

# Statistik

## 03. Nichtparametrische Verfahren

Roland Schäfer

Institut für Germanistische Sprachwissenschaft  
Friedrich-Schiller-Universität Jena

stets aktuelle Fassungen: <https://github.com/rsling/VL-Deutsche-Syntax>

- 1 Testverfahren für Zähldaten
  - Vierfelder- $\chi^2$ -Unterschiedstest
  - Fisher-Exakt-Test
  - Effektstärke: Cramér's  $v$  und  $\phi$

- Chancenverhältnis
- Binomialtest

- 2 Nächste Woche | Überblick

Zähldaten

- Unterschiede in Zähldaten

- Unterschiede in Zähldaten
- Signifikanz und Effektstärke

- Unterschiede in Zähldaten
- Signifikanz und Effektstärke
- Unterschiede bei Ja/Nein-Experimenten

- Gravetter & Wallnau 2007
- Bortz & Lienert 2008

Beobachtungen von zwei kategorialen Variablen.

Auxiliarwahl beim Perfekt: haben, sein

Herkunft des Belegs: nord, sued

Fall	Aux	Region
1	haben	nord
2	haben	nord
3	sein	nord
4	sein	sued
5	sein	sued
6	haben	nord
7	haben	sued
8	haben	sued



# Kreuztabelle

Beobachtungen von zwei kategorialen Variablen.

Auxiliarwahl beim Perfekt: haben, sein

Herkunft des Belegs: nord, sued

Fall	Aux	Region
1	haben	nord
2	haben	nord
3	sein	nord
4	sein	sued
5	sein	sued
6	haben	nord
7	haben	sued
8	haben	sued

	Aux	
Region	haben	sein
nord	3	1
sued	2	2

# Kreuztabelle mit Randsummen

Spaltensumme für Spalte  $i$ :  $\sum_k x_{ik}$

Zeilensumme für Zeile  $j$ :  $\sum_k x_{kj}$

	haben	sein	Zeilensummen
nord	3	1	
sued	2	2	
Spaltensummen			

# Kreuztabelle mit Randsummen

Spaltensumme für Spalte  $i$ :  $\sum_k x_{ik}$

Zeilensumme für Zeile  $j$ :  $\sum_k x_{kj}$

	haben	sein	Zeilensummen
nord	3	1	4
sued	2	2	
Spaltensummen			

# Kreuztabelle mit Randsummen

Spaltensumme für Spalte  $i$ :  $\sum_k x_{ik}$

Zeilensumme für Zeile  $j$ :  $\sum_k x_{kj}$

	haben	sein	Zeilensummen
nord	3	1	4
sued	2	2	4
Spaltensummen			

# Kreuztabelle mit Randsummen

Spaltensumme für Spalte  $i$ :  $\sum_k x_{ik}$

Zeilensumme für Zeile  $j$ :  $\sum_k x_{kj}$

	haben	sein	Zeilensummen
nord	3	1	4
sued	2	2	4
Spaltensummen	5		

# Kreuztabelle mit Randsummen

Spaltensumme für Spalte  $i$ :  $\sum_k x_{ik}$

Zeilensumme für Zeile  $j$ :  $\sum_k x_{kj}$

	haben	sein	Zeilensummen
nord	3	1	4
sued	2	2	4
Spaltensummen	5	3	

# Kreuztabelle mit Randsummen

Spaltensumme für Spalte  $i$ :  $\sum_k x_{ik}$

Zeilensumme für Zeile  $j$ :  $\sum_k x_{kj}$

	haben	sein	Zeilensummen
nord	3	1	4
sued	2	2	4
Spaltensummen	5	3	8

# Beobachtete vs. erwartete Häufigkeiten

$n=100$

50 mal *haben*, 50 mal *sein* (= Spaltensummen)

50 mal Norden, 50 mal Süden (= Zeilensummen)

	haben	sein	Zeilensummen
nord			50
sued			50
Spaltensummen	50	50	100



# Beobachtete vs. erwartete Häufigkeiten

$n=100$

50 mal *haben*, 50 mal *sein* (= Spaltensummen)

50 mal Norden, 50 mal Süden (= Zeilensummen)

- erwartete Häufigkeiten unter Annahme der  $H_0$   
= kein Zusammenhang zwischen Hilfsverb und Region?

	haben	sein	Zeilensummen
nord			50
sued			50
Spaltensummen	50	50	100

# Beobachtete vs. erwartete Häufigkeiten

$n=100$

50 mal *haben*, 50 mal *sein* (= Spaltensummen)

50 mal Norden, 50 mal Süden (= Zeilensummen)

- erwartete Häufigkeiten unter Annahme der  $H_0$   
= kein Zusammenhang zwischen Hilfsverb und Region?

	haben	sein	Zeilensummen
nord	25	25	50
sued	25	25	50
Spaltensummen	50	50	100

# Beobachtete vs. erwartete Häufigkeiten

$n=100$

50 mal *haben*, 50 mal *sein* (= Spaltensummen)

30 mal Norden, 70 mal Süden (= Zeilensummen)

- erwartete Häufigkeiten unter Annahme der  $H_0$ ?

	haben	sein	Zeilensummen
nord			
sued			
Spaltensummen			<b>100</b>

# Beobachtete vs. erwartete Häufigkeiten

$n=100$

50 mal *haben*, 50 mal *sein* (= Spaltensummen)

30 mal Norden, 70 mal Süden (= Zeilensummen)

- erwartete Häufigkeiten unter Annahme der  $H_0$ ?

	haben	sein	Zeilensummen
nord			30
sued			
Spaltensummen			100

# Beobachtete vs. erwartete Häufigkeiten

$n=100$

50 mal *haben*, 50 mal *sein* (= Spaltensummen)

30 mal Norden, 70 mal Süden (= Zeilensummen)

- erwartete Häufigkeiten unter Annahme der  $H_0$ ?

	haben	sein	Zeilensummen
nord			30
sued			70
Spaltensummen			100

# Beobachtete vs. erwartete Häufigkeiten

$n=100$

50 mal *haben*, 50 mal *sein* (= Spaltensummen)

30 mal Norden, 70 mal Süden (= Zeilensummen)

- erwartete Häufigkeiten unter Annahme der  $H_0$ ?

	haben	sein	Zeilensummen
nord			30
sued			70
Spaltensummen	50		100

# Beobachtete vs. erwartete Häufigkeiten

$n=100$

50 mal *haben*, 50 mal *sein* (= Spaltensummen)

30 mal Norden, 70 mal Süden (= Zeilensummen)

