1 Introduction

1.1 The Paradigm of Parametric Statistics

a Probability $\longrightarrow$ Model

Typically, this is a parametric family, like the normal $\mathcal{N}\langle \mu, \sigma^2 \rangle$.

Statistics: Bridge between model and data The 3 basic questions of inferential statistics:

[1.] Which parameter value(s) is (are) most plausible?
$\longrightarrow$ estimator

[2.] Is a given value plausible?
$\longrightarrow$ Test.

[3.] Which values are plausible?
$\longrightarrow$ confidence interval

b **Probability Model** is needed to describe
what "could have happened, too"
and with which "chances" (or odds).
Prob. model should be clear before the data are obtained.
and describes our ideas about the possible results
and their "plausibility".

Model is needed for determining the
statistical uncertainty of an estimate or a test statistic.

c Check assumptions!

d We would dream of methods for which
no assumptions were needed!

e Rank tests assume (almost) nothing!

f The basic idea of resampling:
Use the data to estimate their distribution!

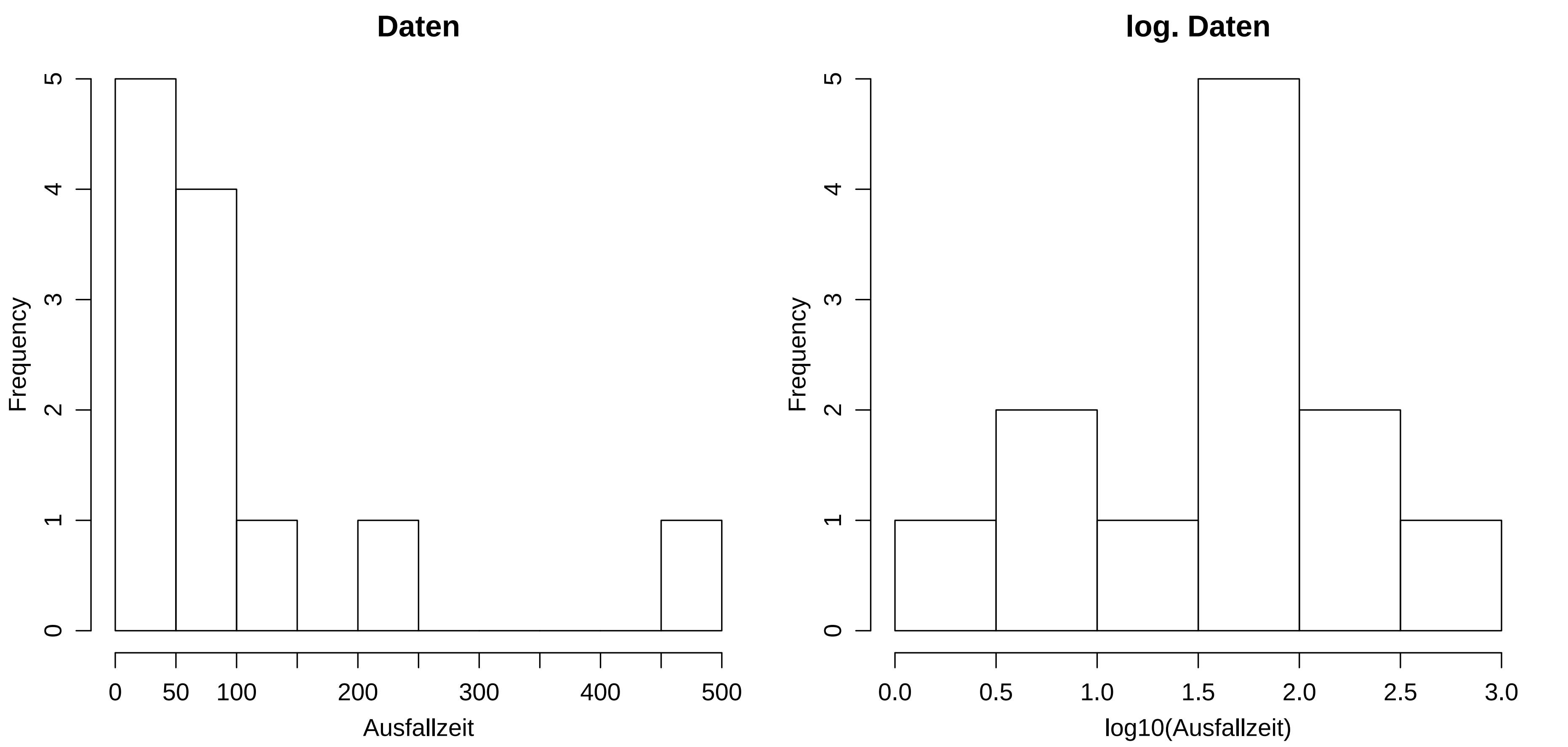
g Complicated estimator $\longrightarrow$ determine precision by

