

BEZIEHUNGEN ZWISCHEN ORDINALEN VARIABLEN I

Martin-Luther-Universität Halle-Wittenberg

Institut für Soziologie

Übung Einführung in die deskriptive Statistik

Was machen wir heute?

- Logik des Paarvergleichs
- symmetrische Zusammenhangsmaße auf Basis der Logik des Paarvergleiches für ordinale Variablen

Schulabschluss und politisches Interesse

Eine Bildungsforscherin interessiert sich dafür, ob es einen Zusammenhang zwischen dem höchsten Schulabschluss und dem politischen Interesse gibt. Mithilfe des ALLBUS 2014 erhält sie folgende Kreuztabelle:

		höchster Schulabschluss (rek)			Gesamtsumme
		1 - max. Hauptschule	2 - mittlere Reife	3 - FH-/HS-Reife	
politisches Interesse (rek)	1 - schwach	312	232	173	717
	2 - mittel	441	530	432	1403
	3 - stark	285	382	640	1307
Gesamtsumme		1038	1144	1245	3427

Quiz: Handelt es sich um eine symmetrische oder eine asymmetrische Fragestellung? Welches Skalenniveau haben die Ausgangsvariablen?

Schulabschluss und politisches Interesse: Analyse

symmetrische
Fragestellung:
Zusammenhang ohne
Kausalrichtung postuliert

Eine Bildungsforscherin interessiert sich dafür, ob es einen Zusammenhang zwischen dem höchsten Schulabschluss und dem politischen Interesse gibt.

Mehrfeldertafel mit zwei
ordinalen Variablen →
größer/kleiner Relation
vorhanden

		höchster Schulabschluss (rek)			Gesamtsumme
		1 - max. Hauptschule	2 - mittlere Reife	3 - FH-/HS- Reife	
politisches Interesse (rek)	1 - schwach	312	232	173	717
	2 - mittel	441	530	432	1403
	3 - stark	285	382	640	1307
Gesamtsumme		1038	1144	1245	3427

geignetes Maß:
symmetrisches Maß auf
Basis der Logik der
Paarvergleichs

Kreuztabellen: Aufbau

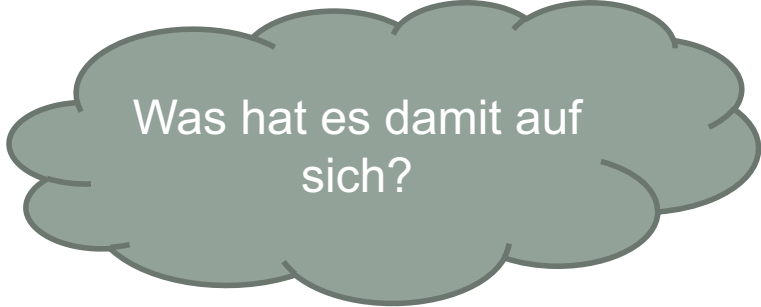
höhere Werte in X

höhere Werte in Y

		höchster Schulabschluss (rek)			Gesamtsumme
		1 - max. Hauptschule	2 - mittlere Reife	3 - FH-/HS- Reife	
politisches Interesse (rek)	1 - schwach	312	232	173	717
	2 - mittel	441	530	432	1403
	3 - stark	285	382	640	1307
Gesamtsumme		1038	1144	1245	3427

Logik des Paarvergleichs I

- bei der Logik des Paarvergleichs werden zwei Fälle A und B hinsichtlich ihrer Ausprägungen auf der X- und Y-Variablen verglichen
- 5 verschiedene Paartypen:
 - Konkordante Paare
 - Diskordante Paare
 - x-verbundene Paare
 - y-verbundene Paare
 - x,y-verbundene Paare



Was hat es damit auf sich?

Logik des Paarvergleichs II: Konkordante Paare

- Formale Beschreibung:
 - $x_a > x_b$ und $y_a > y_b$ (gleichgerichteter Zusammenhang)
 - „Fall A hat sowohl in X einen höheren Wert als Fall B als auch in Y“
 - Indiz für positiven Zusammenhang zwischen X und Y
- Beispiel:

	A: Sven		B: Klaus
X: Abschluss	Realschule (2)	>	Hauptschule (1)
Y: politisches Interesse	mittel (2)	>	niedrig (1)

Hinweis: auch wenn beide Zeichen in die andere Richtung zeigen würden, wäre es ein konkordantes Paar. Wichtig ist hier die gleiche Richtung.

Logik des Paarvergleichs III: Diskordante Paare

- Formale Beschreibung:
 - $x_a > x_b$ und $y_a < y_b$ (gegengerichteter Zusammenhang)
 - „Fall A hat in X einen höheren Wert als Fall B, jedoch einen niedrigeren Wert in Y“
 - Indiz für negativen Zusammenhang zwischen X und Y
- Beispiel:

	A: Sonja		B: Sven
X: Abschluss	Abitur (3)	>	Realschule (2)
Y: politisches Interesse	niedrig (1)	<	mittel (2)

Hinweis: auch wenn beide Zeichen in die andere Richtung zeigen würden, wäre es ein diskordantes Paar. Wichtig ist hier die gegenläufige Richtung.

Logik des Paarvergleichs IV: X-verbundene Paare (Ties in X)

- Formale Beschreibung:
 - $x_a = x_b$ und $y_a \neq y_b$
 - „Fall A hat in X gleichen Wert wie Fall B, aber unterschiedlichen Wert in Y“
- Beispiel:

	A: Sonja		B: Tanja
X: Abschluss	Abitur (3)	=	Abitur (3)
Y: politisches Interesse	niedrig (1)	≠	hoch (3)

Logik des Paarvergleichs V: Y-verbundene Paare (Ties in Y)

- Formale Beschreibung:
 - $x_a \neq x_b$ und $y_a = y_b$
 - „Fall A und B haben unterschiedliche Werte in X, aber gleiche in Y“
- Beispiel:

	A: Klaus		B: Tanja
X: Abschluss	Hauptschule (1)	\neq	Abitur (3)
Y: politisches Interesse	niedrig (1)	$=$	niedrig (1)

Logik des Paarvergleichs VI:

X,Y-verbundene Paare (Ties in X,Y)

- Formale Beschreibung:
 - $x_a = x_b$ und $y_a = y_b$
 - „Fall A und B haben gleiche Werte in X und gleiche Werte in Y“
- Beispiel:

	A: Sven		B: Heidi
X: Abschluss	Realschule (2)	=	Realschule (2)
Y: politisches Interesse	mittel (2)	=	mittel (2)

Logik des Paarvergleiches VII

Konkordante Paare (C)	Diskordante Paare (D)	X-verbundene Paare T_x	Y-verbundene Paare T_y	X,Y-verbundene Paare T_{xy}
$x_a > x_b$ und $y_a > y_b$	$x_a > x_b$ und $y_a < y_b$	$x_a = x_b$ und $y_a \neq y_b$	$x_a \neq x_b$ und $y_a = y_b$	$x_a = x_b$ und $y_a = y_b$
Hinweis auf positive Beziehung	Hinweis auf negative Beziehung	sprechen gegen strikte je-desto-Beziehung	sprechen gegen strikte je-desto-Beziehung	sprechen weder für noch gegen einen Zusammenhang

mithilfe der Paare lassen sich verschiedene symmetrische und asymmetrische ordinale Zusammenhangsmaße konstruieren

Quiz-Time

	A: Tom		B: Tanja
X: Abschluss	Hauptschule (1)		Abitur (3)
Y: politisches Interesse	mittel (2)		hoch (3)

Quiz: Welche Form des Paares liegt hier vor?

