

Statistik I - Sitzung 7

Bernd Schlipphak

Institut für Politikwissenschaft

Sitzung 7

1 Zusammenhangsmaße für ordinal skalierte Variablen

2 Metrische Zusammenhangsmaße

- Einführung
- Formalia
- Pearson's r
- Sonderfall: Spearman's R

Ordinale Zusammenhangsmaße - Übersicht

- Folgende Zusammenhangsmaße gehören zu den wichtigsten ordinalen Zusammenhangsmaßen
 - Kendall's τ
 - Somers' d
 - Goodman and Kruskal's γ
 - Spearman's R
- Grundlage für Kendall's τ , Somers' d und Goodman and Kruskal's γ : Anzahl an konkordanten und diskordanten Paaren!
- Spearman's R : Vergleich der Rangplätze!

Ordinale Zusammenhangsmaße - Grundlagen

- 'In a **concordant** pair, one individual (case) is higher than the other on both variables.' (Johnson/Reynolds 2008: 439)
- Es existieren zwei Fälle i und j . Stellen diese Fälle ein konkordantes Paar dar, so gilt: Wenn $x_j > x_i$, dann $y_j > y_i$ bzw. wenn $x_i > x_j$, dann $y_i > y_j$
- In einem konkordanten Paar weist also ein Fall für beide Variablen eine höhere Kategorie / einen höheren Wert als der andere Fall auf

Ordinale Zusammenhangsmaße - Grundlagen

- 'In a **discordant** pair, one case is lower on one of the variables but higher on the other.' (Johnson/Reynolds 2008: 439f.)
- Es existieren zwei Fälle i und j . Stellen diese Fälle ein diskordantes Paar dar, so gilt: Wenn $x_j > x_i$, dann $y_j < y_i$ bzw. wenn $x_i > x_j$, dann $y_i < y_j$
- In einem diskordanten Paar weist also ein Fall für eine der beiden Variablen eine höhere Kategorie / einen höheren Wert und für die andere Variablen eine niedrigere Kategorie / einen niedrigeren Wert als der andere Fall auf

Ordinale Zusammenhangsmaße - Grundlagen

- 'A **tied** pair is a pair in which both cases have the same value on one or both of the variables.' (Johnson/Reynolds 2008: 439f.)
- Es existieren zwei Fälle i und j . Stellen diese Fälle ein verbundenes Paar dar, so gilt $x_j = x_i$ und / oder $y_j = y_i$
- In einem verbundenen Paar weist also ein Fall für mindestens eine der beiden Variablen den gleichen Wert / die gleiche Kategorie auf wie der andere Fall

Ordinale Zusammenhangsmaße - Grundlagen

TABLE 12-7

Table with Concordant, Discordant, and Tied Pairs

Variable Y	Variable X		
	High	Medium	Low
High	Alex	Dawn	Gus
Medium		Ernesto	Hera
Low	Carl	Fay	Ike Jasmine

Grün = Konkordant

Ordinale Zusammenhangsmaße - Grundlagen

TABLE 12-7

Table with Concordant, Discordant, and Tied Pairs

Variable Y	Variable X		
	High	Medium	Low
High	Alex	Dawn	Gus
Medium		Ernesto	Hera
Low	Carl	Fay	Ike Jasmine

Rot = Diskordant

Ordinale Zusammenhangsmaße - Grundlagen

TABLE 12-7

Table with Concordant, Discordant, and Tied Pairs

Variable Y	Variable X		
	High	Medium	Low
High	Alex	Dawn	Gus
Medium		Ernesto	Hera
Low	Carl	Fay	Ike Jasmine

Blau = Tied/Verbunden

Ordinale Zusammenhangsmaße - Formales

- Für die Berechnung der Zusammenhangsmaße müssen wir uns nun darauf einigen, dass
 - N_C = die Summe/Anzahl aller konkordanten Paare
 - N_D = die Summe/Anzahl aller diskordanten Paare
 - N_T = die Summe aller möglichen Paare
 - T_Y = die Summe/Anzahl aller auf Y verbundenen Paare
 - T_X = die Summe/Anzahl aller auf X verbundenen Paare