- Bootstrap, ...
- Asymptotic distribution

1.2 Example

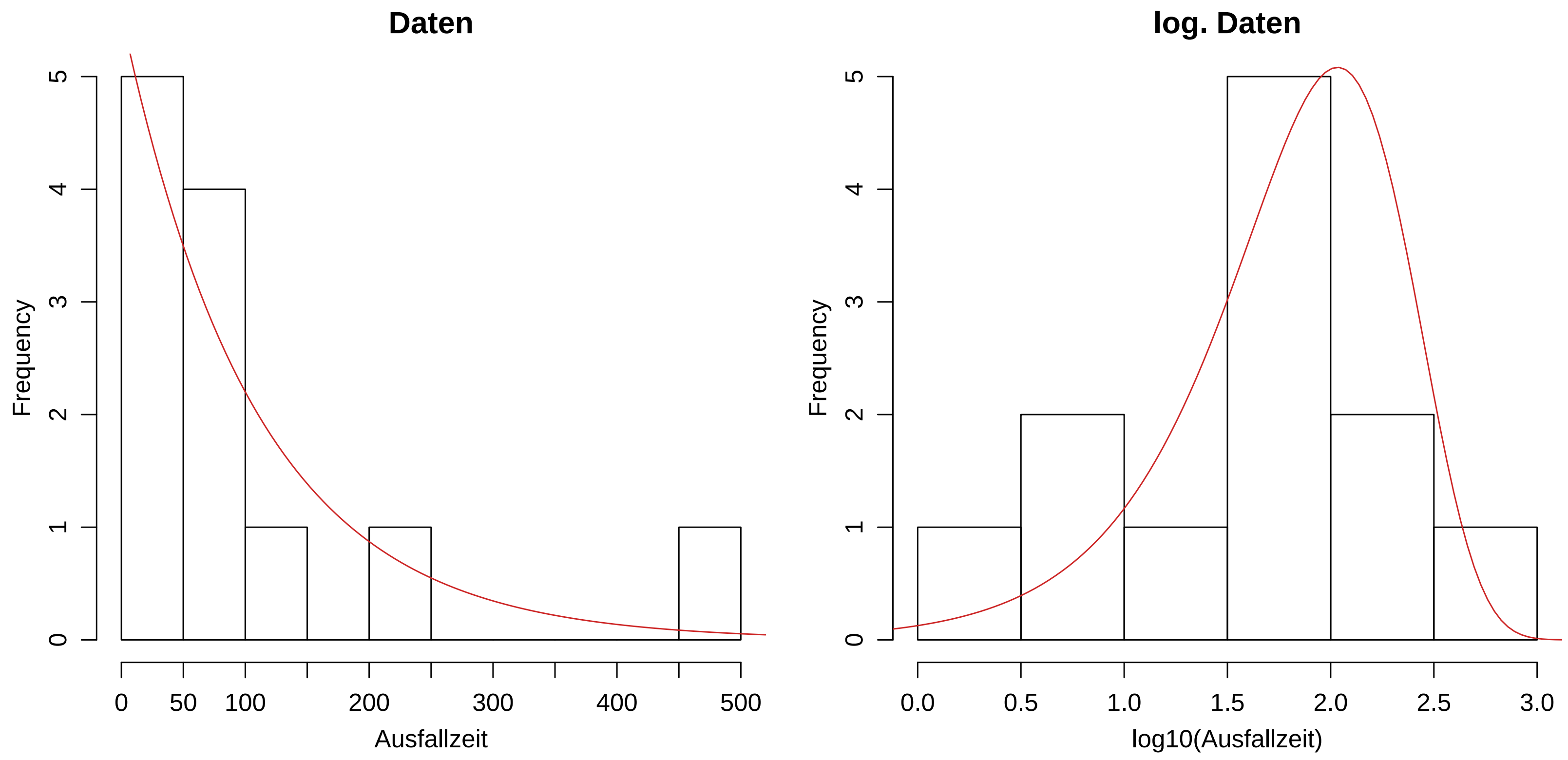
a Failure times of an air conditioning system (Boeing 720)
$n = 12$ observed time intervals between failures (sorted):

