

Sommersemester 2008, Statistik mit SPSS

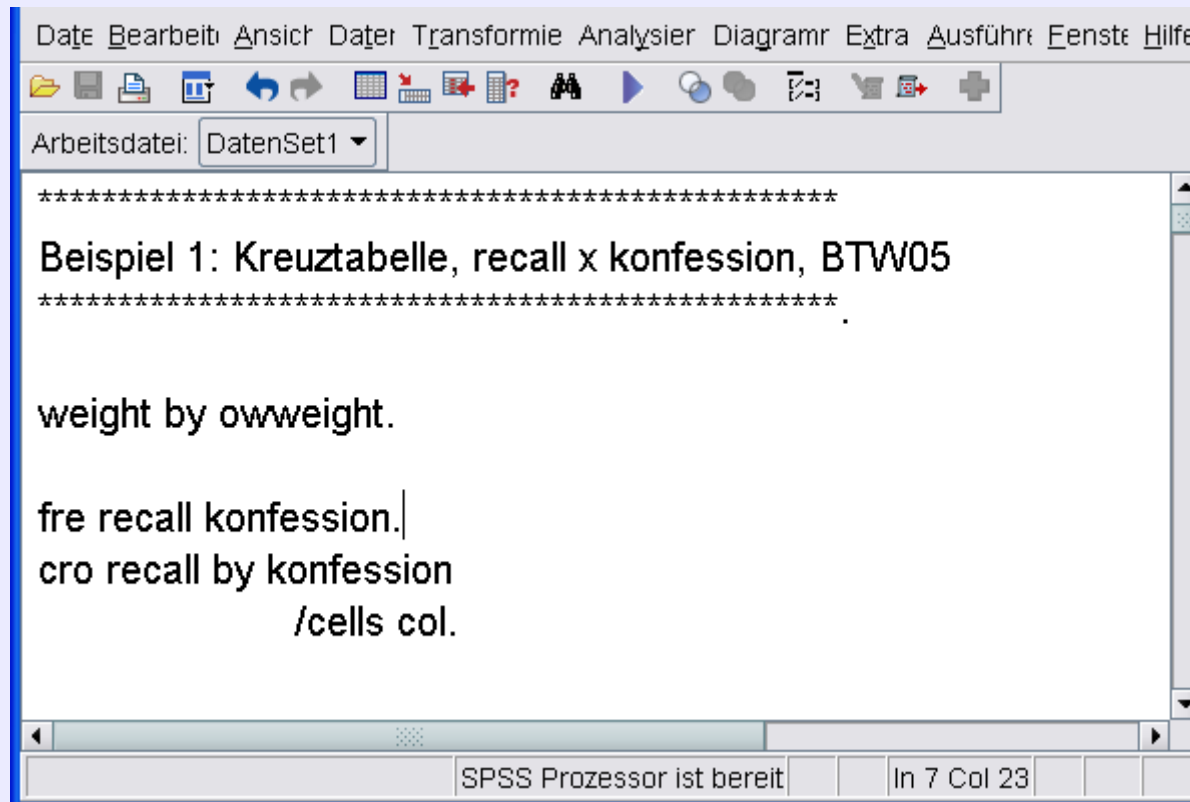


Überblick

1. Mehrfeldertabelle und Zusammenhangsmaße für nominale Daten
2. Mehrfeldertabelle und Zusammenhangsmaße für ordinale Daten

1. Mehrfeldertabelle und Zusammenhangsmaße für **nominale** Daten

Mehrfeldertabellen, nominale Daten



Mehrfeldertabellen, nominale Daten

recall * konfession Konfessionszugehörigkeit Kreuztabelle

% von konfession Konfessionszugehörigkeit

		konfession Konfessionszugehörigkeit			Gesamt
		,00 konfessio nslos	1,00 katholisch	2,00 evangelisch	
recall	1,00 SPD	37,4%	31,8%	46,3%	38,9%
	2,00 CDU/CSU	21,6%	44,2%	29,1%	31,9%
	3,00 B90/Die Grünen	10,7%	7,3%	9,7%	9,2%
	4,00 FDP	9,3%	13,1%	9,5%	10,6%
	5,00 Die Linke.PDS	21,0%	3,5%	5,4%	9,4%
Gesamt		100,0%	100,0%	100,0%	100,0%

Residuen und Chi-Quadrat-Test

Beispiel 2: Residuen recall x konfession

cro recall by konfession

/cells col count s|resid

/stat chi.