Quiz-Time

	A: Tim		B: Heiko
X: Abschluss	Realschule (2)		Abitur (3)
Y: politisches Interesse	mittel (2)		niedrig (1)

Quiz: Welche Form des Paares liegt hier vor?

Schulabschluss und politisches Interesse

		höchster Schulabschluss (rek)			Gesamtsumme
		1 - max. Hauptschule	2 - mittlere Reife	3 - FH-/HS-Reife	
politisches Interesse (rek)	1 - schwach	312	232	173	717
	2 - mittel	441	530	432	1403
	3 - stark	285	382	640	1307
Gesamtsumme		1038	1144	1245	3427

Quiz: Schätzen Sie, wie viele Paare es insgesamt gibt!

$$Gesamt = \binom{n}{2} = \binom{3427}{2} = 5.870.451$$

Logik des Paarvergleiches VIII: Zählregel für konkordante Paare C

- multipliziere die aktuelle Zellenhäufigkeit mit der Summe aller Zellen die rechts unterhalb liegen und verwende dabei alle Zellen für die es rechts unterhalb weitere Tabellenzellen gibt
- bilde anschließend die Summe daraus

Schulabschluss und politisches Interesse: Anzahl konkordanter Paare I

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$C_{(1,1)} = 312 * (530 + 432 + 382 + 640) = 619.008$$

Schulabschluss und politisches Interesse: Anzahl konkordanter Paare II

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$C_{(1,2)} = 232 * (432 + 640) = 248.704$$

Schulabschluss und politisches Interesse: Anzahl konkordanter Paare III

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$C_{(2,1)} = 441 * (382 + 640) = 450.702$$

Schulabschluss und politisches Interesse: Anzahl konkordanter Paare IV

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$C_{(2,2)} = 530 * 640 = 339.200$$

$$C = C_{(1,1)} + C_{(1,2)} + C_{(2,1)} + C_{(2,2)}$$

$$C = 619.008 + 248.704 + 450.702 + 339.200$$

$$C = 1.657.614$$

Logik des Paarvergleiches IX: Zählregel für diskordante Paare D

- verwende alle Zellen, für die es links unterhalb weitere Zellen gibt und multipliziere diese Zellen mit der Summe der Zellen, die links unterhalb liegen
- bilde anschließend die Summe daraus

dark gray	dark gray	red
dark gray	dark gray	light gray
dark gray	dark gray	light gray

Schulabschluss und politisches Interesse: Anzahl diskordanter Paare I

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$D_{(1,3)} = 173 * (441 + 530 + 285 + 382) = 283.374$$

Schulabschluss und politisches Interesse: Anzahl diskordanter Paare II

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$D_{(1,2)} = 232 * (441 + 285) = 168.432$$

Schulabschluss und politisches Interesse: Anzahl diskordanter Paare III

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$D_{(2,3)} = 432 * (285 + 382) = 288.144$$

Schulabschluss und politisches Interesse: Anzahl diskordanter Paare IV

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$D_{(2,3)} = 530 * 285 = 151.050$$

$$D = 891.000$$

Was fangen wir damit an?

Zusammenhangsmaße auf Basis der Logik des Paarvergleichs für ordinale Variablen I

- Grundidee:
 - Überwiegen die konkordanten Paare (positiver Zusammenhang) oder die diskordanten Paare? (C-D)
 - Unterscheiden sich hinsichtlich ihrer Normierung
- allgemeine Interpretation:
 - Stärke analog zu Cramér's V
 - Richtung:
 - -: negativer Zusammenhang
 - 0: kein Zusammenhang
 - +: positiver Zusammenhang

Goodmans und Kruskals γ

- Verwendung:
 - symmetrische Zusammenhänge zwischen ordinalen Variablen
- Idee:
 - Überprüfe, ob die konkordanten oder diskordanten Paare überwiegen und normiere das Ergebnis an der Summe der konkordanten Paare.
 - Normierung an Summe der konkordanten und diskordanten Paare
- Formel:
 - $$\gamma = \frac{C-D}{C+D}$$

Exkurs: alternative Berechnung und Interpretation von Goodmans und Kruskals γ

- Besonderheit:
 - Gamma ist zusätzlich als PRE-Maß interpretierbar
 - (siehe Benninghaus 2012: 170ff.)
- Formel:
 - $PRE = \frac{E_0 - E_1}{E_0}$
 - $E_0 = 0,5(C + D)$
 - $E_1 = \min(C, D)$
 - der Wert wird mit dieser Formel immer positiv
- PRE-Interpretation:
 - Durch Kenntnis von X können wir die Prognose von Y um ... % verbessern und umgekehrt. (Vorzeichen wird hier ignoriert!)

Schulabschluss und politisches Interesse: Berechnung und Interpretation von Gamma

- Berechnung:

- $C = 1.657.614$

- $D = 891.000$

- $\gamma = \frac{C-D}{C+D}$

- $\gamma = \frac{1.657.614 - 891.000}{1.657.614 + 891.000} = 0,301$

- Interpretation:

- Es besteht ein mittlerer positiver Zusammenhang zwischen der Höhe des Schulabschlusses und der Höhe des politischen Interesses. Hohe Werte von X gehen tendenziell mit hohen Werten von Y einher und umgekehrt. Ein höherer Schulabschluss ist folglich tendenziell mit einem höheren politischen Interesse verbunden und umgekehrt.
 - PRE: Durch Kenntnis des Schulabschlusses lässt sich die Prognose des politischen Interesses um 30,1% verbessern und umgekehrt.

Aufgabe 1: Wirtschaftslage BRD und Befragter heute (Gamma)

Ein Wissenschaftler vermutet einen Zusammenhang zwischen der wahrgenommenen allgemeinen Wirtschaftslage und der eigenen Lage heute. Mithilfe des ALLBUS 2014 erhält er folgende Kreuztabelle:

		allgemeine Wirtschaftslage heute			Gesamtsumme
		1 - eher schlecht	2 - teils / teils	3 - eher gut	
persönliche Wirtschaftslage heute	1 - eher schlecht	75	167	99	341
	2 - teils / teils	99	434	399	932
	3 - eher gut	42	634	1508	2184
Gesamtsumme		216	1235	2006	3457

- Bestimmen Sie die konkordanten und diskordanten Paare!
- Berechnen Sie anschließend Gamma!
- Interpretieren Sie ihren Wert sowohl nach der „normalen Assoziationsmaßlogik“ als auch der PRE-Logik!

