STATISTIK FÜR DIE SOZIALWISSENSCHAFTEN

Bivariate Zusammenhangsmaße

Vertiefung Zusammenhangsmaße für ordinale Variablen;

Vervollständigung bivariater Zusammenhänge

Meine Notizen:

Themen für heute

- Mittelwertvergleiche mit ordinalen Daten
- Zusammenhänge zwischen ordinal und metrisch
- Dichotomisierung nominaler Variablen als Strategie
- Übersicht: Bivariate Zusammenhänge

Meine Notizen:

Abstände bei ordinalen Daten

- Grundproblem:
  - Eigentlich liegt es in der Natur des Skalenniveaus, dass die Abstände der Ausprägungen unbekannt sind.
  - Für Zusammenhangsmaße, die über das nominale Messniveau hinausgehen sollen, brauchen wir immer eine Annahme über Abstände.
- Ansatz 1: Behandeln als metrische Variable; Bedeutet letztlich die Annahme:
  - Benachbarte Ausprägungen haben immer gleiche Abstände (Äquidistanz)
  - Wird bei ordinaler Messung latenter Konstrukte oft als sinnvoll angenommen.
  - Auch für Zustimmungs-Items der meist gewählte Weg.
- Ansatz 2: Aufbrechen des Zusammenhangs in Vielzahl von Paarvergleichen
- Ansatz 3: Analyse von Rangplätzen statt von Ausprägungen
  - Statt gleichen Abständen zwischen den Ausprägungen wird gleicher Abstand zwischen den Rangplätzen angenommen (kennen Sie vom IQR).

Meine Notizen:

Analyse von Rangplätzen statt von Ausprägungen

- Lösung des Grundproblems:
  - Umrechnung ordinaler Ausprägungen in Rangwerte und Behandeln von Rangwerten wie metrische Daten!
- Berechnung:
  - Schritt 1: Aufsteigende Sortierung aller Fälle;
  - Schritt 2: kleinster Wert → Rangplatz 1: $x_{(1)}$ → $r_X = 1$
    zweitkleinster Wert → Rangplatz 2: $x_{(2)}$ → $r_X = 2$
    …
    zweitgrößter Wert → Rangplatz n-1: $x_{(n-1)}$ → $r_X = n-1$
    größter Wert → Rangplatz n: $x_{(n)}$ → $r_X = n$
  - Mögliches Problem: Fälle mit gleicher Ausprägung:
    → Verwendung mittlerer Ränge bei gleichen Ausprägungen

Meine Notizen:

Beispiel der Berechnung von Rangplätzen

- Berechnung:
  - Schritt 1: Aufsteigende Sortierung aller Fälle;
  - Schritt 2: mit Rangplätzen belegen
  - Mögliches Problem: Fälle mit gleicher Ausprägung:
    → Verwendung mittlerer Ränge bei gleichen Ausprägungen
- Beispiel: {1,2,3,4,4,5,4,1,4,1,3}
  - Schritt 1: sortieren: {1,1,1,2,3,3,4,4,4,4,5}
  - Schritt 2: Rangwerte 1,2,3,4,5,6,7,8,9,10,11

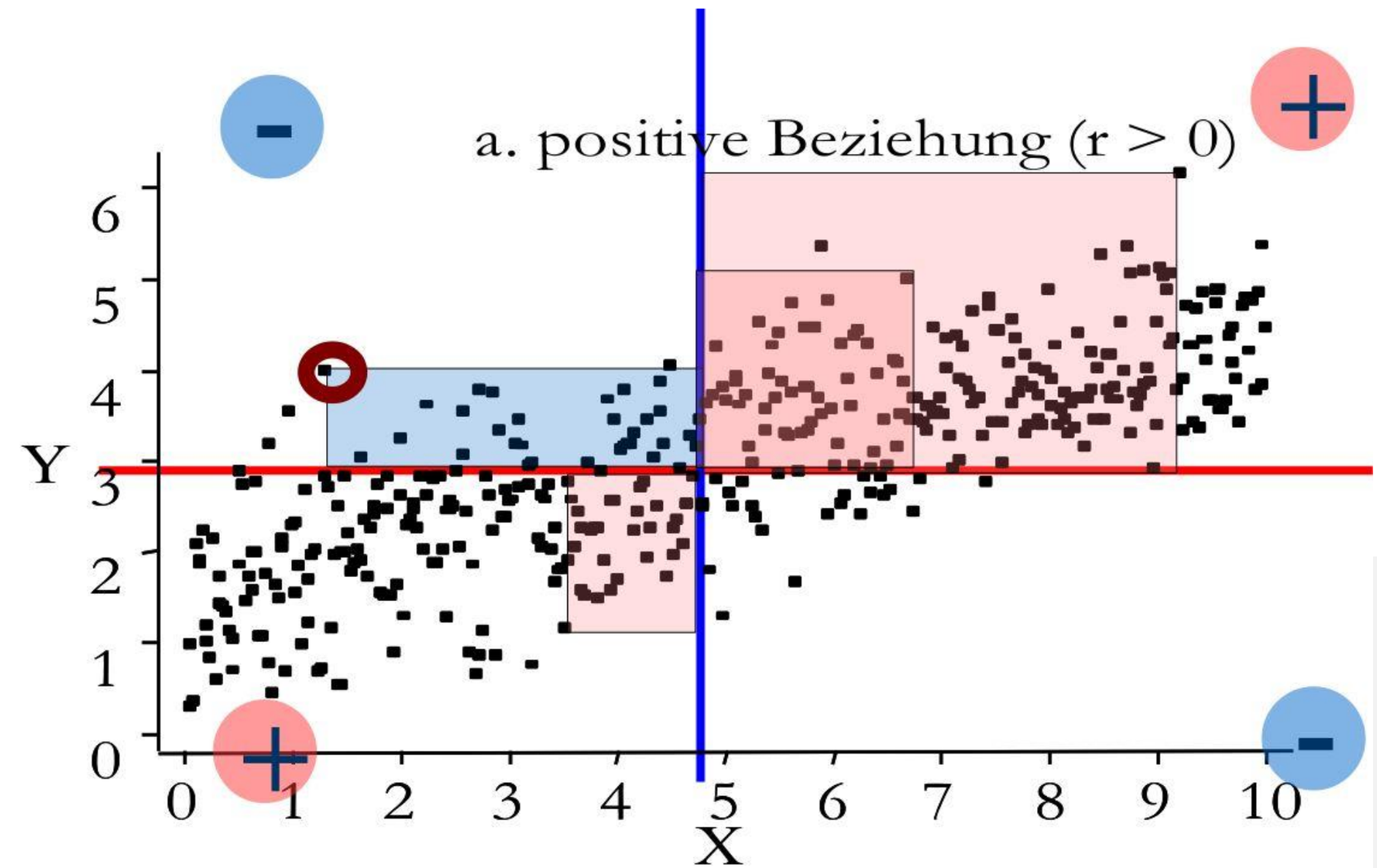
    Fall 1-3 haben nach Sortierung gleichen Rangplatz: (1+2+3)/3 = 2;
    Fall 5-6 haben nach Sortierung gleichen Rangplatz: (5+6)/2 = 5.5;
    Fall 7-10 haben nach Sortierung gleichen Rangplatz: (7+8+9+10)/4= 8,5
- Resultierende sortierte Rangwerte der 11 Fälle:

