

OpenEpi

OpenEpi ist

- ein sehr nützliches, kleines Programm für statistische und epidemiologische Anwendungen,
- einfach zu bedienen und leicht zu erlernen,
- für aggregierte Daten gedacht, eine Eingabe von Einzeldaten ist nicht möglich,
- lauffähig auf Windows und Mac mit und ohne Internetverbindung, auch mit älteren Betriebssystemen,
- ausgerüstet mit wesentlichen Methoden der deskriptiven und induktiven Statistik.

Herunterladen von folgender Webseite:

<http://www.openepi.com/Downloads/Downloads.htm>

Folgen Sie den Anweisungen auf dem Bildschirm:

Ordner OpenEpi unter Dokumente anlegen, **OpenEpi.zip** herunterladen und in den Ordner hineinkopieren, dort entpacken, Datei Index.htm als Icon an Desktop senden, dort evtl. umbenennen, Doppelklick startet das Programm. Internet ist nicht erforderlich.

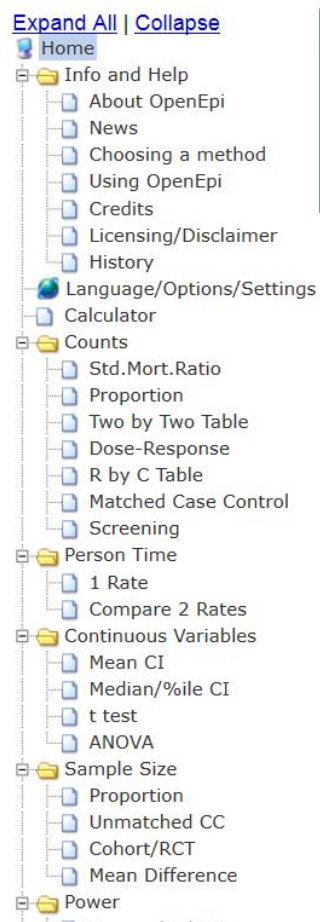
OpenEpi provides statistics for

- counts and measurements in descriptive and analytic studies,
- stratified analysis with exact confidence limits,
- matched pair and person-time analysis,
- sample size and power calculations,
- random numbers,
- sensitivity, specificity and other evaluation statistics,
- R x C tables, chi-square for dose-response

It is always a good idea to check **important results** with software from more than one source.

OpenEpi Startbildschirm

In dieser Lektion sollen Sie nur das Programm ein wenig kennenlernen. Dazu rufen Sie das entsprechende Teilprogramm auf und geben die Daten von den Beispielfolien ein. Nun sehen Sie, was passiert und wie das Programm arbeitet. Unter „Documentation“ erhält man eine kleine statistische Anleitung im pdf-Format, die man sich separat speichern kann.



Open Source Epidemiologic Statistics for Public Health

Now in English, French, Spanish, Italian, and Portuguese

Version 3.01 Updated 2013/04/06 Try it in a Smartphone browser!



OpenEpi provides statistics for counts and measurements in descriptive and analytic studies, stratified analysis with exact confidence limits, matched pair and person-time analysis, sample size and power calculations, random numbers, sensitivity, specificity and other evaluation statistics, R x C tables, chi-square for dose-response, and links to other useful sites.

OpenEpi is free and **open source** software for epidemiologic statistics. It can be run from a web server or downloaded and run without a web connection. A server is not required. The programs are written in JavaScript and HTML, and should be compatible with recent Linux, Mac, and PC browsers, regardless of operating system. (If you are seeing this, your browser settings are allowing JavaScript.) The programs can be run in the browsers of many iPhone and Android cellphones

Test results are provided for each module so that you can judge reliability, although it is always a good idea to check important results with software from more than one source. Links to hundreds of Internet calculators are provided.

The programs have an open source license and can be downloaded, distributed, or translated. Some of the components from other sources have licensing statements in the source code files. Licenses referred to are available in full text at OpenSource.org/licenses. OpenEpi development was supported in part by a grant from the Bill and Melinda Gates Foundation to Emory University, Rollins School of Public Health.