3 5 7 18 43 85 91 98 100 130 230 487



b Most simple param. model for failure times: Exponential distribution $\mathcal{E}xp$ with density

$$f\langle y\rangle = \frac{1}{\mu} e^{-y/\mu} \qquad y > 0$$



c Model fits quite reasonably. Small data set

Better: Do not assume any particular distribution!

Question: Mean? $\overline{X} = 108.1$.

More instructive: 20% trimmed mean?

(= drop 20% smallest and 20% largest obs.

calculate mean of the remaining ones.)

For $n = 12$ drop 2 obs. on each side.

$\longrightarrow$ (7+18+43+85+91+98+100+130)/8=71.5

A number without an indication of its precision is useless!

$\longrightarrow$ Inferential Statistics, confidence intervals!

Probability models needed!

2 Nonparametric Tests

2.1 "Nonparametric" means different things

a Parametric Model:

Regression: $Y_i = h\langle \underline{x}_i; \underline{\theta} \rangle + E_i$

"structural" part $h\langle \underline{x}_i; \underline{\theta} \rangle$ + "random" part $E_i \sim \mathcal{F}\langle \sigma \rangle$

Similarly:

Analysis of Variance, Glim, Discriminant A., Time Series, ...

b Nonparametric Tests:

Distribution of the random variables not assumed to follow a parametric family.

More precise term: "distribution free tests".

c "Distribution free tests" also means:
The test statistic has the same distribution
for all distributional assumptions for the observations.
Goodness of fit tests: Do the observations fit a given distr.?
... or a member of a parametric family? See below.

d Chisquared Test for goodness of fit

e „Nonparametric Regression“:
“structural” part not (primarily) as a formula with parameters but as a smooth, otherwise arbitrary function.
$\longrightarrow$ Block about nonpar. regression

f Density estimation: Estimate the distribution from the data.

2.2 Rank Methods

a Rank tests:

- Signed Rank Test of Wilcoxon
  for matched pairs or a simple sample,
- Rank Sum Test of Wilcoxon, Mann and Whitney
  (U-Test) for 2 independent samples.

Observations X_i with distribution $\mathcal{F}$ $\longrightarrow$
Ranks R_i = always the numbers 1 to n (if there are no ties)

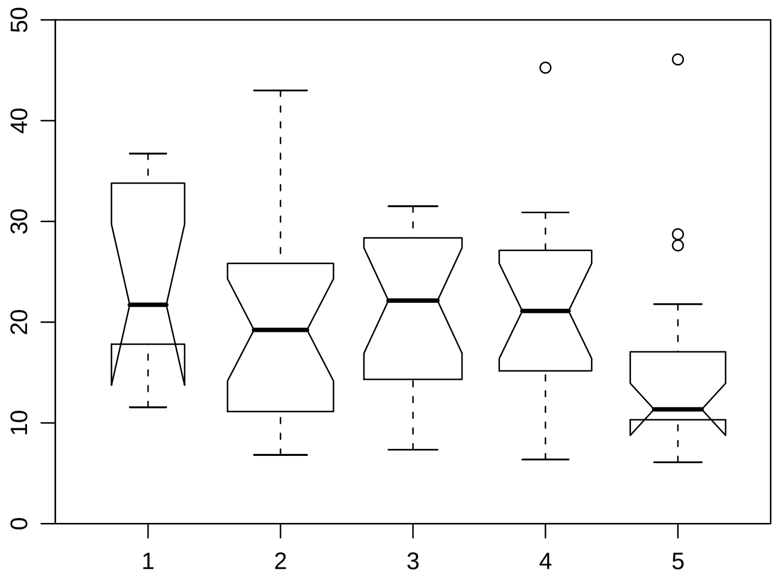
b No assumptions?

