Christian-Albrechts-Universität zu Kiel

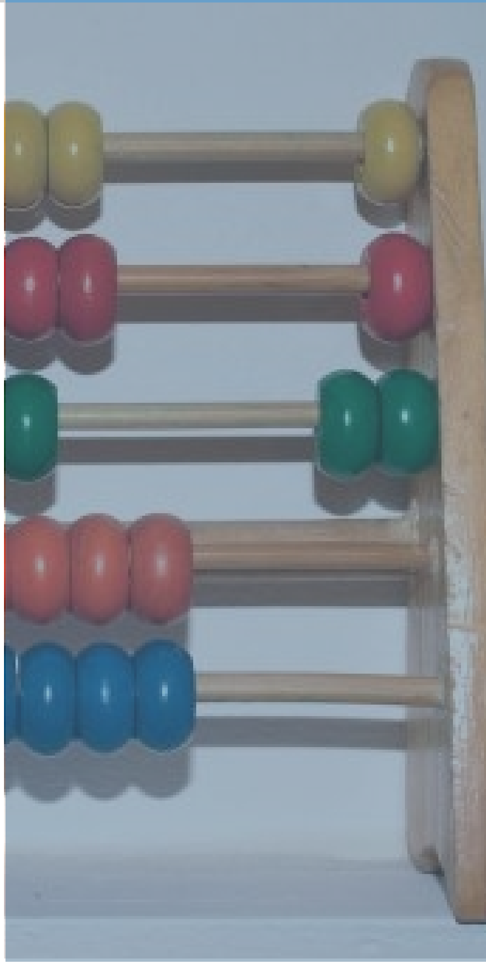
06_chi-quadrat/zusammenhangsmaße

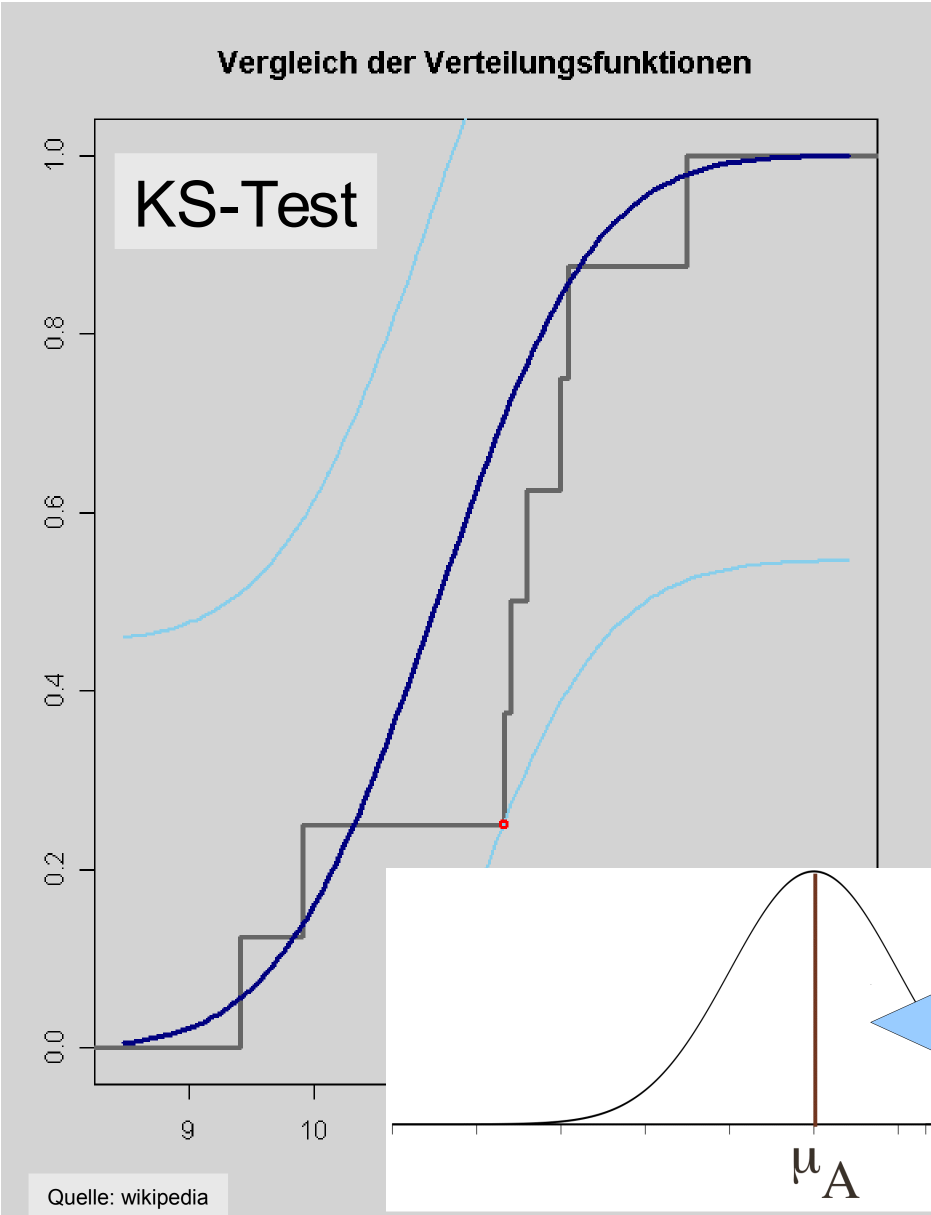
Chi-quadrat, Cramers V

Nachtrag KS-Test, Mann-Whitney-U-Test [1]

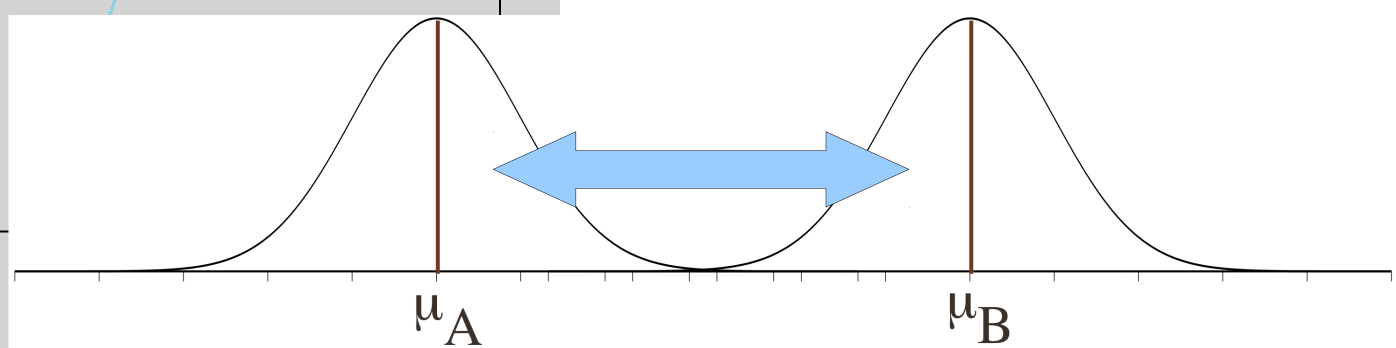
Teststärke
MWU-Test ist generell teststärker.

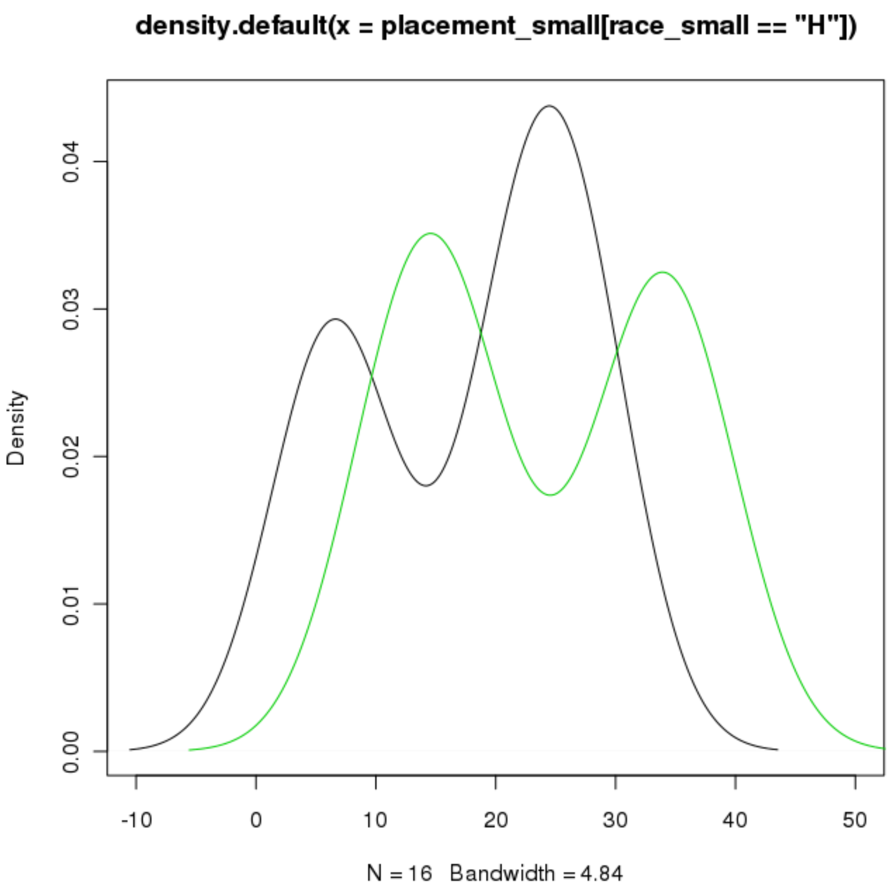
Methode
KS-Test testet auf Form und Lage der Verteilungen,
MWU-Test testet nur auf Lage der Verteilung

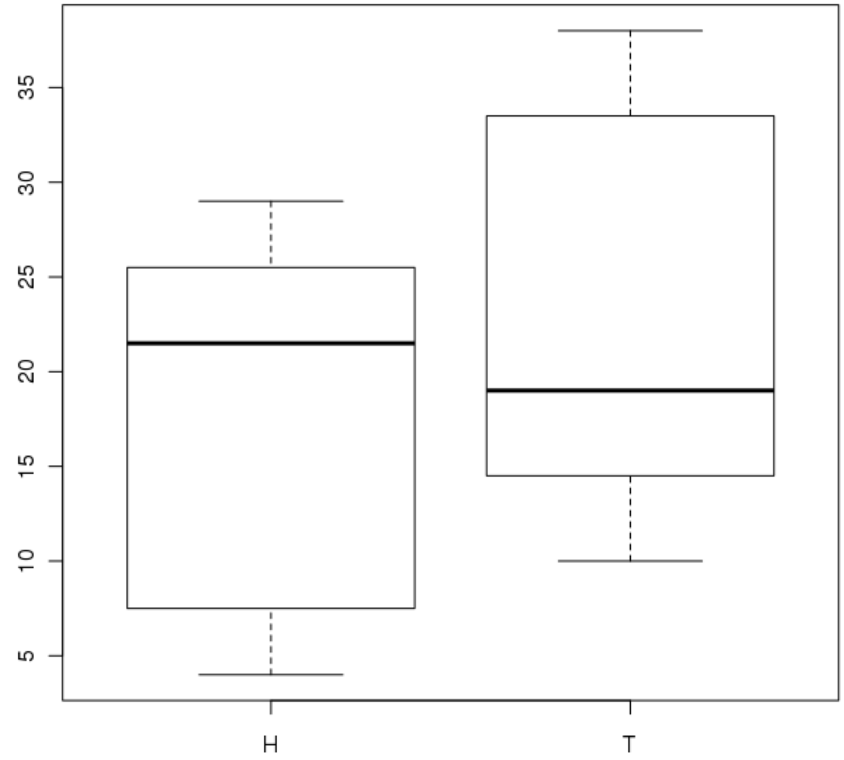




Mann-Withney-U-Test








```
> ks.test(placement_small[race_small=="T"],placement_small[race_small=="H"][-1])

        Two-sample Kolmogorov-Smirnov test

data:  placement_small[race_small == "T"] and placement_small[race_small == "H"]
[-1]
D = 0.4737, p-value = 0.02878
alternative hypothesis: two-sided

> wilcox.test(placement_small[race_small=="T"],placement_small[race_small=="H"][
-1])

        Wilcoxon rank sum test

data:  placement_small[race_small == "T"] and placement_small[race_small == "H"]
[-1]
W = 185, p-value = 0.1469
alternative hypothesis: true location shift is not equal to 0

>
```

