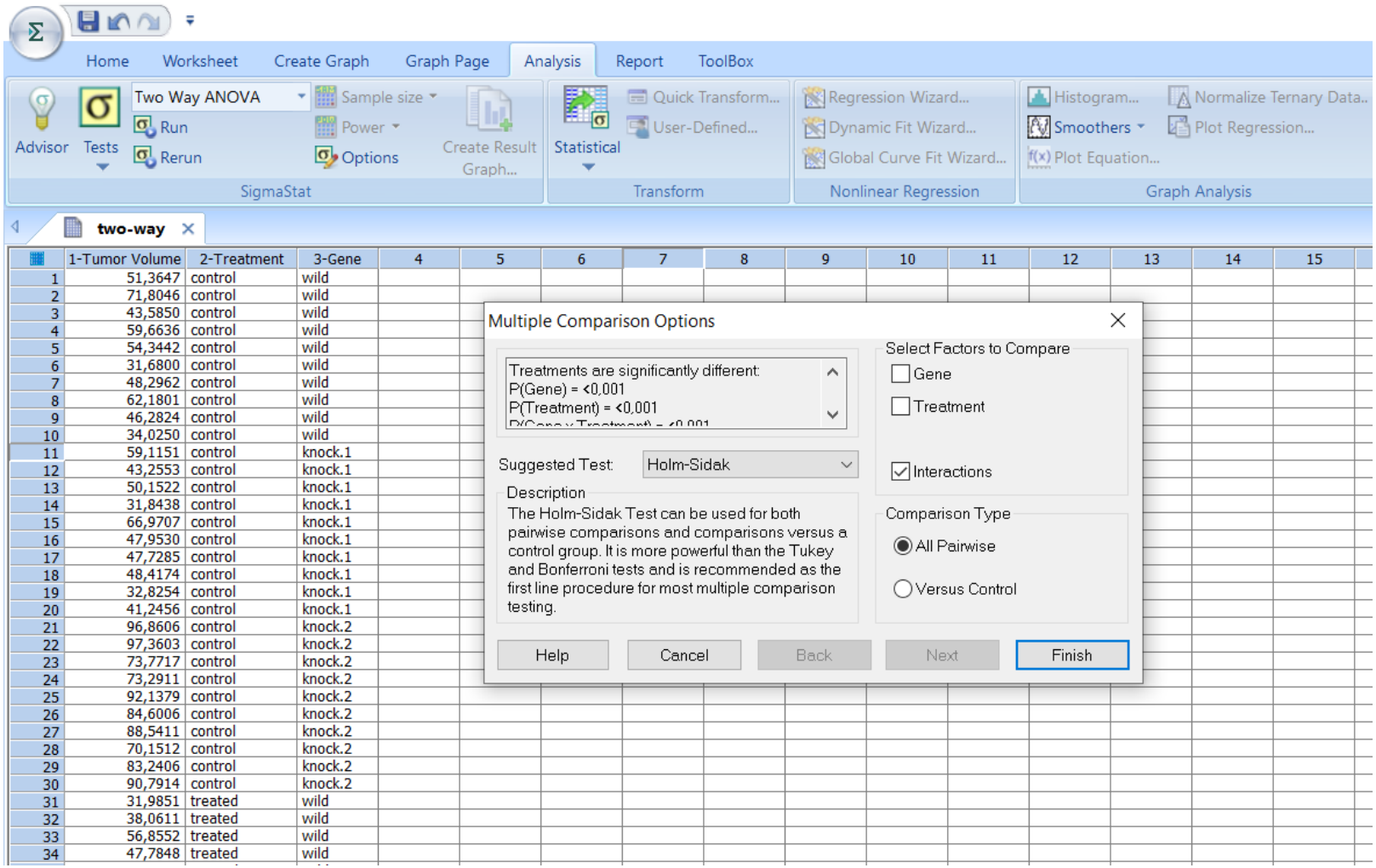


All Pairwise Multiple Comparison Procedures (Holm-Sidak method):
Overall significance level = 0,05

Comparisons for factor: **Gene**

Comparison	**Diff of Means**	**t**	**P**	**P<0,050**
knock.2 vs. knock.1	38,124	7,711	<0,001	Yes
knock.2 vs. wild	34,752	7,029	<0,001	Yes
wild vs. knock.1	3,372	0,682	0,501	No