Aufgabe 1a: Lösung konkordante Paare

		allgemeine Wirtschaftslage heute			Gesamtsumme
		1 - eher schlecht	2 - teils / teils	3 - eher gut	
persönliche Wirtschaftslage heute	1 - eher schlecht	75	167	99	341
	2 - teils / teils	99	434	399	932
	3 - eher gut	42	634	1508	2184
Gesamtsumme		216	1235	2006	3457

- $C_{(1,1)} = 75 * (434 + 399 + 634 + 1508) = 223.125$
- $C_{(1,2)} = 167 * (399 + 1508) = 318.469$
- $C_{(2,1)} = 99 * (634 + 1508) = 212.058$
- $C_{(2,2)} = 434 * 1508 = 654.472$
- **$C = 1.408.124$**

Aufgabe 1a: Lösung diskordante Paare

		allgemeine Wirtschaftslage heute			Gesamtsumme
		1 - eher schlecht	2 - teils / teils	3 - eher gut	
persönliche Wirtschaftslage heute	1 - eher schlecht	75	167	99	341
	2 - teils / teils	99	434	399	932
	3 - eher gut	42	634	1508	2184
Gesamtsumme		216	1235	2006	3457

- $D_{(1,3)} = 99 * (99 + 434 + 42 + 634) = 119.691$
- $D_{(1,2)} = 167 * (99 + 42) = 23.547$
- $D_{(2,3)} = 399 * (42 + 634) = 269.724$
- $D_{(2,2)} = 434 * (42) = 18.228$
- **$D = 431.190$**

Aufgabe 1b: Lösung

- Berechnung Gamma:
 - $C = 1.408.124$
 - $D = 431.190$
 - $\gamma = \frac{C-D}{C+D}$
 - $\gamma = \frac{1.408.124-431.190}{1.408.124+431.190} = 0,531$

Aufgabe 1c: Lösung I

- Interpretation „normales Assoziationsmaß“:
 - $\gamma = 0,531$
 - Es besteht ein hoher positiver Zusammenhang zwischen der Einschätzung der allgemeinen wirtschaftlichen Lage und der persönlichen wirtschaftlichen Lage heute. Hohe Werte in X gehen tendenziell eher mit hohen Werten in Y einher und umgekehrt.
 - Inhaltlich bedeutet gesehen bedeutet das, dass Leute die die allgemeine Wirtschaftslage positiv einstufen auch ihre eigene Wirtschaftslage positiv einstufen und umgekehrt.

Aufgabe 1c: Lösung II

- Interpretation „PRE-Maß“:
 - $\gamma = 0,531$
 - Gamma lässt sich auch als PRE-Maß interpretieren. Durch Kenntnis der Einschätzung der allgemeinen Wirtschaftslage verbessert sich die Prognose der persönlichen Wirtschaftslage um 53,1% und umgekehrt.

Goodmans und Kruskals γ : Eigenschaften

- **Vorteile:**
 - hat neben der normalen Assoziationsmaßinterpretation auch eine PRE-Interpretation
 - leicht zu berechnen
- **Nachteile:**
 - alle Verknüpfungen werden ignoriert
 - anfällig für Eckenkorrelationen
 - stark beeinträchtigt durch Anzahl der Kategorien
 - überschätzt deshalb in der Praxis oft den Zusammenhang

Goodmans und Kruskals γ : Eigenschaften II

	niedrig	mittel	hoch	Gesamt
niedrig	20	0	0	20
mittel	20	0	0	20
hoch	20	2	2	24
Gesamt	60	2	2	64

Quiz: Würden wir hier intuitiv einen Zusammenhang zwischen beiden Variablen vermuten?

Goodmans und Kruskals γ : Eigenschaften III

	niedrig	mittel	hoch	Gesamt
niedrig	20	0	0	20
mittel	20	0	0	20
hoch	20	2	2	24
Gesamt	60	2	2	64

- Berechnung:

- $C = 20 * (0 + 0 + 2 + 2) + 0 * (0 + 2) + 20 * (2 + 2) + 0 * 2)$
- $C = 80 + 80 = 160$
- $D = 0 * (20 + 0 + 20 + 2) + 0 * (20 + 20) + 0 * (20 + 2) + 0 * 20$
- $D = 0$
- $\gamma = \frac{C-D}{C+D} = \frac{160}{160} = +1 \rightarrow$ perfekter positiver Zusammenhang



befriedigendes
Ergebnis?

Kendalls τ_a

- Verwendung:

- symmetrisches Zusammenhangsmaß für ordinale Daten
- berücksichtigt im Gegensatz zu Gamma alle Paare

- Formel:

- $$\tau_a = \frac{C-D}{\frac{n(n-1)}{2}} = \frac{C-D}{\binom{n}{2}} = \frac{C-D}{C+D+T_X+T_Y+T_{X,Y}}$$

- „Bilde die Differenz der konkordanten und diskordanten Paare und normiere das Ergebnis an der Gesamtzahl der möglichen Paare“

Quiz: Warum ist der Betrag vom Tau-a bei Vorliegen von Verknüpfungen immer kleiner als der Betrag von Gamma?

Kendalls τ_a : Schulabschluss

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

- Berechnung:

- $C = 1.657.614$

- $D = 891.000$

- $\tau_a = \frac{C-D}{\frac{n(n-1)}{2}} = \frac{C-D}{\binom{n}{2}}$

- $\tau_a = \frac{1.657.614 - 891.000}{\frac{3427(3427-1)}{2}} = \frac{1.657.614 - 891.000}{\binom{3427}{2}} = \frac{1.657.614 - 891.000}{5.870.451}$

- $\tau_a = 0,131$

Kendalls τ_a : Schulabschluss II

- Berechnung:
 - $\tau_a = 0,131$
- Interpretation:
 - Legt man tau-a als Zusammenhangsmaß zugrunde, so stellt man fest, dass ein niedriger positiver Zusammenhang zwischen dem Schulabschluss und dem politischen Interesse besteht.
 - Befragte mit hohem Schulabschluss haben tendenziell eher ein höhere politisches Interesse und umgekehrt.
 - Der Wert für tau-a ist deutlich niedriger, da tau-a alle Paare berücksichtigt, während Gamma die Normierung nur an der Summe der konkordanten und diskordanten Paare vornimmt.

Aufgabe 2: Wirtschaftslage BRD und Befragter heute (tau-a)

Ein Wissenschaftler vermutet einen Zusammenhang zwischen der wahrgenommenen allgemeinen Wirtschaftslage und der eigenen Lage heute. Mithilfe des ALLBUS 2014 erhält er folgende Kreuztabelle:

		allgemeine Wirtschaftslage heute			Gesamtsumme
		1 - eher schlecht	2 - teils / teils	3 - eher gut	
persönliche Wirtschaftslage heute	1 - eher schlecht	75	167	99	341
	2 - teils / teils	99	434	399	932
	3 - eher gut	42	634	1508	2184
Gesamtsumme		216	1235	2006	3457

$$C = 1.408.124, \quad D = 431.190$$

- a) Berechnen Sie $\tau_a = \frac{C-D}{\frac{n(n-1)}{2}} = \frac{C-D}{\binom{n}{2}}$
- b) Interpretieren Sie Ihr Ergebnis inhaltlich und statistisch.
(Achtung, Kendals tau-Maße sind keine PRE-Maße!)

