

# Wahrscheinlichkeitsrechnung und Statistik für Biologen

## 4. Der t-Test

Martin Hutzenthaler & Dirk Metzler

<http://www.zi.biologie.uni-muenchen.de/evol/StatGen.html>

6./18. Mai 2010

## 1 t-Test für gepaarte Stichproben

- Beispiel: Orientierung bei Trauerschnäppern
- Die t-Verteilung
- Zurück zu: Orientierung bei Trauerschnäppern
- Beispiel: Richtungsabhängige Korkdicke
- Zusammenfassung t-Test
- Prinzip des statistischen Testens

# Wahrscheinlichkeitsrechnung und Statistik für Biologen

## 4. Der t-Test

Martin Hutzenthaler & Dirk Metzler

<http://www.zi.biologie.uni-muenchen.de/evol/StatGen.html>

6./18. Mai 2010

## 1 t-Test für gepaarte Stichproben

- Beispiel: Orientierung bei Trauerschnäppern
- Die t-Verteilung
- Zurück zu: Orientierung bei Trauerschnäppern
- Beispiel: Richtungsabhängige Korkdicke
- Zusammenfassung t-Test
- Prinzip des statistischen Testens

# Inhalt

- 1 t-Test für gepaarte Stichproben
  - Beispiel: Orientierung bei Trauerschnäppern
  - Die t-Verteilung
  - Zurück zu: Orientierung bei Trauerschnäppern
  - Beispiel: Richtungsabhängige Korkdicke
  - Zusammenfassung t-Test
  - Prinzip des statistischen Testens

# Inhalt

- 1 t-Test für gepaarte Stichproben
  - Beispiel: Orientierung bei Trauerschnäppern
  - Die t-Verteilung
  - Zurück zu: Orientierung bei Trauerschnäppern
  - Beispiel: Richtungsabhängige Korkdicke
  - Zusammenfassung t-Test
  - Prinzip des statistischen Testens

# Trauerschnäpper (*Ficedula hypoleuca*)



[http://en.wikipedia.org/wiki/File:Ficedula\\_hypoleuca\\_NRM.jpg](http://en.wikipedia.org/wiki/File:Ficedula_hypoleuca_NRM.jpg)



Foto (c) Simon Eugster



Wiltschko, W.; Gesson, M.; Stapput, K.; Wiltschko, R.

Light-dependent magnetoreception in birds: interaction of at least two different receptors.

*Naturwissenschaften* 91.3, pp. 130-4, 2004.



Wiltschko, R.; Ritz, T.; Stapput, K.; Thalau, P.; Wiltschko, W.

Two different types of light-dependent responses to magnetic fields in birds.

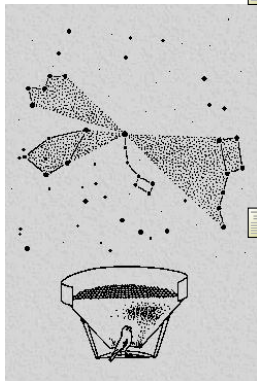
*Curr Biol* 15.16, pp. 1518-23, 2005.



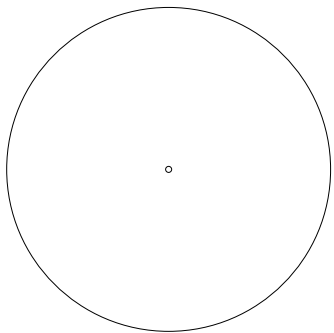
Wiltschko, R.; Stapput, K.; Bischof, H. J.; Wiltschko, W.

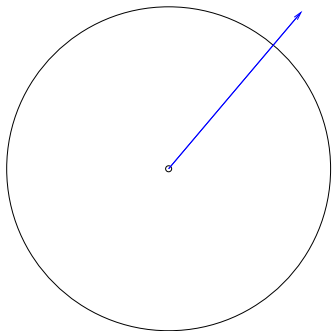
Light-dependent magnetoreception in birds: increasing intensity of monochromatic light changes the nature of the response.

*Front Zool*, 4, 2007.

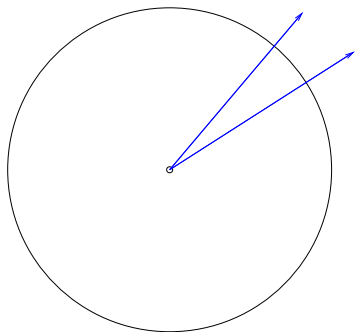






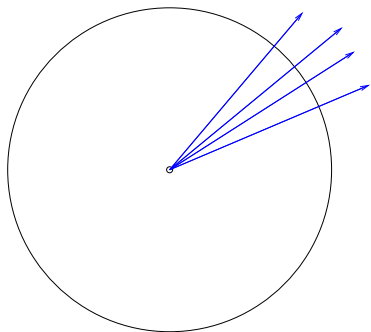


Richtung eines Fluges bei **blauem Licht**.

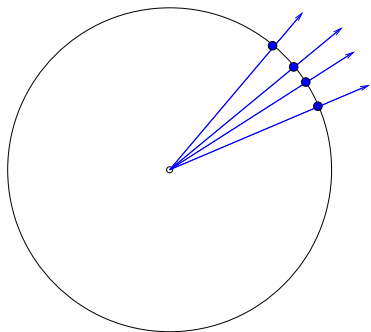


Richtung eines Fluges bei **blauem Licht**.

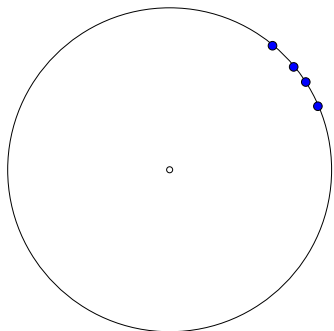
Richtung eines weiteren Fluges desselben Vogels bei **blauem Licht**.



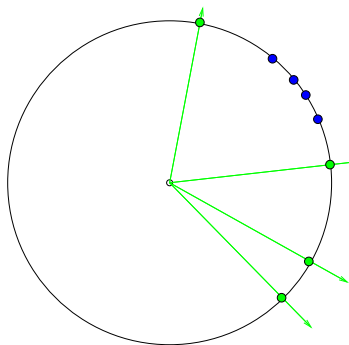
Richtungen aller Flüge dieses Vogels bei blauem Licht.



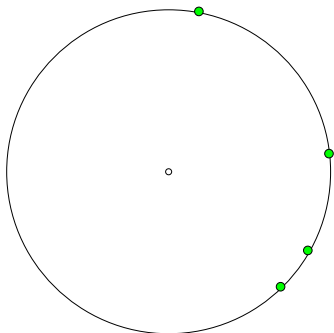
Richtungen aller Flüge dieses Vogels bei blauem Licht.



Richtungen aller Flüge dieses  
Vogels bei **blauem Licht**.  
Zugehörige Austrittspunkte.

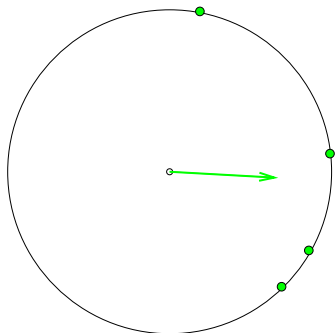


Richtungen aller Flüge dieses Vogels bei grünem Licht.

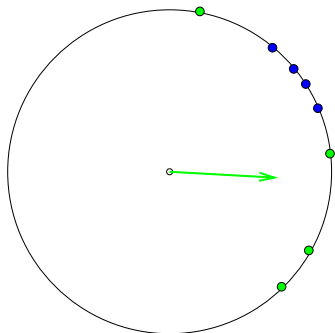


Richtungen aller Flüge dieses  
Vogels bei grünem Licht.  
Zugehörige Austrittspunkte.

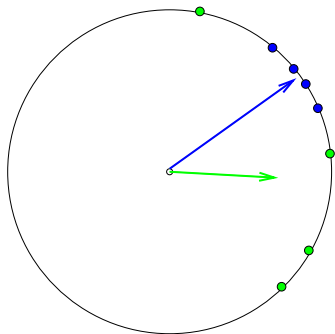




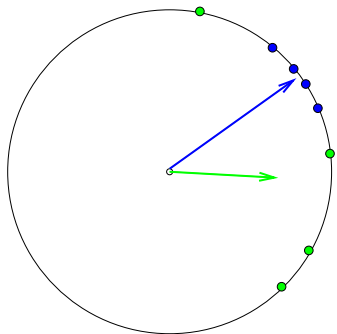
Zugehörige Austrittspunkte.  
Pfeilspitze: Schwerpunkt der Austrittspunkte bei grünem Licht.



Pfeilspitze: Schwerpunkt der Austrittspunkte bei grünem Licht.



Pfeilspitze: Schwerpunkt der Austrittspunkte bei grünem Licht.  
Dasselbe für die "blauen" Austrittspunkte.



Pfeilspitze: Schwerpunkt der Austrittspunkte bei grünem Licht.  
Dasselbe für die "blauen" Austrittspunkte.

Je variabler die Richtungen desto kürzer der Pfeil!

# Fragestellung

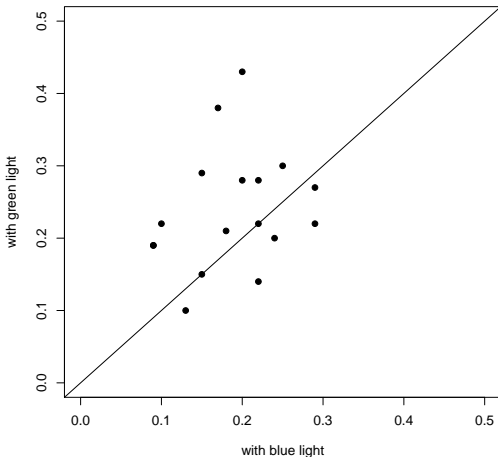
Hat die Farbe der monochromatischen Beleuchtung einen Einfluß auf die Orientierung?

# Fragestellung

Hat die Farbe der monochromatischen Beleuchtung einen  
Einfluß auf die Orientierung?

Experiment: Bei 17 Vögeln wurde die Länge des  
Schwerpunktsvektors sowohl bei **blauem** als auch bei **grünem**  
Licht bestimmt.