Ordinale Zusammenhangsmaße - Formales

- Wir unterscheiden dabei zwischen den **symmetrischen** Assoziationsmaßen, die **keinen** gerichteten Zusammenhang implizieren, und den **asymmetrischen** Assoziationsmaßen, die einen **gerichteten** Zusammenhang implizieren
- Symmetrisch: Goodmans and Kruskals Gamma, Kendalls Tau a,b,c
- Asymmetrisch: Somers' d

Ordinale Zusammenhangsmaße - Formales

- Goodman and Kruskals Gamma: $\gamma = \frac{N_C - N_D}{N_C + N_D}$
- Kendalls tau b: $\tau_B = \frac{N_C - N_D}{\sqrt{(N_C + N_D + T_X)(N_C + N_D + T_Y)}}$
- Somers d: $d_{YX} = \frac{N_C - N_D}{N_C + N_D + T_Y}$

Ordinale Zusammenhangsmaße - Beispiel

- Hypothetisches Beispiel: Wir wollen den Zusammenhang zwischen dem Grad an Euroskeptizismus von Befragten und ihren Stolz auf ihre Nationalität berechnen.
- Der Grad an Euroskeptizismus ist die *abhängige*, der Stolz auf die Nationalität die *unabhängige* Variable, da wir theoretisch erwarten, dass letzteres auf ersteres wirken sollte
- Beide Variablen sind dreistufig (ordinal) skaliert, mit 1 = sehr euroskeptisch / sehr stolz bis 3 = nicht euroskeptisch / nicht stolz

Ordinale Zusammenhangsmaße - Beispiel

	$x_1 = \text{sehr stolz}$	$x_2 = \text{ein wenig stolz}$	$x_3 = \text{nicht stolz}$
$y_1 = \text{sehr euroskeptisch}$	10	4	1
$y_2 = \text{wenig euroskeptisch}$	2	12	5
$y_3 = \text{nicht euroskeptisch}$	2	3	10

Ordinale Zusammenhangsmaße - Beispiel

	$x_1 = \text{sehr stolz}$	$x_2 = \text{ein wenig stolz}$	$x_3 = \text{nicht stolz}$
$y_1 = \text{sehr euroskeptisch}$	10	4	1
$y_2 = \text{wenig euroskeptisch}$	2	12	5
$y_3 = \text{nicht euroskeptisch}$	2	3	10

- N_C (die Summe/Anzahl aller konkordanten Paare) =
 $10 \cdot (12 + 5 + 3 + 10) + 4 \cdot (5 + 10) + 2 \cdot (3 + 10) + 12 \cdot (10) = 506$

Ordinale Zusammenhangsmaße - Beispiel

	$x_1 = \text{sehr stolz}$	$x_2 = \text{ein wenig stolz}$	$x_3 = \text{nicht stolz}$
$y_1 = \text{sehr euroskeptisch}$	10	4	1
$y_2 = \text{wenig euroskeptisch}$	2	12	5
$y_3 = \text{nicht euroskeptisch}$	2	3	10

- N_C (die Summe/Anzahl aller konkordanten Paare) =
 $10 \cdot (12 + 5 + 3 + 10) + 4 \cdot (5 + 10) + 2 \cdot (3 + 10) + 12 \cdot (10) = 506$

Ordinale Zusammenhangsmaße - Beispiel

	$x_1 = \text{sehr stolz}$	$x_2 = \text{ein wenig stolz}$	$x_3 = \text{nicht stolz}$
$y_1 = \text{sehr euroskeptisch}$	10	4	1
$y_2 = \text{wenig euroskeptisch}$	2	12	5
$y_3 = \text{nicht euroskeptisch}$	2	3	10

- N_C (die Summe/Anzahl aller konkordanten Paare) =
 $10 \cdot (12 + 5 + 3 + 10) + 4 \cdot (5 + 10) + 2 \cdot (3 + 10) + 12 \cdot (10) = 506$