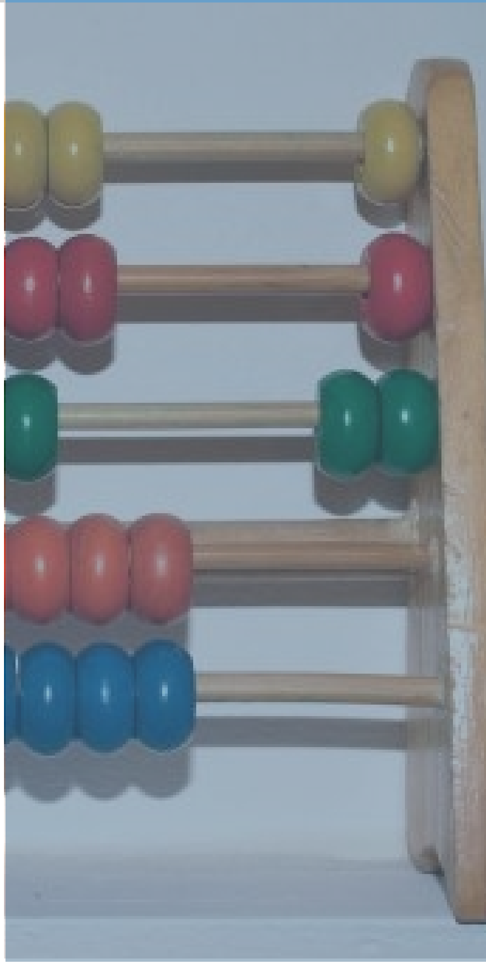
H H H H H H H H H T T T T T T T T T T H H H H H H H H H H T T T T T T T T T

Nachtrag KS-Test, Mann-Whitney-U-Test [2]

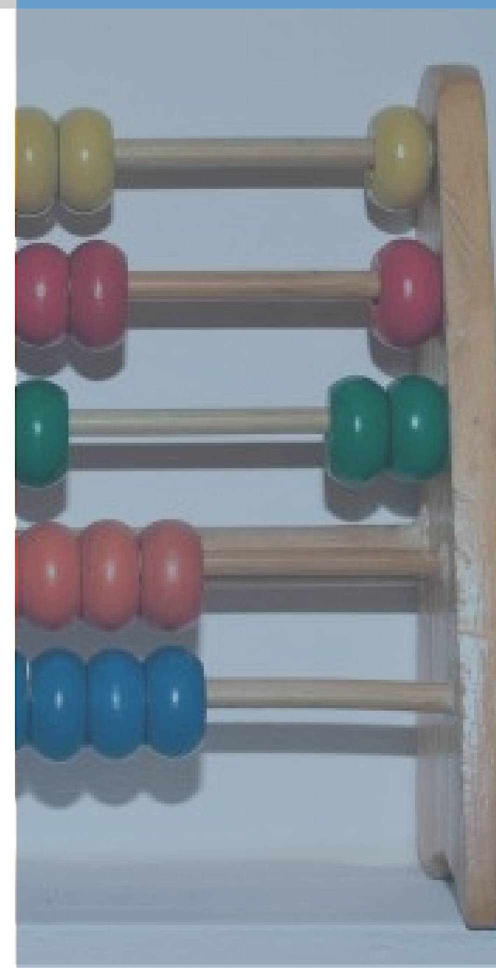
Teststärke
MWU-Test ist generell teststärker.

Methode
KS-Test testet auf Form und Lage der Verteilungen,
MWU-Test testet nur auf Lage der Verteilung

Voraussetzungen
MWU-Test: Form der Verteilungen ähnlich



Hypothesen-Test



Überprüfung von Annahmen über die Grundgesamtheit

Es wird eine Annahme (Hypothese) über die Grundgesamtheit aufgestellt und dann anhand der Stichprobe auf ihre Wahrscheinlichkeit getestet.

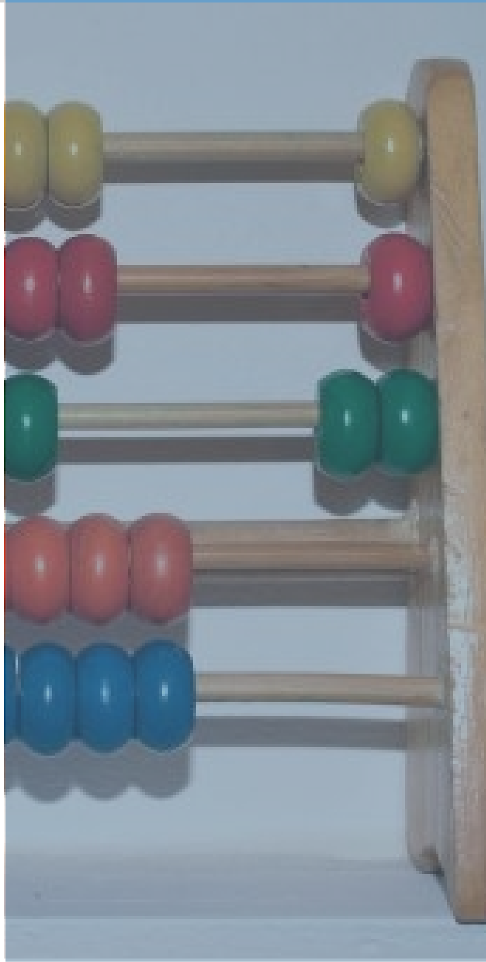
Gängige Fragen:

Wie hoch ist die Wahrscheinlichkeit, das zwei oder mehr Stichproben von unterschiedlichen Grundgesamtheiten stammen?
(Ist die Ausstattungssitte mit Grabbeigaben zwischen Männern und Frauen so unterschiedlich, das sich hier zwei unterschiedliche gesellschaftliche Gruppen zeigen?)
Zwei Stichproben
„Vergleichstest“?

Wie hoch ist die Wahrscheinlichkeit das eine gegebene Stichprobe von einer Grundgesamtheit mit bestimmten Eigenschaften stammt?
(Ist die Anzahl der Grabbeigaben zufällig oder gibt es ein Muster?)
Eine Stichprobe
Anpassungstest

Christian-Albrechts-Universität zu Kiel

Nichtparametrische Tests



parametrisch vs. nicht-parametrisch/parameterfrei:

parametrisch: Werte müssen bestimmter Verteilung folgen (z. B. Normalverteilung); Grundannahmen zur Verteilung sind notwendig

nicht-parametrisch: Annahmen zur Werteverteilung entfallen; keine Grundannahmen notwendig

Nicht-parametrische Tests, Vorteile - Nachteile:

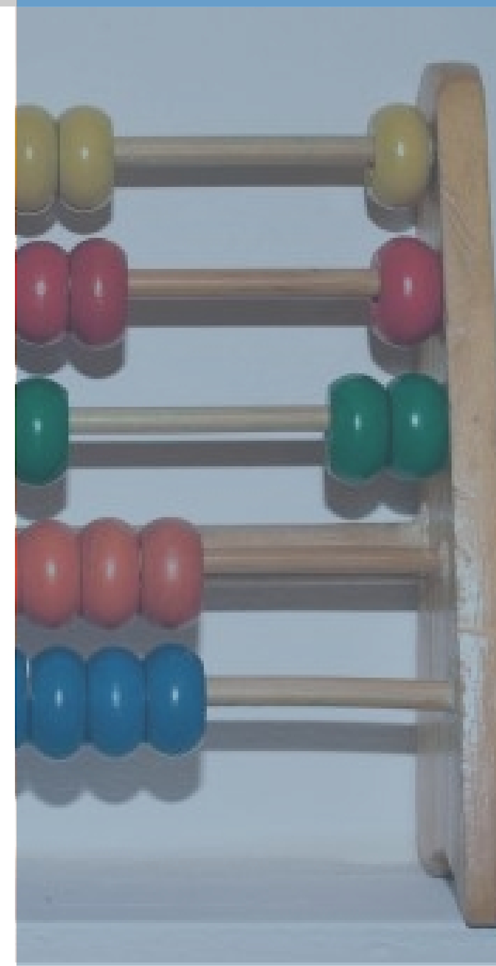
Vorteil: Sind auch anwendbar, wenn keine Aussage über die Verteilung möglich ist oder die Verteilung nicht den für parametrische Tests gegebenen Anforderungen entspricht.
Es können auch relativ kleine Stichproben getestet werden.

Nachteil: Haben meist eine geringere Power (Teststärke),

2

X^2-Test

X^2-Test [1]



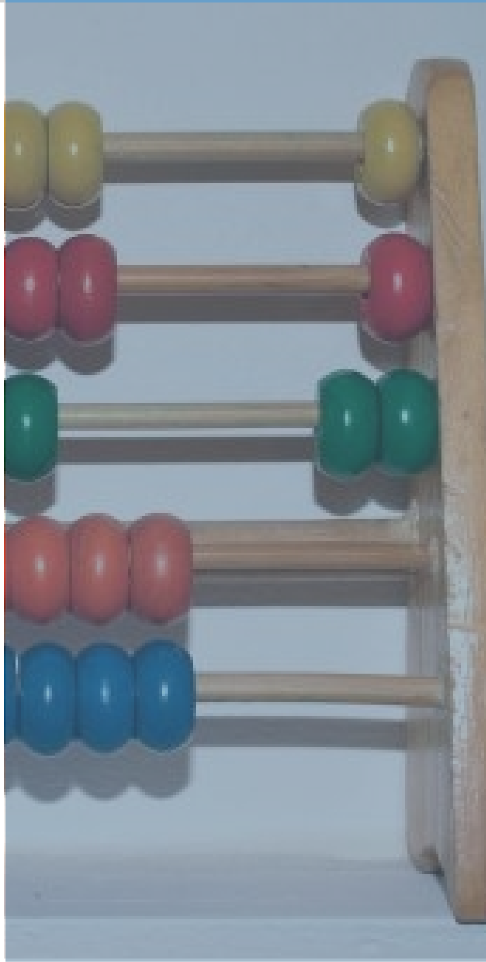
Mögliche Fragestellungen

Befinden sich Siedlungen eher auf besonders guten Böden oder ist die Verteilung zufällig?
Rückschlüsse auf Siedlungsverhalten und Wirtschaftsweise wären möglich.

Haben ältere Personen mehr Schuhleistenkeile als Grabbeigabe als jüngere?
Wenn Schuhleistenkeile Zeichen sozialen Ranges sind, würden sich Rückschlüsse auf Erblichkeit oder Erwerb während der Lebenszeit des sozialen Ranges ergeben.

Tests für nominale Daten sind möglich!
Daher von besonderem Wert für die Archäologie, weil wir häufig mit nominalen Daten arbeiten.

X^2-Test [2]



Test auf Gleichverteilung zweier Verteilungen

Voraussetzung: mindestens 1 nominal skalierte Variable (bei einer Stichprobe) und 1 nominalskalierte Gruppierungsvariable (bei 2 Stichproben)

Vorgehensweise bei einer Stichprobe: beobachtete Werte werden mit erwarteten Werten bei spezifischer Verteilung verglichen; kein Erwartungswert sollte < 5 sein; n sollte > 50 sein

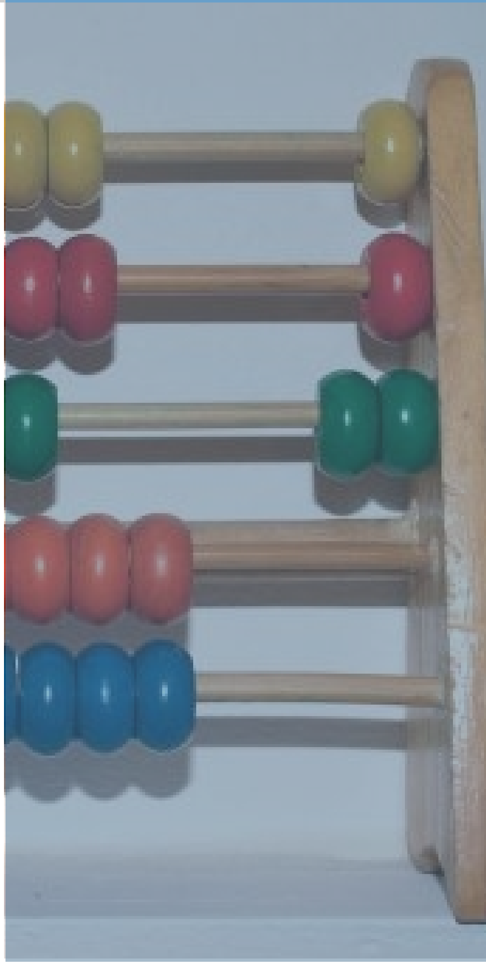
Vorgehensweise bei 2 Stichproben: beobachtete Werte beider Verteilungen werden mit erwarteten Werten bei Gleichverteilung verglichen; kein Erwartungswert sollte < 5 sein; n sollte > 50 sein