A toolkit for creating new modules and for translation is included. Please let us know if you would like to collaborate in this way. Suggestions, comments, and expressions of interest in contributing to this effort should be sent by email to: andy.dean@gmail.com, cdckms@sph.emory.edu, and msoe@cdc.gov

Suggested citation: Dean AG, Sullivan KM, Soe MM. OpenEpi: Open Source Epidemiologic Statistics for Public Health, Version. www.OpenEpi.com, updated 2013/04/06, accessed 2017/09/02.

OpenEpi Settings

Voreinstellung des Konfidenz-Niveaus ist standardmäßig auf 95%. Änderungen (z.B. 90%) bleiben nicht immer erhalten und sollten für jede Rechnung wieder kontrolliert werden. Ein Klick auf „Menü“ legt die Änderung für die nächste Rechnung fest. Beim Layout belassen wir es bei „Epi Info“, links (Zeilen) Exposure und oben (Spalten) Disease.

Menu

OpenEpi Settings

Language

Language

Output Settings

Confidence Level

- Do Not highlight 'significant' p values
- Show column percents
- Show row percents
- Show input datatables
- Show stratum results

Settings for 2 x 2 Input Table

	Disease	(+) on left
(+) at top	(+)	(-)
(+) Exposure		
(-)		

Not all applications use the table layout settings.

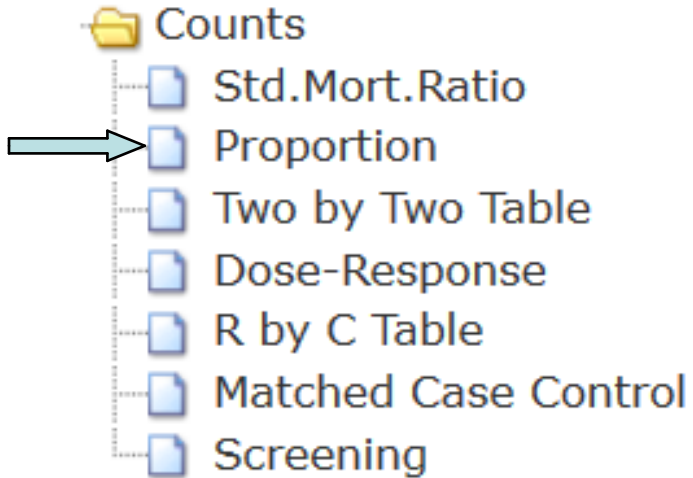
Layouts

- Epi Info; Schlesselman; Lilienfeld; Friis
- Kleinbaum; Breslow/Day
- Custom

Screen

- Laptop/Desktop
- Smartphone
- Auto

Berechnungen mit Zähldaten - Eingabe



Beispiel **Proportion**:

Zahngesundheit von Kindern (rechtes Bild unten)

Population $N=8000$, Stichprobe $n=2707$

Zahngesunde in der Stichprobe $k=1677$

Aufgabe:

Schätzen Sie den Anteil zahngesunder Kinder in der Population mit Konfidenzintervall.

Klicken Sie auf „Proportion“, dann auf „Enter New Data“, dann geben Sie die Daten in die Maske ein und klicken auf „Calculate“

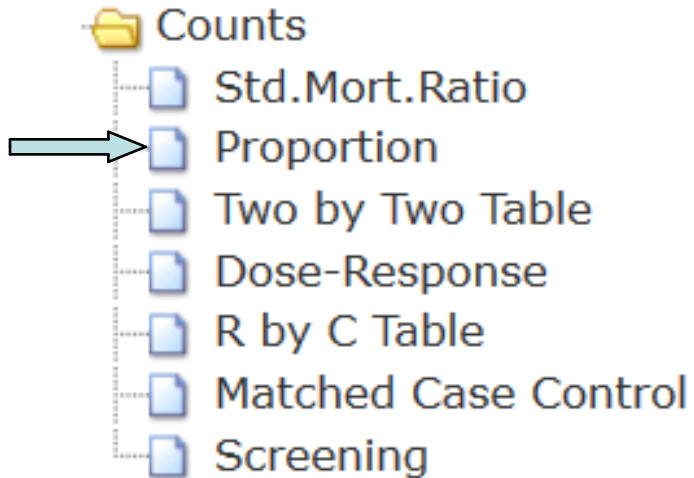
Clear [Settings](#) Conf. level=95% Calculate