Ordinale Zusammenhangsmaße - Beispiel

	$x_1 = \text{sehr stolz}$	$x_2 = \text{ein wenig stolz}$	$x_3 = \text{nicht stolz}$
$y_1 = \text{sehr euroskeptisch}$	10	4	1
$y_2 = \text{wenig euroskeptisch}$	2	12	5
$y_3 = \text{nicht euroskeptisch}$	2	3	10

- N_D (die Summe/Anzahl aller diskordanten Paare) = $4 \cdot (2+2) + 1 \cdot (2+12+2+3) + 12 \cdot (2) + 5 \cdot (2+3) = 84$

Ordinale Zusammenhangsmaße - Beispiel

	$x_1 = \text{sehr stolz}$	$x_2 = \text{ein wenig stolz}$	$x_3 = \text{nicht stolz}$
$y_1 = \text{sehr euroskeptisch}$	10	4	1
$y_2 = \text{wenig euroskeptisch}$	2	12	5
$y_3 = \text{nicht euroskeptisch}$	2	3	10

- N_D (die Summe/Anzahl aller diskordanten Paare) = $4 \cdot (2+2) + 1 \cdot (2+12+2+3) + 12 \cdot (2) + 5 \cdot (2+3) = 84$

Ordinale Zusammenhangsmaße - Beispiel

- T_Y (die Summe/Anzahl aller auf Y verbundenen Paare) = $10 \cdot (4+1) + 4 \cdot (1) + 2 \cdot (12+5) + 12 \cdot (5) + 2 \cdot (3+10) + 3 \cdot (10) = \mathbf{204}$
- T_X (die Summe/Anzahl aller X-verbundenen Paare) = $10 \cdot (2+2) + 2 \cdot (2) + 4 \cdot (12+3) + 12 \cdot (3) + 1 \cdot (5+10) + 5 \cdot (10) = \mathbf{205}$
- N_T (die Summe aller möglichen Paare) = $[(10 + 4 + 1 + 2 + 12 + 5 + 2 + 3 + 10) \cdot (10 + 4 + 1 + 2 + 12 + 5 + 2 + 3 + 10 - 1)] / 2 = \mathbf{1176}$

	$x_1 = \text{sehr stolz}$	$x_2 = \text{ein wenig stolz}$	$x_3 = \text{nicht stolz}$
$y_1 = \text{sehr euroskeptisch}$	10	4	1
$y_2 = \text{wenig euroskeptisch}$	2	12	5
$y_3 = \text{nicht euroskeptisch}$	2	3	10

Ordinale Zusammenhangsmaße - Beispiel

- Goodman and Kruskals Gamma: $\gamma = \frac{N_C - N_D}{N_C + N_D} = \frac{506 - 84}{506 + 84} = 0.72$
- Kendalls tau b: $\tau_B = \frac{N_C - N_D}{\sqrt{(N_C + N_D + T_X)(N_C + N_D + T_Y)}} = \frac{506 - 84}{\sqrt{(506 + 84 + 205)(506 + 84 + 204)}} = 0.53$
- Somers d: $d_{YX} = \frac{N_C - N_D}{N_C + N_D + T_Y} = \frac{506 - 84}{506 + 84 + 204} = 0.53$

Ordinale Zusammenhangsmaße - Beispiel

- Goodman and Kruskals Gamma: $\gamma = 0.72$
- Kendalls tau b: $\tau_B = 0.53$
- Somers d: $d_{YX} = 0.53$
- **ACHTUNG:** Goodman und Kruskals Gamma überschätzt die Größen des Zusammenhangs meist stark, während Kendalls tau b und Somers d realistischere Maße liefern! Dies liegt am (fehlenden) Einbezug der - meist häufig auftretenden - verbundenen Paaren!