# Trauerschnäpper: Länge des Schwerpunktsvektors bei grünem und bei blauem Licht, $n=17$



Wie kann ich  
statistisch testen,  
ob die Farbe  
einen Einfluss hat?



Wir berechnen nun für jeden Vogel den Abstand des Punktes von der Diagonale,

Wir berechnen nun für jeden Vogel den Abstand des Punktes von der Diagonale,

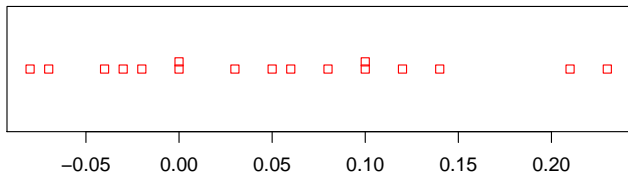
d.h.

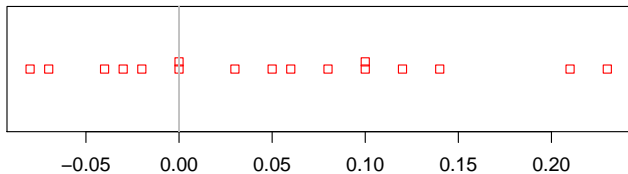
$$x := \text{“Grünwert”} - \text{“Blauwert”}$$

Wir berechnen nun für jeden Vogel den Abstand des Punktes von der Diagonale,

d.h.

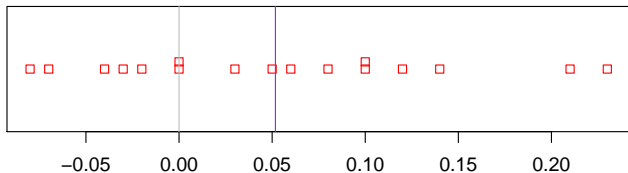
$$x := \text{“Grünwert”} - \text{“Blauwert”}$$





Kann der wahre Mittelwert  $\mu = 0$  sein?

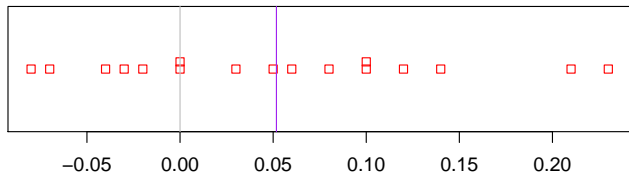




Kann der wahre Mittelwert  $\mu = 0$  sein?

$$\bar{x} = 0.0518$$

$$s = 0.0912$$

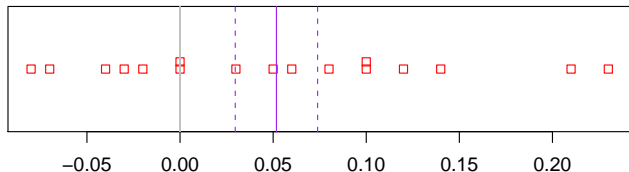


Kann der wahre Mittelwert  $\mu = 0$  sein?

$$\bar{x} = 0.0518$$

$$s = 0.0912$$

$$\text{SEM} = \frac{s}{\sqrt{n}}$$



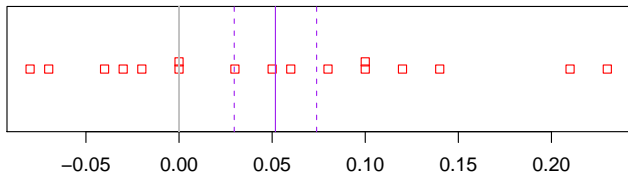
Kann der wahre Mittelwert  $\mu = 0$  sein?

$$\bar{x} = 0.0518$$

$$s = 0.0912$$

$$\text{SEM} = \frac{s}{\sqrt{n}} = \frac{0.0912}{\sqrt{17}}$$





Kann der wahre Mittelwert  $\mu = 0$  sein?

$$\bar{x} = 0.0518$$

$$s = 0.0912$$

$$\text{SEM} = \frac{s}{\sqrt{n}} = \frac{0.0912}{\sqrt{17}} = 0.022$$

Ist  $|\bar{X} - \mu| \approx 0.0518$  eine große Abweichung?

Ist  $|\bar{X} - \mu| \approx 0.0518$  eine große Abweichung?

Groß?

Ist  $|\bar{x} - \mu| \approx 0.0518$  eine große Abweichung?

Groß? Groß im Vergleich zu was?

Ist  $|\bar{x} - \mu| \approx 0.0518$  eine große Abweichung?

Groß? Groß im Vergleich zu was?

In welcher Vergleichseinheit soll  $|\bar{x} - \mu|$  gemessen werden?

Ist  $|\bar{x} - \mu| \approx 0.0518$  eine große Abweichung?

Groß? Groß im Vergleich zu was?

In welcher Vergleichseinheit soll  $|\bar{x} - \mu|$  gemessen werden?

Immer im Vergleich zum  
**Standardfehler!**

Ist  $|\bar{x} - \mu| \approx 0.0518$  eine große Abweichung?

Groß? Groß im Vergleich zu was?

In welcher Vergleichseinheit soll  $|\bar{x} - \mu|$  gemessen werden?

Immer im Vergleich zum

**Standardfehler!**

$$|\bar{x} - \mu|$$

gemessen in der Einheit 'Standardfehler'

heißt **t-Statistik**

Ist  $|\bar{x} - \mu| \approx 0.0518$  eine große Abweichung?

Groß? Groß im Vergleich zu was?

In welcher Vergleichseinheit soll  $|\bar{x} - \mu|$  gemessen werden?

Immer im Vergleich zum

**Standardfehler!**

$$|\bar{x} - \mu|$$

gemessen in der Einheit 'Standardfehler'

heißt **t-Statistik**

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



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

$t = 1$  bedeutet

1 Standardfehler von  $\mu$  entfernt  
(kommt häufig vor)

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

$t = 1$  bedeutet

- 1** Standardfehler von  $\mu$  entfernt  
(kommt häufig vor)

$t = 3$  bedeutet

- 3** Standardfehler von  $\mu$  entfernt  
(kommt selten vor)

In unserem Fall:

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} \approx \frac{0.0518}{0.022} \approx 2.34$$

In unserem Fall:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \approx \frac{0.0518}{0.022} \approx 2.34$$

Also:  $\bar{x}$  ist mehr als 2.3 Standardfehler von  $\mu = 0$  entfernt.

In unserem Fall:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \approx \frac{0.0518}{0.022} \approx 2.34$$

Also:  $\bar{x}$  ist mehr als 2.3 Standardfehler von  $\mu = 0$  entfernt.

Wie wahrscheinlich ist das, wenn 0 der wahre Mittelwert ist?

In unserem Fall:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \approx \frac{0.0518}{0.022} \approx 2.34$$

Also:  $\bar{x}$  ist mehr als 2.3 Standardfehler von  $\mu = 0$  entfernt.

Wie wahrscheinlich ist das, wenn 0 der wahre Mittelwert ist?

anders gefragt:

Ist diese Abweichung signifikant?

In unserem Fall:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \approx \frac{0.0518}{0.022} \approx 2.34$$

Also:  $\bar{x}$  ist mehr als 2.3 Standardfehler von  $\mu = 0$  entfernt.

Wie wahrscheinlich ist das, wenn 0 der wahre Mittelwert ist?

anders gefragt:

## Ist diese Abweichung signifikant?

Für die Antwort benötigen wir die Verteilung der t-Statistik.

# Inhalt

- 1 t-Test für gepaarte Stichproben
  - Beispiel: Orientierung bei Trauerschnäppern
  - Die t-Verteilung
  - Zurück zu: Orientierung bei Trauerschnäppern
  - Beispiel: Richtungsabhängige Korkdicke
  - Zusammenfassung t-Test
  - Prinzip des statistischen Testens



Wir wissen:

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

ist asymptotisch (für große  $n$ ) standardnormalverteilt.

Wir wissen:

$$\frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

ist asymptotisch (für große  $n$ ) standardnormalverteilt.

Die t-Statistik ist jedoch mit  $s$  an Stelle von  $\sigma$  definiert und ist asymptotisch nicht mehr normalverteilt.

Wir wissen:

$$\frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

ist asymptotisch (für große  $n$ ) standardnormalverteilt.

Die t-Statistik ist jedoch mit  $s$  an Stelle von  $\sigma$  definiert und ist asymptotisch nicht mehr normalverteilt.

Aber fast:

## Allgemein gilt

Sind  $X_1, \dots, X_n$  unabhängig aus einer Normalverteilung mit Mittelwert  $\mu$  gezogen, so ist

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

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

Eine t-verteilte Zufallsvariable bezeichnen wir meist mit  $T$ .

## Allgemein gilt

Sind  $X_1, \dots, X_n$  unabhängig aus einer Normalverteilung mit Mittelwert  $\mu$  gezogen, so ist

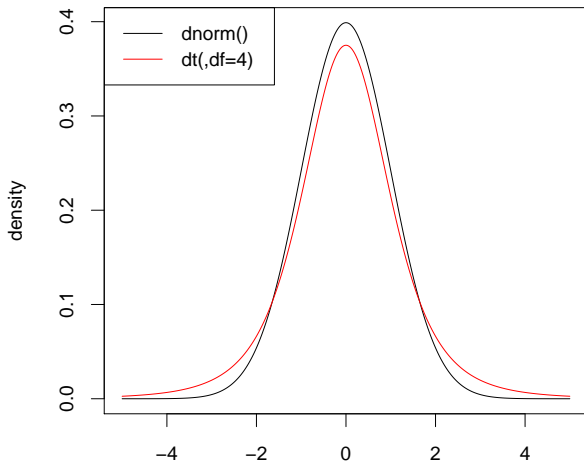
$$\frac{\bar{X} - \mu}{s/\sqrt{n}}$$

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

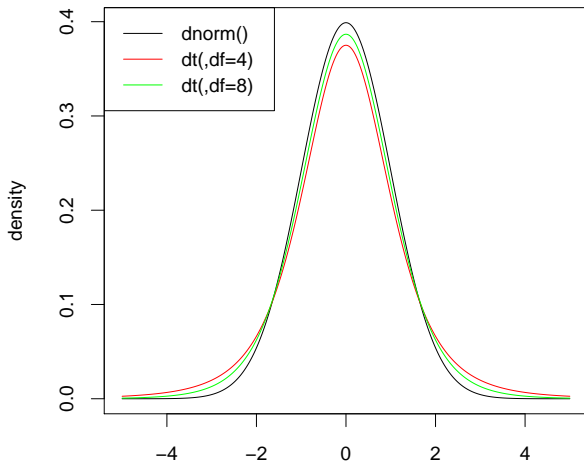
Eine t-verteilte Zufallsvariable bezeichnen wir meist mit  $T$ .