Enter New Data

Simple Proportion			
Sample	Numerator	10	
	Denominator	11	
Multiply results by	100	eg, 100 for %	
Population size	1000000	if large, leave as 1000000	
Compare to %	50.0	for optional statistics	

Simple Proportion			
Sample	Numerator	1677	
	Denominator	2707	
Multiply results by	100	eg, 100 for %	
Population size	8000	if large, leave as 1000000	
Compare to %	50.0	for optional statistics	

Berechnungen mit Zähldaten - Ausgabe



Beispiel **Proportion**:

Zahngesundheit von Kindern

Population N=8000, Stichprobe n=2707

Zahngesunde in der Stichprobe k=1677

Aufgabe:

Schätzen Sie den Anteil zahngesunder Kinder in der Population mit Konfidenzintervall.

95% Confidence Limits for Proportion 1677/2707 Multiplier=100

Wilson Score corrected for population size

	Lower CL	Per 100	Upper CL
Proportion as Percent		61.9505	
Mid-P Exact	60.11		63.77
Fisher Exact(Clopper-Pearson)	60.09		63.78
Wald (Normal Approx.)	60.12		63.78
Modified Wald(Agresti-Coull)	60.11		63.76
Score(Wilson)*	60.11		63.76
Score with Finite Population Correction	60.45		63.42
Score with Continuity Correction (Fleiss Quadratic)	60.09		63.78

*LookFirst items: Editor's choice of items to examine first.

Der Anteil beträgt 61,95%.

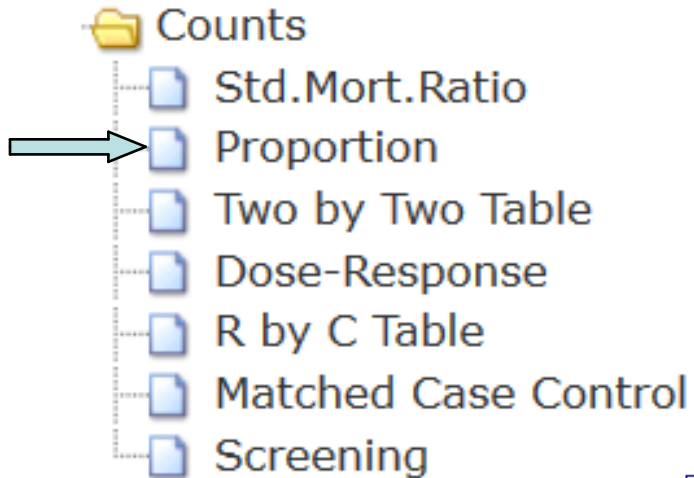
Konfidenzintervalle:

Wald: (60,12 ; 63,78)

Wilson: (60,11 ; 63,76)

Wilson mit Endlichkeitskorrektur: (60,45 ; 63,42)

Berechnungen mit Zähldaten - Ausgabe



Beispiel **Proportion**:

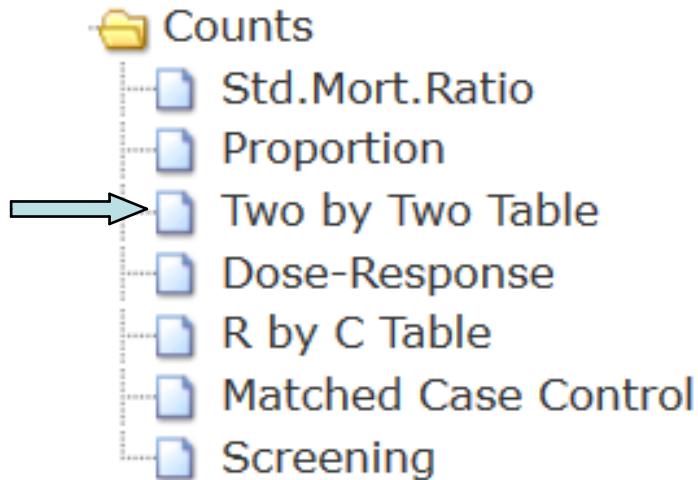
Wollen Sie die Werte korrigieren, drücken Sie „Enter“.

Wollen Sie von vorn beginnen, drücken Sie „Start“.

Allerdings kann man auch unter „Enter“ neue Daten eingeben.