Prüfgröße: $\chi2$
Signifikanz ist abhängig von der Zahl der Freiheitsgrade (df)

Exkurs Freiheitsgrade

Anzahl der frei wählbaren Elemente bei gg. Randsummen

	Männlich	Weiblich	Summe
Brandbestattung			201
Körperbestattung			197
Summe	216	182	398

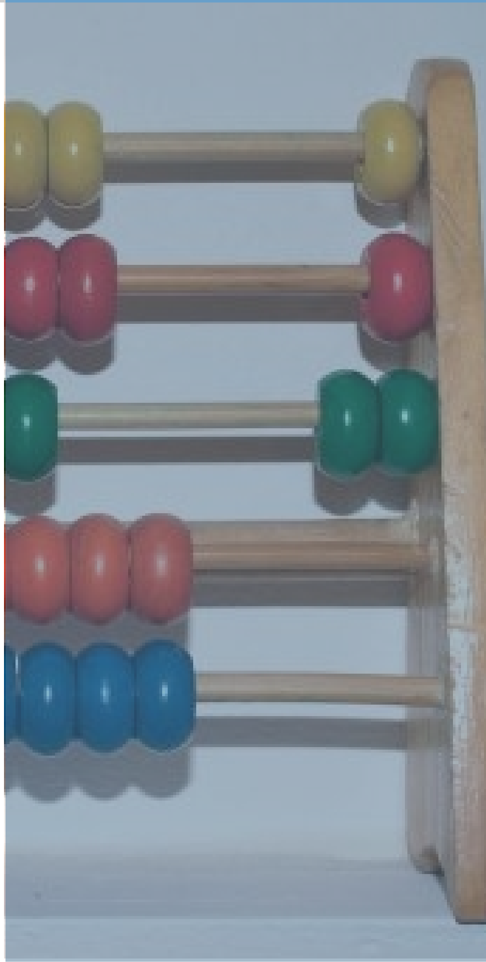


Christian-Albrechts-Universität zu Kiel

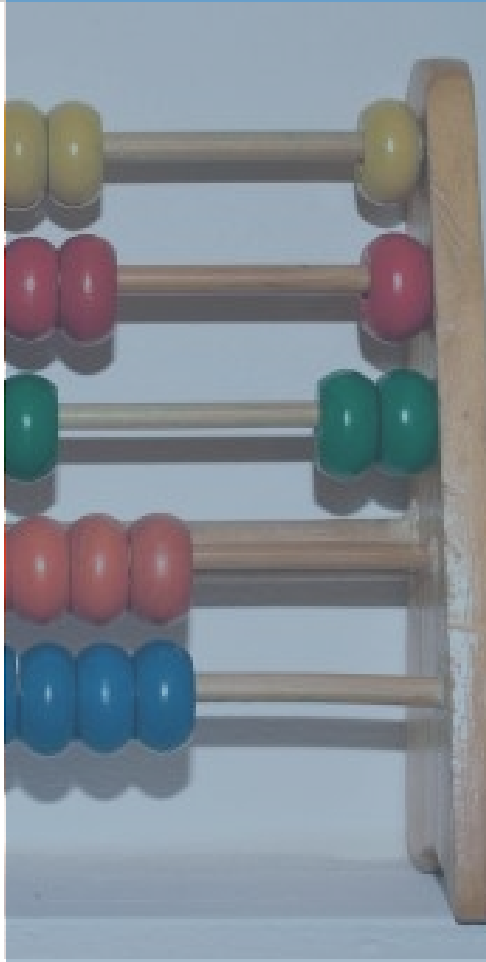
Exkurs Freiheitsgrade

Anzahl der frei wählbaren Elemente bei gg. Randsummen

	Männlich	Weiblich	Summe
Brandbestattung	123		201
Körperbestattung			197
Summe	216	182	398



Exkurs Freiheitsgrade



Anzahl der frei wählbaren Elemente bei gg. Randsummen

	Männlich	Weiblich	Summe
Brandbestattung	123	78	201
Körperbestattung	93	104	197
Summe	216	182	398

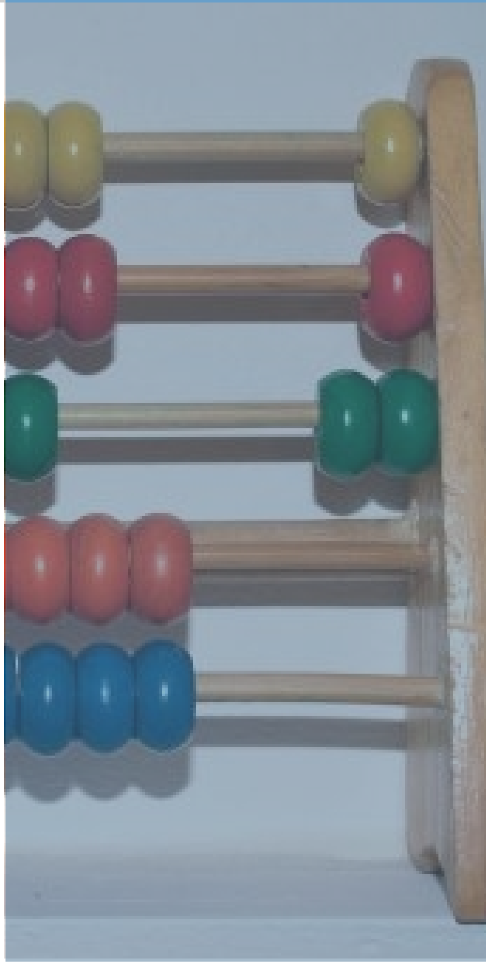
df=1: Wenn ein Wert gewählt ist, ergeben sich die anderen aus den Randsummen

(Zahl der Spalten – 1)*(Zahl der Zeilen – 1)

Exkurs Freiheitsgrade

Anzahl der frei wählbaren Elemente bei gg. Randsummen

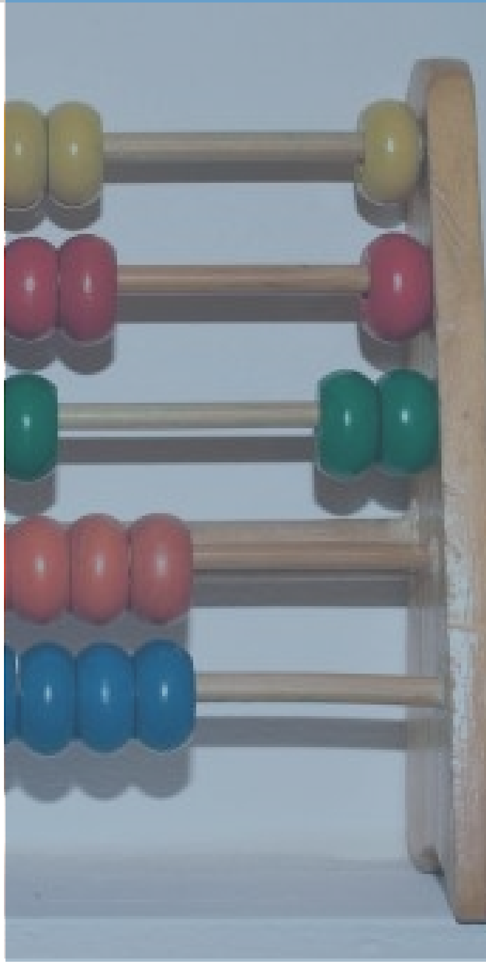
	Männlich	Weiblich	unsicher	Summe
Brandbestattung				201
Körperbestattung				197
Summe	196	179	23	398



Exkurs Freiheitsgrade

Anzahl der frei wählbaren Elemente bei gg. Randsummen

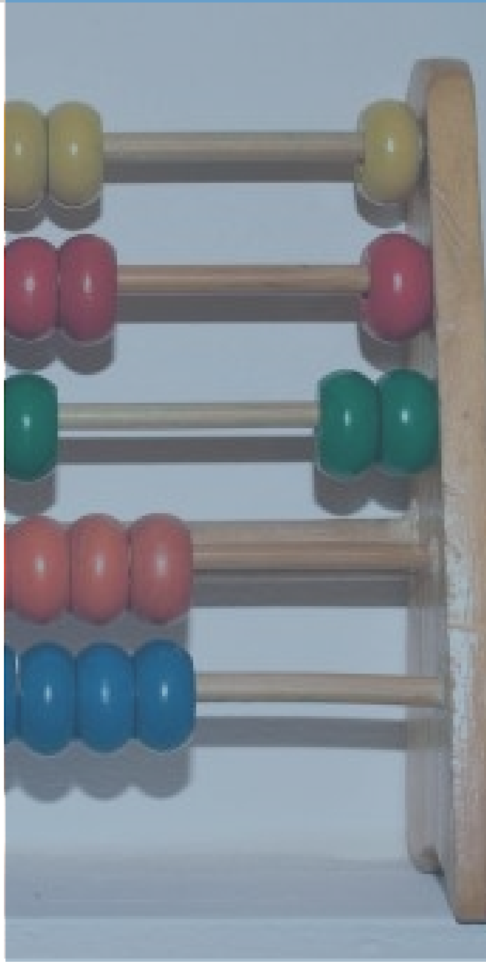
	Männlich	Weiblich	unsicher	Summe
Brandbestattung		78		201
Körperbestattung				197
Summe	196	179	23	398



Exkurs Freiheitsgrade

Anzahl der frei wählbaren Elemente bei gg. Randsummen

	Männlich	Weiblich	unsicher	Summe
Brandbestattung	113	78		201
Körperbestattung				197
Summe	196	179	23	398



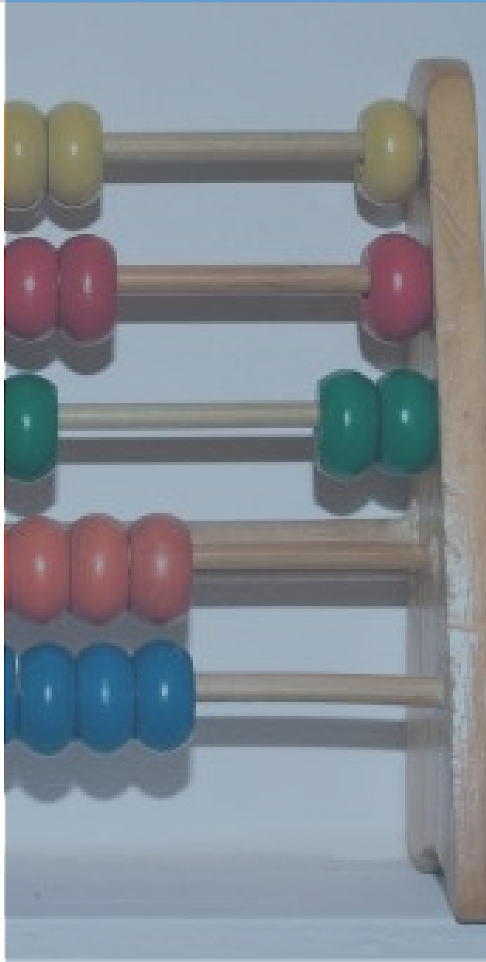
Exkurs Freiheitsgrade

Anzahl der frei wählbaren Elemente bei gg. Randsummen

	Männlich	Weiblich	unsicher	Summe
Brandbestattung	113	78	10	201
Körperbestattung	83	101	13	197
Summe	196	179	23	398

df=2: Wenn zwei Werte gewählt sind, ergeben sich die anderen aus den Randsummen

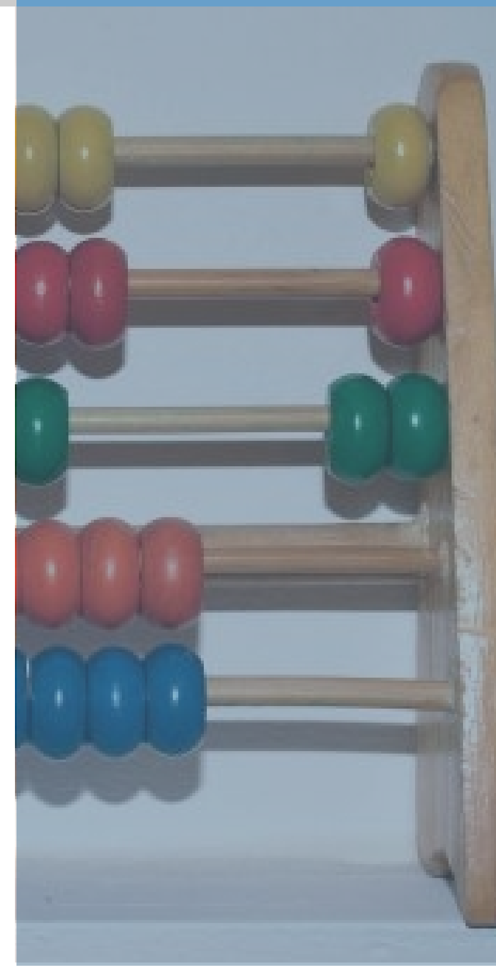
(Zahl der Spalten – 1)*(Zahl der Zeilen – 1)