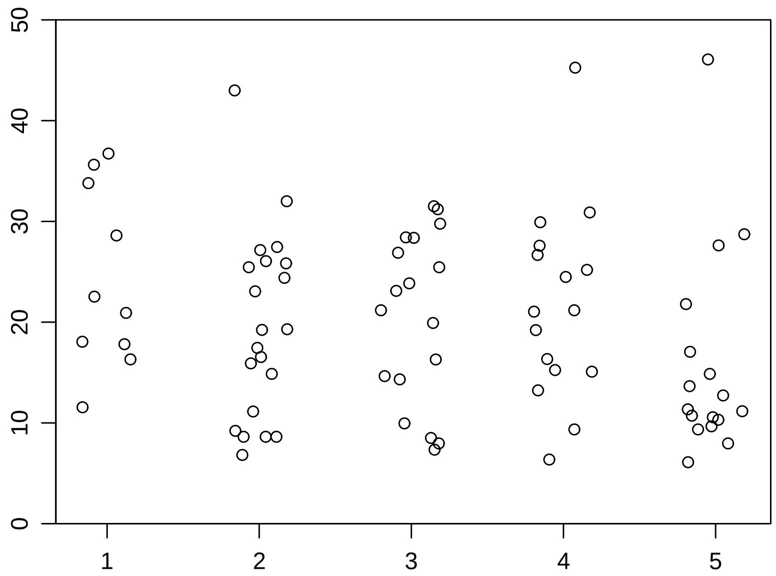
- Signed Rank Test assumes symmetry.
  Matched pairs $\longrightarrow$ differences
  $\longrightarrow$ symmetry is very plausible.

- U test assumes same distr. (shifted) for both samples.
  $\longrightarrow$ Not suitable for testing equality of expected values
  with different scatter!

c More than 2 samples $\longrightarrow$ simple analysis of variance.
Idea: Rank transformation, then simple analysis of variance.
Need to determine new distr. of the test statistic under H_0.

d Example neurons.





e Test-recipe:

H_0 : $Y_{h,i} \sim \mathcal{F}\,(i.i.d)$, $\mathcal{F}$ arbitrary, but the same for all i.

H_A : $Y_{h,i} \sim \mathcal{F}_h$, $F_h\langle x\rangle = F_1\langle x-\delta_h\rangle$, at least one $\delta_h \neq 0$

U : Test statistic: – Determine ranks $R_{h,i}$ (among all observations) – $S_h = \sum_i R_{h,i}$, $T = \frac{12}{n(n+1)} \sum_h \frac{S_h^2}{n_h} - 3(n+1)$

$\mathcal{F}_0\langle U\rangle$: Distribution of T does not depend on $\mathcal{F}$.
Small samples $\longrightarrow$ combinatorics.
Otherwise: Asymptotic χ^2 distribution.

f Example:

```
> kruskal.test(medMaxL~treat, data=d.neurit)
Kruskal-Wallis rank sum test
data:  medMaxL by treat
Kruskal-Wallis chi-squared = 7.2709, df = 4,
  p-value = 0.1222
```

F-Test: $p = 0.208$.

Without 3 outliers: F-Test $p = 0.022$, Kruskal-Wallis $p = 0.044$.

g Two way analysis of variance Block design.

