

Biostatistics week 9

- Wrap-up simple regression
 - standard errors
 - confidence and prediction interval
 - logged response model
- Adjusting for covariables
 - Multiple regression

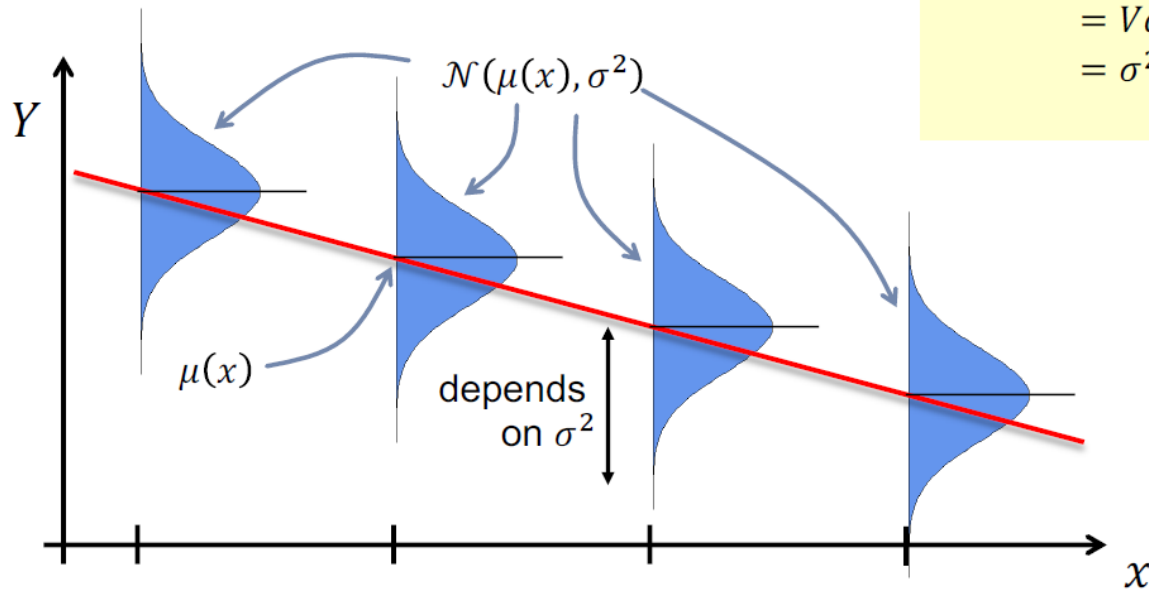
Linear regression: Two possible model definitions

1. $(Y|X = x) \sim N(\underbrace{\beta_0 + \beta_1 x}_{\mu(x)}, \sigma^2)$

2. $Y = \beta_0 + \beta_1 x + \varepsilon$

▪ $\varepsilon \sim N(0, \sigma^2)$

$$\begin{aligned} E(Y) &= E(\beta_0 + \beta_1 x + \varepsilon) \\ &= \beta_0 + \beta_1 x + E(\varepsilon) \\ &= \beta_0 + \beta_1 x \\ \text{Var}(Y) &= \text{Var}(\beta_0 + \beta_1 x + \varepsilon) \\ &= \text{Var}(\varepsilon) \\ &= \sigma^2 \end{aligned}$$



Use linear regression to predict bodyfat from BMI

```
> r.bodyfat <- lm(bodyfat ~ bmi,d.bodyfat)
```

```
> summary(r.bodyfat)
```

Call:

```
lm(formula = bodyfat ~ bmi, data = d.bodyfat)
```

Residuals:

Min	1Q	Median	3Q	Max
-13.5485	-3.5583	0.0785	4.0384	12.7330

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-26.9844	2.7689	-9.746	<2e-16 ***
bmi	1.8188	0.1083	16.788	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.573 on 241 degrees of freedom

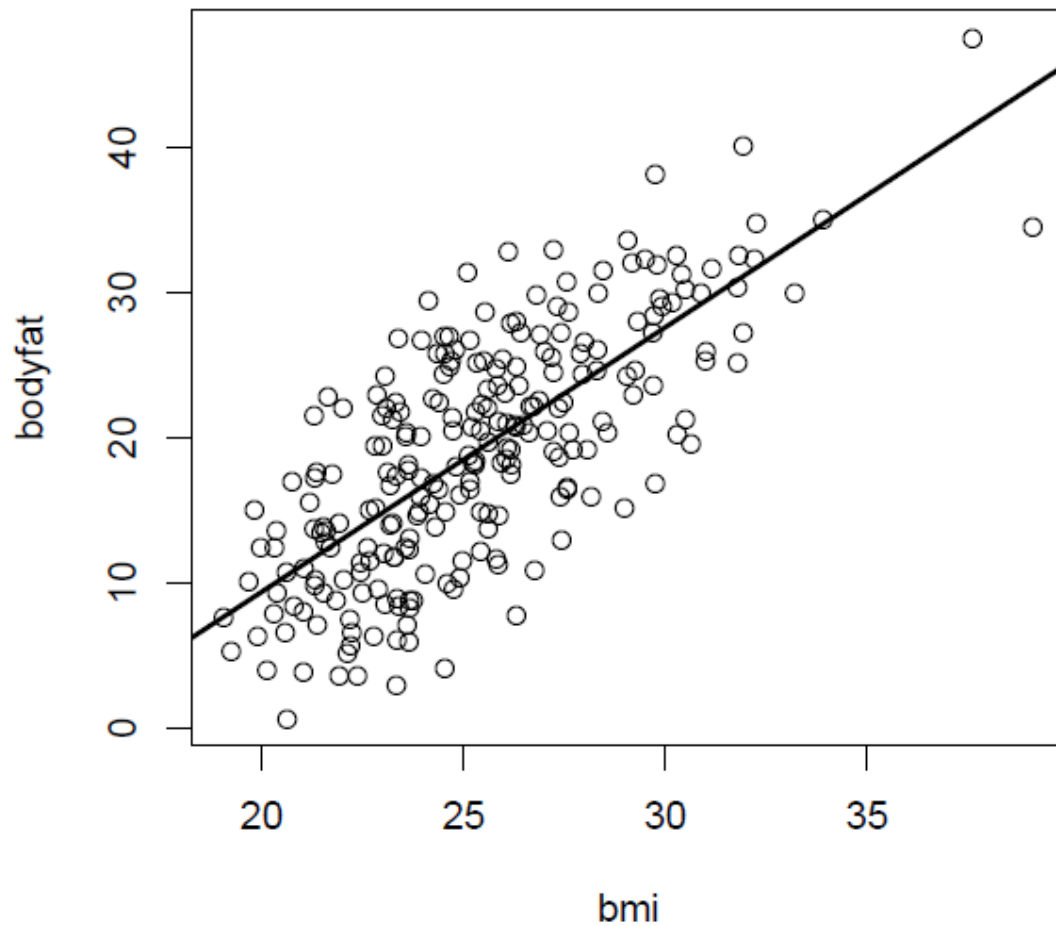
Multiple R-squared: 0.539, Adjusted R-squared: 0.5371

F-statistic: 281.8 on 1 and 241 DF, p-value: < 2.2e-16

$$\Rightarrow \hat{\alpha} = -26.98, \hat{\beta} = 1.82, \hat{\sigma}_e = 5.57$$

Plot the regression line into the scatterplot

```
> plot(bodyfat ~ bmi,d.bodyfat)  
> abline(r.bodyfat,lwd=2)
```



Assumptions of a linear regression model

Before we continue to look into the results, we need to **check if the modelling assumptions are met!**

Why? Because otherwise we draw invalid conclusions from the results.

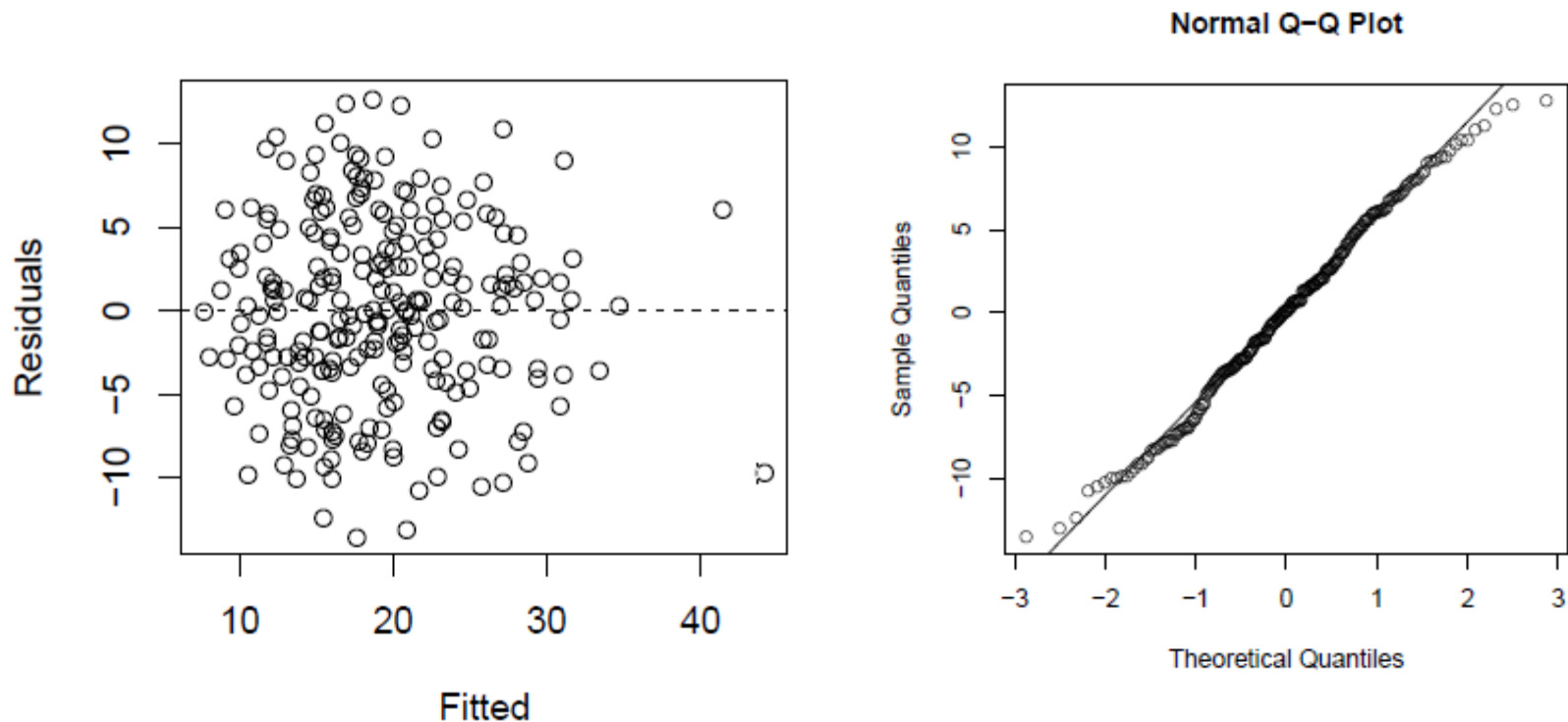
The assumption we took here is that the errors $\varepsilon_i \sim \mathbf{N(0, \sigma^2)}$

This implies four things for the residuals which we use as estimates for the errors:

- a) The expected value of r_i is 0: $E(r_i) = 0$.
- b) All r_i have the same variance: $\text{Var}(r_i) = \hat{\sigma}^2$.
- c) The r_i are normally distributed.
- d) The r_i are independent of each other.