Start	Enter	Results	Examples	Help
95% Confidence Limits for Proportion 1677/2707				
Multiplier=100				
Wilson Score corrected for population size				
		Lower CL	Per 100	Upper CL
Proportion as Percent			61.9505	
Mid-P Exact		60.11		63.77
Fisher Exact(Clopper-Pearson)		60.09		63.78
Wald (Normal Approx.)		60.12		63.78
Modified Wald(Agresti-Coull)		60.11		63.76
Score(Wilson)*		60.11		63.76
Score with Finite Population Correction		60.45		63.42
Score with Continuity				
Correction (Fleiss Quadratic)		60.09		63.78

Berechnungen mit der Vierfeldertafel - Eingabe



Aufgabe: Gibt es einen Zusammenhang zwischen Karies und Geschlecht (W, M)?

Two by Two anklicken,
Enter New Data,
Daten eingeben,
Calculate.

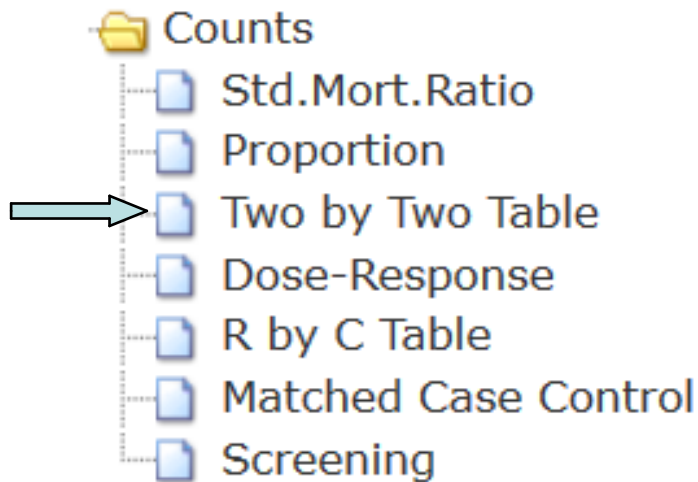
Statt „Calculate“ kann man auch auf „Results“ klicken.

Karies

	1	0	
W	27	14	41
M	26	12	38
	53	26	79

Open Epi 2 x 2 Table				
	Disease			Totals
Exposure		27	14	41
		26	12	38
Totals		53	26	79

Berechnungen mit der Vierfeldertafel - Ausgabe



Karies

	1	0	
W	27	14	41
M	26	12	38
	53	26	79

Aufgabe: Gibt es einen Zusammenhang zwischen Karies und Geschlecht (W, M)?

Chi Square and Exact Measures of Association

Es gibt keinen
Zusammenhang.
Chi²_{Yates} liefert
einen p-Wert von
p = 0,9976 (n.s.)

Test	Value	p-value(1-tail)	p-value(2-tail)
Uncorrected chi square	0.05887	0.4041	0.8083
Yates corrected chi square	0.000009199	0.4988	0.9976
Mantel-Haenszel chi square	0.05813	0.4047	0.8095
Fisher exact		0.4992(P)	0.9984
Mid-P exact		0.4073(P)	0.8146

Dose-Response Berechnungen - Eingabe

Jetzt kennen Sie schon das Prinzip:

- Analyse auswählen
- Enter New Data
- Daten eingeben
- Calculate oder Results

Je nach Analyse kommen noch
Zwischenabfragen.

Manche zusätzlichen Dialogboxen
dienen nur zur Information, mit OK
kommen Sie weiter.

5 Sozialschichten **(1977)**,
Cases=Kinder mit Karies,
Contr.=Kinder ohne Karies

How many EXPOSURE LEVEL CATEGORIES are there (including the baseline)?