Post-hoc tests with sigma-plot: two-way ANOVA



	1-Tumor Volume	2-Treatment	3-Gene
1	51,3647	control	wild
2	71,8046	control	wild
3	43,5850	control	wild
4	59,6636	control	wild
5	54,3442	control	wild
6	31,6800	control	wild
7	48,2962	control	wild
8	62,1801	control	wild
9	46,2824	control	wild
10	34,0250	control	wild
11	59,1151	control	knock.1
12	43,2553	control	knock.1
13	50,1522	control	knock.1
14	31,8438	control	knock.1
15	66,9707	control	knock.1
16	47,9530	control	knock.1
17	47,7285	control	knock.1
18	48,4174	control	knock.1
19	32,8254	control	knock.1
20	41,2456	control	knock.1
21	96,8606	control	knock.2
22	97,3603	control	knock.2
23	73,7717	control	knock.2
24	73,2911	control	knock.2
25	92,1379	control	knock.2
26	84,6006	control	knock.2
27	88,5411	control	knock.2
28	70,1512	control	knock.2
29	83,2406	control	knock.2
30	90,7914	control	knock.2
31	31,9851	treated	wild
32	38,0611	treated	wild
33	56,8552	treated	wild
34	47,7848	treated	wild

Post-hoc tests with sigma-plot: two-way ANOVA

All Pairwise Multiple Comparison Procedures (Holm-Sidak method):

Overall significance level = 0,05

Comparisons for factor: **Treatment within wild**

Comparison	Diff of Means	t	P	P<0,050
control vs. treated	9,703	1,847	0,070	No

Comparisons for factor: **Treatment within knock.1**

Comparison	Diff of Means	t	P	P<0,050
control vs. treated	1,706	0,325	0,747	No

Comparisons for factor: **Treatment within knock.2**

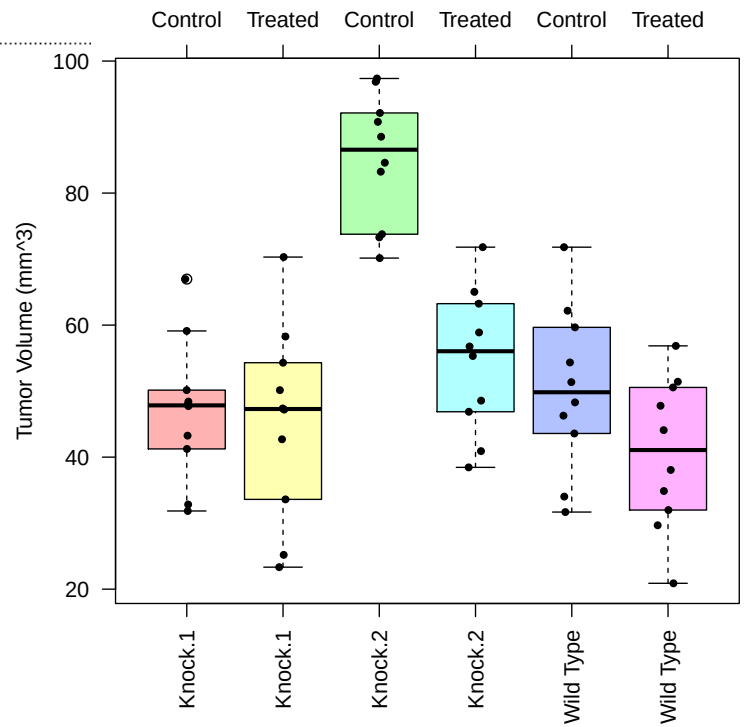
Comparison	Diff of Means	t	P	P<0,050
control vs. treated	30,488	5,804	<0,001	Yes

Comparisons for factor: **Gene within control**

Comparison	Diff of Means	t	P	P<0,050
knock.2 vs. knock.1	38,124	7,258	<0,001	Yes
knock.2 vs. wild	34,752	6,616	<0,001	Yes
wild vs. knock.1	3,372	0,642	0,524	No

Comparisons for factor: **Gene within treated**

Comparison	Diff of Means	t	P	P<0,050
knock.2 vs. wild	13,967	2,659	0,031	Yes
knock.2 vs. knock.1	9,342	1,779	0,155	No
knock.1 vs. wild	4,625	0,880	0,383	No



Assumptions of ANOVA

ANOVA assumes that errors $\epsilon_{ij(z)}$:

Assumptions of ANOVA

ANOVA assumes that errors $\epsilon_{ij(z)}$:

1. are normal;
2. have constant variance (homoscedasticity);
3. are independent.