- erwartete Häufigkeiten unter Annahme der  $H_0$ ?

	haben	sein	Zeilensummen
nord			30
sued			70
Spaltensummen	50	50	100

# Beobachtete vs. erwartete Häufigkeiten

$n=100$

50 mal *haben*, 50 mal *sein* (= Spaltensummen)

30 mal Norden, 70 mal Süden (= Zeilensummen)

- erwartete Häufigkeiten unter Annahme der  $H_0$ ?

	haben	sein	Zeilensummen
nord	15		30
sued			70
Spaltensummen	50	50	100



# Beobachtete vs. erwartete Häufigkeiten

$n=100$

50 mal *haben*, 50 mal *sein* (= Spaltensummen)

30 mal Norden, 70 mal Süden (= Zeilensummen)

- erwartete Häufigkeiten unter Annahme der  $H_0$ ?

	haben	sein	Zeilensummen
nord	15	15	30
sued			70
Spaltensummen	50	50	100

# Beobachtete vs. erwartete Häufigkeiten

$n=100$

50 mal *haben*, 50 mal *sein* (= Spaltensummen)

30 mal Norden, 70 mal Süden (= Zeilensummen)

- erwartete Häufigkeiten unter Annahme der  $H_0$ ?

	haben	sein	Zeilensummen
nord	15	15	30
sued	35		70
Spaltensummen	50	50	100

# Beobachtete vs. erwartete Häufigkeiten

$n=100$

50 mal *haben*, 50 mal *sein* (= Spaltensummen)

30 mal Norden, 70 mal Süden (= Zeilensummen)

- erwartete Häufigkeiten unter Annahme der  $H_0$ ?

	haben	sein	Zeilensummen
nord	15	15	30
sued	35	35	70
Spaltensummen	50	50	100

# Beobachtete vs. erwartete Häufigkeiten

n=100

30 mal Norden, 70 mal Süden

40 mal *haben*, 60 mal *sein*

	haben	sein	Zeilensummen
nord			30
sued			70
Spaltensummen	40	60	100

Allgemein: erwartete Häufigkeit für Zellen:  $\frac{\text{Spaltensumme} \cdot \text{Zeilensumme}}{n}$

$$\text{bzw.: } EH(x_{ij}) = \frac{\sum_k x_{ik} \cdot \sum_k x_{kj}}{n}$$

# Beobachtete vs. erwartete Häufigkeiten

n=100

30 mal Norden, 70 mal Süden

40 mal *haben*, 60 mal *sein*

	haben	sein	Zeilensummen
nord	12		30
sued			70
Spaltensummen	40	60	100

Allgemein: erwartete Häufigkeit für Zellen:  $\frac{\text{Spaltensumme} \cdot \text{Zeilensumme}}{n}$

$$\text{bzw.: } EH(x_{ij}) = \frac{\sum_k x_{ik} \cdot \sum_k x_{kj}}{n}$$

# Beobachtete vs. erwartete Häufigkeiten

n=100

30 mal Norden, 70 mal Süden

40 mal *haben*, 60 mal *sein*

	haben	sein	Zeilensummen
nord	12	18	30
sued			70
Spaltensummen	40	60	100

Allgemein: erwartete Häufigkeit für Zellen:  $\frac{\text{Spaltensumme} \cdot \text{Zeilensumme}}{n}$

$$\text{bzw.: } EH(x_{ij}) = \frac{\sum_k x_{ik} \cdot \sum_k x_{kj}}{n}$$

# Beobachtete vs. erwartete Häufigkeiten

n=100

30 mal Norden, 70 mal Süden

40 mal *haben*, 60 mal *sein*

	haben	sein	Zeilensummen
nord	12	18	30
sued	28		70
Spaltensummen	40	60	100

Allgemein: erwartete Häufigkeit für Zellen:  $\frac{\text{Spaltensumme} \cdot \text{Zeilensumme}}{n}$

$$\text{bzw.: } EH(x_{ij}) = \frac{\sum_k x_{ik} \cdot \sum_k x_{kj}}{n}$$

# Beobachtete vs. erwartete Häufigkeiten

n=100

30 mal Norden, 70 mal Süden

40 mal *haben*, 60 mal *sein*

	haben	sein	Zeilensummen
nord	12	18	30
sued	28	42	70
Spaltensummen	40	60	100

Allgemein: erwartete Häufigkeit für Zellen:  $\frac{\text{Spaltensumme} \cdot \text{Zeilensumme}}{n}$

$$\text{bzw.: } EH(x_{ij}) = \frac{\sum_k x_{ik} \cdot \sum_k x_{kj}}{n}$$



# Beobachtete vs. erwartete Häufigkeiten

beobachtete Häufigkeiten für eine DeReKo-Stichprobe (*geschwebt*):

	haben	sein	Zeilensummen
nord	27	33	60
sued	3	34	37
Spaltensummen	30	67	97

# Beobachtete vs. erwartete Häufigkeiten

beobachtete Häufigkeiten für eine DeReKo-Stichprobe (*geschwebt*):

	haben	sein	Zeilensummen
nord	27	33	60
sued	3	34	37
Spaltensummen	30	67	97

erwartete Häufigkeiten:

	haben	sein	Zeilensummen
nord			60
sued			37
Spaltensummen	30	67	97

# Beobachtete vs. erwartete Häufigkeiten

beobachtete Häufigkeiten für eine DeReKo-Stichprobe (*geschwebt*):

	haben	sein	Zeilensummen
nord	27	33	60
sued	3	34	37
Spaltensummen	30	67	97

erwartete Häufigkeiten:

	haben	sein	Zeilensummen
nord	18.56	41.44	60
sued	11.44	25.56	37
Spaltensummen	30	67	97

- Beobachtete und erwartete Häufigkeit weichen ab.

- Beobachtete und erwartete Häufigkeit weichen ab.
- $H_0$ : kein Zusammenhang zwischen Region und Aux.

- Beobachtete und erwartete Häufigkeit weichen ab.
- $H_0$ : kein Zusammenhang zwischen Region und Aux.
- Ab wann ist der Unterschied „signifikant“?

- Beobachtete und erwartete Häufigkeit weichen ab.
- $H_0$ : kein Zusammenhang zwischen Region und Aux.
- Ab wann ist der Unterschied „signifikant“?
- Ein gemessener Unterschied ist **signifikant**, wenn er angesichts der Stichprobengröße groß genug ist, dass wir das im Experiment gefundene Ergebnis nur sehr selten (typischerweise in unter 5% der Fälle) erwarten würden, wenn er gar nicht bestünde.