Aufgabe 2a: Lösung

- gegeben:

- $C = 1.408.124$

- $D = 431.190$

- $n = 3457$

- Berechnung:

- $\tau_a = \frac{C-D}{\binom{n}{2}}$

- $\tau_a = \frac{1.408.124 - 431.190}{\binom{3457}{2}}$

- $\tau_a = 0,163$

Aufgabe 2b: Lösung

- Interpretation:
 - Zwischen der Einschätzung der allgemeinen wirtschaftlichen Lage und der persönlichen wirtschaftlichen Lage besteht gemäß dieser Maßzahl ein niedriger positiver Zusammenhang. Hohe Werte in X gehen tendenziell mit hohen Werten in Y einher und umgekehrt.
 - Inhaltlich bedeutet dies, dass Befragte, die die allgemeine wirtschaftliche Lage als gut einschätzen, tendenziell auch die persönliche wirtschaftliche Lage als gut einschätzen und umgekehrt.

Kendalls τ_a : Eigenschaften

- Eigenschaften:
 - berücksichtigt sämtliche Verknüpfungen
- Vorteile:
 - „extrem ehrliches Maß“
- Nachteile:
 - kann Extremwerte -1 und +1 nur erreichen, wenn gar keine Verknüpfungen vorliegen
 - dadurch meist sehr niedrige Zusammenhangswerte

Kendalls τ_a : Eigenschaften II

	niedrig	mittel	hoch	Gesamt
niedrig	1	0	0	1
mittel	0	1	0	1
hoch	0	0	1	1
Gesamt	1	1	1	3

- $C=3, D=0$
- $T_X = 0, T_Y = 0$
- $T_{XY} = 0$
- $\tau_a = \frac{C-D}{\frac{n(n-1)}{2}}$
- $\tau_a = \frac{3}{3} = 1$

	niedrig	mittel	hoch	Gesamt
niedrig	10	0	0	10
mittel	0	10	0	10
hoch	0	0	10	10
Gesamt	10	10	10	30

- $C=300, D=0$
- $T_X = 0, T_Y = 0$
- $T_{XY} = 135$
- $\tau_a = \frac{C-D}{\frac{n(n-1)}{2}}$
- $\tau_a = \frac{300}{435} = 0,66$

Suboptimal... Verknüpfungen in X,Y sprechen eigentlich nicht gegen einen Zusammenhang!

Kendalls τ_b

- Verwendung:
 - symmetrische Zusammenhänge zwischen ordinalen Variablen in **quadratischer** Kreuztabellen (Extremwerte -/+1 nur dort möglich)

- Formel:

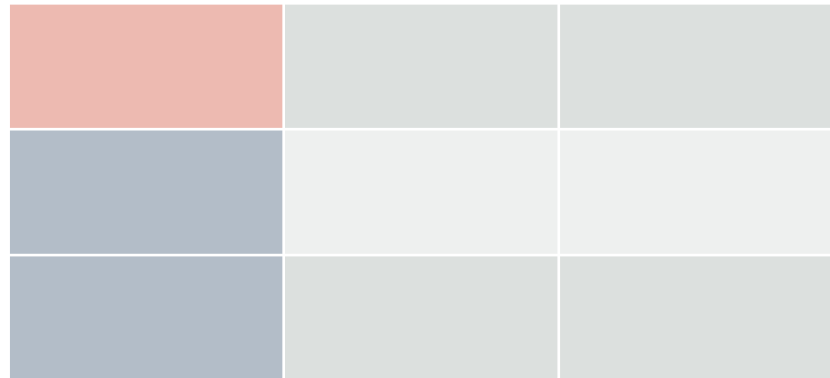
- $$\tau_b = \frac{C-D}{\sqrt{(C+D+T_X)(C+D+T_Y)}}$$

Um die Formel verwenden zu können,
brauchen wir hier zusätzlich Zählregeln für die
Verknüpfungen...

Logik des Paarvergleiches X:

Zählregeln für x-verbundene Paare T_X

- verwende alle Zellen, für die es unterhalb weitere Zellen gibt und multipliziere jede dieser Zellen mit der Summe der Zellen, die unterhalb dieser liegen
- bilde anschließend die Summe daraus



Schulabschluss und politisches Interesse: Anzahl x-verbundener Paare I

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$T_{X(1,1)} = 312 * (441 + 285) = 226.512$$

Schulabschluss und politisches Interesse: Anzahl x-verbundener Paare II

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$T_{X(1,2)} = 232 * (530 + 382) = 211.584$$

Schulabschluss und politisches Interesse: Anzahl x-verbundener Paare III

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$T_{X(1,3)} = 173 * (432 + 640) = 185.456$$

Schulabschluss und politisches Interesse: Anzahl x-verbundener Paare IV

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$T_{X(2,1)} = 441 * 285 = 125.685$$

Schulabschluss und politisches Interesse: Anzahl x-verbundener Paare V

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$T_{X(2,2)} = 530 * 382 = 202.460$$

Schulabschluss und politisches Interesse: Anzahl x-verbundener Paare VI

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$T_{X(2,3)} = 432 * 640 = 276.480$$

$$T_X = 1.228.177$$

Logik des Paarvergleiches XI:

Zählregeln für y -verbundene Paare T_Y

- verwende alle Zellen, für die es rechts weitere Zellen gibt und multipliziere jede dieser Zellen mit der Summe der Zellen, die rechts davon liegen
- bilde anschließend die Summe daraus

Schulabschluss und politisches Interesse: Anzahl y-verbundener Paare I

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$T_{Y(1,1)} = 312 * (232 + 173) = 126.360$$

Schulabschluss und politisches Interesse: Anzahl y-verbundener Paare II

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$T_{Y(1,2)} = 232 * 173 = 40136$$

Schulabschluss und politisches Interesse: Anzahl y-verbundener Paare III

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$T_{Y(2,1)} = 441 * (530 + 432) = 424.242$$

Schulabschluss und politisches Interesse: Anzahl y-verbundener Paare IV

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$T_{Y(2,2)} = 530 * 432 = 228.960$$

Schulabschluss und politisches Interesse: Anzahl y-verbundener Paare V

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$T_{Y(3,1)} = 285 * (382 + 640) = 291.270$$

Schulabschluss und politisches Interesse: Anzahl y-verbundener Paare VI

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$T_{Y(3,2)} = 382 * 640 = 244.480$$

$$T_Y = 1.355.448$$

Logik des Paarvergleichs XII:

Zählregeln für x- und y-verbundene Paare $T_{X,Y}$

- Regel:

