

Biostatistik, WS 2010/2011

Wiederholung

Matthias Birkner

<http://www.mathematik.uni-mainz.de/~birkner/Biostatistik1011/>

4.2. 2011

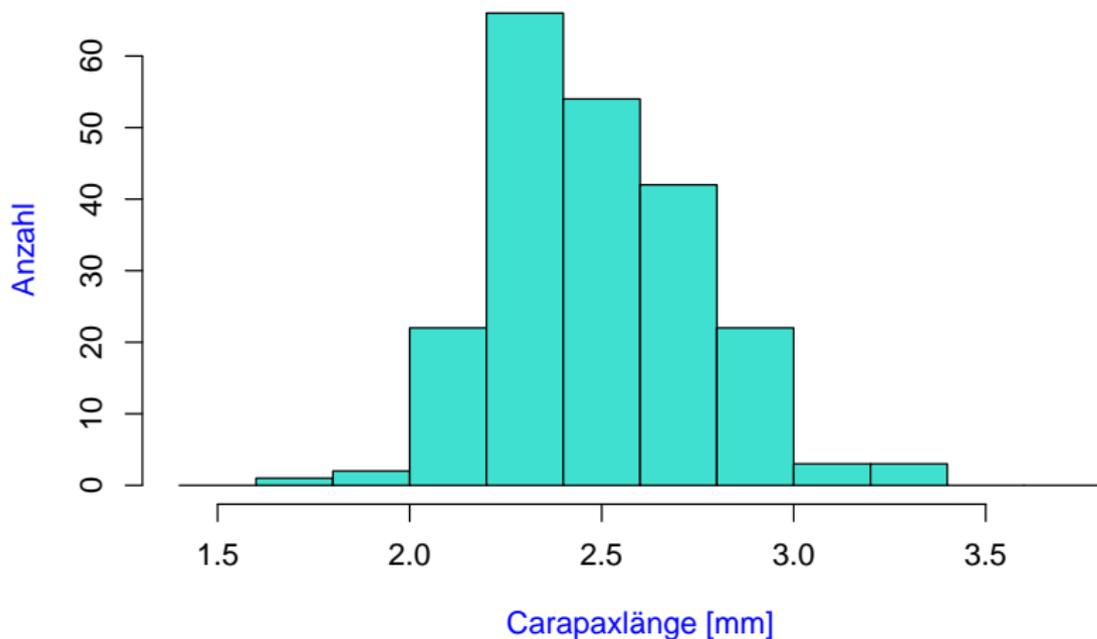


JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

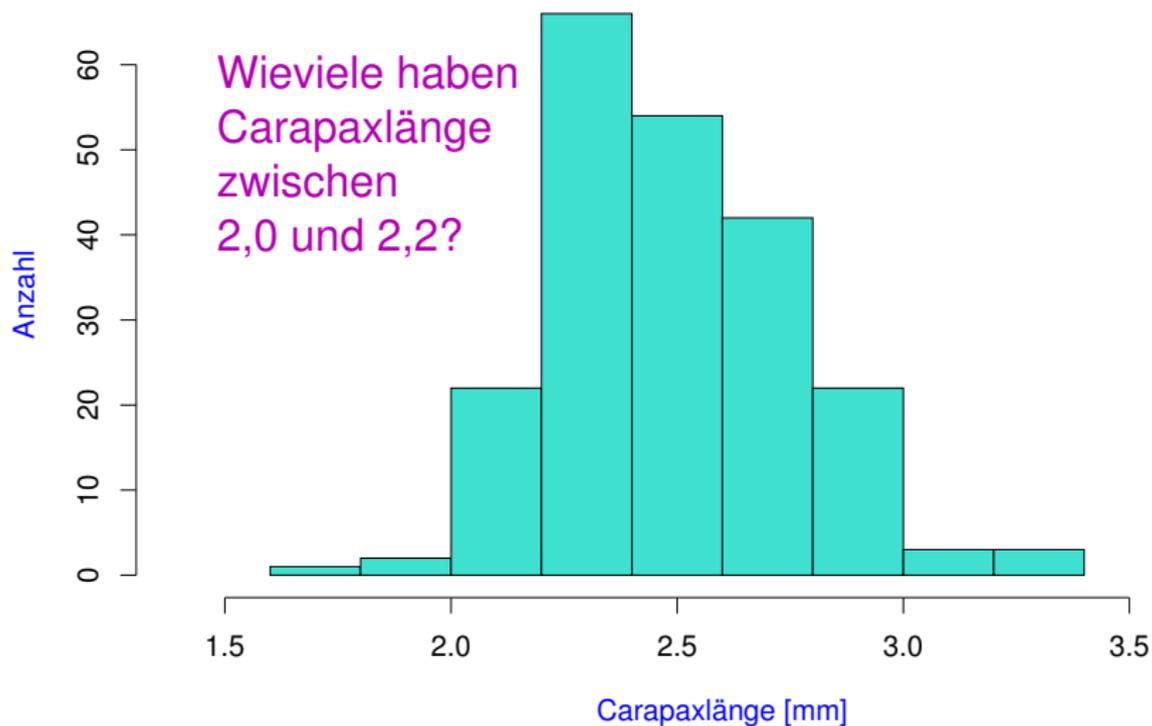
Inhalt

- 1 Deskriptive Statistik
- 2 Standardfehler und t-Tests
- 3 Chi-Quadrat-Tests
 - χ^2 -Test für eine feste Verteilung
 - χ^2 -Test auf Unabhängigkeit (oder Homogenität)
- 4 Konfidenzintervalle
- 5 Lineare Regression
 - Lineare Zusammenhänge
 - t -Test fuer lineare Zusammenhänge

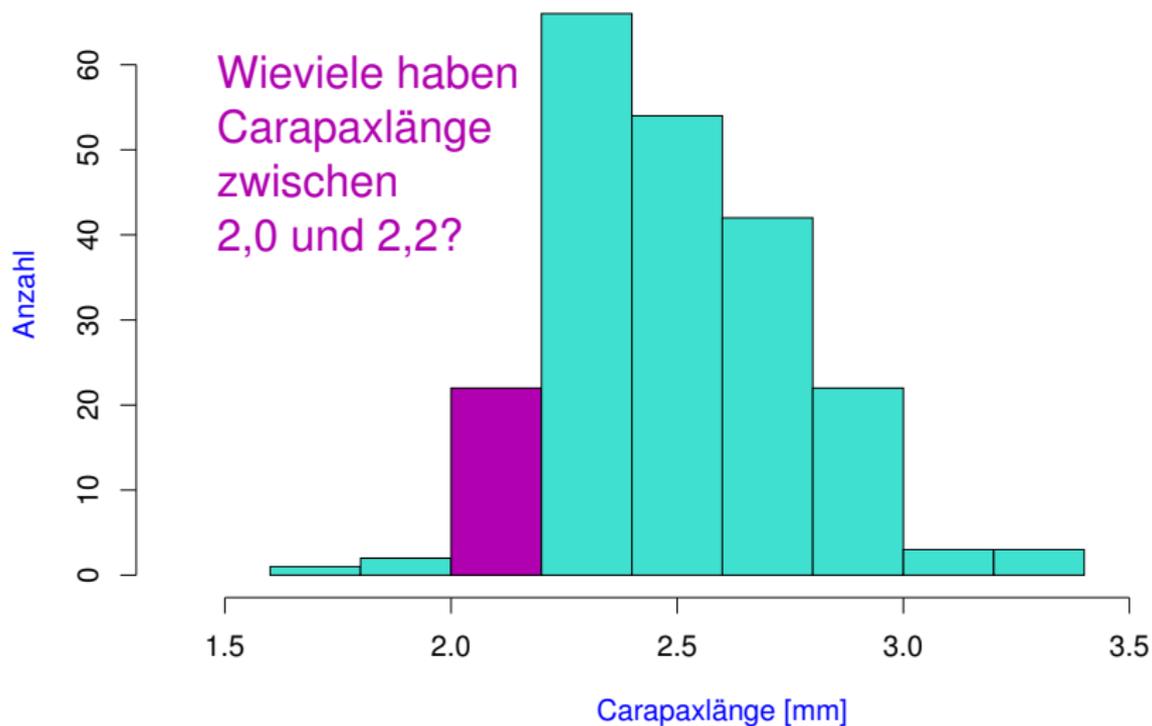
Nichteiertragende Weibchen am 6. Sept. '88, n=215



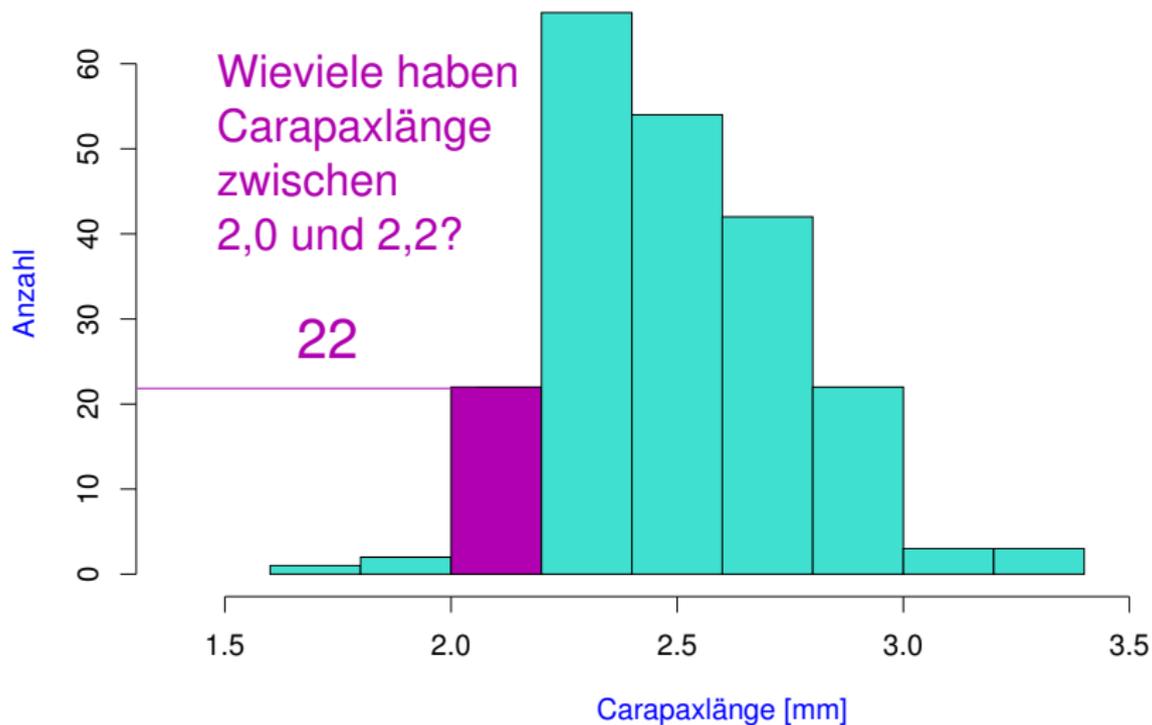
Nichteiertragende Weibchen am 6. Sept. '88, n=215



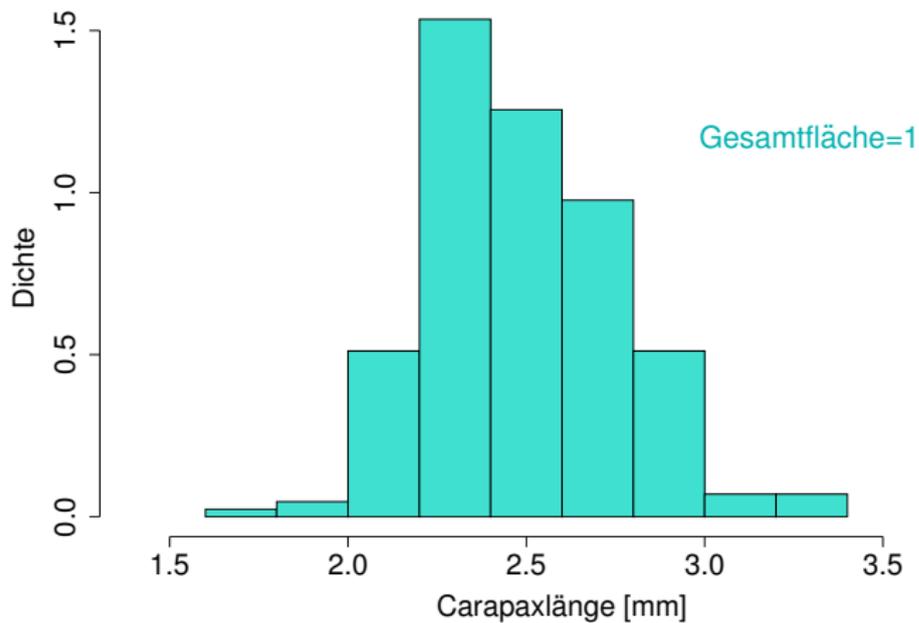
Nichteiertragende Weibchen am 6. Sept. '88, n=215



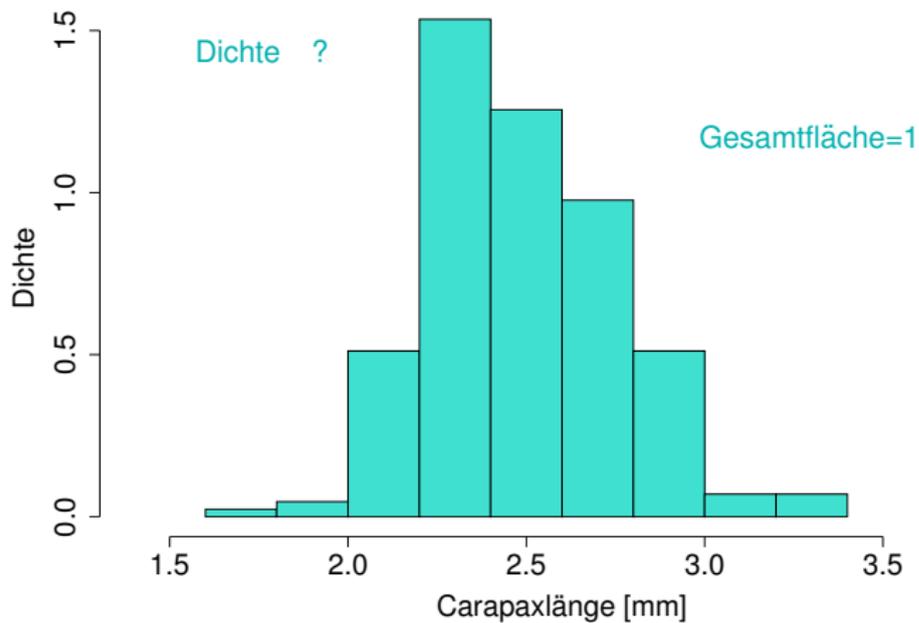
Nichteiertragende Weibchen am 6. Sept. '88, n=215



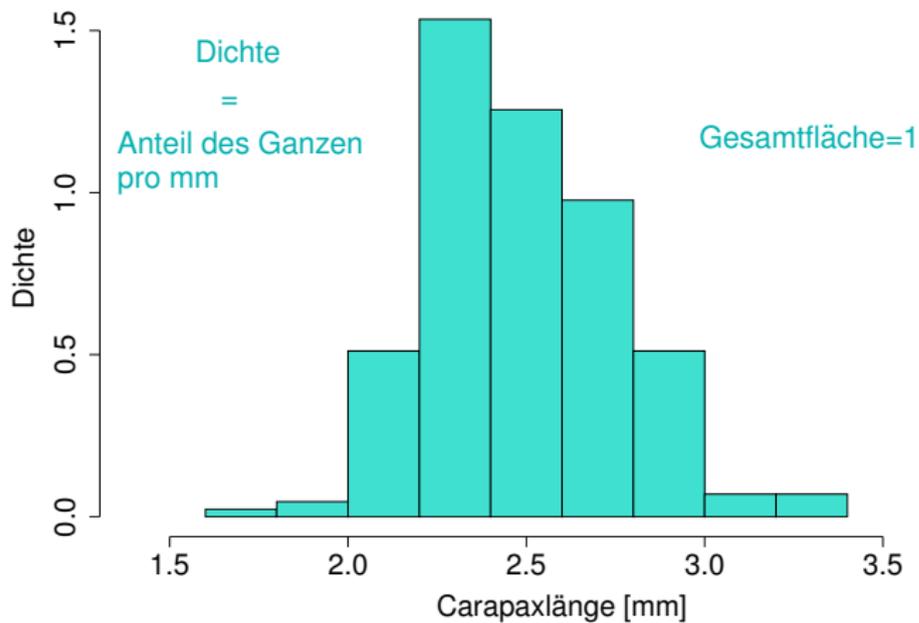
Nichteiertragende Weibchen am 6. Sept. '88, n=215



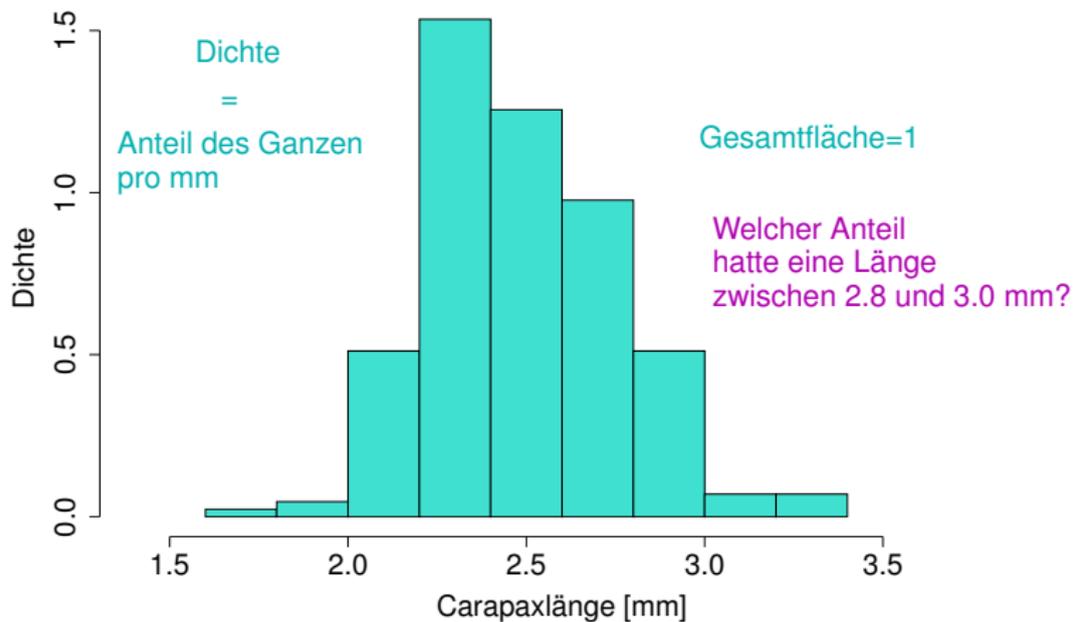
Nichteiertragende Weibchen am 6. Sept. '88, n=215



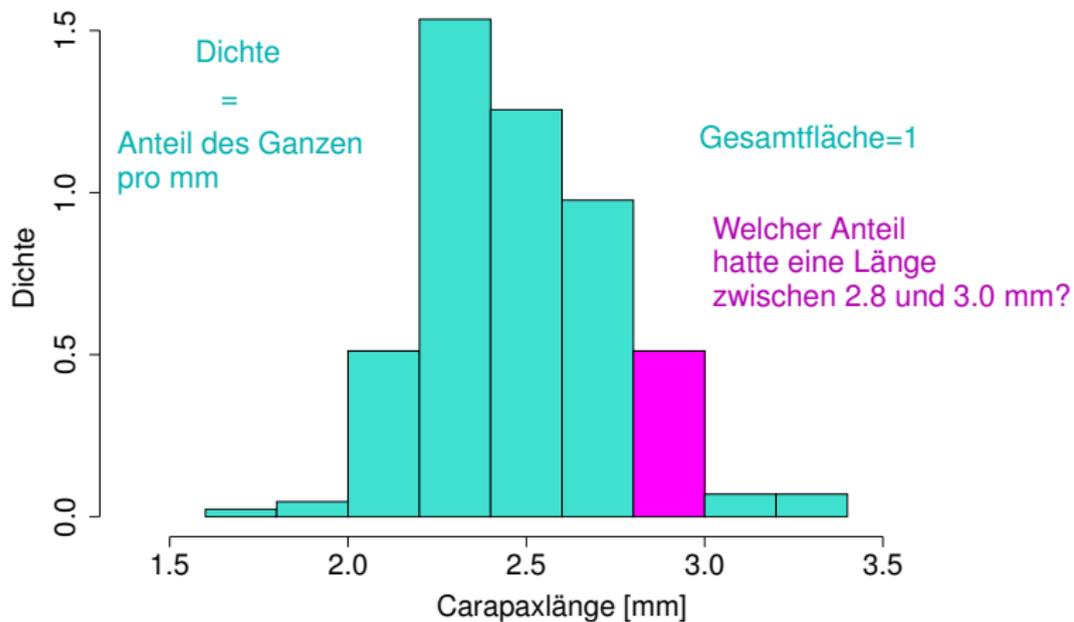
Nichteiertragende Weibchen am 6. Sept. '88, n=215



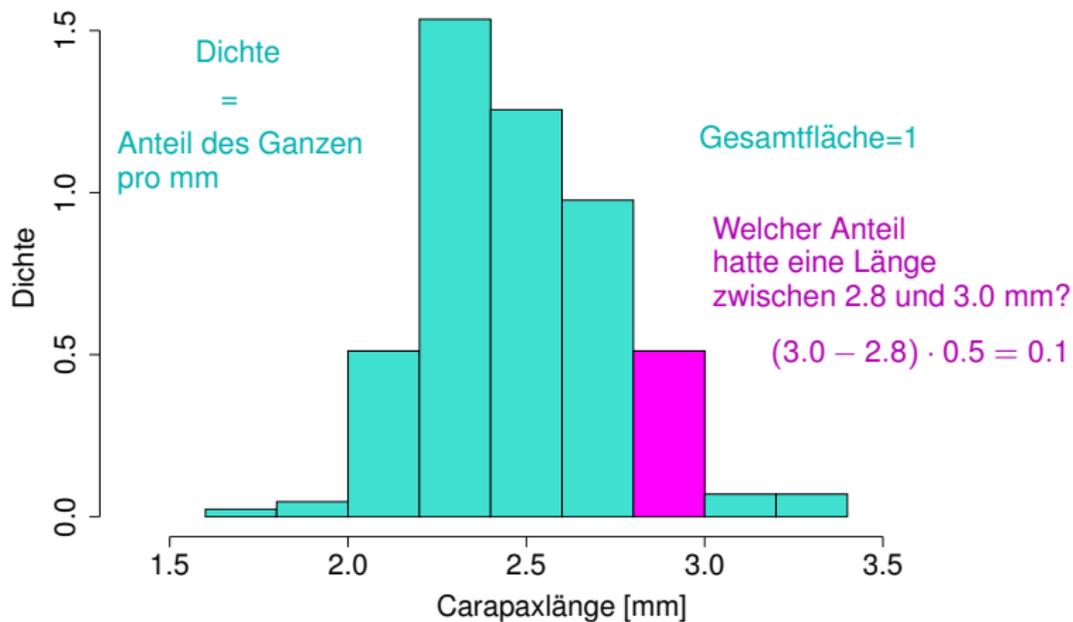
Nichteiertragende Weibchen am 6. Sept. '88, n=215



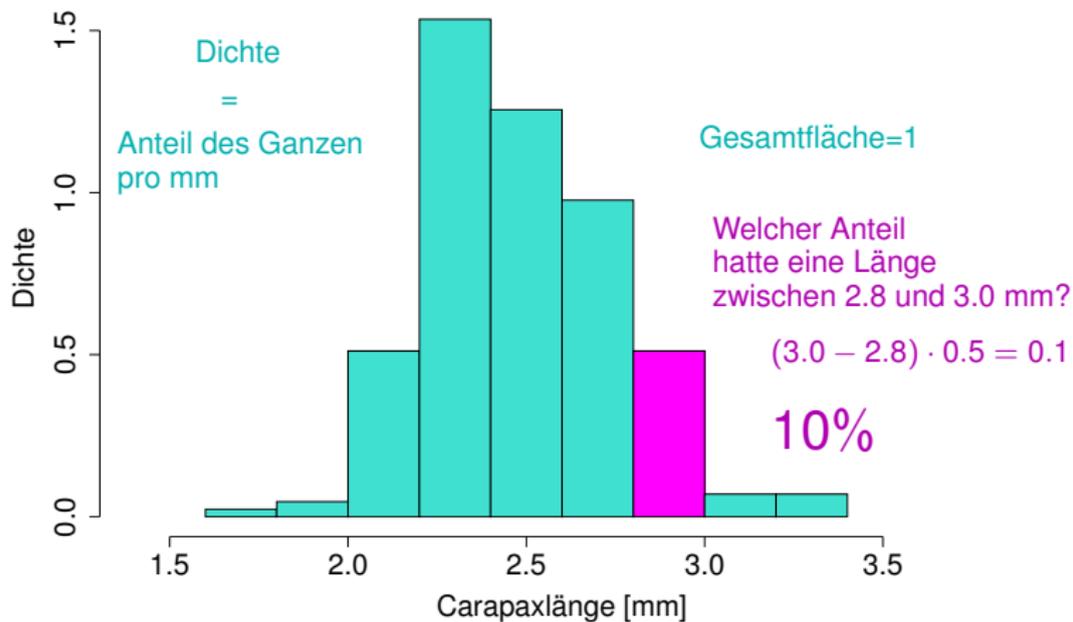
Nichteiertragende Weibchen am 6. Sept. '88, n=215



Nichteiertragende Weibchen am 6. Sept. '88, n=215

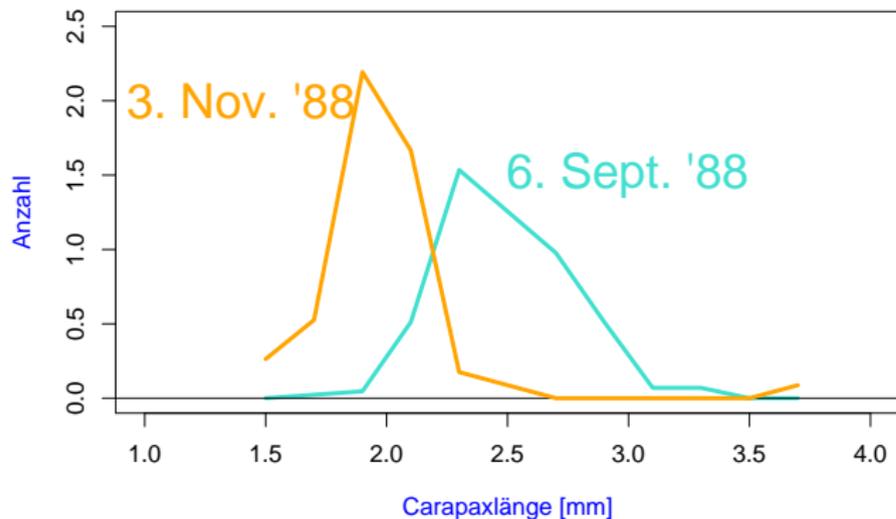


Nichteiertragende Weibchen am 6. Sept. '88, n=215

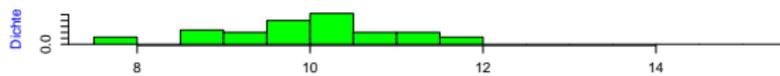
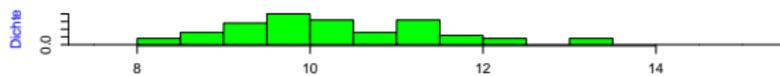
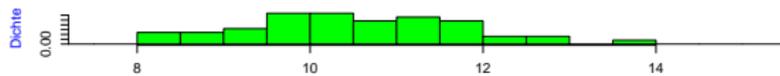
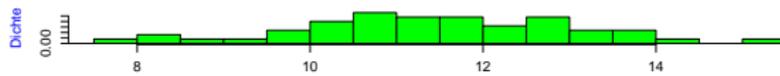


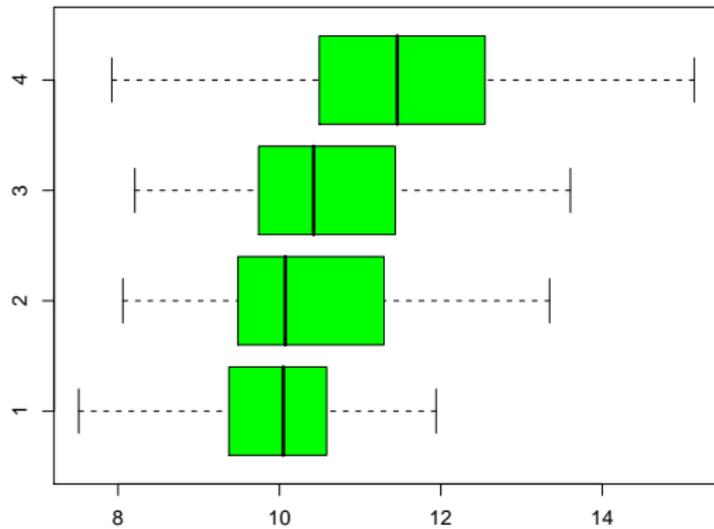
Zwei und mehr Dichtepolygone in einem Plot

Nichteiertragende Weibchen



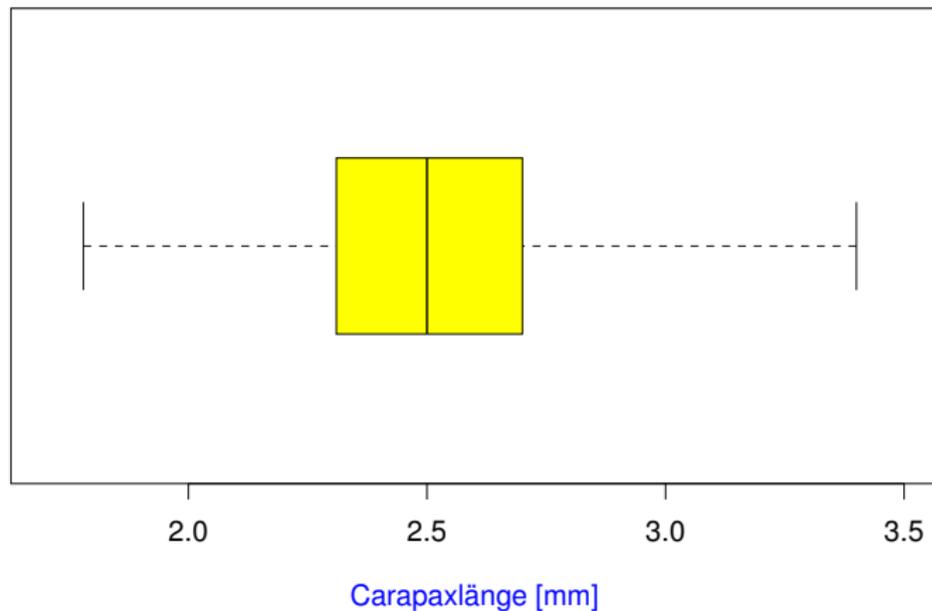
Beispiel: Vergleich von mehreren Gruppen