- Beobachtete und erwartete Häufigkeit weichen ab.
- $H_0$ : kein Zusammenhang zwischen Region und Aux.
- Ab wann ist der Unterschied „signifikant“?
- Ein gemessener Unterschied ist **signifikant**, wenn er angesichts der Stichprobengröße groß genug ist, dass wir das im Experiment gefundene Ergebnis nur sehr selten (typischerweise in unter 5% der Fälle) erwarten würden, wenn er gar nicht bestünde.
- Diese 5% (als Anteil 0.05) sind das Signifikanzniveau.



- Beobachtete und erwartete Häufigkeit weichen ab.
- $H_0$ : kein Zusammenhang zwischen Region und Aux.
- Ab wann ist der Unterschied „signifikant“?
- Ein gemessener Unterschied ist **signifikant**, wenn er angesichts der Stichprobengröße groß genug ist, dass wir das im Experiment gefundene Ergebnis nur sehr selten (typischerweise in unter 5% der Fälle) erwarten würden, wenn er gar nicht bestünde.
- Diese 5% (als Anteil 0.05) sind das **Signifikanzniveau**.
- In Fishers Philosophie abgekürzt *sig*, nicht wie oft zu lesen „ $\alpha$ -Niveau“.

# $\chi^2$ -Unterschiedstest

beobachtet:

	haben	sein
nord	27	33
sued	3	34

erwartet:

	haben	sein
nord	18.56	41.44
sued	11.44	25.56

# $\chi^2$ -Unterschiedstest

beobachtet:

	haben	sein
nord	27	33
sued	3	34

erwartet:

	haben	sein
nord	18.56	41.44
sued	11.44	25.56

$$\chi^2 = \sum \frac{(\text{beobachtet} - \text{erwartet})^2}{\text{erwartet}}$$

# $\chi^2$ -Unterschiedstest

beobachtet:

	haben	sein
nord	27	33
sued	3	34

erwartet:

	haben	sein
nord	18.56	41.44
sued	11.44	25.56

$$\chi^2 = \sum \frac{(\text{beobachtet} - \text{erwartet})^2}{\text{erwartet}}$$

$$\text{bzw.: } \chi^2 = \sum_{ij} \frac{(x_{ij} - EH(x_{ij}))^2}{EH(x_{ij})}$$

# Berechnung des $\chi^2$ -Werts

$$\chi^2 = \sum \frac{(\text{beobachtet} - \text{erwartet})^2}{\text{erwartet}}$$

beobachtet:

	haben	sein
nord	27	33
sued	3	34

erwartet:

	haben	sein
nord	18.56	41.44
sued	11.44	25.56

# Berechnung des $\chi^2$ -Werts

$$\chi^2 = \sum \frac{(\text{beobachtet} - \text{erwartet})^2}{\text{erwartet}}$$

beobachtet:

	haben	sein
nord	27	33
sued	3	34

erwartet:

	haben	sein
nord	18.56	41.44
sued	11.44	25.56

$$\chi^2 = \frac{(27 - 18.56)^2}{18.56}$$

# Berechnung des $\chi^2$ -Werts

$$\chi^2 = \sum \frac{(\text{beobachtet} - \text{erwartet})^2}{\text{erwartet}}$$

beobachtet:

	haben	sein
nord	27	33
sued	3	34

erwartet:

	haben	sein
nord	18.56	41.44
sued	11.44	25.56

$$\chi^2 = \frac{(27 - 18.56)^2}{18.56} + \frac{(33 - 41.44)^2}{41.44}$$

# Berechnung des $\chi^2$ -Werts

$$\chi^2 = \sum \frac{(\text{beobachtet} - \text{erwartet})^2}{\text{erwartet}}$$

beobachtet:

	haben	sein
nord	27	33
sued	3	34

erwartet:

	haben	sein
nord	18.56	41.44
sued	11.44	25.56

$$\chi^2 = \frac{(27-18.56)^2}{18.56} + \frac{(33-41.44)^2}{41.44} + \frac{(3-11.44)^2}{11.44}$$



# Berechnung des $\chi^2$ -Werts

$$\chi^2 = \sum \frac{(\text{beobachtet} - \text{erwartet})^2}{\text{erwartet}}$$

beobachtet:

	haben	sein
nord	27	33
sued	3	34

erwartet:

	haben	sein
nord	18.56	41.44
sued	11.44	25.56

$$\chi^2 = \frac{(27-18.56)^2}{18.56} + \frac{(33-41.44)^2}{41.44} + \frac{(3-11.44)^2}{11.44} + \frac{(34-25.56)^2}{25.56}$$

# Berechnung des $\chi^2$ -Werts

$$\chi^2 = \sum \frac{(\text{beobachtet} - \text{erwartet})^2}{\text{erwartet}}$$

beobachtet:

	haben	sein
nord	27	33
sued	3	34

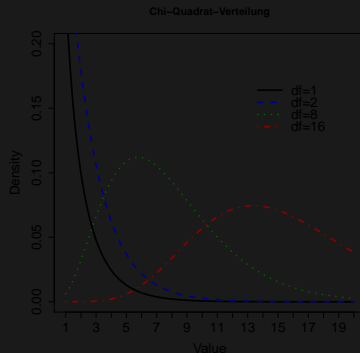
erwartet:

	haben	sein
nord	18.56	41.44
sued	11.44	25.56

$$\begin{aligned}\chi^2 &= \frac{(27-18.56)^2}{18.56} + \frac{(33-41.44)^2}{41.44} + \frac{(3-11.44)^2}{11.44} + \frac{(34-25.56)^2}{25.56} \\ \chi^2 &= 3.84 + 1.72 + 6.23 + 2.79 = 14.58\end{aligned}$$

# Die $\chi^2$ -Verteilung

Die  $\chi^2$ -Verteilung für Stichproben  
aus Grundgesamtheiten ohne Zusammenhang:



Was sind „Freiheitsgrade“ oder *degrees of freedom (df)*?

- Das kommt später noch ausführlicher.

Was sind „Freiheitsgrade“ oder *degrees of freedom (df)*?

- Das kommt später noch ausführlicher.
- Für n-Felder-Tests:  $(\text{Zeilenzahl} - 1) \cdot (\text{Spaltenzahl} - 1)$

Was sind „Freiheitsgrade“ oder *degrees of freedom* (*df*)?

