

Hoofdstuk 9. De betrouwbaarheid

De resultaten van een onderzoek moeten niet alleen met de juiste berekeningen gevonden worden, het is minstens zo belangrijk om te kunnen aantonen dat je metingen betrouwbaar zijn. Dit houdt in dat ze vrij zijn van toevallige fouten. Als het onderzoek nogmaals onder precies dezelfde omstandigheden wordt uitgevoerd, zouden hier dezelfde resultaten uit moeten voortkomen. Is dit niet het geval, dan is de meting niet betrouwbaar.

Om de betrouwbaarheid van een onderzoek te kunnen toetsen en aan te tonen, zijn verschillende statistieken en procedures ontwikkeld. **Interne consistentie**, **equivalentie** en **stabieliteit** meten drie verschillende aspecten van betrouwbaarheid.

Stabiliteit en equivalentie

Een meting is stabiel als bijvoorbeeld dezelfde vraag opnieuw aan dezelfde respondent wordt gesteld en dit hetzelfde antwoord oplevert als bij de eerste meting. Dit heet een test-hertestbetrouwbaarheidsmeting en kan worden uitgevoerd door bijvoorbeeld een vraag twee keer te stellen in een enquête. Vervolgens kan je de samenhang tussen deze twee variabelen meten met behulp van de eerder besproken associatiematen als Spearmans rho en de product moment correlatie r (ordinaal- en interval meetniveau). Kendalls tau-b bij ordinaal meetniveau, Spearmans rho bij intervalniveau en Pearsons productmomentcorrelatie r bij variabelen op ratio meetniveau. Het verband tussen deze variabelen moet zeer sterk (0,9) tot perfect (1) zijn, ze vragen immers naar hetzelfde verschijnsel.

Intra- en intercodeursbetrouwbaarheid

Deze test wordt veel gebruikt bij inhoudsanalyses en wordt op twee manieren toegepast. Hij kan de consistentie tussen verschillende codeurs meten, de intercodeursbetrouwbaarheid. De correlatie hiertussen moet eigenlijk minimaal 0,90 zijn. Is dit niet het geval, dan zullen de codeurs beter opgeleid moeten worden of moet het codeboek waarmee zij werken verder gespecificeerd worden.

Daarnaast kan je de codeur vragen om zijn metingen te herhalen, dit heet intracodeursbetrouwbaarheid. Deze correlatie zou in feite 1 moeten zijn, het gaat hier immers om dezelfde codeur.

Parallele test

Respondenten zullen echter snel geïrriteerd raken wanneer je iedere vraag twee maal stelt. Deze betrouwbaarheidstest wordt daarom niet of nauwelijks gebruikt bij vragenlijsten. Wel kan een vraag naar hetzelfde verschijnsel soms op een andere manier gesteld worden. Dit wordt een parallelle test genoemd. Bijvoorbeeld: je kan respondenten vragen om hun leeftijd in te vullen en daarnaast vragen in welk jaar ze geboren zijn.

Betrouwbaarheid op nominaal niveau

De nog niet eerder besproken maat Cohens kappa en het overeenstemmingspercentage worden gebruikt om de betrouwbaarheid van variabelen op nominaal niveau te meten.

Het overeenstemmingspercentage is het aantal coderingen waarover de codeurs het met elkaar eens zijn, uitgedrukt in percentages. Bijvoorbeeld: twee codeurs hebben beide 10 krantenberichten gecodeerd, 8 daarvan hebben ze in dezelfde categorie ingedeeld. Het overeenstemmingspercentage is dan $8 : 10 \times 100\% = 80\%$.

JoHo Samenvatting – Beschrijvende Statistiek

Cohens kappa maakt eveneens gebruik van het overeenstemmingspercentage, maar houdt daarbij rekening met toevalskans. De berekening is vergelijkbaar met de berekening van chi-kwadraat (zie hiervoor hoofdstuk 4. Associatiematen op nominaal niveau: Cramers V). P_o is in de formule het overeenstemmingspercentage en P_e is het percentage dat je verwacht op basis van toeval. Je verwacht bij P_e eigenlijk geen enkel verband tussen codeur 1 en codeur 2, want dan is het onderzoek betrouwbaarder.

Formule van kappa:

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

Berekening in SPSS

SPSS kan kappa ook berekenen. Ga naar Analyze → Descriptive Statistics → Crosstabs om een kruistabel te maken. Met Statistics kom je in het venster Crosstabs: Statistics. Hierin kan kappa aangevinkt worden.

Interne consistentie

Simpele vragen uit een enquête, bijvoorbeeld over geslacht, zullen geen problemen opleveren. Als je echter een bepaald concept wilt onderzoeken met behulp van meerdere vragen, dan is het berekenen van de betrouwbaarheid van deze factor noodzakelijk.

Je wilt bijvoorbeeld weten of iemand een hang naar sensatie heeft. Je kan respondenten dit niet gewoon vragen, omdat iedereen dit verschijnsel anders zal interpreteren. Daarom meet je het concept 'hang naar sensatie' met verschillende stellingen die je baseert op de theorie of op meetinstrumenten die in ander onderzoek zijn gebruikt. De stellingen "Ik wil een keer bungeejumpen" en "Ik kijk graag spannende films" vragen beide naar sensatie, het is echter nog maar de vraag of deze variabelen intern consistent zijn. Hangen deze variabelen of items inderdaad sterk met elkaar samen?

Cronbachs alpha

Met behulp van Cronbachs alpha worden de correlaties tussen alle items die dezelfde concept meten, berekend. Het kan gezien worden als een gemiddelde correlatie. Uit de formule blijkt bovendien dat het aantal items invloed heeft op de waarde die Cronbachs alpha aanneemt.

Formule voor Cronbachs alpha:

$$\alpha = \frac{k\bar{r}}{1 + (k - 1)\bar{r}}$$

In de formule staat k voor het aantal items, bijvoorbeeld 8 verschillende stellingen over de hang naar sensatie van een persoon. Daarnaast is \bar{r} (uitspraak: r streep) het gemiddelde van de correlaties tussen alle items met elkaar.

JoHo Samenvatting – Beschrijvende Statistiek

Interpretatie alpha

Alpha kleiner dan 0,60	onbetrouwbaar
Alpha tussen 0,60 en 0,80	redelijk
Alpha groter dan 0,80	betrouwbaar

Berekening in SPSS

Via Analyze → Scale → Reliability Analysis bereken je de betrouwbaarheid door het meten van een concept. In het Items scherm geef je aan welke variabelen (items) samen het concept moeten meten. Bijvoorbeeld de 8 variabelen om het concept 'hang naar sensatie' te meten.

Een nieuwe schaal maken

Wanneer uit de berekening van α blijkt dat de 8 variabelen samen het concept 'hang naar sensatie' goed meten, kunnen ze worden samengevoegd tot een schaal. Deze nieuwe variabele maak je door naar Transform → Compute te gaan. In het venster dat verschijnt geef je de nieuwe variabele een naam (Target Variable) en de manier waarop deze berekend moet worden (Numeric Expression).