- $$T_{X,Y} = \sum_{i=1}^I \sum_{j=1}^J \frac{n_{i,j} * (n_{i,j} - 1)}{2}$$

- „Multipliziere für jede Zelle die entsprechende absolute Häufigkeit der Zelle mit dem um eins verminderten Wert und teile das Ergebnis durch zwei. Bilde anschließend die Summe.“

- Alternative Regel:

- $$T_{X,Y} = \sum_{i=1}^I \sum_{j=1}^J \binom{n_{i,j}}{2}$$

Schulabschluss und politisches Interesse: Anzahl x,y-verbundener Paare

Politisches Interesse	1 max Hauptschulabschluss	2 mittlere Reife	3 FH/HS-Reife	Gesamt
1 schwach	312	232	173	717
2 mittel	441	530	432	1403
3 stark	285	382	640	1307
gesamt	1038	1144	1245	3427

$$\begin{aligned}
 T_{X,Y} &= \binom{312}{2} + \binom{232}{2} + \binom{173}{2} + \binom{441}{2} + \binom{530}{2} + \binom{432}{2} + \binom{285}{2} + \binom{382}{2} \\
 &+ \binom{640}{2} = 738.212
 \end{aligned}$$



Kendalls τ_b : Eigenschaften

- Eigenschaften:
 - im Gegensatz zur Berechnung von τ_a sprechen hier Paare in X,Y und nicht gegen einen Zusammenhang sie werden bei der Berechnung nicht negativ ausgelegt
 - erreicht Extremwerte -1 und +1 jedoch nur in Tabellen mit gleicher Zeilen- und Spaltenzahl

Kendalls τ_b : Eigenschaften II

	niedrig	mittel	hoch	Gesamt
niedrig	10	0	0	10
mittel	0	10	0	10
hoch	0	0	10	10
Gesamt	10	10	10	30

- $C=300, D=0$
- $T_X = 0, T_Y = 0$
- $T_{XY} = 135$
- $\tau_a = 0,66$
- $\tau_b = \frac{C-D}{\sqrt{(C+D+T_X)(C+D+T_Y)}}$
- $\tau_b = \frac{300}{\sqrt{(300)(300)}} = 1$

Kendalls τ_c

- Verwendung:
 - symmetrische Zusammenhänge zwischen ordinalen Variablen
 - kann Wert von -1 und +1 auch in nichtquadratischen Tabellen erreichen und wird deswegen meist auch nur dort verwendet

- Formel:

- $$\tau_c = \frac{C - D}{\frac{1}{2}n^2 * \left(\frac{\min(I-1, J-1)}{\min(I, J)} \right)}$$

I: Anzahl Zeilen
J: Anzahl Spalten

Aufgabe 3: Wirtschaftslage BRD und Befragter heute (tau-b)

Ein Wissenschaftler vermutet einen Zusammenhang zwischen der wahrgenommenen allgemeinen Wirtschaftslage und der eigenen Lage heute. Mithilfe des ALLBUS 2014 erhält er folgende Kreuztabelle:

		allgemeine Wirtschaftslage heute			Gesamtsumme
		1 - eher schlecht	2 - teils / teils	3 - eher gut	
persönliche Wirtschaftslage heute	1 - eher schlecht	75	167	99	341
	2 - teils / teils	99	434	399	932
	3 - eher gut	42	634	1508	2184
Gesamtsumme		216	1235	2006	3457

$$C = 1.408.124, \quad D = 431.190$$

- Bestimmen Sie die Verknüpfungen in X sowie in Y.
- Berechnen Sie $\tau_b = \frac{C-D}{\sqrt{(C+D+T_X)(C+D+T_Y)}}$!
- Interpretieren Sie Ihr Ergebnis inhaltlich und statistisch.

Aufgabe 3a: Lösung Verknüpfungen in X

		allgemeine Wirtschaftslage heute			Gesamtsumme
		1 - eher schlecht	2 - teils / teils	3 - eher gut	
persönliche Wirtschaftslage heute	1 - eher schlecht	75	167	99	341
	2 - teils / teils	99	434	399	932
	3 - eher gut	42	634	1508	2184
Gesamtsumme		216	1235	2006	3457

- **Berechnung:**

- $T_X = 75 * (99 + 42) + 167 * (434 + 634) + 99 * (399 + 1508) + 99 * 42 + 434 * 634 + 399 * 1508$
- $T_X = 10.575 + 178.356 + 188.793 + 4.158 + 275.156 + 601.692$
- $T_X = 1.258.730$

Aufgabe 3a: Lösung Verknüpfungen in Y

		allgemeine Wirtschaftslage heute			Gesamtsumme
		1 - eher schlecht	2 - teils / teils	3 - eher gut	
persönliche Wirtschaftslage heute	1 - eher schlecht	75	167	99	341
	2 - teils / teils	99	434	399	932
	3 - eher gut	42	634	1508	2184
Gesamtsumme		216	1235	2006	3457

- **Berechnung:**

- $T_Y = 75 * (167 + 99) + 167 * 99 + 99 * (434 + 399) + 434 * 399 + 42 * (634 + 1508) + 634 * 1508$
- $T_Y = 19.950 + 16.533 + 173.166 + 89.964 + 956.072$
- $T_Y = 1.338.152$

Aufgabe 3b: Lösung

- gegeben:

- $C = 1.408.124$

- $D = 431.190$

- $T_X = 1.258.730$

- $T_Y = 1.338.152$

- Berechnung:

- $$\tau_b = \frac{C-D}{\sqrt{(C+D+T_X)(C+D+T_Y)}}$$

- $$\tau_b = \frac{1.408.124 - 431.190}{\sqrt{(1.408.124+431.190+1.258.730)(1.408.124+431.190+1.338.152)}}$$

