

Hoofdstuk 5. Ordinale associatiematen

Associatiematen op ordinaal niveau

Deze maten zeggen iets over de sterkte van het verband, maar ook iets over de richting. Er is sprake van rangordening. Bijvoorbeeld iemand kan meer of minder radio luisteren. De waarde van ordinale samenhangmaten ligt tussen -1 en +1. Deze maten houden tevens rekening met afhankelijke en onafhankelijke variabelen, dus er is sprake van een asymmetrische relatie.

Er kan bij de verbanden tussen variabelen sprake zijn van een positieve samenhang, oftewel een stijgende lijn (als de waarde op de ene variabele hoger is, is hij op de andere variabele natuurlijk ook hoger). Ten tweede is het mogelijk dat er sprake is van een negatieve samenhang, oftewel een dalende lijn. Ten slotte is er de mogelijkheid dat er geen verband is tussen de variabelen, de lijn is dan horizontaal.

Gamma

Met gamma (γ) kan het verband worden berekend tussen twee ordinale variabelen, waarbij geen rekening wordt gehouden met of de variabelen afhankelijk of onafhankelijk zijn. Waar het vooral om gaat bij gamma zijn de **concordante (nc) en discordante (nd) paren**. Wanneer er meer concordante paren zijn dan discordante paren, spreken we van een positieve samenhang. Als er meer discordante paren zijn, is er een negatieve samenhang.

Formule voor gamma:

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

Er is sprake van een concordant paar als van een paar de ene onderzoekseenheid op beide variabelen hoger scoort dan de andere onderzoekseenheid. Bijvoorbeeld onderzoekseenheid D (2,2) scoort op beide variabelen hoger dan onderzoekseenheid C (1,1). Er is een stijgende lijn te zien als je de punten met elkaar verbindt. Je begint bij het tellen van deze paren altijd bij de cel met de laagste waarden (0,0). Je vermenigvuldigt deze met het paar dat hier concordant aan is. Bijvoorbeeld cel A (0,0) heeft 3 onderzoekseenheden, cel C (1,1) heeft ook 3 onderzoekseenheden en cel E (2,2) heeft 2 onderzoekseenheden.

Je vermenigvuldigt deze om het aantal concordante paren te berekenen, oftewel $3 \times (3+2) = 9$. Je berekent op deze manier alle mogelijke concordante paren binnen een tabel.

Laagste combinatie (aantal n)	x
[Hogere combinatie (aantal n)	+
Hogere combinatie (aantal n)	
Hogere combinatie (aantal n)]	
<hr/>	
Aantal concordante paren	

JoHo Samenvatting – Beschrijvende Statistiek

Er is sprake van een discordant paar als van een paar de ene variabele lager is en de andere variabele hoger. Bijvoorbeeld onderzoekseenheid C (1,0) scoort op de eerste variabele hoger en op de tweede variabele hoger dan onderzoekseenheid D (0,1). Er is een dalende lijn te zien als je de punten met elkaar verbindt. De discordante paren worden op dezelfde manier berekend als de concordante paren, alleen dan andersom.

Hoogste waarde op de ene en de laagste waarde op de andere variabele \times
Lagere waardes op de ene en hogere waardes op de andere variabele
Lagere waardes op de ene en hogere waardes op de andere variabele
Lagere waardes op de ene en hogere waardes op de andere variabele + _____
Aantal discordante paren

Stel dat de uitkomst .79 is, dan betekent dat er een sterke positieve samenhang is tussen de variabelen. Er zijn 79% meer concordante dan discordante paren. Wanneer er een – teken voor de uitkomst staat, is er een negatieve samenhang.

Voor de berekening in SPSS van gamma, zie de uitleg voor de berekening van Somers' d in SPSS hieronder.

Somers' d

Deze associatiemaat heeft veel weg van gamma. Het verschil is dat gamma een symmetrische associatiemaat is en Somers' d een asymmetrische associatiemaat. Hierbij is het dus wel van belang om te weten wat de afhankelijke en wat de onafhankelijke variabele is.

Berekening

De berekening van Somers' d wordt gedaan door middel van concordante en discordante paren, maar daarnaast ook met behulp van geknoopte paren. De waarde op één van de twee variabelen is dan hetzelfde. Je begint met knopen bij cel (1,1), je vermenigvuldigt ze met de paren die rechts van de eerste cel liggen. Dan ga je een rij naar beneden en knoopt daar de eerste cel met de cellen die er rechts van liggen en dus de eerste waarde gelijk hebben. Uiteindelijk tel je de geknoopte paren per rij bij elkaar op, zodat je het totaal aantal geknoopte paren krijgt.