Model checking

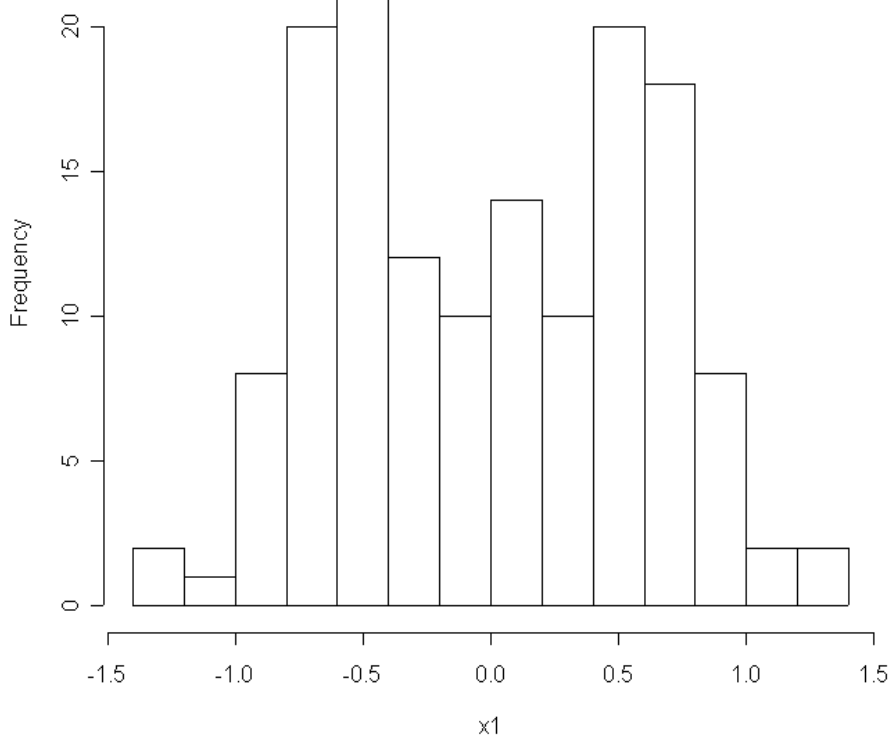
The **Tukey-Anscombe diagram** plots the **residuals against the fitted values** (see left plot) is the most important model checking tool. This plot is ideal to check if assumptions a) and b) (and partially d)) are met. Normality of the residuals is assumption c) and can be easily checked – e.g. by inspecting the Normal-Q-Q plot of the residuals (see right plot). Here, all is o.k..



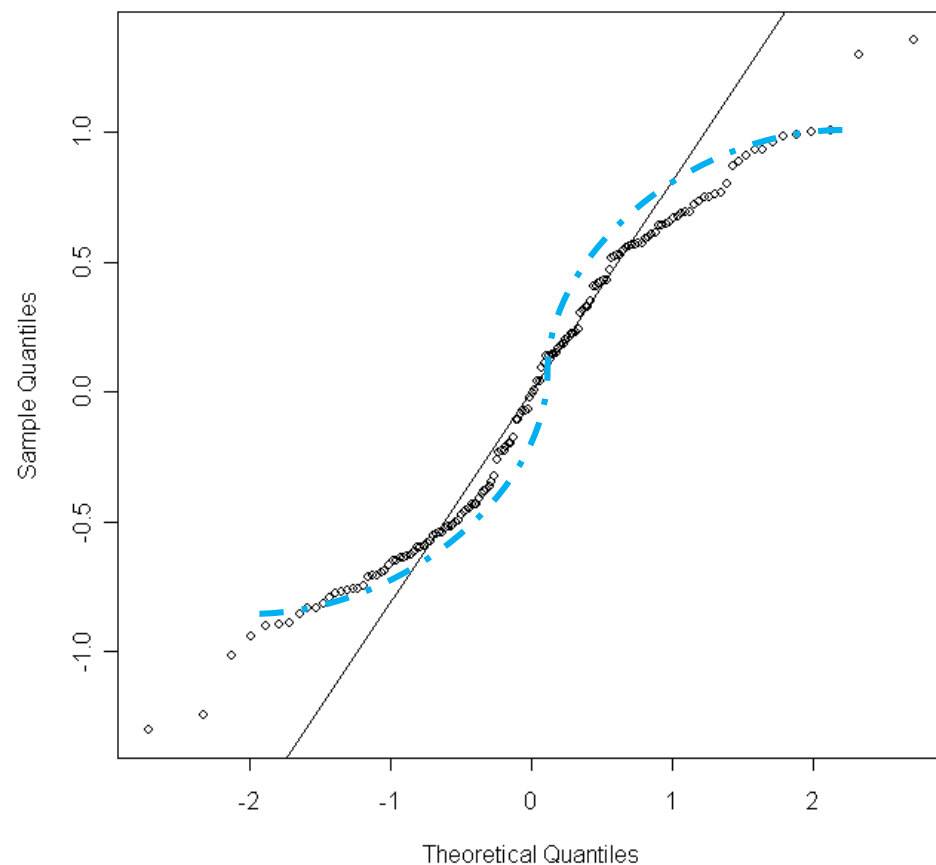
What to learn from a Normal-QQ-Plot?

Short Tails (compared to Gaussian bell)

Histogram of x1



Normal Q-Q Plot



If the qq-Plot shows a “S”-shape the empirical distribution is short-tailed

Examples of Tukey-Anscombe-Plots

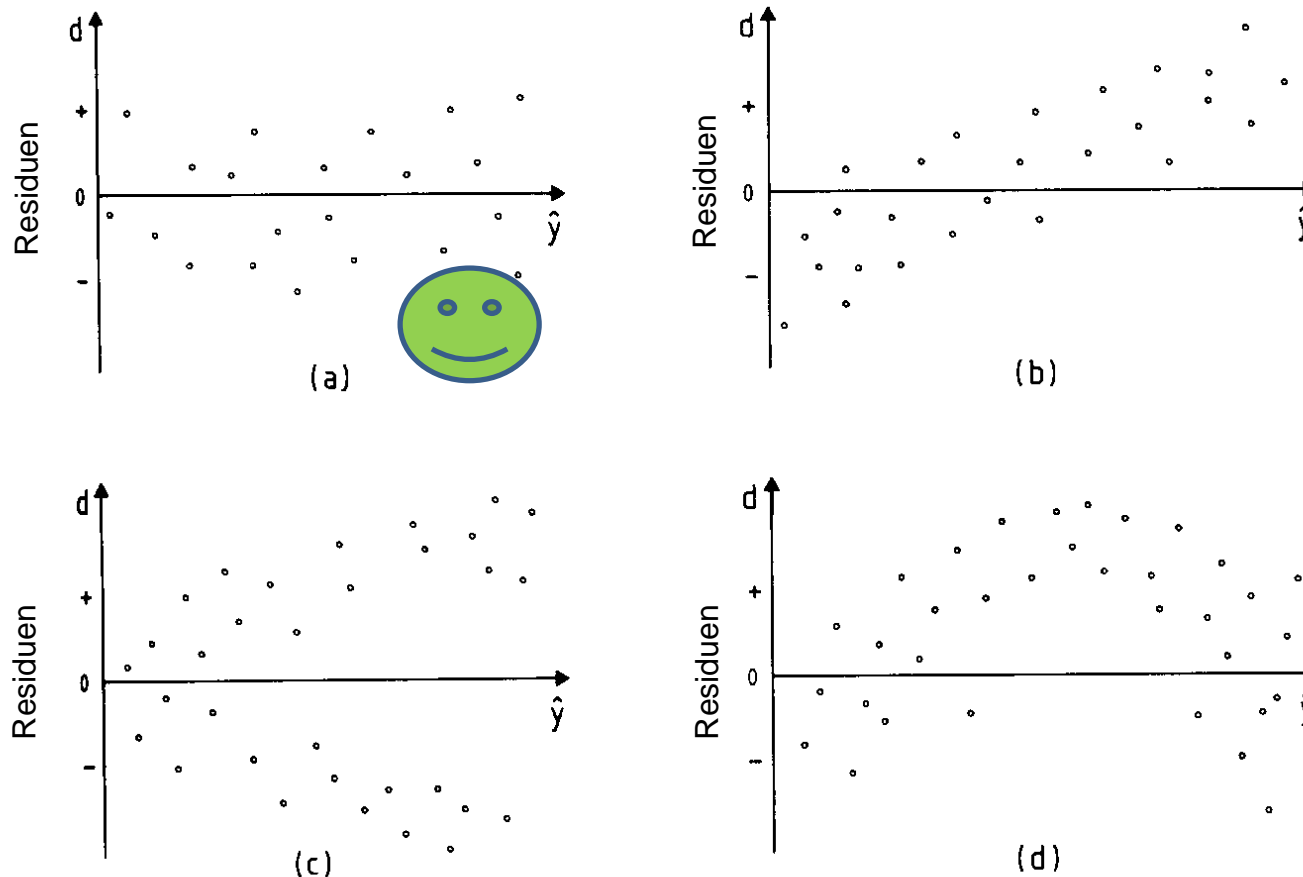


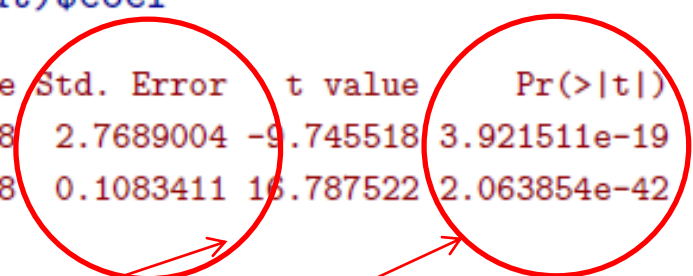
Abb. 12: Graphische Darstellung der normierten Residuen in Abhängigkeit von den geschätzten Werten \hat{y} des Regressanden: (a) idealer Verlauf; (b) linearer Trend, auf einen Rechenfehler hindeutend; (c) ansteigende Varianzen (mit \hat{y}), Wechsel zu einem Modell für ungleiche Varianz der Beobachtungen eventuell angebracht; (d) nichtlinearer Verlauf der Residuen, inadäquates Modell, d. h. Transformation der Daten oder Änderung der Regressionsfunktion angezeigt