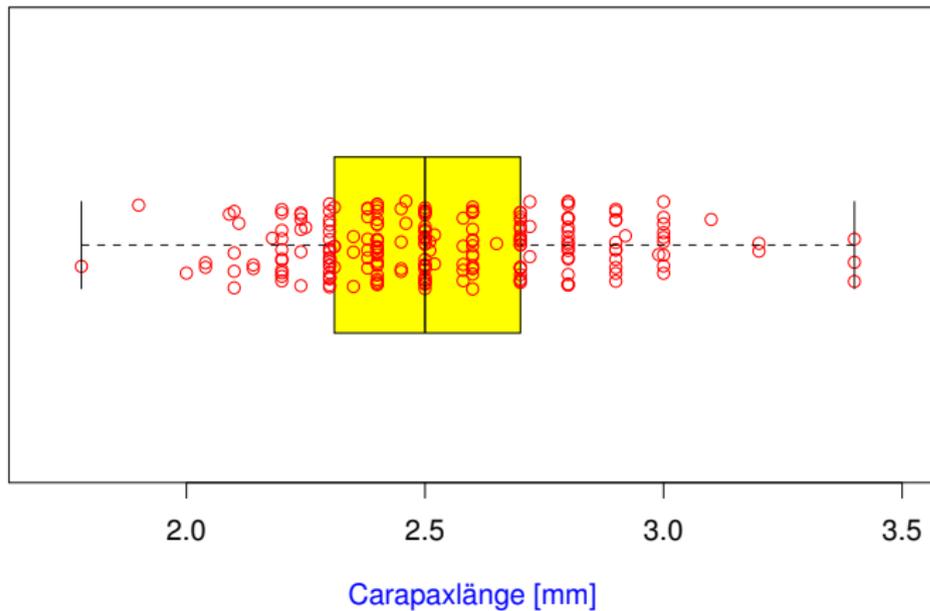
Der Boxplot

Boxplot, einfache Ausführung



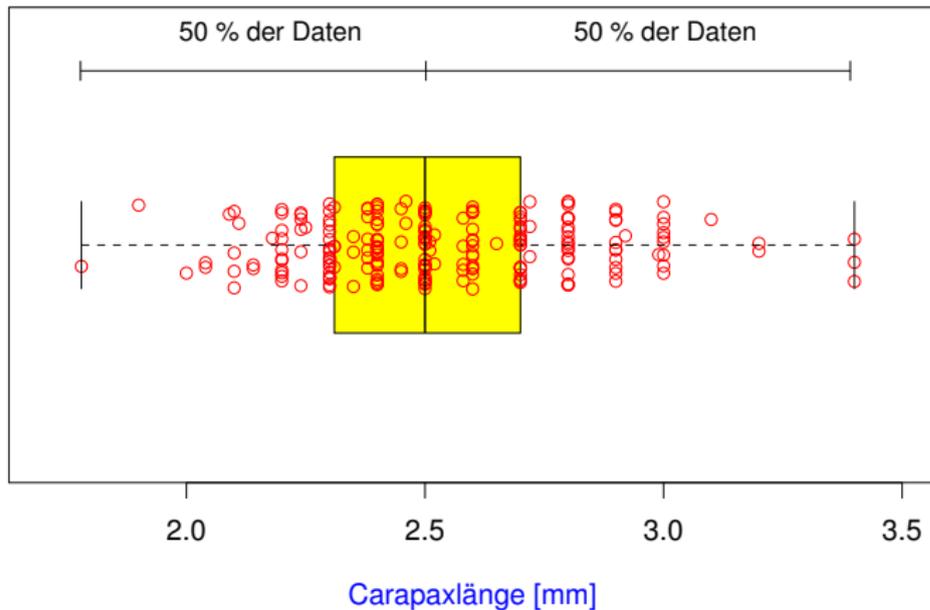
Der Boxplot

Boxplot, einfache Ausführung



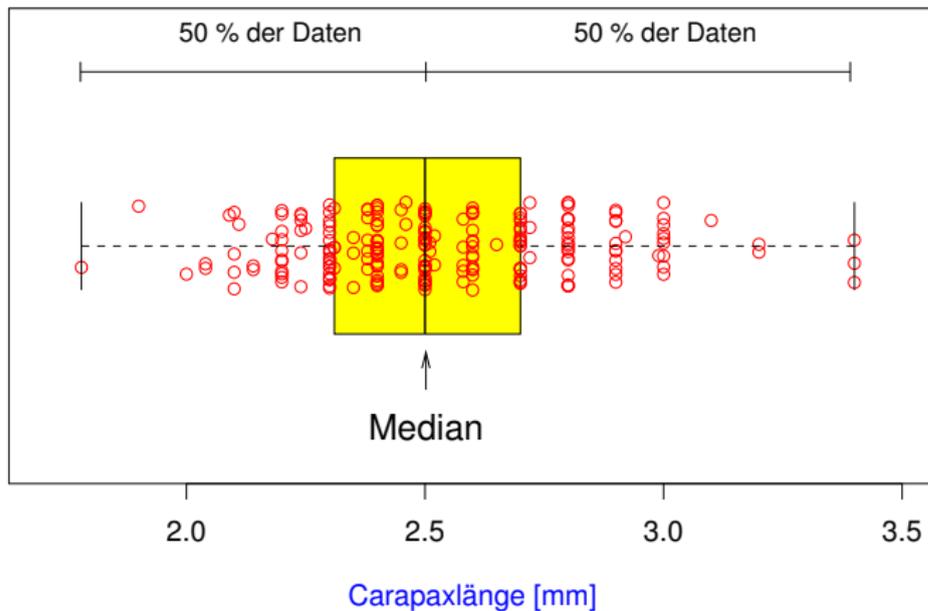
Der Boxplot

Boxplot, einfache Ausführung



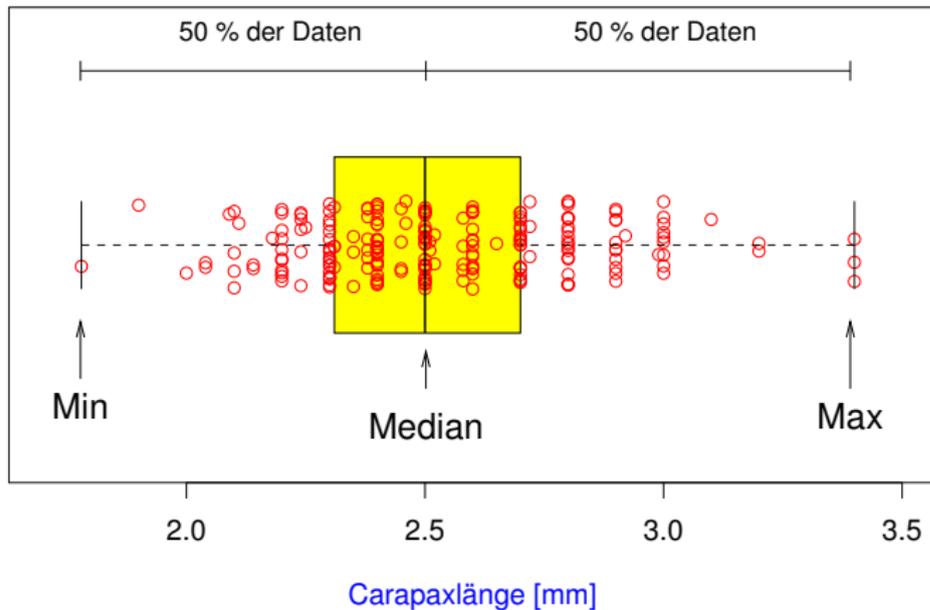
Der Boxplot

Boxplot, einfache Ausführung



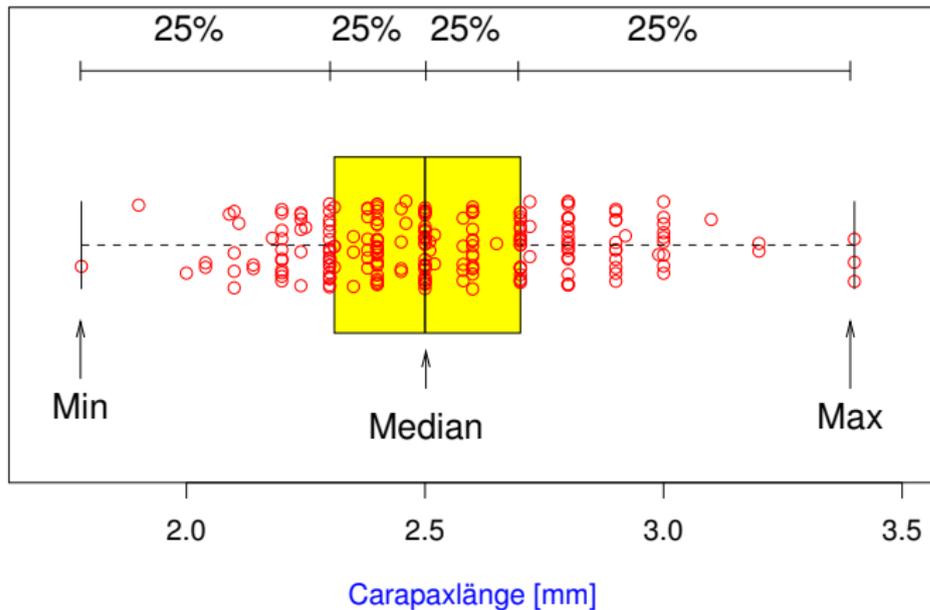
Der Boxplot

Boxplot, einfache Ausführung



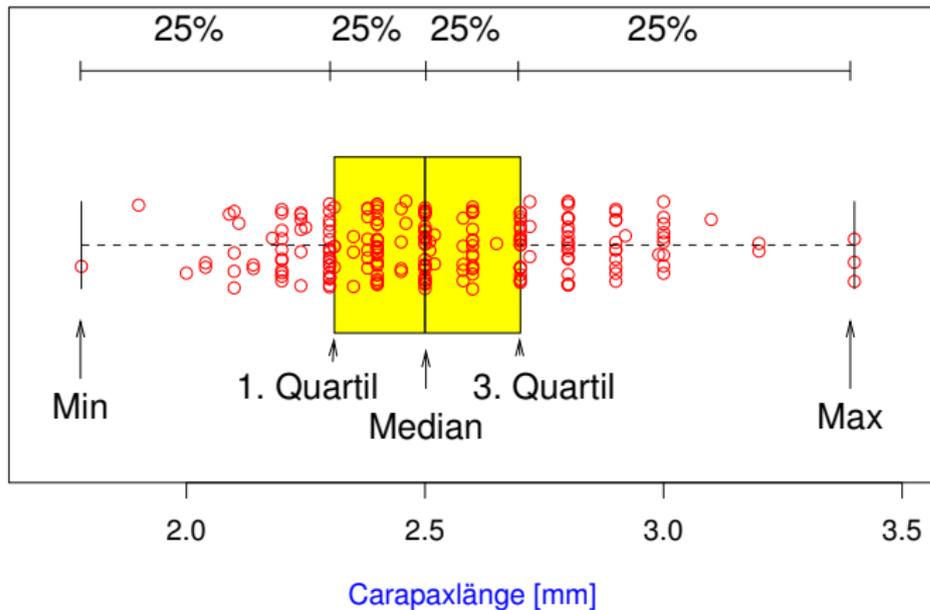
Der Boxplot

Boxplot, einfache Ausführung



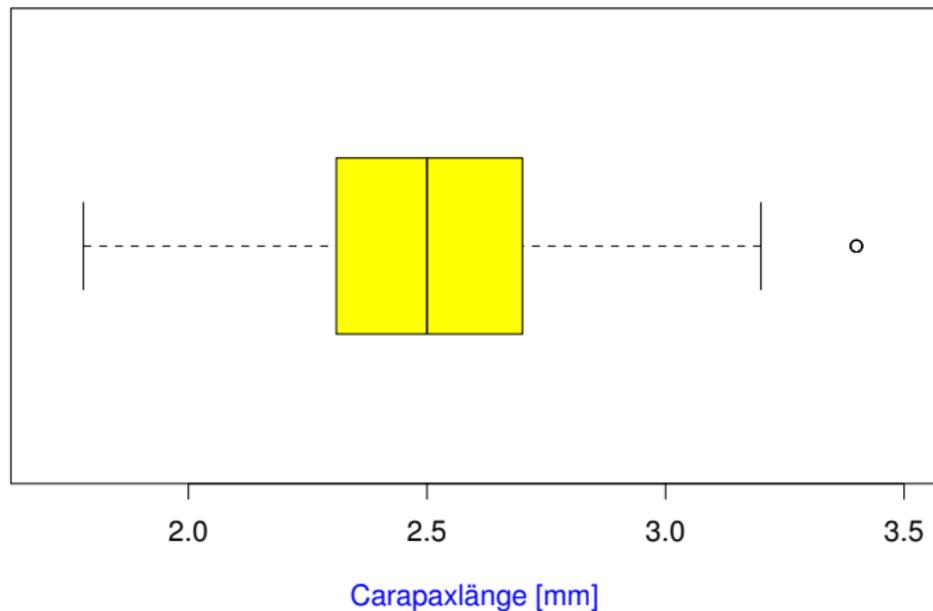
Der Boxplot

Boxplot, einfache Ausführung



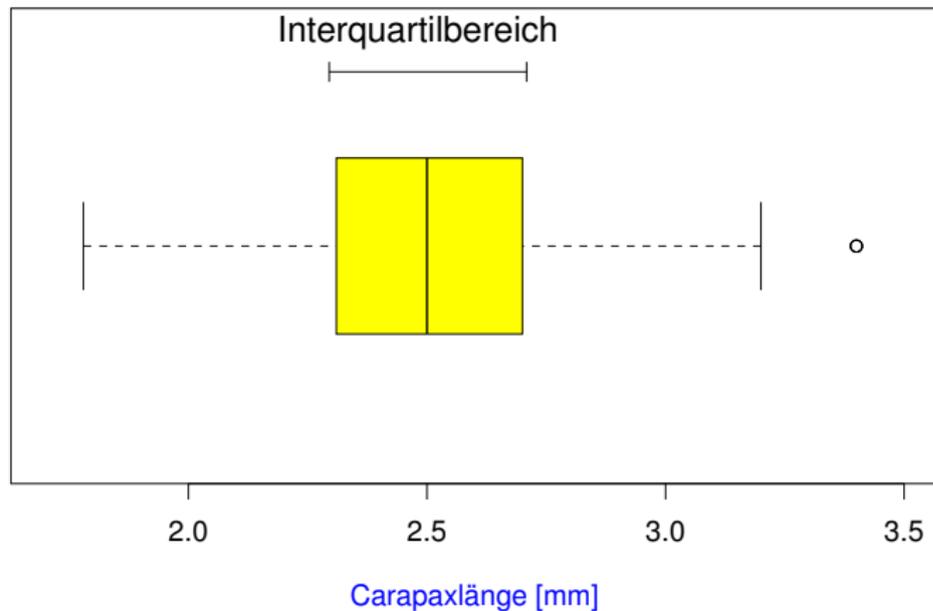
Der Boxplot

Boxplot, Standardausführung



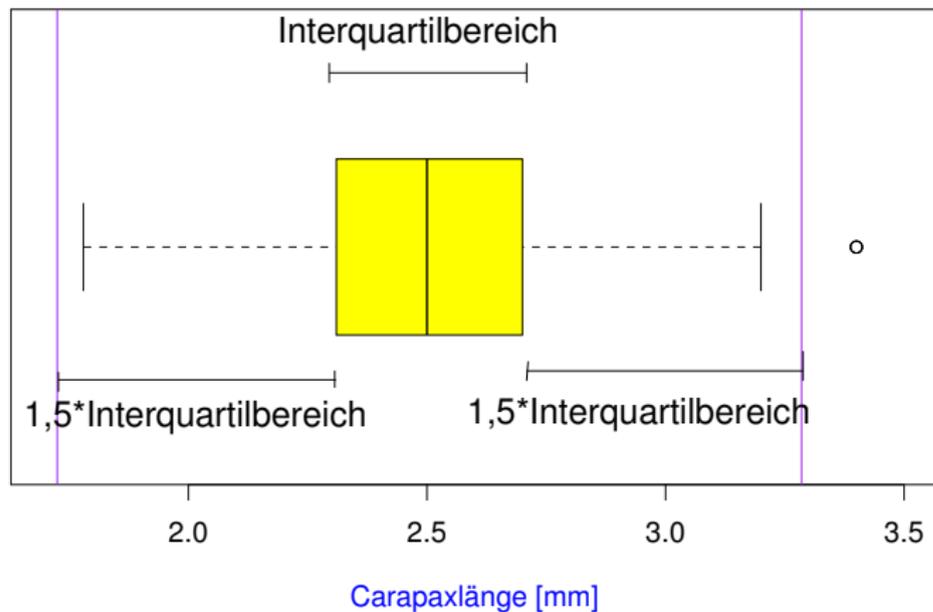
Der Boxplot

Boxplot, Standardausführung



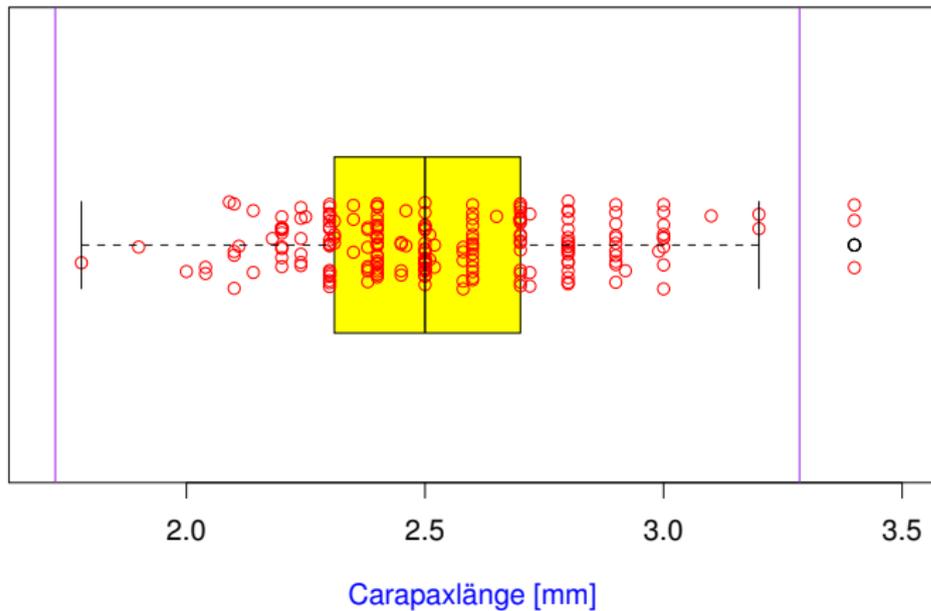
Der Boxplot

Boxplot, Standardausführung



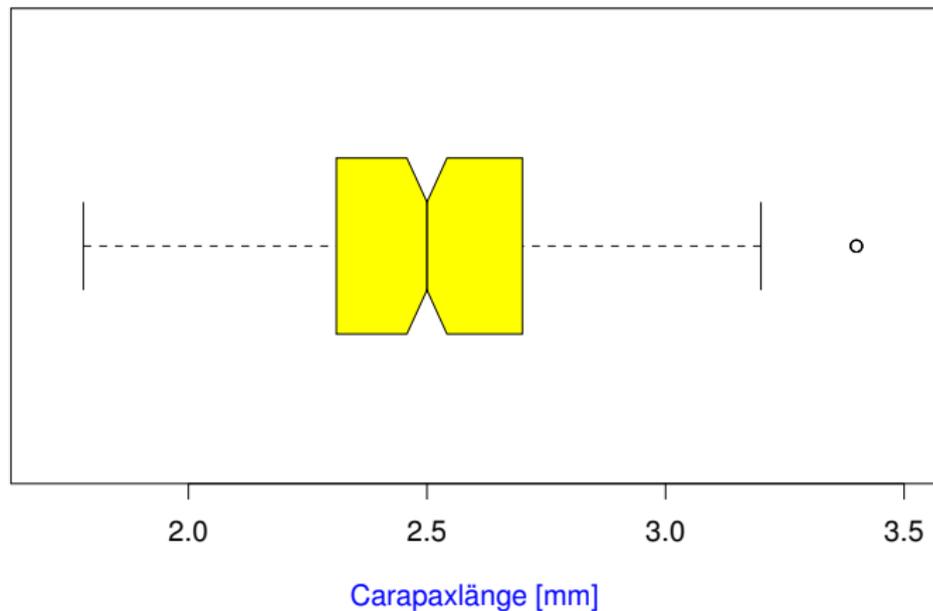
Der Boxplot

Boxplot, Standardausführung



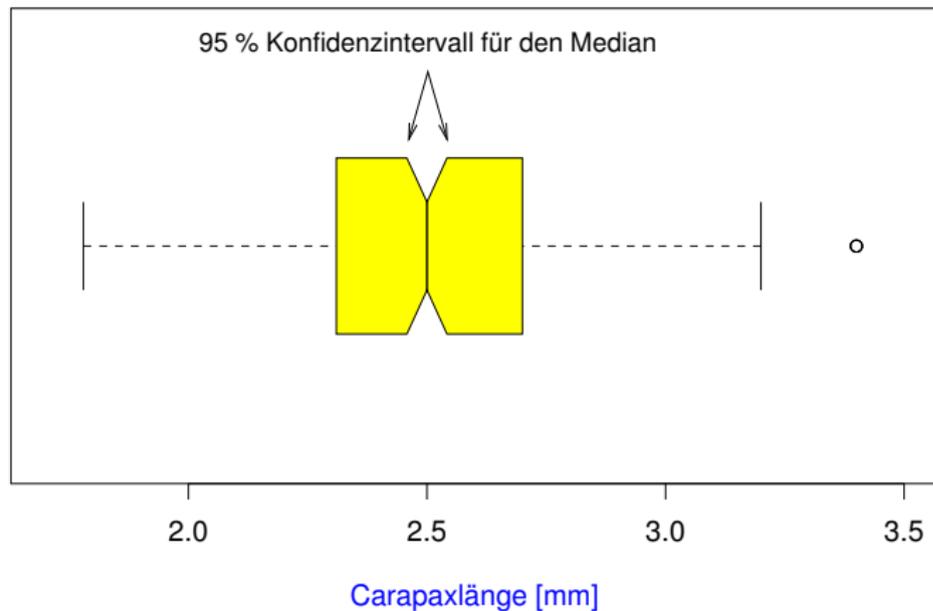
Der Boxplot

Boxplot, Profiausstattung



Der Boxplot

Boxplot, Profiausstattung



Es ist oft möglich,
das Wesentliche
an einer Stichprobe

mit ein paar Zahlen
zusammenzufassen.

Wesentlich:

1. Wie groß?

2. Wie variabel?

Wesentlich:

1. Wie groß?

Lageparameter

2. Wie variabel?

Wesentlich:

1. Wie groß?

Lageparameter

2. Wie variabel?

Streuungsparameter

Eine Möglichkeit
kennen wir schon
aus dem Boxplot:

Lageparameter

Der Median

Lageparameter

Der Median

Streuungsparameter

Lageparameter

Der Median

Streuungsparameter

Der Quartilabstand ($Q_3 - Q_1$)

Der Median:

die Hälfte der Beobachtungen sind kleiner,
die Hälfte sind größer.

Der **Median**:

die Hälfte der Beobachtungen sind kleiner,
die Hälfte sind größer.

Der Median ist
das **50%-Quantil**
der Daten.

Die Quartile

Das erste Quartil, Q_1 :

Die Quartile

Das erste Quartil, Q_1 :
ein Viertel der Beobachtungen
sind kleiner,
drei Viertel sind größer.

Die Quartile

Das erste Quartil, Q_1 :
ein Viertel der Beobachtungen
sind kleiner,
drei Viertel sind größer.

Q_1 ist das
25%-Quantil
der Daten.

Die Quartile

Das dritte Quartil, Q_3 :

Die Quartile

Das dritte Quartil, Q_3 :
drei Viertel der Beobachtungen
sind kleiner,
ein Viertel sind größer.

Die Quartile

Das dritte Quartil, Q_3 :
drei Viertel der Beobachtungen
sind kleiner,
ein Viertel sind größer.

Q_3 ist das
75%-Quantil
der Daten.

Am häufigsten werden benutzt:

Lageparameter

Der Mittelwert \bar{x}

Am häufigsten werden benutzt:

Lageparameter

Der Mittelwert \bar{x}

Streuungsparameter

Die Standardabweichung s

Der Mittelwert

(engl. *mean*)

NOTATION:

Wenn die Beobachtungen

$x_1, x_2, x_3, \dots, x_n$

heißen,

schreibt man oft

\bar{x}

für den Mittelwert.

DEFINITION:

Mittelwert

=

$$\frac{\text{Summe der Messwerte}}{\text{Anzahl der Messwerte}}$$

DEFINITION:

Mittelwert

=

$$\frac{\text{Summe}}{\text{Anzahl}}$$

DEFINITION:

Der Mittelwert von x_1, x_2, \dots, x_n als Formel:

DEFINITION:

Der Mittelwert von x_1, x_2, \dots, x_n als Formel:

$$\bar{x} = (x_1 + x_2 + \dots + x_n)/n$$

DEFINITION:

Der Mittelwert von x_1, x_2, \dots, x_n als Formel:

$$\begin{aligned}\bar{x} &= (x_1 + x_2 + \dots + x_n)/n \\ &= \frac{1}{n} \sum_{i=1}^n x_i\end{aligned}$$

Geometrische Bedeutung des Mittelwerts: Der Schwerpunkt

Die Standardabweichung

Die Standardabweichung

Wie weit weicht
eine typische Beobachtung
vom
Mittelwert
ab ?

Die **Standardabweichung** σ (“sigma”)
ist ein
etwas komisches
gewichtetes Mittel
der Abweichungsbeträge

Die **Standardabweichung** σ (“sigma”)
ist ein

etwas komisches

gewichtetes Mittel
der Abweichungsbeträge
und zwar

$$\sigma = \sqrt{\text{Summe}(\text{Abweichungen}^2)/n}$$

Die **Standardabweichung** von x_1, x_2, \dots, x_n
als Formel:

Die **Standardabweichung** von x_1, x_2, \dots, x_n
als Formel:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

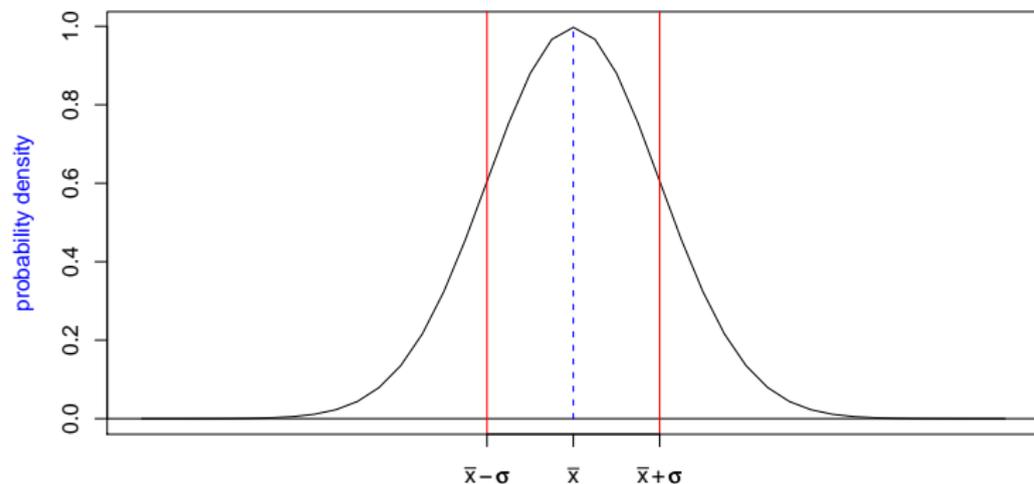
Die **Standardabweichung** von x_1, x_2, \dots, x_n
als Formel:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

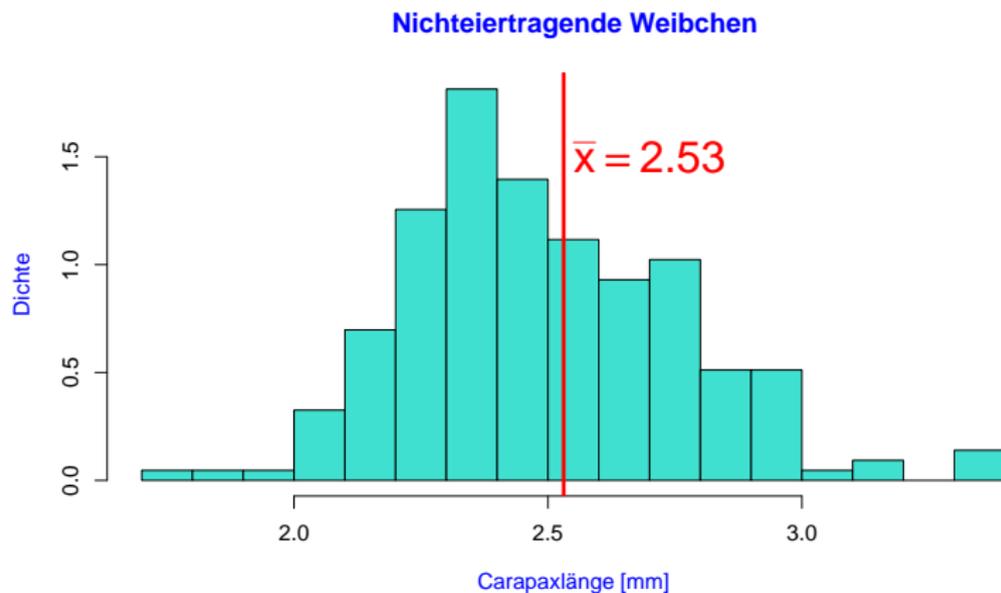
$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ heißt **Varianz**.

Faustregel für die Standardabweichung

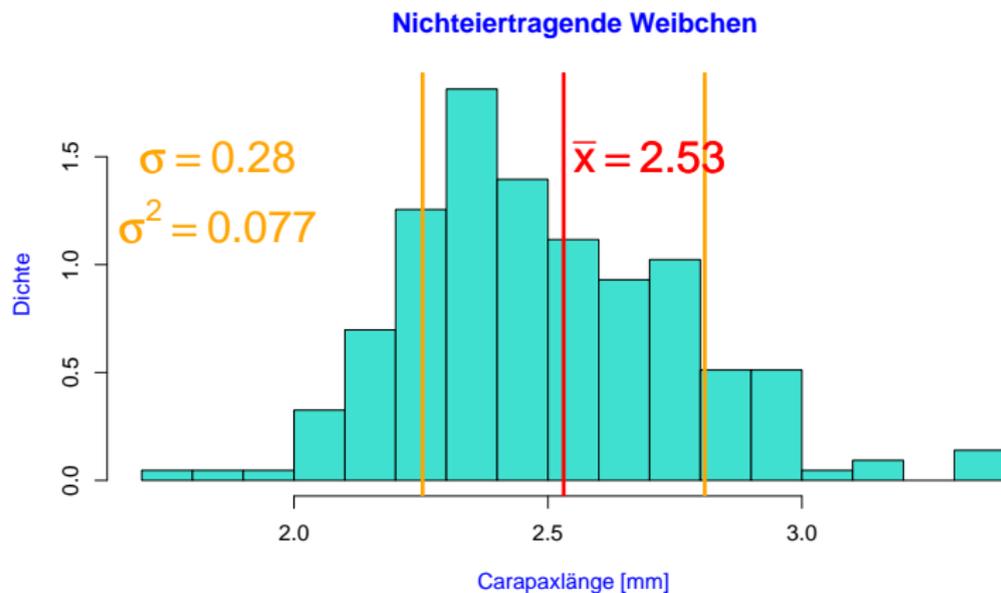
Bei ungefähr glockenförmigen (also eingipfligen und symmetrischen) Verteilungen liegen ca. 2/3 der Verteilung zwischen $\bar{x} - \sigma$ und $\bar{x} + \sigma$.



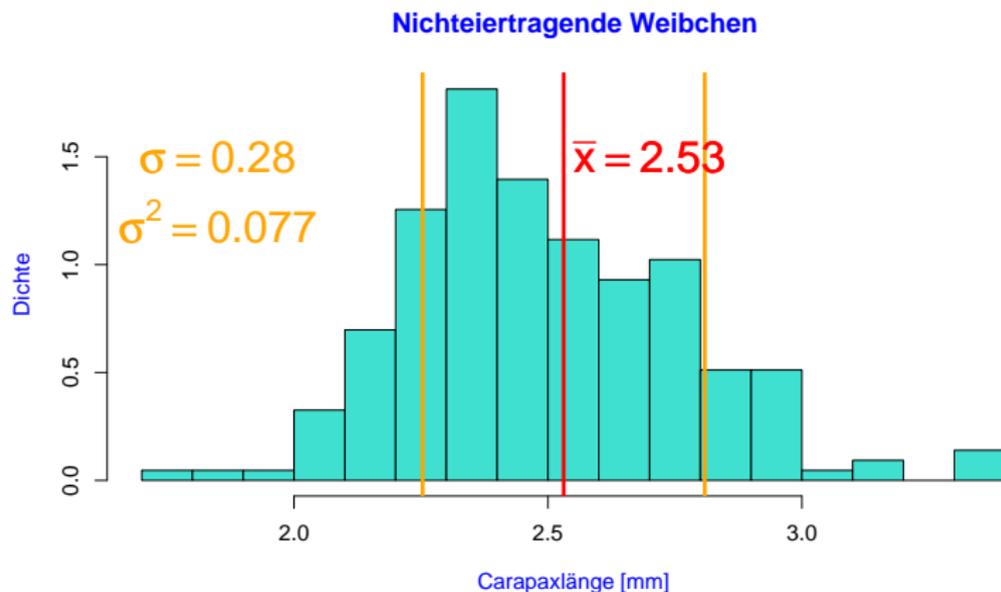
Standardabweichung der Carapaxlängen nichteiertragender Weibchen vom 6.9.88



Standardabweichung der Carapaxlängen nichteiertragender Weibchen vom 6.9.88



Standardabweichung der Carapaxlängen nichteiertragender Weibchen vom 6.9.88



Hier liegt der Anteil zwischen $\bar{x} - \sigma$ und $\bar{x} + \sigma$ bei 72%.