- Das kommt später noch ausführlicher.
- Für n-Felder-Tests:  $(\text{Zeilenzahl} - 1) \cdot (\text{Spaltenzahl} - 1)$
- Bei Vierfelder-Test also:  $df = 1$

- Wahrscheinlichkeit eines bestimmten  $\chi^2$ -Werts unter Annahme der  $H_0$ ?  
VOR dem Experiment! Nach dem Experiment ist die Wahrscheinlichkeit des gemessenen p-Werts immer 1.

- Wahrscheinlichkeit eines bestimmten  $\chi^2$ -Werts unter Annahme der  $H_0$ ?  
VOR dem Experiment! Nach dem Experiment ist die Wahrscheinlichkeit des gemessenen p-Werts immer 1.
- In Fishers Philosophie Entscheidung nach Signifikanzniveau (*sig*):  
Der  $\chi^2$ -Wert muss in den extremen *sig*-Anteilen liegen, um die  $H_0$  zu *sig* zurückzuweisen.



- Wahrscheinlichkeit eines bestimmten  $\chi^2$ -Werts unter Annahme der  $H_0$ ?  
VOR dem Experiment! Nach dem Experiment ist die Wahrscheinlichkeit des gemessenen p-Werts immer 1.
- In Fishers Philosophie Entscheidung nach Signifikanzniveau (*sig*):  
Der  $\chi^2$ -Wert muss in den extremen *sig*-Anteilen liegen,  
um die  $H_0$  zu *sig* zurückzuweisen.

- Wahrscheinlichkeit eines bestimmten  $\chi^2$ -Werts unter Annahme der  $H_0$ ?  
VOR dem Experiment! Nach dem Experiment ist die Wahrscheinlichkeit des gemessenen p-Werts immer 1.
- In Fishers Philosophie Entscheidung nach Signifikanzniveau (*sig*):  
Der  $\chi^2$ -Wert muss in den extremen *sig*-Anteilen liegen, um die  $H_0$  zu *sig* zurückzuweisen.

In R ähnlich wie bei Normalverteilung:

```
> qchisq(0.95, df=1)  $\Rightarrow$  3.84
```

- Wahrscheinlichkeit eines bestimmten  $\chi^2$ -Werts unter Annahme der  $H_0$ ?  
VOR dem Experiment! Nach dem Experiment ist die Wahrscheinlichkeit des gemessenen p-Werts immer 1.
- In Fishers Philosophie Entscheidung nach Signifikanzniveau (*sig*):  
Der  $\chi^2$ -Wert muss in den extremen *sig*-Anteilen liegen, um die  $H_0$  zu *sig* zurückzuweisen.

In R ähnlich wie bei Normalverteilung:

`> qchisq(0.95, df=1)  $\Rightarrow$  3.84`

- Also ist für  $\chi^2 = 14.58$  auf jeden Fall  $p < 0.05$  (weil  $14.58 > 3.84$ ).

# Mehr oder weniger signifikant?

- Oft liest man etwas von „ $\alpha$ -Niveaus“ wie:

# Mehr oder weniger signifikant?

- Oft liest man etwas von „ $\alpha$ -Niveaus“ wie:
  - ▶ 5% („signifikant“)

# Mehr oder weniger signifikant?

- Oft liest man etwas von „ $\alpha$ -Niveaus“ wie:
  - ▶ 5% („signifikant“)
  - ▶ 1%

# Mehr oder weniger signifikant?

- Oft liest man etwas von „ $\alpha$ -Niveaus“ wie:
  - ▶ 5% („signifikant“)
  - ▶ 1%
  - ▶ 0.1% („hochsignifikant“)

# Mehr oder weniger signifikant?

- Oft liest man etwas von „ $\alpha$ -Niveaus“ wie:
  - ▶ 5% („signifikant“)
  - ▶ 1%
  - ▶ 0.1% („hochsignifikant“)
- Diese Niveaus entsprechen einem falsch interpretierten *sig.*



# Mehr oder weniger signifikant?

- Oft liest man etwas von „ $\alpha$ -Niveaus“ wie:
  - ▶ 5% („signifikant“)
  - ▶ 1%
  - ▶ 0.1% („hochsignifikant“)
- Diese Niveaus entsprechen einem falsch interpretierten *sig.*
- Die Idee von „mehr oder weniger signifikant“ ist kompletter Schwachsinn.

# Mehr oder weniger signifikant?

- Oft liest man etwas von „ $\alpha$ -Niveaus“ wie:
  - ▶ 5% („signifikant“)
  - ▶ 1%
  - ▶ 0.1% („hochsignifikant“)
- Diese Niveaus entsprechen einem falsch interpretierten *sig.*
- Die Idee von „mehr oder weniger signifikant“ ist kompletter Schwachsinn.
- Entweder ist das gesetzte Niveau akzeptabel, und dann bringt ein kleineres  $p$  aber auch nicht mehr.

# Mehr oder weniger signifikant?

- Oft liest man etwas von „ $\alpha$ -Niveaus“ wie:
  - ▶ 5% („signifikant“)
  - ▶ 1%
  - ▶ 0.1% („hochsignifikant“)
- Diese Niveaus entsprechen einem falsch interpretierten *sig.*
- Die Idee von „mehr oder weniger signifikant“ ist kompletter Schwachsinn.
- Entweder ist das gesetzte Niveau akzeptabel, und dann bringt ein kleineres  $p$  aber auch nicht mehr.
- Oder es müsste eigtl. ein strengeres *sig*-Niveau gewählt werden, und dann ist  $p < 0.05$  schlicht nicht ausreichend (s. Fishers Sensitivität).

# Mehr oder weniger signifikant?

- Oft liest man etwas von „ $\alpha$ -Niveaus“ wie:
  - ▶ 5% („signifikant“)
  - ▶ 1%
  - ▶ 0.1% („hochsignifikant“)
- Diese Niveaus entsprechen einem falsch interpretierten *sig.*
- Die Idee von „mehr oder weniger signifikant“ ist kompletter Schwachsinn.
- Entweder ist das gesetzte Niveau akzeptabel, und dann bringt ein kleineres  $p$  aber auch nicht mehr.
- Oder es müsste eigtl. ein strengeres *sig*-Niveau gewählt werden, und dann ist  $p < 0.05$  schlicht nicht ausreichend (s. Fishers Sensitivität).
- Die Entscheidung für ein bestimmtes *sig*-Niveau muss auf Basis konzeptueller/inhaltlicher Gründe gefällt werden.