Met behulp van de geknoopte paren en de concordante en discordante paren kan Somers' d worden berekend. Somers' d kan het beste worden gebruikt bij vierkante tabellen, omdat bij andere vormen van tabellen de waarde +1 en -1 niet kan bereiken.

Bij een asymmetrische associatiemaat zoals Somers' d gaat het om de invloed die de onafhankelijke variabele uitoefent op de afhankelijke variabele.

Formule voor Somers' d:

$$d_{yx} = \frac{Nc - Nd}{Nc + Nd + Ty}$$

De uitslag kan op dezelfde wijze worden geïnterpreteerd als gamma.

JoHo Samenvatting – Beschrijvende Statistiek

Berekening in SPSS

Berekenen van zowel Somers' d, als gamma als Kendalls tau-b in SPSS. Je maakt daarvoor eerst een kruistabel zoals in het vorige hoofdstuk uiteengezet. Daarna ga je naar Statistics, waarin je kunt aangeven welke associatiemaat je wilt berekenen. De gewenste associatiemaat kan je hier aanvinken.

Kendalls tau-b

Kendalls tau-b is een symmetrische associatiemaat (net zoals gamma) en houdt rekening met geknoopte paren (net zoals Somers' d). Het verschil met Somers' d is dat deze maat geen onderscheid maakt tussen afhankelijke en onafhankelijke variabele en de geknoopte paren worden op beide variabelen bij de berekening meegeteld. Deze associatiemaat is immers symmetrisch.

Formule voor Kendalls tau-b:

$$\text{tau-b} = \frac{(Nc - Nd)}{\sqrt{(Nc + Nd + Tx)(Nc + Nd + Ty)}}$$

Kendalls tau-b knoopt zowel op de afhankelijke variabele als op de onafhankelijke variabele. Bij Somers' d knoop je alleen op de rijen, bij Kendalls tau-b knoop je zowel op de rijen als op de kolommen.

De berekening van Kendalls tau-b in SPSS gaat via dezelfde stappen als bij gamma en Somers' d, zoals hierboven beschreven.

Bij het uitrekenen van een associatiemaat is sprake van een bivariate analyse. Je kunt in SPSS ook meerdere samenhangsmaten tegelijkertijd uitrekenen. In een **correlatiematrix** worden dan de resultaten van meerdere bivariate analyses weergegeven. Dit kan van belang zijn indien je wilt weten in hoeverre meerdere variabelen met elkaar samenhangen. Voor het berekenen van Kendalls tau-b over meerdere variabelen, ga je via Analyze → Correlate naar Bivariate. Vink vervolgens Kendall's tau-b aan en selecteer de benodigde variabelen.

JoHo Samenvatting – Beschrijvende Statistiek

Spearman's rho

Spearman's rho is een maat om de correlatie tussen rangnummers aan te geven. Deze maat wordt gebruikt tussen variabelen op ordinaal niveau en soms ook op interval- of rationiveau. De maat is gebaseerd op verschillen in de rangorde van de waarden.

Formule voor Spearman's rho:

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

Bij Spearman's rho wordt niet gerekend met de ingevulde data, zoals bij de andere associatiematen wel het geval was, maar met rangordeningen. De onderzoekseenheid met de hoogste score staat dan op één, de op één na hoogste score staat op twee, etc. Nu kan het verschil van waarden van de variabelen x en y worden berekend en worden gekwadrateerd, nadat de rangordeningen zijn vastgesteld.

Een voorbeeld van een berekening waarbij Spearman's rho wordt gebruikt, is om aan te geven wat het verband is tussen tijd gependend aan televisie kijken en het aantal spijkerbroeken in de garderobe. De persoon die het meeste televisie kijkt komt dan op één, etc. In de tweede kolom komt de rangordering in aantal spijkerbroeken. De uitslag van de berekening geeft de samenhang tussen de variabelen.

Er kan ook een gedeelde eerste positie zijn. Om de exacte positie van de respondenten te berekenen wordt een gemiddelde rangpositie genomen. De rangen worden dan opgeteld en gedeeld door het aantal respondenten. Spearman's rho is echter minder geschikt als relatief veel onderzoekseenheden een rangnummer moeten delen, doordat ze dezelfde waarden hebben.

Bij deze associatiemaat geldt een iets andere interpretatie van de uitslag. Deze ligt namelijk tussen de 0 en de 1. Bijvoorbeeld de uitkomst is .010, er is dan een zwakke samenhang tussen de mate waarin naar televisie wordt gekeken en het aantal spijkerbroeken dat men in de kast heeft hangen.