  {2,2,2,4,5.5,5.5,8.5,8.5,8.5,8.5,11}

Meine Notizen:

Zusammenhänge am Streudiagramm messen

- Es werden Rechtecke eingezeichnet, die die Differenzen zum Mittelwert abbilden:



Zusammenhang metrischer Variablen

- Der Zusammenhang wird zunächst gemessen als Summe der Flächen der Rechtecke.
- Diese Summe wird Kovariation genannt und mit SP abgekürzt.
- Sie berechnet sich nach:

$$SP_{XY} = \sum_{i=1}^{n}(x_i - \bar{x}) * (y_i - \bar{y})$$

- Die Kovarianz ist entsprechend:

$$s_{XY} = \frac{SP_{XY}}{n} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}) * (y_i - \bar{y}) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i * y_i\right) - \bar{x} * \bar{y}$$

Meine Notizen:

Korrelation (nach Pearson)

- Als Standardisierung wird die Kovarianz geteilt durch die Standardabweichungen der Variablen X und Y:

$$r_{XY} = \frac{s_{XY}}{s_X * s_Y} = \frac{SP_{XY}}{\sqrt{SS_X * SS_Y}}$$

- Diese Korrelationen können ihrer Größe nach bewertet werden. Dazu gelten folgende Faustregeln:

$-0.05 \leq r_{xy} < 0$	vernachlässigbare Korrelation	$0 < r_{xy} \leq 0{,}05$
$-0{,}20 \leq r_{xy} < -0{,}05$	geringe Korrelation	$0{,}05 < r_{xy} \leq 0{,}20$
$-0{,}50 \leq r_{xy} < -0{,}20$	mittlere Korrelation	$0{,}20 < r_{xy} \leq 0{,}50$
$-0{,}70 \leq r_{xy} < -0{,}50$	hohe Korrelation	$0{,}50 < r_{xy} \leq 0{,}70$
$-1{,}00 < r_{xy} < -0{,}70$	sehr hohe Korrelation	$0{,}70 < r_{xy} < 1{,}00$
$-1{,}00 = r_{xy}$	perfekte Korrelation	$1{,}00 = r_{xy}$

Meine Notizen:

Korrelation: Beispiel

$$r_{XY} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})\cdot(y_i - \bar{y})}{\sqrt{\left(\frac{1}{n}\cdot\sum_{i=1}^{n}(x_i - \bar{x})^2\right)\cdot\left(\frac{1}{n}\cdot\sum_{i=1}^{n}(y_i - \bar{y})^2\right)}}$$

$$= \frac{\left(\sum_{i=1}^{n} x_i \cdot y_i\right) - n\cdot\bar{x}\cdot\bar{y}}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - n\cdot\bar{x}^2\right)\cdot\left(\sum_{i=1}^{n} y_i^2 - n\cdot\bar{y}^2\right)}}$$

$$= \frac{\left(n\cdot\sum_{i=1}^{n} x_i \cdot y_i\right) - \left(\sum_{i=1}^{n} x_i\right)\cdot\left(\sum_{j=1}^{n} y_j\right)}{\sqrt{\left(n\cdot\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right)\cdot\left(n\cdot\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right)}}$$

$$= \frac{185\cdot 243078 - 6402\cdot 5977}{\sqrt{\left(185\cdot 264050 - 6402^2\right)\cdot\left(185\cdot 230417 - 5977^2\right)}}$$

$$= 0{,}91$$

X	Y	n	$\sum x_i$	$\sum y_i$	$\sum x_i^2$	$\sum y_i^2$	$\sum x_i \bullet y_i$
17	46	1	17	46	289	2116	782
18	17	1	18	17	324	289	306
19	16	1	19	16	361	256	304
19	18	2	38	36	722	648	684
19	19	4	76	76	1444	1444	1444
....	....	....	....	....	....	....	....
71	71	1	71	71	5041	5041	5041
72	65	1	72	65	5184	4225	4680
73	71	1	73	71	5329	5041	5183
73	72	1	73	72	5329	5184	5256
82	80	1	82	80	6724	6400	6560
Summe:		185	6402	5977	264050	230417	243078

Meine Notizen:

Spearmans Rang-Korrelation

- Für die Korrelation nennt sich das Verfahren nach Ansatz 3 Spearmans Rang-Korrelation = Produktmomentkorrelation zwischen zwei Variablen auf der Basis der Rangwerte aller Realisationen:

$$r_s = \frac{Kovariation\ nach\ Spearman}{\sqrt{Variation\ von\ X\ nach\ Spearman * Variation\ von\ Y\ nach\ Spearman}} =$$

$$\frac{\sum_{i=1}^{n} r_{X,i} * r_{Y,i} - \frac{\sum_{i=1}^{n} r_{X,i} * \sum_{i=1}^{n} r_{Y,i}}{n}}{\sqrt{\left(\sum_{i=1}^{n} r_{X,i}^2 - \frac{\left(\sum_{i=1}^{n} r_{X,i}\right)^2}{n}\right) * \left(\sum_{i=1}^{n} r_{Y,i}^2 - \frac{\left(\sum_{i=1}^{n} r_{Y,i}\right)^2}{n}\right)}}$$

wobei:

r_s = Spearmans Rang-Korrelationskeoffizient;

$r_{X,i}$ = Rangplatz von Fall i bei Variable X;

$r_{Y,i}$ = Rangplatz von Fall i bei Variable Y;

Meine Notizen:

Spearmans Rang-Korrelation: ein Beispiel

Beispiel:

Fall	X	Y	r_X	r_Y	r_X^2	r_Y^2	$r_X \cdot r_Y$
1	1	2	1.5	2.0	2.25	4.00	3.00
2	3	4	4.0	4.5	16.00	20.25	18.00
3	1	4	1.5	4.5	2.25	20.25	6.75
4	6	3	5.0	3.0	25.00	9.00	15.00
5	2	1	3.0	1.0	9.00	1.00	3.00
Summe:			15.0	15.0	54.50	54.50	45.75

$$r_s = \frac{\sum_{i=1}^{n} r_{X,i} * r_{Y,i} - \frac{\sum_{i=1}^{n} r_{X,i} * \sum_{i=1}^{n} r_{Y,i}}{n}}{\sqrt{\left(\sum_{i=1}^{n} r_{X,i}^2 - \frac{\left(\sum_{i=1}^{n} r_{X,i}\right)^2}{n}\right) * \left(\sum_{i=1}^{n} r_{Y,i}^2 - \frac{\left(\sum_{i=1}^{n} r_{Y,i}\right)^2}{n}\right)}}$$

$$= \frac{45{,}75 - \frac{15 * 15}{5}}{\sqrt{\left(54{,}5 - \frac{15^2}{5}\right) * \left(54{,}5 - \frac{15^2}{5}\right)}} = 0{,}079$$