Assumptions should be based on theoretical grounds or checked:

Assumptions of ANOVA

ANOVA assumes that errors $\epsilon_{ij(z)}$:

1. are normal;
2. have constant variance (homoscedasticity);
3. are independent.

Assumptions should be based on theoretical grounds or checked:

1. Normality: normality tests, e.g. Shapiro-Wilk.

Assumptions of ANOVA

ANOVA assumes that errors $\epsilon_{ij(z)}$:

1. are normal;
2. have constant variance (homoscedasticity);
3. are independent.

Assumptions should be based on theoretical grounds or checked:

1. Normality: normality tests, e.g. Shapiro-Wilk.
2. Variance homogeneity across the groups (homoscedasticity): Levene-Test or (better) Brown-Forsythe-Test.

Assumptions of ANOVA

ANOVA assumes that errors $\epsilon_{ij(z)}$:

1. are normal;
2. have constant variance (homoscedasticity);
3. are independent.

Assumptions should be based on theoretical grounds or checked:

1. Normality: normality tests, e.g. Shapiro-Wilk.
2. Variance homogeneity across the groups (homoscedasticity): Levene-Test or (better) Brown-Forsythe-Test.
3. Independence of the errors: experimental design (spatial, temporal or repeated measures effect).

Assumptions of ANOVA: effect of violations

If assumptions are violated, inferences become inaccurate. If the main focus is testing, in particular:

- Type I error and power of the test may be affected, i.e. the test may not have the nominal probability of type I error α, and power can be increased or decreased.
- Consequences are generally more serious when the assumption of independence is violated, and when the design is not balanced (different number of observations in each cell).

Possible solutions:

- Difference in variance across groups and/or non-normality: **transform the data**.
- Non-normality: resort to **nonparametric tests** (Lecture 5, 4th November).
- Dependence: addressed by the experimental design or via modelling.

Outliers: poll

Outliers: poll

- Check if they are genuinely ‘bad’ measurements.
- Report the analysis with and without.
- Try a nonparametric approach (may have low power)/ transformations.
- Check if the model is appropriate.

Additional applications of ANOVA

- ANOVA can include more than two factors (the output and computations become longer, but the rationale is the same).
- ANOVA can also be extended to situations in which multiple measurements from the same individuals/units (ideally sampled from a larger population) are available: this requires a random effects model. In the simplest scenario, it has the form

$$\begin{aligned} Y_{i,j} &\sim N(a_j, \sigma^2_\epsilon) \\ a_j &\sim N(\mu, \sigma^2_a), \end{aligned}$$

  where i denotes the observation and j the **individual or unit** which is resampled.
- More commonly, you would need a mixture of random and fixed effects: mixed effects model.
- Inference for mixed effects models can also be handled the context of linear mixed models (Lecture 12, 13th January), with some advantages.

Conclusions & Take-home message

- ANOVA is a **parametric** tool to compare the means of different groups.
- When there are only two groups, it is equivalent to a *t*-test.
- The results are based on a number of distributional assumptions (normality, homoscedasticity, independence), which should be checked or accepted on theoretical grounds.
- Always keep in mind your **experimental design** when analysing your data.
- A well designed experiment can avoid some problems, and give you some power (e.g. balanced designs are preferable).
- If the design becomes very complex, it can help to consult a 'friendly statistician' either after, or (better) before, the experiment.

Bibliography

- Scheffé, H. (1999). The analysis of variance (Vol. 72). John Wiley Sons.
- Oehlert, G. (2010). A First Course in Design and Analysis of Experiments. Available at: http://users.stat.umn.edu/ gary/book/fcdae.pdf.

...Up next

14th October: Multiple linear regression
(Dr. Diana Tichy)