# Mehr oder weniger signifikant?

- Oft liest man etwas von „ $\alpha$ -Niveau“ wie:
  - ▶ 5% („signifikant“)
  - ▶ 1%
  - ▶ 0.1% („hochsignifikant“)
- Diese Niveaus entsprechen einem falsch interpretierten *sig.*
- Die Idee von „mehr oder weniger signifikant“ ist kompletter Schwachsinn.
- Entweder ist das gesetzte Niveau akzeptabel, und dann bringt ein kleineres  $p$  aber auch nicht mehr.
- Oder es müsste eigtl. ein strengeres *sig*-Niveau gewählt werden, und dann ist  $p < 0.05$  schlicht nicht ausreichend (s. Fishers Sensitivität).
- Die Entscheidung für ein bestimmtes *sig*-Niveau muss auf Basis konzeptueller/inhaltlicher Gründe gefällt werden.
- EIN signifikantes Testergebnis alleine sagt nicht viel aus!!!

- 1 Die Beobachtungen sind voneinander unabhängig.

# Voraussetzungen für $\chi^2$ -Tests

- 1 Die Beobachtungen sind voneinander unabhängig.
- 2 In jeder Zelle ist die erwartete Häufigkeit mindestens 5.

# Voraussetzungen für $\chi^2$ -Tests

- 1 Die Beobachtungen sind voneinander unabhängig.
- 2 In jeder Zelle ist die erwartete Häufigkeit mindestens 5.
- 3 Keine Beschränkung auf vier Felder!



Mit einer Matrix `my.matrix`:

```
> chisq.test(my.matrix)
```

Eingabe einer einfachen Vierfeldermatrix:

```
> my.matrix <- matrix(c(27,33,3,34), 2, 2, byrow=TRUE)
```

Ausgeben der erwarteten Häufigkeiten:

```
> chisq.test(my.matrix)$expected
```

# Wann und wie Fisher-Exakt?

Der Fisher-Exakt-Test ist eine Alternative zum  $\chi^2$ -Test.

- exakter Test: direkte Berechnung der Wahrscheinlichkeit

# Wann und wie Fisher-Exakt?

Der Fisher-Exakt-Test ist eine Alternative zum  $\chi^2$ -Test.

- exakter Test: direkte Berechnung der Wahrscheinlichkeit
- keine allgemein bessere Alternative zu  $\chi^2$

# Wann und wie Fisher-Exakt?

Der Fisher-Exakt-Test ist eine Alternative zum  $\chi^2$ -Test.

- exakter Test: direkte Berechnung der Wahrscheinlichkeit
- keine allgemein bessere Alternative zu  $\chi^2$
- robuster bei sehr kleinen Stichproben

# Wann und wie Fisher-Exakt?

Der Fisher-Exakt-Test ist eine Alternative zum  $\chi^2$ -Test.

- exakter Test: direkte Berechnung der Wahrscheinlichkeit
- keine allgemein bessere Alternative zu  $\chi^2$
- robuster bei sehr kleinen Stichproben
- aber nur für feststehende Randsummen geeignet!

Fisher-Exakt in R:

```
> fisher.test(my.matrix)
> fisher.test(my.vector.1, my.vector.2)
```

# Wann und wie Fisher-Exakt?

Der Fisher-Exakt-Test ist eine Alternative zum  $\chi^2$ -Test.

- exakter Test: direkte Berechnung der Wahrscheinlichkeit
- keine allgemein bessere Alternative zu  $\chi^2$
- robuster bei sehr kleinen Stichproben
- aber nur für feststehende Randsummen geeignet!
- ohne feste Randsummen: Barnards Test

Fisher-Exakt in R:

```
> fisher.test(my.matrix)
> fisher.test(my.vector.1, my.vector.2)
```

Der  $\chi^2$ -Wert sagt nichts über die Stärke eines Zusammenhangs!

	haben	sein
nord	27	33
sued	3	34

 $\chi^2 =$ 

	haben	sein
nord	54	66
sued	6	68

 $\chi^2 =$

Der  $\chi^2$ -Wert sagt nichts über die Stärke eines Zusammenhangs!  
Bei höheren absoluten Frequenzen wird auch der  $\chi^2$ -Wert größer.

	haben	sein
nord	27	33
sued	3	34

 $\chi^2 =$ 

	haben	sein
nord	54	66
sued	6	68

 $\chi^2 =$



Der  $\chi^2$ -Wert sagt nichts über die Stärke eines Zusammenhangs!  
Bei höheren absoluten Frequenzen wird auch der  $\chi^2$ -Wert größer.

	haben	sein
nord	27	33
sued	3	34

 $\chi^2 =$ 

	haben	sein
nord	54	66
sued	6	68

 $\chi^2 =$

Der  $\chi^2$ -Wert sagt nichts über die Stärke eines Zusammenhangs!  
Bei höheren absoluten Frequenzen wird auch der  $\chi^2$ -Wert größer.

	haben	sein
nord	27	33
sued	3	34

 $\chi^2 = 12,89$ 

	haben	sein
nord	54	66
sued	6	68

 $\chi^2 =$

Der  $\chi^2$ -Wert sagt nichts über die Stärke eines Zusammenhangs!  
Bei höheren absoluten Frequenzen wird auch der  $\chi^2$ -Wert größer.

	haben	sein
nord	27	33
sued	3	34

 $\chi^2 = 12,89$ 

	haben	sein
nord	54	66
sued	6	68

 $\chi^2 = 27,46$

Der  $\chi^2$ -Wert sagt nichts über die Stärke eines Zusammenhangs!  
Bei höheren absoluten Frequenzen wird auch der  $\chi^2$ -Wert größer.

	haben	sein
nord	27	33
sued	3	34

$$\chi^2 = 12,89$$

	haben	sein
nord	27.84%	34.02%
sued	3.09%	35.05%

	haben	sein
nord	54	66
sued	6	68

$$\chi^2 = 27,46$$

Der  $\chi^2$ -Wert sagt nichts über die Stärke eines Zusammenhangs!  
Bei höheren absoluten Frequenzen wird auch der  $\chi^2$ -Wert größer.

	haben	sein
nord	27	33
sued	3	34

$$\chi^2 = 12,89$$

	haben	sein
nord	27.84%	34.02%
sued	3.09%	35.05%

	haben	sein
nord	54	66
sued	6	68

$$\chi^2 = 27,46$$

	haben	sein
nord	27.84%	34.02%
sued	3.09%	35.05%

Pearsons  $\phi$ : Maß für die Stärke des Zusammenhangs in  $2 \times 2$ -Tabellen

$$\phi = \sqrt{\frac{\chi^2}{n}}$$

Pearsons  $\phi$ : Maß für die Stärke des Zusammenhangs in  $2 \times 2$ -Tabellen

$$\phi = \sqrt{\frac{\chi^2}{n}}$$

$\phi$  ist eine Zahl zwischen 0 und 1:

Je größer, desto stärker der Zusammenhang zwischen den Variablen.