Meine Notizen:

Berechnung von Rangwerten bei Kreuztabellen

Eigene wirtsch. Lage	Allgem. wirtschaftliche Lage (1)	(2)	(3)	(4)	(5)	$\sum$
Sehr gut (1)	48	74	32	5	1	160
Gut (2)	73	645	526	72	6	1322
Teils,teils (3)	12	143	361	115	16	647
Schlecht (4)	5	30	68	64	11	178
Sehr schlecht (5)	1	6	14	11	6	38
$\sum$	139	898	1001	267	40	2345
Kumul. Summen	139	1037	2038	2305	2345	
Rangplatz:	70	588.5	1538	2172	2325.5	

70=(1+139)/2; 588.5=(139+1+1037)/2; 1538=(1037+1+2038)/2; 2172=(2038+1+2305)/2; 2325.5=(2305+1+2345)/2

Kumul. Summen	Rang-platz	
160	80.5	= (1+160)/2
1482	821.5	= (160+1+1482)/2
2129	1806	= (1482+1+212)/2
2307	2218.5	= (2129+1+2307)/2
2345	2326.5	= (2307+1+2345)/2

1. Kumulieren der Häufigkeiten
2. Rangplatz einer Kategorie = (Anzahl kumul. Fälle bis zur vorhergehenden Kateg.+ 1 + Anzahl kumul. Fälle bis Kateg.) / 2

Meine Notizen:

Spearmans Rang-Korrelation an größeren Kreuztabellen

Eigene wirtsch. Lage	Allgem. wirtschaftliche Lage (1)	(2)	(3)	(4)	(5)	$\sum$
Sehr gut (1)	48	74	32	5	1	160
Gut (2)	73	645	526	72	6	1322
Teils,teils (3)	12	143	361	115	16	647
Schlecht (4)	5	30	68	64	11	178
Sehr schlecht (5)	1	6	14	11	6	38
$\sum$	139	898	1001	267	40	2345
Rangplatz	70	588.5	1538	2172	2325.5	

Rang-Platz	$\sum r_Y$	$\sum r^2_Y$
80.5	12880	1036840
821.5	1086023	892167894.5
1806	1168482	2110278492.
2218.5	394893	876070120.5
2326.5	88407	205678885.5
Summe:	2750685	4085232233.

$\sum r_X$ = 2750685 = n·(n+1)/2

$\sum r^2_X$ = 4155409843 = 139·70² + … + 40·2325.5²

$\sum r_X \cdot r_Y$ = 3587487381 = 48·80.5·70 + … + 6·2326.5·2325.5

$$r_s = \frac{Kovariation\ nach\ Spearman}{\sqrt{Variation\ X\ nach\ Spearman\ *\ Variation\ Y\ nach\ Spearman}} = \frac{\sum_{i=1}^{n} r_{X,i} * r_{Y,i} - \frac{\sum_{i=1}^{n} r_{X,i} * \sum_{i=1}^{n} r_{Y,i}}{n}}{\sqrt{\left(\sum_{i=1}^{n} r_{X,i}^2 - \frac{\left(\sum_{i=1}^{n} r_{X,i}\right)^2}{n}\right) * \left(\sum_{i=1}^{n} r_{Y,i}^2 - \frac{\left(\sum_{i=1}^{n} r_{Y,i}\right)^2}{n}\right)}}$$

$$= \frac{\sum_{i=1}^{n} r_{X,i} * r_{Y,i} - \frac{n * (n+1)^2}{4}}{\sqrt{\left(\sum_{i=1}^{n} r_{X,i}^2 - \frac{n * (n+1)^2}{4}\right) * \left(\sum_{i=1}^{n} r_{Y,i}^2 - \frac{n * (n+1)^2}{4}\right)}}$$

$$= \frac{3587487381 - \frac{2750685 * 2750685}{2345}}{\sqrt{\left(4155409843 - \frac{2750685^2}{2345}\right) * \left(4085232233 - \frac{2750685^2}{2345}\right)}} = 0{,}404$$

Meine Notizen:

Excel: Ordinale Zusammenhänge über Rangplätze messen

	A	B	C	D	E	F	G	H	I	J
1		Hauptschule	Mittlere Reife	Fachhochschulreife	Abitur	Summe	Kummulierte	Rangplatz	Rangsumme	Rangquadratsumme
2	Hauptschule	732	59	3	16	=SUMME(B2:E2)	=SUMME(F$	=(G2+1)/2	=F2*H2	=H2^2*F2
3	Mittlere Reife	534	192	19	83	=SUMME(B3:E3)	=SUMME(F$	=(G2+G3+1)/2	=F3*H3	=H3^2*F3
4	Fachhochschulreife	82	35	9	26	=SUMME(B4:E4)	=SUMME(F$	=(G3+G4+1)/2	=F4*H4	=H4^2*F4
5	Abiur	223	132	33	239	=SUMME(B5:E5)	=SUMME(F$	=(G4+G5+1)/2	=F5*H5	=H5^2*F5
6	Summe	=SUMME(B2:B5)	=SUMME(C2:C5)	=SUMME(D2:D5)	=SUMME(E2:E5)	=SUMME(F2:F5)			=SUMME(I2:I5)	=SUMME(J2:J5)
7	Kummulierte	=SUMME($B6:B6)	=SUMME($B6:C6)	=SUMME($B6:D6)	=SUMME($B6:E6)					
8	Rangplatz	=(B7+1)/2	=(B7+C7+1)/2	=(C7+D7+1)/2	=(D7+E7+1)/2					
9	Rangsumme	=B8*B6	=C8*C6	=D8*D6	=E8*E6	=SUMME(B9:E9)				
10	Rangquadratsumme	=B8^2*B6	=C8^2*C6	=D8^2*D6	=E8^2*E6	=SUMME(B10:E10)				
11							Kovariation	=F16-F9*I6/F6		
12		=B2*$H2*B$8	=C2*$H2*C$8	=D2*$H2*D$8	=E2*$H2*E$8		Variation X	=J6-I6^2/F6		
13		=B3*$H3*B$8	=C3*$H3*C$8	=D3*$H3*D$8	=E3*$H3*E$8		Variation Y	=F10-F9^2/F6		
14		=B4*$H4*B$8	=C4*$H4*C$8	=D4*$H4*D$8	=E4*$H4*E$8			=WURZEL(H12*H13)		
15		=B5*$H5*B$8	=C5*$H5*C$8	=D5*$H5*D$8	=E5*$H5*E$8		Korrelation	=H11/H14		
16					Kovariation	=SUMME(B12:E15)				
17										
18										
19		Hauptschule	Mittlere Reife	Fachhochschulreife	Abitur	Summe	Kummulierte	Rangplatz	Rangsumme	Rangquadratsumme
20	Hauptschule	732	59	3	16	810	810	405,5	328455	133188502,5
21	Mittlere Reife	534	192	19	83	828	1638	1224,5	1013886	1241503407
22	Fachhochschulreife	82	35	9	26	152	1790	1714,5	260604	446805558
23	Abiur	223	132	33	239	627	2417	2104	1319208	2775613632
24	Summe	1571	418	64	364	2417			2922153	4597111100
25	Kummulierte	1571	1989	2053	2417					
26	Rangplatz	786	1780,5	2021,5	2235,5					
27	Rangsumme	1234806	744249	129376	813722	2922153				
28	Rangquadratsumme	970557516	1325135345	261533584	1819075531	4376301976				
29							Kovariation	442732456,5		
30		233305236,00	42597572,25	2459154,75	14503924,00		Variation X	1064228123		
31		513952038,00	418602672,00	47031208,25	227201689,25		Variation Y	843418998,5		
32		110502954,00	106843353,75	31192755,75	99651883,50			947412379,7		
33		368784912,00	494494704,00	140356788,00	1124134588,00		Korrelation	0,467307021		
34						3975615433,50				