Exkurs Freiheitsgrade

Anzahl der frei wählbaren Elemente bei gg. Randsummen

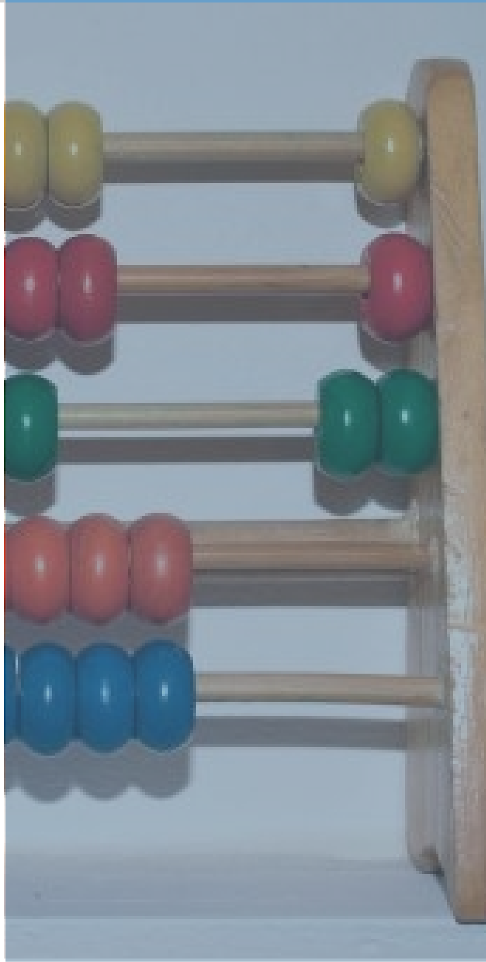
	Männlich	Weiblich	unsicher	Summe
Brandbestattung				201
Körperbestattung				197
unsicher				30
Summe	201	187	40	398



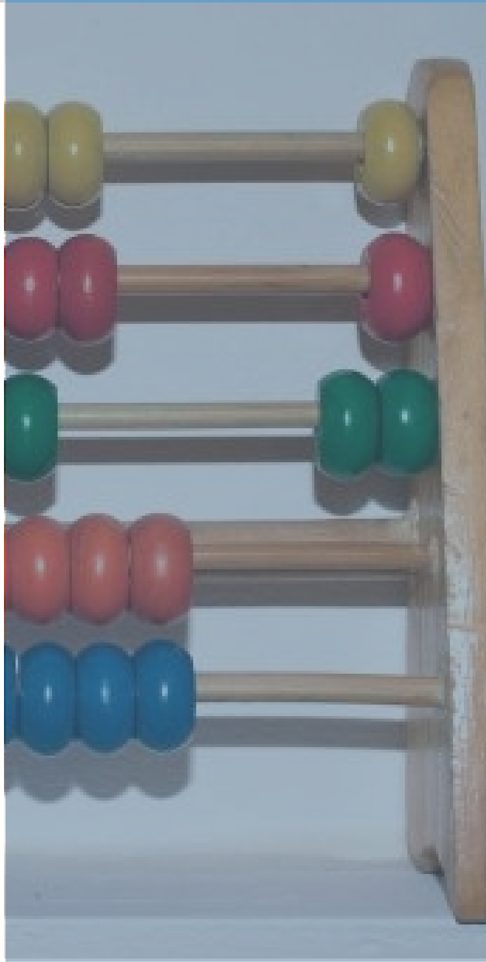
Exkurs Freiheitsgrade

Anzahl der frei wählbaren Elemente bei gg. Randsummen

	Männlich	Weiblich	unsicher	Summe
Brandbestattung		78		201
Körperbestattung	83		13	197
unsicher		8		30
Summe	201	187	40	398



Exkurs Freiheitsgrade



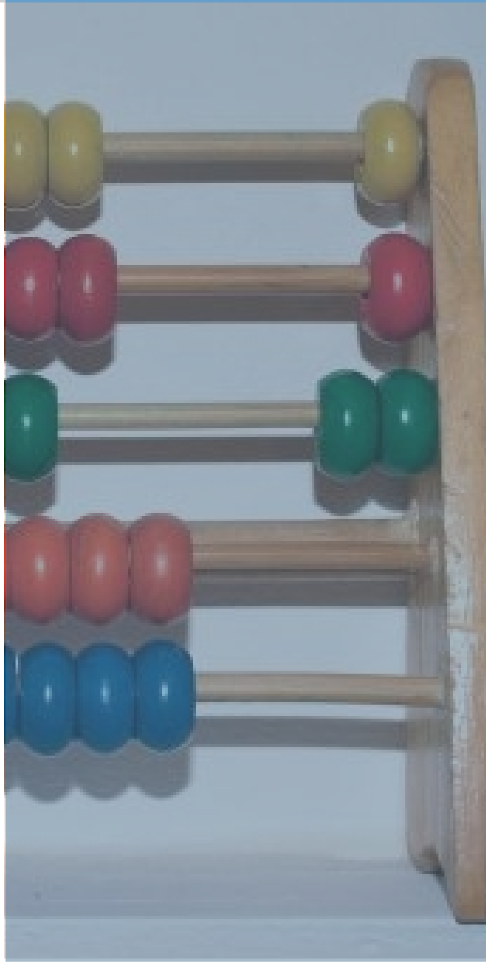
Anzahl der frei wählbaren Elemente bei gg. Randsummen

	Männlich	Weiblich	unsicher	Summe
Brandbestattung	113	78	10	201
Körperbestattung	83	101	13	197
unsicher	5	8	17	30
Summe	201	187	40	398

df=4: Wenn vier Werte gewählt sind, ergeben sich die anderen aus den Randsummen

(Zahl der Spalten – 1)*(Zahl der Zeilen – 1)

X²-Test [3]



Test für eine Stichprobe (Beispiel nach Shennan)

Anzahlen neolithischer Siedlungen nach Bodentyp im östlichen Frankreich.

Bodenart	Anzahl der Siedlungen
Rendzina	26
Alluvial	9
Braunerde	18
Gesamt	**53**

Frage: Gibt es eine signifikante Bevorzugung für einen Bodentyp?
Wir berechnen zwei Varianten:

1. Gleichmäßig verteilt

2. Gleichmäßig verteilt mit Berücksichtigung des Anteils der Bodentypen an der Gesamtfläche

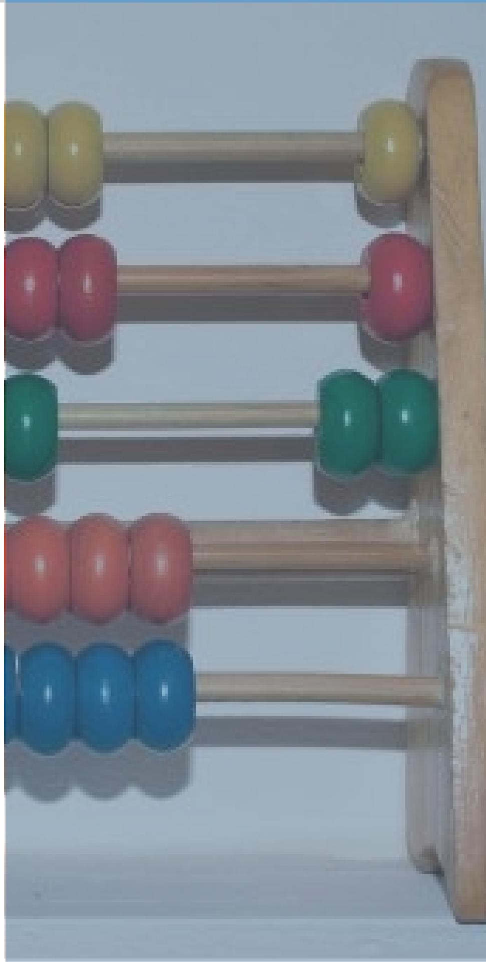
X²-Test [4]

Variante 1: Gleichmäßige Verteilung

Bodenart	Anzahl der Siedlungen	Anteil der Bodenart	Erwartete Anzahl der Siedlungen
Rendzina	26	1/3	17,66667
Alluvial	9	1/3	17,66667
Braunerde	18	1/3	17,66667
Gesamt	53	1	53

1. Gleichmäßig verteilt

H_0: Die Siedlungen sind gleichmäßig über alle Bodentypen verteilt
H_1: Die Siedlungen sind nicht gleichmäßig über alle Bodentypen verteilt



X²-Test [5]

Variante 1: Gleichmäßige Verteilung

Bodenart	Anzahl der Siedlungen	Anteil der Bodenart	Erwartete Anzahl der Siedlungen
Rendzina	26	1/3	17,66667
Alluvial	9	1/3	17,66667
Braunerde	18	1/3	17,66667
Gesamt	**53**	1	53

Formel für X²

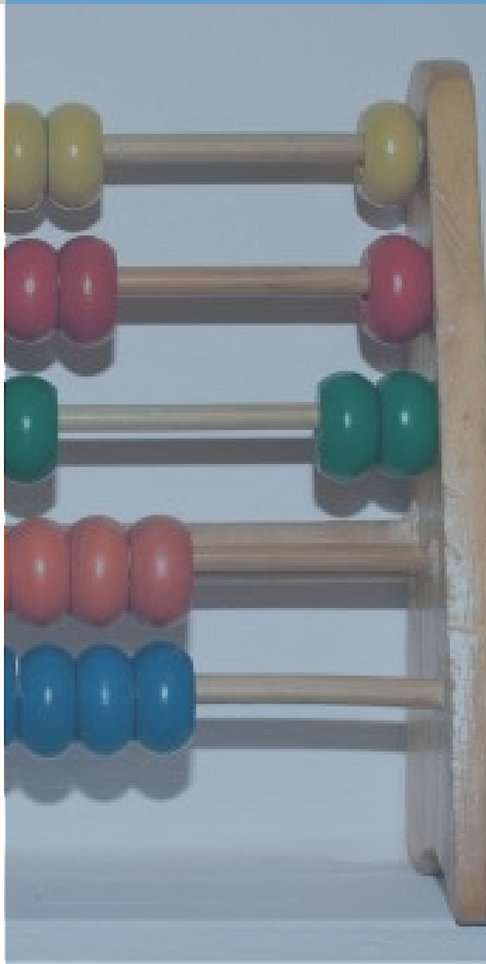
$$\chi^2 = \sum_{i=1}^{k} \frac{(B_i - E_i)^2}{E_i}$$

$$B_i: Anzahl\ der\ Beobachteten\ Fälle$$

$$E_i: Anzahl\ der\ Erwarteten\ Fälle$$

$$k: Anzahl\ der\ Kategorien$$

$$\chi^2: Symbol\ für\ die\ Testgröße\ Chi-Quadrat$$



Christian-Albrechts-Universität zu Kiel

X²-Test [6]

Variante 1: Gleichmäßige Verteilung

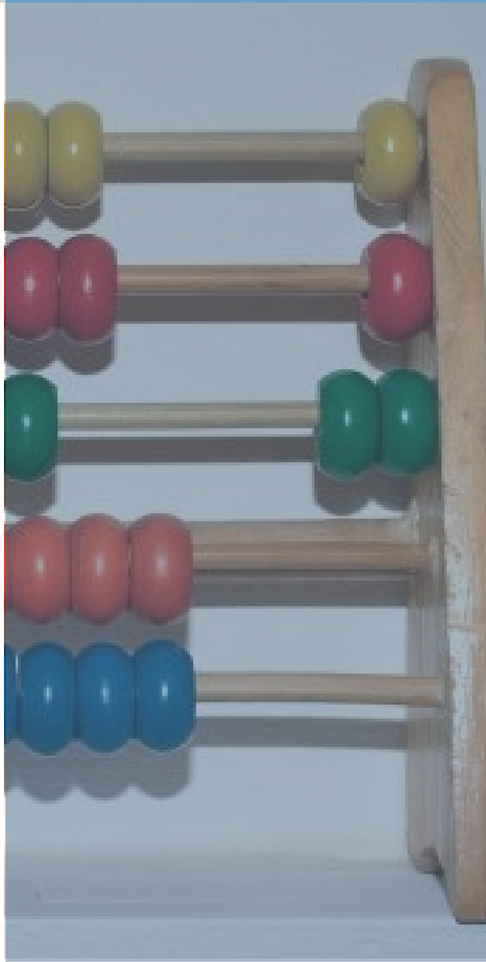
$$\chi^2 = \sum_{i=1}^{k} \frac{(B_i - E_i)^2}{E_i}$$

Berechnen des X²-Wertes

Bodenart	Anzahl der Siedlungen	Erwartete Anzahl der Siedlungen	O_i-E_i	$(O_i$-$E_i)^2$	$(O_i$-$E_i)^2/E_i$
Rendzina	26	17,66667	8,33333	69,44439	3,93081
Alluvial	9	17,66667	-8,66667	75,11117	4,25158
Braunerde	18	17,66667	0,33333	0,11111	0,00629
Gesamt	**53**	53			**8,18868**

Nachschlagen in Tabelle (z.B. Shennan):
Df=2 (2 Spalten (Beobachtet, Erwartet), 3 Kategorien)
Signifikanzniveau: 0,05
Grenzwert: 5,99145

Signifikanter Unterschied: Die Verteilung weicht von einer gleichmäßigen Verteilung ab!