- $\tau_b = 0,311$

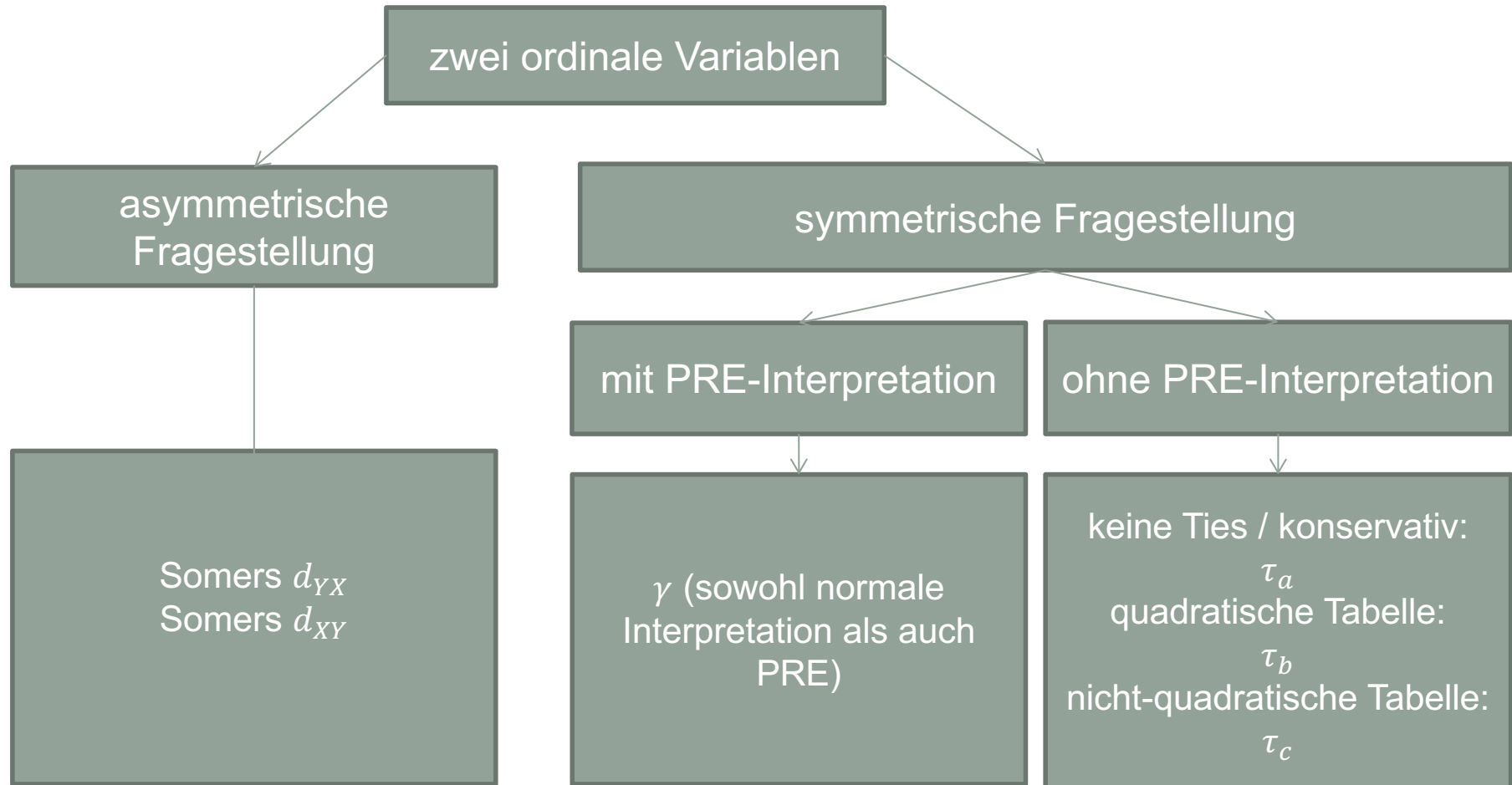
Aufgabe 3b: Lösung

- Berechnung:
 - $\tau_b = 0,311$
- Interpretation:
 - Es besteht ein mittlerer positiver Zusammenhang zwischen der Wahrnehmung der allgemeinen wirtschaftlichen Lage und der persönlichen. Hohe Werte in X gehen tendenziell mit hohen Werten in Y einher und umgekehrt.
 - Dies bedeutet inhaltlich, dass Befragte, die allgemeine Wirtschaftslage als positiv einschätzen tendenziell eher auch ihre eigene Wirtschaftslage positiv einschätzen und umgekehrt.

Daumenregeln: Wann sollte man welchen Koeffizienten anwenden?

- Gamma:
 - bei Interesse an PRE-Interpretation
 - wenn auch Eckenkorrelationen einen perfekten Zusammenhang zeigen sollen
 - wenn Verknüpfungen ignoriert werden sollen
 - wenn man hohe Zusammenhangswerte haben will 😊
- Tau-a
 - wenn keine Verknüpfungen vorliegen
 - wenn man hinsichtlich des Zusammenhangs konservativ sein will
- Tau-b
 - ideal bei quadratischen Tabellen, wenn PRE-Interpretation nicht benötigt wird
- Tau-c
 - bei nicht-quadratischen Tabellen, wenn PRE-Interpretation nicht benötigt wird

Entscheidungsbaum für ordinale Zusammenhangsmaße



Übungsaufgabe 1: Politisches Interesse und Wichtigkeit von Politik und öffentlichem Leben

Eine Politikwissenschaftlerin stellt die Hypothese auf, dass es einen Zusammenhang zwischen dem politischen Interesse der Befragten und der wahrgenommenen Wichtigkeit von Politik und öffentlichem Leben gibt.

		politisches Interesse			Gesamt
		stark	mittelmäßig	schwach	
Wichtigkeit Politik und öffentliches Leben	eher unwichtig	77	297	668	1042
	teils/teils	167	480	274	921
	eher wichtig	739	608	163	1510
Gesamt		983	1385	1105	3473

- Welches Maß ist zur Untersuchung des Zusammenhangs geeignet, wenn Sie eine PRE-Interpretation vornehmen wollen?
- Berechnen Sie das Zusammenhangsmaß.
- Interpretieren Sie das Zusammenhangsmaß statistisch und inhaltlich?
- Welche Nachteile weist das Zusammenhangsmaß auf?

Übungsaufgabe 1a: Analyse

		politisches Interesse			Gesamt
		stark	mittelmäßig	schwach	
Wichtigkeit Politik und öffentliches Leben	eher unwichtig	77	297	668	1042
	teils/teils	167	480	274	921
	eher wichtig	739	608	163	1510
Gesamt		983	1385	1105	3473

ordinale Variable

ordinale Variable

...dass es einen Zusammenhang zwischen dem politischen Interesse der Befragten und der wahrgenommenen Wichtigkeit von Politik und öffentlichem Leben gibt.

keine Hypothese über Richtung → symmetrischer Zusammenhang gesucht

Symmetrisches Zusammenhangsmaß für
ordinale Variablen mit PRE-Interpretation → Gamma

Übungsaufgabe 1b: Lösung

- Konkordante Paare:

- $C_{(1,1)} = 77 * (480 + 274 + 608 + 163) = 117425$
- $C_{(1,2)} = 297 * (274 + 163) = 129789$
- $C_{(2,1)} = 167 * (608 + 163) = 128757$
- $C_{(2,2)} = 480 * 163 = 78240$
- $C = 454211$

- Diskordante Paare:

- $D_{(1,3)} = 668 * (167 + 480 + 739 + 608) = 1331992$
- $D_{(1,2)} = 297 * (167 + 739) = 269082$
- $D_{(2,3)} = 274 * (739 + 608) = 369078$
- $D_{(2,2)} = 480 * (739) = 354720$
- $D = 2324872$

Übungsaufgabe 1b: Lösung II

- Berechnung:

- $N_c = 454211, N_d = 2324872$

- $\gamma = \frac{N_c - N_d}{N_c + N_d}$

- $\gamma = \frac{454211 - 2324872}{454211 + 2324872}$

- $\gamma = -0,673$

- Gamma beträgt -0,673. Es besteht ein starker negativer Zusammenhang zwischen dem politischen Interesse und der wahrgenommenen Wichtigkeit von Politik und öffentlichem Leben

Übungsaufgabe 1c: Lösung

- Interpretation:
 - $\gamma = -0,673$
 - Es besteht ein starker negativer Zusammenhang zwischen dem politischen Interesse (X) und der wahrgenommenen Wichtigkeit von Politik und öffentlichem Leben (Y). Hohe Werte auf der X-Achsen gehen tendenziell mit niedrigen Werten auf der Y-Achsen einher und umgekehrt.
 - Da ein hoher Wert von X für ein schwach ausgeprägtes politisches Interesse steht und ein niedriger Wert von Y für die Einschätzung, dass Politik und öffentliches Leben unwichtig sind, bedeutet dies, dass mit einem schwachen politischen Interesse tendenziell einhergeht, dass Politik und öffentliches Leben als eher unwichtig eingeschätzt werden.

