

DE GEREEDSCHAPSKIST

CONFIRMATIEF VERSUS
EXPLORATIEF TOETSEN

Stel: u bent klinisch psycholoog en u beschikt over symtoomdata van depressieve patiënten. U wilt weten of etniciteit en sekse invloed hebben op de totaalscore van de depressiesymptomen, maar u heeft er vooraf geen echte hypothese over. Maakt het voor de statistische analyse uit of u vooraf wel of geen hypothese(n) heeft? Ja, dat maakt veel uit. Waarom zal ik zo uitleggen.

Even terug naar het voorbeeld: u heeft feitelijk drie zogeheten *nulhypothesen* die een statistisch programma als SPSS automatisch toetst: 1) etniciteit heeft geen invloed op depressie (geen hoofdeffect); 2) sekse heeft geen invloed op depressie (geen hoofdeffect) en 3) etniciteit en sekse hebben in interactie geen invloed op depressie (geen interactie-effect).

Een statistische analyse uitvoeren kan op twee manieren. Met de *exploratieve* manier heeft u vooraf geen hypothese: u weet simpelweg niet wat voor effect u verwacht. In dat geval toetst u de drie hypothesen in het voorbeeld tegelijk. Met de *confirmatieve* manier heeft u vooraf wél een hypothese: u verwacht bijvoorbeeld dat vrouwen hoger zullen sco-

'Leuker kunnen wij het als methodologen niet voor u maken, hopelijk wel makkelijker.' De nieuwe rubriek *De Gereedschapskist* wordt verzorgd door redactieraadlid Angélique Cramer, universitair docent aan de Programmagroep Psychologische Methodenleer van de Universiteit van Amsterdam. Deze eerste aflevering gaat over de vraag of het uitmaakt of je confirmatief of exploratief toetst.



ren dan mannen. In dat geval toetst u slechts één hypothese, namelijk het hoofdeffect van sekse.

De statistische analyse levert een *p*-waarde op voor elke hypothese: de kans dat u deze (of extremere) resultaten vindt indien de nulhypothese waar is. (U hoeft zich trouwens niet te generen als u dit niet goed had: veel onderzoekers en zelfs statistici kunnen niet altijd goed uitleggen wat een *p*-waarde betekent.) Deze *p*-waarde wordt vervolgens vergeleken met een significantieniveau *alpha*, deze wordt

meestal automatisch op .05 gezet. Is de *p*-waarde kleiner dan .05 (dus de kans dat u uw resultaten vindt als de nulhypothese waar zou zijn, is kleiner dan 5%), dan verwerpt u de nulhypothese dat sekse geen invloed heeft op depressie ten gunste van de alternatieve hypothese dat sekse die invloed wel heeft.

Terug naar de vraag: maakt exploratieve of confirmatieve toetsing uit? Jazeker, en wel voor *alpha*. Op de confirmatieve manier toetst u in het voorbeeld één hypothese: dan is het verantwoord *alpha* op

.05 te zetten. Toetst u echter exploratief en dus alle drie hypothesen tegelijk, dan moet *alpha* niet .05 zijn. Waarom niet? U toetst drie hypothesen met dezelfde data. Hierdoor neemt de kans op een toevalsbevinding toe: u verwerpt mogelijk een nulhypothese terwijl deze wel waar is. Daarom moet in het exploratieve geval *alpha* aangepast worden om te voorkomen dat u ten onrechte concludeert dat mannen hoger scoren op depressie dan vrouwen. U toetst dus strenger.

Er bestaan meerdere methoden om het significantieniveau aan te passen. Helaas zijn deze alleen niet geïmplementeerd in het veel gebruikte SPSS. Daardoor worden er in de onderzoekspraktijk helaas veel foute conclusies over data getrokken. Let in uw vakliteratuur dus goed op bij multiplere hypothesetoetsing: zonder aanpassing van *alpha* is de kans dat de gepresenteerde conclusies onterecht zijn echt een stuk groter dan die .05.

Cramer, A.O.J. et al. (2015). Hidden Multiplicity in Exploratory Multi-way ANOVA: Prevalence and Remedies. *Psychonomic Bulletin and Review*, 1-8. doi: <http://dx.doi.org/10.3758/s13423-015-0913-5>