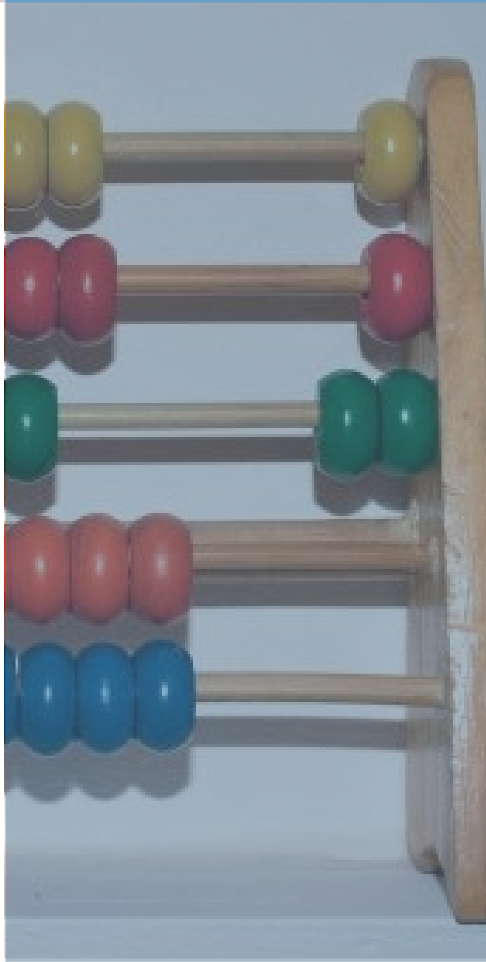


X^2-Test [7]

Variante 2: Gleichmäßig verteilt mit Berücksichtigung des Anteils der Bodentypen an der Gesamtfläche

Bodenart	Anzahl der Siedlungen	Anteil der Bodenart	Erwartete Anzahl der Siedlungen
Rendzina	26	32%	16,96
Alluvial	9	25%	13,25
Braunerde	18	34%	22,79
Gesamt	**53**	1	53

Formel für X² $$\chi^2 = \sum_{i=1}^{k} \frac{(B_i - E_i)^2}{E_i}$$



X^2-Test [8]

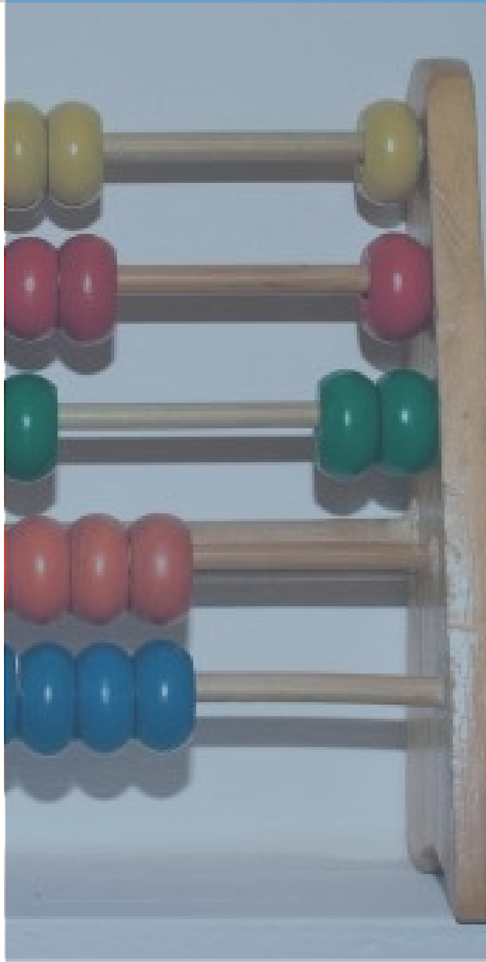
Variante 2: Gleichmäßig verteilt mit Berücksichtigung des Anteils der Bodentypen an der Gesamtfläche

$$\chi^2=\sum_{i=1}^{k}\frac{(B_i-E_i)^2}{E_i}$$

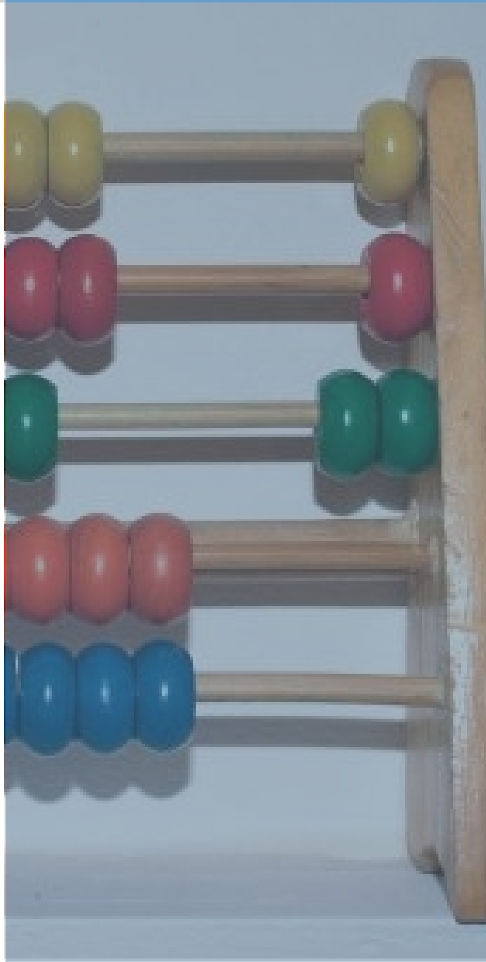
Bodenart	Anzahl der Siedlungen	Erwartete Anzahl der Siedlungen	O_i-E_i	$(O_i$-$E_i)^2$	$(O_i$-$E_i)^2/E_i$
Rendzina	26	16,96	9,04	81,7216	4,81849
Alluvial	9	13,25	-4,25	18,0625	1,36321
Braunerde	18	22,79	-4,79	22,9441	1,00676
Gesamt	**53**	53			**7,18846**

Nachschlagen in Tabelle (z.B. Shennan):
Df=2 (2 Spalten (Beobachtet, Erwartet), 3 Kategorien)
Signifikanzniveau: 0,05
Grenzwert: 5,99145

Signifikanter Unterschied: Die Verteilung weicht von der prozentual erwarteten Verteilung ab!



X^2-Test [9]



Das ganze in R

```
> siedlungen<-c(26,9,18)
> names(siedlungen)<-c("Rendzina","Alluvial","Braunerde")
> siedlungen
 Rendzina  Alluvial Braunerde
       26         9        18
```

Variante 1: Gleichmäßig verteilt

```
> chisq.test(siedlungen)

     Chi-squared test for given probabilities

data:  siedlungen
X-squared = 8.1887, df = 2, p-value = 0.01667
```

Variante 1: Gleichmäßig verteilt nach prozentualen Anteilen

```
> chisq.test(siedlungen,p=c(0.32,0.25,0.43))

     Chi-squared test for given probabilities

data:  siedlungen
X-squared = 7.1885, df = 2, p-value = 0.02748
```

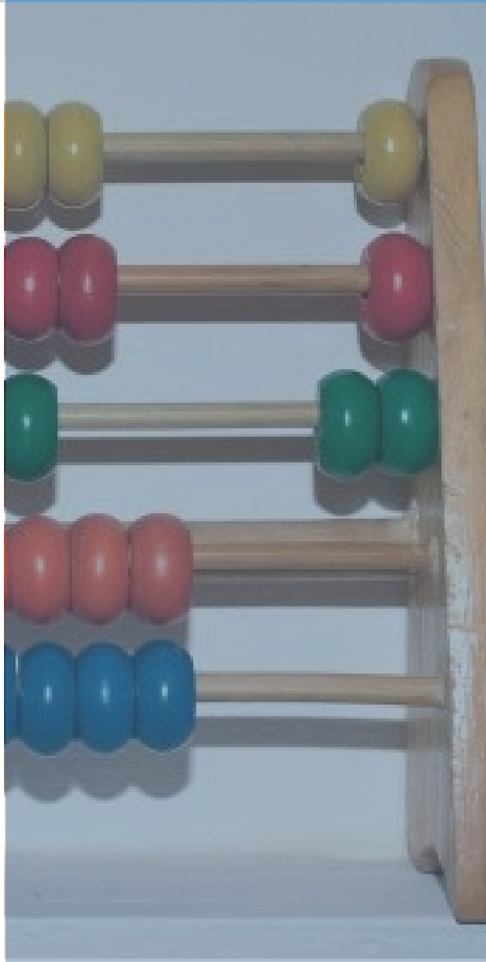
X^2-Test [10]

Test für zwei Stichproben (Test auf Unabhängigkeit)
(Beispiel nach Hinz, geschönt)

Vergleich von Vorhandensein von Bernstein in Gräbern und Siedlungen
Klassischer Vier-Felder-Test

Fundkategorie	Bernstein		Randsumme
	+	-	
Siedlung	6	18	24
Grab	132	44	176
Randsumme	**138**	62	200

Ist Bernstein primär als Grabbeigabe anzusehen?
Df=1
Signifikanzniveau=0.05

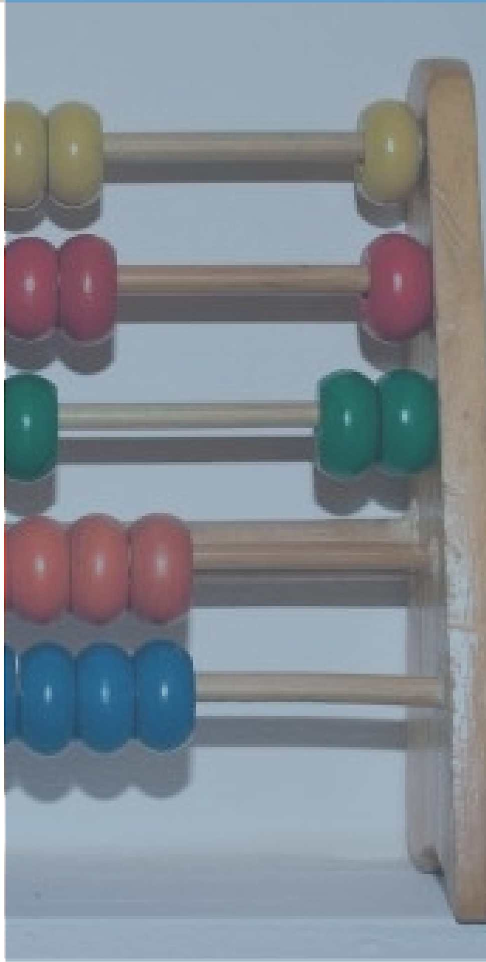


X^2-Test [11]

Berechnen der Erwartungswerte

Multiplikation der Randsummen, geteilt durch Gesamtzahl

Fundkategorie	Bernstein		Randsumme
	+	-	
Siedlung	6 E=24*138/200 =16,56	18 E=24*62/200 =7,44	24
Grab	132 E=138*176/20 0 =121,44	44 E=62*176/200 =54,56	176
Randsumme	**138**	62	200



X^2-Test [12]