Pearsons  $\phi$ : Maß für die Stärke des Zusammenhangs in  $2 \times 2$ -Tabellen

$$\phi = \sqrt{\frac{\chi^2}{n}}$$

$\phi$  ist eine Zahl zwischen 0 und 1:

Je größer, desto stärker der Zusammenhang zwischen den Variablen.

$$\text{Beispiel: } \phi = \sqrt{\frac{\chi^2}{n}} = \sqrt{\frac{12.89}{97}} = 0.3648$$



Cramér's  $v$  für  $n \times n$ -Tabellen mit  $n > 2$  oder  $m > 2$

$$v = \sqrt{\frac{\frac{\chi^2}{n}}{\min(s-1, z-1)}}$$

mit:  $s$  die Spaltenzahl und  $z$  die Zeilenzahl

Beachte: für  $2 \times 2$ -Tabellen:  $s - 1 = 1$  und  $z - 1 = 1$ ,

also  $\min(s - 1, z - 1) = 1$

$$\text{daher: } v = \sqrt{\frac{\frac{\chi^2}{n}}{1}} = \sqrt{\frac{\chi^2}{n}} = \phi$$

Speichern des Test-Objekts:

```
> my.chi2.test <- chisq.test(my.matrix)
```

Speichern des  $\chi^2$ -Werts mit:

```
> my.chi2.value <- as.numeric(my.chi2.test$statistic)
```

Speichern von  $n$ :

```
> my.n <- sum(my.matrix)
```

Also Effektstärke (mit Ausgabe):

```
> my.phi <- sqrt( my.chi2.value / my.n ); my.phi
```

- Die Chance (odds)  $o$  setzt die Wahrscheinlichkeit  $p$  eines Ereignisses  $E$  in Relation zur Gegenwahrscheinlichkeit:

- Die Chance (odds)  $o$  setzt die Wahrscheinlichkeit  $p$  eines Ereignisses  $E$  in Relation zur Gegenwahrscheinlichkeit:

- Die Chance (odds)  $o$  setzt die Wahrscheinlichkeit  $p$  eines Ereignisses  $E$  in Relation zur Gegenwahrscheinlichkeit:

$$o(E) = \frac{p(E)}{1-p(E)}$$

- Die Chance (odds)  $o$  setzt die Wahrscheinlichkeit  $p$  eines Ereignisses  $E$  in Relation zur Gegenwahrscheinlichkeit:

$$o(E) = \frac{p(E)}{1-p(E)}$$

und damit

- Die Chance (odds)  $o$  setzt die Wahrscheinlichkeit  $p$  eines Ereignisses  $E$  in Relation zur Gegenwahrscheinlichkeit:

$$o(E) = \frac{p(E)}{1-p(E)}$$

und damit

$$p(E) = \frac{o(E)}{1+o(E)}$$

- Die Chance (odds)  $o$  setzt die Wahrscheinlichkeit  $p$  eines Ereignisses  $E$  in Relation zur Gegenwahrscheinlichkeit:

$$o(E) = \frac{p(E)}{1-p(E)}$$

und damit

$$p(E) = \frac{o(E)}{1+o(E)}$$

- Ein Ereignis ist in Korpusstudien i. d. R. das Auftreten einer Variablenausprägung.



# Chance (odds)

- Die Chance (odds)  $o$  setzt die Wahrscheinlichkeit  $p$  eines Ereignisses  $E$  in Relation zur Gegenwahrscheinlichkeit:

$$o(E) = \frac{p(E)}{1-p(E)}$$

und damit

$$p(E) = \frac{o(E)}{1+o(E)}$$

- Ein Ereignis ist in Korpusstudien i. d. R. das Auftreten einer Variablenausprägung.
- Die Information in den Maßen Wahrscheinlichkeit und Chance ist dieselbe (s. Umrechenbarkeit ineinander).

Aux	Anzahl
haben	27
sein	33

Aux	Anzahl
haben	27
sein	33

$$p(\textit{haben}) = \frac{27}{27+33} = \frac{27}{60} = 0.45 \text{ (Wahrscheinlichkeit)}$$

Aux	Anzahl
haben	27
sein	33

$$p(\textit{haben}) = \frac{27}{27+33} = \frac{27}{60} = 0.45 \text{ (Wahrscheinlichkeit)}$$

$$1 - p(\textit{haben}) = p(\neg \textit{haben}) = \frac{33}{27+33} = \frac{33}{60} = 0.55 \text{ (Gegenwahrscheinlichkeit)}$$

Aux	Anzahl
haben	27
sein	33

$$p(\text{haben}) = \frac{27}{27+33} = \frac{27}{60} = 0.45 \text{ (Wahrscheinlichkeit)}$$

$$1 - p(\text{haben}) = p(\neg\text{haben}) = \frac{33}{27+33} = \frac{33}{60} = 0.55 \text{ (Gegenwahrscheinlichkeit)}$$

$$\text{Beachte: } p(\text{haben}) + p(\neg\text{haben}) = 1$$

Aux	Anzahl
haben	27
sein	33

$$p(\text{haben}) = \frac{27}{27+33} = \frac{27}{60} = 0.45 \text{ (Wahrscheinlichkeit)}$$

$$1 - p(\text{haben}) = p(\neg\text{haben}) = \frac{33}{27+33} = \frac{33}{60} = 0.55 \text{ (Gegenwahrscheinlichkeit)}$$

$$\text{Beachte: } p(\text{haben}) + p(\neg\text{haben}) = 1$$

$$o(\text{haben}) = \frac{\frac{27}{60}}{\frac{33}{60}} = \frac{27}{60} \cdot \frac{60}{33} = \frac{27}{33} = 0.82$$

Aux	Anzahl
haben	27
sein	33

$$p(\text{haben}) = \frac{27}{27+33} = \frac{27}{60} = 0.45 \text{ (Wahrscheinlichkeit)}$$

$$1 - p(\text{haben}) = p(\neg\text{haben}) = \frac{33}{27+33} = \frac{33}{60} = 0.55 \text{ (Gegenwahrscheinlichkeit)}$$

$$\text{Beachte: } p(\text{haben}) + p(\neg\text{haben}) = 1$$

$$o(\text{haben}) = \frac{\frac{27}{60}}{\frac{33}{60}} = \frac{27}{60} \cdot \frac{60}{33} = \frac{27}{33} = 0.82$$

$$\text{allgemein: } p(E) = \frac{\text{Anzahl}(E)}{\text{Anzahl}(E) + \text{Anzahl}(\neg E)} \text{ und } o(E) = \frac{\text{Anzahl}(E)}{\text{Anzahl}(\neg E)}$$

