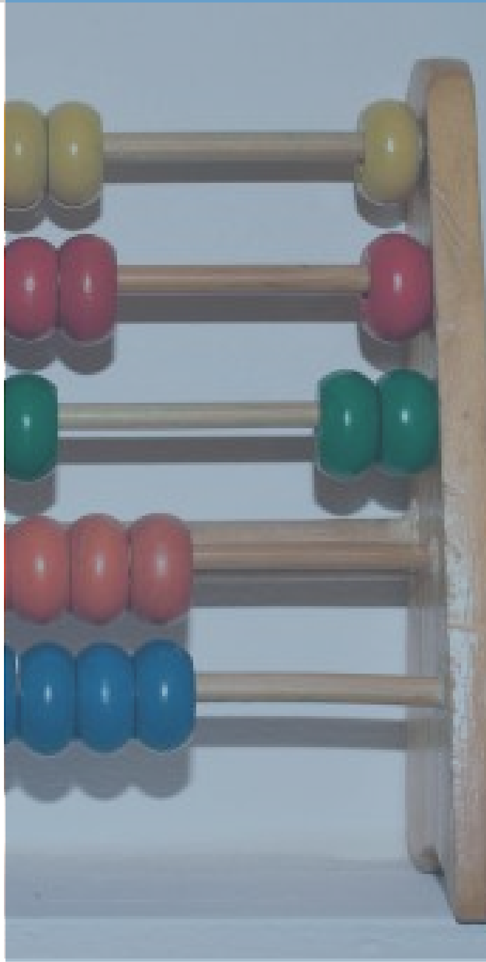
Christian-Albrechts-Universität zu Kiel

06_chi-squared/measurement_of_associations

chi-squared, Cramers V, Yules Q

Repetition: Statistical hypothesis testing



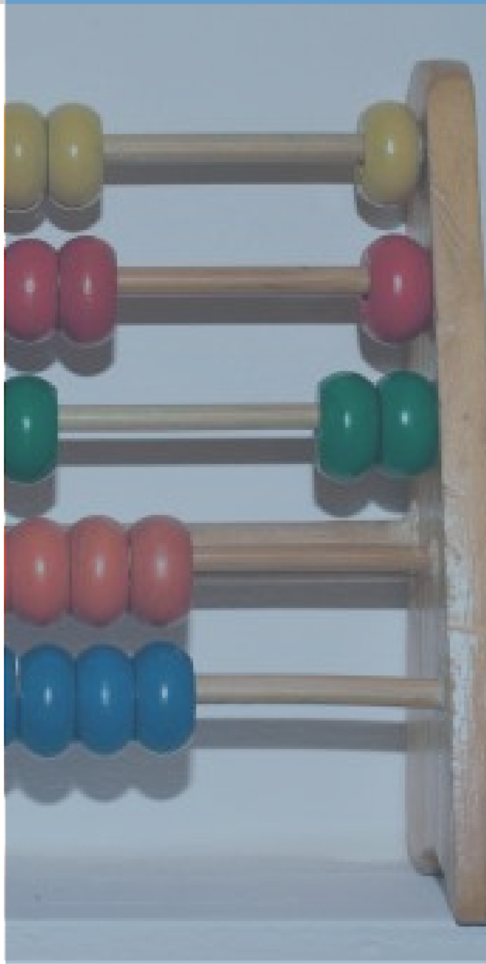
Validation of an assumption about the population
A assumption (hypothesis) about the population is made and than its probability is checked against the sample.

Usual questions:

How probable is it that two or more samples descend from the different/the same population?
(eg. Is the custom of grave goods for man and women so different that two different social groups are visible?)
Two samples
Test for independence

How probable is it that a given sample descend from a population with certain parameters?
(Is the amount of grave goods random or is a pattern visible?)
Two samples
Goodness-of-fit-Test

Repetition: Nonparametric tests



Parametric vs. nonparametric

Parametric: The distribution of the values have to be in a certain form (e.g. normal distribution); assumptions about the distribution of the population are needed

non-parametric: no assumptions about the distribution of the sample and the population are needed

Nonparametric tests, advantages and disadvantages:

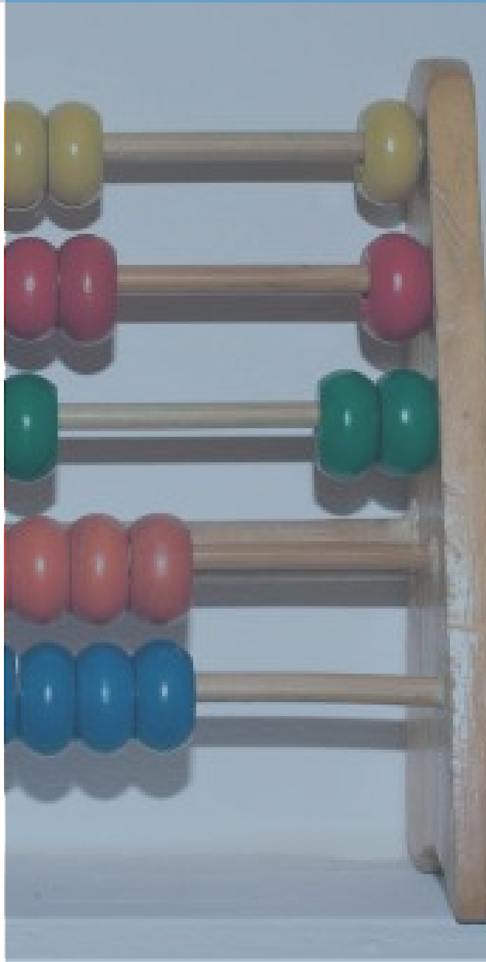
Advantage: Also appropriate if no statements about the distribution are possible or the distribution fits no for parametric tests.
Also smaller samples are possible.

Disadvantages: Tests have general a lesser power.

2

X^2-Test

X^2-Test [1]



Possible Questions

Do settlements tend to be situated on rather good soil or is the distribution random?
Conclusions about settlement behaviour and economy would be possible

Do older individuals have more shoe-last celt as grave goods than younger?
If shoe-last celt would be signs of social rank than this situation would make conclusions possible about heredity or acquisition of social rank during life time.

Tests for nominale scaled variables are possible!
Therefore of particular value for archaeology because we have often to deal with such data.

X^2-Test [2]

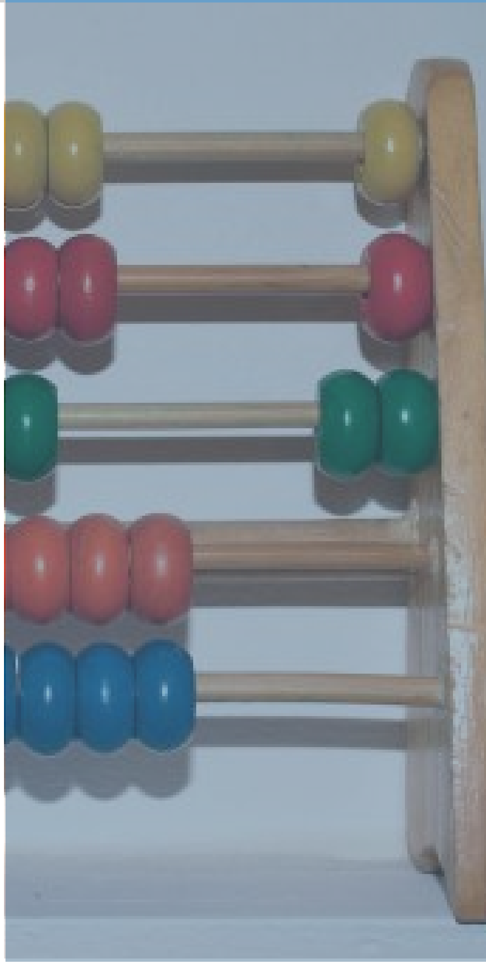
Test for independence of two distributions

Requirements: at least 1 nominal scaled variable (one sample case) and 1 nominal scaled grouping variable (two sample case)

Procedure with one sample: observed values are compared with expected values given a certain distribution, no expected value should be < 5; n should be > 50

Procedure with two samples: observed values of both distributions are compared with expected values if the samples would be even distributed, no expected value should be < 5; n should be > 50