Bivariate Zusammenhänge nach Skalenniveaus

Symmetrisch

	Dichotome	Nominalskala	Ordinalskala	Metrische Skala
Metrische Skala			τ_a, τ_b, γ, r_s	Korrelation
Ordinalskala	τ_a, τ_b, γ, r_s		τ_a, τ_b, γ, r_s	
Nominalskala	Cramers V	Cramers V		
Dichotome V.	Phi, α			

Asym-metrisch

AV \ UV	Dichotome	Nominalskala	Ordinalskala	Metrische Skala
Metrische Skala	Mittelwert-vergleich, Regression		Sommers d	Regression
Ordinalskala	Sommers d, U-Tests		Sommers d	Sommers d
Nominalskala	Odds, λ, R´	Odds, λ, R´		
Dichotome V.	Prozentsatz-differenz, α	Odds, λ, R´	Sommers d	

- Asymmetrische Zusammenhänge gehen von einer Wirkung einer unabhängigen Variablen (UV) auf eine abhängige Variable (AV) aus.
- Symmetrische Zusammenhänge sind geeignet, wenn keine (einseitige) kausale Wirkung unterstellt werden soll oder Wirkung in beide Richtungen angenommen werden muss.

Meine Notizen:

Rangsummentest
(U-Test, Wilcoxon-Test, Mann-Whitney-Test)

- Rangsummentest = Alternative zum T-Test für zwei unabhängige Stichproben bei ordinalen abhängigen Variablen.
- Idee:

  Die Daten lassen sich auf Basis einer Dichotomen Variablen teilen in 2 unabhängige Stichproben des Umfangs n_1 und n_2. Für eine abhängige ordinale Variable Y gilt dann:

  Bei Sortierung aller Fälle nach Rangfolge (Größe) müsste der mittlere Rangplatz in beiden Stichproben tendenziell gleich sein!
- H_0: Lage der Rangverteilungen ist gleich

  (allgemeiner H_0: gleiche Mittelwerte der Ränge).

Meine Notizen:

Rangsummentest (U-Test)

Realisierung:

- Ränge sind von 1 bis $n = n_1 + n_2$ vergeben, bei gleichen Rängen (Ties) erhalten alle Fälle mit gleicher Ausprägung die mittlere Rangnummer
- Die Summe aller Ränge $n = n_1 + n_2$ ist $n*(n+1)/2$.
- Bei zufälliger Zuordnung der Rangnummer r_i zu einer der beiden Stichproben (Gruppen) g_j mit n_j Fällen ist der Erwartungswert jeder Rangsumme:

$$\mu\left(\sum_{i=1\in g_j}^{n_j} r_i \middle| H_0\right) = \frac{n_j}{n_1+n_2}\cdot\frac{n\cdot(n+1)}{2} \;;\; j=1,2$$

Meine Notizen:

Rangsummentest (U-Test)

- Varianz der Rangsummen in jeder der beiden Stichproben:

$$\sigma^2\left(\sum_{i=1\in g_j}^{n_j} r_i \middle| H_0\right) = \begin{cases} \dfrac{n_1 \cdot n_2 \cdot (n_1+n_2)}{12} & \text{keine Ties} \\ \left(\dfrac{n_1 \cdot n_2}{(n_1+n_2)\cdot(n_1+n_2-1)} \cdot \sum_{i=1}^{n_1+n_2} r_i^2\right) - \dfrac{n_1 \cdot n_2 \cdot (n_1+n_2+1)^2}{4\cdot(n_1+n_2-1)} & \text{mit Ties} \end{cases}$$

$$= \frac{n_1 \cdot n_2}{(n_1+n_2)} \cdot \frac{1}{n_1+n_2-1} \cdot \left(\sum_{i=1}^{n_1+n_2} r_i^2 - \frac{\left((n_1+n_2)\cdot(n_1+n_2+1)/2\right)^2}{(n_1+n_2)}\right)$$

- Entscheidungsregel:

  Wenn die Rangsumme in einer Stichprobe signifikant vom Erwartungswert abweicht, dann gilt H1:

  Stichproben kommen aus unterschiedlichen Verteilungen!

Meine Notizen:

Rangsummentest (U-Test): Beispiel Singleeinkommen

Geschlecht	n	$\sum r_i$	$\mu(\sum r_i)$	Differenz	Rangsumme²
männlich	155	21489	21002.5	486.5	3911719
weiblich	115	15096	15582.5	–486.5	2683722
Summe	270	36585	36585.0		6595441

Daten: Allbus 2012

Rangsumme insgesamt	= 270·271/2	= 36585
tatsächliche Rangsumme in Gruppe 1		= 21489
– μ(R.Summe|männlich & H_0)	= 155/270· 36585	= 21002.5
= Differenz		= 486.5
		→ > als bei H_0 erwartet!

$$\hat{\sigma}^2_{\text{Rangsumme}} = \frac{155\cdot 115}{270\cdot 269}\cdot\left(6595441 - \frac{36585^2}{270}\right) = 402043.8199$$

$$= \frac{155\cdot 115}{270\cdot 269}\cdot 6595441 - \frac{155\cdot 115\cdot 271^2}{4\cdot 269} = 402043.8199$$

$$\approx \frac{155\cdot 115\cdot 270}{12} = 402547.9167 \text{ (ohne Bindungen)}$$

Irrtumswahrscheinlichkeit: 5% , zweiseitiger Test: Z = +/- 1,960

Teststatistik: $\rightarrow Z = \frac{486.5}{\sqrt{402043.8199}} = 0.767,$ $\Rightarrow$ H_0: kein sign. Unterschied

→ Die höhere Rangsumme des Einkommen der Männer kann auch zufällig sein.