Metrische Zusammenhangsmaße - Grundlagen

- Um zu verstehen, was wir mit den einfachen Zusammenhangsmaßen für metrische Variablen – der **Kovarianz** und dem **Korrelationskoeffizienten** – überhaupt messen, schauen wir uns zuerst ein Streudiagramm an
- Ein Streudiagramm (oder auch Punktwolke, engl. Scatterplot) beinhaltet alle Fälle eines Datensatzes als Punkte, welche über die Koordinaten der Fälle auf der x- und y-Achse (d.h., über die Werte der Fälle auf den Variablen X und Y) definiert sind

Metrische Zusammenhangsmaße - Grundlagen

- Perfekter Zusammenhang dann, wenn es nur diskordante oder konkordante Paare gibt
- D.h., perfekter Zusammenhang liegt dann vor, wenn Punkte alle auf einer Linie von links unten nach rechts oben oder von links oben nach rechts unten liegen
- Das wird allerdings in den Sozialwissenschaften niemals der Fall sein – eher sehen (starke) Zusammenhänge folgendermaßen aus:

Metrische Zusammenhangsmaße - Grundlagen



Abbildung: Diaz-Bone 2006: 84

Metrische Zusammenhangsmaße - Grundlagen

- Generell ist anhand von Streudiagrammen ein erster 'Trend' zum Zusammenhang zweier Variablen erkennbar

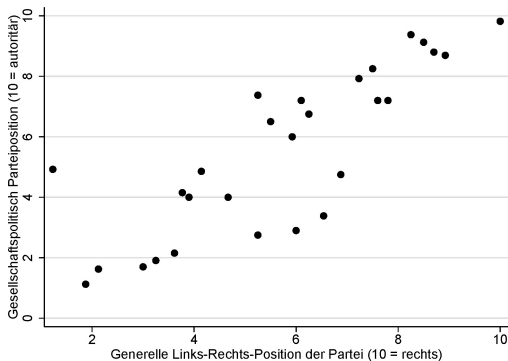


Abbildung: Chapel Hill Expert Survey: Parteienposition aus DE, AUT, CH

Metrische Zusammenhangsmaße - Grundlagen

- Um nun zu erkennen, wie stark die Variablen miteinander (ungerichtet!) zusammenhängen (oder korrelieren), können wir zwei Maße berechnen
 - Die Kovarianz
 - Den Korrelationskoeffizienten r (oder: Pearson's r)

Metrische Zusammenhangsmaße - Formales

- Die Kovarianz wird als $cov(x; y)$ bezeichnet und bildet die Grundlage für den später einzuführenden Korrelationskoeffizienten

$$cov(x; y) = \frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})$$

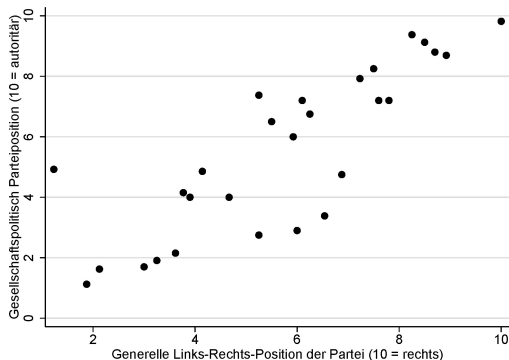
- Für jeden Fall wird damit das Abweichungsprodukt hinsichtlich der Abweichungen des Falles zum Mittelwert von X und Y berechnet
- Bildung des arithmetischen Mittels auf der Grundlage der Summe der Abweichungsprodukte aller Fälle \Rightarrow Kovarianz auch als **'durchschnittliches Abweichungsprodukt'**

Metrische Zusammenhangsmaße - Formales

- Werte der Kovarianz werden dann extrem, wenn die Fälle gleichermaßen und einem systematischen Trend folgend sehr weit von den Mittelwerten von X und Y entfernt sind
- Ein Nullzusammenhang tritt auf, wenn
 - alle Fälle als x - und / oder y -Werte die jeweiligen Mittelwerte für X und / oder Y aufweisen
 - alle Fälle vollkommen unsystematisch (zufällig) von den Mittelwerten abweichen

Metrische Zusammenhangsmaße - Beispiel

- Beispiel: Gibt es einen Zusammenhang zwischen der grundlegenden Links-Rechts-Positionierung einer Partei und ihrer Positionierung in der Gesellschaftspolitik (libertäre versus autoritäre Positionen)? (Datenquelle: Chapel Hill Expert Survey)