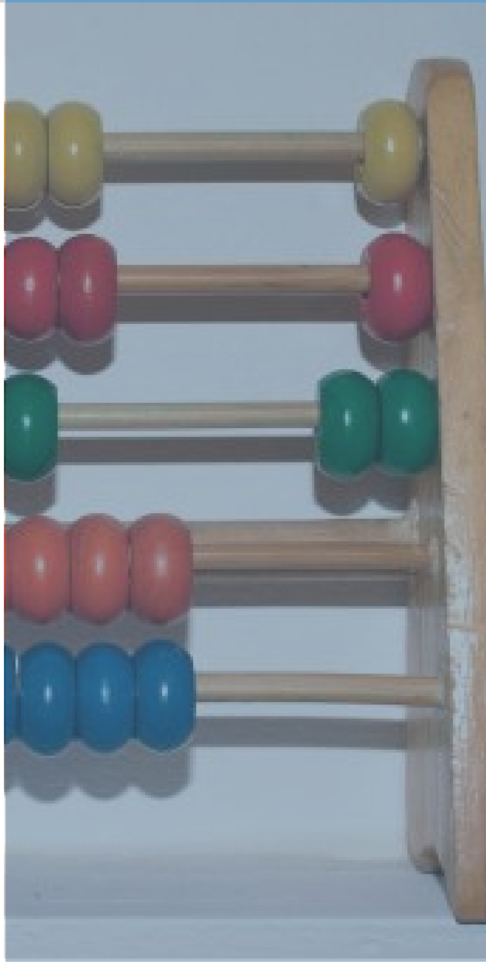
Test statistics: $\chi 2$
Significance depend on degree of freedom (df)



Excursus degree of freedom

Number of slots free to vary given the margin sums

	male	female	total
cremation	123		201
inhumation			197
total	216	182	398



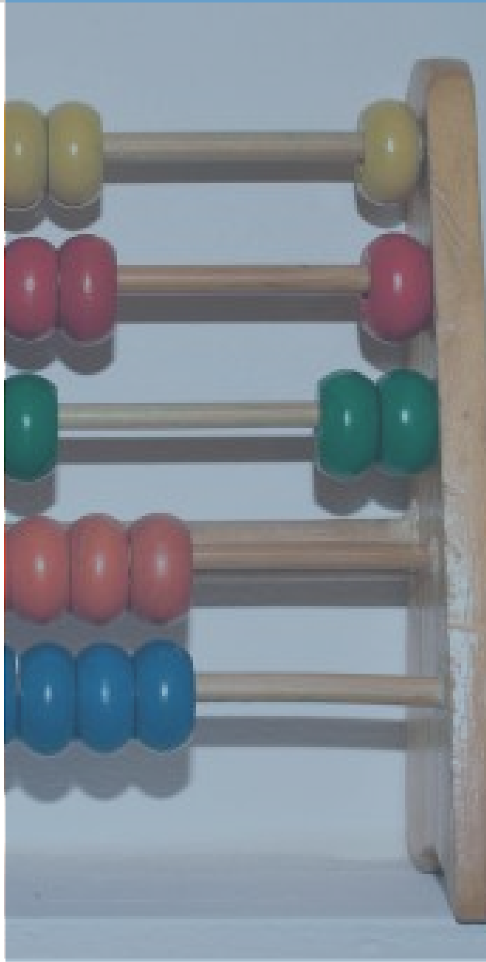
Excursus degree of freedom

Number of slots free to vary given the margin sums

	male	female	total
cremation	123	78	201
inhumation	93	104	197
total	216	182	398

df=1: if one value is chosen all other can be calculated with the help of the margins

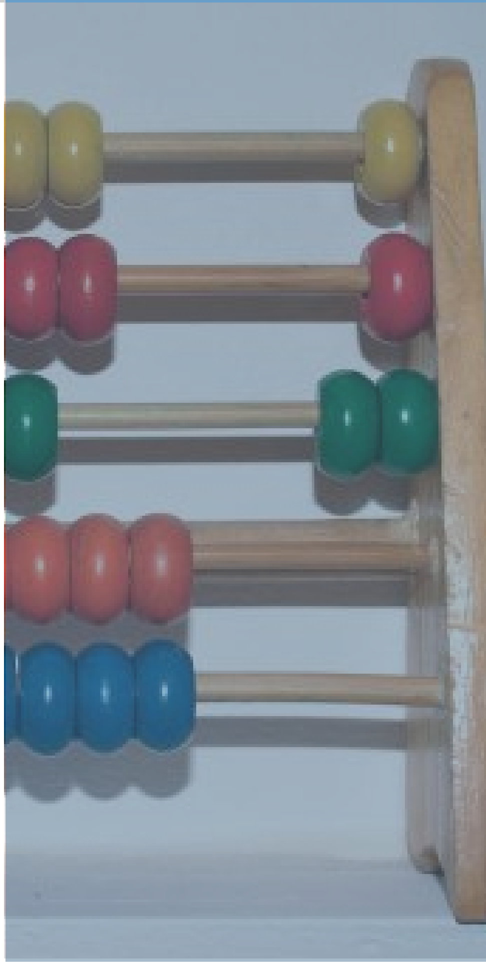
(number of columns – 1)*(number of rows – 1)



Excursus degree of freedom

Number of slots free to vary given the margin sums

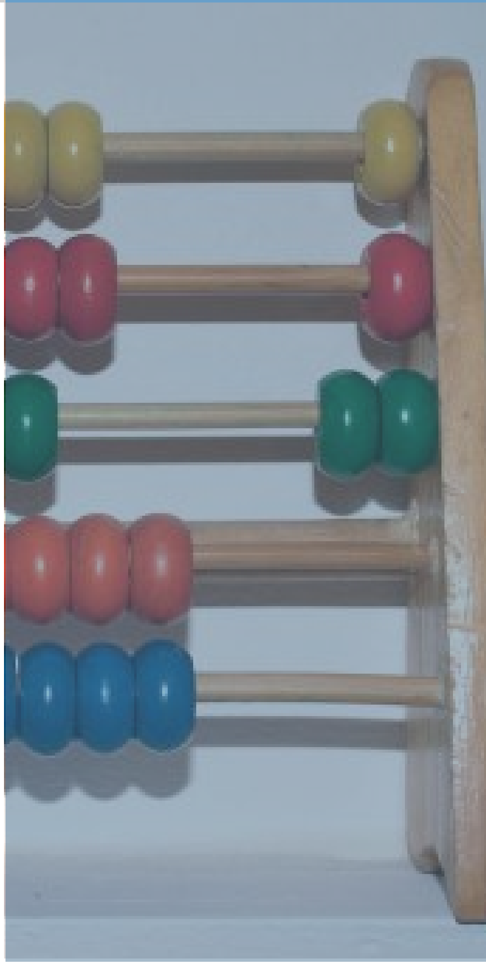
	male	female	uncertain	total
cremation		78		201
inhumation				197
total	196	179	23	398



Excursus degree of freedom

Number of slots free to vary given the margin sums

	male	female	uncertain	total
cremation	113	78		201
inhumation				197
total	196	179	23	398



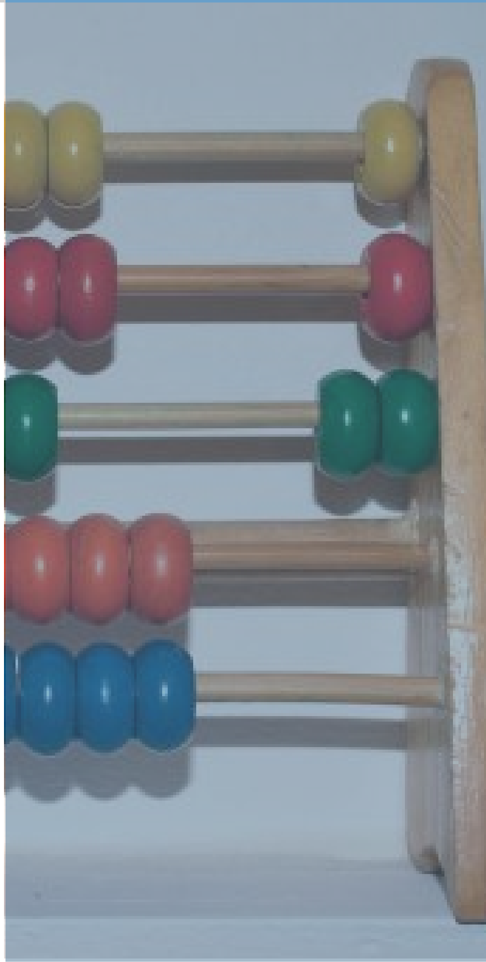
Excursus degree of freedom

Number of slots free to vary given the margin sums

	male	female	uncertain	total
cremation	113	78	10	201
inhumation	83	101	13	197
total	196	179	23	398

df=2: if two values are chosen all other can be calculated with the help of the margins

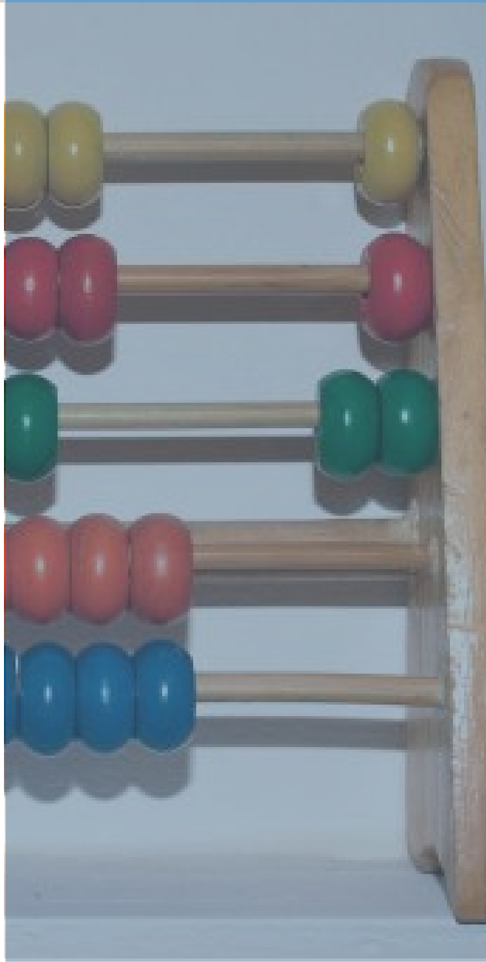
(number of columns – 1)*(number of rows – 1)