Varianz der Carapaxlängen nichtteiertragender Weibchen vom 6.9.88

Alle Carapaxlängen im Meer: $\mathcal{X} = (X_1, X_2, \dots, X_N)$.

Varianz der Carapaxlängen nichtteiertragender Weibchen vom 6.9.88

Alle Carapaxlängen im Meer: $\mathcal{X} = (X_1, X_2, \dots, X_N)$.

Carapaxlängen in unserer Stichprobe: $\mathcal{S} = (S_1, S_2, \dots, S_{n=215})$

Varianz der Carapaxlängen nichteiertragender Weibchen vom 6.9.88

Alle Carapaxlängen im Meer: $\mathcal{X} = (X_1, X_2, \dots, X_N)$.

Carapaxlängen in unserer Stichprobe: $\mathcal{S} = (S_1, S_2, \dots, S_{n=215})$

Stichprobenvarianz:

$$\sigma_S^2 = \frac{1}{n} \sum_{i=1}^{215} (S_i - \bar{S})^2 \approx 0,0768$$

Varianz der Carapaxlängen nichtteiertragender Weibchen vom 6.9.88

Alle Carapaxlängen im Meer: $\mathcal{X} = (X_1, X_2, \dots, X_N)$.

Carapaxlängen in unserer Stichprobe: $\mathcal{S} = (S_1, S_2, \dots, S_{n=215})$

Stichprobenvarianz:

$$\sigma_S^2 = \frac{1}{n} \sum_{i=1}^{215} (S_i - \bar{S})^2 \approx 0,0768$$

Können wir 0,0768 als Schätzwert für die Varianz σ_X^2 in der ganzen Population verwenden?

Varianz der Carapaxlängen nichtteiertragender Weibchen vom 6.9.88

Alle Carapaxlängen im Meer: $\mathcal{X} = (X_1, X_2, \dots, X_N)$.

Carapaxlängen in unserer Stichprobe: $\mathcal{S} = (S_1, S_2, \dots, S_{n=215})$

Stichprobenvarianz:

$$\sigma_S^2 = \frac{1}{n} \sum_{i=1}^{215} (S_i - \bar{S})^2 \approx 0,0768$$

Können wir 0,0768 als Schätzwert für die Varianz σ_X^2 in der ganzen Population verwenden?

Ja, können wir machen.

Varianz der Carapaxlängen nichtteiertragender Weibchen vom 6.9.88

Alle Carapaxlängen im Meer: $\mathcal{X} = (X_1, X_2, \dots, X_N)$.

Carapaxlängen in unserer Stichprobe: $\mathcal{S} = (S_1, S_2, \dots, S_{n=215})$

Stichprobenvarianz:

$$\sigma_S^2 = \frac{1}{n} \sum_{i=1}^{215} (S_i - \bar{S})^2 \approx 0,0768$$

Können wir 0,0768 als Schätzwert für die Varianz $\sigma_{\mathcal{X}}^2$ in der ganzen Population verwenden?

Ja, können wir machen. Allerdings ist σ_S^2 im Durchschnitt um den Faktor $\frac{n-1}{n}$ ($= 214/215 \approx 0,995$) kleiner als $\sigma_{\mathcal{X}}^2$.

Varianzbegriffe

Varianz in der Population: $\sigma_X^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2$

Stichprobenvarianz: $\sigma_S^2 = \frac{1}{n} \sum_{i=1}^n (S_i - \bar{S})^2$

Varianzbegriffe

Varianz in der Population: $\sigma_X^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2$

Stichprobenvarianz: $\sigma_S^2 = \frac{1}{n} \sum_{i=1}^n (S_i - \bar{S})^2$

korrigierte Stichprobenvarianz:

$$s^2 = \frac{n}{n-1} \sigma_S^2$$

Varianzbegriffe

Varianz in der Population: $\sigma_X^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2$

Stichprobenvarianz: $\sigma_S^2 = \frac{1}{n} \sum_{i=1}^n (S_i - \bar{S})^2$

korrigierte Stichprobenvarianz:

$$\begin{aligned} s^2 &= \frac{n}{n-1} \sigma_S^2 \\ &= \frac{n}{n-1} \cdot \frac{1}{n} \cdot \sum_{i=1}^n (S_i - \bar{S})^2 \end{aligned}$$

Varianzbegriffe

Varianz in der Population: $\sigma_X^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2$

Stichprobenvarianz: $\sigma_S^2 = \frac{1}{n} \sum_{i=1}^n (S_i - \bar{S})^2$

korrigierte Stichprobenvarianz:

$$\begin{aligned}
 s^2 &= \frac{n}{n-1} \sigma_S^2 \\
 &= \frac{n}{n-1} \cdot \frac{1}{n} \cdot \sum_{i=1}^n (S_i - \bar{S})^2 \\
 &= \frac{1}{n-1} \cdot \sum_{i=1}^n (S_i - \bar{S})^2
 \end{aligned}$$

Mit “Standardabweichung von S ” ist meistens das korrigierte s gemeint.

Definition (Varianz, Kovarianz und Korrelation)

Die *Varianz* einer \mathbb{R} -wertigen Zufallsgröße X ist

$$\text{Var}X = \sigma_X^2 = \mathbb{E} [(X - \mathbb{E}X)^2] .$$

$\sigma_X = \sqrt{\text{Var} X}$ ist die *Standardabweichung*.

Definition (Varianz, Kovarianz und Korrelation)

Die *Varianz* einer \mathbb{R} -wertigen Zufallsgröße X ist

$$\text{Var}X = \sigma_X^2 = \mathbb{E} [(X - \mathbb{E}X)^2] .$$

$\sigma_X = \sqrt{\text{Var} X}$ ist die *Standardabweichung*.

Ist Y eine weitere reellwertige Zufallsvariable, so ist

$$\text{Cov}(X, Y) = \mathbb{E} [(X - \mathbb{E}X) \cdot (Y - \mathbb{E}Y)]$$

die *Kovarianz* von X und Y .

Definition (Varianz, Kovarianz und Korrelation)

Die *Varianz* einer \mathbb{R} -wertigen Zufallsgröße X ist

$$\text{Var}X = \sigma_X^2 = \mathbb{E} [(X - \mathbb{E}X)^2] .$$

$\sigma_X = \sqrt{\text{Var} X}$ ist die *Standardabweichung*.

Ist Y eine weitere reellwertige Zufallsvariable, so ist

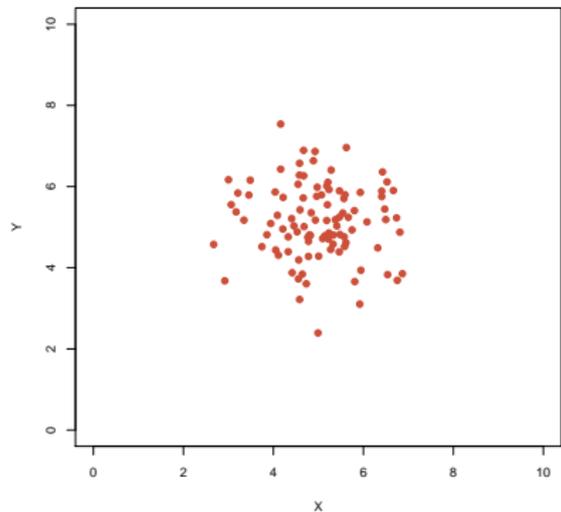
$$\text{Cov}(X, Y) = \mathbb{E} [(X - \mathbb{E}X) \cdot (Y - \mathbb{E}Y)]$$

die *Kovarianz* von X und Y .

Die *Korrelation* von X und Y ist

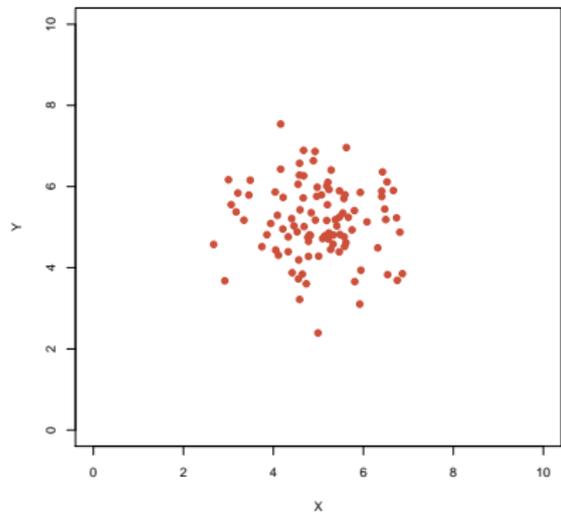
$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y} .$$

$$\sigma_X = 0.95, \sigma_Y = 0.92$$



$$\sigma_X = 0.95, \sigma_Y = 0.92$$

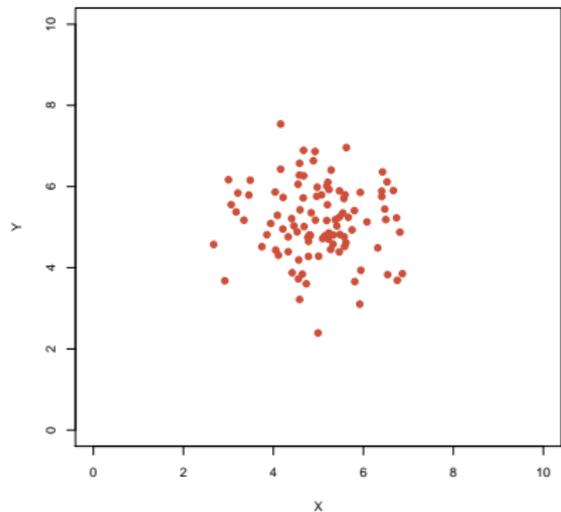
$$\text{Cov}(X, Y) = -0.06$$



$$\sigma_X = 0.95, \sigma_Y = 0.92$$

$$\text{Cov}(X, Y) = -0.06$$

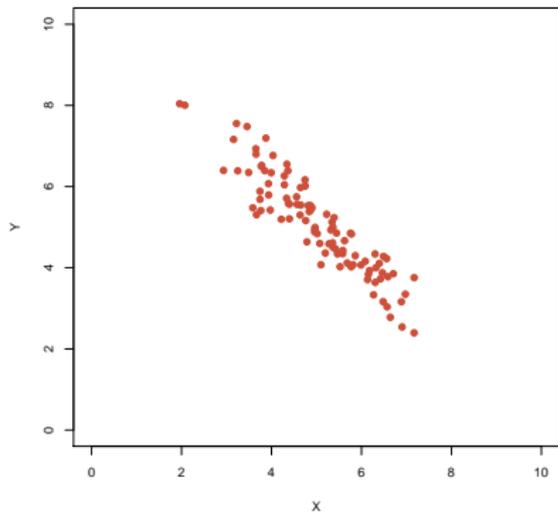
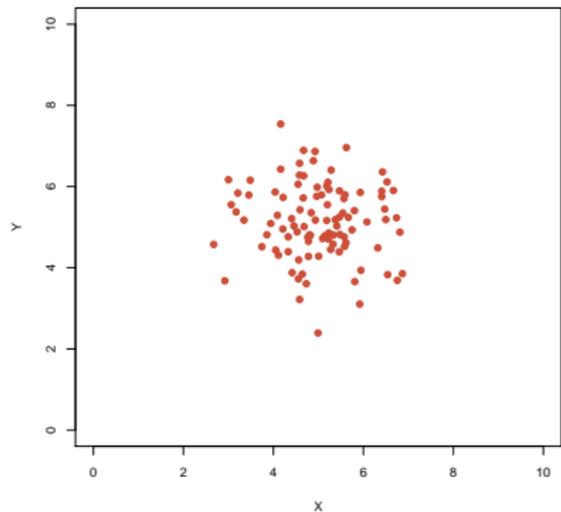
$$\text{Cor}(X, Y) = -0.069$$



$$\sigma_X = 0.95, \sigma_Y = 0.92$$

$$\text{Cov}(X, Y) = -0.06$$

$$\text{Cor}(X, Y) = -0.069$$

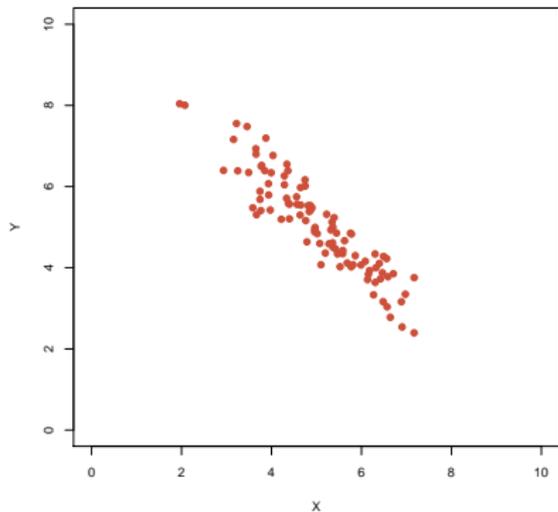
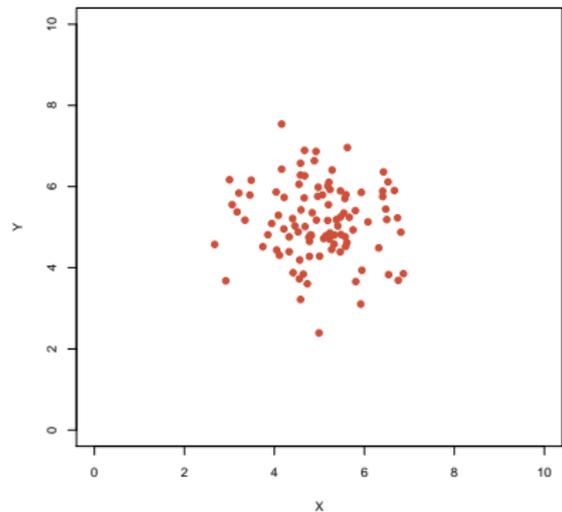


$$\sigma_X = 0.95, \sigma_Y = 0.92$$

$$\text{Cov}(X, Y) = -0.06$$

$$\text{Cor}(X, Y) = -0.069$$

$$\sigma_X = 1.13, \sigma_Y = 1.2$$



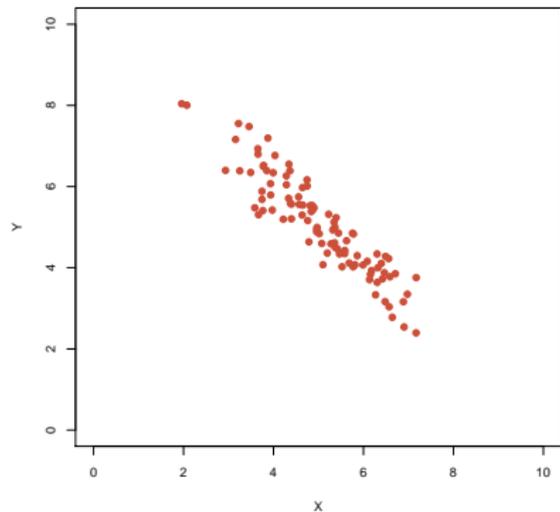
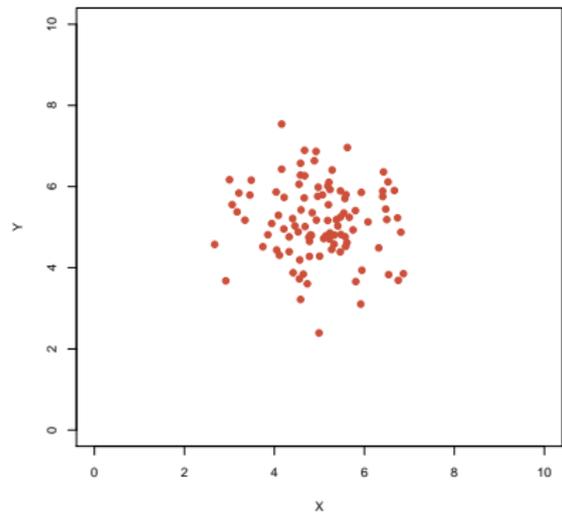
$$\sigma_X = 0.95, \sigma_Y = 0.92$$

$$\text{Cov}(X, Y) = -0.06$$

$$\text{Cor}(X, Y) = -0.069$$

$$\sigma_X = 1.13, \sigma_Y = 1.2$$

$$\text{Cov}(X, Y) = -1.26$$

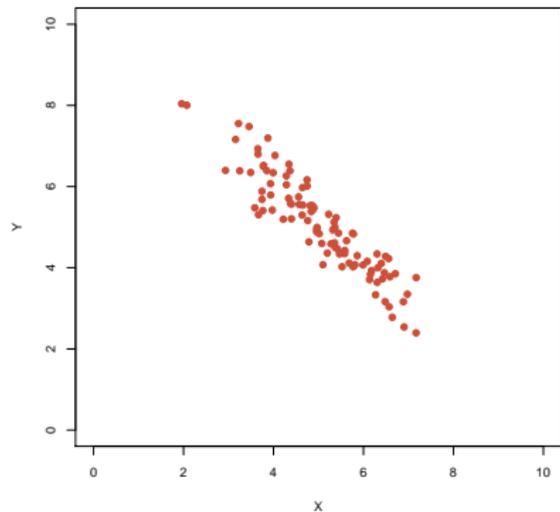
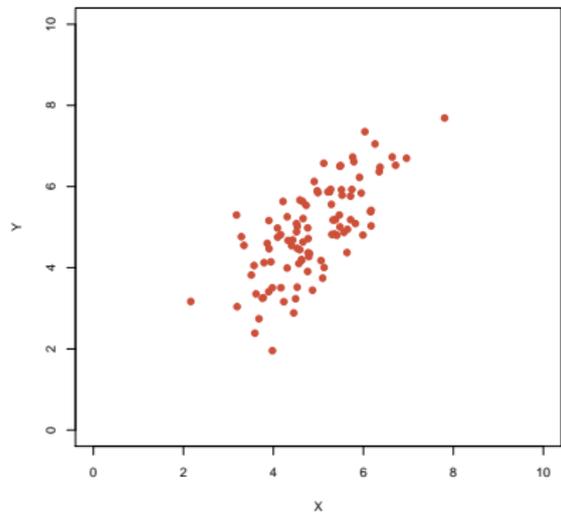


$$\sigma_X = 1.14, \sigma_Y = 0.78$$

$$\sigma_X = 1.13, \sigma_Y = 1.2$$

$$\text{Cov}(X, Y) = -1.26$$

$$\text{Cor}(X, Y) = -0.92$$



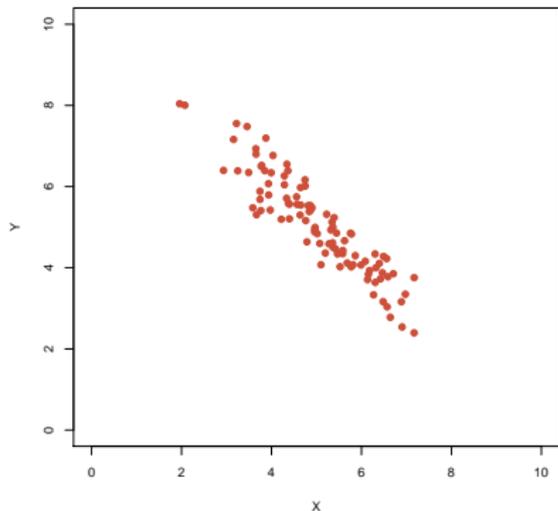
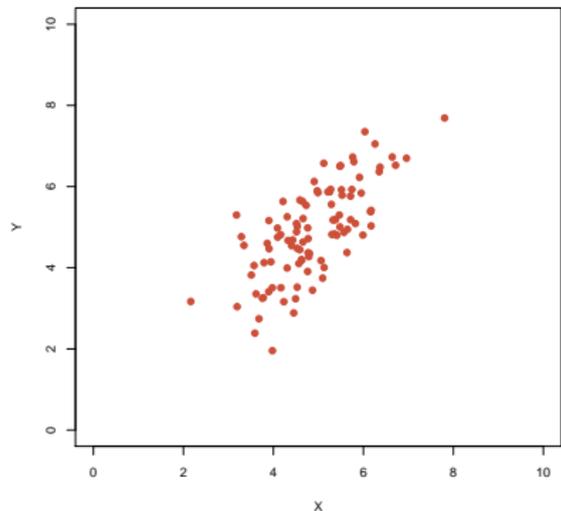
$$\sigma_X = 1.14, \sigma_Y = 0.78$$

$$\text{Cov}(X, Y) = 0.78$$

$$\sigma_X = 1.13, \sigma_Y = 1.2$$

$$\text{Cov}(X, Y) = -1.26$$

$$\text{Cor}(X, Y) = -0.92$$



$$\sigma_X = 1.14, \sigma_Y = 0.78$$

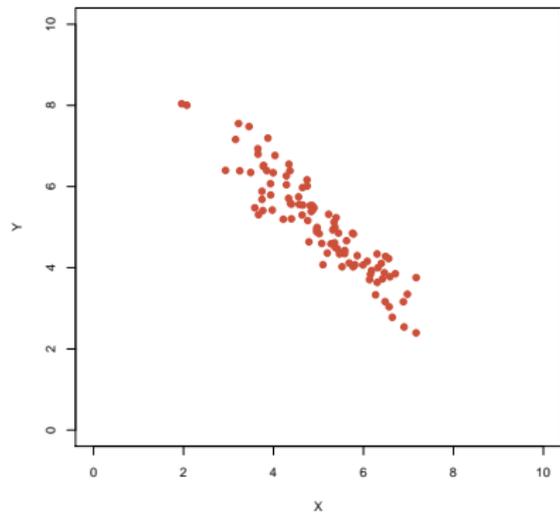
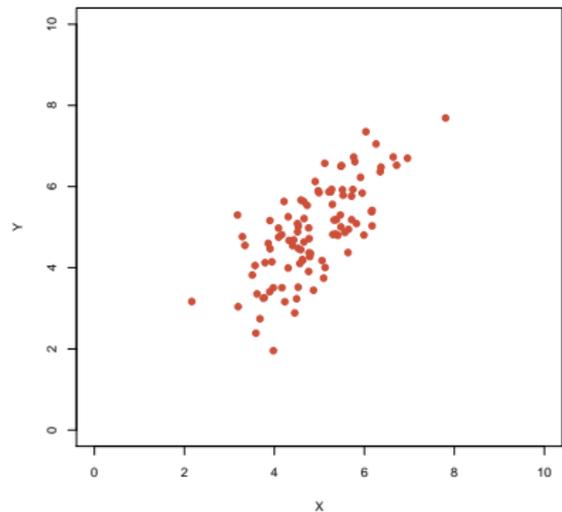
$$\text{Cov}(X, Y) = 0.78$$

$$\text{Cor}(X, Y) = 0.71$$

$$\sigma_X = 1.13, \sigma_Y = 1.2$$

$$\text{Cov}(X, Y) = -1.26$$

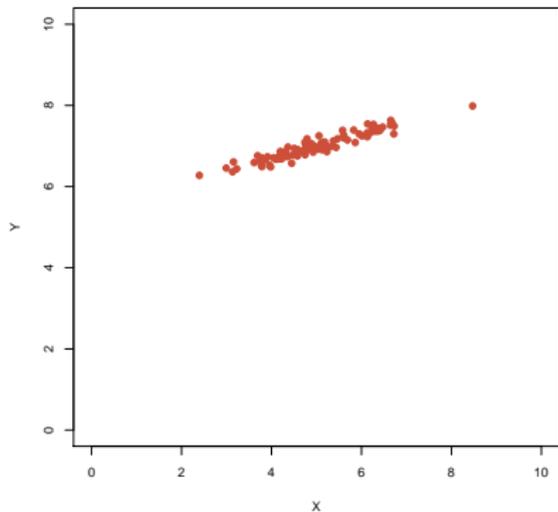
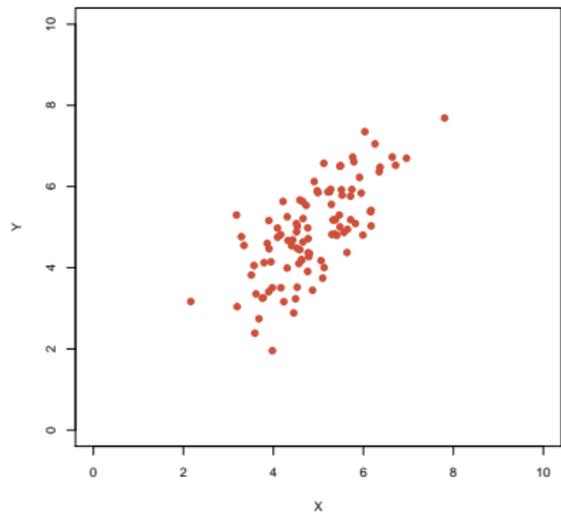
$$\text{Cor}(X, Y) = -0.92$$



$$\sigma_X = 1.14, \sigma_Y = 0.78$$

$$\text{Cov}(X, Y) = 0.78$$

$$\text{Cor}(X, Y) = 0.71$$

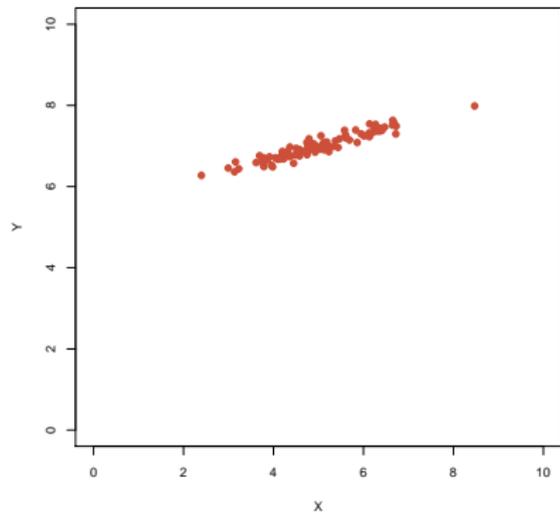
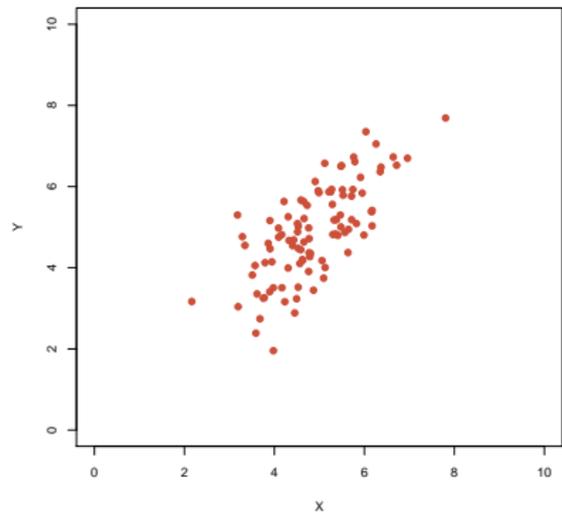


$$\sigma_X = 1.14, \sigma_Y = 0.78$$

$$\text{Cov}(X, Y) = 0.78$$

$$\text{Cor}(X, Y) = 0.71$$

$$\sigma_X = 1.03, \sigma_Y = 0.32$$



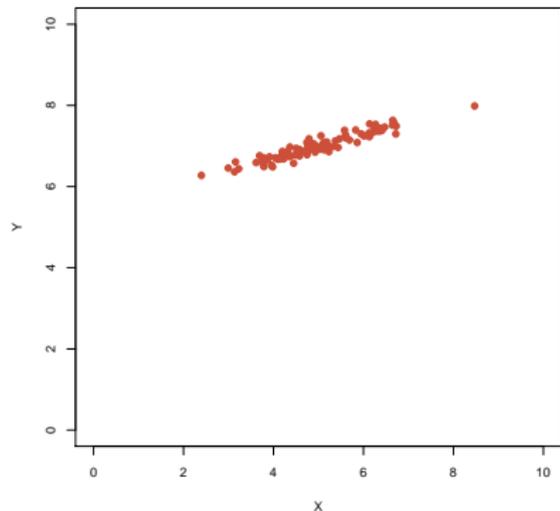
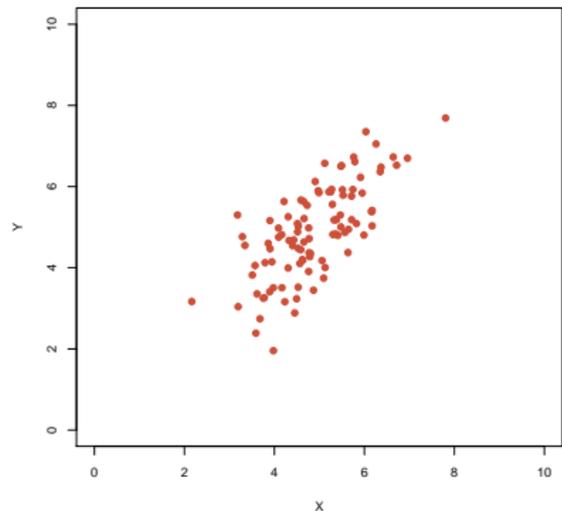
$$\sigma_X = 1.14, \sigma_Y = 0.78$$

$$\text{Cov}(X, Y) = 0.78$$

$$\text{Cor}(X, Y) = 0.71$$

$$\sigma_X = 1.03, \sigma_Y = 0.32$$

$$\text{Cov}(X, Y) = 0.32$$

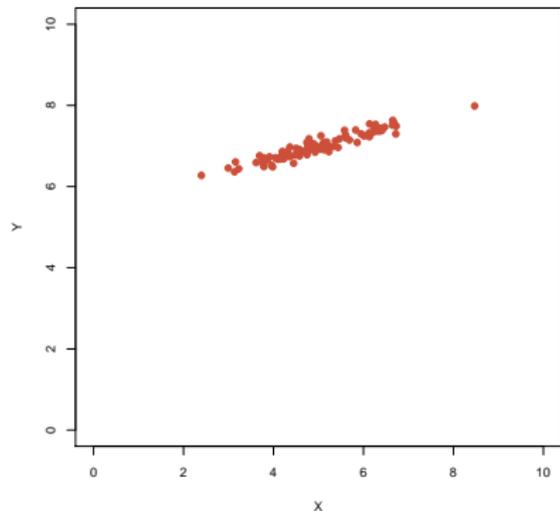
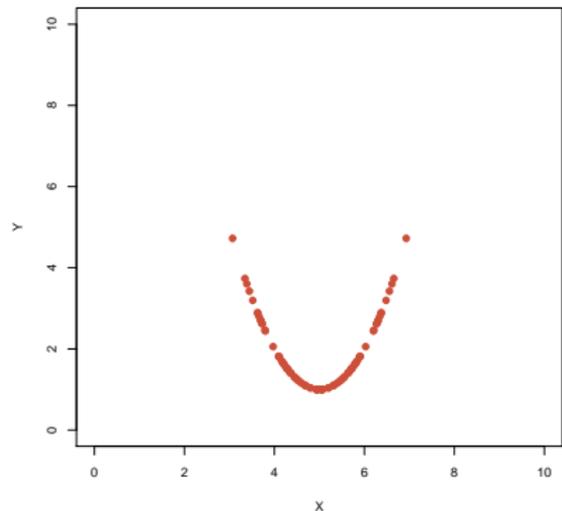


