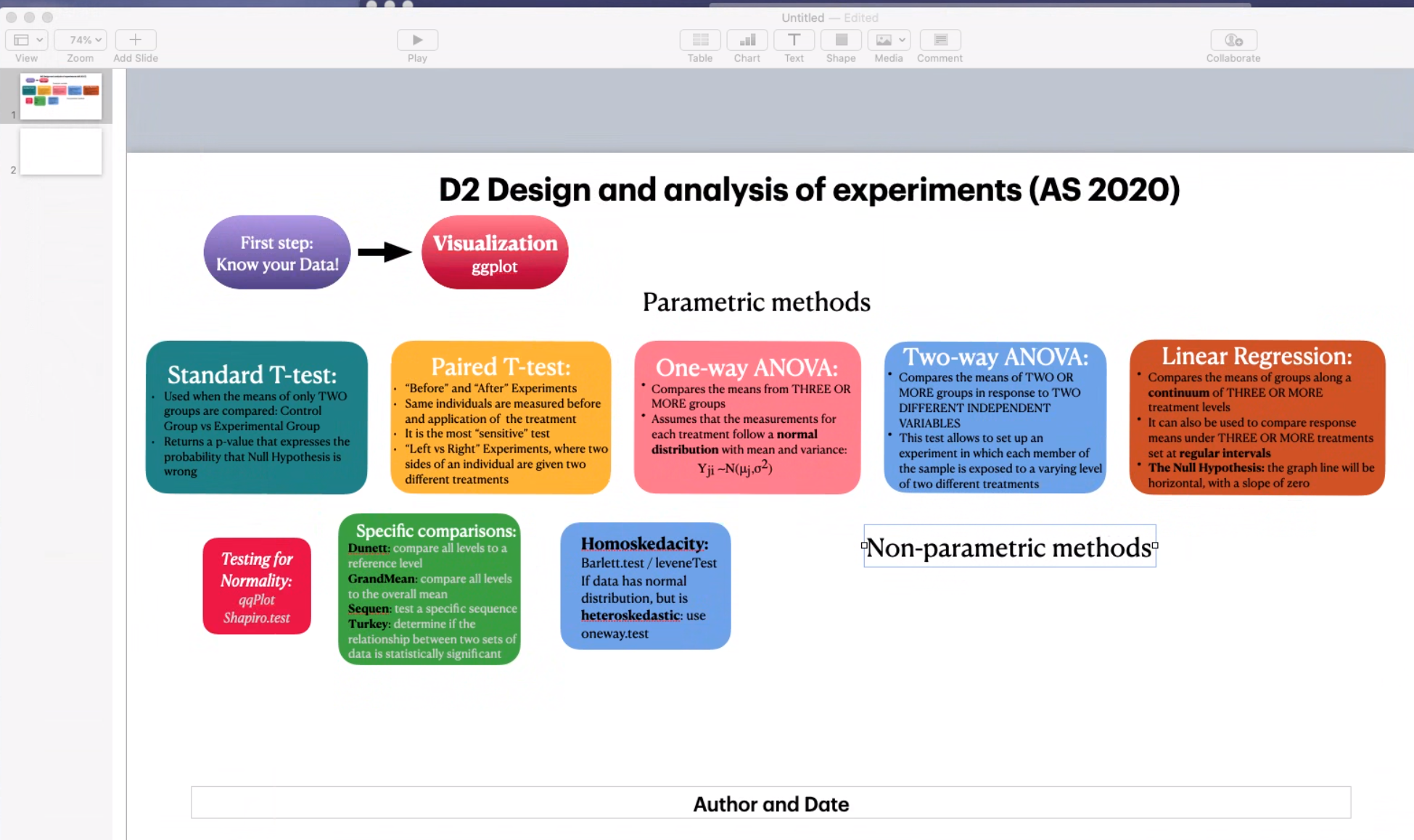
Statistische Versuchsplanung – Design of Experiments (DoE)

Idee: Optimierung von Prozessen und Produkten.

# 1. Script experimental design - einführung

Anschließend wird die Bedeutung der Datenvisualisierung als erster Schritt einer jeden Datenanalyse besprochen. Des Weiteren wird eine erste Analyse mit einem Varianzanalyse (aka ANOVA)-Modell durchgeführt, welches der grundlegende Ansatz ist, um den Unterschied in der mittleren Antwort über mehrere (>2) Gruppen zu analysieren.

Das verwendete ANOVA-Modell trifft einige Annahmen, die in dem als Modelldiagnose bezeichneten Prozess bewertet werden müssen, um zu beurteilen, ob die Annahmen plausibel sind oder ob einige Anpassungen in der Modellierungsstrategie vorgenommen werden müssen.

|  |  |
| --- | --- |
| experimental unit | is the smallest unit to which a treatment can be applied. |
| observational unit | is the smallest unit on which a response will be measured. |
| Dependent variable | The outcome |
| Independent variable | Defining the treatment |

# 2. Completely randomized designs without treatment structure

-independece assumption

-distributional assumption

Assume that we have k treatments and that treatment j is applied to nj plots, j = 1, . . . , k. If treatments are assigned to the plots at random without any restrictions, the design is called a completely randomized design (CRD). In completely randomized designs, observational units were randomly allocated to the treatments, so the only systematic influence on the outcome should be due to the treatment.