Uncertainty of the estimated model coefficients

Lets look at the estimates and the standard errors of the model coefficients:

```
> summary(r.bodyfat)$coef
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-26.984368	2.7689004	-9.745518	3.921511e-19
bmi	1.818778	0.1083411	16.787522	2.063854e-42



The second column shows a standard error of the estimated model coefficient quantify the uncertainty of the estimated coefficients.

The Null Hypothesis is, that the corresponding coefficient is zero, if the p-value is smaller that 0.05, we have evidence to reject H_0 .

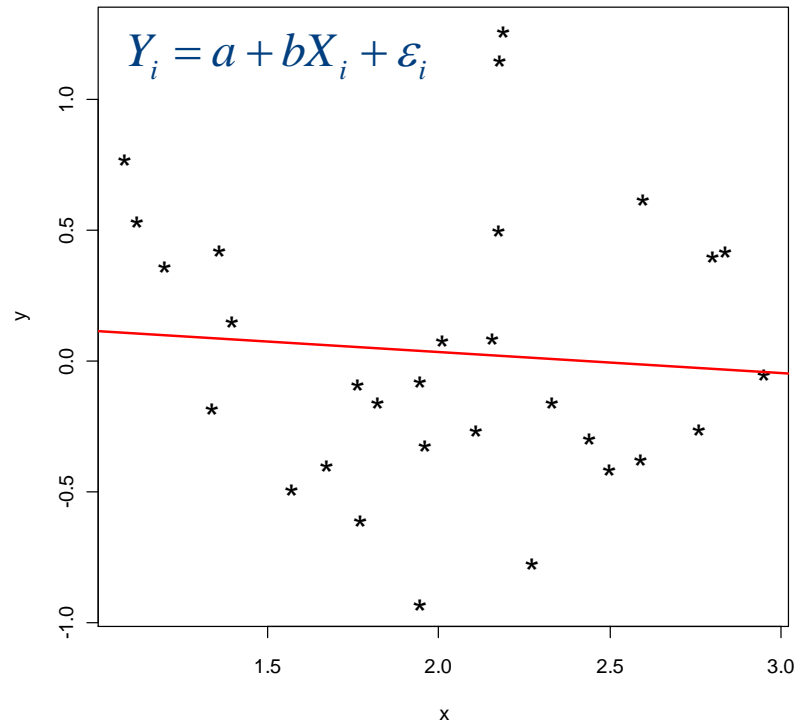
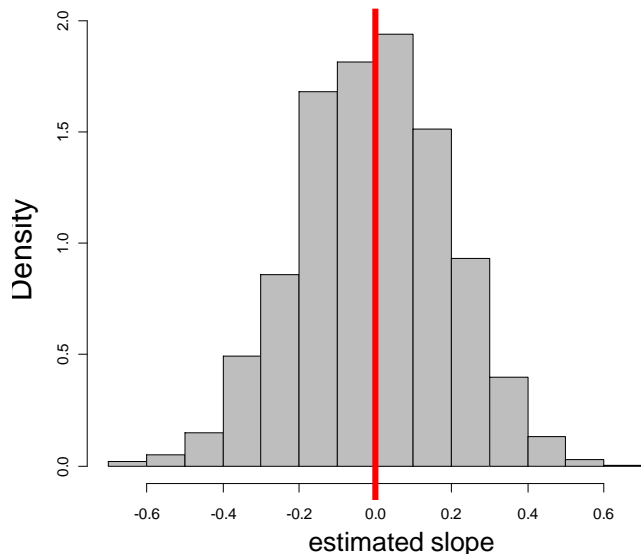
How does the distribution of the estimated coefficients look like?

Determine the distribution of the coefficients under H_0

```
x=runif(n=30,min=1,max=3 )
my.err = rnorm(n=30,mean=0,sd=0.5)
y=0*x+0+my.err
plot(x,y,pch="*",cex=2)
mod.l=lm(y~x)
abline(mod.l,lwd=2,col="red")
coef(mod.l)[2] # -0.0049
```

Repeat simulation under H_0 5000 times:

$$\hat{b} \sim N(0, se(\hat{b})^2)$$



The ordinary least square (OLS) fit yields **unbiased estimates of the coefficients**.

See simulation, where the estimated coefficients are normally distributed with expected value equal to the true coefficients which have been used to simulate the data.

Standard errors and confidence intervals of the coefficients

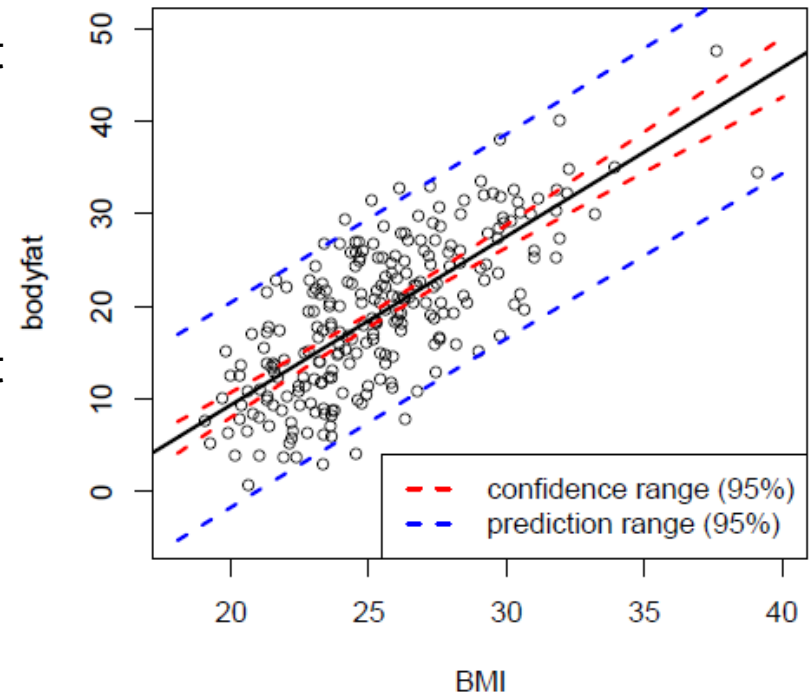
Intercept	$\hat{a} \pm q^{t_{n-2}}_{\alpha/2} \cdot se(\hat{a})$	$se(\hat{a}) = \sqrt{\hat{\sigma}_\varepsilon^2 \cdot \left(\frac{1}{n} + \frac{\bar{x}}{\sum (x_i - \bar{x})^2} \right)}$
Slope	$\hat{b} \pm q^{t_{n-2}}_{\alpha/2} \cdot se(\hat{b})$	$se(\hat{b}) = \sqrt{\frac{\hat{\sigma}_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$

remark: the estimates are normally distributed- still we have to use the quantiles of the t-distribution for the CI since we need to estimate the variance of the normal distribution - the formulas for the se are not relevant for the exam.

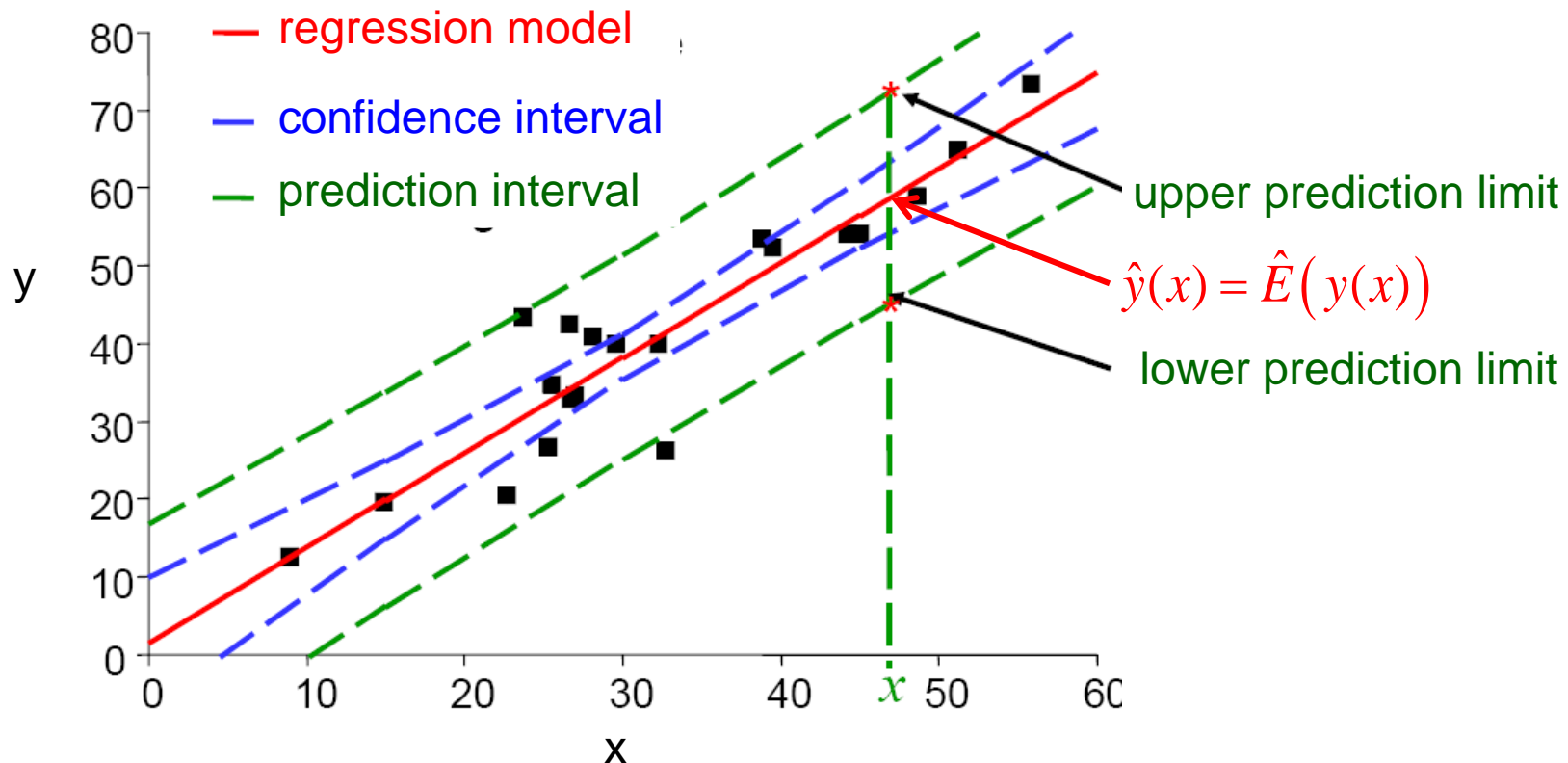
Confidence- and prediction intervals

Remember: We estimate the regression coefficients from a random sample. There are two questions to be asked:

- How to give an interval so that the risk that the **true linear model** runs not within this interval is not larger than 5%?
→ This leads to the **confidence range**.
- How to give an interval so that the risk that an **individual observation** is not within interval is not larger than 5%?
→ This leads to the **prediction range**.



Confidence- and prediction intervals



Regression yields an estimate for the expected value of y for each x value and the true expected value of y at the position x is with 95% percentage certainty covered by the confidence interval.

A single observation y at the position x is with 95% percentage certainty within the prediction interval.

Interesting intervals when doing regression

$$Y_i = a + bX_i + \varepsilon_i$$

CI for
y-intercept:

$$\hat{a} \pm q^{t_{n-2}}_{\alpha/2} \cdot se(\hat{a})$$

```
➤ confint(fit, parm=1, level=0.95)
                2.5 %      97.5 %
(Intercept) -32.438703 -21.530032
```

CI for
the slope:

$$\hat{b} \pm q^{t_{n-2}}_{\alpha/2} \cdot se(\hat{b})$$