$$\sigma_X = 0.91, \sigma_Y = 0.88$$

$$\sigma_X = 1.03, \sigma_Y = 0.32$$

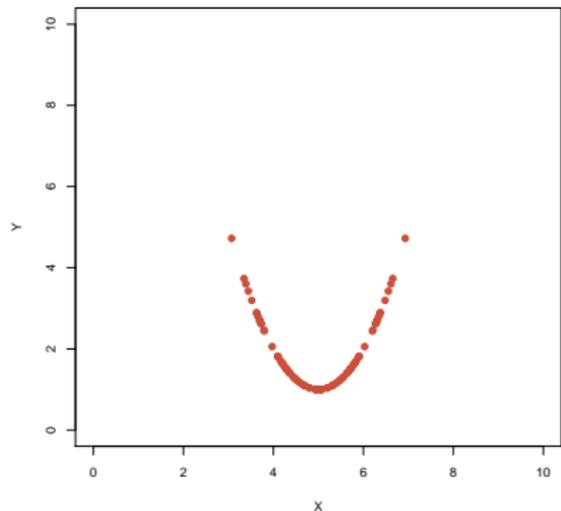
$$\text{Cov}(X, Y) = 0.32$$

$$\text{Cor}(X, Y) = 0.95$$



$$\sigma_X = 0.91, \sigma_Y = 0.88$$

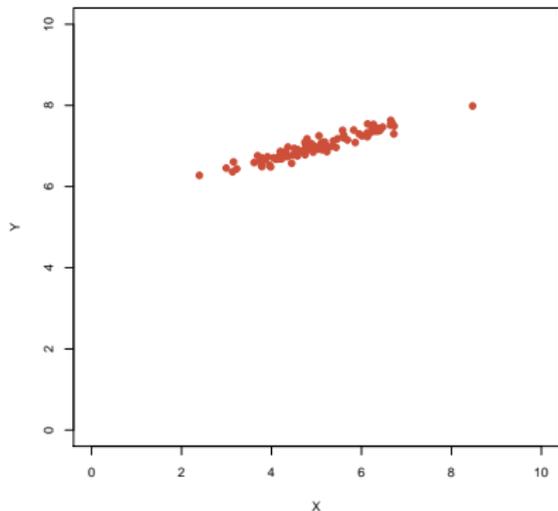
$$\text{Cov}(X, Y) = 0$$



$$\sigma_X = 1.03, \sigma_Y = 0.32$$

$$\text{Cov}(X, Y) = 0.32$$

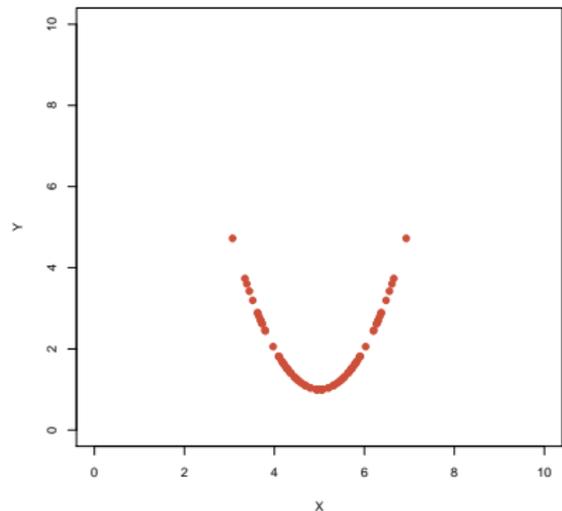
$$\text{Cor}(X, Y) = 0.95$$



$$\sigma_X = 0.91, \sigma_Y = 0.88$$

$$\text{Cov}(X, Y) = 0$$

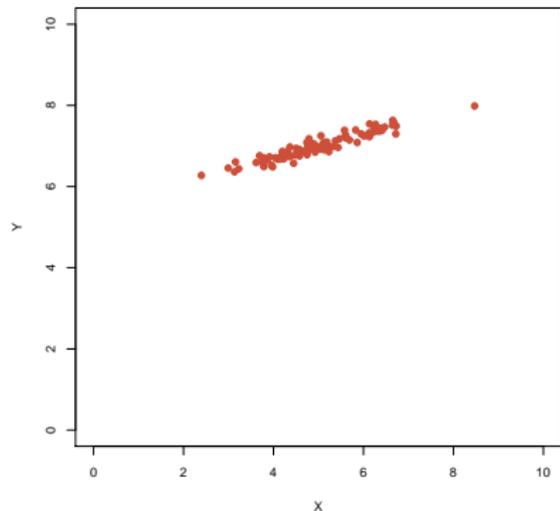
$$\text{Cor}(X, Y) = 0$$



$$\sigma_X = 1.03, \sigma_Y = 0.32$$

$$\text{Cov}(X, Y) = 0.32$$

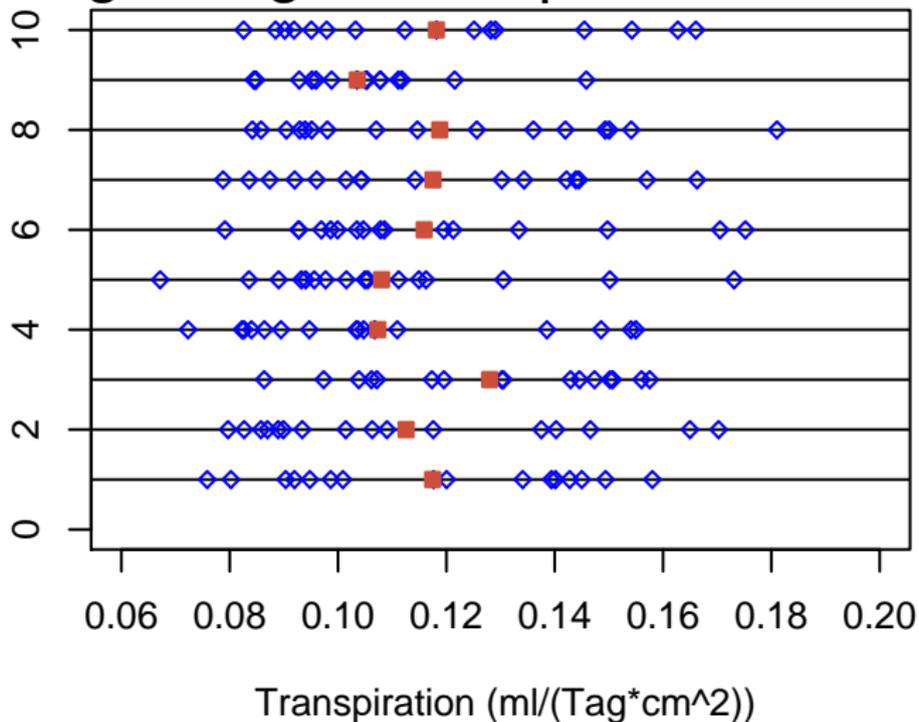
$$\text{Cor}(X, Y) = 0.95$$



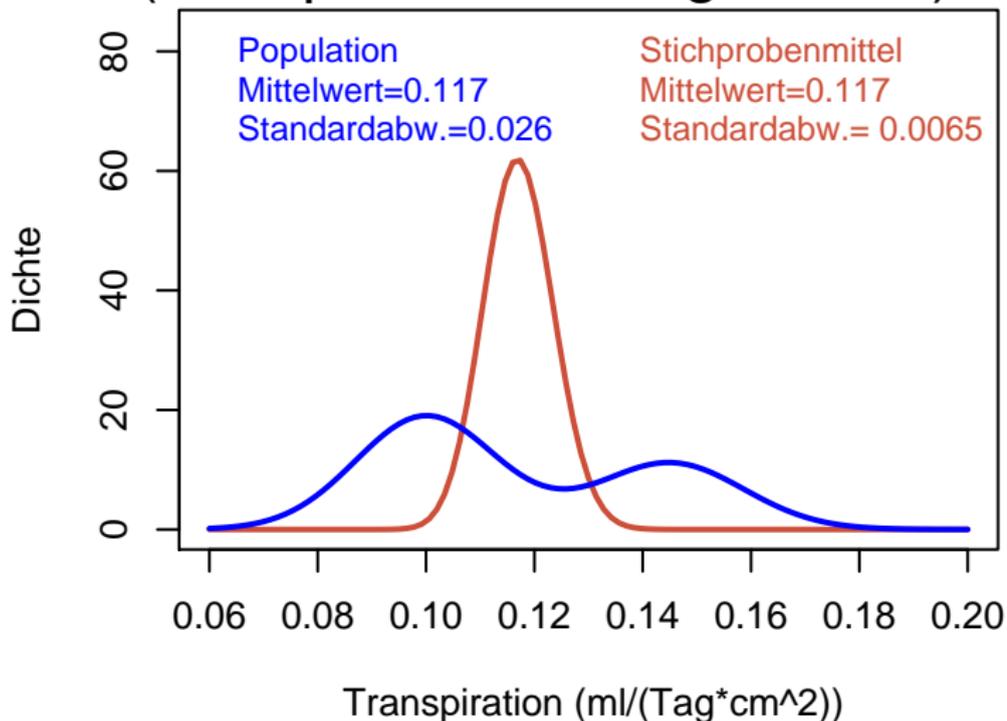
Inhalt

- 1 Deskriptive Statistik
- 2 Standardfehler und t-Tests**
- 3 Chi-Quadrat-Tests
 - χ^2 -Test für eine feste Verteilung
 - χ^2 -Test auf Unabhängigkeit (oder Homogenität)
- 4 Konfidenzintervalle
- 5 Lineare Regression
 - Lineare Zusammenhänge
 - t -Test fuer lineare Zusammenhänge

10 Stichproben vom Umfang 16 und die zugehörigen Stichprobenmittel



Verteilung der Stichprobenmittelwerte (Stichprobenumfang $n = 16$)



Die allgemeine Regel

Die Standardabweichung
des Mittelwerts einer Stichprobe vom
Umfang n

Die allgemeine Regel

Die Standardabweichung
des Mittelwerts einer Stichprobe vom
Umfang n

ist

$$1/\sqrt{n}$$

mal

der Standardabweichung
der Population.

Die Standardabweichung der Population
bezeichnet man mit

σ
(sigma).

Die Standardabweichung der Population
bezeichnet man mit

σ
(sigma).

Die Regel schreibt man häufig so:

$$\sigma(\bar{X}) = \frac{1}{\sqrt{n}}\sigma(X)$$

In der Praxis ist

σ

unbekannt.

In der Praxis ist

σ

unbekannt.

Es wird durch

die Stichproben-Standardabweichung S

geschätzt:

In der Praxis ist

σ

unbekannt.

Es wird durch

die Stichproben-Standardabweichung S

geschätzt:

$$\sigma = ??$$

In der Praxis ist

σ

unbekannt.

Es wird durch

die Stichproben-Standardabweichung S

geschätzt:

$$\sigma \approx S$$

$$s/\sqrt{n}$$

(die geschätzte
Standardabweichung
von \bar{x})
nennt man den
Standardfehler.

$$s/\sqrt{n}$$

(die geschätzte
Standardabweichung
von \bar{x})

nennt man den

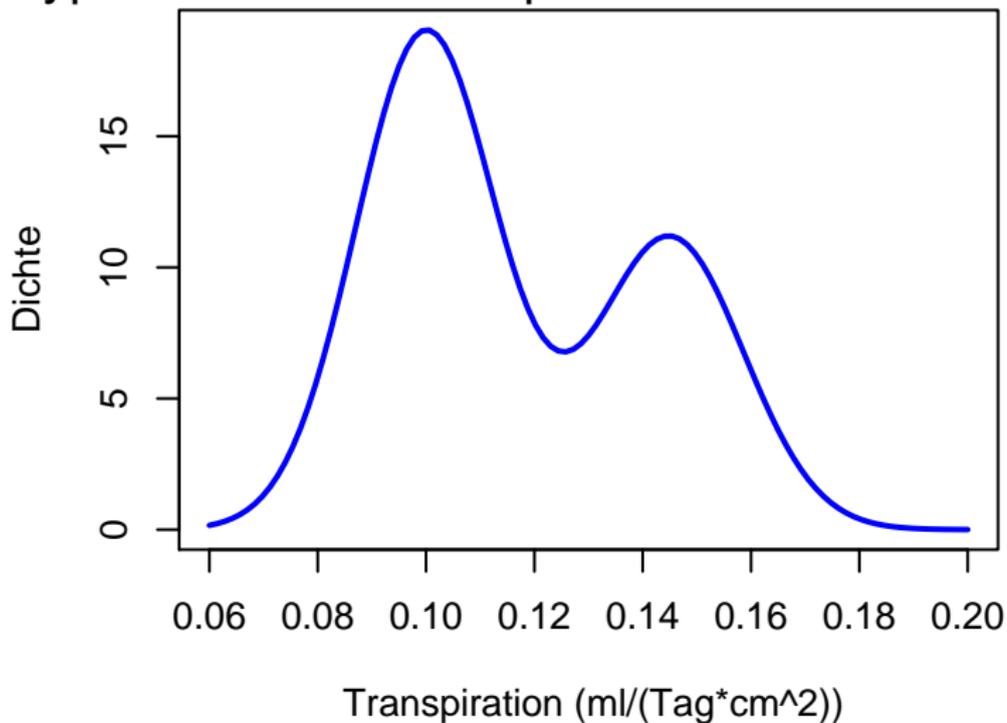
Standardfehler.

(Englisch: *standard error of the mean,
standard error, SEM*)

Wir haben gesehen:

Auch wenn die Verteilung von
 x mehrgipfelig
&
asymmetrisch
ist

Hypothetische Transpirationsratenverteilung



ist die Verteilung von

\bar{X}

trotzdem

(annähernd)

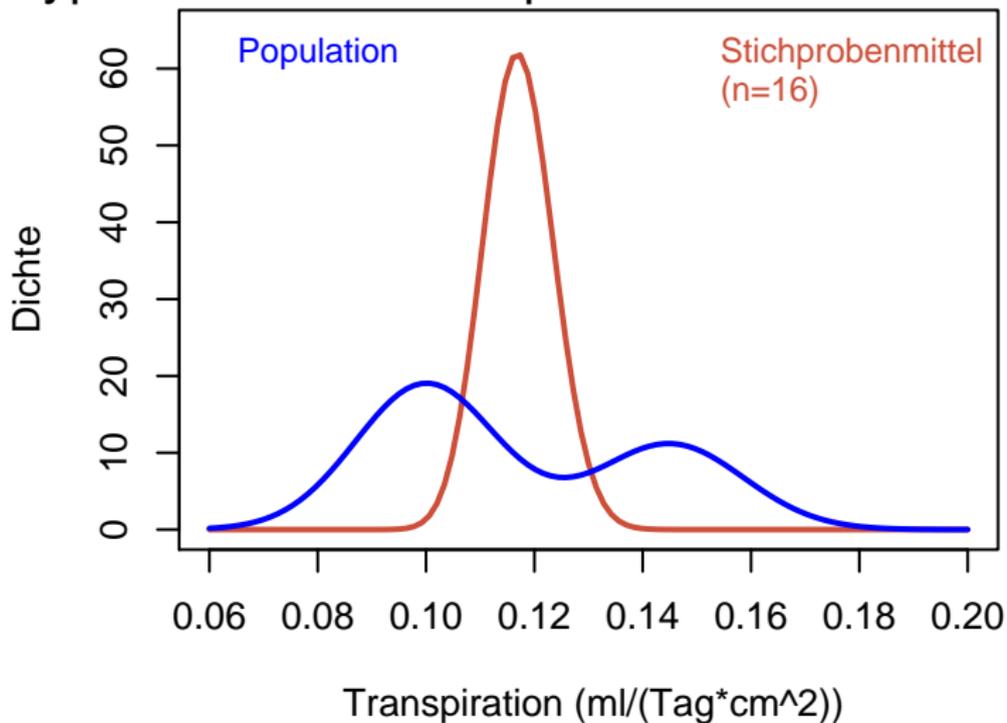
eingipfelig

&

symmetrisch

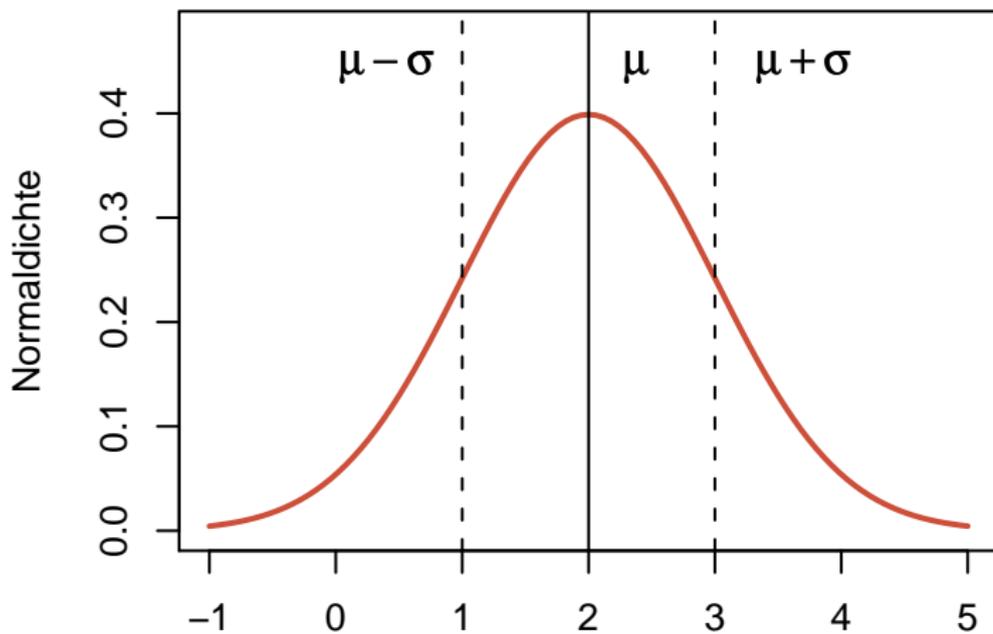
(wenn der Stichprobenumfang n nur groß genug ist)

Hypothetische Transpirationsratenverteilung

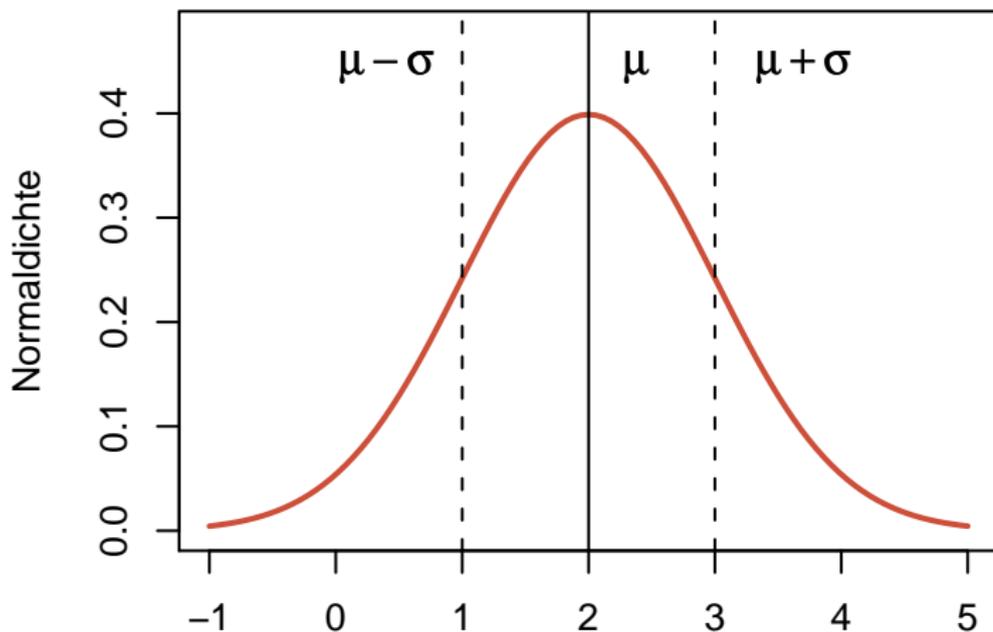


Die Verteilung von \bar{x}
hat annähernd
eine ganz bestimmte Form:
die Normalverteilung.

Dichte der Normalverteilung



Dichte der Normalverteilung



Die Normalverteilungsdichte heisst
auch *Gauß'sche Glockenkurve*

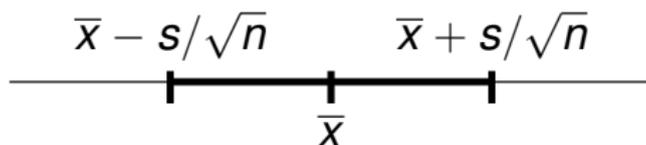
Dichte der Normalverteilung



Die Normalverteilungsdichte heisst
auch *Gauß'sche Glockenkurve*
(nach Carl Friedrich Gauß, 1777-1855)

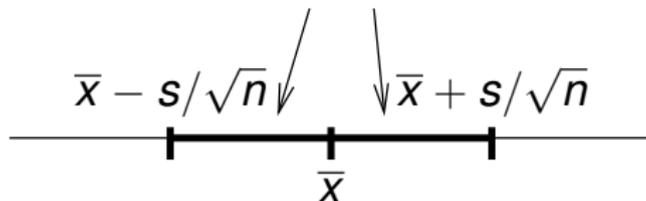
Wichtige Folgerung

Wir betrachten das Intervall



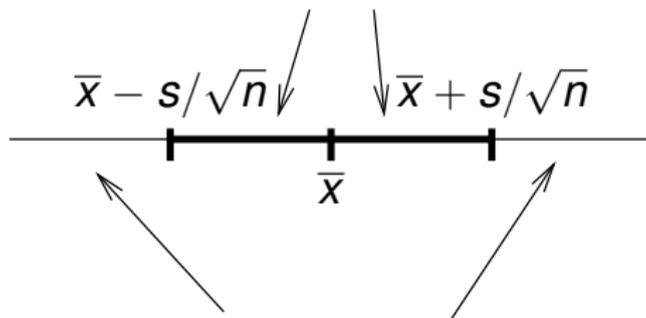
Wichtige Folgerung

Mit Wahrscheinlichkeit ca. $2/3$
liegt μ innerhalb dieses Intervalls



Wichtige Folgerung

Mit Wahrscheinlichkeit ca. $2/3$
liegt μ innerhalb dieses Intervalls



Mit Wahrscheinlichkeit ca. $1/3$
liegt μ **ausserhalb** des Intervalls

Demnach:

Es kommt durchaus vor, dass \bar{x}
von μ
um mehr als
 s/\sqrt{n} abweicht.

- Nehmen wir an, eine Population hat Mittelwert μ und Standardabweichung σ .

- Nehmen wir an, eine Population hat Mittelwert μ und Standardabweichung σ .
- Aus dieser Population ziehen wir eine Zufallsstichprobe vom Umfang n , mit Stichprobenmittelwert \bar{x} .

- Nehmen wir an, eine Population hat Mittelwert μ und Standardabweichung σ .
- Aus dieser Population ziehen wir eine Zufallsstichprobe vom Umfang n , mit Stichprobenmittelwert \bar{x} .
- \bar{x} ist eine Zufallsgröße

- Nehmen wir an, eine Population hat Mittelwert μ und Standardabweichung σ .
- Aus dieser Population ziehen wir eine Zufallsstichprobe vom Umfang n , mit Stichprobenmittelwert \bar{x} .
- \bar{x} ist eine Zufallsgröße mit Mittelwert μ und Standardabweichung σ/\sqrt{n} .

- Nehmen wir an, eine Population hat Mittelwert μ und Standardabweichung σ .
- Aus dieser Population ziehen wir eine Zufallsstichprobe vom Umfang n , mit Stichprobenmittelwert \bar{x} .
- \bar{x} ist eine Zufallsgröße mit Mittelwert μ und Standardabweichung σ/\sqrt{n} .
- Man schätzt die Standardabweichung von \bar{x} mit s/\sqrt{n} .

- Nehmen wir an, eine Population hat Mittelwert μ und Standardabweichung σ .
- Aus dieser Population ziehen wir eine Zufallsstichprobe vom Umfang n , mit Stichprobenmittelwert \bar{x} .
- \bar{x} ist eine Zufallsgröße mit Mittelwert μ und Standardabweichung σ/\sqrt{n} .
- Man schätzt die Standardabweichung von \bar{x} mit s/\sqrt{n} .
- s/\sqrt{n} nennt man den **Standardfehler**.

- Nehmen wir an, eine Population hat Mittelwert μ und Standardabweichung σ .
- Aus dieser Population ziehen wir eine Zufallsstichprobe vom Umfang n , mit Stichprobenmittelwert \bar{x} .
- \bar{x} ist eine Zufallsgröße mit Mittelwert μ und Standardabweichung σ/\sqrt{n} .
- Man schätzt die Standardabweichung von \bar{x} mit s/\sqrt{n} .
- s/\sqrt{n} nennt man den **Standardfehler**.
- Schwankungen in \bar{x} von der Größe s/\sqrt{n} kommen häufig vor.

- Nehmen wir an, eine Population hat Mittelwert μ und Standardabweichung σ .
- Aus dieser Population ziehen wir eine Zufallsstichprobe vom Umfang n , mit Stichprobenmittelwert \bar{x} .
- \bar{x} ist eine Zufallsgröße mit Mittelwert μ und Standardabweichung σ/\sqrt{n} .
- Man schätzt die Standardabweichung von \bar{x} mit s/\sqrt{n} .
- s/\sqrt{n} nennt man den **Standardfehler**.
- Schwankungen in \bar{x} von der Größe s/\sqrt{n} kommen häufig vor.
Solche Schwankungen sind „**nicht signifikant**“: sie könnten Zufall sein.

Allgemein gilt

Sind X_1, \dots, X_n unabhängig aus einer Normalverteilung mit Mittelwert μ gezogen und ist

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2},$$

so ist

$$\frac{\bar{X} - \mu}{s/\sqrt{n}}$$

t-verteilt mit $n - 1$ Freiheitsgraden (df=*degrees of freedom*).

Allgemein gilt

Sind X_1, \dots, X_n unabhängig aus einer Normalverteilung mit Mittelwert μ gezogen und ist

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2},$$

so ist

$$\frac{\bar{X} - \mu}{s/\sqrt{n}}$$

t-verteilt mit $n - 1$ Freiheitsgraden (df=*degrees of freedom*).

Die t-Verteilung heißt auch *Student-Verteilung*, da Gosset sie unter diesem Pseudonym publiziert hat.

Wie (un)wahrscheinlich ist nun eine
so große Abweichung wie 2.34 Standardfehler?

$$\mathbb{P}(T = 2.34) =$$

Wie (un)wahrscheinlich ist nun eine
so große Abweichung wie 2.34 Standardfehler?

$$\mathbb{P}(T = 2.34) = 0$$

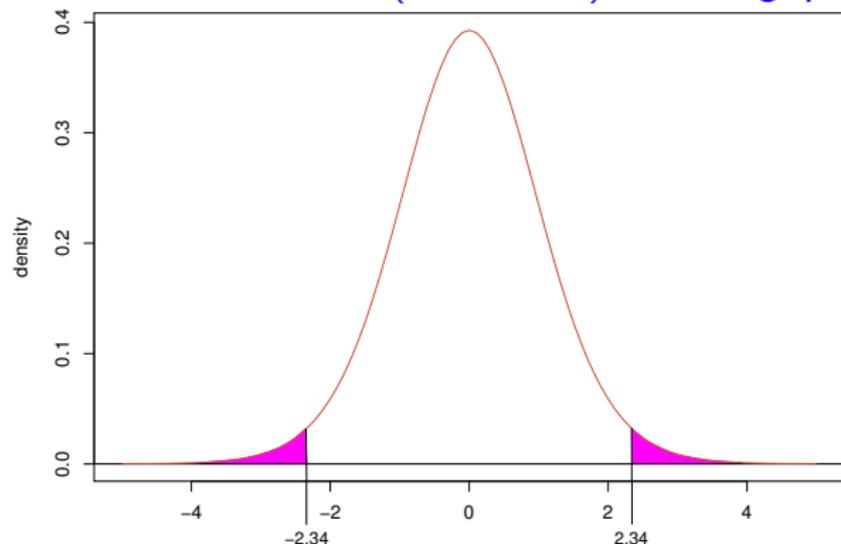
Wie (un)wahrscheinlich ist nun eine **mindestens** so große Abweichung wie 2.34 Standardfehler?

$$\mathbb{P}(T = 2.34) = 0 \quad \text{Das bringt nichts!}$$

Wie (un)wahrscheinlich ist nun eine
so große Abweichung wie 2.34 Standardfehler?

$$\mathbb{P}(T = 2.34) = 0 \quad \text{Das bringt nichts!}$$

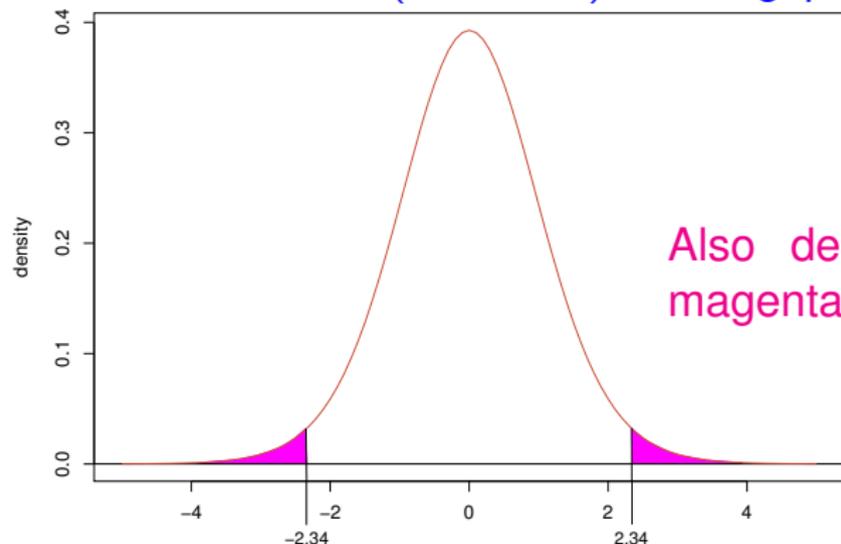
Zu berechnen ist $\mathbb{P}(T \geq 2.34)$, der sog. p -Wert.



Wie (un)wahrscheinlich ist nun eine
so große Abweichung wie 2.34 Standardfehler?

$$\mathbb{P}(T = 2.34) = 0 \quad \text{Das bringt nichts!}$$

Zu berechnen ist $\mathbb{P}(T \geq 2.34)$, der sog. p -Wert.



Also der Gesamthalt der
magentafarbenen Flächen.

Wir halten fest:

$$p - \text{Wert} = 0.03254$$

Wir halten fest:

$$p - \text{Wert} = 0.03254$$

d.h.: Wenn die **Nullhypothese** “alles nur Zufall”, also in diesem Fall die Hypothese $\mu = 0$ gilt, dann ist eine mindestens so große Abweichung sehr unwahrscheinlich.

Wir halten fest:

$$p - \text{Wert} = 0.03254$$

d.h.: Wenn die **Nullhypothese** “alles nur Zufall”, also in diesem Fall die Hypothese $\mu = 0$ gilt, dann ist eine mindestens so große Abweichung sehr unwahrscheinlich.

Wenn wir beschließen, dass wir die Nullhypothese immer verwerfen, wenn der p -Wert unterhalb einem **Signifikanzniveau** von 0.05 liegt, gilt:

Wir halten fest:

$$p - \text{Wert} = 0.03254$$

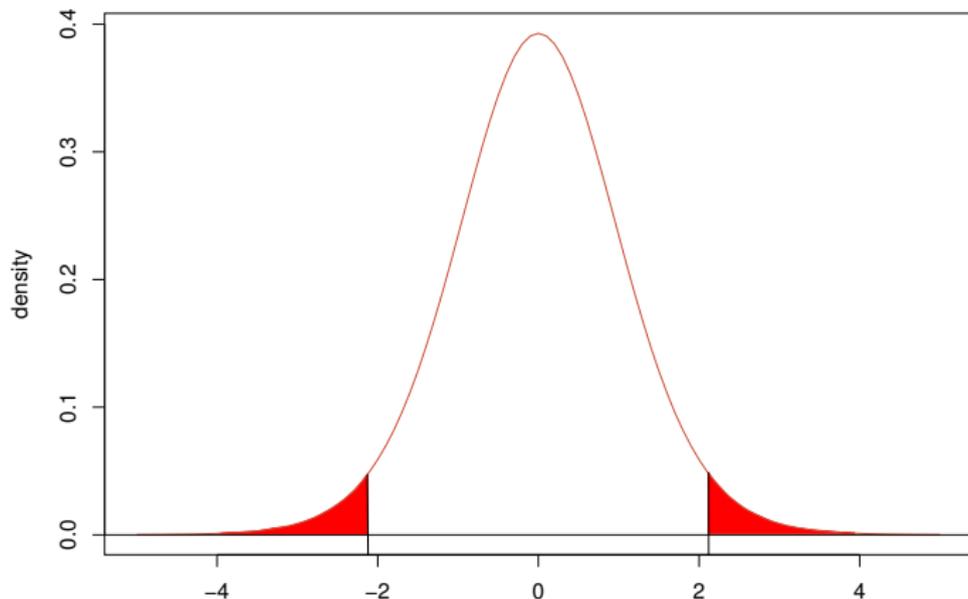
d.h.: Wenn die **Nullhypothese** “alles nur Zufall”, also in diesem Fall die Hypothese $\mu = 0$ gilt, dann ist eine mindestens so große Abweichung sehr unwahrscheinlich.

Wenn wir beschließen, dass wir die Nullhypothese immer verwerfen, wenn der p -Wert unterhalb einem

Signifikanzniveau von 0.05 liegt, gilt:

Falls die Nullhypothese zutrifft, ist die Wahrscheinlichkeit, dass wir sie zu Unrecht verwerfen, lediglich 0.05.