Die t-Verteilung heißt auch **Student-Verteilung**. Die t-Verteilung wurde 1908 von William Gosset veröffentlicht, während Gosset in einer Guinness-Brauerei arbeitete. Da sein Arbeitgeber die Veröffentlichung nicht gestattete, veröffentlichte Gosset sie unter dem Pseudonym *Student*.

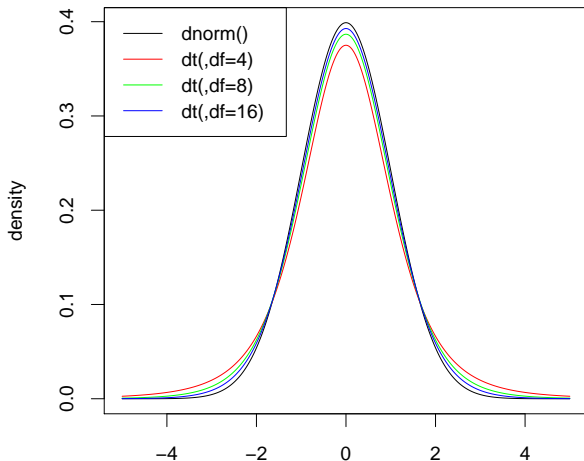
# Dichte der t-Verteilung



# Dichte der t-Verteilung

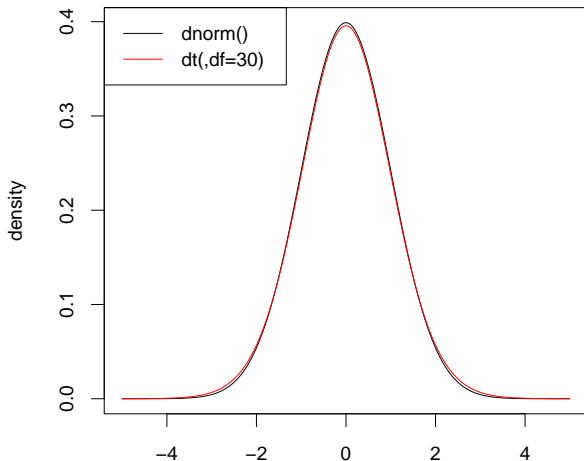


# Dichte der t-Verteilung





# Dichte der t-Verteilung



# Freiheitsgrade

Es gibt 5 Freiheitsgrade im Vektor

$$x = (x_1 \ x_2 \ x_3 \ x_4 \ x_5)$$

da 5 Werte frei wählbar sind. Der Vektor

$$v := x - \bar{x}$$

hat 4 Freiheitsgrade, denn nach Wahl von  $v_1, v_2, v_3, v_4$  ist  $v_5$  festgelegt wegen  $\text{mean}(v) = 0$  und somit  $v_1 + \dots + v_4 + v_5 = 0$ .

# Freiheitsgrade

## Faustregel:

Freiheitsgrade =  
Stichprobenlänge - Anzahl der festgelegten Parameter

# Freiheitsgrade

## Faustregel:

Freiheitsgrade =  
Stichprobenlänge - Anzahl der festgelegten Parameter

**Beispiele:** Sei  $n$  die Länge des Vektors  $x$ .

- Der Mittelwert von  $x - \bar{x}$  ist gleich 0 und somit 'festgelegt'. Es verbleiben  $n - 1$  Freiheitsgrade.

# Freiheitsgrade

## Faustregel:

Freiheitsgrade =  
Stichprobenlänge - Anzahl der festgelegten Parameter

**Beispiele:** Sei  $n$  die Länge des Vektors  $x$ .

- Der Mittelwert von  $x - \bar{x}$  ist gleich 0 und somit 'festgelegt'. Es verbleiben  $n - 1$  Freiheitsgrade.
- Bei  $\bar{x}$  ist kein Parameter festgelegt, also  $n$  Freiheitsgrade.

# Freiheitsgrade

## Faustregel:

Freiheitsgrade =  
Stichprobenlänge - Anzahl der festgelegten Parameter

**Beispiele:** Sei  $n$  die Länge des Vektors  $x$ .

- Der Mittelwert von  $x - \bar{x}$  ist gleich 0 und somit 'festgelegt'. Es verbleiben  $n - 1$  Freiheitsgrade.
- Bei  $\bar{x}$  ist kein Parameter festgelegt, also  $n$  Freiheitsgrade.
- Die Standardabweichung von  $\frac{x}{s}$  ist 1. Also verbleiben für  $\frac{\bar{x} - \mu}{s/\sqrt{n}}$  noch  $n - 1$  Freiheitsgrade.

# Freiheitsgrade

## Faustregel:

Freiheitsgrade =  
Stichprobenlänge - Anzahl der festgelegten Parameter

**Beispiele:** Sei  $n$  die Länge des Vektors  $x$ .

- Der Mittelwert von  $x - \bar{x}$  ist gleich 0 und somit 'festgelegt'. Es verbleiben  $n - 1$  Freiheitsgrade.
- Bei  $\bar{x}$  ist kein Parameter festgelegt, also  $n$  Freiheitsgrade.
- Die Standardabweichung von  $\frac{x}{s}$  ist 1. Also verbleiben für  $\frac{\bar{x} - \mu}{s/\sqrt{n}}$  noch  $n - 1$  Freiheitsgrade.
- Der Vektor  $\frac{x - \bar{x}}{s}$  hat  $n - 2$  Freiheitsgrade, da Mittelwert und Standardabweichung festgelegt sind.

# Inhalt

- 1 t-Test für gepaarte Stichproben
  - Beispiel: Orientierung bei Trauerschnäppern
  - Die t-Verteilung
  - Zurück zu: Orientierung bei Trauerschnäppern
  - Beispiel: Richtungsabhängige Korkdicke
  - Zusammenfassung t-Test
  - Prinzip des statistischen Testens



Wir meinen:  
Die Farbe der Beleuchtung  
hat einen Einfluß auf die Orientierung

Wir meinen:  
Die Farbe der Beleuchtung  
hat einen Einfluß auf die Orientierung

Ein Skeptiker würde erwidern:  
**Alles nur Zufall**

Wir meinen:  
Die Farbe der Beleuchtung  
hat einen Einfluß auf die Orientierung

Ein Skeptiker würde erwidern:  
**Alles nur Zufall**

Wir wollen nun zeigen:  
Unter der Annahme 'Kein Einfluß'  
ist die Beobachtung sehr unwahrscheinlich

Wir meinen:  
Die Farbe der Beleuchtung  
hat einen Einfluß auf die Orientierung

Ein Skeptiker würde erwidern:  
**Alles nur Zufall**

Wir wollen nun zeigen:  
Unter der Annahme 'Kein Einfluß'  
ist die Beobachtung sehr unwahrscheinlich

**Nullhypothese:**  $\mu = 0$

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

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

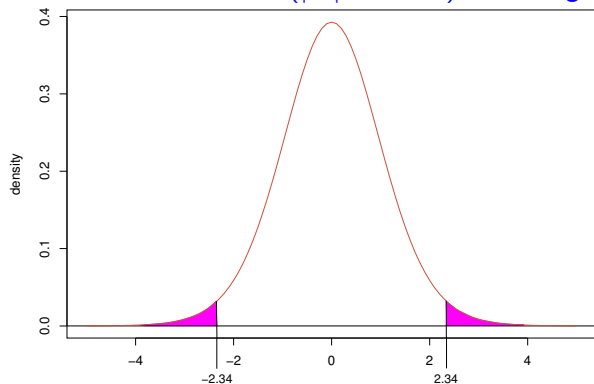
$$\Pr(|T| = 2.34) = 0$$

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

$\Pr(|T| = 2.34) = 0$       **Das bringt nichts!**

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

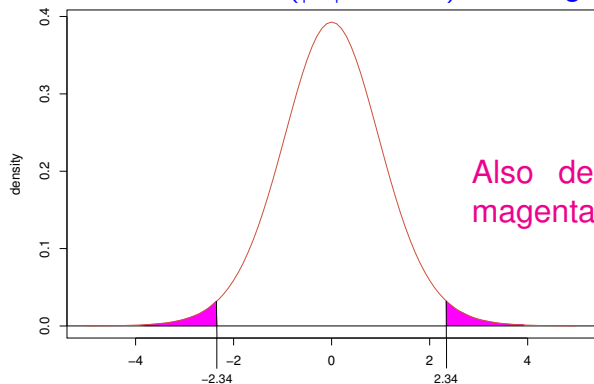
Zu berechnen ist  $\Pr(|T| \geq 2.34)$ , der sog. *p*-Wert.





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

Zu berechnen ist  $\Pr(|T| \geq 2.34)$ , der sog. *p*-Wert.



Also der Gesamthalt der magentafarbenen Flächen.

R macht das für uns:

```
> pt(-2.34,df=16)+pt(2.34,df=16,lower.tail=FALSE)
[1] 0.03257345
```

R macht das für uns:

```
> pt(-2.34,df=16)+pt(2.34,df=16,lower.tail=FALSE)
[1] 0.03257345
```

Beachte:  $\text{pt}(2.34,df=16,lower.tail=FALSE)$  ist dasselbe wie  $1-\text{pt}(2.34,df=16)$

R macht das für uns:

```
> pt(-2.34,df=16)+pt(2.34,df=16,lower.tail=FALSE)
[1] 0.03257345
```

Beachte:  $\text{pt}(2.34, \text{df}=16, \text{lower.tail}=\text{FALSE})$  ist dasselbe wie  $1-\text{pt}(2.34, \text{df}=16)$

Zum Vergleich mal mit der Normalverteilung:

```
> pnorm(-2.34)+pnorm(2.34,lower.tail=FALSE)
[1] 0.01928374
```

# Vollständiger t-Test mit R

```
> x <- trauerschn$gruen-trauerschn$blau  
> t.test(x)
```

One Sample t-test

```
data: x  
t = 2.3405, df = 16, p-value = 0.03254  
alternative hypothesis: true mean is not equal to 0  
95 percent confidence interval:  
 0.004879627 0.098649784  
sample estimates:  
 mean of x  
0.05176471
```

Wir halten fest:

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

Wir halten fest:

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

Wenn die **Nullhypothese** “alles nur Zufall” (hier  $\mu = 0$ ) gilt, dann ist eine mindestens so große Abweichung sehr unwahrscheinlich.