```
➤ confint(fit, parm=2, level=0.95)
                2.5 %      97.5 %
x           1.605362  2.032195
```

CI
for $E(y_k)$:

$$\hat{y}_k \pm q^{t_{n-2}}_{\alpha/2} \cdot se(\hat{y}_k)$$

```
➤ predict(fit, new, se.fit=T,
           interval=c("confidence"))
fit      lwr      upr
23       22.70   23.3
```

Prediction
interval for y_l :

$$\hat{y} \pm q^{t_{n-2}}_{\alpha/2} \cdot se(\widehat{\sigma_\epsilon})$$

```
➤ predict(fit, new, se.fit=T,
           interval=c("prediction"))
fit      lwr      upr
23       21.20   24.80
```

Plotting Confidence and Prediction Interval


Note:

Visualizing the confidence and prediction intervals in R is not straightforward, but requires some tedious handwork.

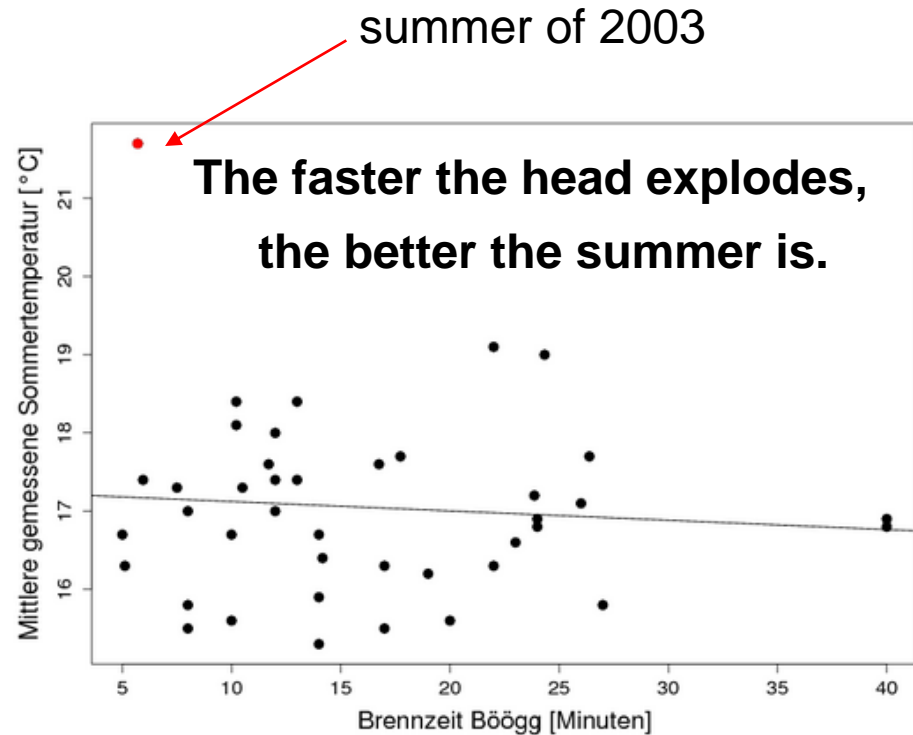
R-Hints:

