Bivariate Zusammenhangsmaße

	Metrisch	Ordinal		Nominal	
		singulär	Rangklassen	Dichotom	Polychotom
Metrisch	Produkt-Moment-Korrelation r	Kendalls τ; Wilsons e	Polyseriale Korrelation	(Punkt-) biseriale Korrelation	Koeffizient η
Ordinal (singulär)		Kendalls τ; Wilsons e	Somers' d_{YX}; Kims $d_{Y\bullet X}$	Rangbiseriale Korrelation; Somers' d_{YX}; Kims $d_{Y\bullet X}$	Rangklassen erstellen
Ordinal (Rangklassen)			Koeffizient γ	Koeffizient γ	Cramérs V; Multiples R^2 einer Probit-Regression
Nominal (Dichotom)				ϕ-Koeffiz.; Yules Q	Cramérs V
Nominal (Polychotom)					Cramérs V

Tabelle aus: Eid, Gollwitzer & Schmitt (2017), S. 569

Geordnete Antwortkategorien

- Hängt die Wahlabsicht mit der Zufriedenheit mit der aktuellen Regierung zusammen?

Kreuztabelle

- Häufigkeitstabelle mit zwei Dimensionen
- Beispiel hier: Zufriedenheit mit der aktuellen Regierung (Zeilen) und Absicht, an der Wahl teilzunehmen (Spalten)

	nein	unentschlossen	ja
unzufrieden	11	22	33
teils-teils	8	34	45
zufrieden	7	14	49

Zusammenhänge bei geordneten Antwortkategorien

- Kovarianz und Korrelation basieren auf Mittelwerten und Varianzen $\rightarrow$ sind für ordinalskalierte Variablen ungeeignet
- Koeffizient γ basiert auf der Strukturierung aller paarweisen Vergleiche

$$\hat{\gamma} = \frac{n_K}{n_K + n_D} - \frac{n_D}{n_K + n_D} = \frac{n_K - n_D}{n_K + n_D}$$

$\hat{\gamma}$ Geschätzter Zusammenhang

n_K Anzahl konkordanter Paare

n_D Anzahl diskordanter Paare

Zusammenhänge bei geordneten Antwortkategorien

	nein	unentschlossen	ja
unzufrieden	11	22	33
teils-teils	8	34	45
zufrieden	7	14	49

Zusammenhänge bei geordneten Antwortkategorien

	nein	unentschlossen	ja
unzufrieden	11	22	33
teils-teils	8	34	45
zufrieden	7	14	49

Konkordante Paare (n_K)

- Paare, bei denen $x_m > x_{m'}$ und $y_m > y_{m'}$, bzw. $x_m < x_{m'}$ und $y_m < y_{m'}$
- Als gleichsinnige Wertekombination bezeichnet

Zusammenhänge bei geordneten Antwortkategorien

	nein	unentschlossen	ja
unzufrieden	11	22	33
teils-teils	8	34	45
zufrieden	7	14	49

Konkordante Paare (n_K)

- Paare, bei denen $x_m > x_{m'}$ und $y_m > y_{m'}$, bzw. $x_m < x_{m'}$ und $y_m < y_{m'}$
- Als gleichsinnige Wertekombination bezeichnet

Diskordante Paare (n_D)

- Paare, bei denen $x_m > x_{m'}$ aber $y_m < y_{m'}$, bzw. $x_m < x_{m'}$ aber $y_m > y_{m'}$
- Als gegensinnige Wertekombination bezeichnet

Zusammenhänge bei geordneten Antwortkategorien

	nein	unentschlossen	ja
unzufrieden	11	22	33
teils-teils	8	34	45
zufrieden	7	14	49

Konkordante Paare (n_K)

- Paare, bei denen $x_m > x_{m'}$ und $y_m > y_{m'}$, bzw. $x_m < x_{m'}$ und $y_m < y_{m'}$
- Als gleichsinnige Wertekombination bezeichnet

Diskordante Paare (n_D)