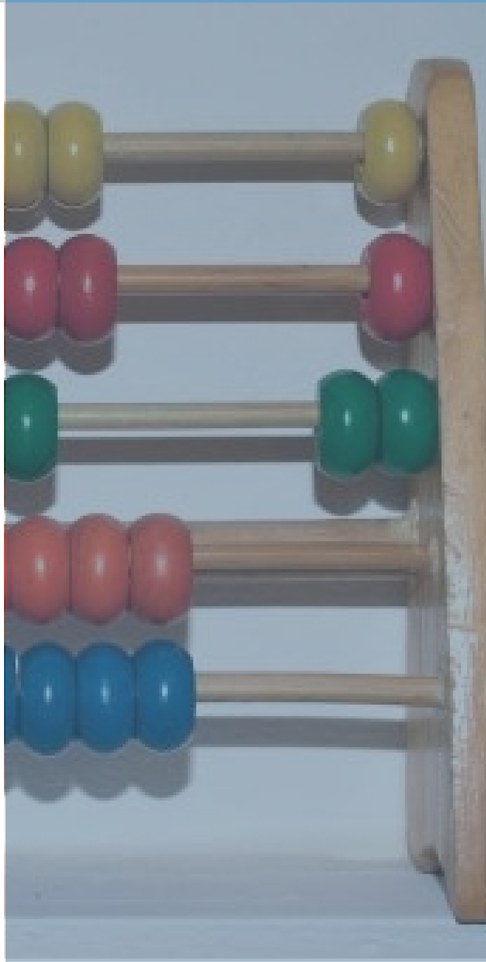
Berechnen der X^2-Werte
(Beobachtet/Erwartet)²/Erwartet

$$\chi^2 = \sum_{i=1}^{k} \frac{(B_i - E_i)^2}{E_i}$$

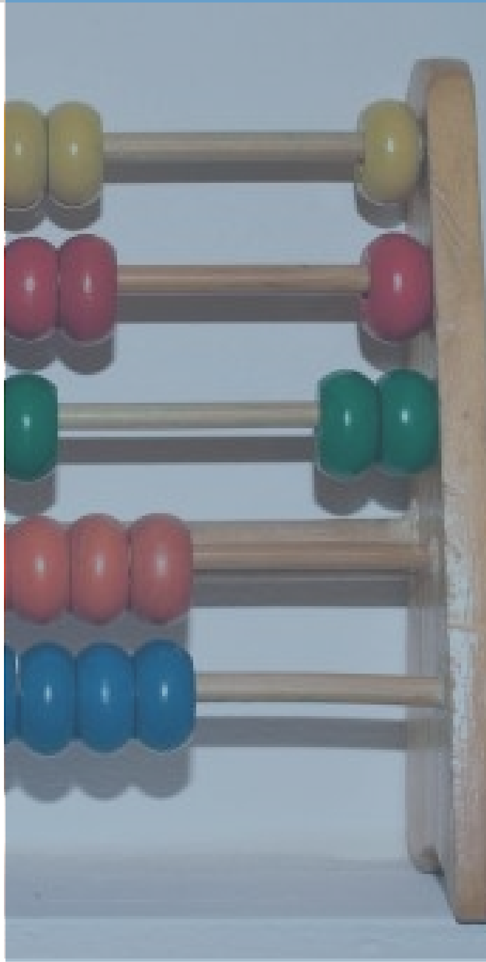
Fundkategorie	Bernstein		Randsumme
	+	-	
Siedlung	(6-16,56)²/16,56 =6,73	(18-7,44)²/7,44 =14,99	24
Grab	(132-121,44)²/121,44 =0,92	(44-54,56)²/54,56 =2,04	176
Randsumme	**138**	62	200

Ist Bernstein primär als Grabbeigabe anzusehen?
Df=1; Signifikanzniveau=0.05
X^2=24,68; Grenzwert bei df=1 und p=0.05: 3,84146

Der Unterschied in der Verteilung ist statistisch signifikant nicht zufällig! Die beiden Variablen sind abhängig voneinander!



X²-Test [13]



Das ganze in R

```
> vergleich<-matrix(c(6,132,18,44),ncol=2)
> colnames(vergleich)<-c("mit Bernstein","ohne Bernstein")
> rownames(vergleich)<-c("Siedlung","Grab")
> vergleich
         mit Bernstein ohne Bernstein
Siedlung             6             18
Grab               132             44
> chisq.test(vergleich)

     Pearson's Chi-squared test with Yates' continuity correction

data:  vergleich
X-squared = 22.4022, df = 1, p-value = 2.211e-06
> chisq.test(vergleich,correct=F)

     Pearson's Chi-squared test

data:  vergleich
X-squared = 24.6844, df = 1, p-value = 6.753e-07
```

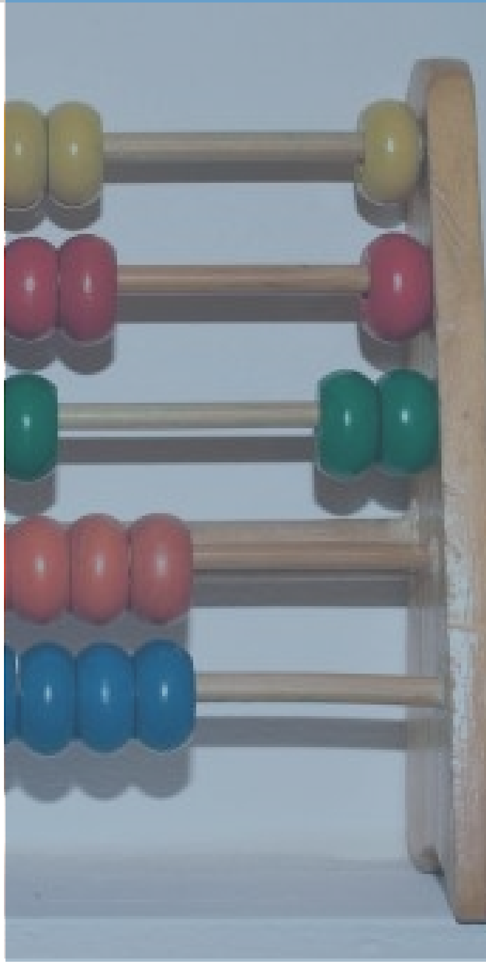
Correct: Yates Korrektur für kleine Datenmengen → (|O-E|-**0,5**)²/E

X^2-Test Aufgabe

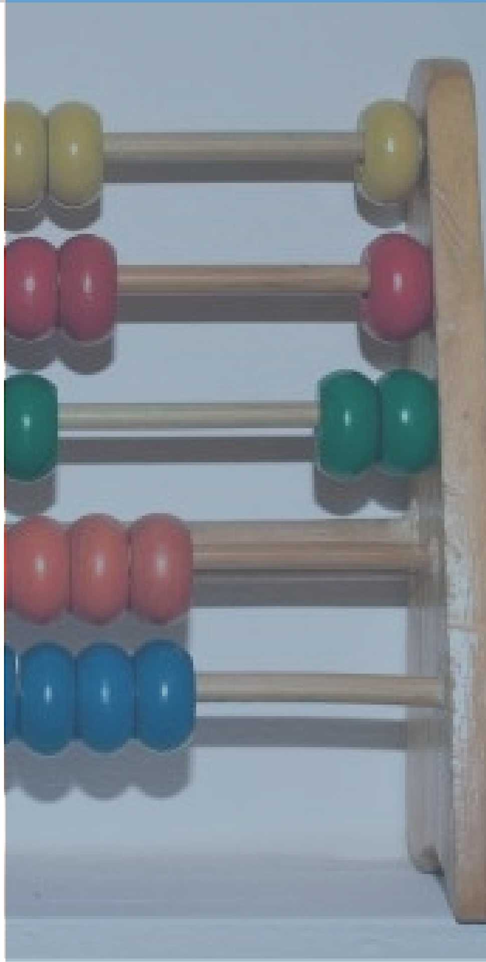
Tierknochen aus den Mittel- und Spätneolithischen Schichten in Wolkenwehe (Mischka et al. 2005)
Gegeben sind folgende Werte:

Schicht	Haustierknochen	Wildtierknochen
202 (Spätneolithikum)	159	32
203 (Mittelneolithikum)	84	54

Prüfen Sie, ob die beobachteten Unterschiede statistisch signifikant sind!



X^2-Test Aufgabe



Tierknochen aus den Mittel- und Spätneolithischen Schichten in Wolkenwehe (Mischka et al. 2005)

Schicht	Haustierknochen	Wildtierknochen
202 (Spätneolithikum)	159	32
203 (Mittelneolithikum)	84	54

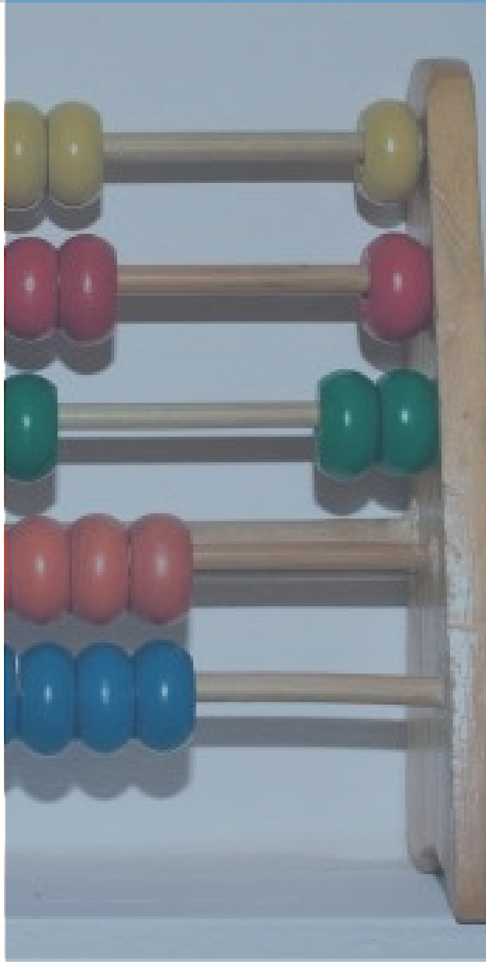
```
> test<-matrix(c(159,84,32,54),ncol=2)
> colnames(test)<-c("schicht 202","schicht 203")
> rownames(test)<-c("Haustier","Wildtier")
> test
         schicht 202 schicht 203
Haustier         159          32
Wildtier          84          54
> chisq.test(test)

     Pearson's Chi-squared test with Yates' continuity correction

data:  test
X-squared = 19.6344, df = 1, p-value = 9.376e-06
```

Christian-Albrechts-Universität zu Kiel

Zusammenhangsmaße [1]



Messen der Stärke des Zusammenhangs zweier Variablen

X^2 ist bereits ein Maß für die Stärke des Zusammenhangs:

Zusammenhang ↑ X^2 ↑ ↔ Zusammenhang ↓ X^2 ↓

Aber: X^2 Abhängig von N

Fundkategorie	Bernstein		Randsumme
	+	-	
Siedlung	6	18	24
Grab	132	44	176
Randsumme	**138**	62	200

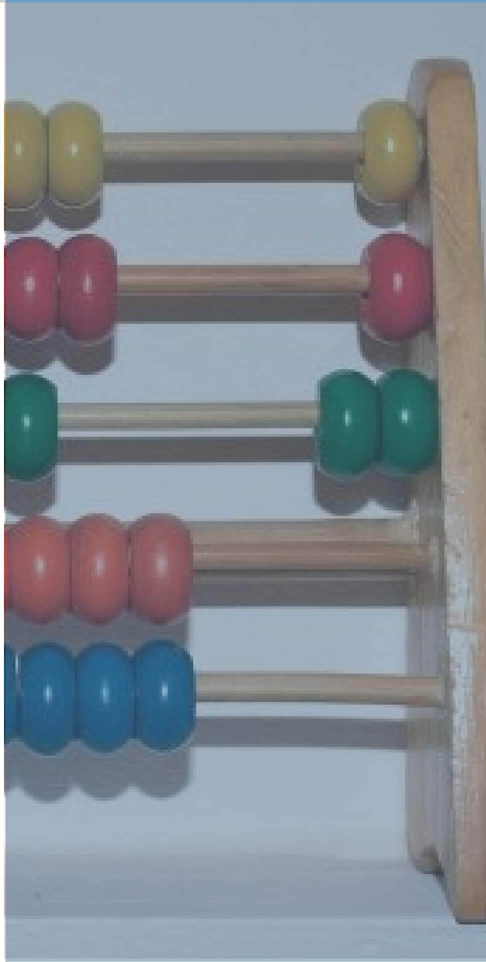
$X^2 = 24.6844$

Fundkategorie	Bernstein		Randsumme
	+	-	
Siedlung	12	36	48
Grab	264	88	352
Randsumme	**276**	124	400

$X^2 = 49.3689$

Cramers V
(oder φ)

Zusammenhangsmaße [2]



Cramers V

Normalisieren von X² auf die Anzahl der Beobachtungen N,
Wurzel ziehen,
Teilen durch den kleineren Wert von (Spaltenzahl, Zeilenzahl) -1

Fundkategorie	Bernstein		Randsumme
	+	-	
Siedlung	6	18	24
Grab	132	44	176
Randsumme	**138**	62	200