- Das Chancenverhältnis (odds ratio) gibt das Verhältnis an, wie sich die Chancen einer Variablenausprägung  $E$  unter Bedingung  $A$  – also  $o(E|A)$  – und unter Bedingung  $B$  – also  $o(E|B)$  – zueinander Verhalten:

$$r(E|A, E|B) = \frac{o(E|A)}{o(E|B)}$$



# Beispiel zum Chancenverhältnis (1)

- Wir haben Texte aus Süddeutschland und Norddeutschland auf das Auftreten des Perfektauxiliars *haben* und *sein* bei bestimmten Verben untersucht.
- Die Kreuztabelle:

	nord	sued
haben	27	3
sein	33	34

## Beispiel zum Chancenverhältnis (2)

	nord	sued
haben	27	3
sein	33	34

- $o(haben|nord) =$
- $o(haben|sued) =$

## Beispiel zum Chancenverhältnis (2)

	nord	sued
haben	27	3
sein	33	34

- $o(haben|nord) = \frac{27}{33} = 0.82$
- $o(haben|sued) =$

## Beispiel zum Chancenverhältnis (2)

	nord	sued
haben	27	3
sein	33	34

- $o(\text{haben}|\text{nord}) = \frac{27}{33} = 0.82$
- $o(\text{haben}|\text{sued}) = \frac{3}{34} = 0.09$

## Beispiel zum Chancenverhältnis (2)

	nord	sued
haben	27	3
sein	33	34

- $o(haben|nord) = \frac{27}{33} = 0.82$
- $o(haben|sued) = \frac{3}{34} = 0.09$
- Verhältnis zwischen den Chancen:

## Beispiel zum Chancenverhältnis (2)

	nord	sued
haben	27	3
sein	33	34

- $o(haben|nord) = \frac{27}{33} = 0.82$
- $o(haben|sued) = \frac{3}{34} = 0.09$
- Verhältnis zwischen den Chancen:  $or = \frac{0.82}{0.09} = 9.11$

## Beispiel zum Chancenverhältnis (2)

	nord	sued
haben	27	3
sein	33	34

- $o(haben|nord) = \frac{27}{33} = 0.82$
- $o(haben|sued) = \frac{3}{34} = 0.09$
- Verhältnis zwischen den Chancen:  $or = \frac{0.82}{0.09} = 9.11$
- D. h. die Chance von *haben* ist 9.11 mal größer, wenn die Region *nord* ist.

## Beispiel zum Chancenverhältnis (2)

	nord	sued
haben	27	3
sein	33	34

- $o(haben|nord) = \frac{27}{33} = 0.82$
- $o(haben|sued) = \frac{3}{34} = 0.09$
- Verhältnis zwischen den Chancen:  $or = \frac{0.82}{0.09} = 9.11$
- D. h. die Chance von *haben* ist 9.11 mal größer, wenn die Region *nord* ist.
- Ersatz für Effektstärke bei Fisher-Test



- binäre Daten: Ereignis vs. Nicht-Ereignis bzw. Ja/Nein

- binäre Daten: Ereignis vs. Nicht-Ereignis bzw. Ja/Nein
- Vgl. Behauptung: „Gen/Dat alternieren frei bei *wegen*.“

- binäre Daten: Ereignis vs. Nicht-Ereignis bzw. Ja/Nein
- Vgl. Behauptung: „Gen/Dat alternieren frei bei *wegen*.“
  - ▶ „frei alternieren“ = beide Kasus haben den gleichen Anteil.

- binäre Daten: Ereignis vs. Nicht-Ereignis bzw. Ja/Nein
- Vgl. Behauptung: „Gen/Dat alternieren frei bei *wegen*.“
  - ▶ „frei alternieren“ = beide Kasus haben den gleichen Anteil.
  - ▶ Grundgesamtheit per Null-Hypothese: 50% Genitive und 50% Dative

- binäre Daten: Ereignis vs. Nicht-Ereignis bzw. Ja/Nein
- Vgl. Behauptung: „Gen/Dat alternieren frei bei *wegen*.“
  - ▶ „frei alternieren“ = beide Kasus haben den gleichen Anteil.
  - ▶ Grundgesamtheit per Null-Hypothese: 50% Genitive und 50% Dative
- Korpusstichprobe:  $F(\text{Genitiv})=41$  und  $F(\text{Dativ})=59$

- binäre Daten: Ereignis vs. Nicht-Ereignis bzw. Ja/Nein
- Vgl. Behauptung: „Gen/Dat alternieren frei bei *wegen*.“
  - ▶ „frei alternieren“ = beide Kasus haben den gleichen Anteil.
  - ▶ Grundgesamtheit per Null-Hypothese: 50% Genitive und 50% Dative
- Korpusstichprobe:  $F(\text{Genitiv})=41$  und  $F(\text{Dativ})=59$
- Stimmt das mit der Null überein bei  $\text{sig} = 0.05$ ?



$H_0$ : Es gibt keine Abweichung  
von den erwarteten gleich großen Anteilen.



Ho: Es gibt keine Abweichung  
von den erwarteten gleich großen Anteilen.

Ho:  $p(\text{Dativ}) = 0.5$  (p für proportion)

Benötigte Größen:

- Stichproben der Größe  $n$

Benötigte Größen:

- Stichproben der Größe  $n$
- Proportion  $p$  (hier  $p = 0.5$ )

Benötigte Größen:

- Stichproben der Größe  $n$
- Proportion  $p$  (hier  $p = 0.5$ )
- Anzahl der beobachteten Ereignisse:  $X$  (hier  $X(\text{Dativ}) = 59$ )

- Wenn  $p \cdot n > 10$  und  $(1 - p) \cdot n > 10$   
approximiert die Binomialverteilung die Normalverteilung.