```
dat  <- data.frame(x=seq(...,..., length=200))
pred <- predict(fit, newdata=dat, interval=...)
plot(..., ..., main="...")
lines(dat$x, pred[,2], col=...)
lines(dat$x, pred[,3], col=...)
```

New data for prediction must be provided in a data frame and the columns with the predictors must have the same names as in the data frame that has been used in fit.

A red arrow originates from the text and points to the 'x' parameter in the first line of the R code, which is circled in red.

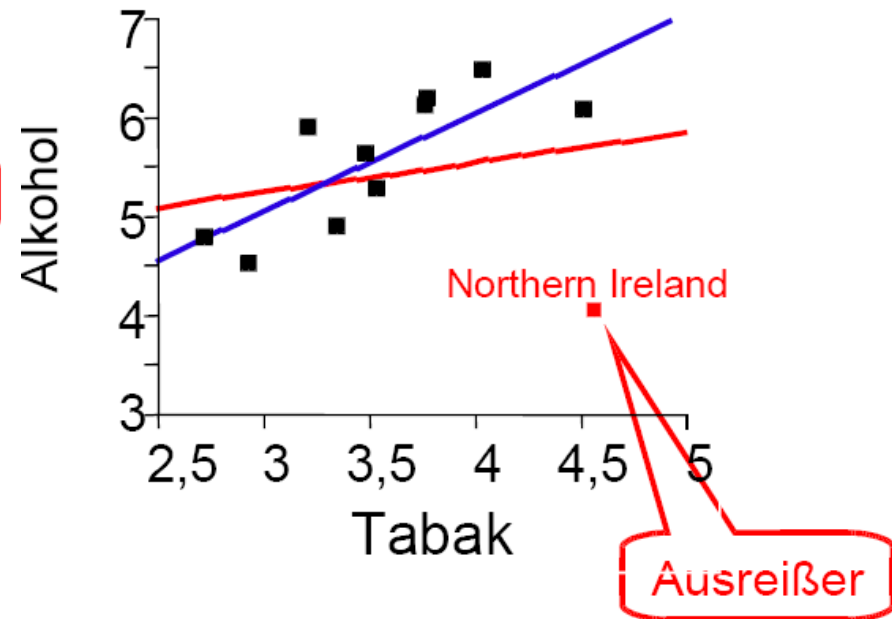
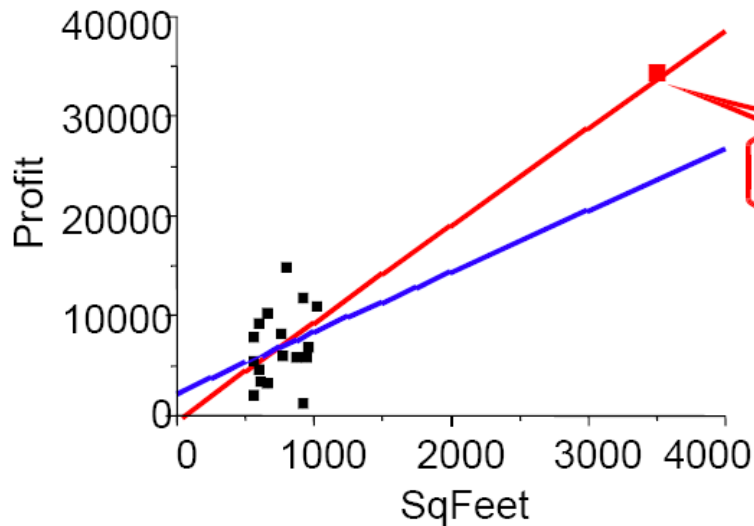
The „Böögg“ as weather oracle



outliers violate the model assumption of i.i.d. normal distributed residuals! i.e. linear regression as we have seen until now is not a suitable paradigm anymore.

Outlier are dangerous!

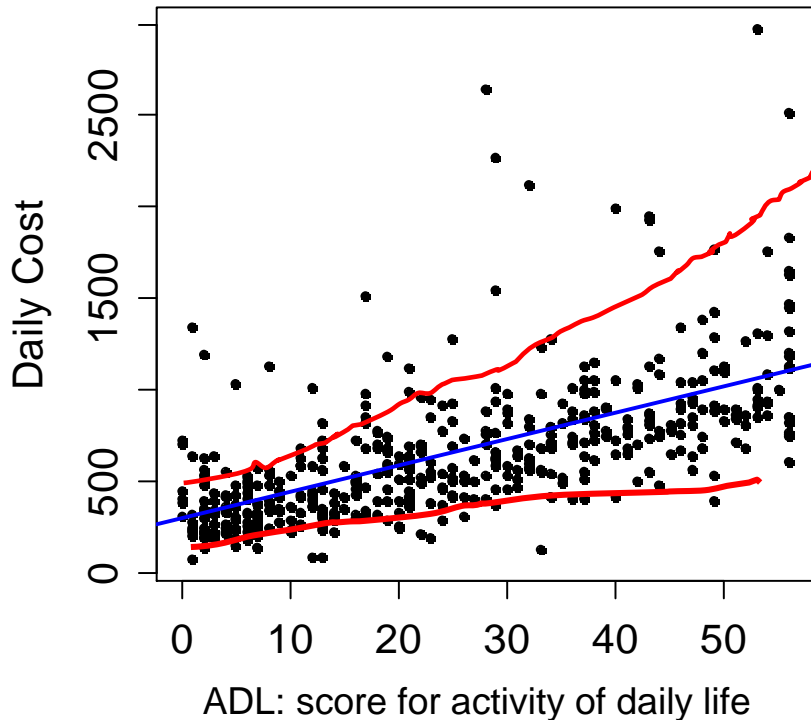
especially when they occur at the border of the x-range



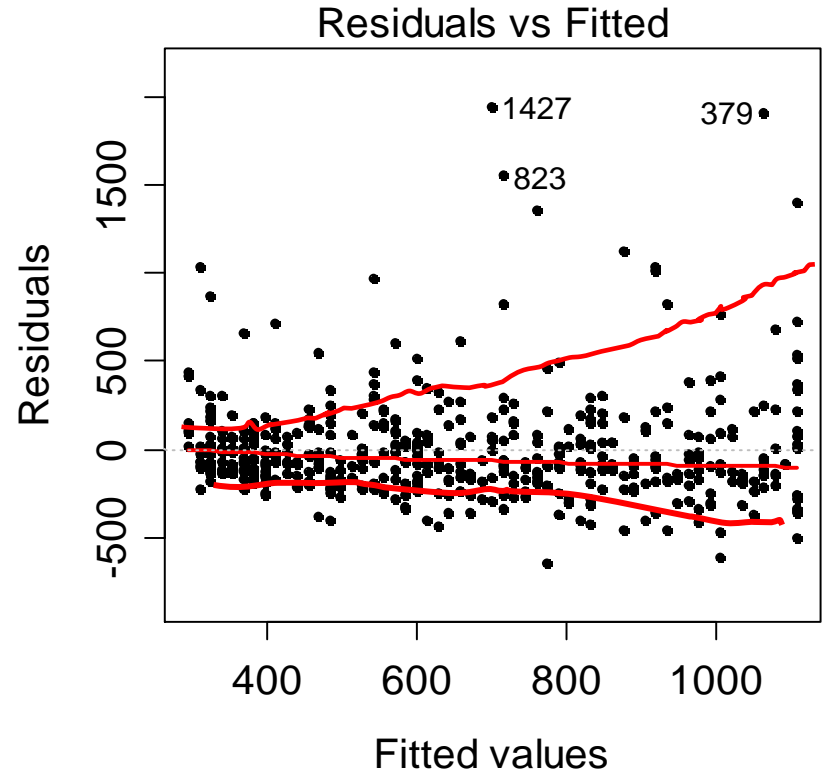
An **Outlier** can have giant impact on the result, especially when it occurs at the border of the x-range (**leverage point**). The result can get totally influenced by such an observation and therefore be useless. In such cases one can apply **robust regression** (`rlm(...)`).

What can we do, if residual variance is not constant?

Daily Cost in Rehab vs. ADL



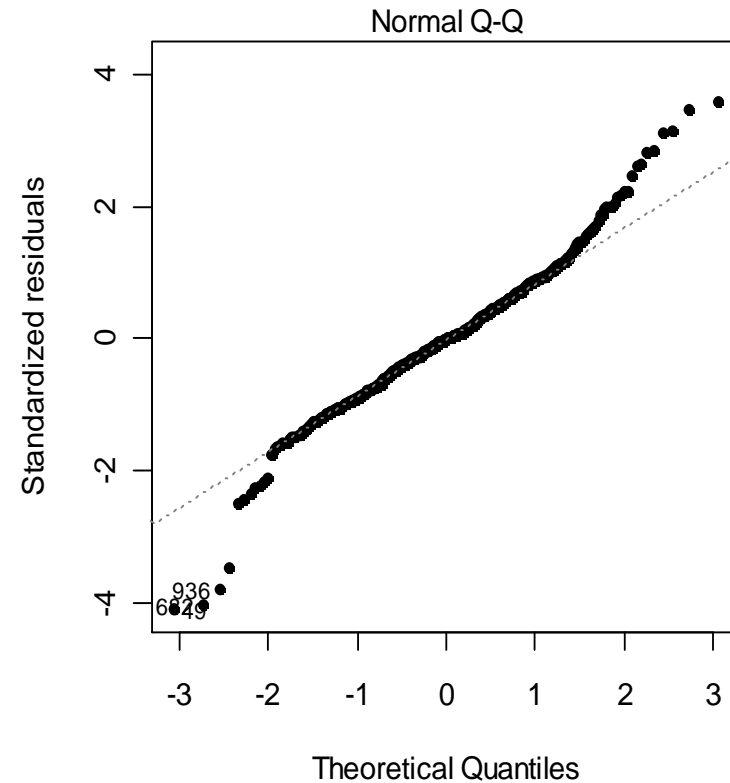
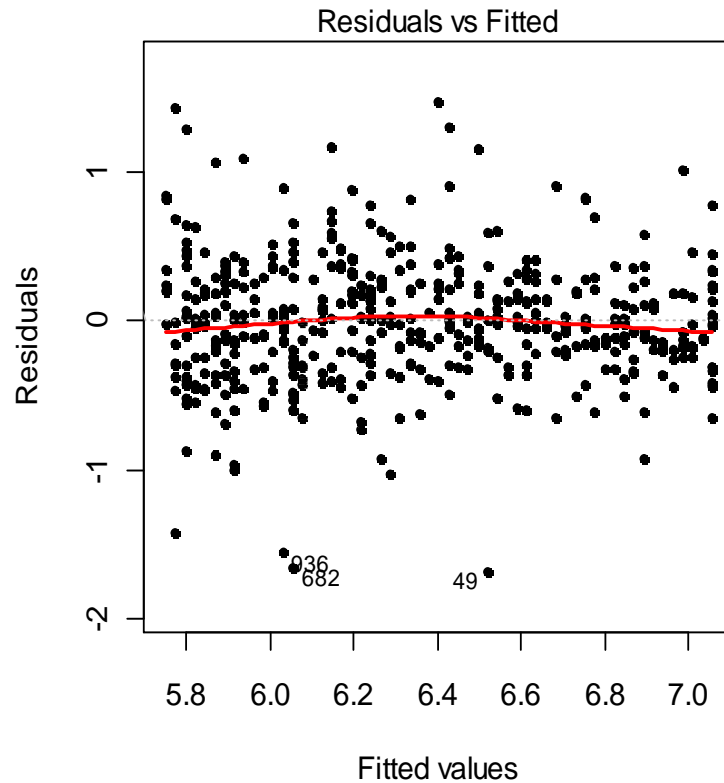
Residuals vs. Fitted Values



In case of **multiplicative errors** we see usually a **funnel shape in the data/residuals**. We already see in the original data that the variance is not constant and the points lay within a funnel. A **funnel is visible in the residual plot** together with some extreme residuals. Hence, **the assumption for the linear regression are violated**.

Log-Transformation has stabilized the variance!

$$Y' = \log(Y) = \beta_0 + \beta_1 x + E$$



Back Transforming the Fitted Values

- In principle, we can „simply back transform“ – under this transformation we get from a straight line to an exponential curve

$$\hat{y} = \exp(\hat{y}')$$

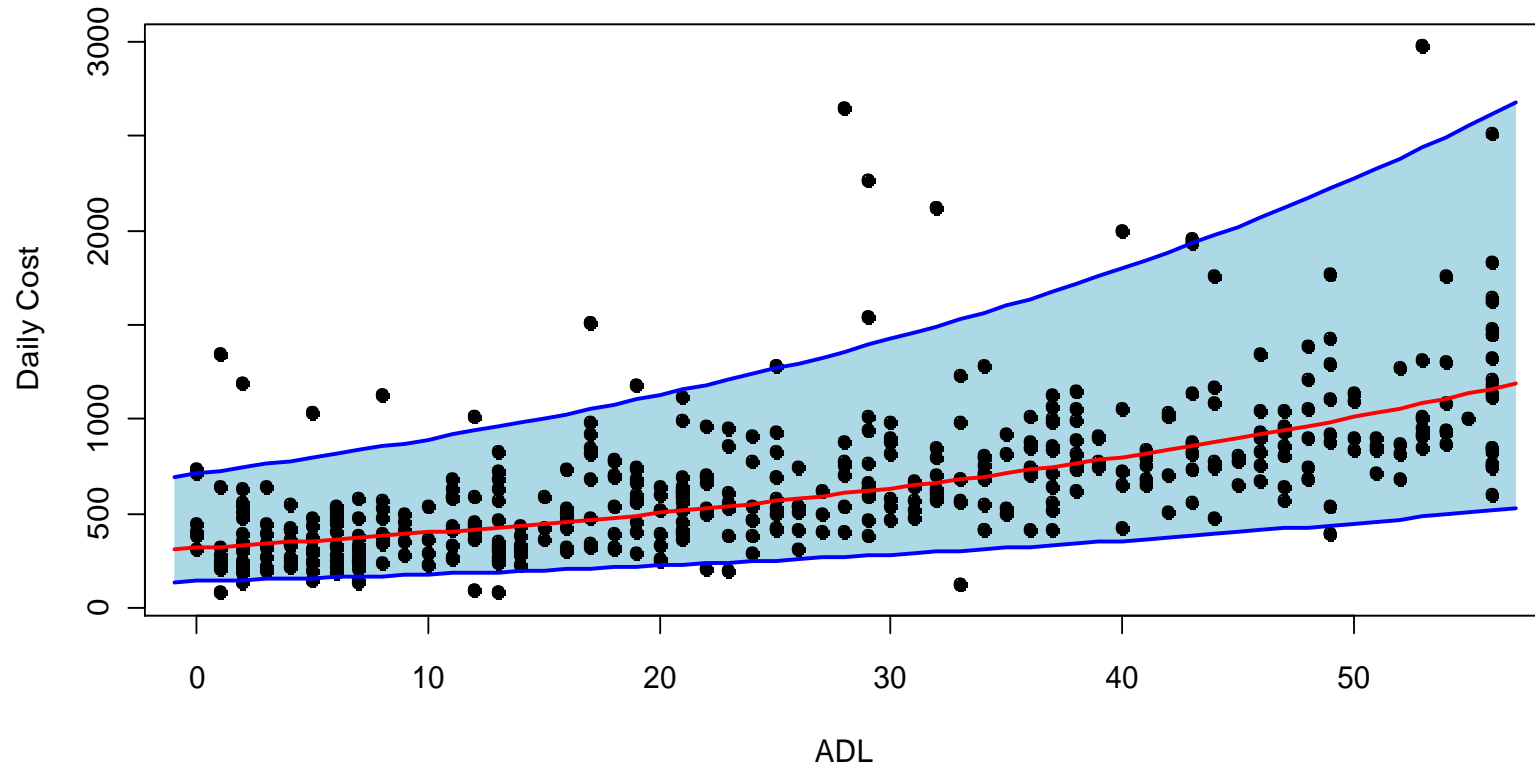
- Since the exp-transformation is a monotonic function, the quantile transform also with exp and the limits of confidence/prediction intervals are determined accordingly

$$[l, u] \rightarrow [\exp(l), \exp(u)]$$

- \hat{y}' is an estimate for the mean (=median because of symmetric distribution) of the logged outcome $\log(y)$, however (since only quantiles as the median, but not the mean, transform under a monotonic non-linear function as exp-function) \hat{y} is an estimate for the median (not the mean) of the outcome y (here: cost) conditioned on the values of the input variables (here: ADL)

Model and Prediction Intervall after Back-Transformation

Daily Cost in Rehabilitation vs. ADL-Score



First Aid Transformations

Variance-stabilizing transformations

These transformations intended to stabilize the variance

First-Aid Transformations:

→ do always apply these (if no practical reasons against it)

→ to both response and predictors

Absolute values and concentrations:

log-transformation: $y' = \log(y)$

Count data:

square-root transformation: $y' = \sqrt{y}$

Proportions:

arcsine transformation: $y' = \arcsin(\sqrt{y})$

How can we adjust for other influences?

From Table 3

Entire sample (n=140) Group with $\geq 85\%$ consumption (n=130)

Model 1 Unadjusted

Intercept	-0.31 (-0.49 to -0.13)*	-0.33 (-0.53 to -0.13)†
Challenge type		
Mix A vs placebo	0.20 (0.01 to 0.40)‡	0.24 (0.02 to 0.47)‡
Mix B vs placebo	0.16 (-0.04 to 0.35)	0.16 (-0.07 to 0.38)

Model 2 Adjusted

Intercept	-0.54 (-0.89 to -0.18)*	-0.51 (-0.92 to -0.11)
Challenge type		
Mix A vs placebo	0.20 (0.01 to 0.39)‡	0.28 (0.05 to 0.51)‡
Mix B vs placebo	0.17 (-0.03 to 0.36)	0.19 (-0.04 to 0.41)

What does
“adjusted”
mean?

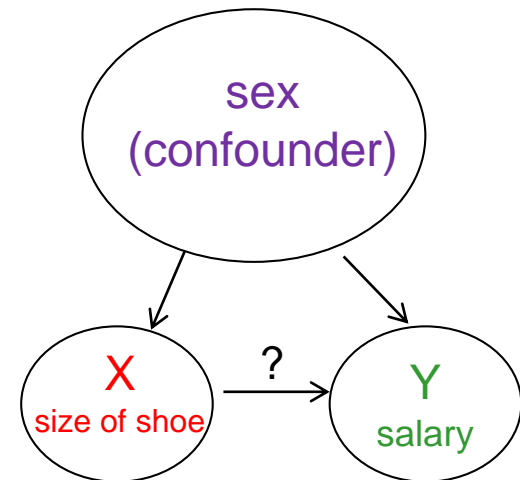
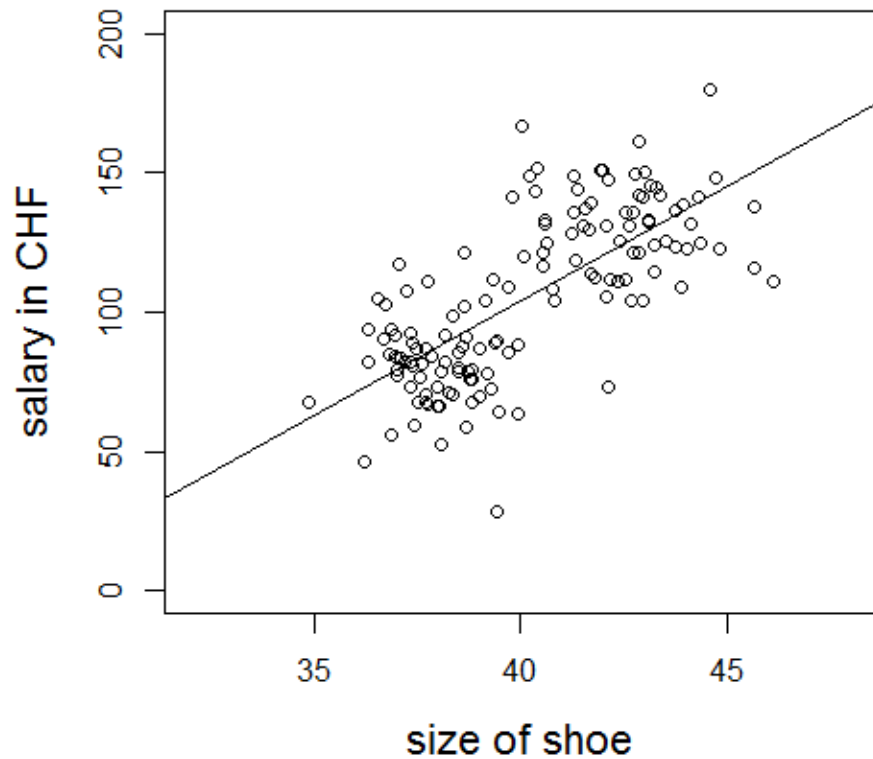
How is it done?

In model 2, in addition to challenge type, the effects of the following factors were adjusted for: week during study, sex, GHA in baseline week, number of additives in pretrial diet, maternal educational level, and social class.

How to adjust for possible confounders if working with observational data?

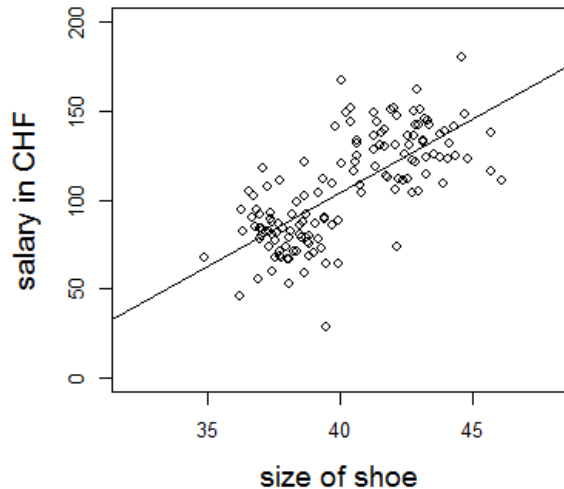
- Multiple regression model approaches
- Stratification / Matching (study desing)
- Propensity Score Adjustment (not covered in this lecture)

**Stratification (or Matching) is most often used
in case of few confounders with few levels**



Association between salary and shoe size before and after taking the sex factor into account

unadjusted on full data set: $\text{lm}(\text{salary} \sim \text{shoe})$

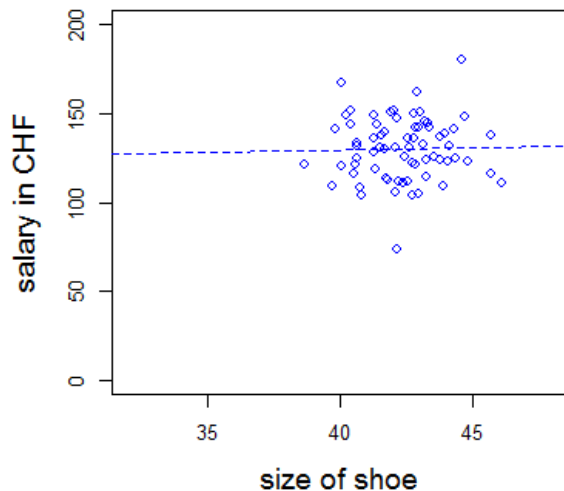


adjusted on full data set: $\text{lm}(\text{salary} \sim \text{shoe} + \text{sex})$

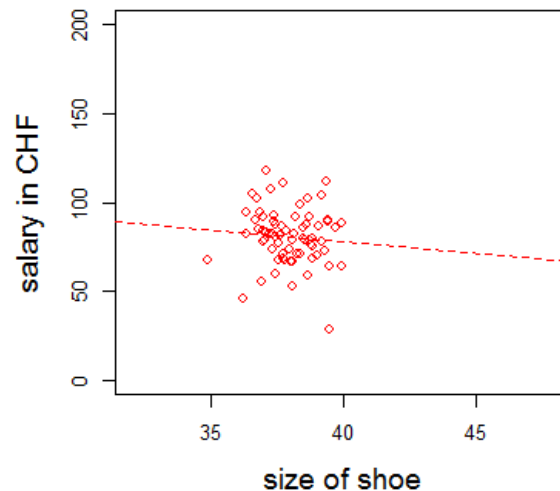


Adjusting via multiple regression without interaction leads only similar results as a stratified analysis if the interaction is not significant -> parallel regression lines are «feature of model»

males stratum: $\text{lm}(\text{salary} \sim \text{shoe})$



females stratum $\text{lm}(\text{salary} \sim \text{shoe})$



Stratified analysis -> different models for male and females are possible, but here not necessary.

Multiple linear regression: interpretation of coefficient

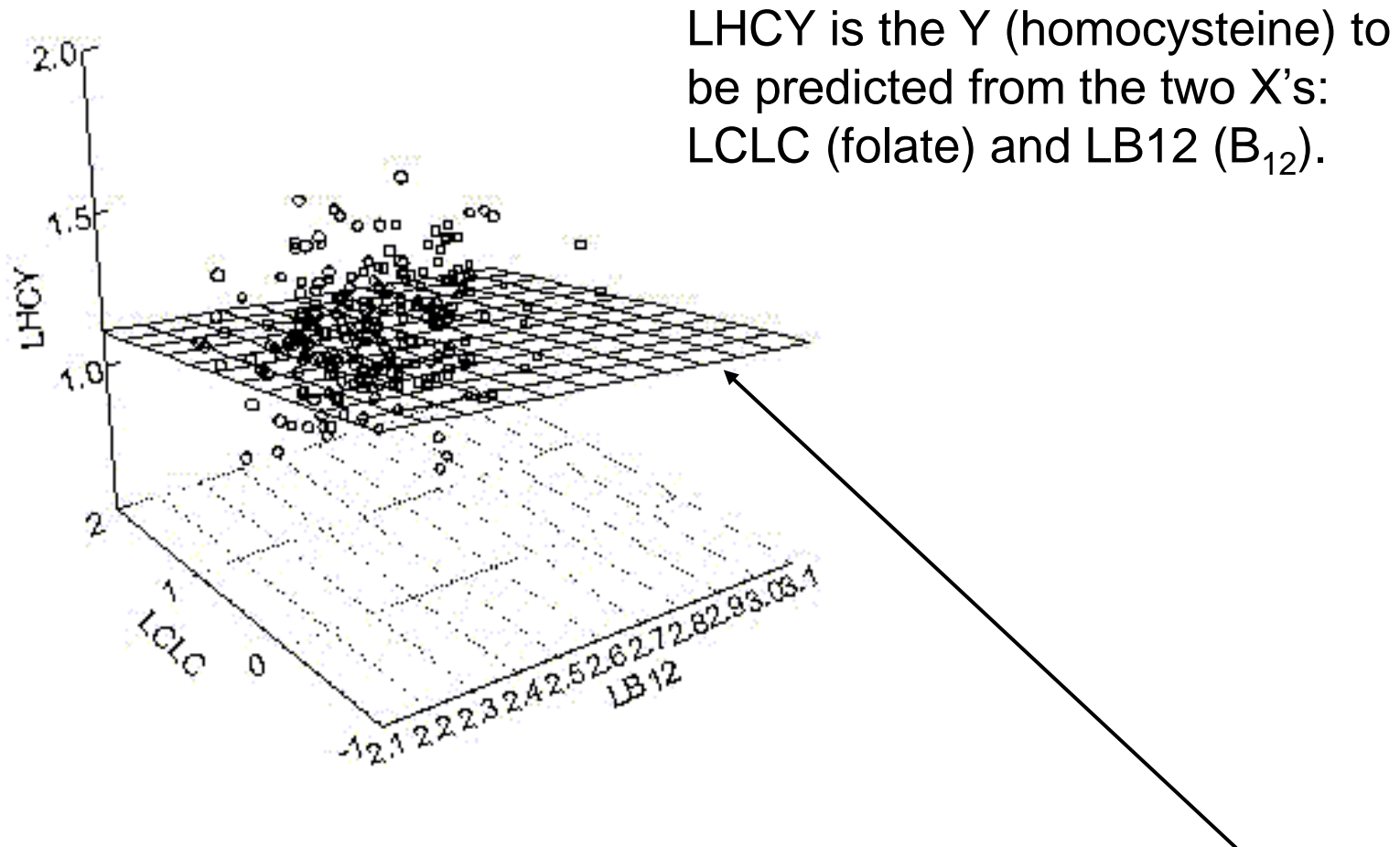
$$y_i = \beta_0 + \beta_1 \cdot x_{i1} + \dots + \beta_p \cdot x_{ip} + \varepsilon_i \quad , \quad \varepsilon_i \sim N(0, \sigma^2)$$

Matrix Notation:

$$\mathbf{y} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \text{with } \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix}, \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

The coefficient β_k gives the change of the outcome y , given the explanatory variable x_k is increased by one unit and all other variables are held constant.

Regression with tow predictor variables



$LHCY = b_0 + b_1LCLC + b_2LB12$ is the equation of the plane

Remark: If we have more than 2 predictors the model is a hyper-plane

Example for multiple Regression

Multiple Lineare Regression

Outcome we want to model:

High density lipoprotein (HDL)

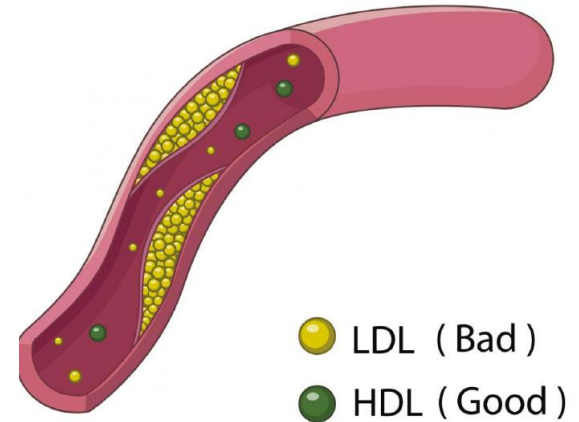
Research question:

Which predictors have an impact on HDL?

Preparation:

Make analysis plan

Collect data



Example for multiple regression: HDL example

	Estimate	Std. Error	t-value	Pr(> t)
Intercept	1.16448	0.28804	4.04	<.0001
AGE	-0.00092	0.00125	-0.74	0.4602
BMI	-0.01205	0.00295	-4.08	<.0001
BLC	0.05055	0.02215	2.28	0.0239
PRSSY	-0.00041	0.00044	-0.95	0.3436
DIAST	0.00255	0.00103	2.47	0.0147
GLUM	-0.00046	0.00018	-2.50	0.0135
SKINF	0.00147	0.00183	0.81	0.4221
LCHOL	0.31109	0.10936	2.84	0.0051

The predictors of log(HDL) are age, body mass index, blood vitamin C, systolic and diastolic blood pressures, skinfold thickness, and the log of total cholesterol. The equation is:

$$\text{Log(HDL)} = 1.16 - 0.00092(\text{Age}) - 0.012(\text{BMI}) + \dots + 0.311(\text{LCHOL})$$

Linear regression: interpretation of coefficient

$$y_i = \beta_0 + \beta_1 \cdot x_{i1} + \dots + \beta_p \cdot x_{ip} + \varepsilon_i$$

The coefficient β_k gives the change of the outcome y , given the explanatory variable x_k is increased by one unit and all other variables are hold constant.

$$\beta_k = y_{x_k+1} - y_{x_k} = y_{x_k \rightarrow x_{k+1}}$$

E.g., expected $\log(\text{LHDL})$ is 0.012 lower in a subject whose BMI is 1 unit greater, but is the same as the other subject on other factors.

$$\log(\text{HDL}) = 1.16 - 0.00092(\text{Age}) - \mathbf{0.012(\text{BMI})} + \dots + 0.311(\text{LCHOL})$$

The meanings of the p-values and the coefficients in the multiple linear regression output

The p-values measure the significance of the association of a factor with Log(HDL) in the presence of all other predictors of the model – meaning “after accounting for other factors” or “adjusting for other factors”, and is called independent association.

SKINF alone probably is associated. However, its $p=0.42$ says that it provides no *additional* information that helps to predict LogHDL, after accounting for other factors such as BMI.

The p-value and also the coefficient-value of a predictor depend i.g. not only on the association with the outcome variable but also on the other predictors in the model.

Only if all predictors are independent multiple regression leads the same p-values and coefficients than p simple regression each with only one predictor.

Very high correlated predictors can cause problems, such as large CIs.

Significance vs. Relevance

The larger a sample, the smaller the p-values for the very same predictor effect. Thus do not confuse a small p-values with an important predictor effect!!!

More important than p-values:

- **Look at absolute values of (significant) coefficients.**
- **Look at confidence intervals!**

Linear Regression with continuous and factorial predictors

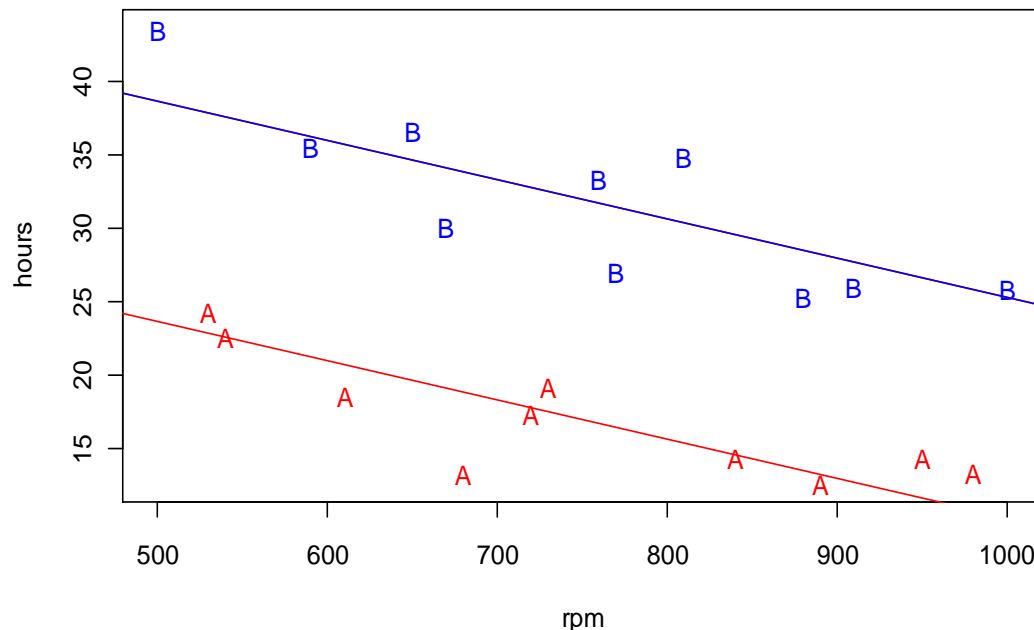
Output: **hours:** lifetime of a cutting tool

Predictor 1: **rpm:** speed of the machine in rpm

Predictor 2: **tool:** tool type A or B

↓

```
fit1 <- lm(hours ~ rpm + tool, data=my.dat)
```

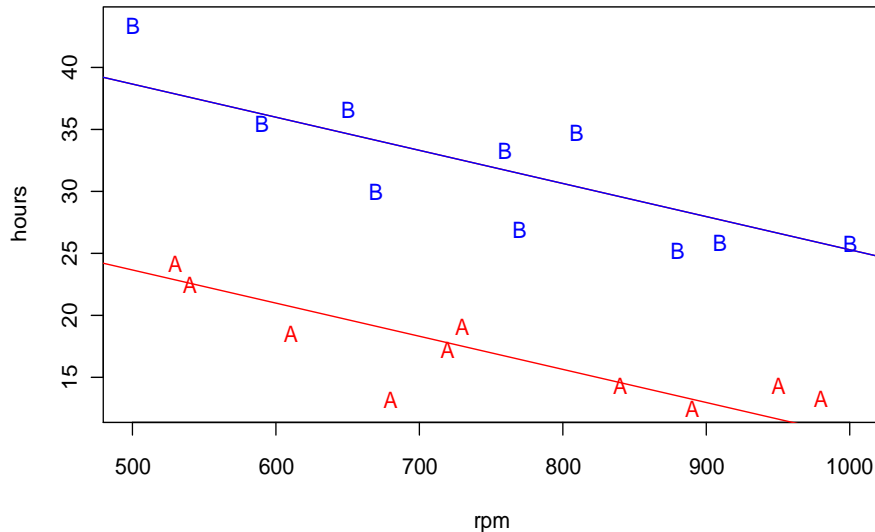


We have an **additive model**: the difference between the tools is a **shift**.

What does interaction mean?

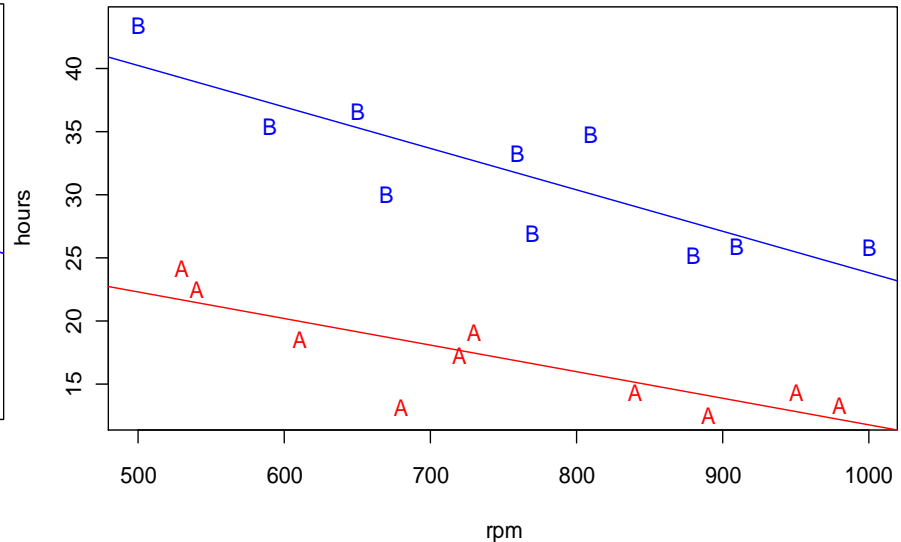
Different slopes of continuous variables at different levels of a factor

Do not allow for interaction



↓
`fit1=lm(hours ~ rpm + tool,
data=my.dat)`

Interaction as allowed



↓
`fit2=lm(hours ~ rpm * tool,
data=my.dat)`

In case of interaction, the slope of the predictor “rpm” changes for different levels of the second predictor “tool”.

Do we get the same slope in rpm for tool A and tool B?
Is there an interaction between rpm and tool?



```
fit2 <- lm(hours ~ rpm * tool, data=my.dat)
```

```
> summary(fit2)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	32.774760	4.633472	7.073	2.63e-06	***
rpm	-0.020970	0.006074	-3.452	0.00328	**
toolB	23.970593	6.768973	3.541	0.00272	**
rpm:toolB	-0.011944	0.008842	-1.351	0.19553	

Residual standard error: 2.968 on 16 degrees of freedom
Multiple R-squared: 0.9105, Adjusted R-squared: 0.8937
F-statistic: 54.25 on 3 and 16 DF, p-value: 1.319e-08

$$\text{hour} = 32.8 + -0.02 \cdot \text{rpm} + 24 \cdot \text{toolB} - 0.01 \cdot (\text{rpm} \cdot \text{toolB})$$

The main effects are hard to interpret in case of interactions.

Here the **interactions seems not to be significant**. With ANOVA we can test for nested models if the more complex model leads to a significant improvement:

How to read a model with interaction?

$$\text{hour} = 32.8 - 0.02 \cdot \text{rpm} + 24 \cdot \text{toolB} - 0.01 \cdot (\text{rpm} \cdot \text{toolB})$$

toolB (toolB=1):

$$\text{hour} = 32.8 - 0.02 \cdot \text{rpm} + 24 \cdot 1 - 0.01 \cdot (\text{rpm} \cdot 1)$$

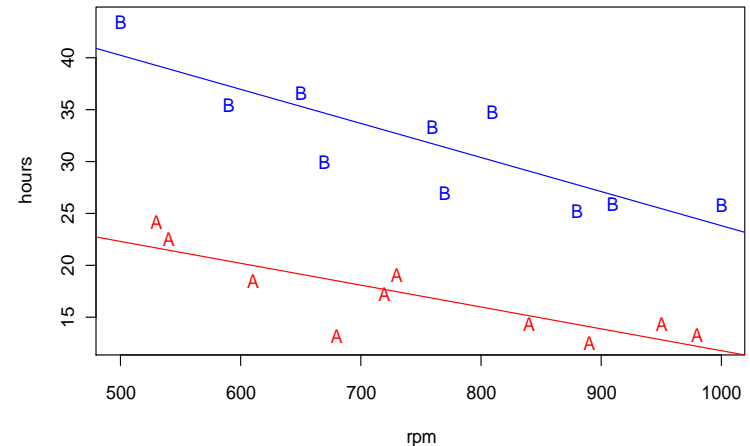
$$\text{hour} = 56.9 - 0.03 \cdot \text{rpm}$$

toolA (toolB=0):

$$\text{hour} = 32.8 - 0.02 \cdot \text{rpm} + 24 \cdot 0 - 0.01 \cdot (\text{rpm} \cdot 0)$$

$$\text{hour} = 32.8 - 0.02 \cdot \text{rpm}$$

Interaction is allowed

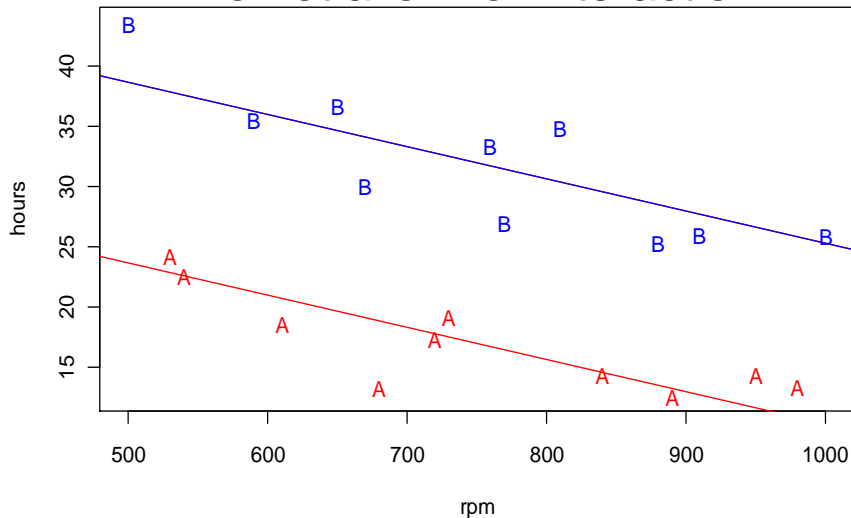


In case of interaction, the slope of the predictor “rpm” changes for different levels of the second predictor “tool” – also the intercept is changing for the two tools.

Remark: In case of interaction between two continuous predictors, slope (and intercept) of one predictor changes continuously with a continuous changing value of the other predictor and vice versa.

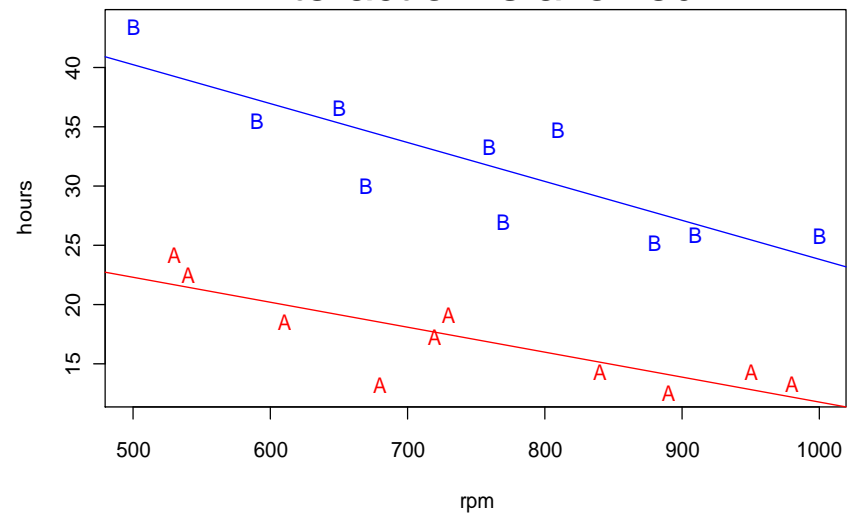
Do we need the complex model with the interaction?

Do not allow for interaction



**fit1=lm(hours ~ rpm + tool,
data=my.dat)**

Interaction is allowed



**fit2=lm(hours ~ rpm * tool,
data=my.dat)**

anova(fit2, fit1, test="F")

p>5%, therefore we have no evidence that an interaction is needed

How to decide between two nested models?

Use the F-test in case of linear models that are fitted with the same data.

This compares only the **in-sample** goodness of fit!

```
m.large=lm(Sepal.Length~., data=iris)
m.small=lm(Sepal.Length~Sepal.Width, data=iris)
anova(m.large, m.small, test="F")
# Analysis of Variance Table
#
# Model 1:
# Sepal.Length ~ Sepal.Width + Petal.Length + Petal.Width + Species
# Model 2:
# Sepal.Length ~ Sepal.Width
#
```

#		Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)					
#	1	144	13.556									
#	2	148	100.756	-4	-87.2	231.56	< 2.2e-16 ***					
#	---											
#	Signif. codes:	0	'***'	0.001	'**'	0.01	'*'	0.05	'.'	0.1	' '	1

Remark: nested means that the larger model contains all predictors of the smaller model plus some additional predictors.

Variable Selection

Goal: We want to **develop a simple model** by dropping all predictors from the regression model which are not necessary.

How: In a step-by-step manner, e.g. the least significant predictor is dropped from the model, as long as we have significant p-values.

In R:

```
> fit <- update(fit, . ~ . - colx)
> summary(fit)
```

→ See exercise

Warning: The p-values of the individual hypothesis tests are based on the assumption that the other predictors remain in the model and do not change. Therefore, you must not drop more than one single non-significant predictor at a time! Moreover, after variable selection the remaining coefficients and p-values are biased leading to an overestimation of effect size and significance.

Why coefficients estimates are not unbiased after model selection

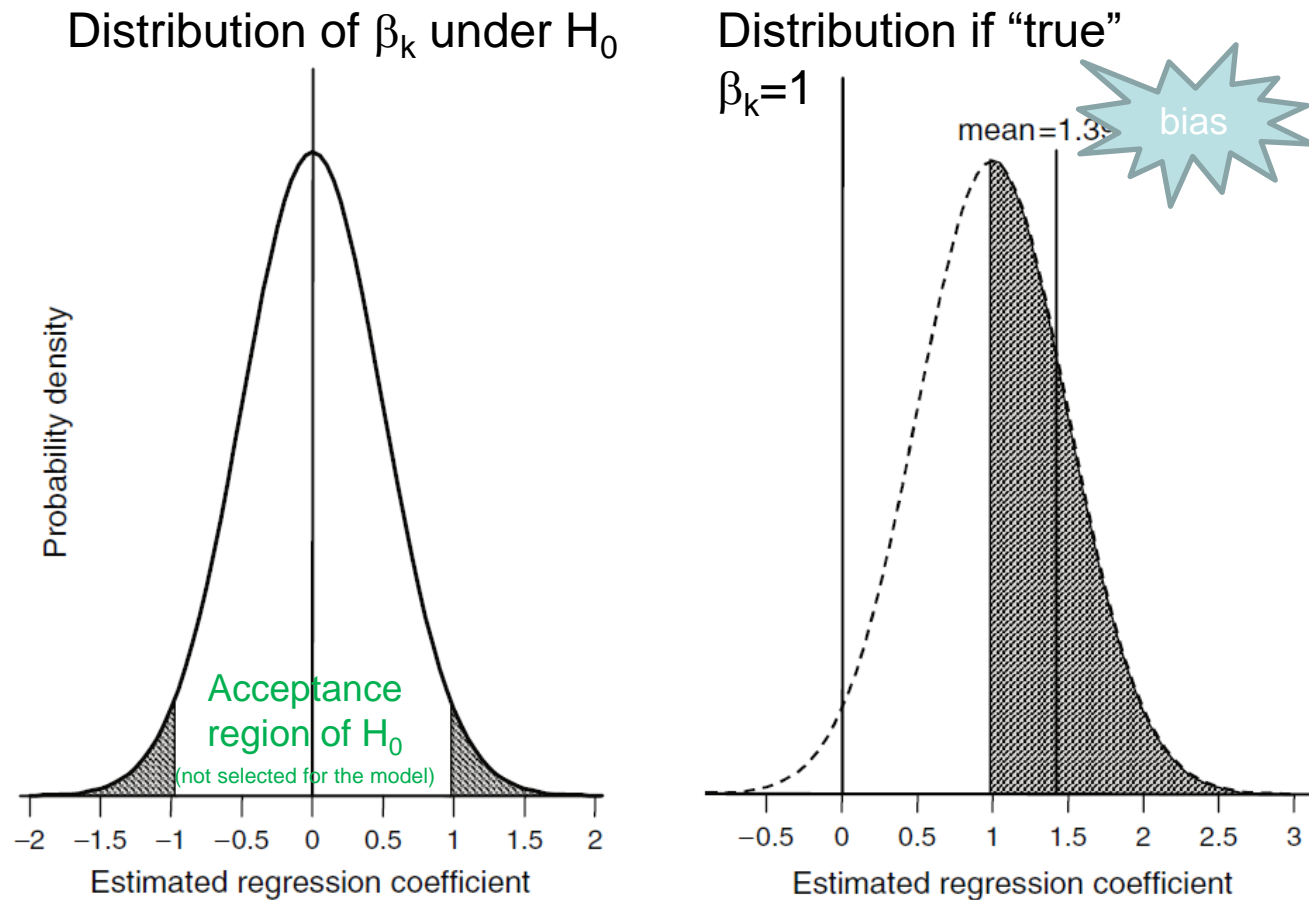


Fig. 5.5 Illustration of testimation bias. In case of a noise variable, the average of estimated regression coefficients is zero, and 2.5% of the coefficients is below -0.98 ($1.96 \times \text{SE of } 0.5$), and 2.5% of the coefficients is larger than $+0.98$ ($1.96 \times \text{SE of } 0.5$). In case of a true coefficient of 1, the estimated coefficients are statistically significant in 52%. For these cases, the average of estimated coefficients is 1.39 instead of 1

Steps in linear modelling

0) Preprocessing

- learning the meaning of all variables, check for correlations
- give short and informative names
- check for impossible values, errors
- if they exist (missing, error): set them to NA
- consider imputation methods, but be careful

1) First-aid transformations

- bring all variables to a suitable scale (use also field knowledge)
- routinely apply the first-aid transformations

2) Find a good model

- start with a model including important confounders
- perform a residual analysis
- improve model by transformations or adding better predictors
- reduce step by step complexity (be aware of introduced biases)
- use your specific knowledge to choose between variables

Limits of linear Regression

If your **residuals do not follow a Normal distribution** (even after transformations) use generalized linear modeling (glm – e.g. logistic regression)

If your **predictors show a strong correlation** use shrinkage methods (e.g. lasso)

If your **data are not independent** use mixed models or methods for time-series.

If you **do not have a linear relation**, use non-linear regression (e.g. nlm) or generalized additive models (e.g. gam) or tree models

Summary

- Minimizing squared errors yields an unbiased model describing the train data.
- The 95%-CI of the modeled y gives a range so that we only have a risk of 5% that the true expected value of y is not within this range
- The 95% prediction interval of the modeled y gives a range so that 95% of the fitted data is within this range.
- For a model with p parameters fitted to n observations the 95%CI of each parameter is given by $\hat{b} \pm q^{t_{n-p}}_{\alpha/2} \cdot se(\hat{b})$ which is significantly different from all values not within this CI.
- The larger n the smaller the CI and the more significant parameters we get.
- The parameter size indicates the change of y if the respective variable is increased by one unit and all other variables kept constant (often not possible)
- The size of the parameter and the scale of the variable determine the relevance of a significant parameter
- The size and the p-value of a parameter in a multiple regression model can change if other variables are added to or deleted from the model