Wenn wir uns auf ein Signifikanzniveau von $\alpha = 0.05$ festlegen, verwerfen wir die Nullhypothese also, wenn der t -Wert in den roten Bereich fällt:



(hier am Beispiel der t -Verteilung mit $df = 16$ Freiheitsgraden)

Welche t -Werte sind “auf dem 5%-Niveau” signifikant?

Anzahl Freiheitsgrade	$ t \geq \dots$
5	2.57
10	2.23
20	2.09
30	2.04
100	1.98
∞	1.96

- Wir möchten belegen, dass eine Abweichung in den Daten vermutlich nicht allein auf Zufallsschwankung beruht.

- Wir möchten belegen, dass eine Abweichung in den Daten vermutlich nicht allein auf Zufallsschwankung beruht.
- Dazu spezifizieren wir zunächst eine **Nullhypothese H_0** , d.h. wir konkretisieren, was “allein auf Zufall beruhen” bedeutet.

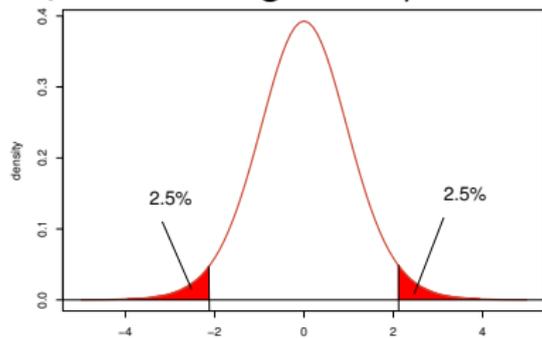
- Wir möchten belegen, dass eine Abweichung in den Daten vermutlich nicht allein auf Zufallsschwankung beruht.
- Dazu spezifizieren wir zunächst eine **Nullhypothese H_0** , d.h. wir konkretisieren, was “allein auf Zufall beruhen” bedeutet.
- Dann versuchen wir zu zeigen: Wenn H_0 gilt, dann ist eine **Abweichung** wie, die mindestens so groß sind wie die beobachtete, sehr unwahrscheinlich.

- Wir möchten belegen, dass eine Abweichung in den Daten vermutlich nicht allein auf Zufallsschwankung beruht.
- Dazu spezifizieren wir zunächst eine **Nullhypothese H_0** , d.h. wir konkretisieren, was “allein auf Zufall beruhen” bedeutet.
- Dann versuchen wir zu zeigen: Wenn H_0 gilt, dann ist eine **Abweichung** wie, die mindestens so groß sind wie die beobachtete, sehr unwahrscheinlich.
- Wenn uns das gelingt, verwerfen wir H_0 .

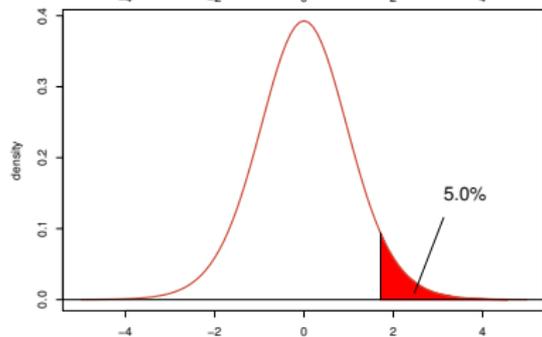
- Wir möchten belegen, dass eine Abweichung in den Daten vermutlich nicht allein auf Zufallsschwankung beruht.
- Dazu spezifizieren wir zunächst eine **Nullhypothese H_0** , d.h. wir konkretisieren, was “allein auf Zufall beruhen” bedeutet.
- Dann versuchen wir zu zeigen: Wenn H_0 gilt, dann ist eine **Abweichung** wie, die mindestens so groß sind wie die beobachtete, sehr unwahrscheinlich.
- Wenn uns das gelingt, verwerfen wir H_0 .
- Was wir als **Abweichung** auffassen, sollten klar sein, bevor wir die Daten sehen.

Zweiseitig oder einseitig testen?

Wir beobachten einen Wert x , der deutlich größer als der H_0 -Erwartungswert μ ist.



$$p\text{-Wert} = \mathbb{P}_{H_0}(|X - \mu| \geq |x - \mu|)$$



$$p\text{-Wert} = \mathbb{P}_{H_0}(X \geq x)$$

Reine Lehre des statistischen Testens

- Formuliere eine **Nullhypothese H_0** , z.B. $\mu = 0$.

Reine Lehre des statistischen Testens

- Formuliere eine **Nullhypothese** H_0 , z.B. $\mu = 0$.
- Lege ein **Signifikanzniveau** α fest; üblich ist $\alpha = 0.05$.

Reine Lehre des statistischen Testens

- Formuliere eine **Nullhypothese** H_0 , z.B. $\mu = 0$.
- Lege ein **Signifikanzniveau** α fest; üblich ist $\alpha = 0.05$.
- Lege ein Ereignis \mathcal{A} fest, so dass

$$\mathbb{P}_{H_0}(\mathcal{A}) = \alpha$$

(oder zumindest $\mathbb{P}_{H_0}(\mathcal{A}) \leq \alpha$).

Reine Lehre des statistischen Testens

- Formuliere eine **Nullhypothese** H_0 , z.B. $\mu = 0$.
- Lege ein **Signifikanzniveau** α fest; üblich ist $\alpha = 0.05$.
- Lege ein Ereignis \mathcal{A} fest, so dass

$$\mathbb{P}_{H_0}(\mathcal{A}) = \alpha$$

(oder zumindest $\mathbb{P}_{H_0}(\mathcal{A}) \leq \alpha$).

z.B. $\mathcal{A} = \{\bar{X} > q\}$ oder $\mathcal{A} = \{|\bar{X} - \mu| > r\}$

Reine Lehre des statistischen Testens

- Formuliere eine **Nullhypothese** H_0 , z.B. $\mu = 0$.
- Lege ein **Signifikanzniveau** α fest; üblich ist $\alpha = 0.05$.
- Lege ein Ereignis \mathcal{A} fest, so dass

$$\mathbb{P}_{H_0}(\mathcal{A}) = \alpha$$

(oder zumindest $\mathbb{P}_{H_0}(\mathcal{A}) \leq \alpha$).

z.B. $\mathcal{A} = \{\bar{X} > q\}$ oder $\mathcal{A} = \{|\bar{X} - \mu| > r\}$

allgemein: $H_0 = \{\text{p-Wert} \leq \alpha\}$

Reine Lehre des statistischen Testens

- Formuliere eine **Nullhypothese** H_0 , z.B. $\mu = 0$.
- Lege ein **Signifikanzniveau** α fest; üblich ist $\alpha = 0.05$.
- Lege ein Ereignis \mathcal{A} fest, so dass

$$\mathbb{P}_{H_0}(\mathcal{A}) = \alpha$$

(oder zumindest $\mathbb{P}_{H_0}(\mathcal{A}) \leq \alpha$).

z.B. $\mathcal{A} = \{\bar{X} > q\}$ oder $\mathcal{A} = \{|\bar{X} - \mu| > r\}$

allgemein: $H_0 = \{p\text{-Wert} \leq \alpha\}$

- **ERST DANN:** Betrachte die Daten und überprüfe, **ob** \mathcal{A} eintritt.

Reine Lehre des statistischen Testens

- Formuliere eine **Nullhypothese** H_0 , z.B. $\mu = 0$.
- Lege ein **Signifikanzniveau** α fest; üblich ist $\alpha = 0.05$.
- Lege ein Ereignis \mathcal{A} fest, so dass

$$\mathbb{P}_{H_0}(\mathcal{A}) = \alpha$$

(oder zumindest $\mathbb{P}_{H_0}(\mathcal{A}) \leq \alpha$).

z.B. $\mathcal{A} = \{\bar{X} > q\}$ oder $\mathcal{A} = \{|\bar{X} - \mu| > r\}$

allgemein: $H_0 = \{\text{p-Wert} \leq \alpha\}$

- **ERST DANN:** Betrachte die Daten und überprüfe, **ob** \mathcal{A} eintritt.
- Dann ist die Wahrscheinlichkeit, dass H_0 verworfen wird, wenn H_0 eigentlich richtig ist ("**Fehler erster Art**"), lediglich α .

Verstöße gegen die reine Lehre

“Beim zweiseitigen Testen kam ein p -Wert von 0.06 raus. Also hab ich einseitig getestet, da hat's dann funktioniert.”

Verstöße gegen die reine Lehre

“Beim zweiseitigen Testen kam ein p -Wert von 0.06 raus. Also hab ich einseitig getestet, da hat's dann funktioniert.”

genauso problematisch:

Verstöße gegen die reine Lehre

“Beim zweiseitigen Testen kam ein p -Wert von 0.06 raus. Also hab ich einseitig getestet, da hat's dann funktioniert.”

genauso problematisch:

“Beim ersten Blick auf die Daten habe ich sofort gesehen, dass \bar{x} größer ist als μ_{H_0} . Also habe ich gleich einseitig getestet”

Wichtig

Die Entscheidung ob einseitig oder zweiseitig getestet wird darf nicht von den konkreten Daten, die zum Test verwendet werden, abhängen.

Wichtig

Die Entscheidung ob einseitig oder zweiseitig getestet wird darf nicht von den konkreten Daten, die zum Test verwendet werden, abhängen.

Allgemeiner: Ist \mathcal{A} das Ereignis, dass zum Verwerfen von H_0 führt (falls es eintritt), so muss die Festlegung von H_0 stattfinden bevor man in den Daten herumgeschnüffelt hat.

Die **Wahl von \mathcal{A}** sollte von der **Alternative H_1** abhängen, also für das, was wir eigentlich zeigen wollen, indem wir H_0 durch einen Test verwerfen. Es muss gelten:

$$\mathbb{P}_{H_0}(\mathcal{A}) = \alpha$$

und

$$\mathbb{P}_{H_1}(\mathcal{A}) = \text{möglichst groß,}$$

Die **Wahl von \mathcal{A}** sollte von der **Alternative H_1** abhängen, also für das, was wir eigentlich zeigen wollen, indem wir H_0 durch einen Test verwerfen. Es muss gelten:

$$\mathbb{P}_{H_0}(\mathcal{A}) = \alpha$$

und

$$\mathbb{P}_{H_1}(\mathcal{A}) = \text{möglichst groß,}$$

damit die **W'keit eines Fehlers zweiter Art**, dass also H_0 nicht verworfen wird, obwohl H_1 zutrifft, möglichst klein.

Beispiele

- Wenn wir von Anfang an unsere Vermutung belegen wollten, dass sich die Trauerschnäpper bei grünem Licht stärker auf eine Richtung konzentrieren als bei blauem, dürfen wir einseitig testen.

Beispiele

- Wenn wir von Anfang an unsere Vermutung belegen wollten, dass sich die Trauerschnäpper bei grünem Licht stärker auf eine Richtung konzentrieren als bei blauem, dürfen wir einseitig testen.
- Wenn dann aber noch so deutlich herauskommt, dass die Richtungswahl bei blauem Licht deutlicher war, so ist das dann nicht als signifikant zu betrachten.

Beispiele

- Wenn wir von Anfang an unsere Vermutung belegen wollten, dass sich die Trauerschnäpper bei grünem Licht stärker auf eine Richtung konzentrieren als bei blauem, dürfen wir einseitig testen.
- Wenn dann aber noch so deutlich herauskommt, dass die Richtungswahl bei blauem Licht deutlicher war, so ist das dann nicht als signifikant zu betrachten.
- Wenn wir von Anfang an die Vermutung belegen wollten, dass der Kork an der Nordseite des Baumes dicker war, dürfen wir einseitig testen.

Beispiele

- Wenn wir von Anfang an unsere Vermutung belegen wollten, dass sich die Trauerschnäpper bei grünem Licht stärker auf eine Richtung konzentrieren als bei blauem, dürfen wir einseitig testen.
- Wenn dann aber noch so deutlich herauskommt, dass die Richtungswahl bei blauem Licht deutlicher war, so ist das dann nicht als signifikant zu betrachten.
- Wenn wir von Anfang an die Vermutung belegen wollten, dass der Kork an der Nordseite des Baumes dicker war, dürfen wir einseitig testen.
- Wenn dann aber noch so deutlich herauskommt, dass der Kork im Westen dicker ist, ist das nicht mehr signifikant.

Angenommen, H_0 wird auf dem 5%-Niveau verworfen. Welche Aussage gilt dann?

- Die Nullhypothese ist falsch.
-
-

Angenommen, H_0 wird auf dem 5%-Niveau verworfen. Welche Aussage gilt dann?

- ~~Die Nullhypothese ist falsch.~~
-
-

Angenommen, H_0 wird auf dem 5%-Niveau verworfen. Welche Aussage gilt dann?

- ~~Die Nullhypothese ist falsch.~~
- H_0 ist mit 95%-iger Wahrscheinlichkeit falsch.
-

Angenommen, H_0 wird auf dem 5%-Niveau verworfen. Welche Aussage gilt dann?

- ~~Die Nullhypothese ist falsch.~~
- ~~H_0 ist mit 95% iger Wahrscheinlichkeit falsch.~~
-

Angenommen, H_0 wird auf dem 5%-Niveau verworfen. Welche Aussage gilt dann?

- ~~Die Nullhypothese ist falsch.~~
- ~~H_0 ist mit 95% iger Wahrscheinlichkeit falsch.~~
- Falls die Nullhypothese wahr ist, beobachtet man ein so extremes Ergebnis nur in 5% der Fälle.

Angenommen, H_0 wird auf dem 5%-Niveau verworfen. Welche Aussage gilt dann?

- ~~Die Nullhypothese ist falsch.~~
- ~~H_0 ist mit 95% iger Wahrscheinlichkeit falsch.~~
- Falls die Nullhypothese wahr ist, beobachtet man ein so extremes Ergebnis nur in 5% der Fälle. ✓

Angenommen, H_0 konnte durch den Test nicht verworfen werden. Welche Aussagen sind dann richtig?

- Wir müssen die Alternative H_1 verwerfen.
-
-
-
-
-

Angenommen, H_0 konnte durch den Test nicht verworfen werden. Welche Aussagen sind dann richtig?

- ~~Wir müssen die Alternative H_1 verwerfen.~~
-
-
-
-
-

Angenommen, H_0 konnte durch den Test nicht verworfen werden. Welche Aussagen sind dann richtig?

- ~~Wir müssen die Alternative H_1 verwerfen.~~
- H_0 ist wahr.
-
-
-
-

Angenommen, H_0 konnte durch den Test nicht verworfen werden. Welche Aussagen sind dann richtig?

- ~~Wir müssen die Alternative H_1 verwerfen.~~
- ~~H_0 ist wahr.~~
-
-
-
-

Angenommen, H_0 konnte durch den Test nicht verworfen werden. Welche Aussagen sind dann richtig?

- ~~Wir müssen die Alternative H_1 verwerfen.~~
- ~~H_0 ist wahr.~~
- H_0 ist wahrscheinlich wahr.
-
-
-

Angenommen, H_0 konnte durch den Test nicht verworfen werden. Welche Aussagen sind dann richtig?

- ~~Wir müssen die Alternative H_1 verwerfen.~~
- ~~H_0 ist wahr.~~
- ~~H_0 ist wahrscheinlich wahr.~~
-
-
-

Angenommen, H_0 konnte durch den Test nicht verworfen werden. Welche Aussagen sind dann richtig?

- ~~Wir müssen die Alternative H_1 verwerfen.~~
- ~~H_0 ist wahr.~~
- ~~H_0 ist wahrscheinlich wahr.~~
- Es ist ungefährlich, davon auszugehen, dass H_0 zutrifft.
-
-

Angenommen, H_0 konnte durch den Test nicht verworfen werden. Welche Aussagen sind dann richtig?

- ~~Wir müssen die Alternative H_1 verwerfen.~~
- ~~H_0 ist wahr.~~
- ~~H_0 ist wahrscheinlich wahr.~~
- ~~Es ist ungefährlich, davon auszugehen, dass H_0 zutrifft.~~
-
-

Angenommen, H_0 konnte durch den Test nicht verworfen werden. Welche Aussagen sind dann richtig?

- ~~Wir müssen die Alternative H_1 verwerfen.~~
- ~~H_0 ist wahr.~~
- ~~H_0 ist wahrscheinlich wahr.~~
- ~~Es ist ungefährlich, davon auszugehen, dass H_0 zutrifft.~~
- Auch wenn H_0 zutrifft, ist es nicht sehr unwahrscheinlich, dass unsere Teststatistik einen so extrem erscheinenden Wert annimmt.
-

Angenommen, H_0 konnte durch den Test nicht verworfen werden. Welche Aussagen sind dann richtig?

- ~~Wir müssen die Alternative H_1 verwerfen.~~
- ~~H_0 ist wahr.~~
- ~~H_0 ist wahrscheinlich wahr.~~
- ~~Es ist ungefährlich, davon auszugehen, dass H_0 zutrifft.~~
- Auch wenn H_0 zutrifft, ist es nicht sehr unwahrscheinlich, dass unsere Teststatistik einen so extrem erscheinenden Wert annimmt. ✓
-

Angenommen, H_0 konnte durch den Test nicht verworfen werden. Welche Aussagen sind dann richtig?

- ~~Wir müssen die Alternative H_1 verwerfen.~~
- ~~H_0 ist wahr.~~
- ~~H_0 ist wahrscheinlich wahr.~~
- ~~Es ist ungefährlich, davon auszugehen, dass H_0 zutrifft.~~
- Auch wenn H_0 zutrifft, ist es nicht sehr unwahrscheinlich, dass unsere Teststatistik einen so extrem erscheinenden Wert annimmt. ✓
- Die Nullhypothese ist in dieser Hinsicht mit den Daten verträglich.

Angenommen, H_0 konnte durch den Test nicht verworfen werden. Welche Aussagen sind dann richtig?

- ~~Wir müssen die Alternative H_1 verwerfen.~~
- ~~H_0 ist wahr.~~
- ~~H_0 ist wahrscheinlich wahr.~~
- ~~Es ist ungefährlich, davon auszugehen, dass H_0 zutrifft.~~
- Auch wenn H_0 zutrifft, ist es nicht sehr unwahrscheinlich, dass unsere Teststatistik einen so extrem erscheinenden Wert annimmt. ✓
- Die Nullhypothese ist in dieser Hinsicht mit den Daten verträglich. ✓

Frage

Hipparion:
Laubfresser \longrightarrow Grasfresser

Frage

Hipparion:

Laubfresser \longrightarrow Grasfresser

andere Nahrung \longrightarrow andere Zähne?

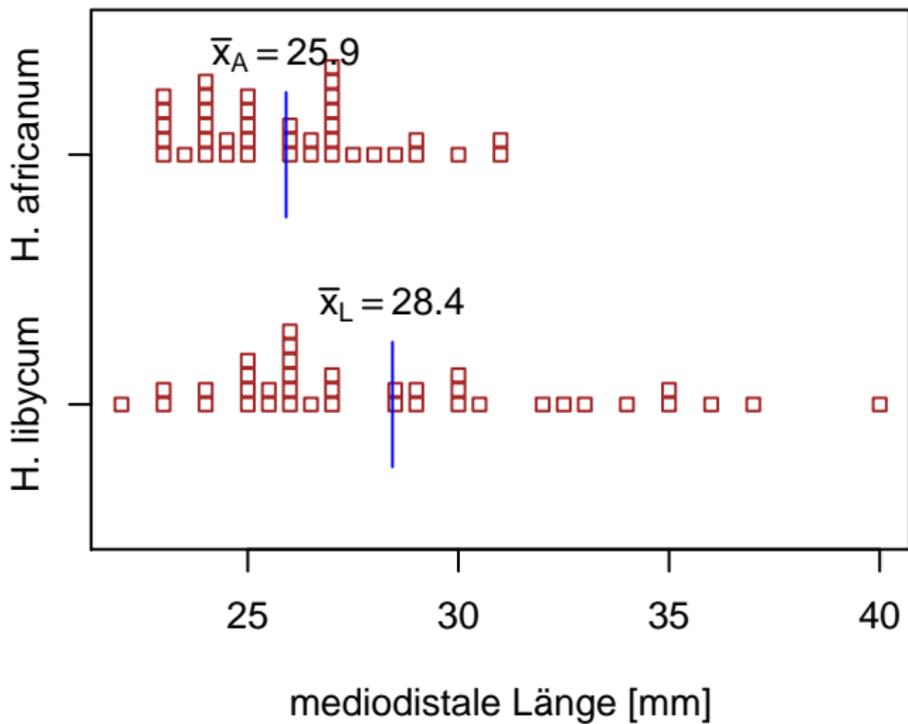
Frage

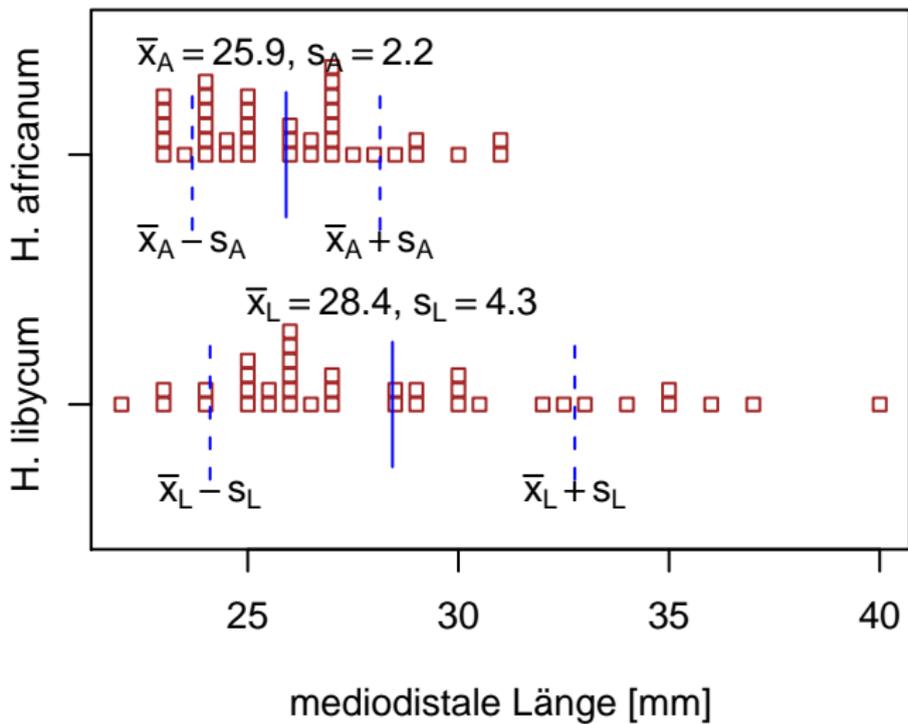
Hipparion:

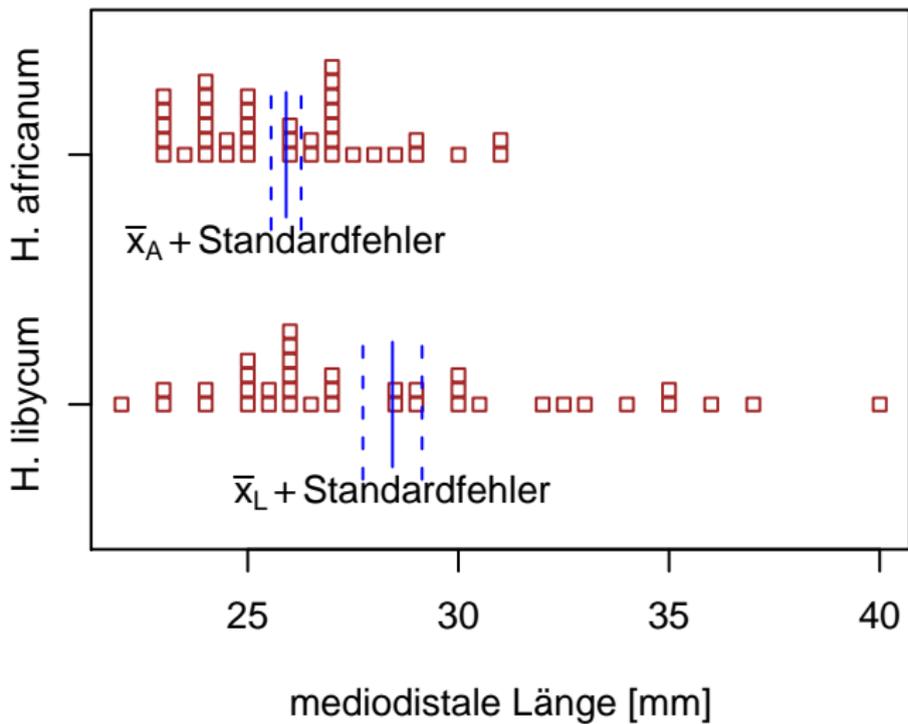
Laubfresser \longrightarrow Grasfresser

andere Nahrung \longrightarrow andere Zähne?

Messungen: mesiodistale Länge







Wir beobachten ($n_A = 39$, $n_L = 38$):

$$\bar{x}_A = 25,9, s_A = 2,2$$

(unser Schätzwert für die Streuung von \bar{x}_A ist also

$$f_A = s_A/\sqrt{n_A} = 2,2/\sqrt{39} = 0,36 \text{ (Standardfehler)})$$

Wir beobachten ($n_A = 39$, $n_L = 38$):

$$\bar{x}_A = 25,9, s_A = 2,2$$

(unser Schätzwert für die Streuung von \bar{x}_A ist also
 $f_A = s_A/\sqrt{n_A} = 2,2/\sqrt{39} = 0,36$ (Standardfehler)),

$$\bar{x}_L = 28,4, s_L = 4,3$$

(unser Schätzwert für die Streuung von \bar{x}_L ist also
 $f_L = s_L/\sqrt{n_L} = 4,3/\sqrt{38} = 0,70$).

Wir beobachten ($n_A = 39$, $n_L = 38$):

$$\bar{x}_A = 25,9, s_A = 2,2$$

(unser Schätzwert für die Streuung von \bar{x}_A ist also
 $f_A = s_A/\sqrt{n_A} = 2,2/\sqrt{39} = 0,36$ (Standardfehler)),

$$\bar{x}_L = 28,4, s_L = 4,3$$

(unser Schätzwert für die Streuung von \bar{x}_L ist also
 $f_L = s_L/\sqrt{n_L} = 4,3/\sqrt{38} = 0,70$).

Ist die beobachtete Abweichung $\bar{x}_L - \bar{x}_A = 2,5$ mit der
Nullhypothese verträglich, dass $\mu_L = \mu_A$?

Ungepaarter t -Test (mit Annahme gleicher Varianzen)

Die „Bilderbuchsituation“: Wir haben zwei unabhängige Stichproben

$x_{1,1}, \dots, x_{1,n_1}$ und $x_{2,1}, \dots, x_{2,n_2}$.

Die $x_{1,i}$ stammen aus einer Normalverteilung mit (unbekanntem) Mittelwert μ_1 und (unbekannter) Varianz $\sigma^2 > 0$, die $x_{2,i}$ aus einer Normalverteilung mit (unbekanntem) Mittelwert μ_2 und derselben Varianz σ^2 .

Ungepaarter t -Test (mit Annahme gleicher Varianzen)

Die „Bilderbuchsituation“: Wir haben zwei unabhängige Stichproben

$x_{1,1}, \dots, x_{1,n_1}$ und $x_{2,1}, \dots, x_{2,n_2}$.

Die $x_{1,i}$ stammen aus einer Normalverteilung mit (unbekanntem) Mittelwert μ_1 und (unbekannter) Varianz $\sigma^2 > 0$, die $x_{2,i}$ aus einer Normalverteilung mit (unbekanntem) Mittelwert μ_2 und derselben Varianz σ^2 .

Seien

$$\bar{x}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} x_{1,i}, \quad \bar{x}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} x_{2,i}$$

die jeweiligen Stichprobenmittelwerte,

$$s_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_{1,i} - \bar{x}_1)^2, \quad s_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (x_{2,i} - \bar{x}_2)^2,$$

die (korrigierten) Stichprobenvarianzen.

Wir möchten die Hypothese $H_0 : \mu_1 = \mu_2$ prüfen.

Wir möchten die Hypothese $H_0 : \mu_1 = \mu_2$ prüfen.

Wenn $\mu_1 = \mu_2$ gilt, so sollte $\bar{x}_1 = \bar{x}_2$ „bis auf

Zufallsschwankungen“ gelten, denn $\mathbb{E}[\bar{X}_1] = \mu_1$, $\mathbb{E}[\bar{X}_2] = \mu_2$.

Wir möchten die Hypothese $H_0 : \mu_1 = \mu_2$ prüfen.

Wenn $\mu_1 = \mu_2$ gilt, so sollte $\bar{x}_1 = \bar{x}_2$ „bis auf
Zufallsschwankungen“ gelten, denn $\mathbb{E}[\bar{X}_1] = \mu_1$, $\mathbb{E}[\bar{X}_2] = \mu_2$.

Was ist die Skala der typischen Schwankungen von $\bar{x}_1 - \bar{x}_2$?

Wir möchten die Hypothese $H_0 : \mu_1 = \mu_2$ prüfen.

Wenn $\mu_1 = \mu_2$ gilt, so sollte $\bar{x}_1 = \bar{x}_2$ „bis auf
Zufallsschwankungen“ gelten, denn $\mathbb{E}[\bar{X}_1] = \mu_1$, $\mathbb{E}[\bar{X}_2] = \mu_2$.

Was ist die Skala der typischen Schwankungen von $\bar{X}_1 - \bar{X}_2$?

$$\text{Var}(\bar{X}_1 - \bar{X}_2) = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

Wir möchten die Hypothese $H_0 : \mu_1 = \mu_2$ prüfen.

Wenn $\mu_1 = \mu_2$ gilt, so sollte $\bar{x}_1 = \bar{x}_2$ „bis auf
Zufallsschwankungen“ gelten, denn $\mathbb{E}[\bar{X}_1] = \mu_1$, $\mathbb{E}[\bar{X}_2] = \mu_2$.

Was ist die Skala der typischen Schwankungen von $\bar{X}_1 - \bar{X}_2$?

$$\text{Var}(\bar{X}_1 - \bar{X}_2) = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

Problem (wie bereits im ein-Stichproben-Fall): Wir kennen σ^2 nicht.

Wir möchten die Hypothese $H_0 : \mu_1 = \mu_2$ prüfen.

Wenn $\mu_1 = \mu_2$ gilt, so sollte $\bar{x}_1 = \bar{x}_2$ „bis auf
Zufallsschwankungen“ gelten, denn $\mathbb{E}[\bar{X}_1] = \mu_1$, $\mathbb{E}[\bar{X}_2] = \mu_2$.

Was ist die Skala der typischen Schwankungen von $\bar{X}_1 - \bar{X}_2$?

$$\text{Var}(\bar{X}_1 - \bar{X}_2) = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

Problem (wie bereits im ein-Stichproben-Fall): Wir kennen σ^2 nicht.

Wir schätzen es im zwei-Stichproben-Fall durch die gepoolte
Stichprobenvarianz

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \left(= \frac{1}{n_1 + n_2 - 2} \left(\sum_{i=1}^{n_1} (x_{1,i} - \bar{x}_1)^2 + \sum_{i=1}^{n_2} (x_{2,i} - \bar{x}_2)^2 \right) \right)$$

und bilden die Teststatistik

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

ungepaarter zwei-Stichproben t-Test (mit der Annahme gleicher Varianzen)

Seien X_1, \dots, X_n und Y_1, \dots, Y_m unabhängige normalverteilte Zufallsvariablen mit der selben Varianz σ^2 . Als **gepoolte Stichprobenvarianz** definieren wir

$$s_p^2 = \frac{(n-1) \cdot s_X^2 + (m-1) \cdot s_Y^2}{m+n-2}.$$

Unter der Nullhypothese gleicher Erwartungswerte $\mu_X = \mu_Y$ folgt die Statistik

$$t = \frac{\bar{X} - \bar{Y}}{s_p \cdot \sqrt{\frac{1}{n} + \frac{1}{m}}}$$

einer t -Verteilung mit $n + m - 2$ mit Freiheitsgraden.

Im Hipparion-Beispiel war $\bar{x}_L = 28,4$, $\bar{x}_A = 25,9$, $s_A = 2,2$,
 $s_A = 4,3$.

Wir finden $t = 3,2$, das 99,5%-Quantil der Student-Vert. mit 75
Freiheitsgraden ist 2,64.

Im Hipparion-Beispiel war $\bar{x}_L = 28,4$, $\bar{x}_A = 25,9$, $s_A = 2,2$,
 $s_A = 4,3$.

Wir finden $t = 3,2$, das 99,5%-Quantil der Student-Vert. mit 75 Freiheitsgraden ist 2,64.