Meine Notizen:

Bivariate Zusammenhänge nach Skalenniveaus

Symmetrisch

	Dichotome	Nominalskala	Ordinalskala	Metrische Skala
Metrische Skala			τ_a, τ_b, γ, r_s	Korrelation
Ordinalskala	τ_a, τ_b, γ, r_s		τ_a, τ_b, γ, r_s	
Nominalskala	Cramers V	Cramers V		
Dichotome V.	Phi, α			

Asym-metrisch

AV \ UV	Dichotome	Nominalskala	Ordinalskala	Metrische Skala
Metrische Skala	Mittelwert-vergleich, Regression		Sommers d	Regression
Ordinalskala	Sommers d, U-Tests		Sommers d	Sommers d
Nominalskala	Odds, λ, R´	Odds, λ, R´		
Dichotome V.	Prozentsatz-differenz, α	Odds, λ, R´	Sommers d	

- Asymmetrische Zusammenhänge gehen von einer Wirkung einer unabhängigen Variablen (UV) auf eine abhängige Variable (AV) aus.
- Symmetrische Zusammenhänge sind geeignet, wenn keine (einseitige) kausale Wirkung unterstellt werden soll oder Wirkung in beide Richtungen angenommen werden muss.

Meine Notizen:

Ansatz 2: Aufbrechen des Zusammenhangs in Vielzahl von Paarvergleichen

- Wir gehen von einem Datensatz mit n Befragten aus.
- Wir interessieren uns für den Zusammenhang zwischen zwei ordinalen Variablen.
- Wir wählen 2 Befragte aus und vergleichen die Ausprägungen in X:

Ist: $x_i > x_j$ oder $x_i < x_j$ oder $x_i = x_j$?

- Wir schauen uns die Ausprägungen in Y an. Es liegt einer von 5 möglichen Paartypen vor:

- $x_i > x_j$ & $y_i > y_j$ $\rightarrow$ ***konkordantes*** (gleichgerichtetes) Paar;
- $x_i > x_j$ & $y_i < y_j$ $\rightarrow$ ***diskordantes*** (gegengerichtetes) Paar;
- $x_i = x_j$ & $y_i \neq y_j$ $\rightarrow$ ***x-verbundenes*** (engl.: tied x) Paar;
- $x_i \neq x_j$ & $y_i = y_j$ $\rightarrow$ ***y-verbundenes*** (engl.: tied y) Paar;
- $x_i = x_j$ & $y_i = y_j$ $\rightarrow$ ***x,y-verbundenes*** (engl.: tied x,y) Paar.

Wir müssen jedes mögliche Paar (es gibt 0,5*(n-1)*n) einem der 5 Typen zuordnen!

Meine Notizen:

Ansatz 2: Paarvergleiche am Beispiel

Eigene wirtsch. Lage (Y)	Allg. wirtschaftliche Lage (X) (1)	(2)	(3)	(4)	(5)
Sehr gut (1)	48	74	32	5	1
Gut (2)	73	645	526	72	6
Teils,teils (3)	12	143	361	115	16
Schlecht (4)	5	30	68	64	11
Sehr schlecht (5)	1	6	14	11	6

Bei gleicher Anordnung beider Variablen:

Konkordante Paare:	Fälle in Zellen rechts unterhalb
Diskordante Paare:	Fälle in Zellen links unterhalb
X-verbundene P.:	Fälle in Zellen gleicher Spalte
Y-verbundene P.:	Fälle in Zellen gleicher Zeile
X,Y-verbundene P.:	Fälle in gleicher Zelle

Fall i: X=2 Y=1
Fall j: X=4 Y=3
→ konkordant

Fall i: X=2 Y=1
Fall j: X=1 Y=3
→ diskordant

Fall i: X=2 Y=1
Fall j: X=2 Y=4
→ x-verbunden

Fall i: X=2 Y=1
Fall j: X=5 Y=1
→ y-verbunden

Fall i: X=2 Y=1
Fall j: X=2 Y=1
→ x,y-verbunden

Meine Notizen:

Ansatz 2: Paarvergleiche am Beispiel

Eigene wirtsch. Lage (Y)	Allg. wirtschaftliche Lage (X) (1)	(2)	(3)	(4)	(5)
Sehr gut (1)	48	74	32	5	1
Gut (2)	73	645	526	72	6
Teils,teils (3)	12	143	361	115	16
Schlecht (4)	5	30	68	64	11
Sehr schlecht (5)	1	6	14	11	6

Bei gleicher Anordnung beider Variablen:

Konkordante Paare:	Fälle in Zellen rechts unterhalb
Diskordante Paare:	Fälle in Zellen links unterhalb
X-verbundene P.:	Fälle in Zellen gleicher Spalte
Y-verbundene P.:	Fälle in Zellen gleicher Zeile
X,Y-verbundene P.:	Fälle in gleicher Zelle

Anzahl konkordanter Paare C:

C = 48·(645+526+72+6 +143+361+115+16 +30+68+64+11 +6+14+11+6)
+ 74·(526+72+6 +361+115+16 +68+64+11 +14+11+6)
+ 32·(72+6 +115+16 +64+11+11+6) + 5·(6+16+11+6)
+ 73·(143+361+115+16 +30+68+64+11 +6+14+11+6)
+ 645·(361+115+16 +68+64+11 +14+11+6)
+ 526·(115+16+64+11+11+6) + 72·(16 +11+6)
+ 12·(30+68+64+11 +6+14+11+6) +143·(68+64+11 +14+11+6)
+ 361·(64+11+11+6) + 115·(11+6)
+5·(6+14+11+6) +30·(14+11+6) +68·(11+6) +64·6

= 880472

Meine Notizen:

Ansatz 2: Paarvergleiche am Beispiel

Eigene wirtsch. Lage (Y)	Allg. wirtschaftliche Lage (X) (1)	(2)	(3)	(4)	(5)
Sehr gut (1)	48	74	32	5	1
Gut (2)	73	645	526	72	6
Teils,teils (3)	12	143	361	115	16
Schlecht (4)	5	30	68	64	11
Sehr schlecht (5)	1	6	14	11	6

Bei gleicher Anordnung beider Variablen:

Konkordante Paare:	Fälle in Zellen rechts unterhalb
Diskordante Paare:	Fälle in Zellen links unterhalb
X-verbundene P.:	Fälle in Zellen gleicher Spalte
Y-verbundene P.:	Fälle in Zellen gleicher Zeile
X,Y-verbundene P.:	Fälle in gleicher Zelle

Anzahl diskordanter Paare D:

$$
\begin{aligned}
D = {} & 74\cdot(73+12+5+1) +32\cdot(73+645+12+143+5+30+1+6) \\
& +5\cdot(73+645+526+12+143+361+5+30+68+1+6+14) \\
& +1\cdot(73+645+526+72+12+143+361+115+5+30+68+64+1+6+14+11) \\
& +645\cdot(12+5+1) +526\cdot(12+143+5+30+1+6) +72\cdot(12+143+361+5+30+68+1+6+14) \\
& +6\cdot(12+143+361+115+5+30+68+64+1+6+14+11) \\
& +143\cdot(5+1) +361\cdot(5+30+1+6) +115\cdot(5+30+68+1+6+14) \\
& +16\cdot(5+30+68+64+1+6+14+11) \\
& +30\cdot 1 +68\cdot(1+6) +64\cdot(1+6+14) +11\cdot(1+6+14+11)
\end{aligned}
$$