*Boxplot.*

Look for

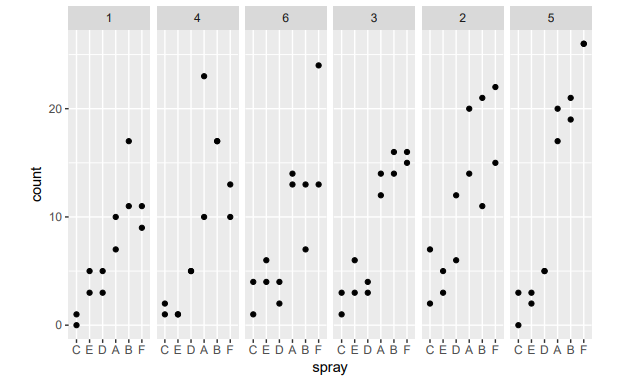
• location differences (compare the location of the bold lines representing the sample medians);

• variability differences (compare the box heights containing the central 50% of the data);

• skewness (look at the whiskers and at the location of the median in the box);

• outliers (look at single points).

Block structure



• Do blocks themselves have an effect?

• Do treatment effects depend on the blocks?

For me, it is not visually clear whether the block effects are important. The second question is difficult to assess with only two observations for each treatment in each block. We will see how to answer these questions when we discuss blocked data.

# 2.2 Parametric models: one-way ANOVA

# 2.2.2 The ANOVA table and the overall F test

We test the null hypothesis that all theoretical treatment means are the same, H0 : µ1 = . . . = µk = µ against the alternative that at least two theoretical treatment means differ.

Under the null hypothesis, µ should be estimated as the mean of all observations, while under the alternative the means µj for each treatment should be estimated by using observations from that treatment

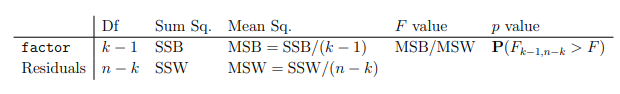


Figure 1: ANOVA table: Df = degree of freedom and it is used to keep track of the number of treatments and the overall sample size. Sums of squares and mean sums of squares. MSB quantifies the variability between treatment means. F value = MSB/MSW, big values of F are evidence against the null hypothesis of equal treatment means. Last column it is the p-value to denote the probability of a random variable with an Fk-1, n-k distribution exceeding the observed value F. the bigger the observed F, the smaller the probability

**Lm** command is to fit a linear model.

**Anova** is used to produce the ANOVA table based on the model.

Top-down approach

# 2.3 Pairwise treatment comparisons

1. estimates a big model – overall F test is performed

2. pairwise treatment comparisons

# 2.3.2 Directly adjusting a set of p values

It features several adjustment methods, the holm method is a good general method. This stepwise procedure is guaranteed to control the familywise error rate, but is less conservative than e. g. the Bonferroni method.

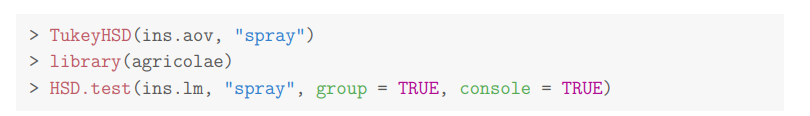
# 2.3.3 Pairwise t and Wilcoxon tests

with(InsectSprays, pairwise.t.test(count, spray, "holm")) -> for adjusted p values

with(InsectSprays, pairwise.wilcox.test(count, spray, "holm"))-> for Holm-corrected p value

If one doubts the normality assumption, it is safer to use the pairwise Wilcoxon tests or another nonparametric or robust procedure.

# 2.3.4 Tukey’s honest significant difference (HSD)



# 3 Completely randomized designs with control

Let us now suppose that one among the treatments is special, we call it the control treatment; an example is the standard treatment in clinical studies.

# 3.1 Comparing all treatments with the control - multiple linear regression model

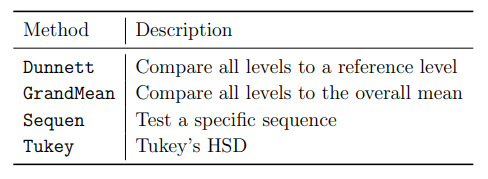
Assume that treatment 1 is the control treatment and the aim is to compare all the other treatments to it (but not among themselves). If we denote by µj the mean of treatment j, then the aim is to estimate µj − µ1 and to test whether µj − µ1 = 0 for j = 2, . . . , k. There are several ways to accomplish this in R. The simplest way is to redefine the control treatment as the reference level of the factor in question, to fit the model with lm and to look at its summary.

library(multcomp)

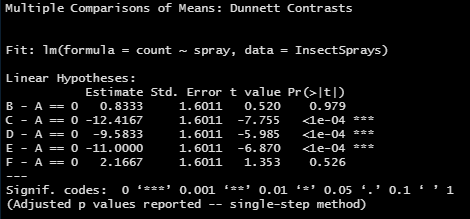
summary(glht(ins.lm, mcp(spray = "Dunnett")))

# 3.3 Specific comparisons

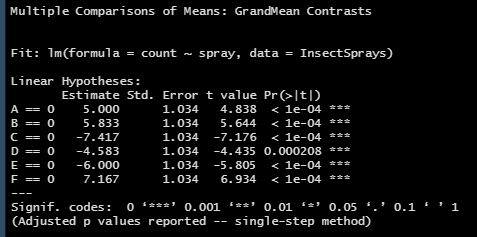
Reducing the number of comparisons is attractive because it makes the p value correction less strict, such that there is a gain in power.