Wir können die Nullhypothese „die mittlere mesiodistale Länge bei *H. lybicum* und bei *H. africanum* sind gleich“ zum Signifikanzniveau 1% ablehnen.

Im Hipparion-Beispiel war $\bar{x}_L = 28,4$, $\bar{x}_A = 25,9$, $s_A = 2,2$,
 $s_A = 4,3$.

Wir finden $t = 3,2$, das 99,5%-Quantil der Student-Vert. mit 75 Freiheitsgraden ist 2,64.

Wir können die Nullhypothese „die mittlere mesiodistale Länge bei *H. lybicum* und bei *H. africanum* sind gleich“ zum Signifikanzniveau 1% ablehnen.

Mögliche Formulierung:

„Die mittlere mesiodistale Länge war signifikant größer (28,4 mm) bei *H. libycum* als bei *H. africanum* (25,9 mm) (t -Test, $\alpha = 0,01$).“

Idee der Rangsummentests

Beobachtungen:

$$X : x_1, x_2, \dots, x_m$$
$$Y : y_1, y_2, \dots, y_n$$

Idee der Rangsummentests

Beobachtungen:

$X : x_1, x_2, \dots, x_m$

$Y : y_1, y_2, \dots, y_n$

- Sortiere alle Beobachtungen der Größe nach.

Idee der Rangsummentests

Beobachtungen:

$$X : x_1, x_2, \dots, x_m$$

$$Y : y_1, y_2, \dots, y_n$$

- Sortiere alle Beobachtungen der Größe nach.
- Bestimme die Ränge der m X -Werte unter allen $m + n$ Beobachtungen.

Idee der Rangsummentests

Beobachtungen:

$$X : x_1, x_2, \dots, x_m$$

$$Y : y_1, y_2, \dots, y_n$$

- Sortiere alle Beobachtungen der Größe nach.
- Bestimme die Ränge der m X -Werte unter allen $m + n$ Beobachtungen.
- Wenn die Nullhypothese zutrifft, sind die m X -Ränge eine rein zufällige Wahl aus $\{1, 2, \dots, m + n\}$.

Idee der Rangsummentests

Beobachtungen:

$X : x_1, x_2, \dots, x_m$

$Y : y_1, y_2, \dots, y_n$

- Sortiere alle Beobachtungen der Größe nach.
- Bestimme die Ränge der m X -Werte unter allen $m + n$ Beobachtungen.
- Wenn die Nullhypothese zutrifft, sind die m X -Ränge eine rein zufällige Wahl aus $\{1, 2, \dots, m + n\}$.
- Berechne die Summe der X -Ränge, prüfe, ob dieser Wert untypisch groß oder klein.

Wilcoxon's Rangsummenstatistik

Beobachtungen:

$X : x_1, x_2, \dots, x_m$

$Y : y_1, y_2, \dots, y_n$



Frank Wilcoxon,
1892–1965

$W =$ Summe der X -Ränge $- (1 + 2 + \dots + m)$
heißt

Wilcoxon's Rangsummenstatistik

Inhalt

- 1 Deskriptive Statistik
- 2 Standardfehler und t-Tests
- 3 Chi-Quadrat-Tests**
 - χ^2 -Test für eine feste Verteilung
 - χ^2 -Test auf Unabhängigkeit (oder Homogenität)
- 4 Konfidenzintervalle
- 5 Lineare Regression
 - Lineare Zusammenhänge
 - t -Test fuer lineare Zusammenhänge

Inhalt

- 1 Deskriptive Statistik
- 2 Standardfehler und t-Tests
- 3 Chi-Quadrat-Tests**
 - χ^2 -Test für eine feste Verteilung
 - χ^2 -Test auf Unabhängigkeit (oder Homogenität)
- 4 Konfidenzintervalle
- 5 Lineare Regression
 - Lineare Zusammenhänge
 - t -Test fuer lineare Zusammenhänge

Szenario:

- Ein Experiment habe r mögliche Ausgänge (z.B. $r = 6$ beim Werfen eines Würfels).

Szenario:

- Ein Experiment habe r mögliche Ausgänge (z.B. $r = 6$ beim Werfen eines Würfels).
- Unter der Nullhypothese H_0 habe Ausgang i Wahrscheinlichkeit p_i .

Szenario:

- Ein Experiment habe r mögliche Ausgänge (z.B. $r = 6$ beim Werfen eines Würfels).
- Unter der Nullhypothese H_0 habe Ausgang i Wahrscheinlichkeit p_i .
- Unter n unabhängigen Wiederholungen des Experiments beobachten wir B_i mal Ausgang i .

Szenario:

- Ein Experiment habe r mögliche Ausgänge (z.B. $r = 6$ beim Werfen eines Würfels).
- Unter der Nullhypothese H_0 habe Ausgang i Wahrscheinlichkeit p_i .
- Unter n unabhängigen Wiederholungen des Experiments beobachten wir B_i mal Ausgang i . Unter H_0 erwarten wir $E_i := \mathbb{E}[B_i] = np_i$ mal Ausgang i zu beobachten.
-

Szenario:

- Ein Experiment habe r mögliche Ausgänge (z.B. $r = 6$ beim Werfen eines Würfels).
- Unter der Nullhypothese H_0 habe Ausgang i Wahrscheinlichkeit p_i .
- Unter n unabhängigen Wiederholungen des Experiments beobachten wir B_i mal Ausgang i . Unter H_0 erwarten wir $E_i := \mathbb{E}[B_i] = np_i$ mal Ausgang i zu beobachten.
-

Frage: Geben die Beobachtungen Anlass, an der Nullhypothese zu zweifeln?

Erwarte $E_i = np_i$ mal Ausgang i , beobachte B_i mal.

Geben diese Beobachtungen Anlass, an der Nullhypothese zu zweifeln?

Erwarte $E_i = np_i$ mal Ausgang i , beobachte B_i mal.

Geben diese Beobachtungen Anlass, an der Nullhypothese zu zweifeln?

Vorgehen:

- Berechne $X^2 = \sum_i \frac{(B_i - E_i)^2}{E_i}$

Erwarte $E_i = np_i$ mal Ausgang i , beobachte B_i mal.

Geben diese Beobachtungen Anlass, an der Nullhypothese zu zweifeln?

Vorgehen:

- Berechne $X^2 = \sum_i \frac{(B_i - E_i)^2}{E_i}$
- X^2 ist unter (approximativ, sofern n genügend groß) χ^2_{r-1} -verteilt („Chi-Quadrat-verteilt mit $r - 1$ Freiheitsgraden“)

Erwarte $E_i = np_i$ mal Ausgang i , beobachte B_i mal.

Geben diese Beobachtungen Anlass, an der Nullhypothese zu zweifeln?

Vorgehen:

- Berechne $X^2 = \sum_i \frac{(B_i - E_i)^2}{E_i}$
- X^2 ist unter (approximativ, sofern n genügend groß) χ^2_{r-1} -verteilt („Chi-Quadrat-verteilt mit $r - 1$ Freiheitsgraden“)
- Lehne H_0 zum Signifikanzniveau α ab, wenn $X^2 \geq q_{1-\alpha}$, wo $q_{1-\alpha}$ das $(1 - \alpha)$ -Quantil der χ^2 -Verteilung mit $r - 1$ Freiheitsgraden ist.

95%-Quantil der χ^2 -Verteilung in Abhängigkeit der Anzahl Freiheitsgrade

F.g.	1	2	3	4	5	6	7	8	9
Quantil	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92

Beispiel: Unter 12.000 Würfeln eines Würfels beobachten wir folgende Häufigkeiten der Augenzahlen:

i	1	2	3	4	5	6
B_i	2014	2000	2017	1925	1998	2046

Beispiel: Unter 12.000 Würfeln eines Würfels beobachten wir folgende Häufigkeiten der Augenzahlen:

i	1	2	3	4	5	6
B_i	2014	2000	2017	1925	1998	2046

Ist der Würfel fair ($H_0: p_1 = \dots = p_6 = 1/6$)?

Beispiel: Unter 12.000 Würfeln eines Würfels beobachten wir folgende Häufigkeiten der Augenzahlen:

i	1	2	3	4	5	6
B_i	2014	2000	2017	1925	1998	2046

Ist der Würfel fair ($H_0: p_1 = \dots = p_6 = 1/6$)?

Es ist $E_1 = \dots = E_6 = 12.000 \cdot 1/6 = 2000$,

$$\begin{aligned} \chi^2 &= \frac{(2014 - 2000)^2}{2000} + \frac{(2000 - 2000)^2}{2000} + \frac{(2017 - 2000)^2}{2000} \\ &+ \frac{(1925 - 2000)^2}{2000} + \frac{(1998 - 2000)^2}{2000} + \frac{(2046 - 2000)^2}{2000} = 4,115. \end{aligned}$$

Beispiel: Unter 12.000 Würfeln eines Würfels beobachten wir folgende Häufigkeiten der Augenzahlen:

i	1	2	3	4	5	6
B_i	2014	2000	2017	1925	1998	2046

Ist der Würfel fair ($H_0: p_1 = \dots = p_6 = 1/6$)?

Es ist $E_1 = \dots = E_6 = 12.000 \cdot 1/6 = 2000$,

$$\chi^2 = \frac{(2014 - 2000)^2}{2000} + \frac{(2000 - 2000)^2}{2000} + \frac{(2017 - 2000)^2}{2000} + \frac{(1925 - 2000)^2}{2000} + \frac{(1998 - 2000)^2}{2000} + \frac{(2046 - 2000)^2}{2000} = 4,115.$$

Das 95%-Quantil der χ^2 -Verteilung mit 5 Freiheitsgraden ist $9,49 > 4,115$, wir lehnen H_0 nicht ab (zum Signifikanzniveau 5%).

95%-Quantil der χ^2 -Verteilung in Abh.keit d. Anz. Freiheitsgrade

F.g.	1	2	3	4	5	6	7	8	9	10
Quantil	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92	18.31

Beispiel: Unter 12.000 Würfeln eines Würfels beobachten wir folgende Häufigkeiten der Augenzahlen:

i	1	2	3	4	5	6
B_i	2014	2000	2017	1925	1998	2046

Ist der Würfel fair ($H_0: p_1 = \dots = p_6 = 1/6$)?

Es ist $E_1 = \dots = E_6 = 12.000 \cdot 1/6 = 2000$,

$$\chi^2 = \frac{(2014 - 2000)^2}{2000} + \frac{(2000 - 2000)^2}{2000} + \frac{(2017 - 2000)^2}{2000} + \frac{(1925 - 2000)^2}{2000} + \frac{(1998 - 2000)^2}{2000} + \frac{(2046 - 2000)^2}{2000} = 4,115.$$

Das 95%-Quantil der χ^2 -Verteilung mit 5 Freiheitsgraden ist $9,49 > 4,115$, wir lehnen H_0 nicht ab (zum Signifikanzniveau 5%).

95%-Quantil der χ^2 -Verteilung in Abh.keit d. Anz. Freiheitsgrade

F.g.	1	2	3	4	5	6	7	8	9	10
Quantil	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92	18.31

Bemerkung: $\chi_5^2([4,115, \infty)) = 0,533$, d.h. wir finden einen p -Wert von 53%, der Test gibt keinen Anlass zu Zweifel an H_0 .

Inhalt

- 1 Deskriptive Statistik
- 2 Standardfehler und t-Tests
- 3 Chi-Quadrat-Tests**
 - χ^2 -Test für eine feste Verteilung
 - χ^2 -Test auf Unabhängigkeit (oder Homogenität)
- 4 Konfidenzintervalle
- 5 Lineare Regression
 - Lineare Zusammenhänge
 - t -Test fuer lineare Zusammenhänge

Beispiel: 48 Teilnehmer eines Management-Kurses entscheiden über Beförderung:

	Weiblich	Männlich	Summe
Befördern	14	21	35
Ablegen	10	3	13
Summe	24	24	48

Beispiel: 48 Teilnehmer eines Management-Kurses entscheiden über Beförderung:

	Weiblich	Männlich	Summe
Befördern	14	21	35
Ablegen	10	3	13
Summe	24	24	48

Kann das Zufall sein? Testen wir H_0 : „Geschlecht und Beförderungsentscheidung sind unabhängig“.

Beispiel: 48 Teilnehmer eines Management-Kurses entscheiden über Beförderung:

	Weiblich	Männlich	Summe
Befördern	14	21	35
Ablegen	10	3	13
Summe	24	24	48

Kann das Zufall sein? Testen wir H_0 : „Geschlecht und Beförderungsentscheidung sind unabhängig“.

Anteil Weiblich= $24/48=0.5$, Anteil befördert= $35/48=0.73$, also erwartete Zahlen unter H_0 :

Beispiel: 48 Teilnehmer eines Management-Kurses entscheiden über Beförderung:

	Weiblich	Männlich	Summe
Befördern	14	21	35
Ablegen	10	3	13
Summe	24	24	48

Kann das Zufall sein? Testen wir H_0 : „Geschlecht und Beförderungsentscheidung sind unabhängig“.

Anteil Weiblich= $24/48=0.5$, Anteil befördert= $35/48=0.73$, also erwartete Zahlen unter H_0 :

	Weiblich	Männlich	Summe
Befördern	17.5 ($= 48 \cdot \frac{24}{48} \cdot \frac{35}{48}$)	17.5 ($= 48 \cdot \frac{24}{48} \cdot \frac{35}{48}$)	35
Ablegen	6.5 ($= 48 \cdot \frac{24}{48} \cdot \frac{13}{48}$)	6.5 ($= 48 \cdot \frac{24}{48} \cdot \frac{13}{48}$)	13
Summe	24	24	48

H_0 : „Geschlecht und Beförderungentscheidung sind unabhängig“

Beobachtete Anzahlen:

	Weiblich	Männlich	Summe
Befördern	14	21	35
Ablegen	10	3	13
Summe	24	24	48

Unter H_0 erwartete Anzahlen:

	Weiblich	Männlich	Summe
Befördern	17.5	17.5	35
Ablegen	6.5	6.5	13
Summe	24	24	48

H_0 : „Geschlecht und Beförderungsentscheidung sind unabhängig“

Beobachtete Anzahlen:

	Weiblich	Männlich	Summe
Befördern	14	21	35
Ablegen	10	3	13
Summe	24	24	48

Unter H_0 erwartete Anzahlen:

	Weiblich	Männlich	Summe
Befördern	17.5	17.5	35
Ablegen	6.5	6.5	13
Summe	24	24	48

Die X^2 -Statistik ist

$$X^2 = \frac{(17.5 - 14)^2}{17.5} + \frac{(21 - 17.5)^2}{17.5} + \frac{(10 - 6.5)^2}{6.5} + \frac{(3 - 6.5)^2}{6.5} = 5.17.$$

H_0 : „Geschlecht und Beförderungentscheidung sind unabhängig“

Beobachtete Anzahlen:

	Weiblich	Männlich	Summe
Befördern	14	21	35
Ablegen	10	3	13
Summe	24	24	48

Unter H_0 erwartete Anzahlen:

	Weiblich	Männlich	Summe
Befördern	17.5	17.5	35
Ablegen	6.5	6.5	13
Summe	24	24	48

Die X^2 -Statistik ist

$$X^2 = \frac{(17.5 - 14)^2}{17.5} + \frac{(21 - 17.5)^2}{17.5} + \frac{(10 - 6.5)^2}{6.5} + \frac{(3 - 6.5)^2}{6.5} = 5.17.$$

Unter H_0 ist X^2 (approximativ) χ^2 -verteilt mit einem Freiheitsgrad $(1 = 4 - 1 - 1 - 1 = (2 - 1) \cdot (2 - 1))$: 4 Zellen, ein Freiheitsgrad geht für die feste Gesamtsumme, einer für das (prinzipiell) unbekannte Geschlechterverhältnis und einer für die (prinzipiell) unbekannte Beförderungswahrscheinlichkeit „verloren“.

95%-Quantil der χ^2 -Verteilung in Abh.keit d. Anz. Freiheitsgrade

F.g.	1	2	3	4	5	6	7	8	9	10
Quantil	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92	18.31

Wir können H_0 zum Signifikanzniveau 5% ablehnen.

H_0 : „Geschlecht und Beförderungentscheidung sind unabhängig“

Beobachtete Anzahlen:

	Weiblich	Männlich	Summe
Befördern	14	21	35
Ablegen	10	3	13
Summe	24	24	48

Unter H_0 erwartete Anzahlen:

	Weiblich	Männlich	Summe
Befördern	17.5	17.5	35
Ablegen	6.5	6.5	13
Summe	24	24	48

Die X^2 -Statistik ist

$$X^2 = \frac{(17.5 - 14)^2}{17.5} + \frac{(21 - 17.5)^2}{17.5} + \frac{(10 - 6.5)^2}{6.5} + \frac{(3 - 6.5)^2}{6.5} = 5.17.$$

Unter H_0 ist X^2 (approximativ) χ^2 -verteilt mit einem Freiheitsgrad ($1 = 4 - 1 - 1 - 1 = (2 - 1) \cdot (2 - 1)$): 4 Zellen, ein Freiheitsgrad geht für die feste Gesamtsumme, einer für das (prinzipiell) unbekannte Geschlechterverhältnis und einer für die (prinzipiell) unbekannte Beförderungswahrscheinlichkeit „verloren“.

95%-Quantil der χ^2 -Verteilung in Abh.keit d. Anz. Freiheitsgrade

F.g.	1	2	3	4	5	6	7	8	9	10
Quantil	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92	18.31

Wir können H_0 zum Signifikanzniveau 5% ablehnen.

(Es ist $\chi_1^2([5.17, \infty)) = 0.023$, d.h. wir finden einen p -Wert von ca. 2%.)

Chi-Quadrat-Test auf Unabhängigkeit, allgemeine Situation:

- 2 Merkmale mit r bzw. s Ausprägungen ($r \times s$ -Kontingenztafel), n Beobachtungen
- Bestimme erwartete Anzahlen unter H_0 als Produkt der (normierten) Zeilen- und Spaltensummen
- X^2 ist unter H_0 (approximativ) χ^2 -verteilt mit $rs - 1 - (r - 1) - (s - 1) = (r - 1)(s - 1)$ Freiheitsgraden.

Inhalt

- 1 Deskriptive Statistik
- 2 Standardfehler und t-Tests
- 3 Chi-Quadrat-Tests
 - χ^2 -Test für eine feste Verteilung
 - χ^2 -Test auf Unabhängigkeit (oder Homogenität)
- 4 Konfidenzintervalle**
- 5 Lineare Regression
 - Lineare Zusammenhänge
 - t -Test fuer lineare Zusammenhänge

Wir beobachten in einer Stichprobe der Größe n aus einer Population X Exemplare mit einer gewissen Eigenschaft (z.B. „ist männlich“) und möchten den (unbekannten) Anteil θ dieser Eigenschaft in der Gesamtpopulation schätzen.

Wir beobachten in einer Stichprobe der Größe n aus einer Population X Exemplare mit einer gewissen Eigenschaft (z.B. „ist männlich“) und möchten den (unbekannten) Anteil θ dieser Eigenschaft in der Gesamtpopulation schätzen.

Der offensichtliche Schätzer ist $\hat{\theta} := \frac{X}{n}$

Wir beobachten in einer Stichprobe der Größe n aus einer Population X Exemplare mit einer gewissen Eigenschaft (z.B. „ist männlich“) und möchten den (unbekannten) Anteil θ dieser Eigenschaft in der Gesamtpopulation schätzen.

Der offensichtliche Schätzer ist $\hat{\theta} := \frac{X}{n}$

Frage: Wie verlässlich ist die Schätzung?

Wir beobachten in einer Stichprobe der Größe n aus einer Population X Exemplare mit einer gewissen Eigenschaft (z.B. „ist männlich“) und möchten den (unbekannten) Anteil θ dieser Eigenschaft in der Gesamtpopulation schätzen.

Der offensichtliche Schätzer ist $\hat{\theta} := \frac{X}{n}$

Frage: Wie verlässlich ist die Schätzung?

Gewünscht: Ein in Abhängigkeit von den Beobachtungen konstruiertes (und möglichst kurzes) Intervall $[\hat{\theta}_u, \hat{\theta}_o]$ mit der Eigenschaft

$$\mathbb{P}_\theta \left([\hat{\theta}_u, \hat{\theta}_o] \text{ überdeckt } \theta \right) \geq 1 - \alpha$$

für *jede Wahl* von θ .

Wir beobachten in einer Stichprobe der Größe n aus einer Population X Exemplare mit einer gewissen Eigenschaft (z.B. „ist männlich“) und möchten den (unbekannten) Anteil θ dieser Eigenschaft in der Gesamtpopulation schätzen.

Der offensichtliche Schätzer ist $\hat{\theta} := \frac{X}{n}$

Frage: Wie verlässlich ist die Schätzung?

Gewünscht: Ein in Abhängigkeit von den Beobachtungen konstruiertes (und möglichst kurzes) Intervall $[\hat{\theta}_u, \hat{\theta}_o]$ mit der Eigenschaft

$$\mathbb{P}_\theta \left([\hat{\theta}_u, \hat{\theta}_o] \text{ überdeckt } \theta \right) \geq 1 - \alpha$$

für *jede Wahl* von θ .

Ein solches Intervall heißt ein *Konfidenzintervall* (zum Irrtumsniveau α), engl. *confidence interval*.

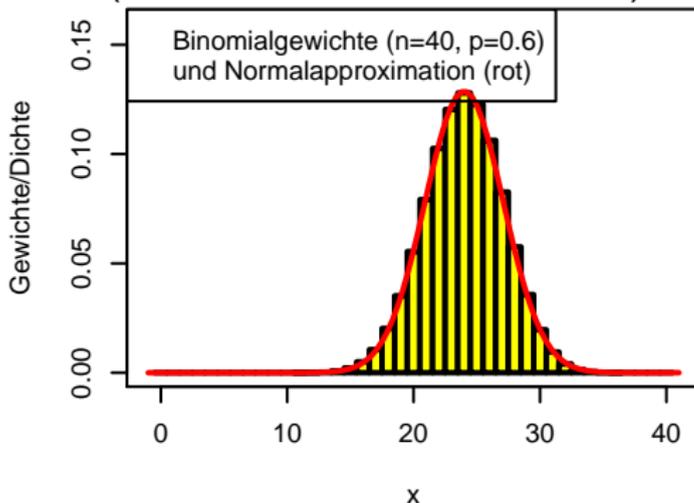
Für gegebenes θ ist X Binomial(n, θ)-verteilt,
 $\mathbb{E}[X] = n\theta$, $\text{Var}[X] = n\theta(1 - \theta)$.

Für gegebenes θ ist X Binomial(n, θ)-verteilt,

$$\mathbb{E}[X] = n\theta, \text{Var}[X] = n\theta(1 - \theta).$$

Für (genügend) großes n ist X ungefähr normalverteilt
mit Mittelwert $n\theta$ und Varianz $n\theta(1 - \theta)$

(„zentraler Grenzwertsatz“):

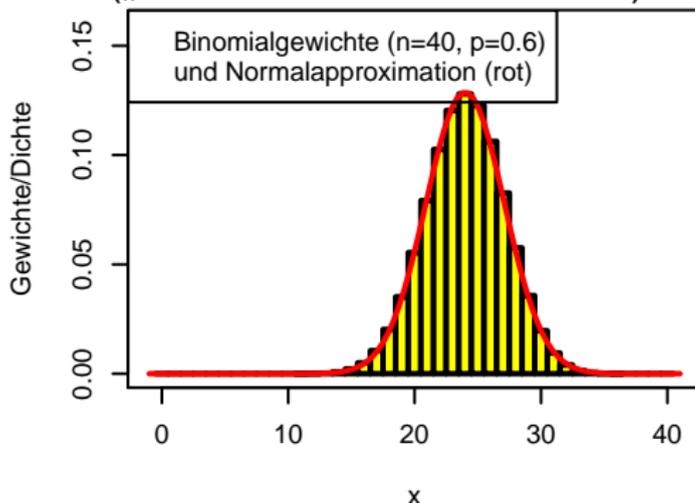


Für gegebenes θ ist X Binomial(n, θ)-verteilt,

$$\mathbb{E}[X] = n\theta, \text{Var}[X] = n\theta(1 - \theta).$$

Für (genügend) großes n ist X ungefähr normalverteilt
mit Mittelwert $n\theta$ und Varianz $n\theta(1 - \theta)$

(„zentraler Grenzwertsatz“):



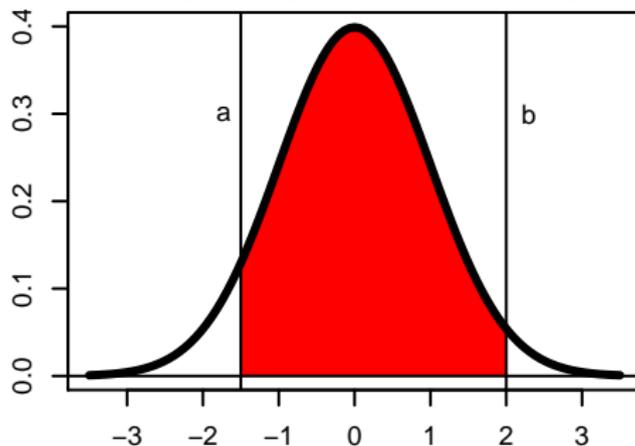
Also ist $\hat{\theta} = \frac{X}{n}$ (ungefähr) normalverteilt mit Mittelwert θ und
Varianz $\frac{1}{n}\theta(1 - \theta)$

$\hat{\theta} = \frac{X}{n}$ ist (ungefähr) normalverteilt
mit Mittelwert θ und Varianz $\frac{1}{n}\theta(1 - \theta)$:

$$\mathbb{P}_{\theta} \left(a \leq \frac{\hat{\theta} - \theta}{\sqrt{\frac{1}{n}\theta(1 - \theta)}} \leq b \right) \approx \mathbb{P}(a \leq Z \leq b)$$

(mit standard-normalverteiletem Z)

Dichte der Standard-Normalverteilung



Man schätzt die (unbekannte) Streuung von \hat{p}
durch $\sqrt{\frac{1}{n}\hat{p}(1 - \hat{p})}$:

Man schätzt die (unbekannte) Streuung von \hat{p}
durch $\sqrt{\frac{1}{n}\hat{p}(1 - \hat{p})}$:

Wähle $z_{1-\alpha/2}$ so dass $\mathbb{P}(-z_{1-\alpha/2} \leq Z \leq z_{1-\alpha/2}) = 1 - \alpha$, dann ist

$$\left[\hat{p} - z_{1-\alpha/2} \frac{\sqrt{\hat{p}(1 - \hat{p})}}{\sqrt{n}}, \hat{p} + z_{1-\alpha/2} \frac{\sqrt{\hat{p}(1 - \hat{p})}}{\sqrt{n}} \right]$$

ein (approximatives) Konfidenzintervall für p zum Niveau $1 - \alpha$.

Man schätzt die (unbekannte) Streuung von \hat{p}
 durch $\sqrt{\frac{1}{n}\hat{p}(1-\hat{p})}$:

Wähle $z_{1-\alpha/2}$ so dass $\mathbb{P}(-z_{1-\alpha/2} \leq Z \leq z_{1-\alpha/2}) = 1 - \alpha$, dann ist

$$\left[\hat{p} - z_{1-\alpha/2} \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}, \hat{p} + z_{1-\alpha/2} \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \right]$$

ein (approximatives) Konfidenzintervall für p zum Niveau $1 - \alpha$.

$z_{1-\alpha/2}$ ist das $(1 - \alpha/2)$ -Quantil der Standardnormalverteilung,
 für die Praxis wichtig sind die Werte

$z_{0,975} \doteq 1,96$ (für $\alpha = 0,05$) und $z_{0,995} \doteq 2,58$ (für $\alpha = 0,01$).

In einem Fang von 53 Porzellankrabben waren 23 Weibchen und 30 Männchen, d.h. der Männchenanteil in der Stichprobe war $30/53 = 0,57$.

In einem Fang von 53 Porzellankrabben waren 23 Weibchen und 30 Männchen, d.h. der Männchenanteil in der Stichprobe war $30/53 = 0,57$.

(Approximatives) 95%-Konfidenzintervall für θ , den Männchenanteil in der Gesamtpopulation:

In einem Fang von 53 Porzellankrabben waren 23 Weibchen und 30 Männchen, d.h. der Männchenanteil in der Stichprobe war $30/53 = 0,57$.

(Approximatives) 95%-Konfidenzintervall für θ , den Männchenanteil in der Gesamtpopulation:

$$\begin{aligned} I &= \left[\frac{30}{53} - 1.96 \sqrt{\frac{(30/53)(23/53)}{53}}, \frac{30}{53} + 1.96 \sqrt{\frac{(30/53)(23/53)}{53}} \right] \\ &= [0.43, 0.70] \end{aligned}$$

Anmerkungen

$\left[\hat{\theta} \pm z_{1-\alpha/2} \frac{\sqrt{\hat{\theta}(1-\hat{\theta})}}{\sqrt{n}} \right]$ ist ein (approximatives) Konfidenzintervall für θ zum Irrtumsniveau α .

Anmerkungen

$\left[\hat{\theta} \pm z_{1-\alpha/2} \frac{\sqrt{\hat{\theta}(1-\hat{\theta})}}{\sqrt{n}} \right]$ ist ein (approximatives) Konfidenzintervall für θ zum Irrtumsniveau α .

- Für die Gültigkeit der Approximation muss n genügend groß und θ nicht zu nahe an 0 oder 1 sein. (Eine häufig zitierte „Faustregel“ ist “ $n\theta(1 - \theta) \geq 9$ ”.)

Anmerkungen

$\left[\hat{\theta} \pm z_{1-\alpha/2} \frac{\sqrt{\hat{\theta}(1-\hat{\theta})}}{\sqrt{n}} \right]$ ist ein (approximatives) Konfidenzintervall für θ zum Irrtumsniveau α .

- Für die Gültigkeit der Approximation muss n genügend groß und θ nicht zu nahe an 0 oder 1 sein. (Eine häufig zitierte „Faustregel“ ist “ $n\theta(1 - \theta) \geq 9$ ”.)
- Die Philosophie der Konfidenzintervalle entstammt der *frequentistischen* Interpretation der Statistik: Für jede Wahl des Parameters θ würden wir bei häufiger Wiederholung des Experiments finden, dass in (ca.) $(1 - \alpha) \cdot 100\%$ der Fälle das (zufällige) Konfidenzintervall den „wahren“ (festen) Parameter θ überdeckt.

Anmerkungen

$\left[\hat{\theta} \pm z_{1-\alpha/2} \frac{\sqrt{\hat{\theta}(1-\hat{\theta})}}{\sqrt{n}} \right]$ ist ein (approximatives) Konfidenzintervall für θ zum Irrtumsniveau α .

- Für die Gültigkeit der Approximation muss n genügend groß und θ nicht zu nahe an 0 oder 1 sein. (Eine häufig zitierte „Faustregel“ ist „ $n\theta(1 - \theta) \geq 9$ “.)
- Die Philosophie der Konfidenzintervalle entstammt der *frequentistischen* Interpretation der Statistik: Für jede Wahl des Parameters θ würden wir bei häufiger Wiederholung des Experiments finden, dass in (ca.) $(1 - \alpha) \cdot 100\%$ der Fälle das (zufällige) Konfidenzintervall den „wahren“ (festen) Parameter θ überdeckt.
- Formulierungen, die sich auf eine Wahrscheinlichkeitsverteilung des Parameters θ beziehen (beispielsweise: „~~Wie wahrscheinlich ist es, dass $\theta \leq 0,3$?~~“), sind in der frequentistischen Interpretation sinnlos. (Dies ist anders in der *Bayesschen Interpretation*.)

Student-Konfidenzintervall für den Mittelwert

Sei $\alpha \in (0, 1)$ (oft $\alpha = 0.05$), $t_{n-1, 1-\alpha/2}$ das $(1 - \alpha/2)$ -Quantil der Student-Verteilung mit $n - 1$ Freiheitsgraden (d.h. für Student- $(n - 1)$ -verteiltes T_{n-1} gilt $\mathbb{P}(T_{n-1} \leq t_{n-1, 1-\alpha/2}) = 1 - \alpha/2$).

Student-Konfidenzintervall für den Mittelwert

Sei $\alpha \in (0, 1)$ (oft $\alpha = 0.05$), $t_{n-1, 1-\alpha/2}$ das $(1 - \alpha/2)$ -Quantil der Student-Verteilung mit $n - 1$ Freiheitsgraden (d.h. für Student- $(n - 1)$ -verteiltes T_{n-1} gilt $\mathbb{P}(T_{n-1} \leq t_{n-1, 1-\alpha/2}) = 1 - \alpha/2$). Dann ist die Wahrscheinlichkeit, dass der wahre Mittelwert μ von dem Intervall

$$I := \left[\bar{X} - t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}}, \bar{X} + t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}} \right]$$

überdeckt wird, (approximativ*) $1 - \alpha$.