 **Standardisierte Residuen und Chi-Quadrat-Test**

Die **standardisierten Residuen** werden berechnet, indem man die Differenz aus beobachteter und erwarteter Häufigkeit durch die Quadratwurzel der erwarteten Häufigkeiten dividiert.

$$sr_{ij} = \frac{n_{ij} - e_{ij}}{\sqrt{e_{ij}}}$$

Standardisierte Residuen

recall * konfession Konfessionszugehörigkeit Kreuztabelle

			konfession Konfessionszugehörigkeit			
			,00 konfession slos	1,00 katholisch	2,00 evangelisch	Gesamt
recall	1,00 SPD	Anzahl	210	199	335	744
		% von konfession Konfessionszuge hörigkeit	37,4%	31,8%	46,3%	38,9%
		Standardisierte Residuen	-,6	-2,9	3,2	
	2,00 CDU/CSU	Anzahl	121	277	211	609
		% von konfession Konfessionszuge hörigkeit	21,6%	44,2%	29,1%	31,9%
		Standardisierte Residuen	-4,3	5,5	-1,3	
	3,00 Bdp/Die Grünen	Anzahl	60	46	70	176
		% von konfession Konfessionszuge hörigkeit	10,7%	7,3%	9,7%	9,2%
		Standardisierte Residuen	1,2	-1,5	,4	
	4,00 FDP	Anzahl	52	82	69	203
		% von konfession Konfessionszuge hörigkeit	9,3%	13,1%	9,5%	10,6%
		Standardisierte Residuen	-1,0	1,9	-,9	
	5,00 Die Linke.PDS	Anzahl	118	22	39	179
		% von konfession Konfessionszuge hörigkeit	21,0%	3,5%	5,4%	9,4%
		Standardisierte Residuen	9,0	-4,8	-3,5	

Standardisierte Residuen:
 ≤ -2 bzw. $\geq +2$.

Die standardisierten Residuen
 deuten auf überzufällige
 Abweichungen von bei
 Unabhängigkeit erwarteten
 Werten hin.

Chiquadrat-Test auf Unabhängigkeit

Chiquadrat-Statistik:

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

n_{ij} = beobachtete Häufigkeit

e_{ij} = erwartete Häufigkeit

Chi-Quadrat-Tests

	Wert	df	Asymptotische Signifikanz (2-seitig)
Chi-Quadrat nach Pearson	194,934 ^a	8	,000
Kontinuitätskorrektur			
Likelihood-Quotient	181,999	8	,000
Zusammenhang linear-mit-linear	56,297	1	,000
Anzahl der gültigen Fälle	1911		

a. 0 Zellen (,0%) haben eine erwartete Häufigkeit kleiner 5. Die minimale erwartete Häufigkeit ist 51,67.

Zusammenhangsmaße für nominale Daten

```
[/STATISTICS=[CHISQ] [LAMBDA] [BTAU] [GAMMA] [ETA ]]  
[PHI ] [UC ] [CTAU] [D ] [CORR]  
[CC ] [KAPPA ] [RISK] [MCNEMAR] [CMH [(value)]]  
[ALL ] [NONE]
```

Zwei Arten von Zusammenhangsmaßen für **nominale** sind in SPSS verfügbar:

1. Symmetrische Maße: Phi, Cramer's V und Kontingenzkoeffizient
(basieren auf dem Chisq – Konzept)
2. Richtungsmaße: Lambda und der Unsicherheitskoeffizient (basieren auf dem Konzept der proportionalen Fehlerreduktion)

Zusammenhangsmaße für nominale Daten

Beispiel 2, Zusammenhangsmaße recall x konfession

*****.

cro recall by konfession

/cells col

/stat phi lambda cc uc.

Berechnet folgende Zusammenhangsmaße:

Phi, **Cramer's V** (phi)

Lambda, Tau (lambda)

Kontingenzkoeffizient (cc)

Unsicherheitskoeffizient (uc)



Symmetrische Maße für nominale Daten

Symmetrische Maße				
		Wert	Asymptotischer Standardfehler ^a	Näherungsweise Signifikanz ^b
Nominal- bzgl. Nominalmaß	Phi	,313		,000
	Cramer-V	,226		,000
	Kontingenzkoeffizient	,304		,000
Anzahl der gültigen Fälle		1911		

Symmetrisches Zusammenhangsmaß, Interpretation von Cramer's V erfolgt analog zu Phi, d.h. Cramers' V = 0,226 = mittlerer Zusammenhang.

Bei einer Irrtumswahrscheinlichkeit von $\leq 0,1\%$ ($p \leq 0,001$) wird die Nullhypothese abgelehnt.

Lambda; Richtungsmaße, die Logik von PRE-Maßen

- Möchte man die Realisation einer abhängigen Variable voraussagen, so kann man dabei Fehler machen.
- Die Anzahl der Fehler lässt sich verringern, wenn die AV mit einer UV zusammenhängt und man die Werte der UV bei den einzelnen Fällen kennt, man also Zusatzinformationen hat.