Example: sales of 5 products in 7 stores

	A	B	C	D	E
1	5	4	7	10	12
2	1	3	1	0	2
3	16	12	22	22	35
4	5	4	3	5	4
5	10	9	7	13	10
6	19	18	28	37	58
7	10	7	6	8	7

= several matched samples

h The model was

$$Y_{h,i} = \mu + \alpha_h + \beta_i + E_{h,i}$$

If the treatments (products) differ clearly,
then the rankings within all blocks i are similar.
$\longrightarrow$ Friedman Test

H_0 : $Y_{h,i} \sim \mathcal{F}_i$, independent of h.

H_A : $Y_{h,i} = \mu + \alpha_h + \beta_i + E_{h,i}, \quad E_{h,i} \sim \mathcal{F}$.
at least one $\alpha_h \neq 0$.

U : Determine ranks $R_{h,i}$ within each block i.
$S_h = \sum_i R_{h,i}$,
$U = \sum_h (S_h - b(g+1)/2)^2$

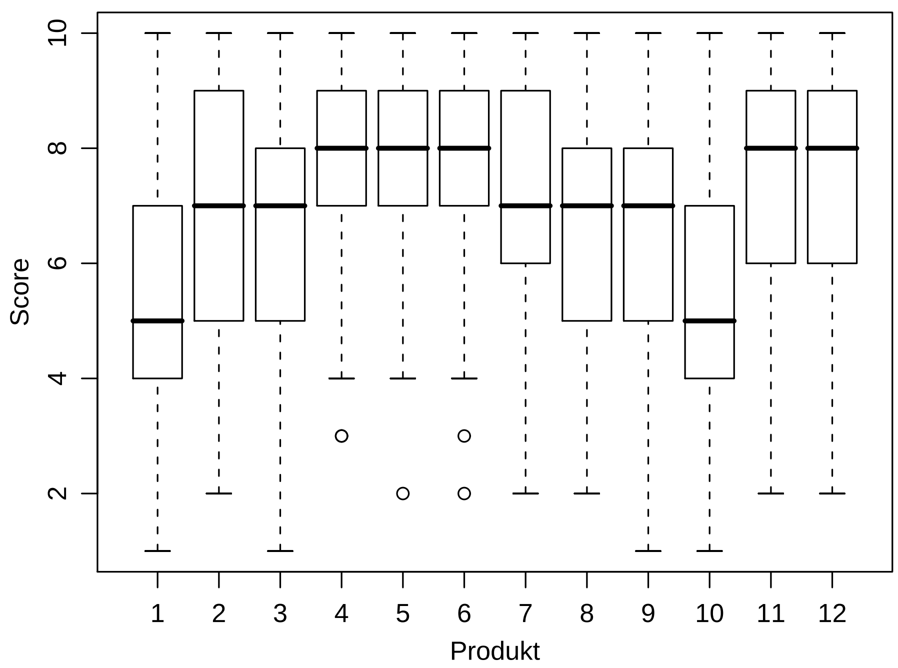
$\mathcal{F}_0\langle U \rangle$: $T = \frac{12}{bg(g+1)} U$
Combinatorics, asymptotic approximation.