Wir halten fest:

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

Wenn die **Nullhypothese** “alles nur Zufall” (hier  $\mu = 0$ ) gilt, dann ist eine mindestens so große Abweichung sehr unwahrscheinlich.

Sprechweise:

Wir verwerfen die Nullhypothese  
auf dem 5%-Signifikanzniveau.



Wir halten fest:

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

Wenn die **Nullhypothese** “alles nur Zufall” (hier  $\mu = 0$ ) gilt, dann ist eine mindestens so große Abweichung sehr unwahrscheinlich.

Sprechweise:

Wir verwerfen die Nullhypothese  
auf dem 5%-Signifikanzniveau.

Oder:

Die Differenz zwischen grün und blau  
ist auf dem 5%-Niveau signifikant.

Die Nullhypothese wurde also auf dem 5%-Niveau verworfen.  
Welche Aussagen sind wahr/sinnvoll?

- Die Nullhypothese ist falsch.
- 
- 
- 
-

Die Nullhypothese wurde also auf dem 5%-Niveau verworfen.  
Welche Aussagen sind wahr/sinnvoll?

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

Die Nullhypothese wurde also auf dem 5%-Niveau verworfen.  
Welche Aussagen sind wahr/sinnvoll?

- ~~Die Nullhypothese ist falsch.~~
- Die Nullhypothese ist mit 95%-iger Ws falsch.
- 
- 
-

Die Nullhypothese wurde also auf dem 5%-Niveau verworfen.  
Welche Aussagen sind wahr/sinnvoll?

- ~~Die Nullhypothese ist falsch.~~
- ~~Die Nullhypothese ist mit 95% iger Ws falsch.~~
- 
- 
-

Die Nullhypothese wurde also auf dem 5%-Niveau verworfen.  
Welche Aussagen sind wahr/sinnvoll?

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

Die Nullhypothese wurde also auf dem 5%-Niveau verworfen.  
Welche Aussagen sind wahr/sinnvoll?

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

Die Nullhypothese wurde also auf dem 5%-Niveau verworfen.  
Welche Aussagen sind wahr/sinnvoll?

- ~~Die Nullhypothese ist falsch.~~
- ~~Die Nullhypothese ist mit 95% iger Ws falsch.~~
- Falls die Nullhypothese wahr ist, beobachtet man ein so extremes Ergebnis nur in 5% der Fälle. ✓
- Die Orientierung der Vögel ist bei blau und grün verschieden.
-



Die Nullhypothese wurde also auf dem 5%-Niveau verworfen.  
Welche Aussagen sind wahr/sinnvoll?

- ~~Die Nullhypothese ist falsch.~~
- ~~Die Nullhypothese ist mit 95% iger Ws falsch.~~
- Falls die Nullhypothese wahr ist, beobachtet man ein so extremes Ergebnis nur in 5% der Fälle. ✓
- ~~Die Orientierung der Vögel ist bei blau und grün verschieden.~~
-

Die Nullhypothese wurde also auf dem 5%-Niveau verworfen.  
Welche Aussagen sind wahr/sinnvoll?

- ~~Die Nullhypothese ist falsch.~~
- ~~Die Nullhypothese ist mit 95% iger Ws falsch.~~
- Falls die Nullhypothese wahr ist, beobachtet man ein so extremes Ergebnis nur in 5% der Fälle. ✓
- ~~Die Orientierung der Vögel ist bei blau und grün verschieden.~~
- Die Orientierung bei grün und blau ist auf dem 5%-Niveau signifikant verschieden.

Die Nullhypothese wurde also auf dem 5%-Niveau verworfen.  
Welche Aussagen sind wahr/sinnvoll?

- ~~Die Nullhypothese ist falsch.~~
- ~~Die Nullhypothese ist mit 95% iger Ws falsch.~~
- Falls die Nullhypothese wahr ist, beobachtet man ein so extremes Ergebnis nur in 5% der Fälle. ✓
- ~~Die Orientierung der Vögel ist bei blau und grün verschieden.~~
- Die Orientierung bei grün und blau ist auf dem 5%-Niveau signifikant verschieden. ✓

Man könnte auch ein anderes Signifikanzniveau  $\alpha$  wählen.  
Dann müsste man zeigen, dass der p-Wert kleiner als  $\alpha$  ist.

Man könnte auch ein anderes Signifikanzniveau  $\alpha$  wählen. Dann müsste man zeigen, dass der p-Wert kleiner als  $\alpha$  ist.

Wichtig: Wähle zuerst das Signifikanzniveau und ermittle erst dann den p-Wert! Das Signifikanzniveau je nach p-Wert zu wählen ist geschummelt.

Man könnte auch ein anderes Signifikanzniveau  $\alpha$  wählen. Dann müsste man zeigen, dass der p-Wert kleiner als  $\alpha$  ist.

Wichtig: Wähle zuerst das Signifikanzniveau und ermittle erst dann den p-Wert! Das Signifikanzniveau je nach p-Wert zu wählen ist geschummelt.

In der Literatur wird üblicherweise 5% als Signifikanzniveau gewählt.

Man könnte auch ein anderes Signifikanzniveau  $\alpha$  wählen. Dann müsste man zeigen, dass der p-Wert kleiner als  $\alpha$  ist.

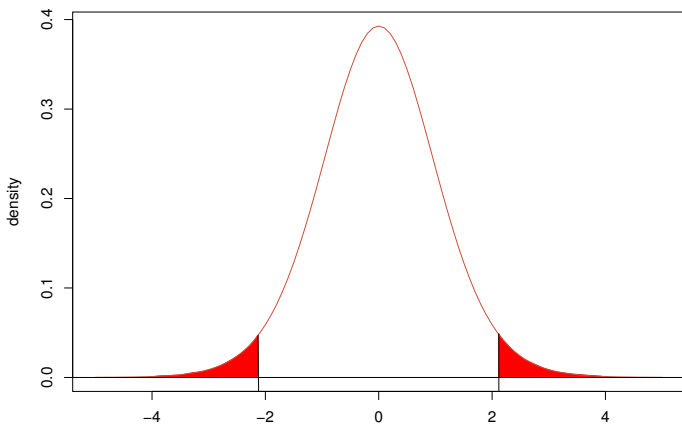
Wichtig: Wähle zuerst das Signifikanzniveau und ermittle erst dann den p-Wert! Das Signifikanzniveau je nach p-Wert zu wählen ist geschummelt.

In der Literatur wird üblicherweise  
5% als Signifikanzniveau gewählt.

Beachte:

Falls die Nullhypothese zutrifft,  
ist die Wahrscheinlichkeit,  
dass wir sie zu Unrecht auf dem 5%-Niveau verwerfen,  
höchstens 5%.

Wir verwerfen also die Nullhypothese auf 5%-Niveau, wenn der Wert der  $t$ -Statistik 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
$\infty$	1.96

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
$\infty$	1.96

```
> qt(0.025,df=c(5,10,20,30,100,1e100))
```

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
$\infty$	1.96

```
> qt(0.025,df=c(5,10,20,30,100,1e100))  
[1] -2.570582 -2.228139 -2.085963 -2.042272 -1.983972  
-1.959964
```

# Inhalt

- 1 t-Test für gepaarte Stichproben
  - Beispiel: Orientierung bei Trauerschnäppern
  - Die t-Verteilung
  - Zurück zu: Orientierung bei Trauerschnäppern
  - **Beispiel: Richtungsabhängige Korkdicke**
  - Zusammenfassung t-Test
  - Prinzip des statistischen Testens

# Korkeiche (*Quercus suber*)



Foto (c) Hannes Grobe



Foto (c) Manfred Werner

Fragestellung: Hängt die Korkdicke von der Himmelsrichtung ab?

# Achtung: simulierte Daten!

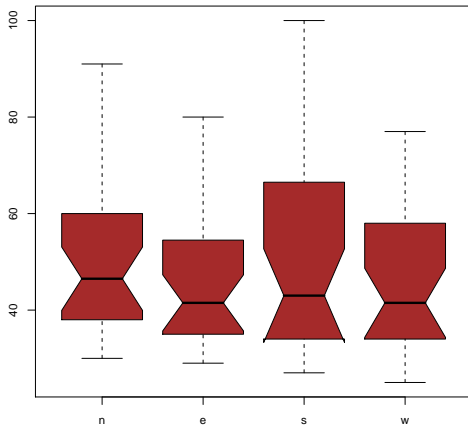
Im Beispiel mit den Korkdicken verwenden wir wieder simulierte Daten, die aber Daten aus echten Studien nachempfunden sind, auch im Ergebnis.

Bei  $n = 28$  Bäumen wurden die Korkdicken [mm] in den vier Himmelsrichtungen gemessen:

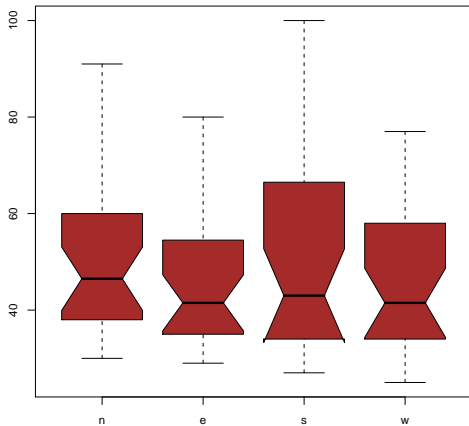
n	e	s	w
72	66	76	77
60	53	66	63
5	57	64	58
41	29	36	38
32	32	35	36
30	35	34	26
39	39	31	27
.	.	.	.
.	.	.	.