- Wenn  $p \cdot n > 10$  und  $(1 - p) \cdot n > 10$   
approximiert die Binomialverteilung die Normalverteilung.
- Es gilt dann (unter Annahme der  $H_0$ !) für die Normalverteilung:

- Wenn  $p \cdot n > 10$  und  $(1 - p) \cdot n > 10$   
approximiert die Binomialverteilung die Normalverteilung.
- Es gilt dann (unter Annahme der Ho!) für die Normalverteilung:
  - ▶ Mittel:  $\mu = p \cdot n$

- Wenn  $p \cdot n > 10$  und  $(1 - p) \cdot n > 10$   
approximiert die Binomialverteilung die Normalverteilung.
- Es gilt dann (unter Annahme der Ho!) für die Normalverteilung:
  - ▶ Mittel:  $\mu = p \cdot n$
  - ▶ Standardabweichung:  $s = \sqrt{n \cdot p \cdot (1 - p)}$



- Wenn  $p \cdot n > 10$  und  $(1 - p) \cdot n > 10$   
approximiert die Binomialverteilung die Normalverteilung.
- Es gilt dann (unter Annahme der Ho!) für die Normalverteilung:
  - ▶ Mittel:  $\mu = p \cdot n$
  - ▶ Standardabweichung:  $s = \sqrt{n \cdot p \cdot (1 - p)}$
  - ▶ Wir können für den gemessenen Wert den z-Wert ausrechnen.

- Wenn  $p \cdot n > 10$  und  $(1 - p) \cdot n > 10$   
approximiert die Binomialverteilung die Normalverteilung.
- Es gilt dann (unter Annahme der Ho!) für die Normalverteilung:
  - ▶ Mittel:  $\mu = p \cdot n$
  - ▶ Standardabweichung:  $s = \sqrt{n \cdot p \cdot (1 - p)}$
  - ▶ Wir können für den gemessenen Wert den z-Wert ausrechnen.

- Wenn  $p \cdot n > 10$  und  $(1 - p) \cdot n > 10$  approximiert die Binomialverteilung die Normalverteilung.
- Es gilt dann (unter Annahme der Ho!) für die Normalverteilung:
  - ▶ Mittel:  $\mu = p \cdot n$
  - ▶ Standardabweichung:  $s = \sqrt{n \cdot p \cdot (1 - p)}$
  - ▶ Wir können für den gemessenen Wert den z-Wert ausrechnen.

$$z = \frac{x - \mu}{s} = \frac{x - p \cdot n}{\sqrt{n \cdot p \cdot (1 - p)}}$$

$$z = \frac{59 - (0.5 \cdot 100)}{\sqrt{100 \cdot 0.5 \cdot 0.5}} = \frac{59 - 50}{\sqrt{25}} = \frac{9}{5} = 1.8$$

$$z = \frac{59 - (0.5 \cdot 100)}{\sqrt{100 \cdot 0.5 \cdot 0.5}} = \frac{59 - 50}{\sqrt{25}} = \frac{9}{5} = 1.8$$

- Der gemessene Wert liegt 1.8 Standardabweichungen vom Ho-Mittel entfernt.

$$z = \frac{59 - (0.5 \cdot 100)}{\sqrt{100 \cdot 0.5 \cdot 0.5}} = \frac{59 - 50}{\sqrt{25}} = \frac{9}{5} = 1.8$$

- Der gemessene Wert liegt 1.8 Standardabweichungen vom  $H_0$ -Mittel entfernt.
- Wir kennen bereits die kritischen Werte für Normalverteilungen und  $\text{sig} = 0.05$ :  $-1.96..1.96$

$$z = \frac{59 - (0.5 \cdot 100)}{\sqrt{100 \cdot 0.5 \cdot 0.5}} = \frac{59 - 50}{\sqrt{25}} = \frac{9}{5} = 1.8$$

- Der gemessene Wert liegt 1.8 Standardabweichungen vom Ho-Mittel entfernt.
- Wir kennen bereits die kritischen Werte für Normalverteilungen und  $\text{sig} = 0.05$ :  $-1.96..1.96$
- Die Ho kann also nicht zurückgewiesen werden bei  $\text{sig} = 0.05$ .

$$z = \frac{59 - (0.5 \cdot 100)}{\sqrt{100 \cdot 0.5 \cdot 0.5}} = \frac{59 - 50}{\sqrt{25}} = \frac{9}{5} = 1.8$$

- Der gemessene Wert liegt 1.8 Standardabweichungen vom Ho-Mittel entfernt.
- Wir kennen bereits die kritischen Werte für Normalverteilungen und  $\text{sig} = 0.05$ :  $-1.96..1.96$
- Die Ho kann also nicht zurückgewiesen werden bei  $\text{sig} = 0.05$ .
- Interpretation: Entweder ist die Variation nicht genau gleich verteilt oder ein seltenes Ereignis ist eingetreten.



```
> binom.test(59, 100, 0.5)
```

```
Exact binomial test
```

```
data: 59 and 100
```

```
number of successes = 59, number of trials = 100, p-value = 0.08863
```

```
alternative hypothesis: true probability of success is not equal to 0.5
```

```
95 percent confidence interval:
```

```
0.4871442 0.6873800 sample estimates:
```

```
probability of success 0.59
```

Nächste Woche | Überblick

- 1 Statistik, Inferenz und probabilistische Grammatik
- 2 Deskriptive Statistik
- 3 Nichtparametrische Verfahren
- 4 z-Test und t-Test
- 5 ANOVA
- 6 Freiheitsgrade und Effektstärken
- 7 Power
- 8 Lineare Modelle
- 9 Generalisierte Lineare Modelle
- 10 Gemischte Modelle

Bortz, Jürgen & Gustav Lienert. 2008. *Kurzgefasste Statistik für die klinische Forschung*. Heidelberg: Springer.

Gravetter, Frederick J. & Larry B. Wallnau. 2007. *Statistics for the Behavioral Sciences*. 7. Aufl. Belmont: Thomson.

## Kontakt

Prof. Dr. Roland Schäfer  
Institut für Germanistische Sprachwissenschaft  
Friedrich-Schiller-Universität Jena  
Fürstengraben 30  
07743 Jena

<https://rolandschaefer.net>  
[roland.schaefer@uni-jena.de](mailto:roland.schaefer@uni-jena.de)

## Creative Commons BY-SA-3.0-DE

Dieses Werk ist unter einer Creative Commons Lizenz vom Typ *Namensnennung - Weitergabe unter gleichen Bedingungen 3.0 Deutschland* zugänglich. Um eine Kopie dieser Lizenz einzusehen, konsultieren Sie

<http://creativecommons.org/licenses/by-sa/3.0/de/> oder wenden Sie sich brieflich an Creative Commons, Postfach 1866, Mountain View, California, 94042, USA.