I heißt ein **Konfidenzintervall für μ zum Niveau $1 - \alpha$** oder kurz ein $(1 - \alpha)$ -Konfidenzintervall.

Student-Konfidenzintervall für den Mittelwert

Sei $\alpha \in (0, 1)$ (oft $\alpha = 0.05$), $t_{n-1, 1-\alpha/2}$ das $(1 - \alpha/2)$ -Quantil der Student-Verteilung mit $n - 1$ Freiheitsgraden (d.h. für Student- $(n - 1)$ -verteiltes T_{n-1} gilt $\mathbb{P}(T_{n-1} \leq t_{n-1, 1-\alpha/2}) = 1 - \alpha/2$). Dann ist die Wahrscheinlichkeit, dass der wahre Mittelwert μ von dem Intervall

$$I := \left[\bar{X} - t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}}, \bar{X} + t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}} \right]$$

überdeckt wird, (approximativ*) $1 - \alpha$.

I heißt ein **Konfidenzintervall für μ zum Niveau $1 - \alpha$** oder kurz ein $(1 - \alpha)$ -Konfidenzintervall.

* Die Aussage ist wörtlich korrekt, wenn die Daten als normalverteilt angenommen werden dürfen, die Näherung ist sehr gut und für die Praxis ausreichend, wenn die Daten ungefähr symmetrisch und glockenförmig verteilt sind oder n genügend groß.

Dualität von Konfidenzintervallen und zweiseitigen Tests

Beispiel:

Stichprobe der Größe $n = 29$, in der wir Stichprobenmittelwert $\bar{x} = 3.23$ und Stichprobenstreuung $s = 0.9$ beobachtet haben. Konfidenzintervall für den wahren Mittelwert μ zum Irrtumsniveau 5% ($t_{28,0.975} = 2.048$)

$$\left[\bar{x} - t_{n-1,0.975} \frac{s}{\sqrt{n}}, \bar{x} + t_{n-1,0.975} \frac{s}{\sqrt{n}} \right] = [2.88, 3.58]$$

Dualität von Konfidenzintervallen und zweiseitigen Tests

Beispiel:

Stichprobe der Größe $n = 29$, in der wir Stichprobenmittelwert $\bar{x} = 3.23$ und Stichprobenstreuung $s = 0.9$ beobachtet haben. Konfidenzintervall für den wahren Mittelwert μ zum Irrtumsniveau 5% ($t_{28,0.975} = 2.048$)

$$\left[\bar{x} - t_{n-1,0.975} \frac{s}{\sqrt{n}}, \bar{x} + t_{n-1,0.975} \frac{s}{\sqrt{n}} \right] = [2.88, 3.58]$$

Nehmen wir an, wir wollten (anhand derselben Beobachtungen) die Nullhypothese „ $\mu = 3.2$ “ zum Signifikanzniveau 5% testen:

Dualität von Konfidenzintervallen und zweiseitigen Tests

Beispiel:

Stichprobe der Größe $n = 29$, in der wir Stichprobenmittelwert $\bar{x} = 3.23$ und Stichprobenstreuung $s = 0.9$ beobachtet haben. Konfidenzintervall für den wahren Mittelwert μ zum Irrtumsniveau 5% ($t_{28,0.975} = 2.048$)

$$\left[\bar{x} - t_{n-1,0.975} \frac{s}{\sqrt{n}}, \bar{x} + t_{n-1,0.975} \frac{s}{\sqrt{n}} \right] = [2.88, 3.58]$$

Nehmen wir an, wir wollten (anhand derselben Beobachtungen) die Nullhypothese „ $\mu = 3.2$ “ zum Signifikanzniveau 5% testen:

Verwende den t -Test: $t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = 0.18$

Dualität von Konfidenzintervallen und zweiseitigen Tests

Beispiel:

Stichprobe der Größe $n = 29$, in der wir Stichprobenmittelwert $\bar{x} = 3.23$ und Stichprobenstreuung $s = 0.9$ beobachtet haben. Konfidenzintervall für den wahren Mittelwert μ zum Irrtumsniveau 5% ($t_{28,0.975} = 2.048$)

$$\left[\bar{x} - t_{n-1,0.975} \frac{s}{\sqrt{n}}, \bar{x} + t_{n-1,0.975} \frac{s}{\sqrt{n}} \right] = [2.88, 3.58]$$

Nehmen wir an, wir wollten (anhand derselben Beobachtungen) die Nullhypothese „ $\mu = 3.2$ “ zum Signifikanzniveau 5% testen:

Verwende den t -Test: $t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = 0.18$,

$|t| = |0.18| \leq t_{28,0.975} = 2.048$, d.h. wir würden die Nullhypothese nicht verwerfen (der p -Wert ist $0.86 (= \mathbb{P}_{\mu=3.2}(|t| \geq 0.18))$).

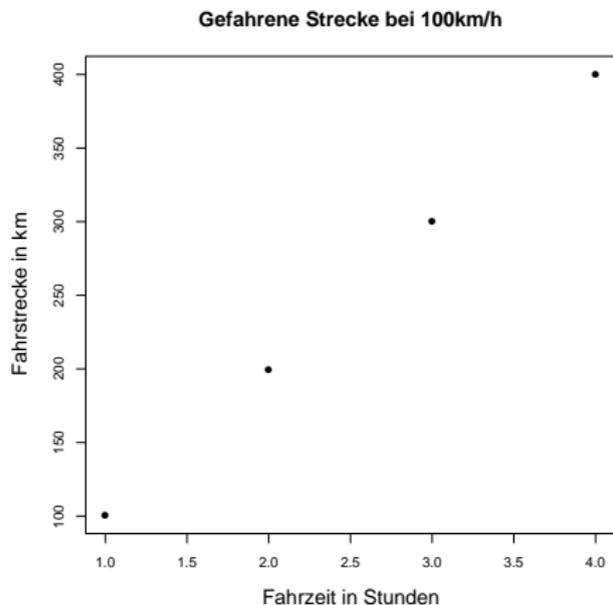
Inhalt

- 1 Deskriptive Statistik
- 2 Standardfehler und t-Tests
- 3 Chi-Quadrat-Tests
 - χ^2 -Test für eine feste Verteilung
 - χ^2 -Test auf Unabhängigkeit (oder Homogenität)
- 4 Konfidenzintervalle
- 5 **Lineare Regression**
 - **Lineare Zusammenhänge**
 - **t-Test fuer lineare Zusammenhänge**

Inhalt

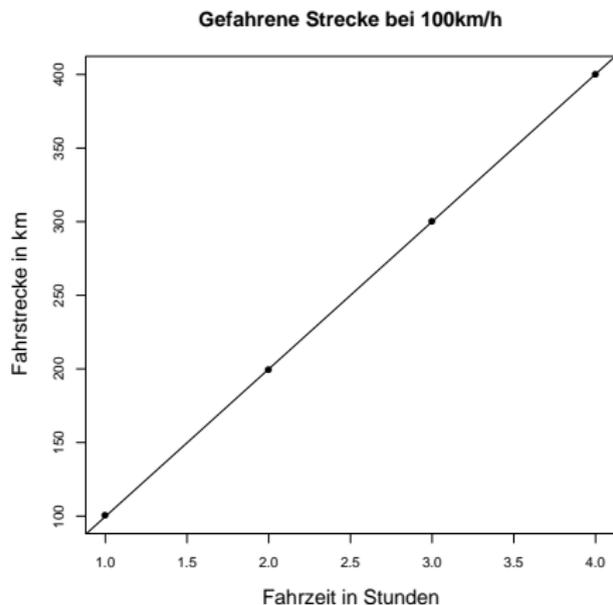
- 1 Deskriptive Statistik
- 2 Standardfehler und t-Tests
- 3 Chi-Quadrat-Tests
 - χ^2 -Test für eine feste Verteilung
 - χ^2 -Test auf Unabhängigkeit (oder Homogenität)
- 4 Konfidenzintervalle
- 5 **Lineare Regression**
 - **Lineare Zusammenhänge**
 - *t*-Test fuer lineare Zusammenhänge

Erst mal ohne Zufallsschwankungen



Gemessene Fahrstrecke bei exakt 100 km/h

Erst mal ohne Zufallsschwankungen



Gemessene Fahrtstrecke bei exakt 100 km/h

Zusammenhang:
Strecke s in km, Zeit t in Stunden

$$s = 100 \frac{\text{km}}{\text{h}} \cdot t$$

Problem und Lösung:

- Problem: Strecke ist schwer zu messen.

Problem und Lösung:

- Problem: Strecke ist schwer zu messen.
- Beobachtung: Zeit ist leicht zu messen (Blick auf Uhr)

Problem und Lösung:

- Problem: Strecke ist schwer zu messen.
- Beobachtung: Zeit ist leicht zu messen (Blick auf Uhr)
- Lösung: Linearer Zusammenhang zwischen Strecke und Zeit ermöglicht leichte Berechnung der Strecke (\rightsquigarrow Problem gelöst)



photo (c) by Jörg Hempel

Englisch: Grif-
fon Vulture
Gypus fulvus
Gänsegeier

Beispiel: Herzfrequenz und Stoffwechselrate beim Gänsegeier

- Frage: Was ist die Stoffwechselrate bei Gänsegeiern im Alltag (zB im Flug)?

Beispiel: Herzfrequenz und Stoffwechselrate beim Gänsegeier

- Frage: Was ist die Stoffwechselrate bei Gänsegeiern im Alltag (zB im Flug)?
- Problem: Stoffwechselrate ist aufwändig zu messen (eigentlich nur im Labor)

Beispiel: Herzfrequenz und Stoffwechselrate beim Gänsegeier

- Frage: Was ist die Stoffwechselrate bei Gänsegeiern im Alltag (zB im Flug)?
- Problem: Stoffwechselrate ist aufwändig zu messen (eigentlich nur im Labor)
- Beobachtung: Herzfrequenz ist leicht zu messen.

Beispiel: Herzfrequenz und Stoffwechselrate beim Gänsegeier

- Frage: Was ist die Stoffwechselrate bei Gänsegeiern im Alltag (zB im Flug)?
- Problem: Stoffwechselrate ist aufwändig zu messen (eigentlich nur im Labor)
- Beobachtung: Herzfrequenz ist leicht zu messen.
- Lösung: Nutze linearen Zusammenhang zwischen Stoffwechselrate und Herzfrequenz.

Beispiel: Herzfrequenz und Stoffwechselrate beim Gänsegeier

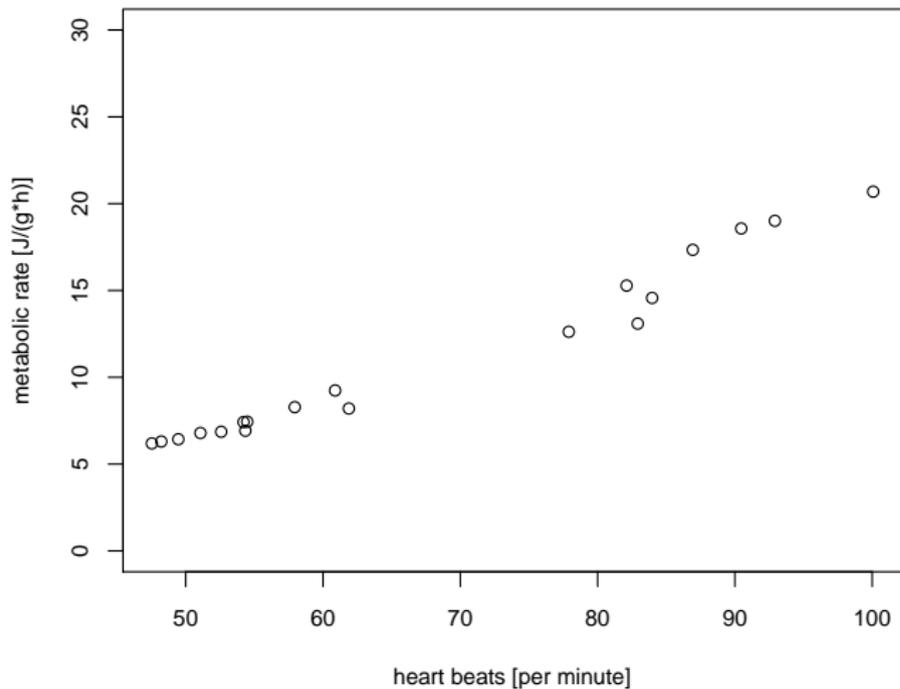
- Frage: Was ist die Stoffwechselrate bei Gänsegeiern im Alltag (zB im Flug)?
- Problem: Stoffwechselrate ist aufwändig zu messen (eigentlich nur im Labor)
- Beobachtung: Herzfrequenz ist leicht zu messen.
- Lösung: Nutze linearen Zusammenhang zwischen Stoffwechselrate und Herzfrequenz.
- Komplikation: Der lineare Zusammenhang ist nicht deterministisch, sondern zufallsbehaftet auf Grund von Messfehlern und da Stoffwechselrate von der „Tagesform“ abhängt.

Beispiel: Herzfrequenz und Stoffwechselrate beim Gänsegeier

- Frage: Was ist die Stoffwechselrate bei Gänsegeiern im Alltag (zB im Flug)?
- Problem: Stoffwechselrate ist aufwändig zu messen (eigentlich nur im Labor)
- Beobachtung: Herzfrequenz ist leicht zu messen.
- Lösung: Nutze linearen Zusammenhang zwischen Stoffwechselrate und Herzfrequenz.
- Komplikation: Der lineare Zusammenhang ist nicht deterministisch, sondern zufallsbehaftet auf Grund von Messfehlern und da Stoffwechselrate von der „Tagesform“ abhängt.
- Lineare Regression löst diese Komplikation in Wohlgefallen auf.

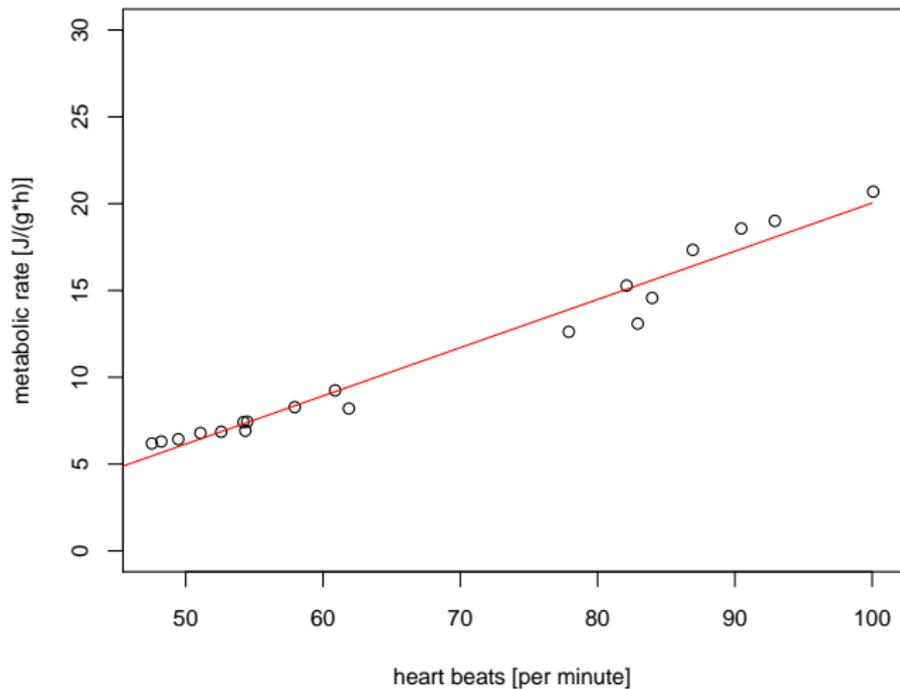
Beispiel: Herzfrequenz und Stoffwechselrate beim Gänsegeier

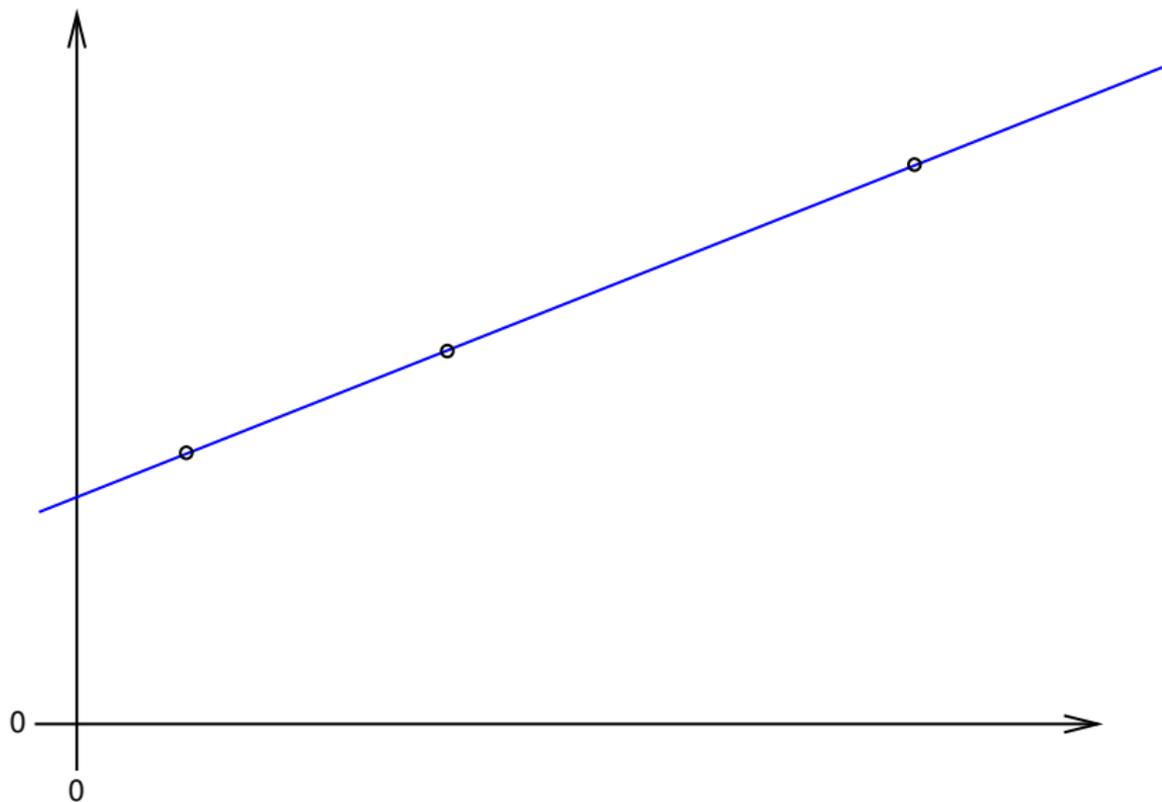
griffon vulture, 17.05.99, 16 degrees C

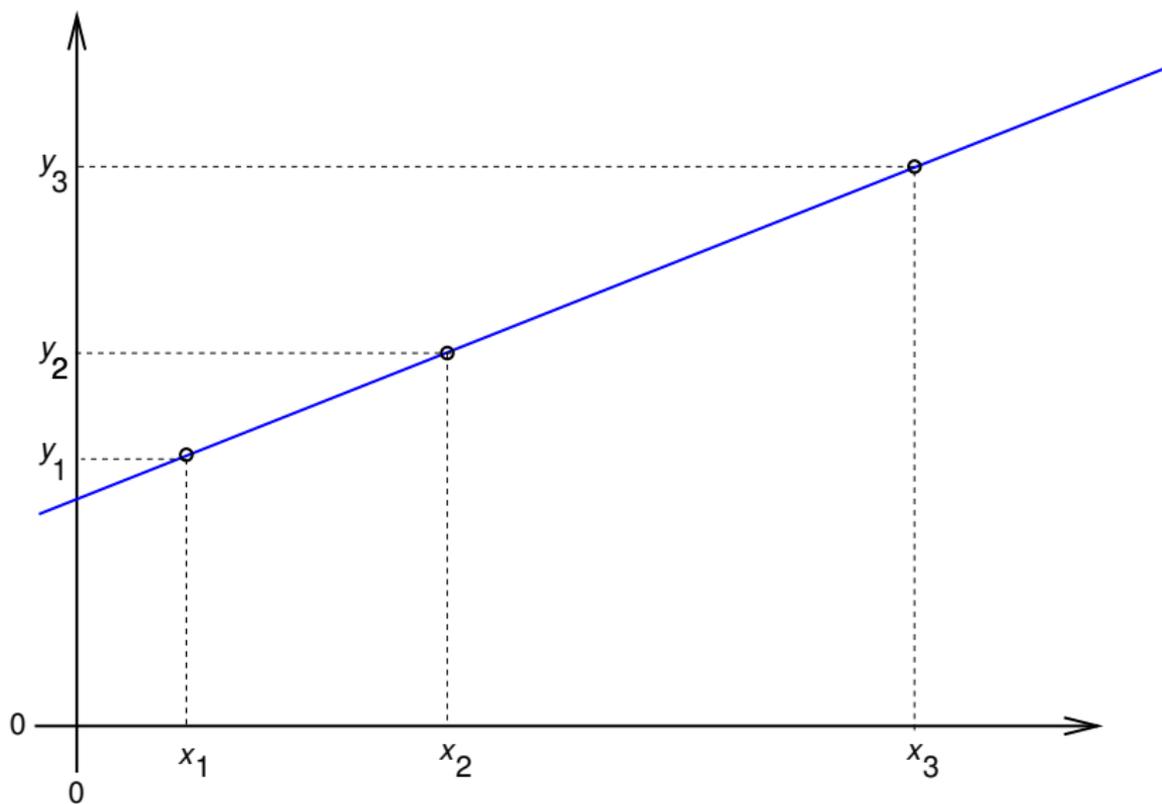


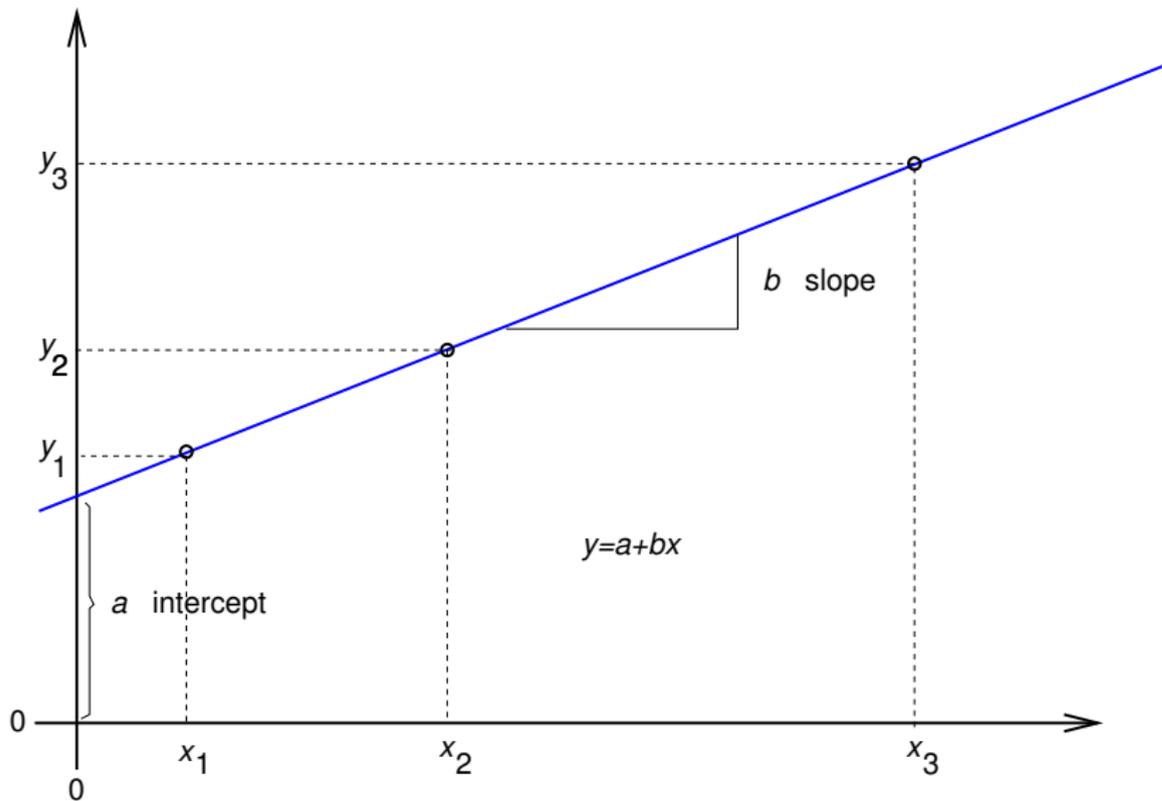
Beispiel: Herzfrequenz und Stoffwechselrate beim Gänsegeier