Excursus degree of freedom

Number of slots free to vary given the margin sums

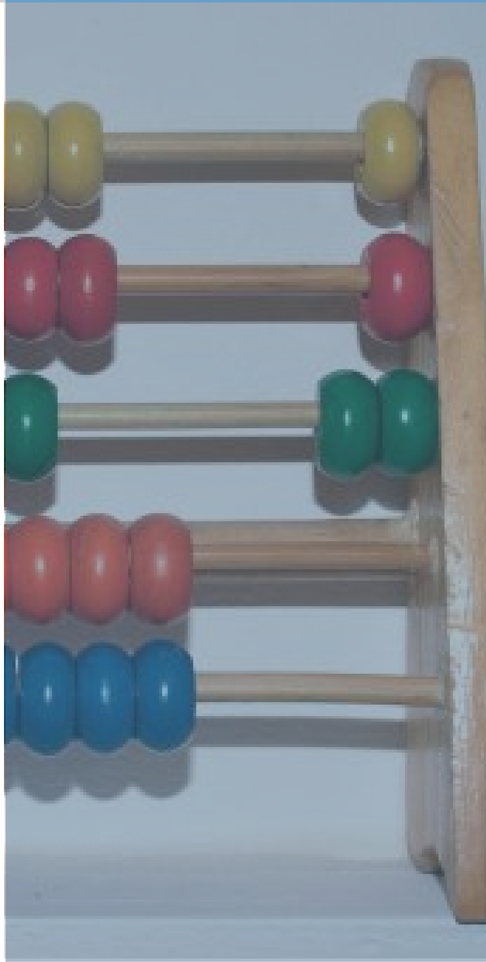
	male	female	uncertain	total
cremation				201
inhumation				197
uncertain				30
total	201	187	40	398



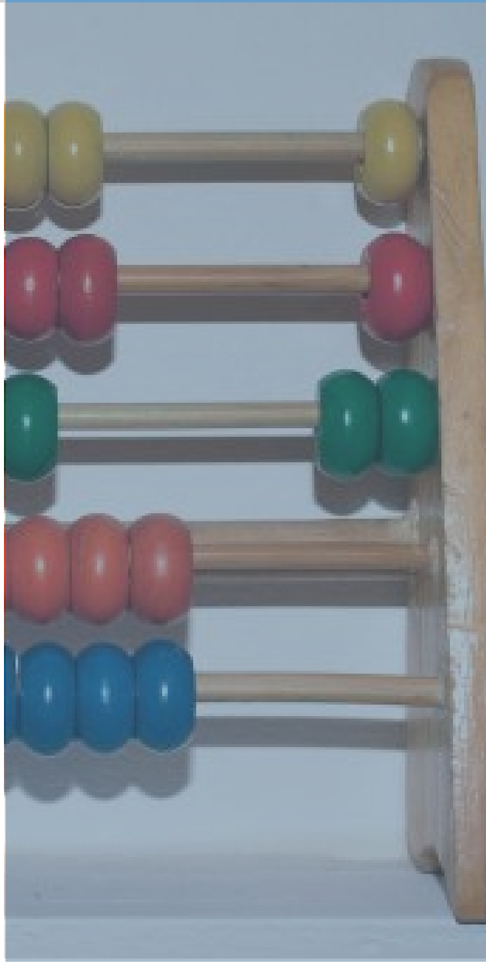
Exkurs Freiheitsgrade

Number of slots free to vary given the margin sums

	male	female	uncertain	total
cremation		78		201
inhumation	83		13	197
uncertain		8		30
total	201	187	40	398



Exkurs Freiheitsgrade



Number of slots free to vary given the margin sums

	male	female	uncertain	total
cremation	113	78	10	201
inhumation	83	101	13	197
uncertain	5	8	17	30
total	201	187	40	398

df=4: if four values are chosen all other can be calculated with the help of the margins

(number of columns – 1)*(number of rows – 1)

X^2-Test [3]

Test for one sample (example after Shennan)

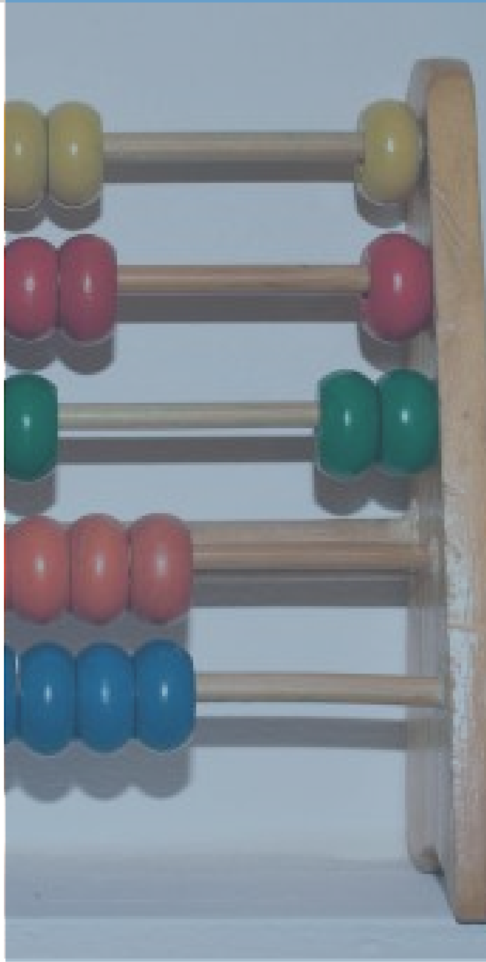
Numbers of neolithic settlements by soil type in eastern france

Soil type	Number of settlements
Rendzina	26
Alluvial	9
Brown earth	18
total	**53**

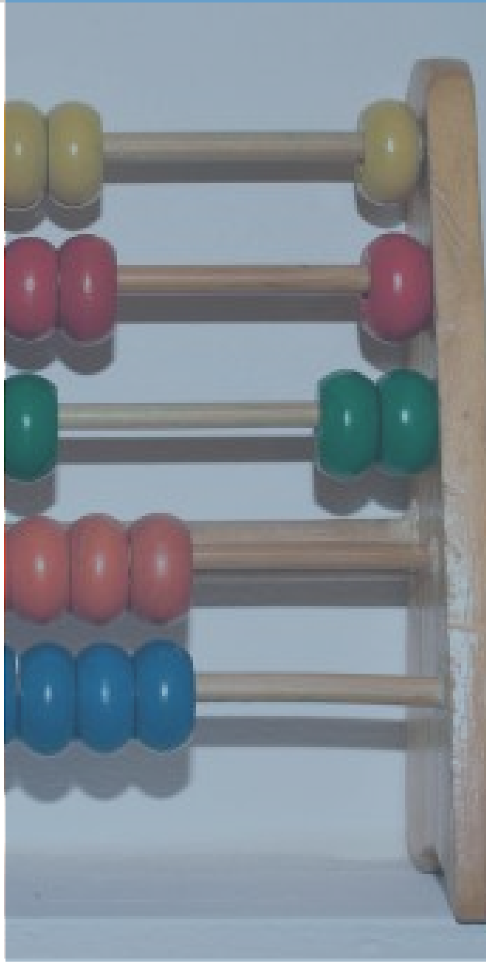
Question: Is there a significant preference for a soil type?
We calculate two versions:

1. even distributed

2. even distributed with consideration of the proportion of the soil types on the total area



X^2-Test [4]



Version 1: even distributed

Soil type	Number of settlements	Proportion of soil type	Expected number of settlements
Rendzina	26	1/3	17,66667
Alluvial	9	1/3	17,66667
Brown earth	18	1/3	17,66667
total	53	1	53

1. even distributed

H_0: The settlements are evenly distributed on all soil types.

H_1: The settlements are **not** evenly distributed on all soil types.