OK

Abbrechen

Exposure Level	Cases	Controls	Total
0(Baseline)	123	31	154
1	135	24	159
2	455	50	505
3	400	22	422
4	243	9	252
Total	1356	136	1492

Dose-Response Berechnungen

5 Sozialschichten, Cases=Kinder mit Karies, Contr.=Kinder ohne Karies (1977)
Exposure Level 0 - obere Sozialschicht, 1 - obere Mittelschicht, 2 - Mittelschicht,
3 - untere Mittelschicht und Exposure Level 4 - untere Sozialschicht

Dose Response Analysis

Stratum 1						
Exposure Level	Cases	Controls	Total	Odds of Exp.	Odds Ratio	
0	123	31	154	3.97	1	
1	135	24	159	5.63	1.42	
2	455	50	505	9.1	2.29	
3	400	22	422	18.18	4.58	
4	243	9	252	27	6.8	
Total	1356	136	1492			

Extended Mantel-Haenszel chi square for linear trend=
p-value(1 degree of freedom)=

44.59
<0.0000001

Das Odds Ratio (Chancenverhältnis) für Zahnkaries erhöht sich 1977 stark vom Referenzwert 1,0 (Level 0) auf 6,8 (Level 4) mit steigendem Sozialindex (0 bis 4). Kinder der unteren Sozialschicht haben gegenüber Kindern der oberen Schicht eine 6,8-fache Chance Karies zu bekommen.

Dose-Response Berechnungen

5 Sozialschichten, Cases=Kinder mit Karies, Contr.=Kinder ohne Karies (1998)

Dose Response Analysis

Stratum 1					
Exposure Level	Cases	Controls	Total	Odds of Exp.	Odds Ratio
0	41	103	144	0.4	1
1	40	172	212	0.23	0.58
2	230	589	819	0.39	0.98
3	119	130	249	0.92	2.3
4	164	126	290	1.3	3.27

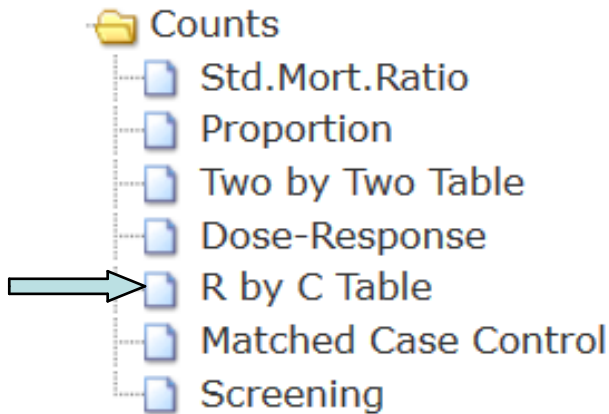
Total	594	1120	1714		
-------	-----	------	------	--	--

Extended Mantel-Haenszel chi square for linear trend=
p-value(1 degree of freedom)=

91.63
<0.0000001

Das Odds Ratio (Chancenverhältnis) für Zahnkaries erhöht sich 1998 nicht mehr so stark von 1,0 auf 3,27 mit steigendem Sozialindex (0 bis 4)

Berechnungen mit der R x C - Tafel



Zahnpflege

Alter	1	2	3	Gesamt
3	79	359	44	482
4	184	731	122	1037
5	205	855	128	1188
Gesamt	468	1945	294	2707

Chi Square for R by C Table

Zusammenhang
zwischen Zahnpflege
und Alter?

Kein Zusammenhang
 $p = 0,5245$

OpenEpi für RxC
Tafeln ist wenig
ergiebig.

Chi Square= 3.203
Degrees of Freedom= 4
p-value= 0.5245

Cochran recommends accepting the chi square if:

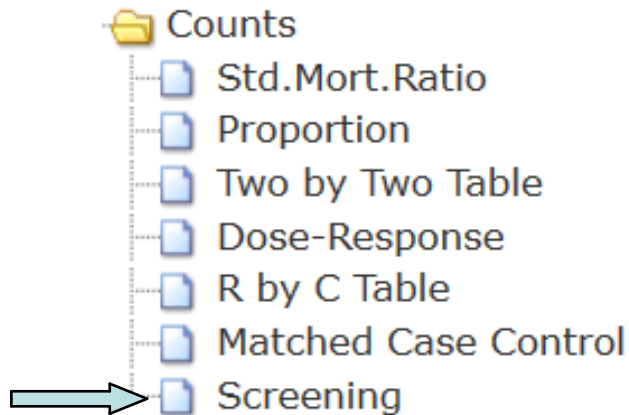
1. No more than 20% of cells have expected < 5 .
2. No cell has an expected value < 1 .

In this table:

None of 9 cells have expected values < 5 .
No cells have expected values < 1 .

Using these criteria, this chi square can be accepted.

