

Vorlesung: Statistik I

Prof. Dr. Simone Abendschön

8. Einheit

- **Zusammenhangsmaß für ordinalskalierte Variablen:
Spearman's Rho**

- Auch für zwei **ordinalskalierte Merkmale** lassen sich Zusammenhänge berechnen
- Es empfehlen sich zunächst auch hier **Kreuztabellen** zur Anschauung der kombinierten Häufigkeiten
- **Spearman's Rho als sog. „Rangkorrelationskoeffizient“:** Merkmalsausprägungen zweier Merkmale werden in jeweils geordnete Rangfolgen gebracht und in Beziehung gesetzt
- Weitere Maßzahlen: Kendall's Tau (werden hier nicht behandelt)

Vorgehen:

- 1) Merkmalsausprägungen zweier Merkmale werden in jeweils geordnete Rangfolgen gebracht
- 2) Anschließend werden die jeweiligen Rangpositionen der gepaarten Merkmalsausprägungen in Beziehung gesetzt
- 3) Rho mit Formel berechnen

$$\rho = 1 - \frac{6 * \sum_{i=1}^n d_i^2}{n * (n^2 - 1)}$$

wobei $d_i = R(x_i) - R(y_i)$ (Differenz der Ränge)

Kann Werte zwischen -1 und +1 annehmen, wobei 0 keinen Zusammenhang bedeutet

- **Positiver Koeffizient** weist auf einen gleichgerichteten bzw. positiven Zusammenhang hin
- **Negativer Koeffizient** auf einen negativen Zusammenhang

Beispiel: Schulbildung und politisches Interesse (beide ordinal skaliert)

- **Interpretation positiver Wert:** Mit steigendem Bildungsabschluss lässt sich auch ein höheres politisches Interesse beobachten
- **Interpretation negativer Wert:** Mit steigendem Bildungsabschluss lässt sich ein niedrigeres politisches Interesse beobachten

- **Zusammenhang zwischen sozialer Schicht und Gesundheitszustand** (vgl. auch Lehrbrief S. 60f.)
- **Hypothese:** je höher die soziale Schicht desto besser auch der Gesundheitszustand

ID	Soziale Schicht	Gesundheitszustand
1	Mittelschicht (3)	Schlecht (2)
2	Arbeiterschicht (2)	Sehr schlecht (1)
3	Oberschicht (5)	Sehr gut (5)
4	Obere Mittelschicht (4)	Gut (4)
5	Unterschicht (1)	Zufriedenstellend (3)

Quelle: Eigene Darstellung

Spearman's Rho: Beispiel

9

- Schritt 1: Überführung der Werte auf den Merkmalen in Rangpositionen - Arbeitstabelle

ID	Soziale Schicht (x)	Gesundheitszustand (y)	$d_i = R(x_i) - R(y_i)$
1	3	2	1
2	2	1	1
3	5	5	0
4	4	4	0
5	1	3	-2

Quelle: Eigene Darstellung

„Rangplätze“

- Überführung der Werte auf den Merkmalen in Rangpositionen – Arbeitstabelle

ID	Soziale Schicht (x)	Gesundheitszustand (y)	$d_i = R(x_i) - R(y_i)$
1	3	2	1
2	2	1	1
3	5	5	0
4	4	4	0
5	1	3	-2

Quelle: Eigene Darstellung

$$\rho = 1 - \frac{6 * \sum_{i=1}^n d_i^2}{n * (n^2 - 1)} = \frac{6 * (1^2 + 1^2 + 0^2 + 0^2 + (-2)^2)}{5 * (5^2 - 1)} = 1 - \frac{6 * 6}{120} = 1 - 0,3 = 0,7$$

Spearman's Rho	Interpretation
$\leq 0,05$	kein Zusammenhang
$> 0,05$ bis $\leq 0,20$	schwacher Zusammenhang
$> 0,20$ bis $\leq 0,50$	mittelstarker Zusammenhang
$> 0,50$ bis $\leq 0,70$	starker Zusammenhang
$> 0,70$	sehr starker Zusammenhang

Quelle: Eigene Darstellung

- **Bei größeren Stichproben:** mindestens zwei oder mehr Fälle weisen die gleiche Merkmalsausprägung auf → Rangpositionen mit sogenannten Bindungen (engl. ties)
- **Pragmatisches Vorgehen:** Den verbundenen Beobachtungseinheiten wird das arithmetische Mittel derjenigen Rangplätze zugewiesen, die ohne Bindungen vergeben worden wären
- **Achtung:** Übersteigt die Anzahl der gebundenen Rangplätze den Anteil von 20% aller Ränge, muss Spearman's Rho mit einer komplexeren Formel berechnet werden (siehe Lehrbrief; Statistikprogramm nutzt automatisch die korrekte Formel)

ID	Soziale Schicht	Gesundheitszustand
1	Mittelschicht (3)	Schlecht (2)
2	Arbeiterschicht (2)	Sehr schlecht (1)
3	Oberschicht (5)	Sehr gut →
4	Obere Mittelschicht (4)	4,5 für beide
5	Unterschicht (1)	Zufriedenstellend (3)

Quelle: Eigene Darstellung

Mittelwert aus 4.
und 5. Rangplatz

Bildungsabschluss und Einkommen:

Bitte berechnen Sie auf Basis der Urliste

Spearman's Rho

ID	Bildung	Einkommen
1	Hauptschule (1)	Mittel (2)
2	Realschule (2)	Niedrig (1)
3	Abitur (3)	Hoch (3)
4	Realschule (2)	Hoch (3)
5	Studium (4)	Sehr hoch (4)