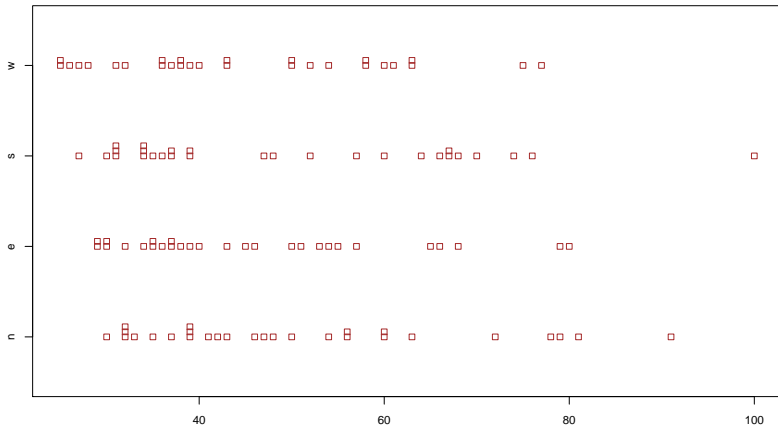
## Korkdicken nach Himmelsrichtung getrennt



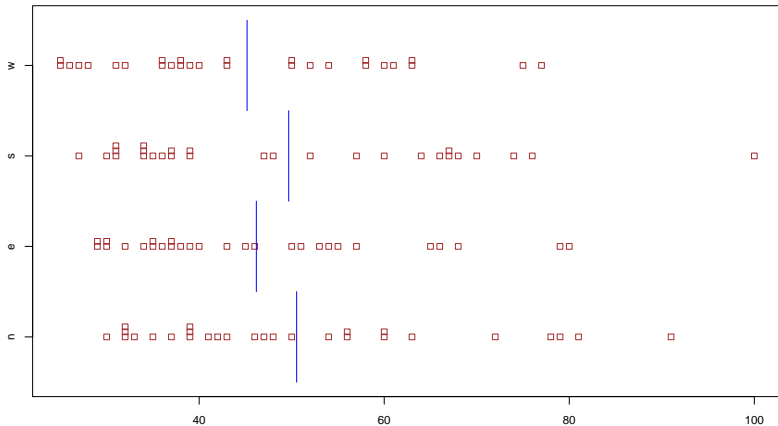
## Korkdicken nach Himmelsrichtung getrennt



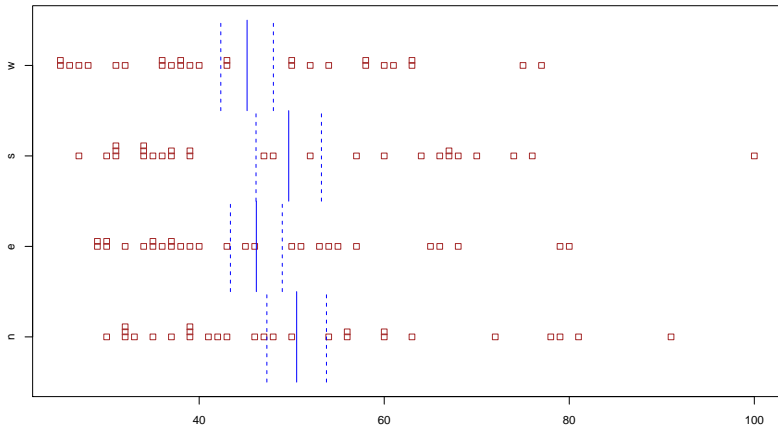
Kann da was signifikant unterschiedlich sein???



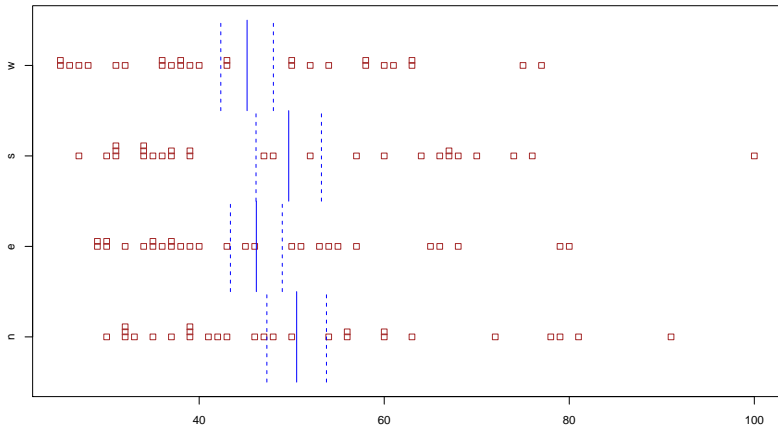
Stripchart der Korkdicken je nach Himmelsrichtung



Stripchart der Korkdicken je nach Himmelsrichtung **mit Mittelwerten**



Stripchart der Korkdicken je nach Himmelsrichtung **mit Mittelwerten und Mittelwerten  $\pm$  Standardfehler**



Stripchart der Korkdicken je nach Himmelsrichtung mit  
 Mittelwerten und Mittelwerten  $\pm$  Standardfehler  
 Kann da was signifikant unterschiedlich sein???

Haben wir irgend etwas übersehen?

Haben wir irgend etwas übersehen?

Wir haben bisher vernachlässigt  
welche Werte von demselben Baum kommen!



Haben wir irgend etwas übersehen?

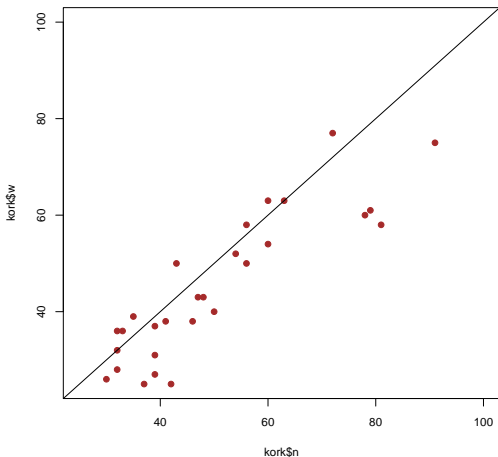
Wir haben bisher vernachlässigt  
welche Werte von demselben Baum kommen!

Die Bäume unterscheiden sich sehr in ihrer Größe und Dicke.

Haben wir irgend etwas übersehen?

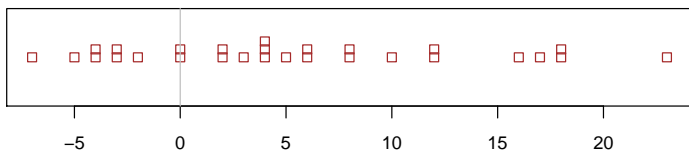
Wir haben bisher vernachlässigt  
welche Werte von demselben Baum kommen!  
Die Bäume unterscheiden sich sehr in ihrer Größe und Dicke.

Vergleiche also jeweils Paare von Korkdicken,  
die von demselben Baum kommen!  
( $\rightsquigarrow$  gepaarter t-Test)

Korkdicken [mm] bei  $n = 28$  BäumenKorkdicke an  
der Westseite

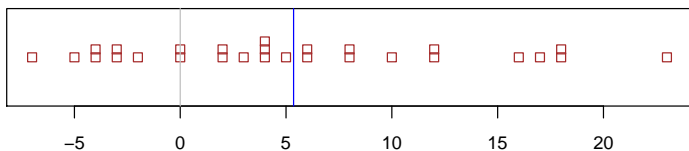
Korkdicke an der Nordseite

Differenz der Korkdicken an der Nord- und der Westseite für  
 $n = 28$  Bäume



Ist die Differenz signifikant von 0  
verschieden?

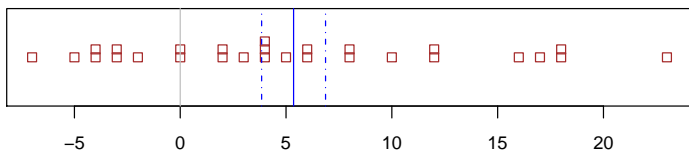
# Differenz der Korkdicken an der Nord- und der Westseite für $n = 28$ Bäume



mit Mittelwert

Ist die Differenz signifikant von 0  
verschieden?

## Differenz der Korkdicken an der Nord- und der Westseite für $n = 28$ Bäume



mit Mittelwert und Mittelwert  $\pm$  Standardfehler  
Ist die Differenz signifikant von 0  
verschieden?

$$x := (\text{Korkdicke Nordseite}) - (\text{Korkdicke Westseite})$$

$$x := (\text{Korkdicke Nordseite}) - (\text{Korkdicke Westseite})$$
$$\bar{x} \approx 5.36$$



$x := (\text{Korkdicke Nordseite}) - (\text{Korkdicke Westseite})$

$$\bar{x} \approx 5.36$$

$$s_x \approx 7.99$$

$x :=$  (Korkdicke Nordseite)  $-$  (Korkdicke Westseite)

$$\bar{x} \approx 5.36$$

$$s_x \approx 7.99$$

$$\frac{s_x}{\sqrt{n}} \approx 1.51$$

$x :=$  (Korkdicke Nordseite)  $-$  (Korkdicke Westseite)

$$\bar{x} \approx 5.36$$

$$s_x \approx 7.99$$

$$\frac{s_x}{\sqrt{n}} \approx 1.51$$

$t$  – Wert =

$x :=$  (Korkdicke Nordseite)  $-$  (Korkdicke Westseite)

$$\bar{x} \approx 5.36$$

$$s_x \approx 7.99$$

$$\frac{s_x}{\sqrt{n}} \approx 1.51$$

$$t\text{-Wert} = \frac{\bar{x}}{s_x/\sqrt{n}} \approx 3.547$$

$x :=$  (Korkdicke Nordseite)  $-$  (Korkdicke Westseite)

$$\bar{x} \approx 5.36$$

$$s_x \approx 7.99$$

$$\frac{s_x}{\sqrt{n}} \approx 1.51$$

$$t\text{-Wert} = \frac{\bar{x}}{s_x/\sqrt{n}} \approx 3.547$$

Anzahl Freiheitsgrade:  $df =$

$x :=$  (Korkdicke Nordseite)  $-$  (Korkdicke Westseite)

$$\bar{x} \approx 5.36$$

$$s_x \approx 7.99$$

$$\frac{s_x}{\sqrt{n}} \approx 1.51$$

$$t\text{-Wert} = \frac{\bar{x}}{s_x/\sqrt{n}} \approx 3.547$$

$$\text{Anzahl Freiheitsgrade: } df = n - 1 = 27$$

```
> t.test(kork$n-kork$w)
```