Screening und diagnostischer Test

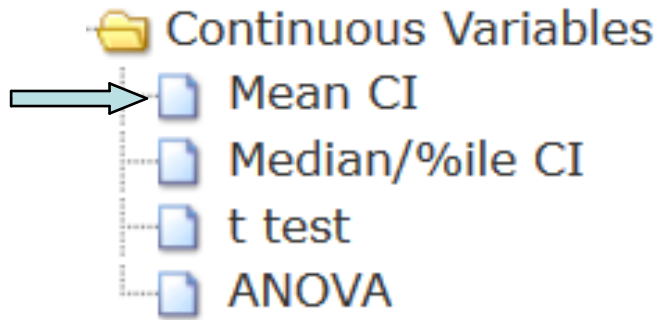


		Histologie	
		+	-
TEST	+	40	63
	-	60	837

Beurteilung von 1000 Zähnen auf vorhandene Karies im Rahmen einer Kalibrierung von Zahnärzten und Vergleich mit der bekannten feingeweblichen Untersuchung der Zähne (Histologie).

Parameter	Estimate	Lower - Upper 95% CIs	Method
Sensitivity	40%	(30.94, 49.8 ¹)	Wilson Score
Specificity	93%	(91.14, 94.49 ¹)	Wilson Score
Positive Predictive Value	38.83%	(29.99, 48.49 ¹)	Wilson Score
Negative Predictive Value	93.31%	(91.48, 94.77 ¹)	Wilson Score
Diagnostic Accuracy	87.7%	(85.52, 89.59 ¹)	Wilson Score
Likelihood ratio of a Positive Test	5.714	(5.147 - 6.344)	
Likelihood ratio of a Negative Test	0.6452	(0.6243 - 0.6667)	
Diagnostic Odds	8.857	(5.508 - 14.24)	
Cohen's kappa (Unweighted)	0.3257	(0.2637 - 0.3876)	

Konfidenzintervall für einen Mittelwert



Gegeben sei ein Mittelwert mit Standardabweichung, außerdem n (Stichprobe) und N (Population)
 $\bar{x} = 1,815$, $SD = 3,3043$, $n = 2707$, $N = 7978$

Achtung, Dezimalzahlen mit PUNKT nicht KOMMA eingeben.

Confidence Intervals for a Sample Mean

Input Data			
Sample Mean	1.815		
Sample Std. Deviation	3.3043	Std. Error	0.063509
Sample size	2707		Variance 10.9184
Population size	7978		
Confidence Interval	95	% (100)	

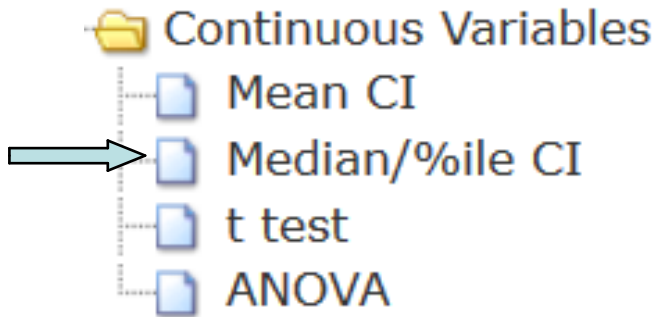
95% Confidence Limits for the Mean of 1.815

Based on:	Lower Limit	Upper Limit
z-test	1.71382	1.91618
t-test	1.71377	1.91623

Für **große Fallzahlen** ($n > 500$) wählt man die Konfidenzgrenzen des **z - test**.

Für **kleine Fallzahlen** ($n < 500$) **t - test**

Quartile mit Konfidenzintervall



90 Daten in aufsteigender Reihenfolge

Welche Werte ergeben sich für Q_1 , Q_2 , Q_3 ?