Metrische Zusammenhangsmaße - (vereinfachtes) Beispiel

	Mittelwert LR	Mittelwert Gesellschaftspolitik	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$
CDU	5.9	6	0.43	0.63	0.27
SPD	3.8	4.2	-1.67	-1.17	1.96
Grüne	3.6	2.2	-1.87	-3.17	5.92
FDP	6.5	3.4	1.03	-1.97	-2.04
Linke	1.2	4.9	-4.27	-0.47	2.02
AfD	8.9	8.7	3.43	3.33	11.42
SPÖ	3.9	4	-1.57	-1.37	2.25
ÖVP	6.1	7.2	0.63	1.83	1.16
FPÖ	8.7	8.8	3.23	3.43	11.08
GRÖ	3	1.7	-2.47	-3.67	9.06
TeamST	7.6	7.2	2.13	1.83	3.90
SVP	8.3	9.4	2.83	4.03	11.41
SP	2.1	1.6	-3.37	-3.77	12.70
FDP	6.9	4.8	1.43	-0.57	-0.82
CVP	5.5	6.5	0.03	1.13	0.04
$\bar{x} \approx 5.47/s_x \approx 2.35$		$\bar{y} \approx 5.37/s_y \approx 2.49$			$\sum 70.23$

Metrische Zusammenhangsmaße - Beispiel

- In die Formel eingesetzt, ergibt das anhand unseres (realen) Beispiels für den Zusammenhang von genereller LR-Position und gesellschaftspolitischer Position einer Partei

$$\text{Kovarianz: } cov(x; y) = \frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{15}(70.23) \approx \mathbf{4.68}$$

- Die Kovarianz variiert von 0 bis +/- unendlich, wobei 0 einen nicht vorhandenen Zusammenhang bezeichnet.
- Je größer der negative oder positive Wert der Kovarianz, desto stärker ist der negative oder positive Zusammenhang zwischen zwei Variablen.

Metrische Zusammenhangsmaße - Pearson's r

- Aber: Größe der Kovarianz auch von der Dimensionierung der beiden metrisch skalierten Variablen abhängig!
- Daher: Normierung der Kovarianz durch ein anderes Maß auf den Bereich zwischen -1 und $+1$ notwendig
- Dieses neue Maß ist der Korrelationskoeffizient (Pearson's) r

Metrische Zusammenhangsmaße - Pearson's r

- Die Normierung erfolgt durch das Teilen der Kovarianz durch die Standardabweichungen von X und Y

$$r = \frac{\text{cov}(x; y)}{s_x * s_y}$$

- Auch der Korrelationskoeffizient r weist den Wert 0 auf, wenn kein Zusammenhang existiert. Ein perfekter negativer Zusammenhang ist aber = -1, ein perfekter positiver Zusammenhang = +1

Metrische Zusammenhangsmaße - Pearson's r

- Setzen wir die Werte aus dem Beispiel ein, so erhalten wir:

$$r = \frac{\text{cov}(x; y)}{s_x * s_y} = \frac{4.68}{2.35 * 2.49} \approx 0.79$$

$0,00 \leq r \leq 0,05$	keine Korrelation
$0,05 < r < 0,20$	schwache Korrelation
$0,20 < r < 0,50$	mittlere Korrelation
$0,50 < r < 0,70$	starke Korrelation
$0,70 < r < 1,00$	sehr starke Korrelation

Abbildung: Diaz-Bone 2006: 91

Ordinale Zusammenhangsmaße: Spearman's R

- Spearman's R ist ein Sonderfall, weil er eigentlich ein ordinales Zusammenhangsmaß ist, aber im Kern doch metrisches Skalenniveau voraussetzt
- Grundlegend definiert sich Spearman's R über die Differenz zwischen den Rangplätzen, die ein Fall im Hinblick auf zwei unterschiedliche Variablen einnimmt \Rightarrow ordinales Zusammenhangsmaß!
- Diese Differenz nennen wir d bzw. für jeden individuellen Fall d_i
- Die Formel: Spearman's $R = 1 - \frac{6 * \sum d_i^2}{n(n^2 - 1)}$