One Sample t-test

```
data: kork$n - kork$w
```

```
t = 3.5471, df = 27, p-value = 0.001447
```

```
alternative hypothesis: true mean is not equal to 0
```

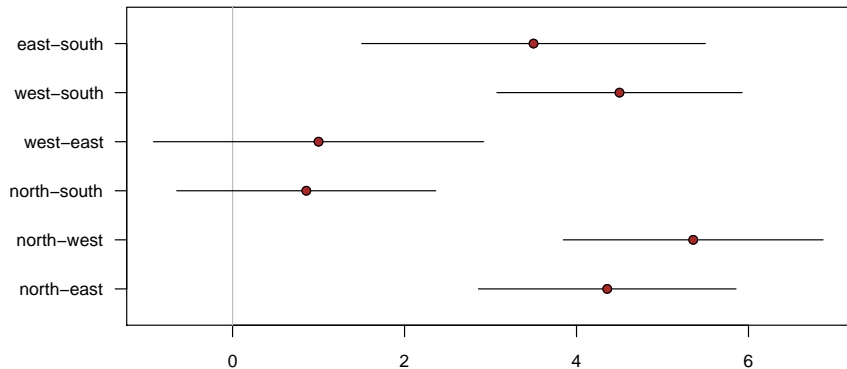
```
95 percent confidence interval:
```

```
 2.258274 8.456012
```

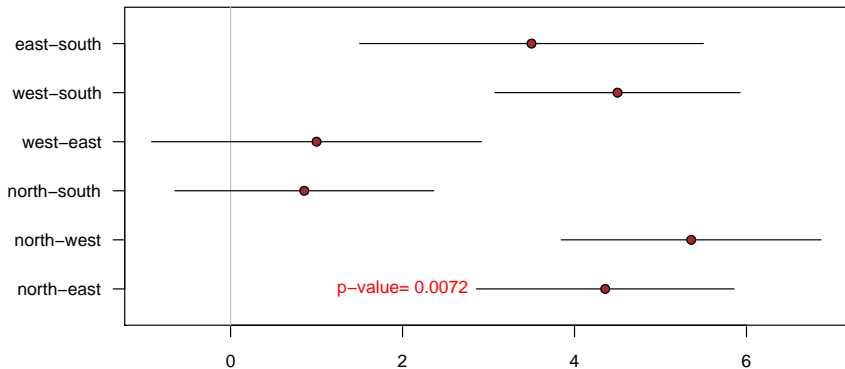
```
sample estimates:
```

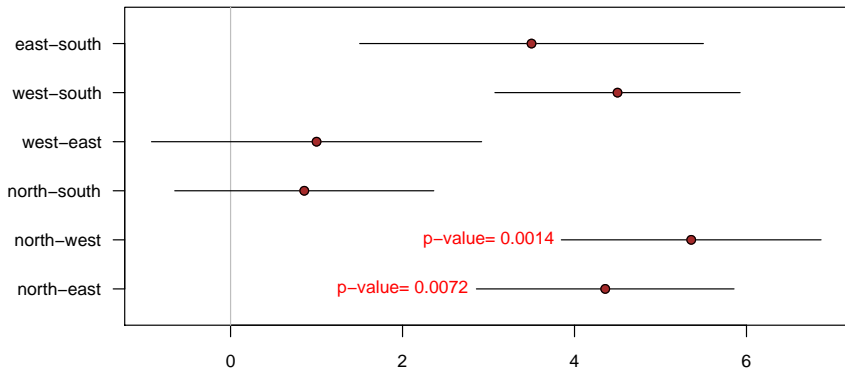
```
mean of x
```

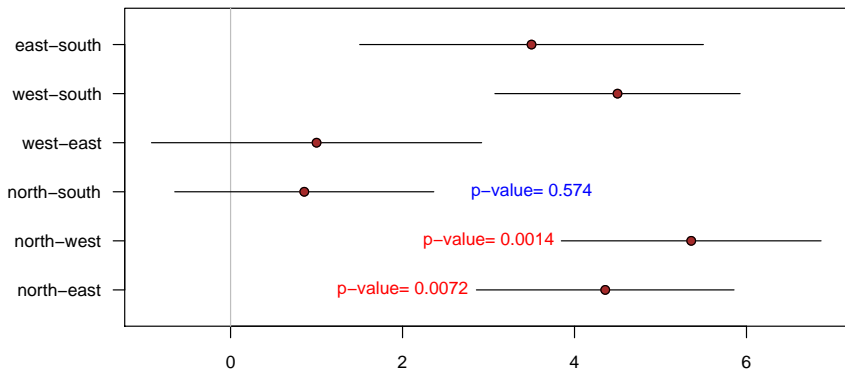
```
5.357143
```

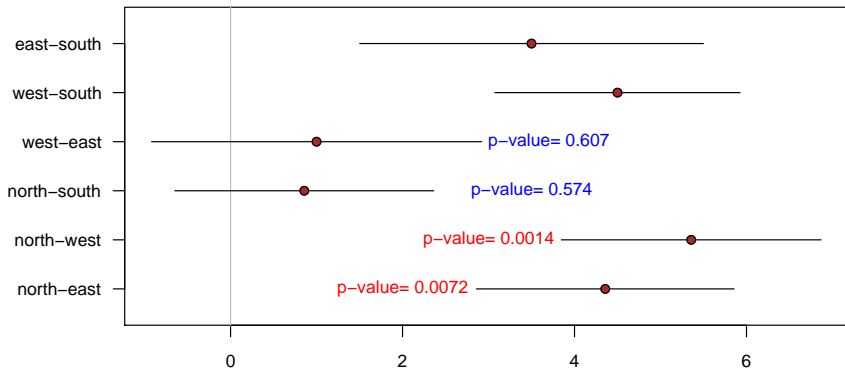


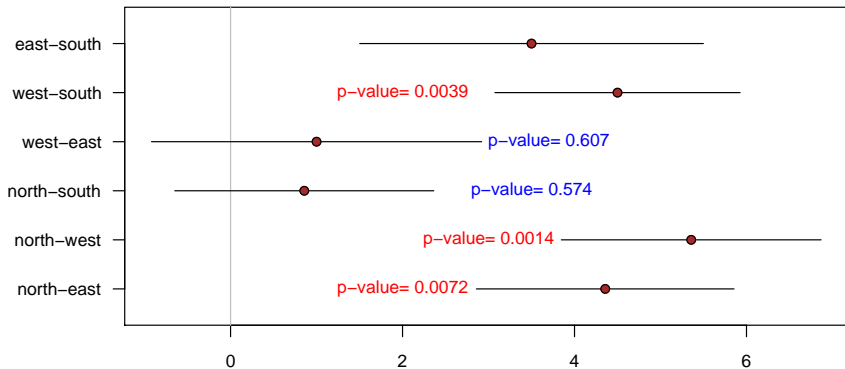


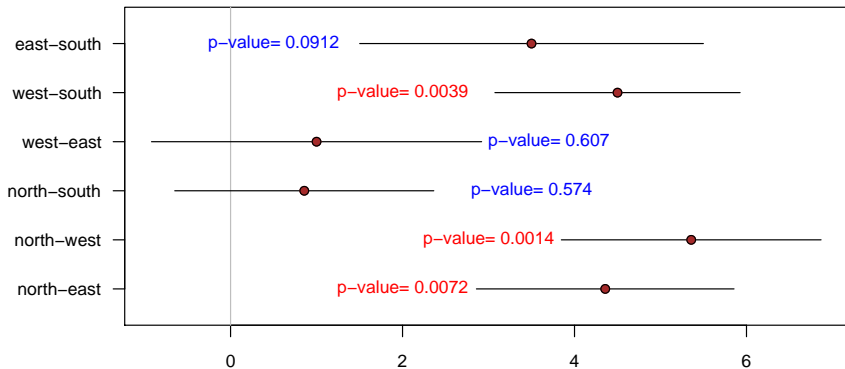












# Inhalt

## 1 t-Test für gepaarte Stichproben

- Beispiel: Orientierung bei Trauerschnäppern
- Die t-Verteilung
- Zurück zu: Orientierung bei Trauerschnäppern
- Beispiel: Richtungsabhängige Korkdicke
- **Zusammenfassung t-Test**
- Prinzip des statistischen Testens

# Zusammenfassung gepaarter t-Test

**Gegeben:** gepaarte Beobachtungen

$$(Y_1, Z_1), (Y_2, Z_2), \dots, (Y_n, Z_n)$$



# Zusammenfassung gepaarter t-Test

**Gegeben:** gepaarte Beobachtungen

$$(Y_1, Z_1), (Y_2, Z_2), \dots, (Y_n, Z_n)$$

**Nullhypothese**  $H_0: \mu_Y = \mu_Z$

**Signifikanzniveau:**  $\alpha$  (meist  $\alpha = 5\%$ )

# Zusammenfassung gepaarter t-Test

**Gegeben:** gepaarte Beobachtungen

$$(Y_1, Z_1), (Y_2, Z_2), \dots, (Y_n, Z_n)$$

**Nullhypothese**  $H_0: \mu_Y = \mu_Z$

**Signifikanzniveau:**  $\alpha$  (meist  $\alpha = 5\%$ )

**Test:** gepaarter t-Test (genauer: zweiseitiger gepaarter t-Test)

Berechne Differenz  $X := Y - Z$

Berechne Teststatistik

$$t := \frac{\bar{X}}{s(X)/\sqrt{n}}$$

# Zusammenfassung gepaarter t-Test

**Gegeben:** gepaarte Beobachtungen

$$(Y_1, Z_1), (Y_2, Z_2), \dots, (Y_n, Z_n)$$

**Nullhypothese**  $H_0: \mu_Y = \mu_Z$

**Signifikanzniveau:**  $\alpha$  (meist  $\alpha = 5\%$ )

**Test:** gepaarter t-Test (genauer: zweiseitiger gepaarter t-Test)

Berechne Differenz  $X := Y - Z$

Berechne Teststatistik

$$t := \frac{\bar{X}}{s(X)/\sqrt{n}}$$

p-Wert =  $\Pr(|T_{n-1}| \geq |t|)$  ( $n - 1$  Freiheitsgrade)