- Paare, bei denen $x_m > x_{m'}$ aber $y_m < y_{m'}$, bzw. $x_m < x_{m'}$ aber $y_m > y_{m'}$
- Als gegensinnige Wertekombination bezeichnet

Rangbindungen

- Paare, bei denen $x_m = x_{m'}$ oder $y_m = y_{m'}$

Anzahl der Paare

	nein	unentschlossen	ja
unzufrieden	11	22	33
teils-teils	8	34	45
zufrieden	7	14	49

$$n_K = \underbrace{\sum_{i=1}^{k-1} \sum_{j=1}^{l-1} n_{ij}}_{\text{Relevante Zellen der Tabelle}} \cdot \underbrace{\left(\sum_{a=i+1}^{k} \sum_{b=j+1}^{l} n_{ab} \right)}_{\text{Jede konkordante Zelle}}$$

- Für $i = 1$, $j = 2$ (blau markierte Zelle)

$$n_K = 22 \cdot (45 + 49) = 2068$$

Anzahl der Paare

	nein	unentschlossen	ja
unzufrieden	11	22	33
teils-teils	8	34	45
zufrieden	7	14	49

$$n_D = \underbrace{\sum_{i=1}^{k-1}\sum_{j=2}^{l} n_{ij}}_{\text{Relevante Zellen der Tabelle}} \cdot \underbrace{\left(\sum_{a=i+1}^{k}\sum_{b=1}^{j-1} n_{ab}\right)}_{\text{Jede diskordante Zelle}}$$

- Für $i = 1$, $j = 2$ (blau markierte Zelle)

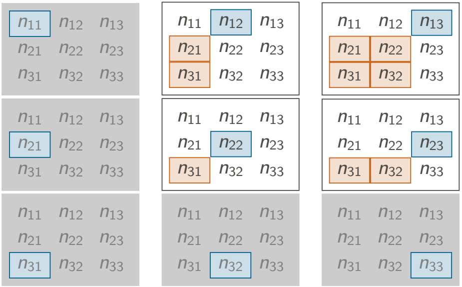
$$n_D = 22 \cdot (8 + 7) = 330$$

Vorgehen zum Zählen der Paare

Anzahl konkordante Paare

$$n_K = \underbrace{\sum_{i=1}^{k-1}\sum_{j=1}^{l-1} n_{ij}}_{\text{Relevante Zellen der Tabelle}} \cdot \underbrace{\left(\sum_{a=i+1}^{k}\sum_{b=j+1}^{l} n_{ab}\right)}_{\text{Jede konkordante Zelle}}$$

Anzahl diskordante Paare



$$n_D = \underbrace{\sum_{i=1}^{k-1}\sum_{j=2}^{l} n_{ij}}_{\text{Relevante Zellen der Tabelle}} \cdot \underbrace{\left(\sum_{a=i+1}^{k}\sum_{b=1}^{j-1} n_{ab}\right)}_{\text{Jede diskordante Zelle}}$$

Zusammenhangsmaß

	nein	unentschlossen	ja
unzufrieden	11	22	33
teils-teils	8	34	45
zufrieden	7	14	49

$$\hat{\gamma} = \frac{n_K}{n_K + n_D} - \frac{n_D}{n_K + n_D} = \frac{n_K - n_D}{n_K + n_D}$$
$$= \frac{5800 - 3592}{9392} = 0.235$$

Interpretation von γ

- Wertebereich $[-1; 1]$
- $\gamma = -1$, wenn alle (nicht gebundenen) Paare diskordant sind
- $\gamma = 0$, wenn gleich viele konkordante und diskordante Paare vorliegen
- $\gamma = 1$, wenn alle (nicht gebundenen) Paare konkordant sind