X²-Test [5]

Version 1: even distributed

Soil type	Number of settlements	Proportion of soil type	Expected number of settlements
Rendzina	26	1/3	17,66667
Alluvial	9	1/3	17,66667
Brown earth	18	1/3	17,66667
total	53	1	53

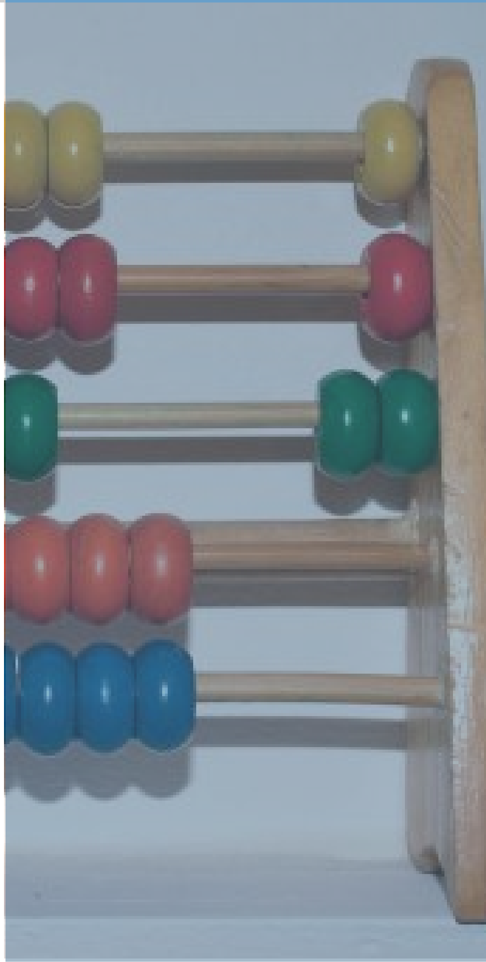
Formular for X²

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

$$O_i: number\ of\ observed\ cases$$

$$E_i: number\ of\ expected\ cases$$

$$\chi^2: symbol\ for\ the\ test\ statistic\ chi-squared$$



X²-Test [6]

Version 1: even distributed

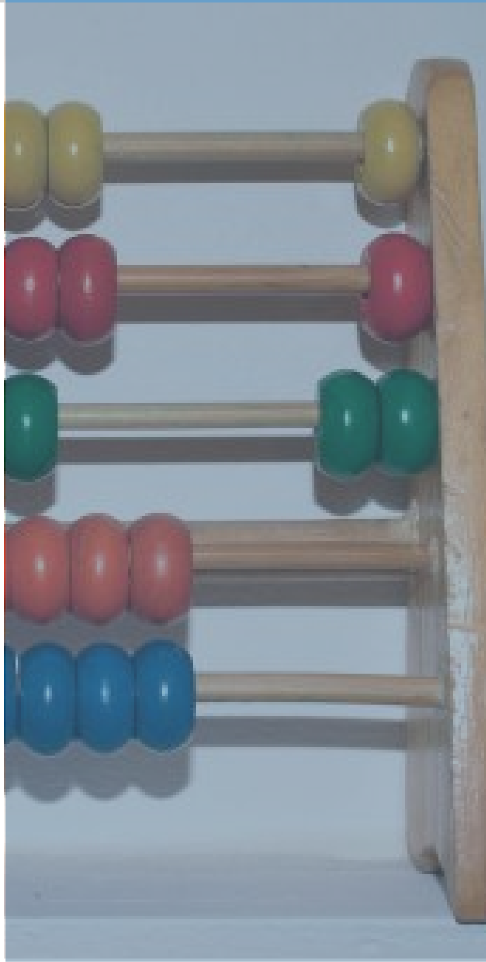
$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

Procedure: Calculation of the X²-value

Soil type	Observed number of settlements	Expected number of settlements	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
Brown earth	18	17,66667	0,33333	0,11111	0,00629
total	**53**	53			**8,18868**

Look up in a table (e.g. Shennan):
Df=2 (2 colums (expected, observed), 3 categories)
Level of significance: 0,05
Boundary value: 5,99145

Significant result: The distribution is uneven!



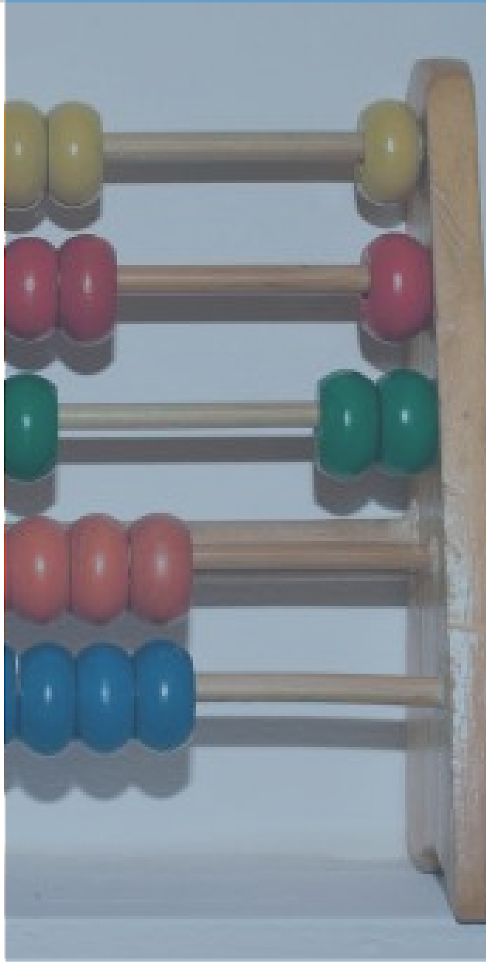
X^2-Test [7]

Version 2: even distributed with consideration of the proportion of the soil types on the total area

Soil type	Number of settlements	Proportion of soil type	Expected number of settlements
Rendzina	26	32%	16,96
Alluvial	9	25%	13,25
Brown earth	18	34%	22,79
Gesamt	**53**	1	53

Formular for X^2

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



X^2-Test [8]

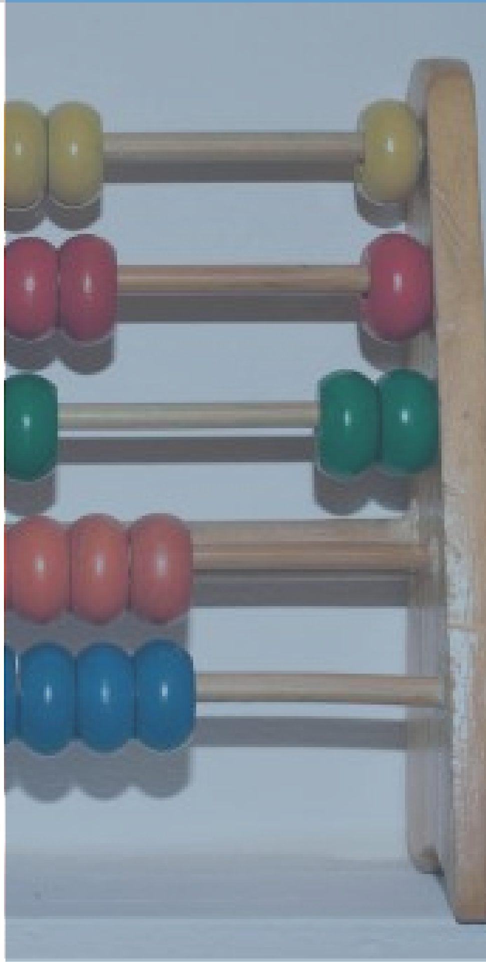
Version 2: even distributed with consideration of the proportion of the soil types on the total area

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

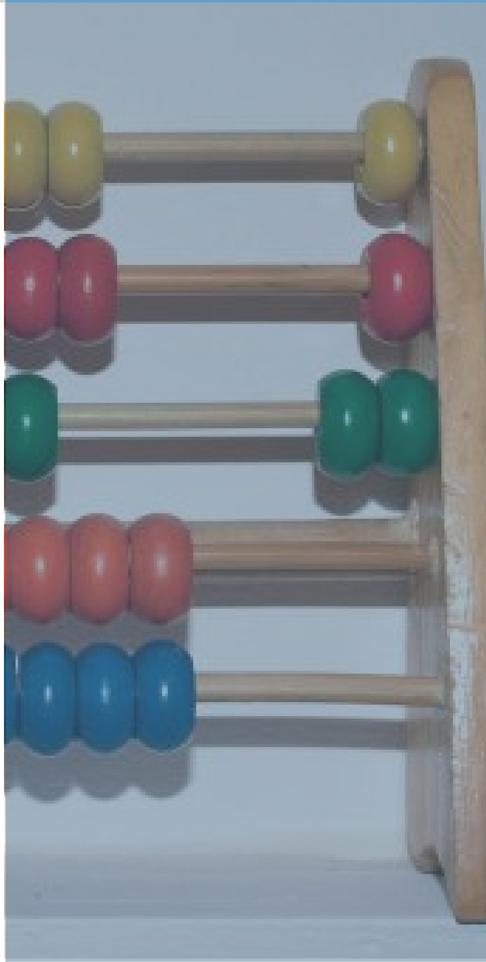
Soil type	Number of settlements	Expected number of settlements	O_i-E_i	$(O_i-E_i)^2$	$(O_i-E_i)^2/E_i$
Rendzina	26	16,96	9,04	81,7216	4,81849
Alluvial	9	13,25	-4,25	18,0625	1,36321
Brown earth	18	22,79	-4,79	22,9441	1,00676
total	**53**	53			**7,18846**

Look up in a table (e.g. Shennan):
Df=2 (2 colums (expected, observed), 3 categories)
Level of significance: 0,05
Boundary value: 5,99145

Significant result: The distribution is uneven also if we consider the proportions of the soil types!