Das Ausmaß, indem sich die Fehler bei Kenntnis einer UV reduzieren, ergibt sich über folgende Formel

$$\text{PRE (proportional reduction in error)} = \frac{E_0 - E_1}{E_0} = 1 - \frac{E_1}{E_0}$$

Richtungsmaße

		Wert	Asymptotischer Standardfehler ^a	Näherungsweise T ^b	Näherungsweise Signifikanz
Nominal- bzgl. Nominalmaß	Lambda				
	Symmetrisch	,100	,017	5,587	,000
	recall abhängig	,067	,018	3,587	,000

Interpretation von Lambda: Bei Kenntnis der Konfessionszugehörigkeit lässt sich die Wahlentscheidung mit einer um 6,7 % geringeren Fehlerquote vorhersagen als ohne Kenntnis der Konfessionszugehörigkeit

Übungsaufgaben 1 und 2

2. Mehrfeldertabelle und Zusammenhangsmaße für **ordinale** Daten

Mehrfeldertabellen für ordinale Daten

Beispiel 3: Bildung x Politisches Interesse

****Rekodierung.

recode s05 (2,3=1) (4=2) (5=3) (6=4) into schule.

val lab schule 1 'höchstens Hauptschule'

2 'Mittlere Reife'

3 'Abi, Fachabi'

4 'Hochschulabschluss'.

cro f005 by schule

/cells col.

Mehrfeldertabellen für ordinale Daten

005 Staerke politisches Interesse * schule Schulbildung, dreistufig Kreuztabelle

% von schule Schulbildung, dreistufig

		schule Schulbildung, dreistufig				Gesamt
		1,00 höchstens Hauptschule	2,00 Mittlere Reife	3,00 Abi, Fachabi	4,00 Hochschul abschluss	
f005 Staerke politisches Interesse	0 ueberhaupt nicht	7,8%	3,3%	1,7%	,6%	4,5%
	1 wenig	20,3%	11,5%	6,1%	3,6%	13,1%
	2 mittel	43,4%	41,2%	39,0%	23,7%	39,3%
	3 ziemlich stark	19,9%	31,9%	34,1%	40,1%	28,5%
	4 sehr stark	8,7%	12,2%	19,1%	32,0%	14,6%
Gesamt		100,0%	100,0%	100,0%	100,0%	100,0%

Zusammenhangsmaße für ordinale Daten

```
[/STATISTICS=[CHISQ] [LAMBDA] [BTAU] [GAMMA] [ETA ]]  
            [PHI  ] [UC      ] [CTAU] [D      ] [CORR]  
            [CC    ] [KAPPA  ] [RISK] [MCNEMAR] [CMH [(value)]]  
            [ALL   ] [NONE]
```

Folgende Zusammenhangsmaßen für **ordinale** sind in SPSS verfügbar:

1. Symmetrische Maße: Gamma, Tau B, Tau C, Spearmans rho
2. Richtungsmaß: Somers' d

Zusammenhangsmaße für ordinale Daten

```
cro f005 by schule  
  /cells col  
  /stat gamma d btau.
```



Berechnet folgende Zusammen-
hangsmaße:

Gamma (gamma)

Somer's D (d)

btau (btau)

Symmetrische Maße für ordinale Daten

Symmetric Measures

		Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Ordinal by Ordinal	Kendall's tau-b	,279	,016	17,226	,000
	Gamma	,400	,022	17,226	,000
N of Valid Cases		2510			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Richtungsmaße für ordinale Daten

Directional Measures

			Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Ordinal by Ordinal	Somers' d	Symmetric	,279	,016	17,226	,000
		f005 - Staerke politisches Interesse Dependent	,293	,017	17,226	,000
		schule Schulbildung, dreistufig Dependent	,266	,015	17,226	,000

a. Not assuming the null hypothesis.

Interpretation: Es besteht eine positive Beziehung, d.h. je höher die Schulbildung ist, desto stärker ist das politische Interesse. Die Koeffizienten sind bei einer Irrtumswahrscheinlichkeit von kleiner als 0,1% höchst signifikant.

Übungsaufgabe 3

Übersicht: Zusammenhangsmaße bei SPSS

Symmetrische Maße

Nominale Daten:

Phi (zwei dichotome Variablen)

Cramers V, Kontingenzkoeffizient,
(Polytome Variablen)

Ordinale Daten:

Gamma

Tau b

Tau c

Asymmetrische Maße

Nominale Daten:

Lambda

Unsicherheitskoeffizient

Ordinale Daten:

Somers D