X² = 24.6844

$$\phi=\sqrt{\frac{\chi^2}{n*(min(zeilen,spalten)-1)}}$$

$$\phi=\sqrt{\frac{22{,}6844}{200*(min(2{,}2)-1)}}$$

$$\phi=0{,}351314901$$

Fundkategorie	Bernstein		Randsumme
	+	-	
Siedlung	12	36	48
Grab	264	88	352
Randsumme	**276**	124	400

X² = 49.3689

$$\phi=\sqrt{\frac{\chi^2}{n*(min(zeilen,spalten)-1)}}$$

$$\phi=\sqrt{\frac{49{,}3689}{400*(min(2{,}2)-1)}}$$

$$\phi=0{,}351314901$$

Zusammenhangsmaße [3]

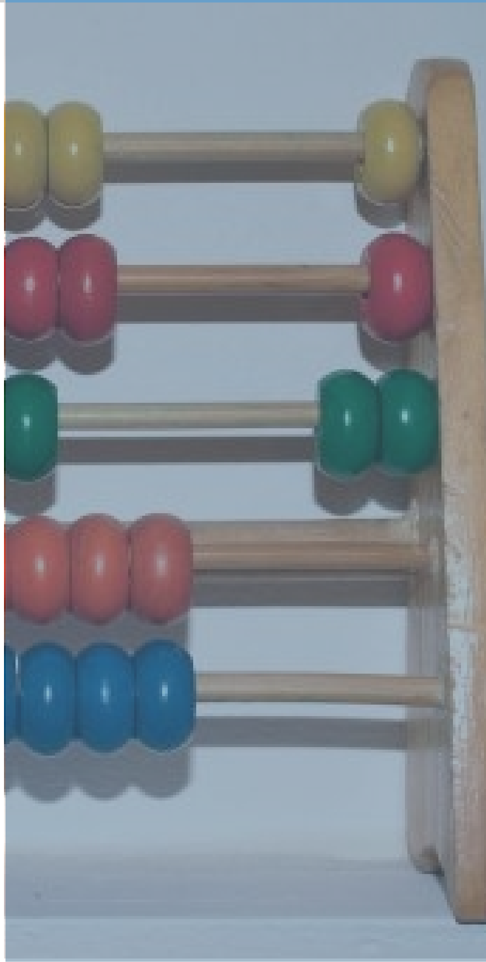
Cramers V

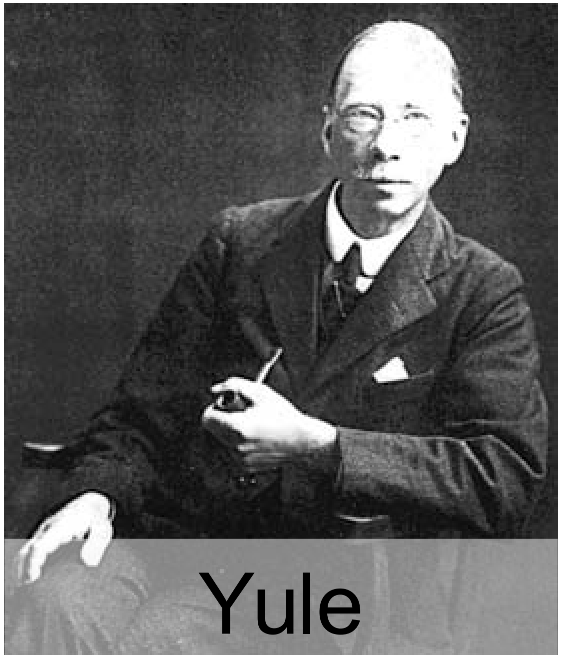
$$\phi=\sqrt{\frac{\chi^2}{n*(min(zeilen, spalten)-1)}}$$

Ist immer ein Wert zwischen 0 und 1
0: kein Zusammenhang
1: perfekter Zusammenhang

In R:

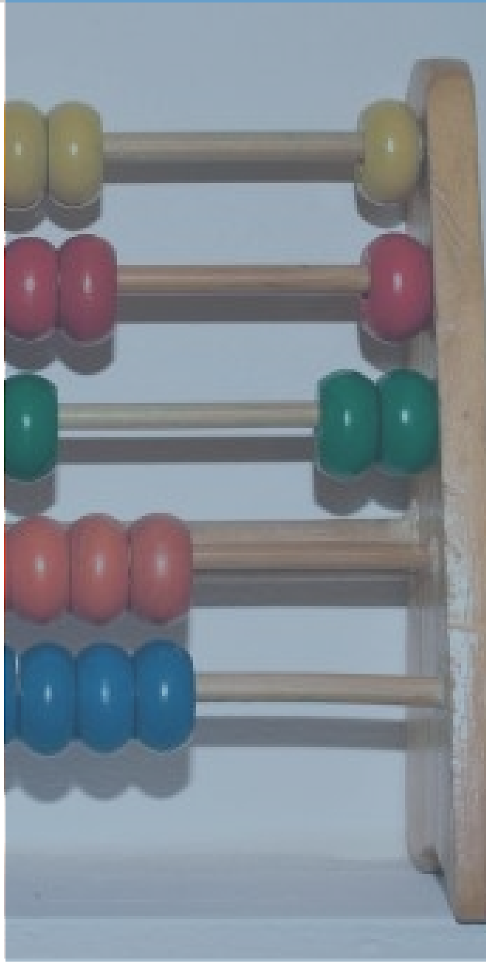
```
calc.CV <- function(x)
{
  CV <- sqrt(chisq.test(x, correct = FALSE)$statistic /
             (sum(x) * min(dim(x) - 1)))
  as.numeric(CV)
}
> calc.CV(test)
[1] 0.3513149
```





Yules Q

Zusammenhangsmaße [4]



Yule's Q

Ein anderes, einfacheres Maß für Zusammenhang, nur bei 2x2-Tafeln anwendbar

$$Q=\frac{ad-bc}{ad+bc}$$

Idee: Je größer die Zahl im linken oberen Feld im Verhältnis zur Gesamtzahl, desto größer der positive Zusammenhang

Fundkategorie	Bernstein		Randsumme
	+	-	
Siedlung	6	18	24
Grab (=nicht Siedlung)	132	44	176
Randsumme	**138**	62	200

$$Q=\frac{6*44-18*132}{6*44+18*132}=-0{,}8$$

Starker negativer Zusammenhang: Gräber (nicht Siedlungen) haben eine deutlich höhere Chance, Bernstein zu enthalten.
Allerdings: nicht anwendbar, wenn in einer der Zellen ein Null-Wert vorkommt!

Zusammenhangsmaße [3]

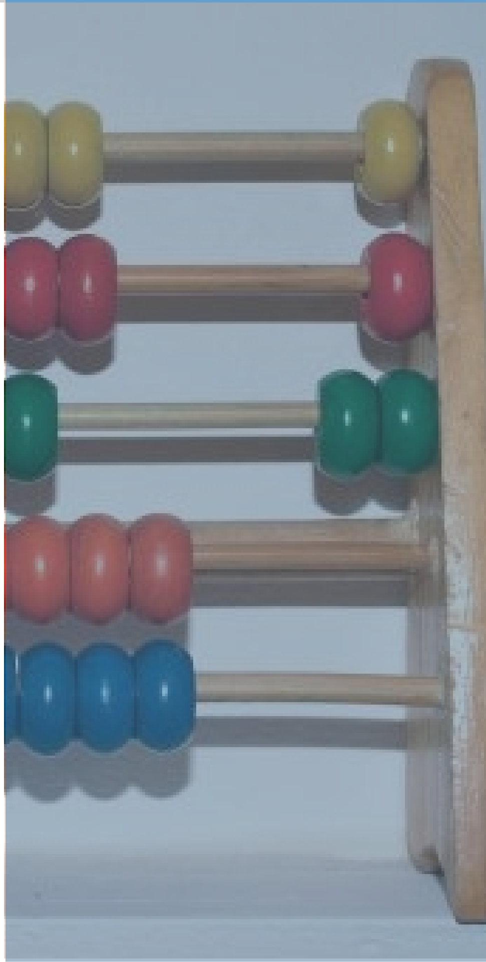
Yules Q

$$Q=\frac{ad-bc}{ad+bc}$$

Ist immer ein Wert zwischen -1 und 1
-1: perfekter negativer Zusammenhang
0: kein Zusammenhang
1: perfekter positiver Zusammenhang

In R:

```
calc.YQ <- function(x)
{
YQ <- (x[1,1]*x[2,2]-x[1,2]*x[2,1])/(x[1,1]*x[2,2]+x[1,2]*x[2,1])
as.numeric(YQ)
}
> calc.YQ(matrix(c(6,132,18,44),ncol=2))
[1] -0.8
```

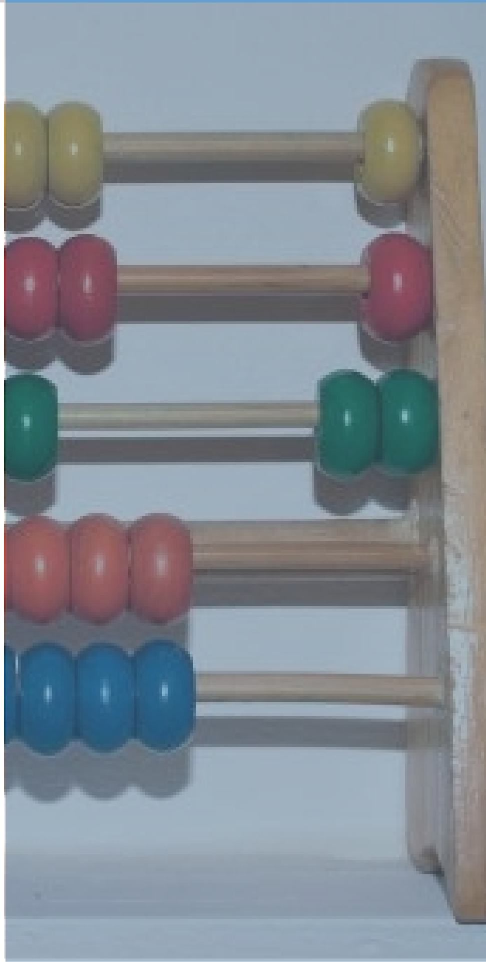


X^2-Test Aufgabe

Tierknochen aus den Mittel- und Spätneolithischen Schichten in Wolkenwehe (Mischka et al. 2005)
Gegeben sind folgende Werte:

Schicht	Haustierknochen	Wildtierknochen
202 (Spätneolithikum)	159	32
203 (Mittelneolithikum)	84	54

Prüfen Sie, wie stark der Zusammenhang (und damit die Änderung in der Wirtschaftsweise) ist!



X²-Test Aufgabe

Tierknochen aus den Mittel- und Spätneolithischen Schichten in Wolkenwehe (Mischka et al. 2005)

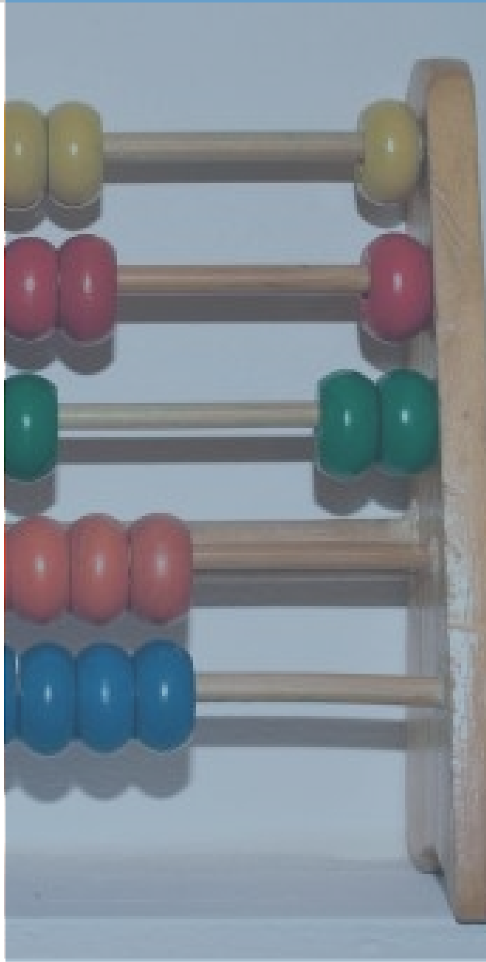
Schicht	Haustierknochen	Wildtierknochen
202 (Spätneolithikum)	159	32
203 (Mittelneolithikum)	84	54

```
> calc.CV(test)
[1] 0.2513021
```

Cramers V ist 0,25, schwache Assoziation von Haustierknochen mit dem Spät-, Wildtierknochen mit dem Mittelneolithikum

```
> calc.YQ(test)
[1] 0.5231506
```

Yules Q ist 0,52, positiver Zusammenhang zwischen Haustierknochen und dem Spätneolithikum

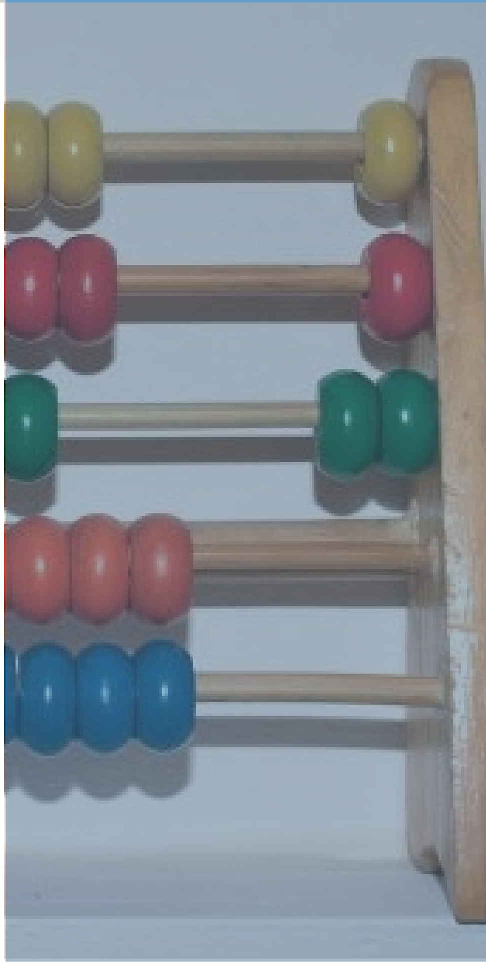


Fishers Test [1]

Problem bei zu kleinen Erwartungswerten

Fundkategorie	Bernstein		Randsumme
	+	-	
Siedlung	3 E=12*69/100 =8,28	9 E=24*62/200 **=3,72**	12
Grab	66 E=138*176/20 0 =60,72	22 E=62*176/200 =27,28	88
Randsumme	**69**	31	200

Kleiner als 5!



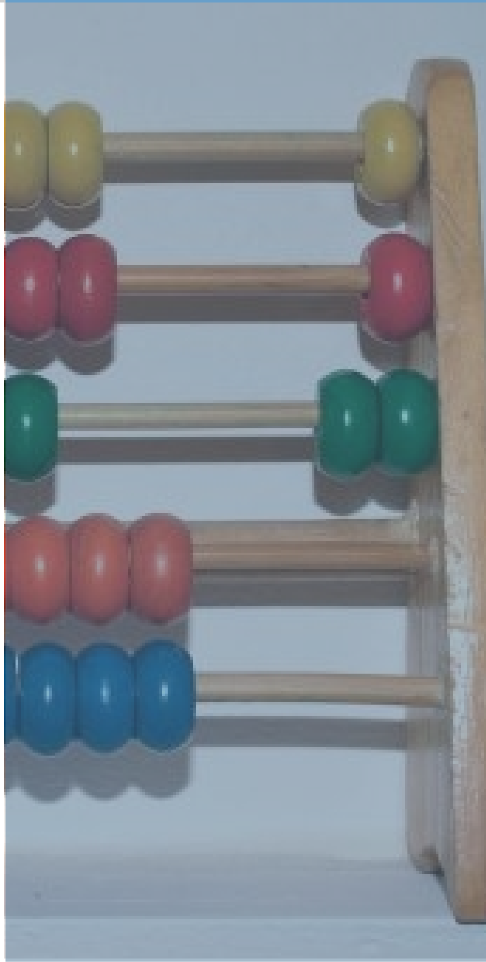
Fishers Test [2]

Test für zwei Stichproben (Test auf Unabhängigkeit) (Beispiel nach Hinz, orginal)

Exakter Test nach Fisher!

Fundkategorie	Bernstein		Randsumme
	+	-	
Siedlung	a: 3	b: 9	12
Grab	c: 66	d: 22	88
Randsumme	**69**	31	n: 100

$$\varphi=\frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!}=\frac{(3+9)!(66+22)!(3+66)!(9+22)!}{100!3!9!66!22!}$$



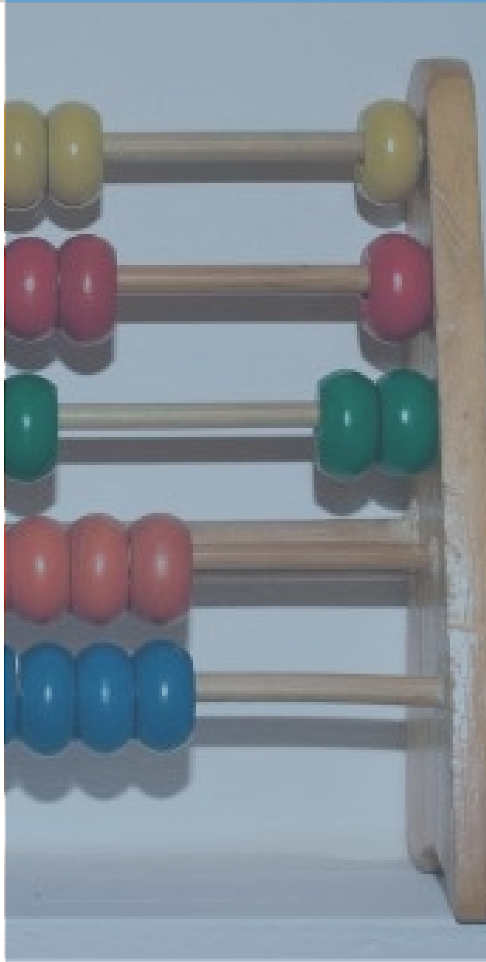
Fishers Test [3]

Das ganze in R

```
> vergleich<-matrix(c(3,66,9,22),ncol=2)
> colnames(vergleich)<-c("mit Bernstein","ohne Bernstein")
> rownames(vergleich)<-c("Siedlung","Grab")
> vergleich
         mit Bernstein ohne Bernstein
Siedlung             3              9
Grab                66             22
> fisher.test(vergleich)

     Fisher's Exact Test for Count Data

data:  vergleich
p-value = 0.001110
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.01825286 0.50879869
sample estimates:
odds ratio
 0.1141018
```



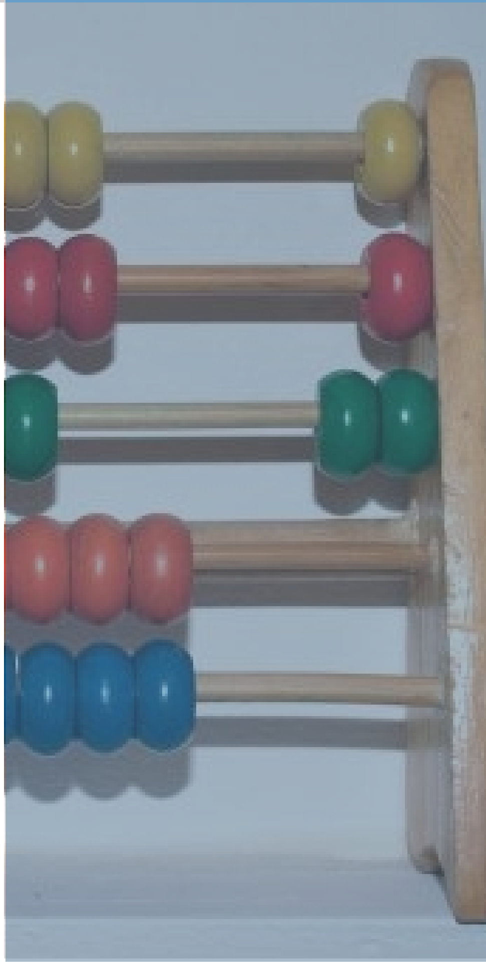
X^2-Test Aufgabe

Eberzähne bei Kugelamphorengräbern (Müller 2001, Zahlen verändert)

Gegeben sind folgende Werte:

Geschlecht	Eberzähne	
	Ja	Nein
Mann	11	7
Frau	1	6

Prüfen Sie, ob ein stat. Signifikanter Zusammenhang besteht!



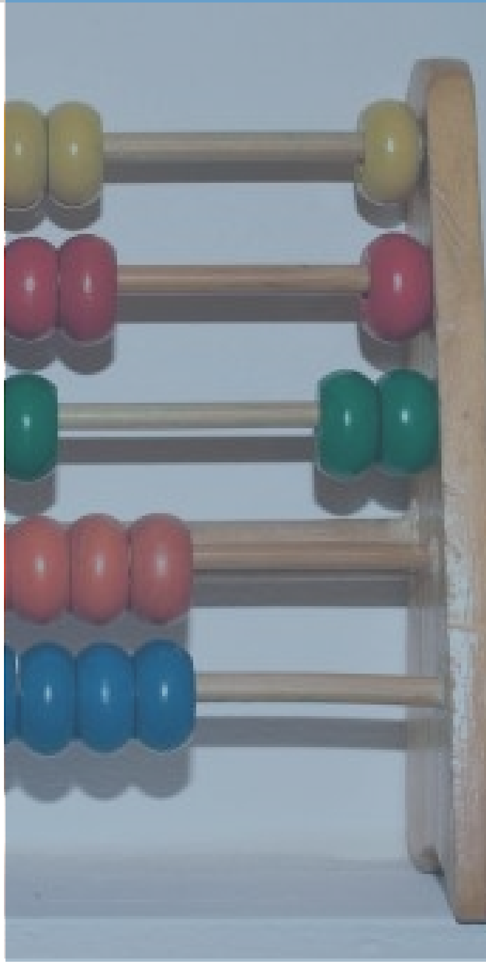
X^2-Test Aufgabe

Eberzähne bei Kugelamphorengräbern (Müller 2001, Zahlen verändert)

Gegeben sind folgende Werte:
Prüfen Sie, ob ein stat. Signifikanter Zusammenhang besteht!

Geschlecht	Eberzähne	
	Ja	Nein
Mann	11	7
Frau	1	6

```
> test<-matrix(c(11,7,1,6),ncol=2)
> chisq.test(test)
…
X-squared = 2.7501, df = 1, p-value = 0.09725
...
> chisq.test(test,correct=F)
...
X-squared = 4.4274, df = 1, p-value = 0.03537
…
> fisher.test(test)
…
p-value = 0.07304
```



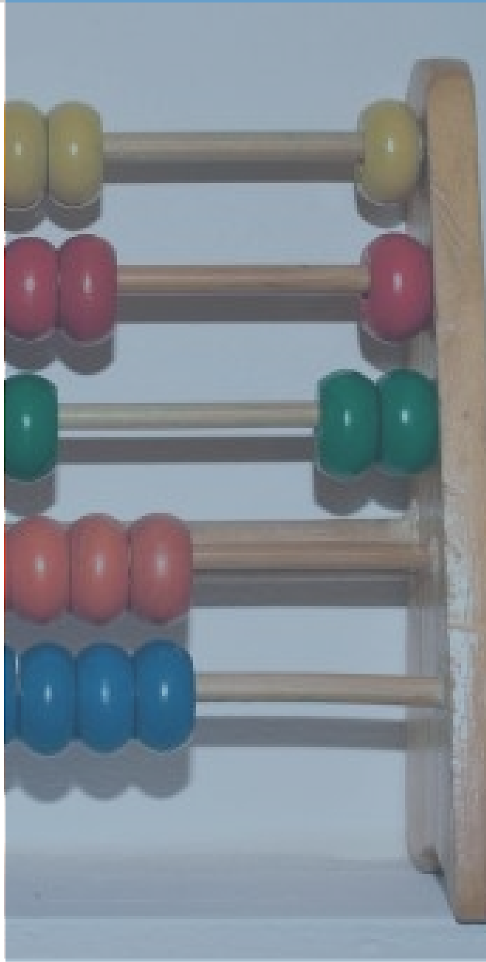
Interpretation von Tests

Statistische Assoziation bedeutet nicht Kausaler Zusammenhang!

Beispiel nach Shennan: Grabgröße und Geschlecht

Auch wenn ein statistisch signifikanter Zusammenhang zwischen Grabgröße und Geschlecht besteht, kann dieser durch einen dritten Faktor (hier Körpergröße) bedingt sein.

Eine Schlußfolgerung, wonach Grabgröße durch das Geschlecht bestimmt wird, wäre daher falsch!



Nächste Sitzung 9. Dezember 2014:
Theorie: Stichprobe und Population,
Wahrscheinlichkeitstheorie

Bitte lesen Sie Shennan Kapitel 7