Ordinale Zusammenhangsmaße: Spearman's R

- Was ist denn diese Differenz zwischen den Rangplätzen und woher bekommen wir sie?
- Jeder Fall kann auf zwei Variablen unterschiedliche Rangplätze haben. Der Unterschied zwischen diesen Rangplätzen ist die Differenz.
- Beispiel Sport: Jemand kann die schnellste sein (Rang 1), aber nur die zehntbeste Weite im Weitsprung erreichen (Rang 10)
- Die Rangplatzdifferenz d ist in diesem Fall 9, da die Differenz zwischen 1 und 10 $(-)$ 9 beträgt.

Ordinale Zusammenhangsmaße: Spearman's R

- Um Spearman's R zu berechnen, benötigen wir also eine andere Tabelle als für die anderen ordinalen Zusammenhangsmaße
- Die Tabelle für Spearman's R enthält für jeden Fall (=Zeile) die
 - absoluten Werte auf den beiden Variablen,
 - den Rang des Falles für beide Variablen, und
 - die Differenz sowie die quadrierte Differenz.

Ordinale Zusammenhangsmaße: Spearman's R

- X = Sympathie Trump (0 = sehr schlecht bis 100 = sehr gut), Y = Vorhergesagter AfD-Wähleranteil

Fall	X	Y	Rang(X)	Rang (Y)	d_i	d_i^2
1	10	2				
2	25	15				
3	35	10				
4	40	20				
5	30	12				
6	20	5				
7	50	9				
8	15	7				
9	55	13				
10	80	90				
Σ						

Ordinale Zusammenhangsmaße: Spearman's R

- X = Sympathie Trump (0 = sehr schlecht bis 100 = sehr gut), Y = Vorhergesagter AfD-Wähleranteil

Fall	X	Y	Rang(X)	Rang (Y)	d_i	d_i^2
1	10	2	1	1		
2	25	15	4	8		
3	35	10	6	5		
4	40	20	7	9		
5	30	12	5	6		
6	20	5	3	2		
7	50	9	8	4		
8	15	7	2	3		
9	55	13	9	7		
10	80	90	10	10		
Σ						

Ordinale Zusammenhangsmaße: Spearman's R

- X = Sympathie Trump (0 = sehr schlecht bis 100 = sehr gut), Y = Vorhergesagter AfD-Wähleranteil

Fall	X	Y	Rang(X)	Rang (Y)	d_i	d_i^2
1	10	2	1	1	0	0
2	25	15	4	8	4	16
3	35	10	6	5	1	1
4	40	20	7	9	2	4
5	30	12	5	6	1	1
6	20	5	3	2	1	1
7	50	9	8	4	4	16
8	15	7	2	3	1	1
9	55	13	9	7	2	4
10	80	90	10	10	0	0
Σ						44

Ordinale Zusammenhangsmaße: Spearman's R

- Die Formel: Spearman's $R = 1 - \frac{6 * \sum d_i^2}{n(n^2 - 1)}$
- Aus Beispiel eingesetzt:
 - Spearman's $R = 1 - \frac{6 * 44}{10(10^2 - 1)}$
 - Spearman's $R = 1 - \frac{264}{990}$
 - Spearman's $R = 1 - 0.27 = 0.73$
- Spearman's $R = 0.73 \Rightarrow$ sehr starke Korrelation!