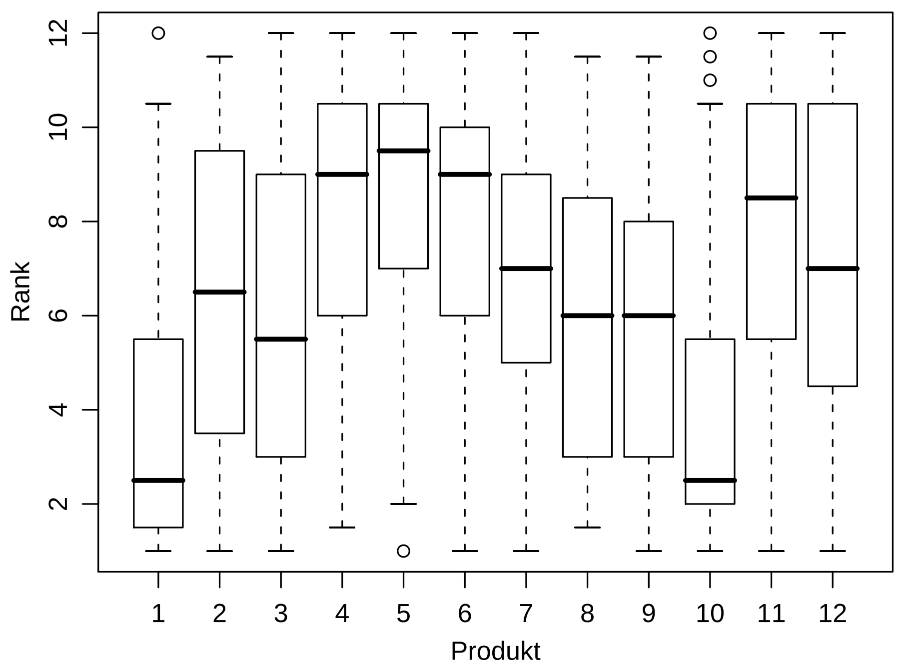
```
        Friedman rank sum test
data:  sales and brand and store
Friedman chi-squared = 8.3284, df = 4,
  p-value = 0.08026
```