Merke: Immer auf die Kodierung der Variablen achten! Die Tabellen sind so aufgebaut, dass links niedrigere Werte als rechts stehen und oben niedrigere Werte als unten.

Übungsaufgabe 1c: Lösung

- Interpretation II:
 - $\gamma = -0,673$
 - Da es sich bei Gamma um ein PRE-Maß handelt, lässt sich darüber hinaus sagen, dass sich durch Kenntnis des politischen Interesses die Prognose der Wichtigkeit von Politik und öffentlichem Leben um 67,3 Prozent verbessern lässt und umgekehrt.

Übungsaufgabe 2: EU-Verlustängste

Eine EU-Integrationsforscherin interessiert sich für die Ängste der Befragten hinsichtlich der Europäischen Union. Sie nimmt an, dass Leute, die befürchten durch die europäische Integration ihren Job zu verlieren, auch befürchten, ihre nationale Identität / Kultur zu verlieren. Beide Variablen sollen miteinander einhergehen.

Angst vor Verlust nationale Identität / Kultur * Angst vor Jobverlust Kreuztabelle

Anzahl

		Angst vor Jobverlust			Gesamt
		große Angst	teils/teils	kaum Angst	
Angst vor Verlust nationale Identität / Kultur	große Angst	603	79	18	700
	teils/teils	478	207	39	724
	kaum Angst	296	174	131	601
Gesamt		1377	460	188	2025

- Berechnen Sie ein geeignetes Zusammenhangsmaß für die oben aufgestellte Hypothese, welches Ties berücksichtigt.
- Interpretieren Sie dieses inhaltlich wie statistisch.

Übungsaufgabe 2a: Lösung

- Analyse:
 - Beide Variablen sind ordinal, die Forscherin geht von einem symmetrischen Zusammenhang aus und möchte Ties berücksichtigen. Deshalb sollte hier tau-b berechnet werden.

Übungsaufgabe 2b: Lösung

- Berechnung der notwendigen Paare:

- $C = 603 * (207 + 39 + 174 + 131) + 478 * (174 + 131) + 79 * (39 + 131) + 207 * 131 = 518590$

- $D = 296 * (207 + 79 + 39 + 18) + 478 * (79 + 18) + 174 * (39 + 18) + 207 * 18 = 161538$

- $T_x = 603 * (478 + 296) + 478 * 296 + 79 * (207 + 174) + 207 * 174 + 18 * (39 + 131) + 39 * 131 = 682496$

- $T_y = 603 * (79 + 18) + 79 * 18 + 478 * (207 + 39) + 207 * 39 + 296 * (174 + 131) + 174 * 131 = 298648$

Übungsaufgabe 2b: Lösung II

- Berechnung von tau-b:

- $$\tau_b = \frac{C-D}{\sqrt{(C+D+T_X)*(C+D+T_Y)}}$$

- $$\tau_b = \frac{(518590-161538)}{\sqrt{(518590+161538+682496)*(518590+161538+298648)}}$$

- $$\tau_b = \frac{357052}{1154860,89} = 0,31$$

Übungsaufgabe 2c: Lösung

- Interpretation:
 - Tau-b beträgt im vorliegenden Fall 0,31.
 - Dies bedeutet, dass zwischen beiden Variablen ein mittlerer positiver Zusammenhang besteht. Eine große Angst vor Jobverlust durch den EU-Integrationsprozess geht tendenziell einher mit einer großer Angst vor Verlust der nationalen Identität und umgekehrt.

Übungsaufgabe 3: Gesundheit und Lebenszufriedenheit 2014 (USA)

Mithilfe des GSS 2014 möchte eine US-Gesundheitssoziologin herausfinden, ob es einen Zusammenhang zwischen der Wahrnehmung des eigenen Gesundheitszustands und der Lebenszufriedenheit eines Befragten gibt. Mithilfe von SPSS erhält sie hierfür folgende Kreuztabelle:

		CONDITION OF HEALTH				Gesamtsumme
		EXCELLENT	GOOD	FAIR	POOR	
GENERAL HAPPINESS	VERY HAPPY	215	257	57	16	545
	PRETTY HAPPY	185	473	202	57	917
	NOT TOO HAPPY	39	72	79	52	242
Gesamtsumme		439	802	338	125	1704

- Welches Maß ist zur Untersuchung des Zusammenhangs geeignet, wenn Sie an einer Aussage zur proportionalen Fehlerreduktion nicht interessiert sind?
- Berechnen Sie das Maß!
- Interpretieren Sie Ihr Ergebnis statistisch und inhaltlich.

Übungsaufgabe 3a: Analyse

- geeignetes Zusammenhangsmaß:
 - Da ein tau-Maß für einen symmetrischen Zusammenhang gesucht ist, Ties in X, Y vorliegen und es sich um eine nicht-quadratische Tabelle handelt, ist tau-c am besten geeignet.

Übungsaufgabe 3b: Lösung

		CONDITION OF HEALTH				Gesamtsumme
		EXCELLENT	GOOD	FAIR	POOR	
GENERAL HAPPINESS	VERY HAPPY	215	257	57	16	545
	PRETTY HAPPY	185	473	202	57	917
	NOT TOO HAPPY	39	72	79	52	242
Gesamtsumme		439	802	338	125	1704

- Konkordante Paare:

- $C = 215 * (473 + 202 + 57 + 72 + 79 + 52) + 257 * (202 + 57 + 79 + 52) + 57 * (57 + 52) + 185 * (72 + 79 + 52) + 473 * (79 + 52) + 202 * 52$
- $C = 417.490$

Übungsaufgabe 3b: Lösung II

		CONDITION OF HEALTH				Gesamtsumme
		EXCELLENT	GOOD	FAIR	POOR	
GENERAL HAPPINESS	VERY HAPPY	215	257	57	16	545
	PRETTY HAPPY	185	473	202	57	917
	NOT TOO HAPPY	39	72	79	52	242
Gesamtsumme		439	802	338	125	1704

- Diskordante Paare:

- $D = 16 * (185 + 473 + 202 + 39 + 72 + 79) + 57 * (185 + 473 + 39 + 72) + 257 * (185 + 39) + 57 * (39 + 72 + 79) + 202 * (39 + 72) + 473 * 39$
- $D = 169.900$