Input Data

Sample Size	90
Desired percentile	25
Confidence Interval (%)	95

Confidence Interval for 25th percentile of sample size 90

Method	Lower Limit	Rank	Upper Limit
Normal Approximation	14	23	32

Input Data

Sample Size	90
Desired percentile	75
Confidence Interval (%)	95

Confidence Interval for 75th percentile of sample size 90

Method	Lower Limit	Rank	Upper Limit
Normal Approximation	59	68	77

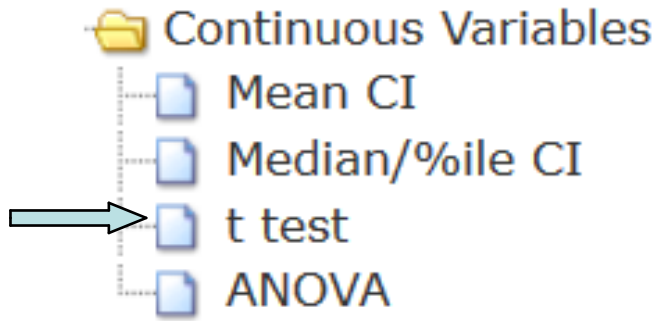
Input Data

Sample Size	90
Desired percentile	50
Confidence Interval (%)	95

Confidence Interval for 50th percentile of sample size 90

Method	Lower Limit	Rank	Upper Limit
Normal Approximation	36	45.5	55

t - Test



Parodontalindex für zwei Berufsgruppen:

Beamte $\bar{x} = 2,376$, SD = 0,959 , n = 415

Facharbeiter $\bar{x} = 2,729$, SD = 0,872 , n = 443

Unterschied ? **deutlich, p = 0,000**

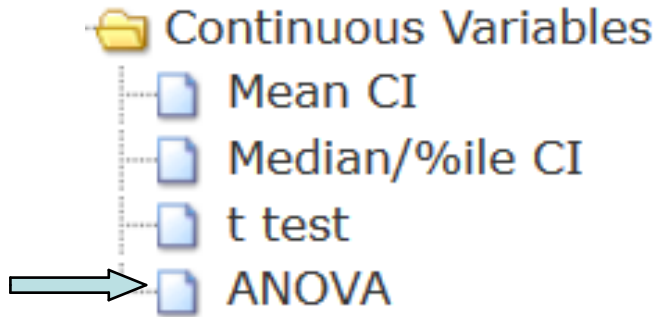
Input Data

Two-sided confidence interval		95%		
	Sample size	Mean	Std. Dev.	Std. Error
Group-1	415	2.376	0.959	
Group-2	443	2.729	0.872	

Result	t statistics	df	p-value ¹	Mean Difference	Lower Limit	Upper Limit
Equal variance	-5.64655	856	<0.0000001	-0.353	-0.475702	-0.230298
Unequal variance	-5.62909	835	<0.0000001	-0.353	-0.476087	-0.229913

	F statistics	df(numerator,denominator)	p-value ¹
Test for equality of variance ²	1.2095	414,442	0.04920

ANOVA



Parodontalindex für vier Berufsgruppen:

Beamte $\bar{x} = 2,38$, SD = 0,96 , n = 415
 Facharbeiter $\bar{x} = 2,73$, SD = 0,87 , n = 443
 Med.Angest. $\bar{x} = 2,04$, SD = 1,04 , n = 39
 Militär $\bar{x} = 2,02$, SD = 0,76 , n = 233

Analysis of Variance (ANOVA)

Input Data

Group	N (count)	Mean	Std. Dev.	Std. error
1	415	2.38	0.96	
2	443	2.73	0.87	
3	39	2.04	1.04	
4	233	2.02	0.76	

ANOVA Table

Source of variation	Sum of squares	d.f	Mean square	F statistics	p-value ¹
Between Groups	86.0049	3	28.6683	36.2216	0.0000000000000000-444089
Within Groups	891.196	1126	0.791471		
Total	977.201	1129			
	Chi square	d.f	p-value ¹		
Test for equality of variance	17.869	3	0.00046807		



Analysis of variance:

Source of variation	Sum of squares	df	Mean square	F	P
Between groups	86.00	3	28.67	36.22	0.000*
Within groups	891.20	1126	0.79		
Total	977.20	1129			

* A low value means significant variation among the means.

Es sind keine multiplen Vergleiche mit OpenEpi möglich.

Hierfür ist z.B. WinPepi geeignet.
www.brixtonhealth.com
 Dieses Programm bietet umfangreiche Möglichkeiten. Einen Überblick findet man in der Datei **FINDER.PDF**
 Großer Umfang, daher ist es nicht Gegenstand dieser Vorlesung.