Bivariate Zusammenhangsmaße

	Metrisch	**Ordinal**		**Nominal**	
		singulär	Rangklassen	Dichotom	Polychotom
Metrisch	Produkt-Moment-Korrelation r	Kendalls τ; Wilsons e	Polyseriale Korrelation	(Punkt-) biseriale Korrelation	Koeffizient η
Ordinal (singulär)		Kendalls τ; Wilsons e	Somers' d_{YX}; Kims $d_{Y \bullet X}$	Rangbiseriale Korrelation; Somers' d_{YX}; Kims $d_{Y \bullet X}$	Rangklassen erstellen
Ordinal (Rangklassen)			Koeffizient γ	Koeffizient γ	Cramérs V; Multiples R^2 einer Probit-Regression
Nominal (Dichotom)				ϕ-Koeffiz.; Yules Q	Cramérs V
Nominal (Polychotom)					Cramérs V

Tabelle aus Eid, Gollwitzer & Schmitt (2017), S. 569

Sonderfall der Nominalskala: Dichotome Variablen

- Anzahl der Kassensitze für Psychotherapeut*innen ist weit unter bedarfsdeckend
- Forschungsfrage: Hängt die Form des Versicherungsstatus (privat vs. gesetzlich) mit dem Wahrnehmen einer Psychotherapie bei einer diagnostizierten Depression zusammen?
- Beide Variablen sind nominalskaliert und zweistufig
- Dichotome Variablen sollten idealerweise dummy kodiert werden (mit 0 und 1)

Sonderfall der Nominalskala: Dichotome Variablen

	keine Therapie (0)	Therapie (1)	Sum
Gesetzlich (0)	94	55	149
Privat (1)	24	34	58
Sum	118	89	207

Mittelwert dichotomer Variablen

$$\bar{x} = \frac{1}{n} \sum_{j=1}^{k} n_j \cdot a_j = \sum_{j=1}^{k} h_j \cdot a_j$$

- Mittelwert entspricht der relativen Häufigkeit der Kategorie 1
- Therapie: $\bar{x} = 0.43$, Versicherung: $\bar{y} = 0.28$

Korrelationskoeffizient ϕ

$$\hat{\phi} = \frac{s_{XY}}{s_X \cdot s_Y} = r_{XY}$$

$$= \frac{n_{11} \cdot n_{22} - n_{12} \cdot n_{21}}{\sqrt{(n_{11} + n_{12}) \cdot (n_{11} + n_{21}) \cdot (n_{12} + n_{22}) \cdot (n_{21} + n_{22})}}$$

- Bei dummy-kodierten Variablen kann die Produkt-Moment-Korrelation als $\hat{\phi}$ interpretiert werden
- Grenzwerte (-1 oder 1) nur bei perfektem Zusammenhang und gleicher Randverteilung möglich
- Im Beispiel: $\phi_{XY} = 0.197$

Yules Q

	keine Therapie (0)	Therapie (1)	Sum
Gesetzlich (0)	94	55	149
Privat (1)	24	34	58
Sum	118	89	207

$$Q = \frac{n_{11} \cdot n_{22} - n_{12} \cdot n_{21}}{n_{11} \cdot n_{22} + n_{12} \cdot n_{21}} = \frac{n_K - n_D}{n_K + n_D}$$

$$Q = \frac{94 \cdot 34 - 55 \cdot 24}{94 \cdot 34 + 55 \cdot 24} = \frac{1876}{4516} = 0.415$$

- Entspricht dem vereinfachten Fall von $\hat{\gamma}$
- Wertebereich: $[-1; 1]$, übliche Interpretation

Vergleich von ϕ und Yules Q

$Q = -1,\ \phi = -.5$

	keine Therapie (0)	Therapie (1)	Sum
Gesetzlich (0)	30	30	60
Privat (1)	30	0	30
Sum	60	30	90

$Q = 1,\ \phi = .5$

	keine Therapie (0)	Therapie (1)	Sum
Gesetzlich (0)	30	30	60
Privat (1)	0	30	30
Sum	30	60	90

- Q gibt strikte Ordnung der drei besetzten Kategorien wieder
- ϕ gibt wieder, dass Eindeutigkeit nur in eine Richtung besteht