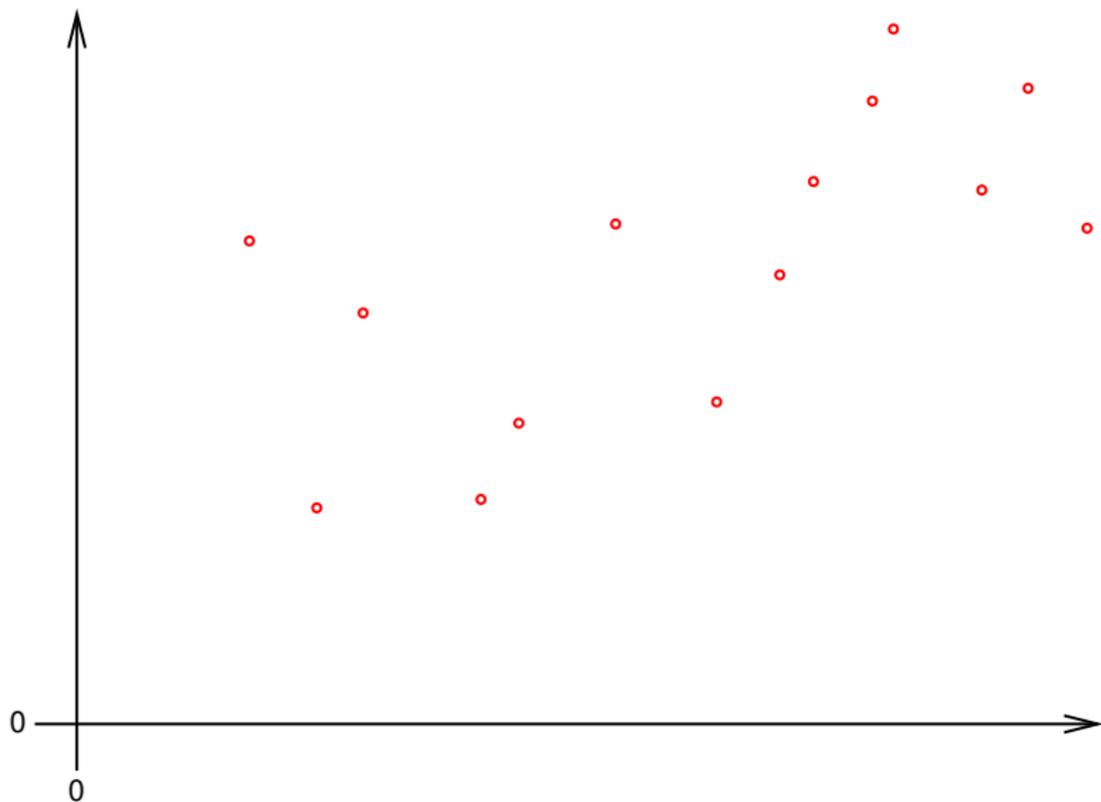
griffon vulture, 17.05.99, 16 degrees C

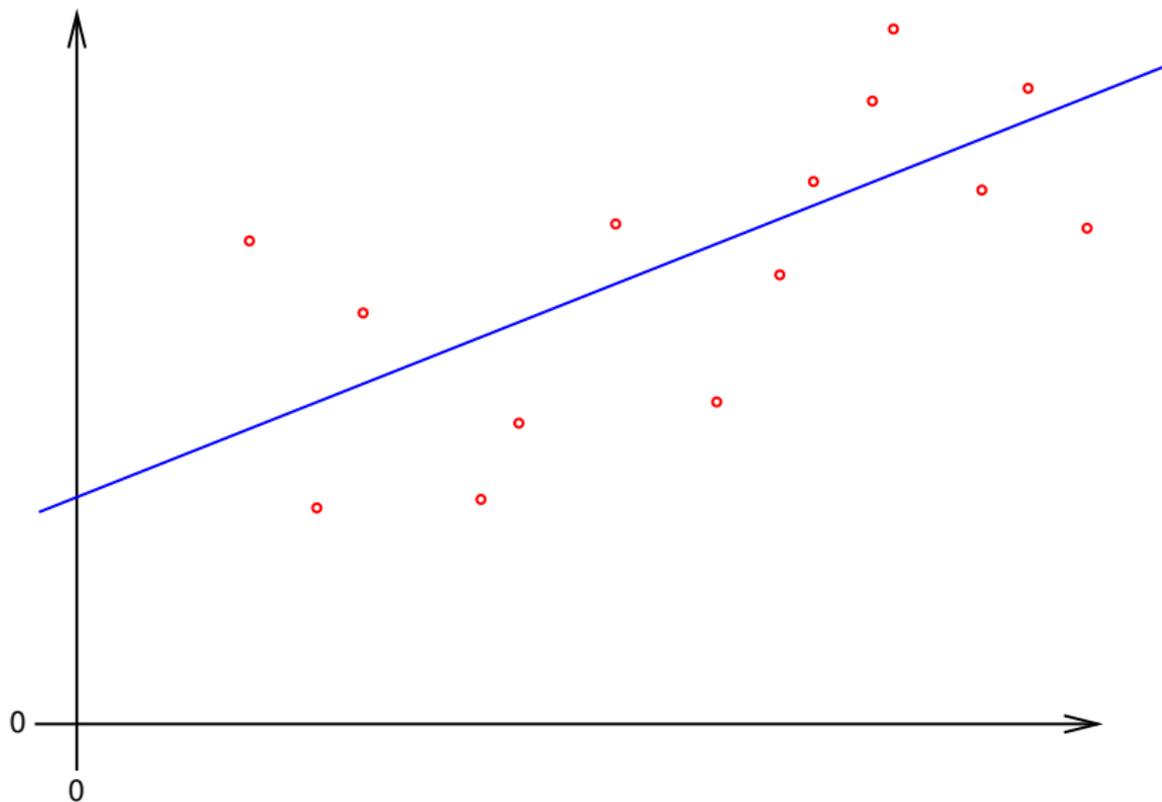


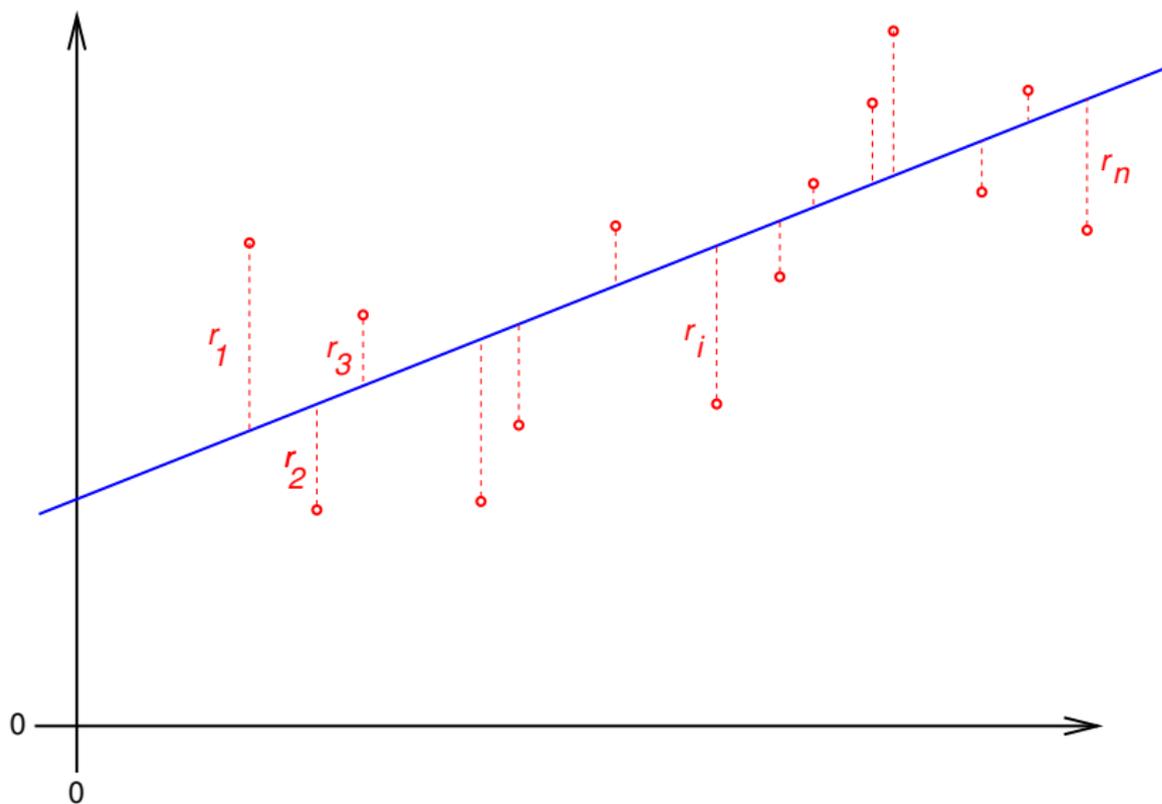


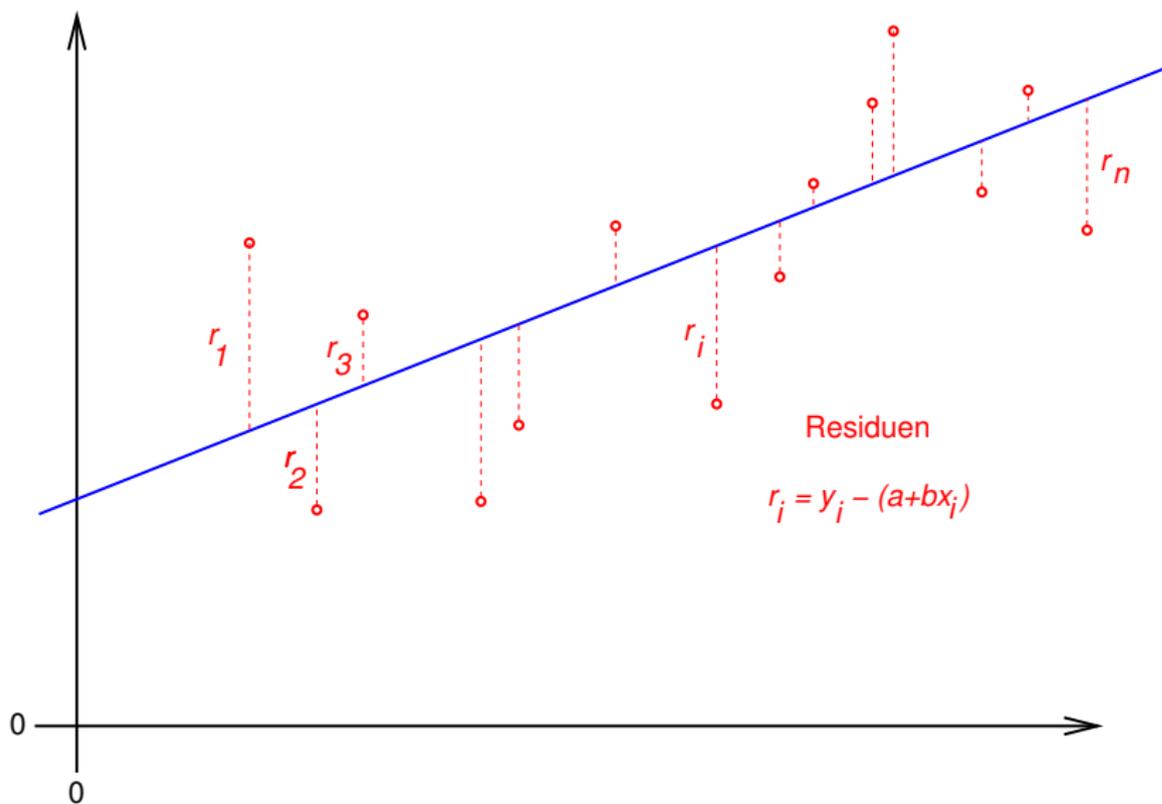


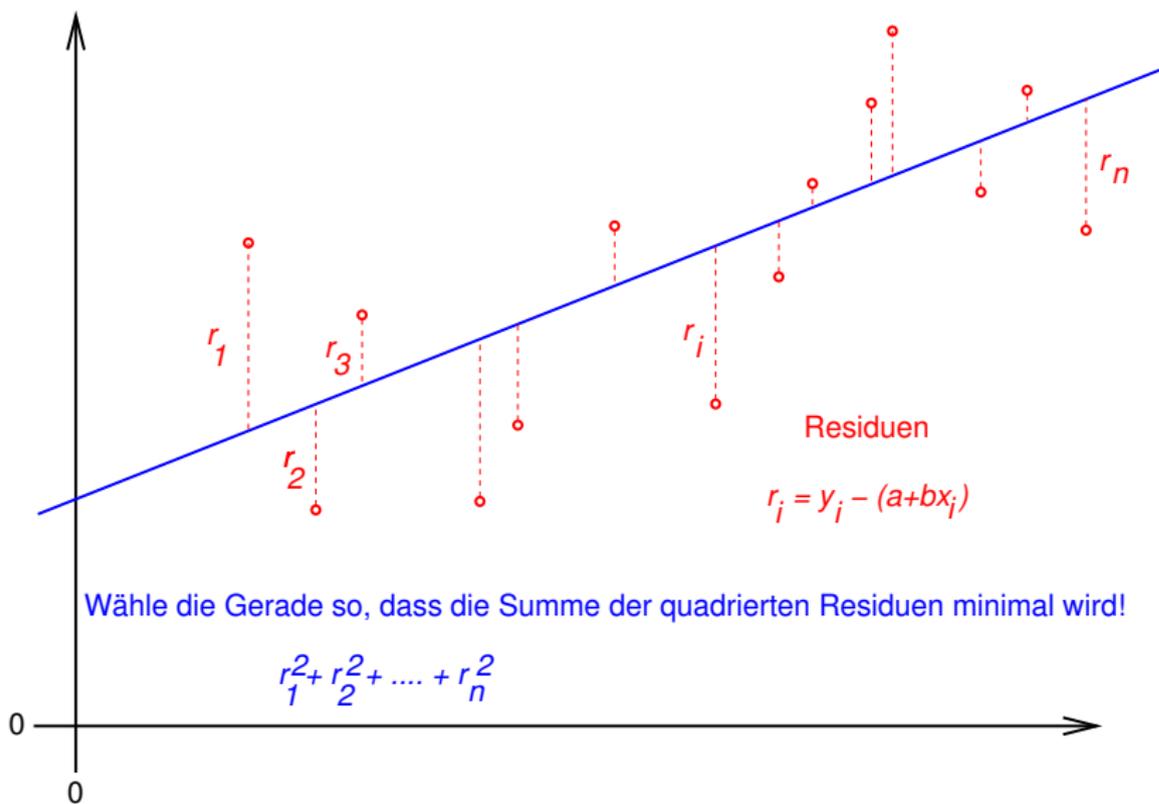












Definiere die Regressionsgerade

$$y = \hat{a} + \hat{b} \cdot x$$

durch die Minimierung der Summe der quadrierten Residuen:

$$(\hat{a}, \hat{b}) = \arg \min_{(a,b)} \sum_i (y_i - (a + b \cdot x_i))^2$$

Dahinter steckt die Modellvorstellung, dass Werte a, b existieren, so dass für alle Datenpaare (x_i, y_i) gilt

$$y_i = a + b \cdot x_i + \varepsilon_i,$$

wobei alle ε_i unabhängig und normalverteilt sind und alle dieselbe Varianz σ^2 haben.

gegebene Daten:

Y	X
y_1	x_1
y_2	x_2
y_3	x_3
\vdots	\vdots
y_n	x_n

gegebene Daten:

Y	X
y_1	x_1
y_2	x_2
y_3	x_3
\vdots	\vdots
y_n	x_n

Modell: es gibt Zahlen
 a, b, σ^2 , so dass

$$y_1 = a + b \cdot x_1 + \varepsilon_1$$

$$y_2 = a + b \cdot x_2 + \varepsilon_2$$

$$y_3 = a + b \cdot x_3 + \varepsilon_3$$

$$\vdots \quad \quad \quad \vdots$$

$$y_n = a + b \cdot x_n + \varepsilon_n$$

gegebene Daten:

Y	X
y_1	x_1
y_2	x_2
y_3	x_3
\vdots	\vdots
y_n	x_n

Modell: es gibt Zahlen
 a, b, σ^2 , so dass

$$\begin{aligned}y_1 &= a + b \cdot x_1 + \varepsilon_1 \\y_2 &= a + b \cdot x_2 + \varepsilon_2 \\y_3 &= a + b \cdot x_3 + \varepsilon_3 \\&\vdots \\y_n &= a + b \cdot x_n + \varepsilon_n\end{aligned}$$

Dabei sind $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ unabhängig $\sim \mathcal{N}(0, \sigma^2)$.

gegebene Daten:

Y	X
y_1	x_1
y_2	x_2
y_3	x_3
\vdots	\vdots
y_n	x_n

Modell: es gibt Zahlen
 a, b, σ^2 , so dass

$$\begin{aligned}y_1 &= a + b \cdot x_1 + \varepsilon_1 \\y_2 &= a + b \cdot x_2 + \varepsilon_2 \\y_3 &= a + b \cdot x_3 + \varepsilon_3 \\&\vdots \\y_n &= a + b \cdot x_n + \varepsilon_n\end{aligned}$$

Dabei sind $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ unabhängig $\sim \mathcal{N}(0, \sigma^2)$.

$\Rightarrow y_1, y_2, \dots, y_n$ sind unabhängig $y_i \sim \mathcal{N}(a + b \cdot x_i, \sigma^2)$.

gegebene Daten:

Y	X
y_1	x_1
y_2	x_2
y_3	x_3
\vdots	\vdots
y_n	x_n

Modell: es gibt Zahlen
 a, b, σ^2 , so dass

$$\begin{aligned}
 y_1 &= a + b \cdot x_1 + \varepsilon_1 \\
 y_2 &= a + b \cdot x_2 + \varepsilon_2 \\
 y_3 &= a + b \cdot x_3 + \varepsilon_3 \\
 &\vdots \\
 y_n &= a + b \cdot x_n + \varepsilon_n
 \end{aligned}$$

Dabei sind $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ unabhängig $\sim \mathcal{N}(0, \sigma^2)$.

$\Rightarrow y_1, y_2, \dots, y_n$ sind unabhängig $y_i \sim \mathcal{N}(a + b \cdot x_i, \sigma^2)$.

a, b, σ^2 sind unbekannt, aber **nicht zufällig**.

Wir schätzen a und b , indem wir

$$(\hat{a}, \hat{b}) := \arg \min_{(a,b)} \sum_i (y_i - (a + b \cdot x_i))^2 \quad \text{berechnen.}$$

Wir schätzen a und b , indem wir

$$(\hat{a}, \hat{b}) := \arg \min_{(a,b)} \sum_i (y_i - (a + b \cdot x_i))^2 \quad \text{berechnen.}$$

Theorem

\hat{a} und \hat{b} sind gegeben durch

$$\hat{b} = \frac{\text{cov}(x, y)}{\sigma_x^2} = \frac{\sum_i (y_i - \bar{y}) \cdot (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2} = \frac{\sum_i y_i \cdot (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

und

$$\hat{a} = \bar{y} - \hat{b} \cdot \bar{x}.$$

Wir schätzen a und b , indem wir

$$(\hat{a}, \hat{b}) := \arg \min_{(a,b)} \sum_i (y_i - (a + b \cdot x_i))^2 \quad \text{berechnen.}$$

Theorem

\hat{a} und \hat{b} sind gegeben durch

$$\hat{b} = \frac{\text{cov}(x, y)}{\sigma_x^2} = \frac{\sum_i (y_i - \bar{y}) \cdot (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2} = \frac{\sum_i y_i \cdot (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

und

$$\hat{a} = \bar{y} - \hat{b} \cdot \bar{x}.$$

Bitte merken:

Die Gerade $y = \hat{a} + \hat{b} \cdot x$ geht genau durch den Schwerpunkt der Punktwolke $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.

Inhalt

- 1 Deskriptive Statistik
- 2 Standardfehler und t-Tests
- 3 Chi-Quadrat-Tests
 - χ^2 -Test für eine feste Verteilung
 - χ^2 -Test auf Unabhängigkeit (oder Homogenität)
- 4 Konfidenzintervalle
- 5 **Lineare Regression**
 - Lineare Zusammenhänge
 - **t-Test fuer lineare Zusammenhänge**

Modell:

$$Y = a + b \cdot X + \varepsilon \quad \text{mit } \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

Modell:

$$Y = a + b \cdot X + \varepsilon \quad \text{mit } \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

Wie berechnet man die Signifikanz eines Zusammenhangs zwischen dem *erklärenden Merkmal* X und der *Zielgröße* Y ?

Modell:

$$Y = a + b \cdot X + \varepsilon \quad \text{mit } \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

Wie berechnet man die Signifikanz eines Zusammenhangs zwischen dem *erklärenden Merkmal* X und der *Zielgröße* Y ?

Anders formuliert: Mit welchem Test können wir der Nullhypothese $b = 0$ zu Leibe rücken?

Modell:

$$Y = a + b \cdot X + \varepsilon \quad \text{mit } \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

Wie berechnet man die Signifikanz eines Zusammenhangs zwischen dem *erklärenden Merkmal* X und der *Zielgröße* Y ?

Anders formuliert: Mit welchem Test können wir der Nullhypothese $b = 0$ zu Leibe rücken?

Wir haben b durch \hat{b} geschätzt (und gehen jetzt mal von $\hat{b} \neq 0$ aus). Könnte das wahre b auch 0 sein?

Modell:

$$Y = a + b \cdot X + \varepsilon \quad \text{mit } \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

Wie berechnet man die Signifikanz eines Zusammenhangs zwischen dem *erklärenden Merkmal* X und der *Zielgröße* Y ?

Anders formuliert: Mit welchem Test können wir der Nullhypothese $b = 0$ zu Leibe rücken?

Wir haben b durch \hat{b} geschätzt (und gehen jetzt mal von $\hat{b} \neq 0$ aus). Könnte das wahre b auch 0 sein?

Wie groß ist der Standardfehler unserer Schätzung \hat{b} ?

$$y_i = a + b \cdot x_i + \varepsilon \quad \text{mit } \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

nicht zufällig: a, b, x_i, σ^2 zufällig: ε, y_i

$$\text{var}(y_i) = \text{var}(a + b \cdot x_i + \varepsilon) = \text{var}(\varepsilon) = \sigma^2$$

und y_1, y_2, \dots, y_n sind stochastisch unabhängig.

$$y_i = a + b \cdot x_i + \varepsilon \quad \text{mit } \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

nicht zufällig: a, b, x_i, σ^2 zufällig: ε, y_i

$$\text{var}(y_i) = \text{var}(a + b \cdot x_i + \varepsilon) = \text{var}(\varepsilon) = \sigma^2$$

und y_1, y_2, \dots, y_n sind stochastisch unabhängig.

$$\hat{b} = \frac{\sum_i y_i (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

$$y_i = a + b \cdot x_i + \varepsilon \quad \text{mit } \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

nicht zufällig: a, b, x_i, σ^2 zufällig: ε, y_i

$$\text{var}(y_i) = \text{var}(a + b \cdot x_i + \varepsilon) = \text{var}(\varepsilon) = \sigma^2$$

und y_1, y_2, \dots, y_n sind stochastisch unabhängig.

$$\hat{b} = \frac{\sum_i y_i (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

$$\text{var}(\hat{b}) = \text{var}\left(\frac{\sum_i y_i (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}\right) = \frac{\text{var}(\sum_i y_i (x_i - \bar{x}))}{(\sum_i (x_i - \bar{x})^2)^2}$$

$$y_i = a + b \cdot x_i + \varepsilon \quad \text{mit } \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

nicht zufällig: a, b, x_i, σ^2 zufällig: ε, y_i

$$\text{var}(y_i) = \text{var}(a + b \cdot x_i + \varepsilon) = \text{var}(\varepsilon) = \sigma^2$$

und y_1, y_2, \dots, y_n sind stochastisch unabhängig.

$$\hat{b} = \frac{\sum_i y_i (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

$$\begin{aligned} \text{var}(\hat{b}) &= \text{var} \left(\frac{\sum_i y_i (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2} \right) = \frac{\text{var}(\sum_i y_i (x_i - \bar{x}))}{(\sum_i (x_i - \bar{x})^2)^2} \\ &= \frac{\sum_i \text{var}(y_i) (x_i - \bar{x})^2}{(\sum_i (x_i - \bar{x})^2)^2} = \sigma^2 \cdot \frac{\sum_i (x_i - \bar{x})^2}{(\sum_i (x_i - \bar{x})^2)^2} \end{aligned}$$

$$y_i = a + b \cdot x_i + \varepsilon \quad \text{mit } \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

nicht zufällig: a, b, x_i, σ^2 zufällig: ε, y_i

$$\text{var}(y_i) = \text{var}(a + b \cdot x_i + \varepsilon) = \text{var}(\varepsilon) = \sigma^2$$

und y_1, y_2, \dots, y_n sind stochastisch unabhängig.

$$\hat{b} = \frac{\sum_i y_i (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

$$\begin{aligned} \text{var}(\hat{b}) &= \text{var} \left(\frac{\sum_i y_i (x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2} \right) = \frac{\text{var}(\sum_i y_i (x_i - \bar{x}))}{(\sum_i (x_i - \bar{x})^2)^2} \\ &= \frac{\sum_i \text{var}(y_i) (x_i - \bar{x})^2}{(\sum_i (x_i - \bar{x})^2)^2} = \sigma^2 \cdot \frac{\sum_i (x_i - \bar{x})^2}{(\sum_i (x_i - \bar{x})^2)^2} \\ &= \sigma^2 / \sum_i (x_i - \bar{x})^2 \end{aligned}$$

Tatsächlich ist \hat{b} Normalverteilt mit Mittelwert b und

$$\text{var}(\hat{b}) = \sigma^2 / \sum_i (x_i - \bar{x})^2$$

Tatsächlich ist \hat{b} Normalverteilt mit Mittelwert b und

$$\text{var}(\hat{b}) = \sigma^2 / \sum_i (x_i - \bar{x})^2$$

Problem: Wir kennen σ^2 nicht.

Tatsächlich ist \hat{b} Normalverteilt mit Mittelwert b und

$$\text{var}(\hat{b}) = \sigma^2 / \sum_i (x_i - \bar{x})^2$$

Problem: Wir kennen σ^2 nicht.

Wir schätzen σ^2 mit Hilfe der beobachteten Residuenvarianz durch

$$s^2 := \frac{\sum_i (y_i - \hat{a} - \hat{b} \cdot x_i)^2}{n - 2}$$

Tatsächlich ist \hat{b} Normalverteilt mit Mittelwert b und

$$\text{var}(\hat{b}) = \sigma^2 / \sum_i (x_i - \bar{x})^2$$

Problem: Wir kennen σ^2 nicht.

Wir schätzen σ^2 mit Hilfe der beobachteten Residuenvarianz durch

$$s^2 := \frac{\sum_i (y_i - \hat{a} - \hat{b} \cdot x_i)^2}{n - 2}$$

Zu beachten ist, dass durch $n - 2$ geteilt wird. Das hat damit zu tun, dass zwei Modellparameter a und b bereits geschätzt wurden, und somit 2 Freiheitsgrade verloren gegangen sind.

$$\text{var}(\hat{b}) = \sigma^2 / \sum_i (x_i - \bar{x})^2$$

Schätze σ^2 durch

$$s^2 = \frac{\sum_i (y_i - \hat{a} - \hat{b} \cdot x_i)^2}{n - 2}.$$

$$\text{var}(\hat{b}) = \sigma^2 / \sum_i (x_i - \bar{x})^2$$

Schätze σ^2 durch

$$s^2 = \frac{\sum_i (y_i - \hat{a} - \hat{b} \cdot x_i)^2}{n - 2}$$

Dann ist

$$\frac{\hat{b} - b}{s / \sqrt{\sum_i (x_i - \bar{x})^2}}$$

Student- t -verteilt mit $n - 2$ Freiheitsgraden und wir können den t -Test anwenden, um die Nullhypothese $b = 0$ zu testen.

$$\text{var}(\hat{b}) = \sigma^2 / \sum_i (x_i - \bar{x})^2$$

Schätze σ^2 durch

$$s^2 = \frac{\sum_i (y_i - \hat{a} - \hat{b} \cdot x_i)^2}{n - 2}$$

Dann ist

$$\frac{\hat{b} - b}{s / \sqrt{\sum_i (x_i - \bar{x})^2}}$$

Student- t -verteilt mit $n - 2$ Freiheitsgraden und wir können den t -Test anwenden, um die Nullhypothese $b = 0$ zu testen.

Verwerfe $H_0: „b = 0“$ zum Signifikanzniveau α , wenn

$$\left| \frac{\hat{b}}{s / \sqrt{\sum_i (x_i - \bar{x})^2}} \right| \geq q_{1-\alpha/2}, \text{ wo } q_{1-\alpha/2} \text{ das } (1 - \alpha/2)\text{-Quantil der}$$

Student-Verteilung mit $n - 2$ Freiheitsgraden ist.

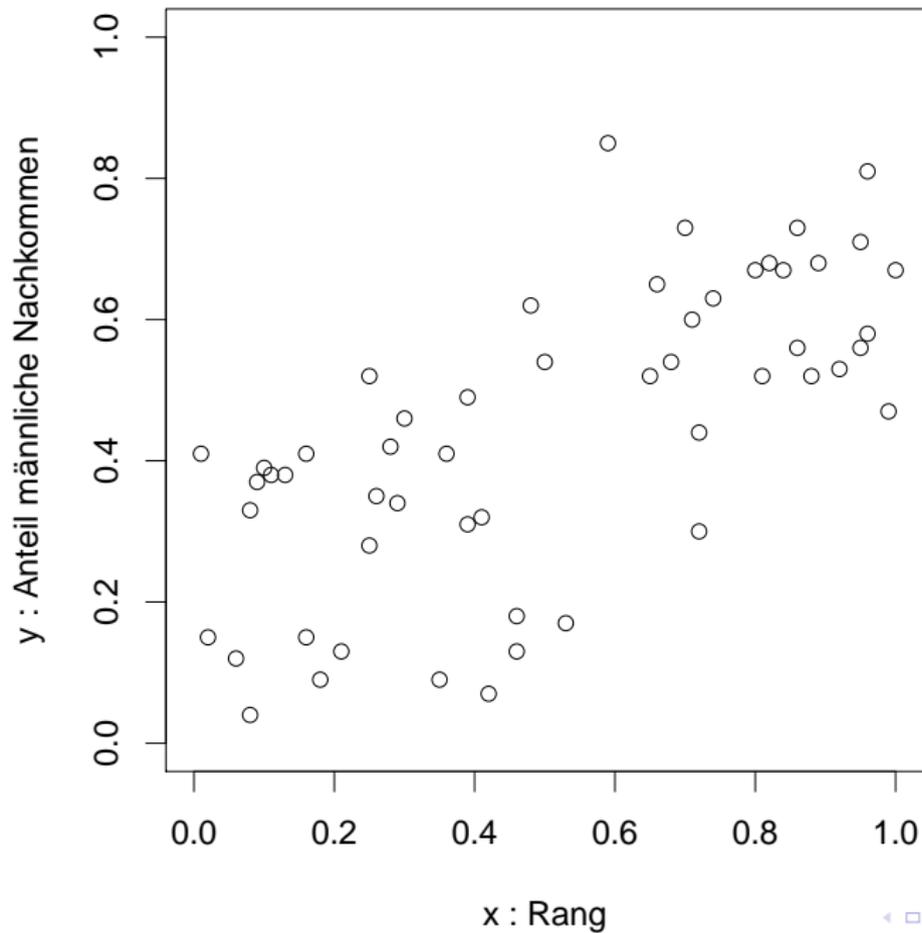
Beispiel: Rothirsch (*Cervus elaphus*)

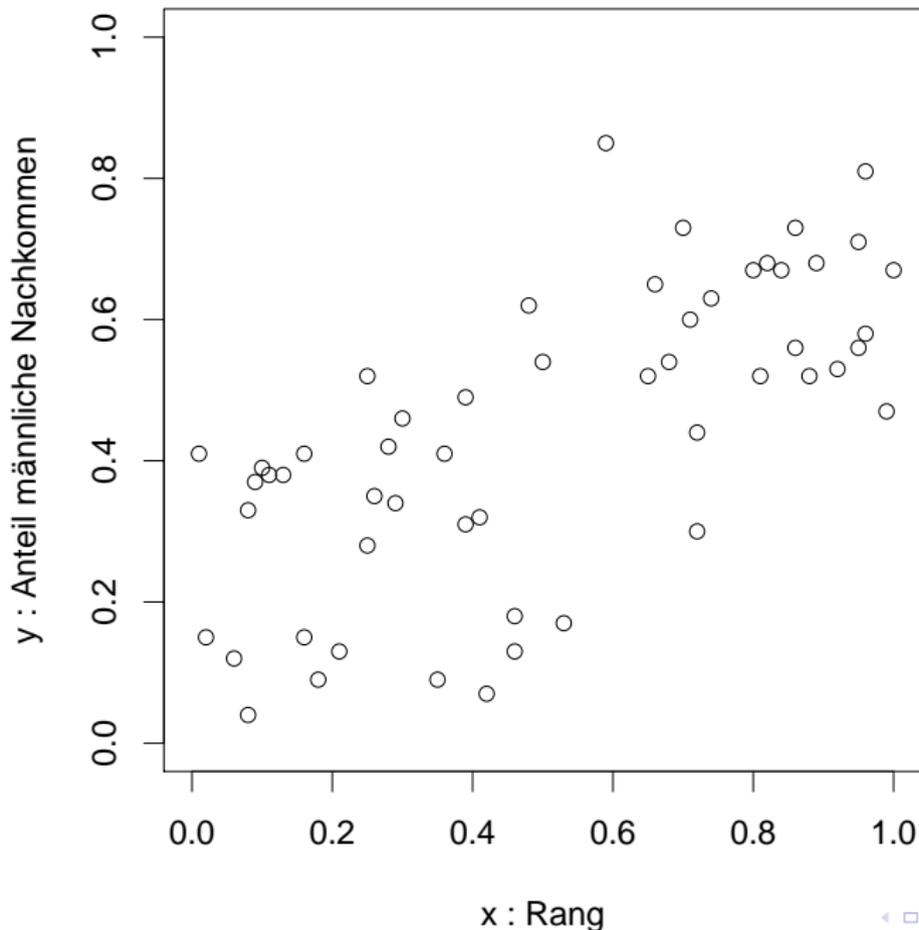
Theorie: Hirschkühe können das Geschlecht ihrer Nachkommen beeinflussen.

Beispiel: Rothirsch (*Cervus elaphus*)

Theorie: Hirschkühe können das Geschlecht ihrer Nachkommen beeinflussen.

Unter dem Gesichtspunkt evolutionär stabiler Strategien ist zu erwarten, dass schwache Tiere eher zu weiblichem und starke Tiere eher zu männlichem Nachwuchs tendieren.





Es ist

$$\bar{x} = 0.51$$

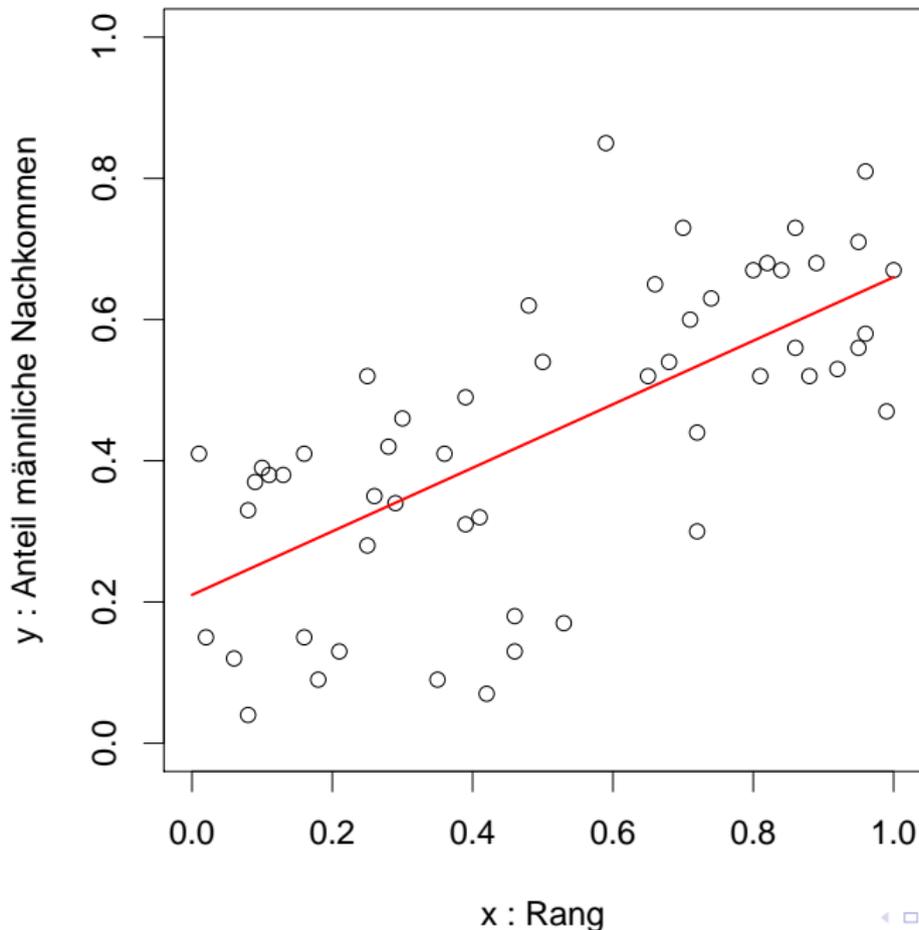
(mittl. Rang)

$$\bar{y} = 0.44$$

(mittl. Ant. männl.
Nachk.)

$$\sigma_x^2 = 0.097$$

$$\text{cov}(x, y) = 0.044$$



Es ist

$\bar{x} = 0.51$
(mittl. Rang)

$\bar{y} = 0.44$
(mittl. Ant. männl.
Nachk.)

$\sigma_x^2 = 0.097$

$cov(x, y) = 0.044$

$\hat{b} = \frac{0.044}{0.097} \doteq 0.45$

\hat{a}

$= 0.44 - 0.51 \cdot 0.45$

$\doteq 0.21$

Im Rothirschkuhe-Beispiel:

$$\hat{b} = 0.45, \frac{s_{\text{res}}}{\sqrt{\sum_i (x_i - \bar{x})^2}} = 0.0673,$$

also beobachten wir $t = \frac{\hat{b} - 0}{s_{\text{res}} / \sqrt{\sum_i (x_i - \bar{x})^2}} = 6.7$

Im Rothirschkühe-Beispiel:

$$\hat{b} = 0.45, \frac{s_{\text{res}}}{\sqrt{\sum_i (x_i - \bar{x})^2}} = 0.0673,$$

also beobachten wir $t = \frac{\hat{b} - 0}{s_{\text{res}} / \sqrt{\sum_i (x_i - \bar{x})^2}} = 6.7$

Einer Tabelle entnehmen wir: Das 99.95%-Quantil der Student-Verteilung mit 50 Freiheitsgraden ist 3.496 (und das der Student-Vert. mit 60 Freiheitsgraden ist 3.460).

Wir können also die Nullhypothese „das wahre $b = 0$ “ zum Signifikanzniveau 0.1% ablehnen.

Im Rothirschkühe-Beispiel:

$$\hat{b} = 0.45, \frac{s_{\text{res}}}{\sqrt{\sum_i (x_i - \bar{x})^2}} = 0.0673,$$

also beobachten wir $t = \frac{\hat{b} - 0}{s_{\text{res}} / \sqrt{\sum_i (x_i - \bar{x})^2}} = 6.7$

Einer Tabelle entnehmen wir: Das 99.95%-Quantil der Student-Verteilung mit 50 Freiheitsgraden ist 3.496 (und das der Student-Vert. mit 60 Freiheitsgraden ist 3.460).

Wir können also die Nullhypothese „das wahre $b = 0$ “ zum Signifikanzniveau 0.1% ablehnen.

Bemerkung: Das beweist natürlich nicht, dass Hirschkühe das Geschlecht ihrer Nachkommen willentlich bestimmen können. Es scheint eher plausibel anzunehmen, dass es Faktoren gibt, die den Rang und die Geschlechterverteilung der Nachkommen zugleich beeinflussen, siehe die Diskussion in dem zitierten Artikel von T. H. Clutton-Brock et. al.

Viel Erfolg beim Lernen!