$= 249538$

Meine Notizen:

Ansatz 2: Paarvergleiche am Beispiel

Eigene wirtsch. Lage (Y)	Allg. wirtschaftliche Lage (X) (1)	(2)	(3)	(4)	(5)
Sehr gut (1)	48	74	32	5	1
Gut (2)	73	645	526	72	6
Teils,teils (3)	12	143	361	115	16
Schlecht (4)	5	30	68	64	11
Sehr schlecht (5)	1	6	14	11	6

Bei gleicher Anordnung beider Variablen:

Konkordante Paare:	Fälle in Zellen rechts unterhalb
Diskordante Paare:	Fälle in Zellen links unterhalb
X-verbundene P.:	Fälle in Zellen gleicher Spalte
Y-verbundene P.:	Fälle in Zellen gleicher Zeile
X,Y-verbundene P.:	Fälle in gleicher Zelle

Anzahl x-verbundener Paare Tx:

$$
\begin{aligned}
T_X = \; & 48\cdot(73+12+5+1) +73\cdot(12+5+1) +12\cdot(5+1) +5\cdot 1 \\
& +74\cdot(645+143+30+6) +645\cdot(143+30+6) +143\cdot(30+6) +30\cdot 6 \\
& +32\cdot(526+361+68+14) +526\cdot(361+68+14) +361\cdot(68+14)+68\cdot 14 \\
& +5\cdot(72+115+64+11) +72\cdot(115+64+11) +115\cdot(64+11) +64\cdot 11 \\
& +1\cdot(6+16+11+6) +6\cdot(16+11+6) +16\cdot(11+6) +11\cdot 6 \\
= \; & 506992
\end{aligned}
$$

Meine Notizen:

Ansatz 2: Paarvergleiche am Beispiel

Eigene wirtsch. Lage (Y)	Allg. wirtschaftliche Lage (X) (1)	(2)	(3)	(4)	(5)
Sehr gut (1)	48	74	32	5	1
Gut (2)	73	645	526	72	6
Teils,teils (3)	12	143	361	115	16
Schlecht (4)	5	30	68	64	11
Sehr schlecht (5)	1	6	14	11	6

Bei gleicher Anordnung beider Variablen:

Konkordante Paare:	Fälle in Zellen rechts unterhalb
Diskordante Paare:	Fälle in Zellen links unterhalb
X-verbundene P.:	Fälle in Zellen gleicher Spalte
Y-verbundene P.:	Fälle in Zellen gleicher Zeile
X,Y-verbundene P.:	Fälle in gleicher Zelle

Anzahl x,y-verbundener Paare Txy:

$$
\begin{aligned}
T_{XY} = \; & 48\cdot(48-1)/2 + 74\cdot(74-1)/2 + 32\cdot(32-1)/2 + 5\cdot(5-1)/2 + 1\cdot(1-1)/2 \\
& + 73\cdot(73-1)/2 + 645\cdot(645-1)/2 + 526\cdot(526-1)/2 + 72\cdot(72-1)/2 + 6\cdot(6-1)/2 \\
& + 12\cdot(12-1)/2 + 143\cdot(143-1)/2 + 361\cdot(361-1)/2 + 115\cdot(115-1)/2 + 16\cdot(16-1)/2 \\
& + 5\cdot(5-1)/2 + 30\cdot(30-1)/2 + 68\cdot(68-1)/2 + 64\cdot(64-1)/2 + 11\cdot(11-1)/2 \\
& + 1\cdot(1-1)/2 + 6\cdot(6-1)/2 + 14\cdot(14-1)/2 + 11\cdot(11-1)/2 + 6\cdot(6-1)/2 \\
= \; & 442143
\end{aligned}
$$

Meine Notizen:

Ansatz 2: Paarvergleiche am Beispiel

Eigene wirtsch. Lage (Y)	Allg. wirtschaftliche Lage (X) (1)	(2)	(3)	(4)	(5)
Sehr gut (1)	48	74	32	5	1
Gut (2)	73	645	526	72	6
Teils,teils (3)	12	143	361	115	16
Schlecht (4)	5	30	68	64	11
Sehr schlecht (5)	1	6	14	11	6

Bei gleicher Anordnung beider Variablen:

Konkordante Paare:	Fälle in Zellen rechts unterhalb
Diskordante Paare:	Fälle in Zellen links unterhalb
X-verbundene P.:	Fälle in Zellen gleicher Spalte
Y-verbundene P.:	Fälle in Zellen gleicher Zeile
X,Y-verbundene P.:	Fälle in gleicher Zelle

Anzahl Paarvergleiche insgesamt:

Anzahl konkordanter Paare	C	=	880472
Anzahl diskordanter Paare	D	=	249538
Anzahl x-verbundener Paare	T_X	=	506992
Anzahl y-verbundener Paare	T_Y	=	669195
Anzahl x,y-verbundener Paare	T_{XY}	=	442143
	Summe	=	2748340

$= n\cdot(n-1)/2 = 2345\cdot 2344/2 = 2748340$

Meine Notizen:

Ansatz 2: Paarvergleiche am Beispiel

Eigene wirtsch. Lage (Y)	Allg. wirtschaftliche Lage (X) (1)	(2)	(3)	(4)	(5)
Sehr gut (1)	48	74	32	5	1
Gut (2)	73	645	526	72	6
Teils,teils (3)	12	143	361	115	16
Schlecht (4)	5	30	68	64	11
Sehr schlecht (5)	1	6	14	11	6

Bei gleicher Anordnung beider Variablen:

Konkordante Paare:	Fälle in Zellen rechts unterhalb
Diskordante Paare:	Fälle in Zellen links unterhalb
X-verbundene P.:	Fälle in Zellen gleicher Spalte
Y-verbundene P.:	Fälle in Zellen gleicher Zeile
X,Y-verbundene P.:	Fälle in gleicher Zelle

Berechnungsformeln:

$$C = \sum_{i=1}^{I-1}\sum_{j=1}^{J-1} n_{i,j} \cdot \left(\sum_{k=i+1}^{I}\sum_{m=j+1}^{J} n_{k,m} \right); \; D = \sum_{i=1}^{I-1}\sum_{j=2}^{J} n_{i,j} \cdot \left(\sum_{k=i+1}^{I}\sum_{m=1}^{j-1} n_{k,m} \right)$$

$$T_X = \sum_{i=1}^{I}\sum_{j=1}^{J-1} n_{i,j} \cdot \left(\sum_{k=j+1}^{J} n_{i,k} \right); \; T_Y = \sum_{j=1}^{J}\sum_{i=1}^{I-1} n_{i,j} \cdot \left(\sum_{k=i+1}^{I} n_{k,j} \right); \; T_{X,Y} = \sum_{i=1}^{I}\sum_{j=1}^{J} \frac{n_{i,j} \cdot \left(n_{i,j} - 1 \right)}{2}$$

Meine Notizen:

Aus all diesen Paarvergleichen lassen sich Zusammenhangsmaße ableiten:

Idee:

Spricht ein Paarvergleich eher für einen positiven Zusammenhang, eher für einen negativen Zusammenhang oder gegen einen monotonen Zusammenhang?