Ordinale Zusammenhangsmaße: Spearman's R

- Problem: Was passiert, wenn mehrere Fälle den gleichen Wert auf X bzw. Y aufweisen (also: den gleichen Rang) aufweisen würden?
- Zwei Effekte - Berechnung der Ränge und Änderung der Formel
 - Bei der Berechnung sogenannter verbundener Ränge werden die Ränge gemittelt (siehe Beispiel)
 - Die (einfachere) Formel kann nicht mehr verwendet werden. Stattdessen wird eine Formel basierend auf dem Pearson'schen Korrelationskoeffizienten verwendet.
 - Spearman's $R = p_{r_x, r_y} = \frac{cov(r_x, r_y)}{s_{r_x} * s_{r_y}}$

Ordinale Zusammenhangsmaße: Spearman's R

- X = Sympathie Trump (0 = sehr schlecht bis 100 = sehr gut), Y = Vorhergesagter AfD-Wähleranteil

Fall	X	Y	Rang(X)	Rang (Y)
1	10	5	$1,5 = (Rang1 + Rang2/2)$	
2	20	10		
3	30	15	$6 = (Rang5 + Rang6 + Rang7/3)$	
4	30	20	$6 = (Rang5 + Rang6 + Rang7/3)$	
5	30	10	$6 = (Rang5 + Rang6 + Rang7/3)$	
6	20	5		
7	50	9		
8	10	7	$1,5 = (Rang1 + Rang2/2)$	
9	50	20		
10	80	90		
Σ				

Ordinale Zusammenhangsmaße: Spearman's R

- X = Sympathie Trump (0 = sehr schlecht bis 100 = sehr gut), Y = Vorhergesagter AfD-Wähleranteil

Fall	X	Y	Rang(X)	Rang (Y)
1	10	5	$1,5 = (Rang1 + Rang2/2)$	1,5
2	20	10	3,5	5,5
3	30	15	$6 = (Rang5 + Rang6 + Rang7/3)$	7
4	30	20	$6 = (Rang5 + Rang6 + Rang7/3)$	8,5
5	30	10	$6 = (Rang5 + Rang6 + Rang7/3)$	5,5
6	20	5	3,5	1,5
7	50	9	8,5	4
8	10	7	$1,5 = (Rang1 + Rang2/2)$	3
9	50	20	8,5	8,5
10	80	90	10	10
Σ				

Ordinale Zusammenhangsmaße: Spearman's R

- Was passiert, wenn mehrere Fälle den gleichen Wert auf X bzw. Y aufweisen (also: den gleichen Rang) aufweisen würden?
- Für die Berechnung von $R = p_{r_x, r_y} = \frac{cov(r_x, r_y)}{s_{r_x} * s_{r_y}}$ müsste man dann
 - für die Variablen $\text{Rang}(X)$ und $\text{Rang}(Y)$ jeweils die StA ($= s_{r_x}, s_{r_y}$)
 - und die Kovarianz ($= cov(r_x, r_y)$) zwischen beiden Variablen berechnen

Ordinale Zusammenhangsmaße - Spearman's R

$$\text{Spearman's } R = \frac{\text{cov}(r_x; r_y)}{s_{r_x} * s_{r_y}}$$

- Für Spearman's R berechnen wir also die Kovarianz und die jeweiligen Standardabweichungen der Rang-Variablen
- Problem: Die Rangvariablen sind ordinal! Wir dürfen dafür eigentlich keine auf dem arithmetischen Mittelwert basierenden Maße berechnen!
- Eine solche Anwendung von Spearman's R sollte daher immer kritisch reflektiert und mit anderen ordinalen Zusammenhangsmaßen abgeglichen werden!

Zusammenfassung

- In dieser Sitzung haben wir vor allem ungerichtete ordinale und metrische Zusammenhangsmaße kennengelernt
- Die ordinalen Zusammenhangsmaße, die auf dem Verhältnis diskordanter und konkordanter Paare beruhen, unterscheiden sich vor allem in ihrer Aufnahme verbundener Paare
- Die metrischen Zusammenhangsmaße drücken nur ungerichtete Zusammenhänge aus - mit der bivariaten Regression lernen wir in der nächsten Sitzung ein gerichtetes Zusammenhangsmaß kennen
- Das Zusammenhangsmaß Spearman's R ist ein Sonderfall, da eigentlich ordinal, aber in seiner 'normalen' Anwendung metrisch skalierte Maße (Mittelwert) verwendend. **Dieses Maß ist daher eigentlich nur unter großen Vorbehalten anzuwenden!**