X^2-Test [9]



In R

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

Version 1: even distributed

```
> chisq.test(siedlungen)

     Chi-squared test for given probabilities

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

Version 2: even distributed with consideration of the proportion of the soil types on the total area

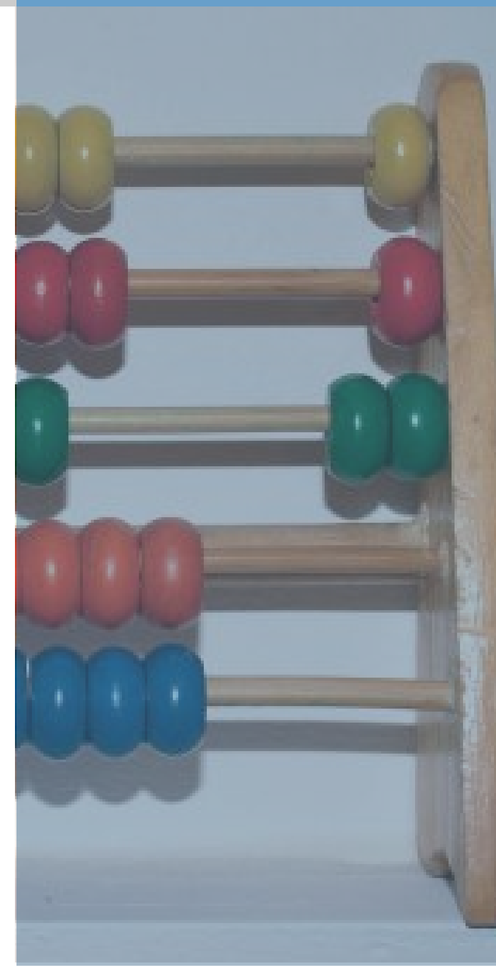
```
> chisq.test(siedlungen,p=c(0.32,0.25,0.43))

     Chi-squared test for given probabilities

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

X^2-Test [10]

Two sample case (Test for independence)
(example after Hinz, beautified)
Comparison of amber in graves and settlements
Classic 2x2 situation

Type of site	amber		total
	+	-	
settlement	6	18	24
grave	132	44	176
total	**138**	62	200

Is amber primary a grave good?
df=1
Level of significance = 0.05

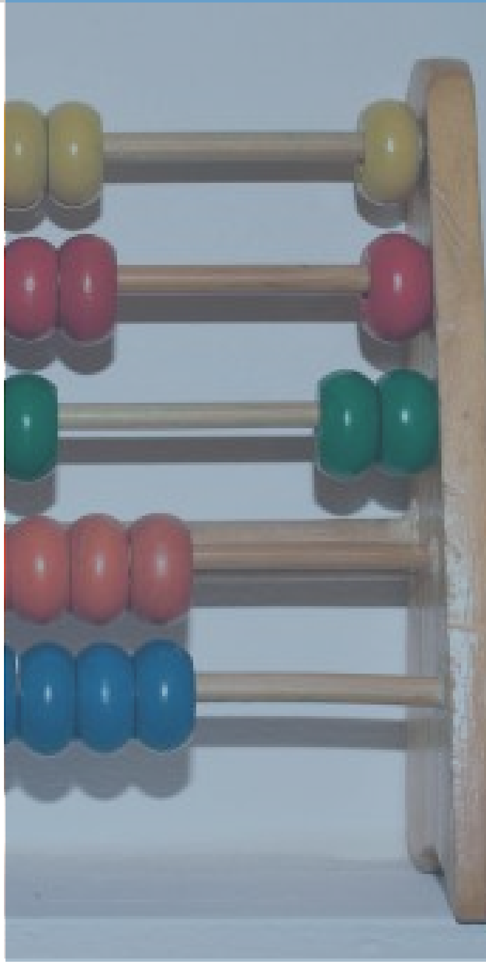


Christian-Albrechts-Universität zu Kiel

X²-Test [11]

Procedure: Calculation of the expected values
Multiply the margins and divide the result by the total number

Type of site	amber		total
	+	-	
settlement	6 E=24*138/200 =16,56	18 E=24*62/200 =7,44	24
grave	132 E=138*176/20 0 =121,44	44 E=62*176/200 =54,56	176
total	**138**	62	200



X^2-Test [12]

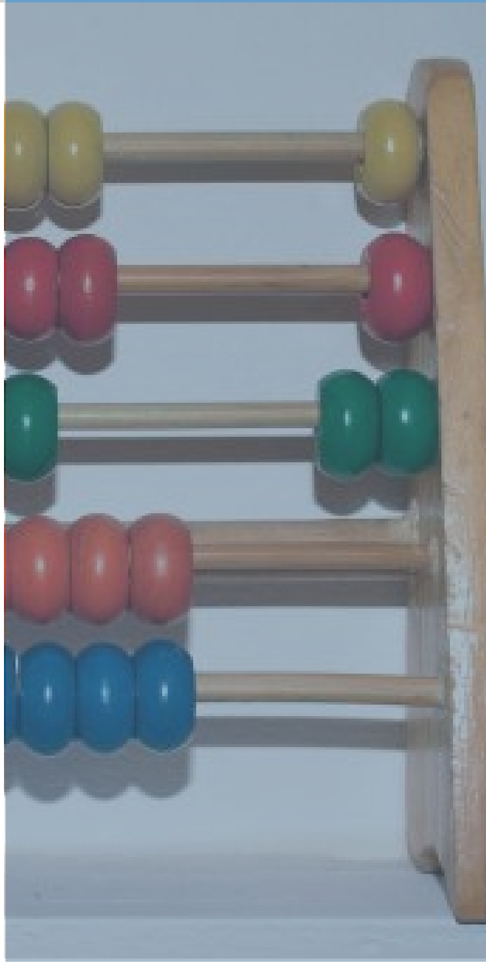
Procedure: Calculation of the X^2-value
(observed/expected)²/expected

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

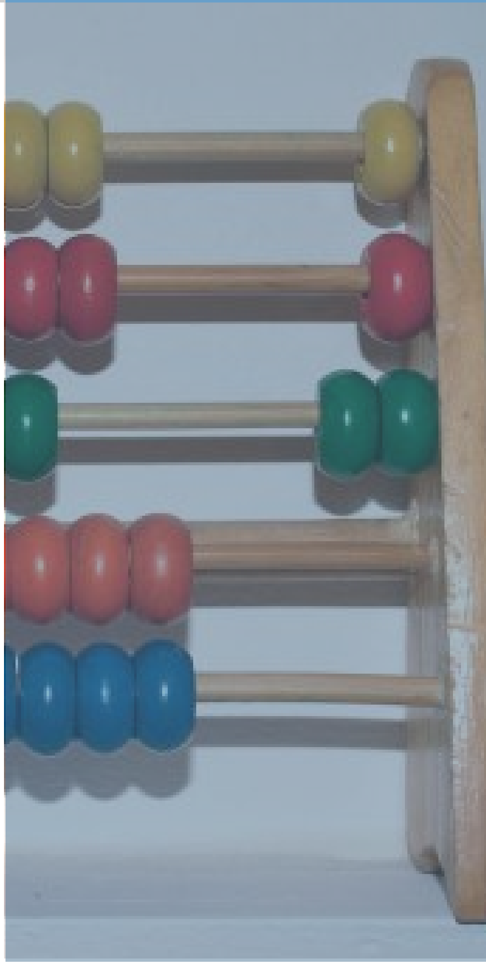
Type of site	amber		total
	+	-	
settlement	(6-16,56)²/16,56 =6,73	(18-7,44)²/7,44 =14,99	24
grave	(132-121,44)²/121,44 =0,92	(44-54,56)²/54,56 =2,04	176
total	**138**	62	200

Is amber primary a grave good?
Df=1, Level of significance = 0.05;
X^2=24,68; boundary value (df=1 and p=0.05): 3,84146

The difference in the distribution is significant not by chance. Both variables are associated!



X^2-Test [13]



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
```

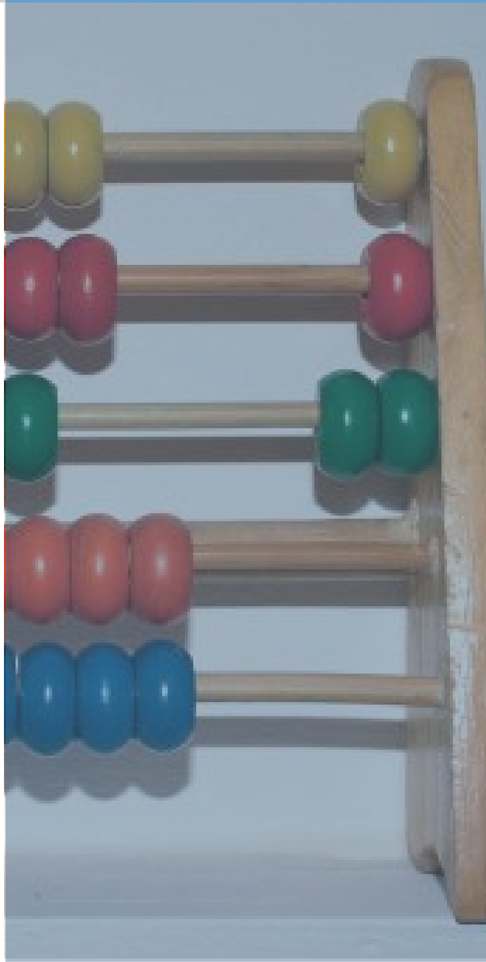
Correct: Yates correction for small samples → $(|O-E|-\mathbf{0{,}5})^2/E$

X²-Test Exercise

Animal bones from middle and late neolithic strata in Wolkenwehe (Mischka et al. 2005)
The following counts are given

layer	Domestic animal	Wild animal
202 (late neolithic)	159	32
203 (middle neolithic)	84	54

Analyse if the observed differences are statistically significant!

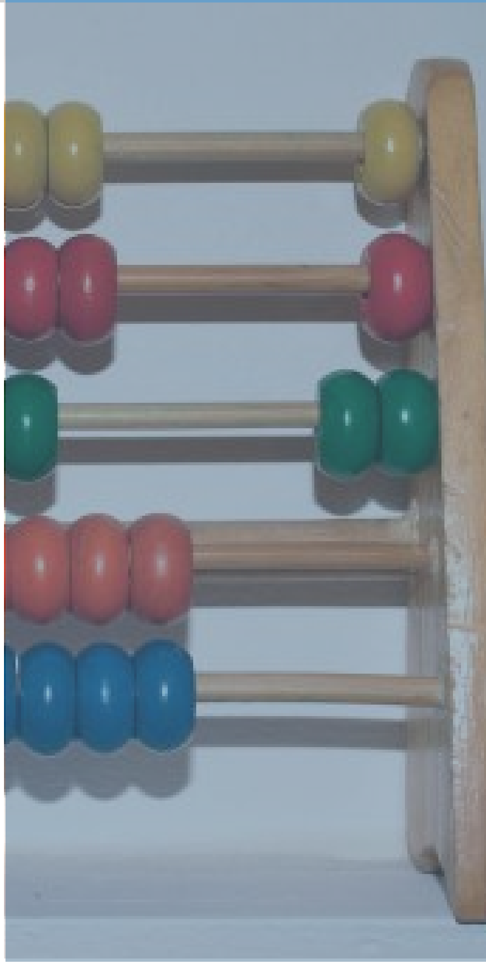


X²-Test Aufgabe

Animal bones from middle and late neolithic strata in Wolkenwehe (Mischka et al. 2005)

layer	Domestic animal	Wild animal
202 (late neolithic)	159	32
203 (middle neolithic)	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
```



Measurement of association [1]

Measurement of the strength of the association of two variables

X^2 is already a Measurment of association:

Association ↑ X^2 ↑ ↔ Association ↓ X^2 ↓

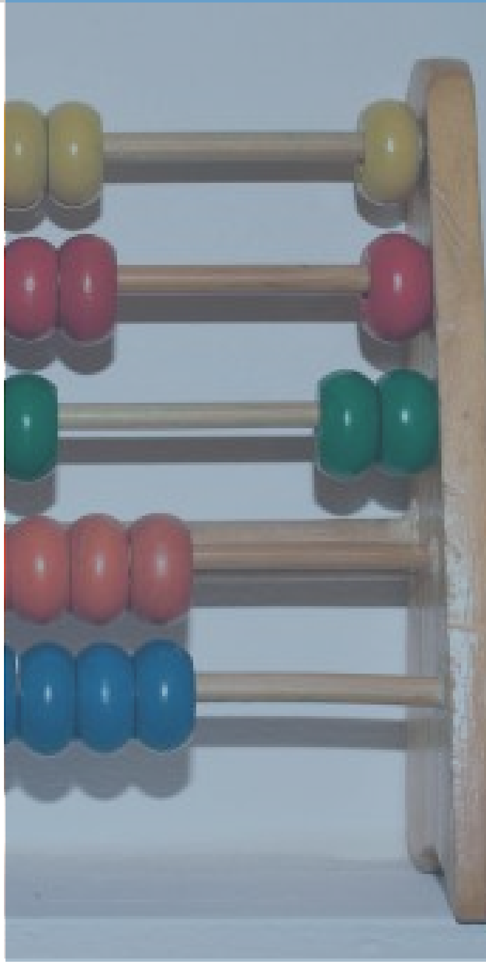
But: X^2 depends on n

Type of site	amber		total
	+	-	
settlement	6	18	24
grave	132	44	176
total	138	62	200

$X^2 = 24.6844$

Type of site	amber		total
	+	-	
settlement	12	36	48
grave	264	88	352
total	276	124	400

$X^2 = 49.3689$



Cramers V
(or φ)

Measurement of association [2]

Cramers V

Normalise X² for the number of observations n,
Square root,
Divide by the smaller value of (number of rows, number of colums) -1

Type of site	amber		total
	+	-	
settlement	6	18	24
grave	132	44	176
total	138	62	200

X² = 24.6844

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

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

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

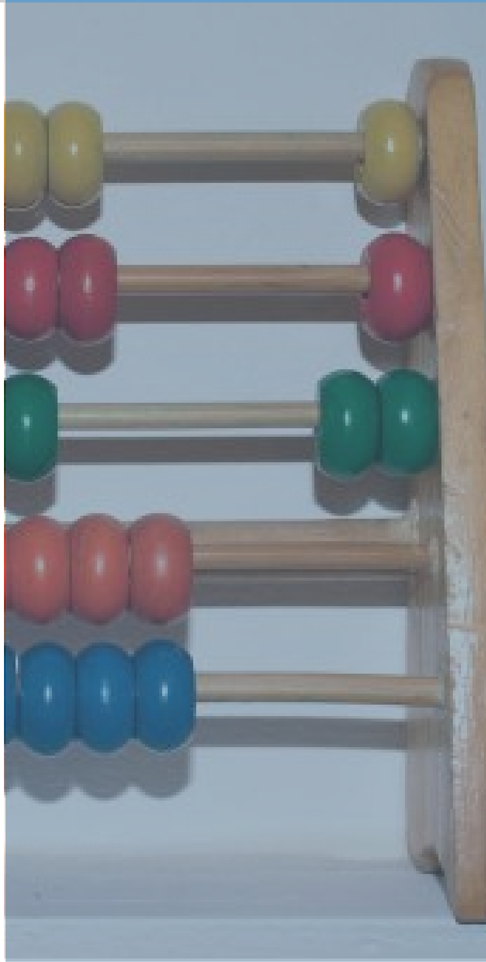
Type of site	amber		total
	+	-	
settlement	12	36	48
grave	264	88	352
total	276	124	400

X² = 49.3689

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

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

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



Measurement of association [3]

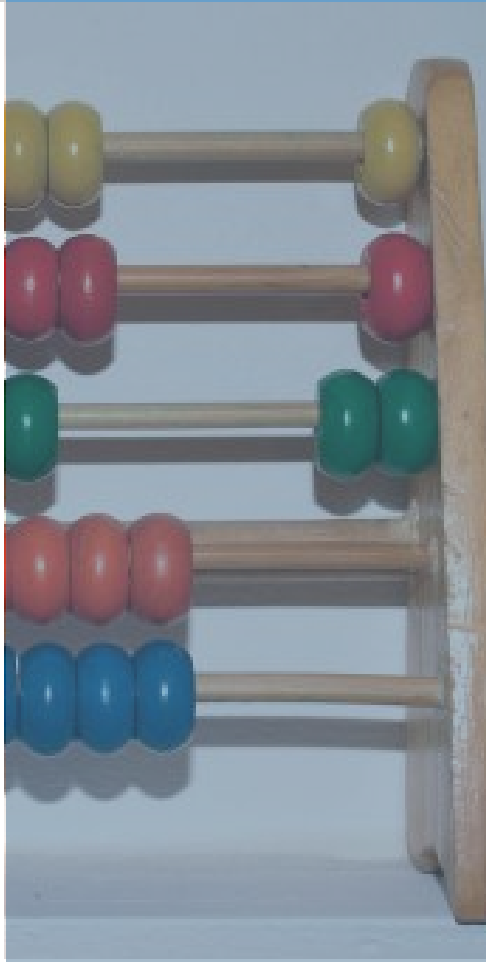
Cramers V

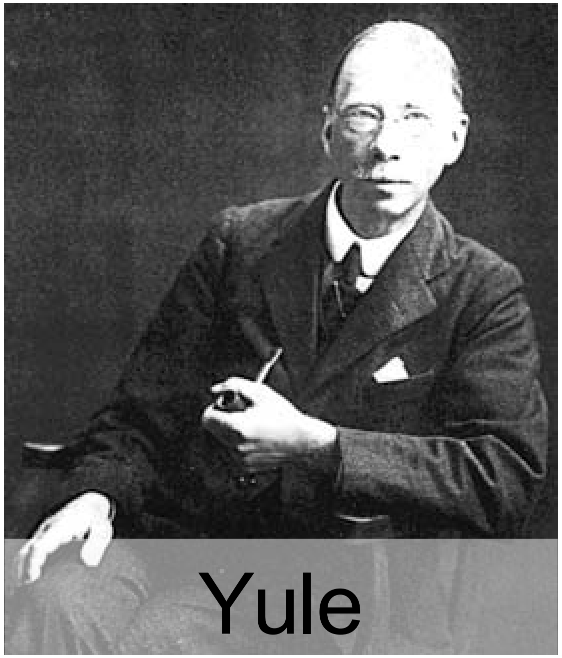
$$\phi=\sqrt{\frac{\chi^2}{n*(min(rows, colums)-1)}}$$

The value is between 0 and 1
0: no association
1: perfect association

In R:

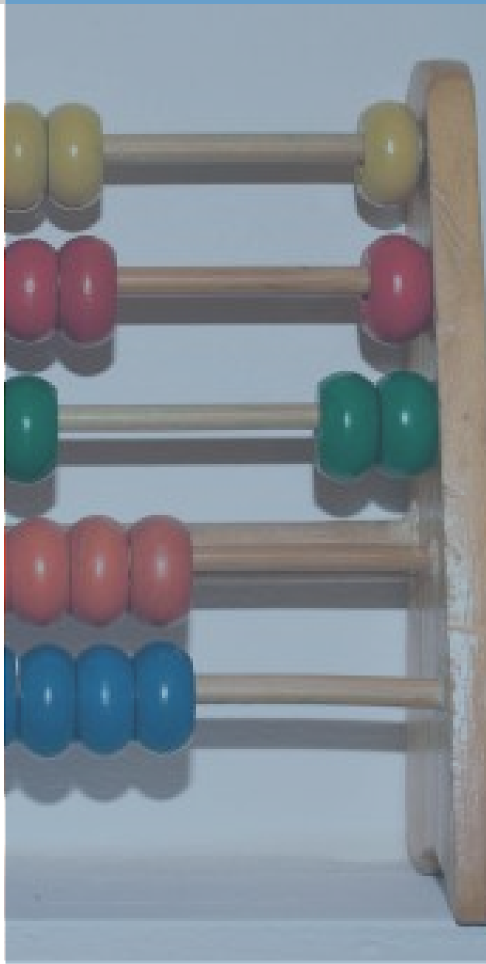
```
calc.CV <- function(x)
{
  CV <- sqrt(chisq.test(x, correct = FALSE)$statistic /
            (sum(x) * min(dim(x) - 1)))
  as.numeric(CV)
}
> calc.CV(test)
[1] 0.3513149
```





Yules Q

Measurement of association [4]



Yule's Q

Another simple measurement of association, only in the 2x2 case applicable

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

Idea: the bigger the number in the left upper field in relation to the total number the higher is the positive association

Type of site	amber		total
	+	-	
settlement	6	18	24
Grave (=no settlement)	132	44	176
total	138	62	200

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

Strong negative association: graves (not settlements) have a bigger possibility to contain amber finds.
But: Yules Q is not suitable for tables with a zero in one field

Measurement of association [5]

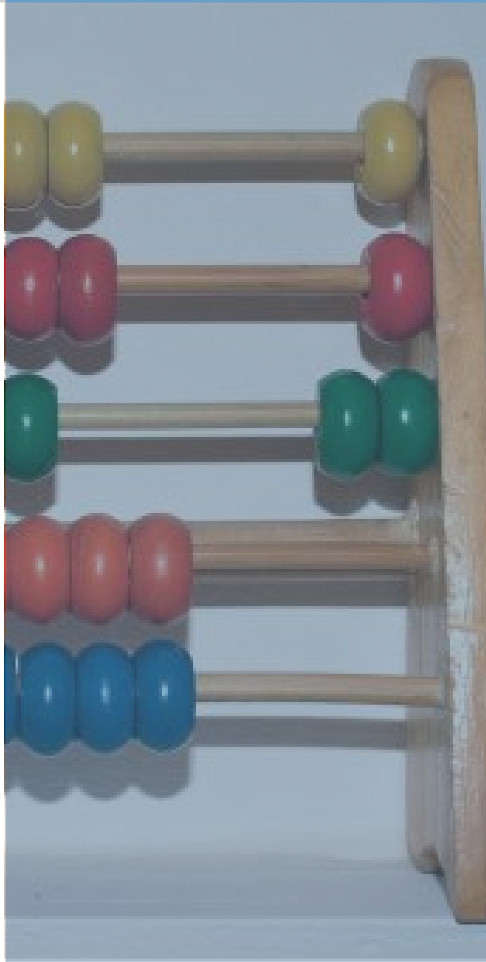
Yules Q

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

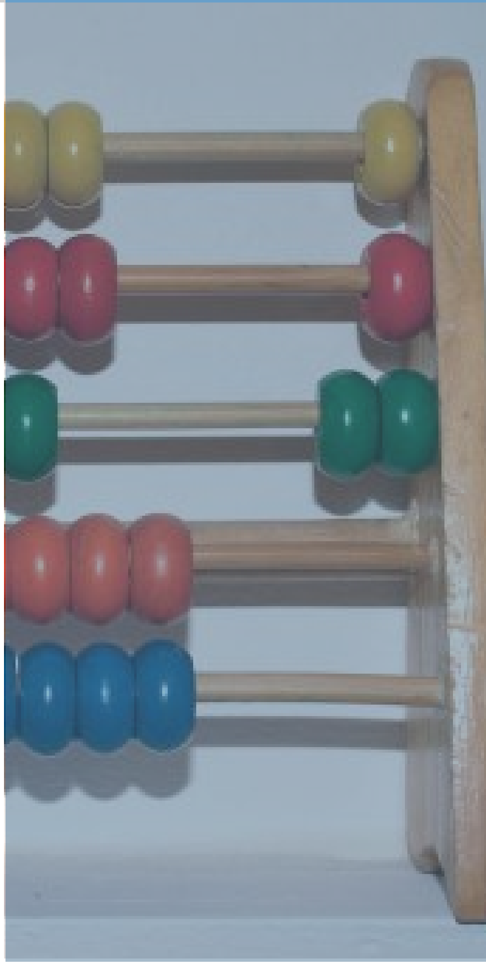
The value is between -1 and 1
-1: perfect negative association
0: no association
1: perfect positive association

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
```



Measurement of association Exercise



Animal bones from middle and late neolithic strata in Wolkenwehe (Mischka et al. 2005)
The following counts are given

layer	Domestic animal	Wild animal
202 (late neolithic)	159	32
203 (middle neolithic)	84	54

Analyse how strong the association is!

Measurement of association Exercise

Animal bones from middle and late neolithic strata in Wolkenwehe (Mischka et al. 2005)

The following counts are given

layer	Domestic animal	Wild animal
202 (late neolithic)	159	32
203 (middle neolithic)	84	54

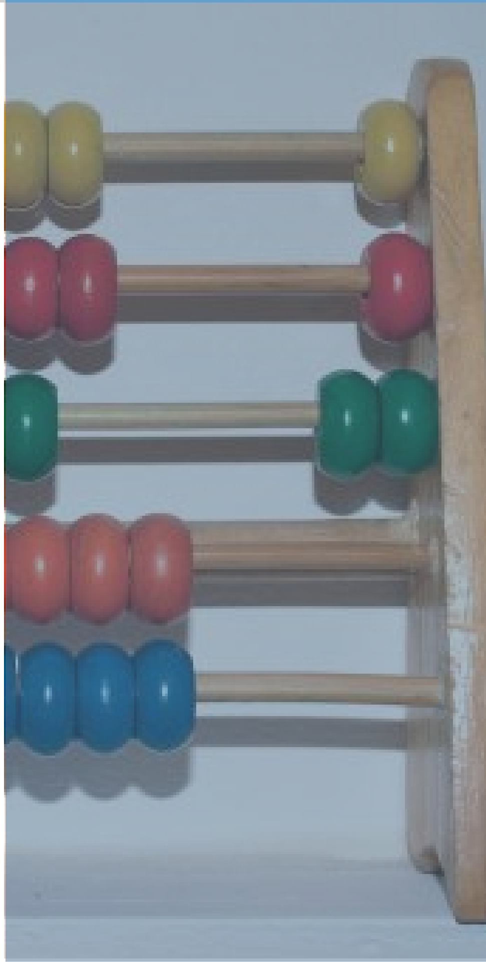
Analyse how strong the association is!

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

Cramers V is 0,25 for the association of domestic animals with late and wild animals with middle Neolithic layers

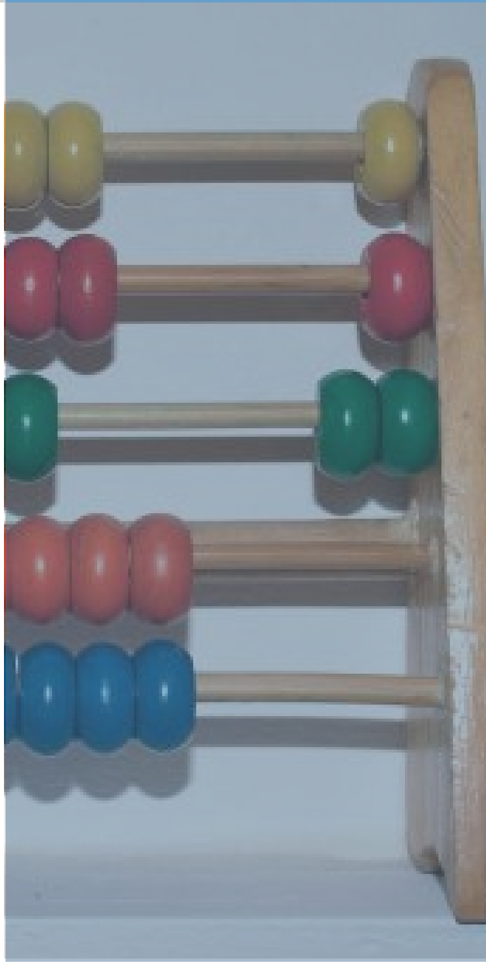
```
> calc.YQ(test)
[1] 0.5231506
```

Yules Q is 0,52 for positive association of domestic animals and late neolithic layers



C | A | U
Christian-Albrechts-Universität zu Kiel

Fishers Test [1]



Problem with to low expected values

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

Smaller than 5!

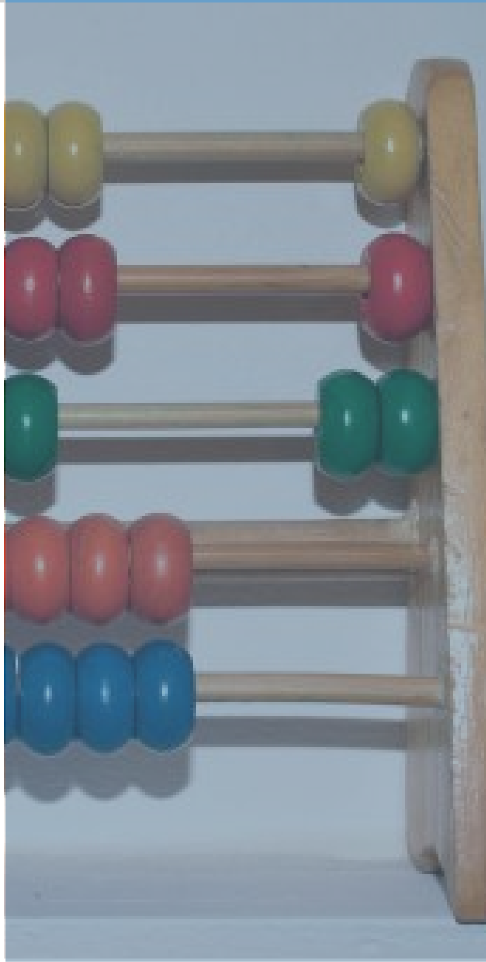
Fishers Test [2]

Test for two samples (test for independence) (example after Hinz, orginal)

Exact test after Fisher!

Type of site	amber		total
	+	-	
settlement	a: 3	b: 9	12
grave	c: 66	d: 22	88
total	69	31	n: 100

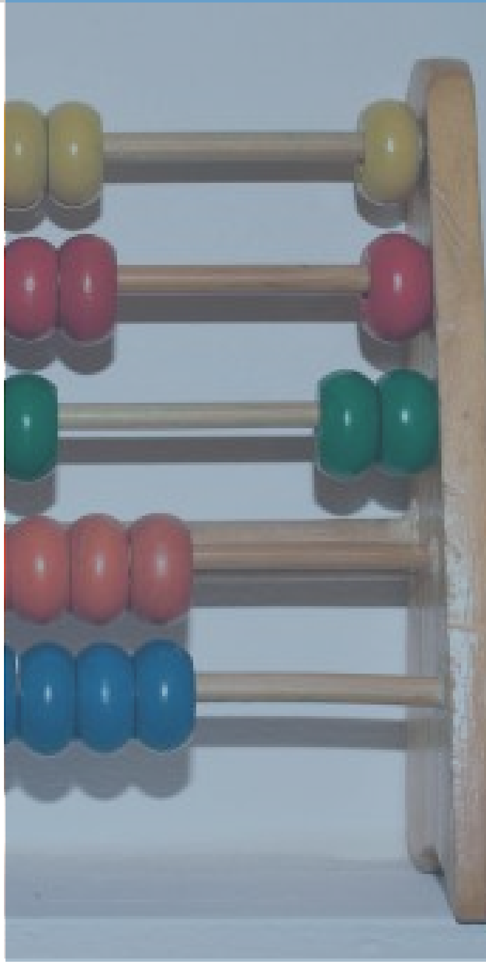
$$\varphi = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!} = \frac{(3+9)!(66+22)!(3+66)!(9+22)!}{100!3!9!66!22!}$$



Fishers Test [3]

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
```

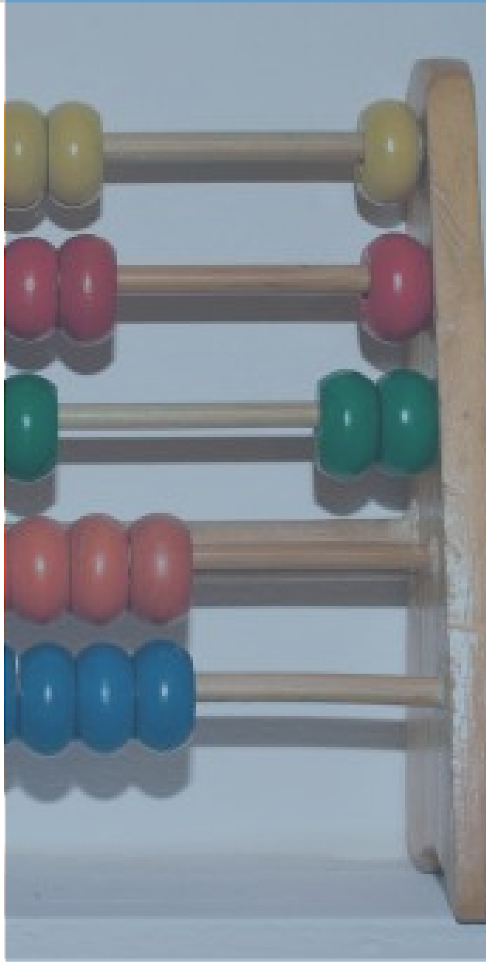


Fishers Test Aufgabe

Boar teeth in globular amphora graves (Müller 2001, numbers changed)
Given are the following numbers

sex	Boar teeth	
	yes	no
male	11	7
female	1	6

Analyse if there is a significant association!



Christian-Albrechts-Universität zu Kiel

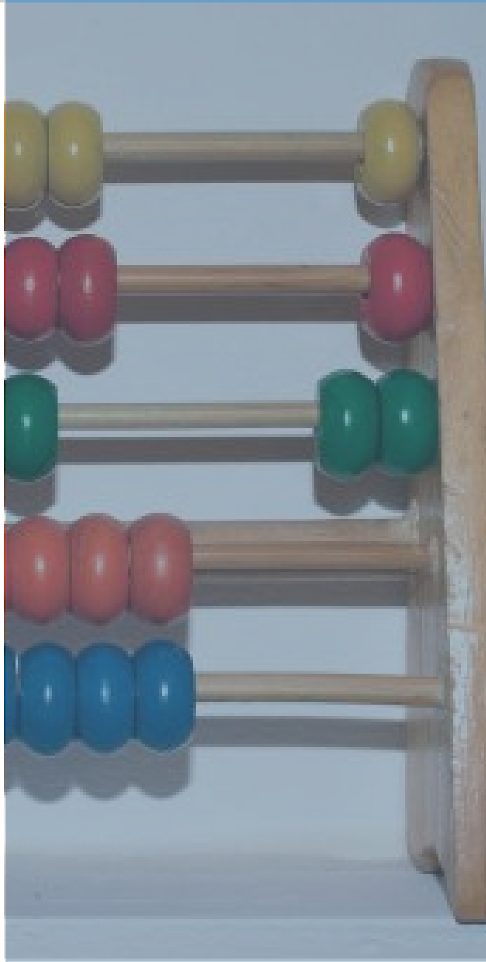
Fishers Test Aufgabe

Boar teeth in globular amphora graves (Müller 2001, numbers changed)

Given are the following numbers
Analyse if there is a significant association!

sex	Boar teeth	
	yes	no
male	11	7
female	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 of Tests

Statistical association not mean causal association!

Example after Shennan: Grave size and sex

Although there is a statistically significant association between grave size and sex this could be caused by a third factor (here height)

A conclusion which says that grave size are determined by sex would be wrong!