Übungsaufgabe 3b: Lösung III

- Berechnung:

- $$\tau_c = \frac{C-D}{\frac{1}{2}n^2 * \left(\frac{\min(I-1, J-1)}{\min(I, J)}\right)}$$

- $$\tau_c = \frac{417.490 - 169.900}{\frac{1}{2} * 1704^2 * \left(\frac{\min(3-1, 4-1)}{\min(3; 4)}\right)}$$

- $$\tau_c = \frac{247590}{\frac{1}{2} * 1704^2 * \frac{2}{3}} = 0,256$$

Übungsaufgabe 3c: Lösung

- Interpretation:
 - Es besteht ein mittlerer positiver Zusammenhang zwischen der Wahrnehmung der des eigenen Gesundheitszustands der Befragten und Lebenszufriedenheit.
 - Da hohe Werte auf der X-Variablen für eine negative Einschätzung der eigenen Gesundheit stehen und hohe Werte der Y-Variablen für eine hohe Unzufriedenheit mit dem Leben, bedeutet dies, dass Befragte, die einen schlechten Gesundheitszustand aufweisen, tendenziell eher unzufrieden sind.

Übungsaufgabe 4: Interesse an Technologie

Ein US-Soziologe vertritt die Hypothese, dass ein hohes technologisches Interesse auch mit einem hohen Interesse hinsichtlich Weltraumerkundungen einhergeht und umgekehrt.

		INTERESTED IN TECHNOLOGIES			Gesamt
		Very interested	Moderately interested	Not at all interested	
INTERESTED IN SPACE EXPLORATION	Very interested	181	38	2	221
	Moderately interested	172	237	24	433
	Not at all interested	67	176	86	329
Gesamt		420	451	112	983

- Der soeben genannte Soziologe interessiert sich vor allem für eine Aussage zur proportionalen Fehlerreduktion. Welches Maß ist hierfür geeignet?
- Berechnen Sie dieses Maß!
- Interpretieren Sie dieses inhaltlich und statistisch! Gehen Sie sowohl auf die normale Assoziationsmaßlogik als auch die PRE-Logik ein.

Übungsaufgabe 3a: Analyse

		INTERESTED IN TECHNOLOGIES			Gesamt
		Very interested	Moderately interested	Not at all interested	
INTERESTED IN SPACE EXPLORATION	Very interested	181	38	2	221
	Moderately interested	172	237	24	433
	Not at all interested	67	176	86	329
Gesamt		420	451	112	983

ordinale Variable

ordinale Variable

symmetrischer
Zusammenhang

Ein US-Soziologe vertritt die Hypothese, dass ein hohes technologisches Interesse auch mit einem hohen Interesse hinsichtlich Weltraumerkundungen **einhergeht und umgekehrt.**

Maß mit PRE-Logik gesucht → Gamma berechnen

Übungsaufgabe 4b: Berechnung

- Formel:

- $\gamma = \frac{N_c - N_d}{N_c + N_d}$

- konkordante Paare:

- $N_c = 181 * (237 + 24 + 176 + 86) + 38 * (24 + 86) + 172 * (176 + 86) + 237 * 86 = 164289$

- diskordante Paare:

- $N_d = 2 * (172 + 237 + 67 + 176) + 38 * (172 + 67) + 24 * (67 + 176) + 237 * 67 = 32097$

- Einsetzen:

- $\gamma = \frac{164289 - 32097}{164289 + 32097} = 0,673$

Übungsaufgabe 4c: Interpretation

- **Assoziationslogik:**
 - $\gamma = 0,673$. Es besteht folglich ein hoher positiver Zusammenhang zwischen beiden Variablen. Hohe Werte beim allgemeinen Technologieinteresse gehen mit hohen Werten beim Interesse in der Weltraumerkundung einher und umgekehrt.
 - Inhaltlich bedeutet dies, dass Befragte mit einem hohen Interesse in Technologie auch ein großes Interesse an Weltraumerkundung haben und umgekehrt.
- **PRE-Logik:**
 - Durch Kenntnis der einen Variablen verbessert sich die Prognose der anderen Variablen um 67,3%. Durch Kenntnis des allgemeinen Technologieinteresses lässt sich der Vorhersagefehler für das Interesse an Weltraumerkundung um 67,3% reduzieren und umgekehrt.

SPSS: Ordinale Zusammenhangsmaße

- Analysieren → Deskriptive Statistiken → Kreuztabellen → Statistiken

Kreuztabellen

Zeile(n):
GENERAL HAPPINESS [happy]

Spalten:
CONDITION OF HEALTH [health]

Schicht 1 von 1
Vorherige Nächste

Gruppiere Balkendiagramme anzeigen
 Keine Tabellen

Schichtvariablen in Tabellenschichten anzeigen

OK Einfügen Zurücksetzen Abbrechen Hilfe

Buttons: Exakt... Statistiken... Zellen... Format... Stil... Bootstrap...

SPSS: Ordinale Zusammenhangsmaße II

Kreuztabellen: Statistik ×

Chi-Quadrat Korrelationen

Nominal

Kontingenzkoeffizient
 Phi und Cramer-V
 Lambda
 Unsicherheitskoeffizient

Ordinal

Gamma
 Somers-d
 Kendall-Tau-b
 Kendall-Tau-c

Nominal bezüglich Intervall

Eta

Kappa
 Risiko
 McNemar

Cochran- und Mantel-Haenszel-Statistik
Gemeinsames Odds-Verhältnis:

SPSS: Ordinale Zusammenhangsmaße III

Symmetrische Maße

		Wert	Asymp. Standardfehler r^a	Näherungsw eise A^b	Näherungsw eise Sig.
Ordinal bezüglich Ordinal	Kendall-Tau-c	,256	,021	12,477	,000
	Gamma	,422	,031	12,477	,000
Anzahl der gültigen Fälle		1704			

- Die Nullhypothese wird nicht vorausgesetzt.
- Unter Annahme der Nullhypothese wird der asymptotische Standardfehler verwendet.

$\gamma = 0,422$. Es besteht folglich ein mittlerer positiver Zusammenhang zwischen der Variablen X und der Variablen Y und umgekehrt, wenn man Gamma als Zusammenhangsmaß verwendet. Hohe Werte von X gehen tendenziell mit hohen Werten von Y einher. Durch Kenntnis von X verbessert sich die Prognose von Y um 42,2%. Legt man $\tau_c = 0,256$ als Zusammenhangsmaß zugrunde, so zeigt sich auch hier ein mittlerer positiver Zusammenhang.

Literaturhinweise

- Kerstin Völkl / Christoph Korb (2018): Deskriptive Statistik. Eine Einführung für Politikwissenschaftlerinnen und Politikwissenschaftler. S. 180-204.
- Steffen-M. Kühnel / Dagmar Krebs (2014): Statistik für die Sozialwissenschaften. Grundlagen, Methoden, Anwendungen. S. 370-384.
- Hans Benninghaus (2007): Deskriptive Statistik. Eine Einführung für Sozialwissenschaftler. S. 137-184.