Comparison of means:

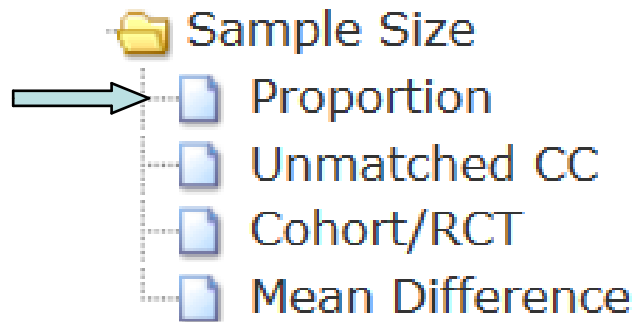
	P#	P*	P**
A,B (2.38 vs 2.73)	0.000	0.000	< 0.01
A,C (2.38 vs 2.04)	0.036	0.056	> 0.05
A,D (2.38 vs 2.02)	0.000	0.000	< 0.01
B,C (2.73 vs 2.04)	0.000	0.000	< 0.01
B,D (2.73 vs 2.02)	0.000	0.000	< 0.01
C,D (2.04 vs 2.02)	0.824	0.909	> 0.05

Simple comparisons; appropriate for planned tests of a priori hypotheses. Assuming that variances are equal.

* Simple comparisons; appropriate for planned tests of a priori hypotheses. Not assuming that variances are equal.

** Taking account of multiple comparisons (Games-Howell procedure).

Fallzahlbestimmung



Hier DEFF = 1 für einfache Zufallsstichproben

Sample Size for Frequency in a Population

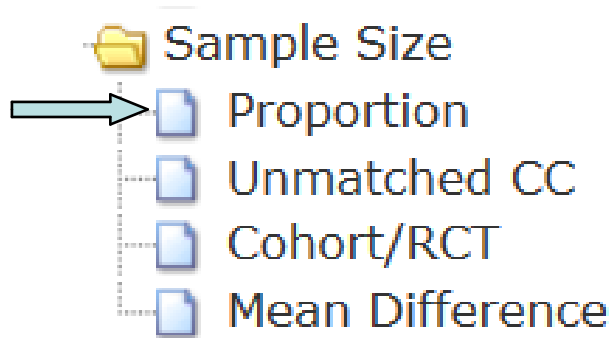
Population size(for finite population correction factor or fpc)(N): 7978
Hypothesized % frequency of outcome factor in the population (p): 62% \pm 1.5
Confidence limits as % of 100(absolute \pm %)(d): 1.5%
Design effect (for cluster surveys- $DEFF$): 1

Sample Size(n) for Various Confidence Levels

ConfidenceLevel(%)
95%

Sample Size
2675

Fallzahlbestimmung



Fallzahlen für Clusterstichproben durch Angabe des DEFF, z.B. DEFF = 1,5

Fallzahl erhöht sich deutlich !

Sample Size for Frequency in a Population

Population size(for finite population correction factor or fpc)(N):	7978
Hypothesized % frequency of outcome factor in the population (p):	62% \pm 1.5
Confidence limits as % of 100(absolute \pm %)(d):	1.5%
Design effect (for cluster surveys- $DEFF$):	1.5

Sample Size(n) for Various Confidence Levels

ConfidenceLevel(%)	Sample Size
95%	4012

Fallzahl für t - Test

Sample Size For Comparing Two Means

- Sample Size
 - Proportion
 - Unmatched CC
 - Cohort/RCT
 - Mean Difference

Input Data			
Confidence Interval (2-sided)	95%		
Power	80%		
Ratio of sample size (Group 2/Group 1)	1		
	Group 1	Group 2	Difference*
Mean	2.38	2.73	-0.35
Standard deviation	0.96	0.87	
Variance	0.9216	0.7569	
Sample size of Group 1	108		
Sample size of Group 2	108		
Total sample size	216		

Jetzt haben Sie [OpenEpi](#) ein wenig kennengelernt. Als weiteres nützliches Programm kann man [WinPepi](#) gelegentlich verwenden. Das besitzt einen weitaus größeren Funktionsumfang und wird daher in einem extra Rahmen näher behandelt werden. Bei Interesse holen Sie sich das Programm von www.brixtonhealth.com. Die Installation auf einem Rechner mit Windows-Betriebssystem geht normalerweise problemlos.