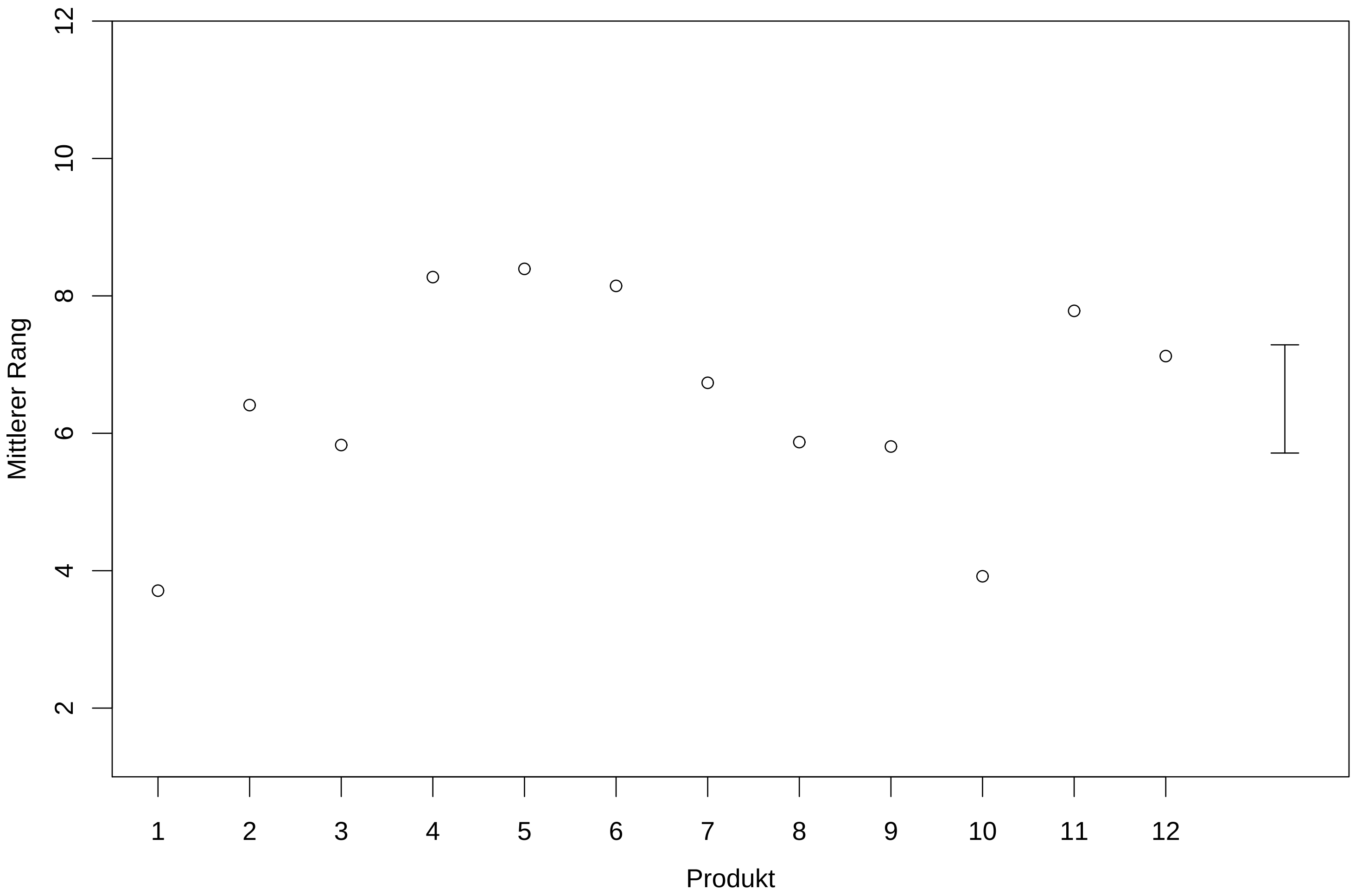
i **Sensory experiment**

Scores (1 to 10) for the visual assessment of a pastury

117 test persons








```
Friedman rank sum test
data:  t.dr
Friedman chi-squared = 270, df = 11, p-value < 2.2e-16
```

j $g = 2 \longrightarrow$ Matched pairs $\longrightarrow$?

k more complicated analyses of variance

l* Rank Regression:

Replace squares by functions of the ranks of residuals.

2.3 Rank correlation

a Rank correlation oft Spearman:

$=$ correlation of rank transformed data

$\longrightarrow$ Test for independence (hence correlation 0)

$\longrightarrow$ Measure for dependence

$\approx$ simple correlation of normally distr. data

Invariant under monotone transformation of the variables.

Similar: Rank correlation of Kendall.

b* Rank methods for multivariate statistics:

There is no “natural ordering” $\longrightarrow$ Ranks in multivariate space.

Research topic.

2.4 Goodness of fit tests

a Are the data compatible with an assumption about their distribution?
QQ plots instead of formal tests, since tests can never prove an model right.

b QQ plot: Empirical quantiles vs. theoretical q.
Quantile $= F^{-1}\langle p_k \rangle$.
F according to H_0 / from data: $\widehat{F}_n\langle x \rangle = \#\{i \mid X_i \leq x\}/n$

Test statistic $T = \max_x \langle |\widehat{F}_n\langle x\rangle - F\langle x\rangle| \rangle$.
Distribution does not depend on F (if F is continuous):

$Y_i = g\langle X_i\rangle \Longrightarrow F^{(Y)}\langle y\rangle = F\langle g^{-1}\langle y\rangle\rangle$.
$\longrightarrow$ T, calculated for Y_i and $F^{(Y)}$ = T for X_i and F.
$g = F^{-1} \Longrightarrow F^{(Y)} =$ uniform distribution between 0 and 1.
$\Longrightarrow$ Distribution of T for $F =$ uniform is valid for all F.

Kolmogorov (1933): asymptotic approximation
obtained by very elegant and fundamental methods.
$\longrightarrow$ important for probability theory!

c Test for shape of distribution: $\mathcal{N}\langle \mu, \sigma\rangle$.
Estimate μ, σ from the data $\longrightarrow$ modif. of the distr. of T.

d **Chisquared Test**. See Section 10.2 in Stahel (200x).

e Chisquared test with suitable choice of classes
has larger power against impotant alternatives (long tails)
than the Kolmogorov test.

2.5 Outlook

- Test against different scatter and general alternatives for 2 independent samples,
  e.g., comparison of 2 empirical distribution functions.
  (Test of Kolmogorov and Smirnov)
- Analysis of variance:
  Tests against specific alternatives, like monotone effects, methods for contrasts
- Methods for Survival Data ($\longrightarrow$ block on Survival)

Literature

Hollander & Wolfe (1999): Comprehensive,
first in form of recipes, then more profound.

Büning & Trenkler (1994): old, in German

Hettmansperger (1984): Rank methods