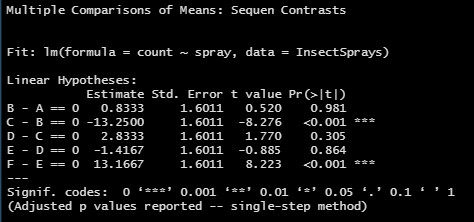
summary(glht(ins.lm, mcp(spray = "Dunnett")))



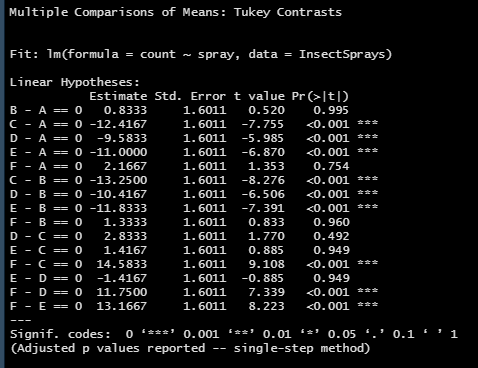
summary(glht(ins.lm, mcp(spray = "GrandMean")))



summary(glht(ins.lm, mcp(spray = "Sequen")))



summary(glht(ins.lm, mcp(spray = "Tukey")))



# 4 Model diagnostics and alternative analysis methods / Modelldiagnose und alternative Analysemethoden

# 4.1 The normality assumption

# 4.1.1 Graphical normality checking

If the normality assumption holds, the points in the normal QQ plot should lie “closely” around the straight line given in the normal QQ plot. I recommend not to use normal QQ plots for less than around 30 observations.

> library(car)

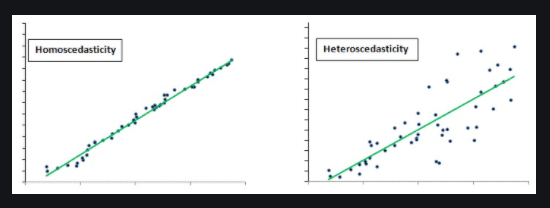
> qqPlot(resid(ins.lm))

# 4.1.2 Testing for normality – saphiro-wilk

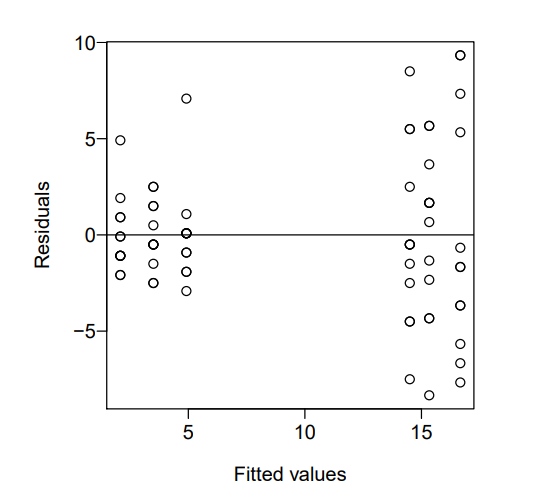
We use the Shapiro-Wilk test and reject the null hypothesis of normality if p < α. Performing the test

Because p < 0.05, we reject the null hypothesis of normality and conclude that the residuals do not come from a normal distribution.

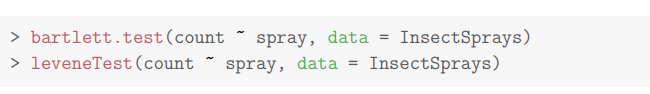
# 4.2 The homoskedasticity assumption



ANOVA assumptions include that the true variance of the dependent variable is the same in each treatment (i. e. that the data are homoskedastic). A visual assessment is possible with residual plots. Here, plotting the residuals vs. the fitted values is interesting.



# 4.2.2 Testing homoskedasticity (equal variances) – Bartlett and Levene test



# 4.2.3 Accounting for heteroskedasticity

Bei der klassischen ANOVA wird angenommen, dass die Varianz für jede Behandlung gleich groß ist. Diese ist oft eine fragwürdige Annahme für reale Daten. Auch die Annahme der Normalität ist für reale Daten oft sehr fragwürdig.

# 4.4 Nonparametric methods

The main use of nonparametric analyses of the one-way layout is that no normality assumptions are necessary.

# 4.4.2 The Kruskal-Wallis test

# 5. Blocking

# 5.2 Complete block designs – randomized complete block design

# 6 Fitting and interpreting mixed effects models

Library(nlme)  
library(lme4)