Paarvergleiche am Beispiel: Sommers d

Eigene wirtsch. Lage	Allgem. wirtschaftliche Lage (1)	(2)	(3)	(4)	(5)
Sehr gut (1)	48	74	32	5	1
Gut (2)	73	645	526	72	6
Teils,teils (3)	12	143	361	115	16
Schlecht (4)	5	30	68	64	11
Sehr schlecht (5)	1	6	14	11	6

Anzahl konkordanter Paare	C	=	880472
Anzahl diskordanter Paare	D	=	249538
Anzahl x-verbundener Paare	T_X	=	506992
Anzahl y-verbundener Paare	T_Y	=	669195
Anzahl x,y-verbundener Paare	T_{XY}	=	442143
	Summe	=	2748340

Asymmetrische Beziehung von X auf Y: Somers d_{YX}

- C spricht für positive, D für negative und T_Y gegen Beziehung von X auf Y

$$\rightarrow d_{YX} = \frac{C-D}{C+D+T_Y} \qquad d_{YX} = \frac{C-D}{C+D+T_y} = \frac{880472-249538}{880472+249538+669195} = 0.351$$

Asymmetrische Beziehung von Y auf X: Somers d_{XY}

- C spricht für positive, D für negative und T_X gegen Beziehung von Y auf X

$$\rightarrow d_{XY} = \frac{C-D}{C+D+T_X} \qquad d_{XY} = \frac{C-D}{C+D+T_x} = \frac{880472-249538}{880472+249538+506992} = 0.385$$

Meine Notizen:

Paarvergleiche am Beispiel: Tau-b

Eigene wirtsch. Lage	Allgem. wirtschaftliche Lage (1)	(2)	(3)	(4)	(5)
Sehr gut (1)	48	74	32	5	1
Gut (2)	73	645	526	72	6
Teils,teils (3)	12	143	361	115	16
Schlecht (4)	5	30	68	64	11
Sehr schlecht (5)	1	6	14	11	6

Anzahl konkordanter Paare	C	=	880472
Anzahl diskordanter Paare	D	=	249538
Anzahl x-verbundener Paare	T_X	=	506992
Anzahl y-verbundener Paare	T_Y	=	669195
Anzahl x,y-verbundener Paare	T_{XY}	=	442143
	Summe	=	2748340

$$\rightarrow |\tau_b| = \sqrt{d_{YX} \cdot d_{XY}} = \sqrt{\frac{C-D}{C+D+T_X} \cdot \frac{C-D}{C+D+T_Y}}; \quad \tau_b = \frac{C-D}{\sqrt{(C+D+T_X)\cdot(C+D+T_Y)}}$$

$$\tau_b = \frac{C-D}{\sqrt{(C+D+T_X)\cdot(C+D+T_y)}} = \frac{880472-249538}{\sqrt{(880472+249538+669195)\cdot(880472+249538+506992)}} = 0.368$$

Meine Notizen:

Symmetrische Zusammenhänge: Alternativen

- Bei empirisch möglichen vollständigen Rangreihen sprechen T_X, T_Y und $T_{X,Y}$ gegen eine monotone Beziehung. Das berücksichtigt Tau-a:

$$\rightarrow \tau_a = \frac{C-D}{C+D+T_Y+T_X+T_{X,Y}} = \frac{C-D}{n\cdot(n-1)/2} \qquad \tau_a = \frac{C-D}{n\cdot(n-1)/2} = \frac{880472-249538}{2748340} = 0.230$$

- Gamma vergleicht nur konkordante und diskordante Paare:

$$\rightarrow \gamma = \frac{C-D}{C+D} \qquad \gamma = \frac{C-D}{C+D} = \frac{880472-249538}{880472+249538} = 0.558$$

- Als Regeln gelten:

$$1 \geq |\gamma| \geq |\tau_b| \geq |\tau_a| \geq 0 \text{ und } |\tau_b| = \sqrt{d_{YX}\cdot d_{XY}}$$

Meine Notizen:

Excel: Ordinale Zusammenhänge über Paarvergleiche messen

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		Hauptschule	Mittlere Reife	Fachhochschulreife	Abitur	Summe				Hauptschule	Mittlere Reife	Fachhochschulreif	Abitur	Summe
2	Hauptschule	732	59	3	16	810			Hauptschule	732	59	3	16	810
3	Mittlere Reife	534	192	19	83	828			Mittlere Reife	534	192	19	83	828
4	Fachhochschulreife	82	35	9	26	152			Fachhochschulreife	82	35	9	26	152
5	Abiur	223	132	33	239	627			Abiur	223	132	33	239	627
6	Summe	1571	418	64	364	2417			Summe	1571	418	64	364	2417
7														
8	C	=B2*SUMME(C3:$E$5)	=C2*SUMME(D3:$E$5)	=D2*SUMME(E3:$E$5)					C	562176	24131	1044		
9		=B3*SUMME(C4:$E$5)	=C3*SUMME(D4:$E$5)	=D3*SUMME(E4:$E$5)						253116	58944	5035		
10		=B4*SUMME(C5:$E$5)	=C4*SUMME(D5:$E$5)	=D4*SUMME(E5:$E$5)						33128	9520	2151		
11						=SUMME(B8:D10)								949245
12														
13	D		=C2*SUMME(B3:$B$5)	=D2*SUMME($B3:C$5)	=E2*SUMME($B3:D$5)				D		49501	3594	20144	
14			=C3*SUMME(B4:$B$5)	=D3*SUMME($B4:C$5)	=E3*SUMME($B4:D$5)						58560	8968	42662	
15			=C4*SUMME(B5:$B$5)	=D4*SUMME($B5:C$5)	=E4*SUMME($B5:D$5)						7805	3195	10088	
16						=SUMME(C13:E15)								204517
17														
18	Tx	=B2*SUMME(B3:B$5)	=C2*SUMME(C3:C$5)	=D2*SUMME(D3:D$5)	=E2*SUMME(E3:E$5)				Tx	614148	21181	183	5568	
19		=B3*SUMME(B4:B$5)	=C3*SUMME(C4:C$5)	=D3*SUMME(D4:D$5)	=E3*SUMME(E4:E$5)					162870	32064	798	21995	
20		=B4*SUMME(B5:B$5)	=C4*SUMME(C5:C$5)	=D4*SUMME(D5:D$5)	=E4*SUMME(E5:E$5)					18286	4620	297	6214	
21						=SUMME(B18:E20)								888224
22														
23	Ty	=B2*SUMME(C2:$E2)	=C2*SUMME(D2:$E2)	=D2*SUMME(E2:$E2)					Ty	57096	1121	48		
24		=B3*SUMME(C3:$E3)	=C3*SUMME(D3:$E3)	=D3*SUMME(E3:$E3)						156996	19584	1577		
25		=B4*SUMME(C4:$E4)	=C4*SUMME(D4:$E4)	=D4*SUMME(E4:$E4)						5740	1225	234		
26		=B5*SUMME(C5:$E5)	=C5*SUMME(D5:$E5)	=D5*SUMME(E5:$E5)						90092	35904	7887		
27						=SUMME(B23:D26)								377504
28														
29	Txy	=B2*(B2-1)/2	=C2*(C2-1)/2	=D2*(D2-1)/2	=E2*(E2-1)/2				Txy	267546	1711	3	120	
30		=B3*(B3-1)/2	=C3*(C3-1)/2	=D3*(D3-1)/2	=E3*(E3-1)/2					142311	18336	171	3403	
31		=B4*(B4-1)/2	=C4*(C4-1)/2	=D4*(D4-1)/2	=E4*(E4-1)/2					3321	595	36	325	
32		=B5*(B5-1)/2	=C5*(C5-1)/2	=D5*(D5-1)/2	=E5*(E5-1)/2					24753	8646	528	28441	
33						=SUMME(B29:E32)								500246
34														
35		Tau-b	=WURZEL(F35*F36)		Sommers dyx	=(F11-F16)/(F11+F16+F27)				Tau-b	0,4211589		Sommers dyx	0,4863479
36		Tau-a	=(F11-F16)/(F11+F16+F21+F27+F33)		Sommers dxy	=(F11-F16)/(F11+F16+F21)				Tau-a	0,255066896		Sommers dxy	0,36470769
37		Gamma	=(F11-F16)/(F11+F16)							Gamma	0,64547801			
38														