Verwirf Nullhypothese, falls p-Wert  $\leq \alpha$

# Zusammenfassung Ein-Stichproben t-Test

**Gegeben:** Beobachtungen

$$X_1, X_2, \dots, X_n$$

# Zusammenfassung Ein-Stichproben t-Test

**Gegeben:** Beobachtungen

$$X_1, X_2, \dots, X_n$$

**Nullhypothese**  $H_0: \mu_X = c$  (Den Wert  $c$  kennt man, oft  $c = 0$ )

**Signifikanzniveau:**  $\alpha$  (meist  $\alpha = 5\%$ )

# Zusammenfassung Ein-Stichproben t-Test

**Gegeben:** Beobachtungen

$$X_1, X_2, \dots, X_n$$

**Nullhypothese**  $H_0: \mu_X = c$  (Den Wert  $c$  kennt man, oft  $c = 0$ )

**Signifikanzniveau:**  $\alpha$  (meist  $\alpha = 5\%$ )

**Test:** t-Test

Berechne Teststatistik

$$t := \frac{\bar{X} - c}{s(X)/\sqrt{n}}$$

# Zusammenfassung Ein-Stichproben t-Test

**Gegeben:** Beobachtungen

$$X_1, X_2, \dots, X_n$$

**Nullhypothese**  $H_0: \mu_X = c$  (Den Wert  $c$  kennt man, oft  $c = 0$ )

**Signifikanzniveau:**  $\alpha$  (meist  $\alpha = 5\%$ )

**Test:** t-Test

Berechne Teststatistik

$$t := \frac{\bar{X} - c}{s(X)/\sqrt{n}}$$

p-Wert =  $\Pr(|T_{n-1}| \geq |t|)$  ( $n - 1$  Freiheitsgrade)

Verwirf Nullhypothese, falls p-Wert  $\leq \alpha$

# Inhalt

## 1 t-Test für gepaarte Stichproben

- Beispiel: Orientierung bei Trauerschnäppern
- Die t-Verteilung
- Zurück zu: Orientierung bei Trauerschnäppern
- Beispiel: Richtungsabhängige Korkdicke
- Zusammenfassung t-Test
- **Prinzip des statistischen Testens**



# Beispiel: Codon Bias

- Wir beobachten 101844 mal CCT und 106159 mal CCA

# Beispiel: Codon Bias

- Wir beobachten 101844 mal CCT und 106159 mal CCA
- Wenn beide eigentlich gleich wahrscheinlich sind, erwarten wir 104001.5 von jedem.

# Beispiel: Codon Bias

- Wir beobachten 101844 mal CCT und 106159 mal CCA
- Wenn beide eigentlich gleich wahrscheinlich sind, erwarten wir 104001.5 von jedem.
- Die Beobachtung weicht um 2156 von diesem Erwartungswert ab

# Beispiel: Codon Bias

- Wir beobachten 101844 mal CCT und 106159 mal CCA
- Wenn beide eigentlich gleich wahrscheinlich sind, erwarten wir 104001.5 von jedem.
- Die Beobachtung weicht um 2156 von diesem Erwartungswert ab
- z-Test: Die Wahrscheinlichkeit einer mindestens so großen Abweichung ist kleiner als  $10^{-20}$

# Beispiel: Codon Bias

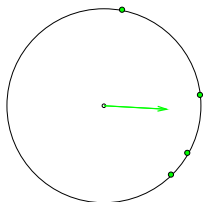
- Wir beobachten 101844 mal CCT und 106159 mal CCA
- Wenn beide eigentlich gleich wahrscheinlich sind, erwarten wir 104001.5 von jedem.
- Die Beobachtung weicht um 2156 von diesem Erwartungswert ab
- z-Test: Die Wahrscheinlichkeit einer mindestens so großen Abweichung ist kleiner als  $10^{-20}$
- Also sind CCT und CCA wohl nicht gleich wahrscheinlich.

# Beispiel: Zugvogelorientierung

- Wie variabel ist die Abflugrichtung bei grünem und bei blauem Licht.

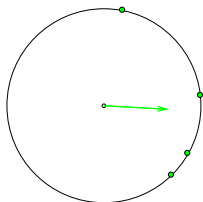
# Beispiel: Zugvogelorientierung

- Wie variabel ist die Abflugrichtung bei grünem und bei blauem Licht.
- Wir messen die Variabilität durch die Länge des Schwerpunktsvektors.



# Beispiel: Zugvogelorientierung

- Wie variabel ist die Abflugrichtung bei grünem und bei blauem Licht.
- Wir messen die Variabilität durch die Länge des Schwerpunktsvektors.

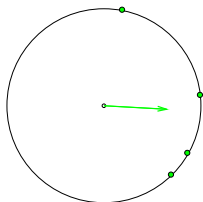


- Quantifiziere Unterschied durch  $X = (\text{Länge grün}) - (\text{Länge blau})$ .



# Beispiel: Zugvogelorientierung

- Wie variabel ist die Abflugrichtung bei grünem und bei blauem Licht.
- Wir messen die Variabilität durch die Länge des Schwerpunktsvektors.



- Quantifiziere Unterschied durch  $X = (\text{Länge grün}) - (\text{Länge blau})$ .
- Wenn das Licht keinen Einfluss hat, gilt  $\mathbb{E}X = 0$ .

# Beispiel: Zugvogelorientierung

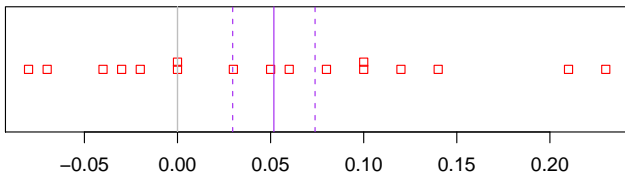
$X = (\text{Länge grün}) - (\text{Länge blau})$

- Wenn das Licht keinen Einfluss hat, gilt  $\mathbb{E}X = 0$ .

# Beispiel: Zugvogelorientierung

$X = (\text{Länge grün}) - (\text{Länge blau})$

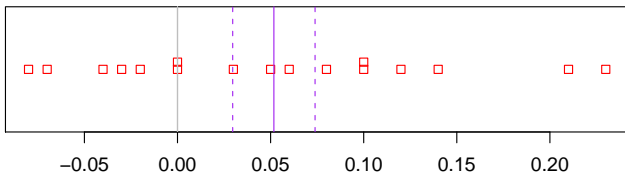
- Wenn das Licht keinen Einfluss hat, gilt  $\mathbb{E}X = 0$ .
- Wir beobachten aber  $\bar{X} = 0.0518$  und  $\text{SEM} = 0.022$



# Beispiel: Zugvogelorientierung

$X = (\text{Länge grün}) - (\text{Länge blau})$

- Wenn das Licht keinen Einfluss hat, gilt  $\mathbb{E}X = 0$ .
- Wir beobachten aber  $\bar{X} = 0.0518$  und  $\text{SEM} = 0.022$

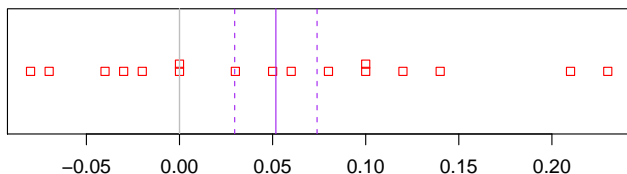


- $t$ -Test:  $p$ -Wert dieser Abweichung ist ca. 3.3%.

# Beispiel: Zugvogelorientierung

$X = (\text{Länge grün}) - (\text{Länge blau})$

- Wenn das Licht keinen Einfluss hat, gilt  $\mathbb{E}X = 0$ .
- Wir beobachten aber  $\bar{X} = 0.0518$  und  $\text{SEM} = 0.022$



- $t$ -Test:  $p$ -Wert dieser Abweichung ist ca. 3.3%.
- Vermutlich hat die Lichtfarbe also doch einen Einfluß

# Beispiel: Dicke des Korks

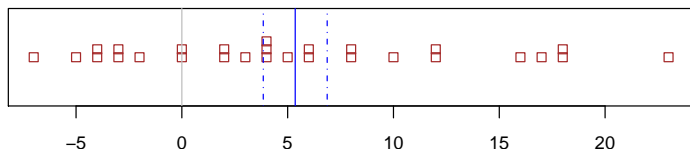
- $X = (\text{Korkdicke an der Nordseite}) - (\text{Korkdicke an der Westseite})$

# Beispiel: Dicke des Korks

- $X = (\text{Korkdicke an der Nordseite}) - (\text{Korkdicke an der Westseite})$
- Wenn die Seite keine Rolle spielt, ist  $\mathbb{E}X = 0$ .

# Beispiel: Dicke des Korks

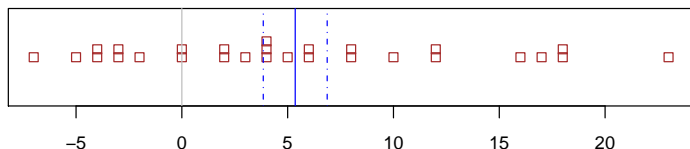
- $X = (\text{Korkdicke an der Nordseite}) - (\text{Korkdicke an der Westseite})$
- Wenn die Seite keine Rolle spielt, ist  $\mathbb{E}X = 0$ .
- Wir sehen aber  $\bar{X} = 5.36$  und  $\text{SEM} = 1.51$





# Beispiel: Dicke des Korks

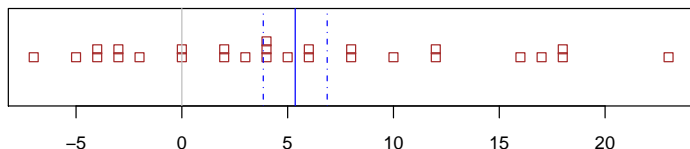
- $X = (\text{Korkdicke an der Nordseite}) - (\text{Korkdicke an der Westseite})$
- Wenn die Seite keine Rolle spielt, ist  $\mathbb{E}X = 0$ .
- Wir sehen aber  $\bar{X} = 5.36$  und  $\text{SEM} = 1.51$



- *t*-Test: *p*-Wert dieser Abweichung ist ca. 0.14%.

# Beispiel: Dicke des Korks

- $X = (\text{Korkdicke an der Nordseite}) - (\text{Korkdicke an der Westseite})$
- Wenn die Seite keine Rolle spielt, ist  $\mathbb{E}X = 0$ .
- Wir sehen aber  $\bar{X} = 5.36$  und  $\text{SEM} = 1.51$



- *t*-Test: *p*-Wert dieser Abweichung ist ca. 0.14%.
- Also hat die Himmelsrichtung wohl doch einen Einfluß.

# Prinzip des statistischen Testens

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

# Prinzip des statistischen Testens

- 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.

# Prinzip des statistischen Testens

- 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 sind **Abweichungen**, die mindestens so groß sind wie die beobachtete, sehr unwahrscheinlich.

# Prinzip des statistischen Testens

- 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 sind **Abweichungen**, die mindestens so groß sind wie die beobachtete, sehr unwahrscheinlich.
- Wenn uns das gelingt, verwerfen wir  **$H_0$** .

# Prinzip des statistischen Testens

- 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 sind **Abweichungen**, die mindestens so groß sind wie die beobachtete, sehr unwahrscheinlich.
- Wenn uns das gelingt, verwerfen wir  **$H_0$** .
- Was wir als **Abweichung** auffassen, sollte klar sein, bevor wir die Daten sehen.

# Nullhypothesen

- $H_0$  bei Codon-Bias: CCT und CCA haben jeweils W'keit  $\frac{1}{2}$



# Nullhypothesen

- $H_0$  bei Codon-Bias: CCT und CCA haben jeweils W'keit  $\frac{1}{2}$   
Außerdem: alle Positionen entscheiden unabhängig  
zwischen CCT und CCA

# Nullhypothesen

- $H_0$  bei Codon-Bias: CCT und CCA haben jeweils  $W$ 'keit  $\frac{1}{2}$   
Außerdem: alle Positionen entscheiden unabhängig  
zwischen CCT und CCA
- $H_0$  bei Vogelorientierung und Korkdicken:  $\mathbb{E}X = 0$ .

# Nullhypothesen

- $H_0$  bei Codon-Bias: CCT und CCA haben jeweils  $W$ 'keit  $\frac{1}{2}$   
Außerdem: alle Positionen entscheiden unabhängig  
zwischen CCT und CCA
- $H_0$  bei Vogelorientierung und Korkdicken:  $\mathbb{E}X = 0$ .  
Außerdem:  $X$  normalverteilt,  $X_i$  unabhängig.

# Abweichungen und $p$ -Werte

- Codon Bias: Anzahl CCT weicht um 2156 vom Mittelwert ab.

# Abweichungen und $p$ -Werte

- Codon Bias: Anzahl CCT weicht um 2156 vom Mittelwert ab.  
Wegen der Binomialverteilungsannahme gehen wir von festem  $\sigma$  aus und berechnen mit dem z-Test

# Abweichungen und $p$ -Werte

- Codon Bias: Anzahl CCT weicht um 2156 vom Mittelwert ab.

Wegen der Binomialverteilungsannahme gehen wir von festem  $\sigma$  aus und berechnen mit dem  $z$ -Test den  $p$ -Wert: Die Wahrscheinlichkeit, dass eine  $\text{bin}(n, \frac{1}{2})$ -verteilte Zufallsgröße um mindestens 2156 von  $n/2$  abweicht.

# Abweichungen und $p$ -Werte

- Codon Bias: Anzahl CCT weicht um 2156 vom Mittelwert ab.

Wegen der Binomialverteilungsannahme gehen wir von festem  $\sigma$  aus und berechnen mit dem  $z$ -Test den  $p$ -Wert: Die Wahrscheinlichkeit, dass eine  $\text{bin}(n, \frac{1}{2})$ -verteilte Zufallsgröße um mindestens 2156 von  $n/2$  abweicht.

- Vogelorientierung und Korkdicke:

$$t\text{-Wert} = \frac{\bar{X}}{s/\sqrt{n}}$$

# Abweichungen und $p$ -Werte

- Codon Bias: Anzahl CCT weicht um 2156 vom Mittelwert ab.

Wegen der Binomialverteilungsannahme gehen wir von festem  $\sigma$  aus und berechnen mit dem  $z$ -Test den  $p$ -Wert: Die Wahrscheinlichkeit, dass eine  $\text{bin}(n, \frac{1}{2})$ -verteilte Zufallsgröße um mindestens 2156 von  $n/2$  abweicht.

- Vogelorientierung und Korkdicke:

$$t\text{-Wert} = \frac{\bar{X}}{s/\sqrt{n}}$$

$p$ -Wert: W'keit, dass  $t$ -Wert bei  $n - 1$  mindestens so stark von 0 abweicht wie beobachtet.



# Zweiseitig oder einseitig testen?

In den meisten Fällen will man testen, ob zwei Stichproben sich signifikant unterscheiden.

↪ **zweiseitiger Test**

# Zweiseitig oder einseitig testen?

In den meisten Fällen will man testen, ob zwei Stichproben sich signifikant unterscheiden.

↪ **zweiseitiger Test**

In manchen Fällen

- kann man von vornherein ausschließen, dass die erste Stichprobe kleinere Werte als die zweite Stichprobe hat. Dann will man testen, ob die erste Stichprobe signifikant größer ist.

# Zweiseitig oder einseitig testen?

In den meisten Fällen will man testen, ob zwei Stichproben sich signifikant unterscheiden.

↪ **zweiseitiger Test**

In manchen Fällen

- kann man von vornherein ausschließen, dass die erste Stichprobe kleinere Werte als die zweite Stichprobe hat. Dann will man testen, ob die erste Stichprobe signifikant größer ist.
- will man nur testen, ob die erste Stichprobe signifikant größer ist.

# Zweiseitig oder einseitig testen?

In den meisten Fällen will man testen, ob zwei Stichproben sich signifikant unterscheiden.

↪ **zweiseitiger Test**

In manchen Fällen

- kann man von vornherein ausschließen, dass die erste Stichprobe kleinere Werte als die zweite Stichprobe hat. Dann will man testen, ob die erste Stichprobe signifikant größer ist.
- will man nur testen, ob die erste Stichprobe signifikant größer ist.
- will man nur testen, ob die erste Stichprobe signifikant kleiner ist.

↪ **einseitiger Test**

## Beispiel für einseitigen Test:

Mann will zeigen,  
dass ein Wachstumshormon wirkt,  
also kein Placebo ist.

## Beispiel für einseitigen Test:

Mann will zeigen,  
dass ein Wachstumshormon wirkt,  
also kein Placebo ist.

Dazu müssen die Größen  $Y$  in der behandelten Gruppe  
signifikant größer sein  
als die Größen  $Z$  in der Kontrollgruppe.

## Beispiel für einseitigen Test:

Mann will zeigen,  
dass ein Wachstumshormon wirkt,  
also kein Placebo ist.

Dazu müssen die Größen  $Y$  in der behandelten Gruppe  
signifikant größer sein  
als die Größen  $Z$  in der Kontrollgruppe.  
Die zu entkräftende Nullhypothese wäre hier:

$$\text{Nullhypothese } \mu_Y \leq \mu_Z$$

## Beispiel für einseitigen Test:

Mann will zeigen,  
dass ein Wachstumshormon wirkt,  
also kein Placebo ist.

Dazu müssen die Größen  $Y$  in der behandelten Gruppe  
signifikant größer sein  
als die Größen  $Z$  in der Kontrollgruppe.  
Die zu entkräftende Nullhypothese wäre hier:

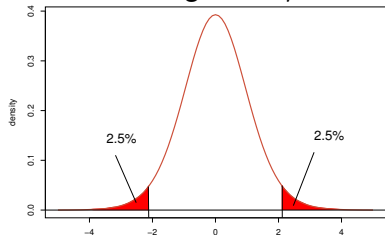
$$\text{Nullhypothese } \mu_Y \leq \mu_Z$$

Definiere die Differenz  $X := Y - Z$ .

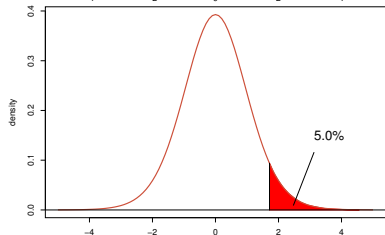


# Zweiseitig oder einseitig testen?

Wir beobachten einen Wert  $x$ , der deutlich größer als der  $H_0$ -Erwartungswert  $\mu$  ist.



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



$$p\text{-Wert} = \Pr_{H_0}(X - \mu \geq x - \mu)$$

# 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**  $\alpha$  fest; üblich ist  $\alpha = 0.05$ .

# Reine Lehre des statistischen Testens

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

$$\Pr_{H_0}(\mathcal{A}) = \alpha$$

(oder zumindest  $\Pr_{H_0}(\mathcal{A}) \leq \alpha$ ).

# Reine Lehre des statistischen Testens

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

$$\Pr_{H_0}(\mathcal{A}) = \alpha$$

(oder zumindest  $\Pr_{H_0}(\mathcal{A}) \leq \alpha$ ).

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**  $\alpha$  fest; üblich ist  $\alpha = 0.05$ .
- Lege ein Ereignis  $\mathcal{A}$  fest, so dass

$$\Pr_{H_0}(\mathcal{A}) = \alpha$$

(oder zumindest  $\Pr_{H_0}(\mathcal{A}) \leq \alpha$ ).

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

- **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**  $\alpha$  fest; üblich ist  $\alpha = 0.05$ .
- Lege ein Ereignis  $\mathcal{A}$  fest, so dass

$$\Pr_{H_0}(\mathcal{A}) = \alpha$$

(oder zumindest  $\Pr_{H_0}(\mathcal{A}) \leq \alpha$ ).

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

- **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  $\alpha$ .

# 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  $\mu_{H_0}$ . Also habe ich gleich einseitig getestet”

## Wichtig

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

## Wichtig

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

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 die Daten betrachtet hat.

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

$$\Pr_{H_0}(\mathcal{A}) = \alpha$$

und

$$\Pr_{H_1}(\mathcal{A}) = \text{möglichst groß,}$$

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

$$\Pr_{H_0}(\mathcal{A}) = \alpha$$

und

$$\Pr_{H_1}(\mathcal{A}) = \text{möglichst groß,}$$

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

# 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. ✓