Zusammenhänge zwischen ordinal und metrisch

- Ansatz 3: Analyse von Rangplätzen statt von Ausprägungen
  - Statt gleichen Abständen zwischen den Ausprägungen wird gleicher Abstand zwischen den Rangplätzen angenommen (kennen Sie vom IQR).
- In Ansatz 3 werden beide ordinalen Variablen umkodiert, sodass Rangplätze die neuen Ausprägungen werden. Diese Ausprägungen werden dann wie metrisch behandelt.
- Setzt man eine ordinale und eine metrische Variable in Beziehung bedient man sich dieser Logik nur für die ordinale Variable und erhält die metrische, deren Abstände ja bereits sinnvoll sind.
- ! In der Interpretation sind die Einheiten zu beachten. Die metrische Variable liegt noch im Original vor, die ordinale muss in Rängen interpretiert werden !

Meine Notizen:

Bivariate Zusammenhänge nach Skalenniveaus

Symmetrisch

	Dichotome	Nominalskala	Ordinalskala	Metrische Skala
Metrische Skala			τ_a, τ_b, γ, r_s	Korrelation
Ordinalskala	τ_a, τ_b, γ, r_s		τ_a, τ_b, γ, r_s	
Nominalskala	Cramers V	Cramers V		
Dichotome V.	Phi, α			

Asym-metrisch

AV \ UV	Dichotome	Nominalskala	Ordinalskala	Metrische Skala
Metrische Skala	Mittelwert-vergleich, Regression		Sommers d	Regression
Ordinalskala	Sommers d, U-Tests		Sommers d	Sommers d
Nominalskala	Odds, λ, R´	Odds, λ, R´		
Dichotome V.	Prozentsatz-differenz, α	Odds, λ, R´	Sommers d	

- Asymmetrische Zusammenhänge gehen von einer Wirkung einer unabhängigen Variablen (UV) auf eine abhängige Variable (AV) aus.
- Symmetrische Zusammenhänge sind geeignet, wenn keine (einseitige) kausale Wirkung unterstellt werden soll oder Wirkung in beide Richtungen angenommen werden muss.

Meine Notizen:

Zusammenhänge zwischen nominal und ordinal/metrisch

- Ist eine Variable in Nominalskala und die Zweite in einem höheren Skalenniveau gemessen, liefern die nominalen Maße Cramers V und McFaddens Pseudo R´ akzeptable Ergebnisse.
- Die einzige Alternative liegt darin, aus der nominalen Skala eine Dichotome Variable zu machen. Dazu gibt es zwei Wege:
  1. Man wählt zwei der Ausprägungen, vergleicht nur diese und ignoriert alle weiteren Fälle.
  2. Man wählt eine Ausprägung und fasst alle Weiteren zur Vergleichsgruppe zusammen.
- In beiden Varianten wird entweder theoriebasiert entschieden, welche Ausprägung(en) für einen Vergleich interessant sind, oder es wird eine Vielzahl von Vergleichen angestellt, die dann gemeinsam interpretiert werden.

Meine Notizen:

Bivariate Zusammenhänge nach Skalenniveaus

Symmetrisch

	Dichotome	Nominalskala	Ordinalskala	Metrische Skala
Metrische Skala	Statistik III	Cramers V, Dichotomi-sieren	τ_a, τ_b, γ, r_s	Korrelation
Ordinalskala	τ_a, τ_b, γ, r_s	Cramers V, Dichotomi-sieren	τ_a, τ_b, γ, r_s	
Nominalskala	Cramers V	Cramers V		
Dichotome V.	Phi, α			

Asym-metrisch

AV \ UV	Dichotome	Nominalskala	Ordinalskala	Metrische Skala
Metrische Skala	Mittelwert-vergleich, Regression	R´, Dicho-tomisieren	Sommers d	Regression
Ordinalskala	Sommers d, U-Tests	R´, Dicho-tomisieren	Sommers d	Sommers d
Nominalskala	Odds, λ, R´	Odds, λ, R´	R´, Dicho-tomisieren	R´, Dicho-tomisieren
Dichotome V.	Prozentsatz-differenz, α	Odds, λ, R´	Sommers d	Statistik III

- Asymmetrische Zusammenhänge gehen von einer Wirkung einer unabhängigen Variablen (UV) auf eine abhängige Variable (AV) aus.
- Symmetrische Zusammenhänge sind geeignet, wenn keine (einseitige) kausale Wirkung unterstellt werden soll oder Wirkung in beide Richtungen angenommen werden muss.

Meine Notizen:

Was Sie am Ende der Woche können sollten

- **Kern**: Sie beschreiben und berechnen Zusammenhänge für zwei beliebige Variablen.
- Sie bestimmen die Rangplätze ordinaler Daten (in Kreuztabellen).
- Sie nutzen Rangplätze zur Berechnung sonst metrischer Zusammenhangsmaße.
- Insbesondere berechnen Sie Spearmans Rangkorrelation.
- Sie nutzen Rangplätze in Mittelwertvergleichen im Sinne eines U-Tests.
- Sie verwenden die Idee von Rangplätzen in der Analyse von Zusammenhängen zwischen einer ordinalen und einer metrischen Variablen.
- Sie dichotomisieren nominale Variablen theoriebasiert, um Zusammenhänge mit höheren Skalenniveaus zu messen.

Meine Notizen: