

# Hoe bepaal je relaties in kwantitatief onderzoek?

## Grip op wetenschappelijk onderzoek (5)

In de afgelopen nummers heeft Anja Visser ons ingeleid in onderzoek doen naar geestelijke verzorging. Ze heeft daarin steeds de basisbegrippen en benaderingen van onderzoek uitgelegd, zodat je wetenschappelijke artikelen beter kunt begrijpen. Dit artikel is het laatste deel van de serie. Het gaat over complexe materie: het vinden van verbanden tussen meerdere variabelen in kwantitatief onderzoek.

### Anja Visser

**D**IT IS HET LAATSTE DEEL van een vijfdelige serie over empirisch onderzoek in de geestelijke verzorging. Doel van deze serie is je meer bekend te maken met verschillende vormen van onderzoek, zodat je inzichten uit onderzoek op waarde kunt schatten en kunt toepassen in de eigen praktijk. En misschien wil je zelf wel onderzoek doen? Dan helpt deze serie om je te oriënteren op welke onderzoeksbenadering zou kunnen passen bij wat jij wilt onderzoeken en geeft het een aantal aandachtspunten mee bij de planning van zo'n onderzoek.

In het vorige deel (zie nummer 2021/104) besprak ik hoe je vanuit gegevens die bijvoorbeeld met een vragenlijst zijn verzameld (zie nummer 2021/103) kunt onderzoeken of groepen verschillen op bepaalde eigenschappen. Zo kan bijvoorbeeld onderzocht worden welke vorm van begeleiding de grootste invloed heeft op spirituele

ontwikkeling. Of je kunt onderzoeken of begeleiding door geestelijk verzorgers een grotere of andere bijdrage levert aan het (spiritueel) welzijn van mensen dan begeleiding door bijvoorbeeld verpleegkundigen. Ik besprak de meest gebruikte toetsen hiervoor: de *t*-toets, de ANOVA, de Wilcoxon signed-rank toets of Mann-Whitney U toets en de Kruskal-Wallis toets en de Chi<sup>2</sup>-toets.

In dit artikel bespreek ik een andere familie statistische toetsen, namelijk de correlatieve toetsen. Doel van deze toetsen is te bepalen of er een relatie bestaat tussen twee of meer eigenschappen van mensen. Bijvoorbeeld of er een relatie bestaat tussen de mate van religiositeit of spiritualiteit van mensen en hun kwaliteit van leven, of tussen de hoeveelheid spirituele zorg die zorgverleners verlenen en hun gevoel van competentie hierin. Daarnaast kun je met deze toetsen ook onderzoeken wat een relatie tussen twee eigen-

## Websites en boeken over statistische toetsen

Versnellingsplan Onderwijsinnovatie met ICT, *Statistisch handboek studiedata* (<https://sh-studiedata.nl>).

Wikipedia, *Statistiek* (<https://nl.wikipedia.org/wiki/Categorie:Statistiek>).

Andy Field, *Discovering statistics using IBM SPSS* (2017).

Deborah Rumsey, *Statistiek voor dummies* (2014).

Nel Verhoeven, *Statistiek in stappen* (2021).

schappen mogelijk verklaart. Ervaren mensen die hoger scoren op religiositeit of spiritualiteit misschien meer kwaliteit van leven, doordat ze meer sociale steun ervaren, minder angstig zijn, gemakkelijker situaties in hun leven kunnen accepteren, lichamelijk gezonder zijn, enzovoort?

Ik bespreek twee typen analyses: de Pearson correlatie en de lineaire regressieanalyse. Ik zal niet uitgebreid ingaan op de formules die onderliggend zijn aan deze toetsen; als je daar meer over wilt weten, verwijst ik je graag naar enkele boeken en websites (zie kader). Ik zal vooral ingaan op wat er met deze toetsen getoetst wordt, wanneer deze toetsen wel en niet toegepast mogen worden en hoe je de getallen die in artikelen gerapporteerd worden hierover kunt interpreteren. Zo ben je goed toegerust om artikelen over dergelijke onderzoeken te begrijpen en kritisch te evalueren.

Dit keer gebruik ik een artikel van mezelf om een en ander te illustreren, namelijk 'Does spirituality reduce the impact of somatic symptoms on distress in cancer patients? Cross-sectional and longitudinal findings' (Visser, De Jager Meezenbroek & Garssen, 2018).<sup>1</sup> In dit onderzoek hebben mijn collega's en ik een mogelijke verklaring onderzocht voor de bevinding uit meerdere onderzoeken dat onder mensen met kanker spiritualiteit samenhangt met minder gevoelens van 'distress'.<sup>2</sup> Onze hypothese was dat mensen die hoger scoren op spiritualiteit beter kunnen omgaan met lichamelijke klachten en de ervaring dat de ziekte het leven bedreigt, waardoor ze minder distress ervaren. In termen van correlaties zou de relatie tussen lichamelijke klachten en distress dus zwakker moeten zijn onder mensen die hoog scoren op spiritualiteit, dan onder mensen die laag scoren op spiritualiteit.

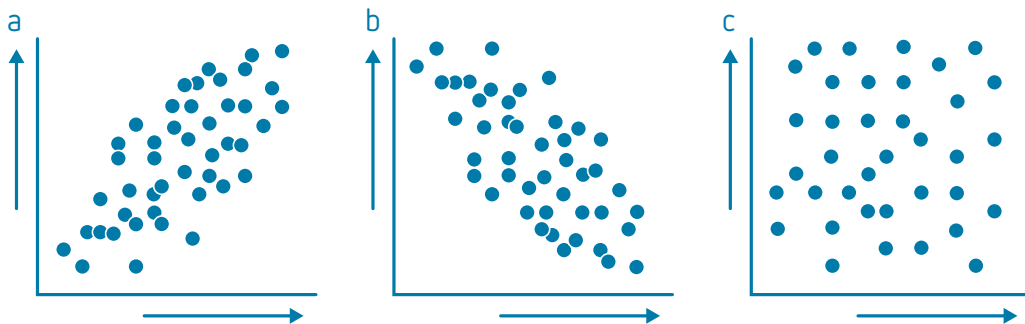


Figuur 1. Scan deze QR-code om naar het artikel van Visser e.a. (2018) te gaan.

## Pearson correlaties

Een correlatie (weergegeven met  $r$ ) wordt gebruikt om te bepalen of en hoe twee eigenschappen met elkaar samenhangen. Bij het berekenen van een correlatie zet het statistiekprogramma alle scores van mensen op de twee eigenschappen bij elkaar in een puntenwolk (zie figuur 2a-c). Dan gaat het na of er een patroon in de punten zit: is het zo dat mensen die hoger scoren op de ene eigenschap over het algemeen ook hoger scoren op de andere eigenschap (zie figuur 2a)? Of dat mensen die hoger scoren op de ene eigenschap over het algemeen juist lager scoren op de andere eigenschap (zie figuur 2b)?

Het getal dat uit deze berekening komt, de correlatiecoëfficiënt, kan variëren tussen -1 en 1. Is de correlatie tussen de -1 en 0, dan betekent dit dat er een negatief verband bestaat tussen de twee eigenschappen: als iemand hoog scoort op de een, dan scoort deze laag op de andere (of andersom). Is de correlatie tussen de 0 en 1, dan is er een positief verband tussen de twee eigenschappen: scoort iemand hoog op de een, dan scoort deze ook hoog op de ander. De grootte van het getal geeft de sterkte van het verband weer: tussen de 0 en 0,2 is het een zwak verband, van 0,2 tot 0,5 een matig verband, van



Figuur 2. Puntenwolken van a) een positieve correlatie ( $r = .60$ ), b) een negatieve correlatie ( $r = -.60$ ), en c) geen correlatie ( $r = .00$ ).

0,5 tot 0,8 een sterk verband en van 0,8 tot 1 een zeer sterk verband (Ferguson, 2009).

Bij een zeer sterk verband is het patroon overduidelijk te zien: iedereen met dezelfde score op de ene eigenschap heeft ook (ongeveer) dezelfde score op de andere eigenschap. Bij een zwak verband is dat patroon diffuser: mensen met dezelfde score op de ene eigenschap scores ongeveer hetzelfde op de andere eigenschap, maar er zit toch wel aardig wat verschil tussen mensen.<sup>3</sup> Als de correlatie dicht bij 0 is, dan betekent dit vaak dat er geen verband is tussen de eigenschappen (zie verderop voor een nuance hierop); mensen met een bepaalde score op de ene eigenschap verschillen veel in hun score op de andere eigenschap (zie figuur 2c).

Naast de effectgrootte (de correlatiecoëfficiënt) wordt bij correlaties ook altijd de  $p$ -waarde gerapporteerd; deze geeft aan of het verband statistisch significant is of niet. Bij  $p \leq 0,05$  is er sprake van een significant verband. Ruwweg betekent dit dat het onwaarschijnlijk is dat het gevonden verband op toeval berust.<sup>4</sup>

Het is belangrijk om goed te letten op hoe de onderzoekers de gevonden verbanden interpreteren. Correlaties worden namelijk vaak gebruikt om te onderzoeken of het waarschijnlijk is dat twee eigenschappen elkaar beïnvloeden. Dan wordt er geconcludeerd dat er een oorzakelijk verband is tussen eigenschappen. Strikt genomen zeggen correlaties echter niets over de vraag of er een oorzakelijk verband is tussen eigenschappen.

Het enige wat we kunnen vaststellen met een correlatie is of en in welke mate eigenschappen mogelijk verband houden met elkaar. Hoe of waarom ze verband houden blijft onbekend. Daarom wordt in lessen statistiek vaak de uitdrukking 'correlatie is geen causatie' gebruikt. Er kunnen allerlei andere redenen zijn waarom eigenschappen met elkaar samenhangen.

Een voorbeeld dat een van mijn statistiekdocenten gebruikte, was het verband tussen de hoeveelheid branden en brandweerauto's in een plaats. Hiertussen bestaat een sterk verband, maar dat is niet omdat brandweerauto's branden veroorzaken of andersom. Er is hier een derde eigenschap in het spel (mogelijk zijn er zelfs meer), bijvoorbeeld de grootte van de plaats: hoe groter de plaats, hoe vaker er branden zullen zijn en hoe meer brandweerauto's er nodig zullen zijn om het gebied te kunnen bestrijken. Ook kan het zijn dat de twee eigenschappen elkaar wederzijds beïnvloeden, zoals bij hongergevoelens en eten: hoe meer honger je hebt hoe meer je zult eten, maar hoe meer je eet hoe vaker je ook honger zult hebben (omdat je maag groter wordt). In beide gevallen is er sprake van een positief verband, maar er is geen duidelijke oorzaak of een duidelijk gevolg.

Het is ook belangrijk om alert te zijn op de conclusie die onderzoekers trekken over zwakke correlaties. Een van de aannames van een Pearson correlatieanalyse is namelijk dat het verband tussen de twee eigenschappen lineair is. Met andere woorden: er wordt gezocht naar een pa-

troon in de vorm van een rechte lijn in de data. Het kan echter heel goed zijn dat het verband tussen twee eigenschappen curvilineair is ofwel in een u-, n-, s-, of zelfs w-, m- of x-vorm. Er zijn bijvoorbeeld enige aanwijzingen dat de relatie tussen religiositeit/spiritualiteit en depressie u-vormig is: in twee studies rapporteerden mensen voor wie religiositeit of spiritualiteit niet belangrijk was en mensen voor wie dit erg belangrijk was ernstiger depressie dan mensen voor wie religiositeit of spiritualiteit matig belangrijk was (King e.a., 2007; Schnittker, 2001). Een correlatie tussen religiositeit/spiritualiteit en depressie zou in zo'n geval dicht bij 0 liggen, maar dat betekent niet dat er geen relatie is tussen de twee eigenschappen, enkel dat er geen lineaire relatie is. Mocht je zelf correlaties gebruiken, maak dan dus altijd ook een puntenwolk (een 'scatterplot') om te kijken of er inderdaad geen relatie lijkt te zijn tussen de eigenschappen of dat de relatie misschien een andere vorm heeft.

## Met deze toetsen kun je bepalen of er een relatie is tussen eigenschappen van mensen

Naast de aanname van lineariteit geldt voor Pearson correlatiecoëfficiënten ook de aanname dat de scores op beide eigenschappen een normaalverdeling volgen en dat beide eigenschappen op een continue schaal (ofwel op interval- of rationiveau) gemeten zijn.<sup>5</sup> Een normaalverdeling betekent dat de meeste mensen op of rond het gemiddelde scoren op de gemeten eigenschap, terwijl steeds minder mensen hoger of lager scoren. Als een eigenschap op een continue schaal gemeten is, dan lopen de scores op de eigenschap in een logische volgorde op van weinig naar veel (zoals bij leeftijd of temperatuur; zie ook nummer 2021/104). Als dit niet het geval is, dan wordt de schatting van de correlatie minder betrouwbaar. Er zijn dan twee alternatieve correlatieanalyses: de Spearman correlatie en Kendalls tau. Spearman correlaties (ook wel Spearman's rho genoemd, weergegeven met  $\rho$ ) en Kendalls tau

(weergegeven met  $\tau$ ) zijn beide gebaseerd op de rangordes van de scores. Beide kunnen ook gebruikt worden als de eigenschappen op ordinaal niveau zijn gemeten, in plaats van op een continue schaal (zie ook nummer 2021/104).<sup>6</sup> Het verschil tussen deze twee alternatieve correlatiecoëfficiënten zit in de eigenschap van de rangorde waarop de berekening gebaseerd is. (Voor meer informatie zie de bronnen in het kader.)

### Lineaire regressieanalyse

Heel vaak wil je niet weten hoe een enkele eigenschap samenhangt met een andere, maar wil je dat van meerdere tegelijk weten. Als je bijvoorbeeld wilt weten welke factoren beïnvloeden hoe spiritueel iemand is, wil je rekening houden met leeftijd, geslacht, socio-economische status, mate van spiritualiteit van het gezin waarin iemand opgegroeid is, mentale en fysieke gezondheid, enzovoort. Al deze factoren kunnen de mate van spiritualiteit van een persoon verklaren. In zo'n geval gebruik je lineaire regressieanalyse.

Je kunt een lineaire regressieanalyse vergelijken met een serie correlaties. In een regressieanalyse onderzoek je of en in welke mate de scores op een eigenschap (de afhankelijke variabele) verklaard kunnen worden door de scores op een x-aantal andere eigenschappen (de onafhankelijke variabelen).<sup>7</sup> Hier doet de onderzoeker dus meteen een aanname over wat mogelijk oorzaak (de onafhankelijke variabelen) en gevolg (de afhankelijke variabele) zijn. Dit betekent echter niet dat dit ook zo is. Net als bij correlaties kan het heel goed zijn dat de eigenschappen elkaar wederzijds beïnvloeden en dat, als je dus een van de onafhankelijke variabelen zou verwisselen met de afhankelijke variabele, je dezelfde relatie tussen deze twee eigenschappen vindt. Let dus ook bij conclusies over regressieanalyses goed op hoe de onderzoekers de resultaten interpreteren. Ook hier geldt: correlatie is geen causatie.

Daarnaast geldt bij lineaire regressieanalyse de aanname van lineariteit. Een zwakke relatie betekent dus niet per definitie geen relatie, enkel geen lineaire relatie. Ten slotte geldt voor regressieanalyse dat de afhankelijke variabele als een continue schaal gemeten moet zijn.<sup>8</sup> Een kracht van regressieanalyse is daarbij dat de onafhanke-



lijke variabelen op allerlei verschillende niveaus gemeten kunnen zijn. Je kunt dus zowel onderzoeken of de mate van spiritualiteit verklaard wordt door continue variabelen (zoals leeftijd, aantal jaren religieus onderwijs en mate van welbevinden), als door categorische variabelen (zoals gender, huwelijkse status en opleidingsniveau).

In een regressieanalyse wordt dus de relatie tussen een afhankelijke variabele en meerdere onafhankelijke variabelen onderzocht. Niet alle onafhankelijke variabelen in een analyse hebben echter dezelfde status. Laat ik mijn eigen onderzoek (Visser e.a., 2018) hier als voorbeeld nemen. In dit onderzoek hebben we meerdere regressieanalyses uitgevoerd. De afhankelijke variabele was altijd de mate van distress die mensen met kanker rapporteerden. De onafhankelijke variabelen wisselden per analyse, maar bestonden

altijd uit sociodemografische eigenschappen, medische eigenschappen, kankergelateerde ervaringen (pijn, vermoeidheid en/of ervaren levensbedreiging), aspecten van spiritualiteit en zogenaamde interactievariabelen, waarmee we de interactie tussen de kankergelateerde ervaringen en spiritualiteit nader wilden onderzoeken (ik leg straks uit waarom we dat deden).

Wat we echt wilden weten, was de relatie tussen de interactievariabelen en distress. Al die andere relaties vonden we helemaal niet zo interessant. Waarom hebben we dan toch al die variabelen in de analyse meegenomen, vooral als je bedenkt dat de schatting van de effectgrootte (de regressiecoëfficiënt) steeds onnauwkeuriger wordt naarmate er meer variabelen in een analyse zitten? Dit deden we om twee redenen: controle en 'confounding'. Controlevariabelen zijn eigen-

schappen waarvan je weet dat die de score op de afhankelijke variabele ook bepalen en die je dus als het ware wilt uifilteren; gevoelens van distress hangen immers met een heleboel dingen samen. Dus, gegeven de invloed van deze variabelen op distress, wilden we weten welk resterend deel van de variatie in distress verklaard kon worden door de interactie tussen symptomen en spiritualiteit. In ons onderzoek bestonden de controlevariabelen bijvoorbeeld uit het hebben van een partner en de ervaren prognose.

## Je kunt ook onderzoeken wat een relatie tussen twee eigenschappen mogelijk verklaart

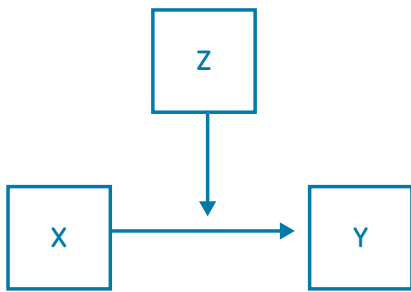
Confounding-variabelen zijn eigenschappen die een alternatieve verklaring kunnen vormen voor de gevonden relaties, de beruchte derde eigenschap waar ik al over sprak. Een voorbeeld hiervan is het ervaren van andere levensgebeurtenissen ten tijde van de diagnose en behandeling van kanker. Als dit een rol had gespeeld, dan waren de conclusies van ons onderzoek misschien alleen van toepassing op mensen met kanker die ook andere levensgebeurtenissen hadden meegemaakt. Dan was er dus veel onzekerheid geweest over de zeggenschap van de resultaten. Al deze variabelen worden daarom meegenomen in de analyse, om de schatting van de effectgrootte van de relatie waarin je geïnteresseerd bent nauwkeuriger te maken.

Wat je graag wilt zien in een wetenschappelijk artikel waarin regressieanalyse wordt toegepast, is dat de auteurs deze verschillende doelen van de variabelen in hun analyse beschrijven. Zoals gezegd, worden de resultaten van analyses namelijk steeds onnauwkeuriger naarmate er meer variabelen in worden opgenomen (dit wordt 'power' genoemd). De effectgrootten worden dan steeds moeilijker te interpreteren. Onderzoekers moeten dus doelbewuste keuzes maken. Ook ligt het gevaar van 'vissen' op de loer, als dit niet gedaan wordt; net zolang analyseren of variabelen

aan je analyse toevoegen tot er een statistisch significante relatie gevonden wordt. Deze is tegen die tijd echter compleet betekenisloos, vanwege de kanskapitalisatie – een statistisch fenomeen waarin de kans steeds groter wordt dat je per toeval een statistisch significant verband vindt in plaats van een 'werkkelijk' verband – en is ook theoretisch vaak moeilijk te plaatsen. Daar wordt de wetenschap niet mee verder geholpen en de praktijk al helemaal niet.

Naast dat er verschillende soorten variabelen om verschillende redenen in regressieanalyse worden opgenomen, kunnen er ook verschillende soorten relaties onderzocht worden in regressieanalyse. Het meest voorkomende zijn directe relaties. Heeft de mate van zin in het leven die mensen rapporteren bijvoorbeeld een relatie met de gerapporteerde mate van distress? Er wordt echter steeds vaker gezocht naar andere soorten relaties, om nog meer inzicht te krijgen in hoe de mate van (bijvoorbeeld) distress van mensen verklaard kan worden. Een van deze manieren is moderatie. Bij moderatie wordt de sterkte van de invloed van een eigenschap op de afhankelijke variabele beïnvloed door een andere eigenschap. Dit is waar we in ons onderzoek naar op zoek waren: wordt de sterkte van de relatie tussen pijn en distress beïnvloed door (onder andere) de mate waarin mensen goede en minder goede dingen in het leven accepteren? Als we moderatie schematisch zouden weergeven, dan zou dat er uitzien als in figuur 3: de relatie tussen X en Y wordt beïnvloed door Z.

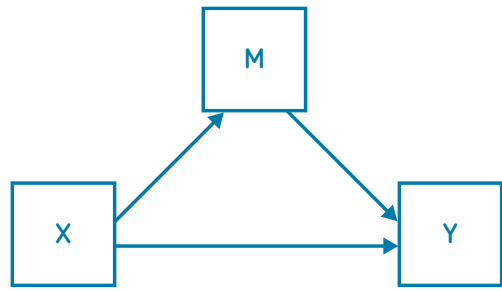
Een andere (veel onderzochte) relatievorm is mediatie. Bij mediatie wordt gezocht via welke andere eigenschappen de onafhankelijke variabele invloed heeft op de afhankelijke variabele. We weten bijvoorbeeld dat kerkbezoek samenhangt met minder depressieve klachten of sneller herstel van depressie, doordat mensen die de kerk bezoeken meer sociale steun vinden. Als we mediatie schematisch zouden weergeven, zou dat er uitzien als in figuur 4: X beïnvloedt M, wat vervolgens Y beïnvloedt. Er zijn ook allerlei combinaties mogelijk tussen deze soorten relaties: zowel mediatie als een directe relatie, gemodererde mediatie, enzovoort, maar dat voert voor nu te ver.



Figuur 3. Moderatie van de relatie tussen X en Y door Z.

Veel onderzoekers die dit soort relaties onderzoeken, geven in hun artikelen figuren weer zoals in figuur 3 en 4. Ze zetten dan de effectgrootte en de statistische significantie bij de verschillende paden (lijnen) die ze onderzocht hebben, zodat de lezer in één oogopslag het resultaat van het onderzoek kan begrijpen. Het is heel handig om deze figuren te bestuderen, omdat ze je ook helpen te bepalen of de onderzoekers de juiste conclusies uit hun onderzoek trekken, zonder dat je de hele tekst van het artikel hoeft te lezen.

In het ideale geval wordt er in artikelen over regressieanalyse een aantal cijfers gerapporteerd. Allereerst de verklaarde variantie ( $R^2$ ), die loopt van 0 tot 1. De verklaarde variantie is een indicatie van hoeveel procent van de spreiding tussen de scores op de afhankelijke variabele verklaard wordt door de onafhankelijke variabele(n). Is dit heel laag, dan betekent dit dat er nog veel andere factoren bijdragen aan de eigenschap die onderzocht wordt, dan datgene wat in dit onderzoek is meegenomen en dat dit dus maar een beperkte ‘verklaring’ van het fenomeen biedt. In ons onderzoek liep dit bijvoorbeeld uiteen van 15 tot 38 procent. Hoewel dat normaal is voor psychologisch onderzoek, betekent dit dat het overgrote deel van de verschillen tussen de deelnemers in hun mate van distress onverklaard blijft door dit onderzoek. Soms wordt er een *F*-toets gerapporteerd over de verklaarde variantie, die aangeeft of deze mate van verklaring statistisch significant (enigszins betekenisvol) is. Als deze toets niet statistisch significant is ( $p > .05$ ), is het zeer de vraag wat de praktische en theoretische waarde is van de resultaten.



Figuur 4. De relatie tussen X en Y wordt gedeeltelijk gemedieerd door M.

Vervolgens wil je kijken naar de regressiecoëfficiënten en hun statistische significantie. Er zijn twee regressiecoëfficiënten die gerapporteerd kunnen worden: de ongestandaardiseerde ( $B$ ) en de gestandaardiseerde ( $\beta$  ofwel  $\beta$ ). De  $B$  kan in principe alle mogelijke waarden aannemen en in de schatting hiervan is geen rekening gehouden met de relatie die de andere variabelen in de analyse hebben met de afhankelijke variabele. Deze waarde is dus niet heel informatief. Waar je bij een  $B$  op wilt letten is of deze positief of negatief is en of deze statistisch significant is ( $p \leq .05$ ).

Een positieve of negatieve regressiecoëfficiënt wordt op dezelfde manier geïnterpreteerd als een correlatiecoëfficiënt: positief betekent dat een hogere score op de ene eigenschap samen gaat met een hogere score op de andere, en negatief betekent dat een hogere score op de ene eigenschap samen gaat met een lagere score op de andere. Een  $\beta$  kan, net als een correlatiecoëfficiënt variëren tussen -1 en 1 en de schatting is gecorrigeerd voor de relaties van de andere variabelen in de analyse met de afhankelijke variabele. De waarde van de  $\beta$  geeft dus informatie over hoe sterk de relatie tussen deze variabele en de afhankelijke variabele is, gecontroleerd voor de invloed van de andere variabelen. Zowel de precieze waarde als de statistische significantie van  $\beta$  is dus informatief.

Tot slot zie je in de beste artikelen ook de betrouwbaarheidsintervallen rond  $B$  gerapporteerd staan (95 procent CI). Betrouwbaarheidsintervallen geven informatie over de mogelijke waarden van  $B$  in de populatie. Als deze 0 (‘nul’)

Toets	Wanneer toepassen	Voorwaarden	Minimaal gerapporteerde getallen
(Pearson) correlatie	Relatie tussen twee eigenschappen	<ul style="list-style-type: none"> <li>Eigenschappen zijn gemeten op een schaal</li> <li>Scores volgen ongeveer een normaalverdeling</li> <li>Verwachte relatie is lineair</li> </ul>	$r$ = effectgrootte $p$ = statistische significantie
(Lineaire) regressie-analyse	Relatie tussen meer dan twee eigenschappen, met een vermoeden van causaliteit	<ul style="list-style-type: none"> <li>De afhankelijke variabele is gemeten op een schaal</li> <li>Scores op de afhankelijke variabele volgen ongeveer een normaalverdeling</li> <li>Verwachte relatie is lineair</li> </ul>	$R^2$ = verklaarde variantie in de afhankelijke variabele $F$ = toetswaarde bij $R^2$ $B$ = ongestandaardiseerde effectgrootte $\beta$ = gestandaardiseerde effectgrootte $p$ = statistische significantie 95 procent CI = betrouwbaarheidsinterval rond $B$

Tabel 1. Overzicht van de besproken toetsen voor relaties met hun eigenschappen.

bevat, dan geeft dit aan dat de relatie niet statistisch significant is. Als de waarde van  $B$  in de populatie o zou kunnen zijn, dan geeft dit namelijk aan dat er een kans is dat er geen lineaire relatie is tussen de eigenschappen die onderzocht worden. Er is onder statistici veel discussie

## Zo kun je artikelen over dergelijke onderzoeken begrijpen en kritisch evalueren

over of het niet beter zou zijn als onderzoekers altijd betrouwbaarheidsintervallen zouden rapporteren in plaats van (of in ieder geval samen met)  $p$ -waarden. Dat gesprek is nog niet ten einde. Mijn advies is om, als ze gerapporteerd worden, hier ook zeker naar te kijken, want soms verschillen de conclusies op basis hiervan van de conclusies op basis van de  $p$ -waarde. De consensus is dat betrouwbaarheidsintervallen in dat geval nauwkeuriger zijn.

### Tot besluit

In tabel 1 staat een overzicht van de twee besproken toetsen en hun belangrijkste eigenschappen. Het is handig om dit artikel (en het vorige in

nummer 2021/104) te gebruiken als referentie bij het lezen van wetenschappelijke artikelen waarin verbanden worden onderzocht. Zo kun je controleren of de gekozen toets past bij de vraag die de onderzoekers stellen en of de resultaten die in de tabellen staan passen bij de conclusie. Ik hoop dat je met deze serie van artikelen weer en/of meer grip hebt gekregen op empirisch wetenschappelijk onderzoek en dat je erdoor geïnspireerd bent geraakt om voorbij de samenvatting en/of conclusie van wetenschappelijke artikelen te lezen.

### Dankwoord

Graag wil ik de redactie van dit tijdschrift bedanken voor hun steun aan het idee voor deze serie en hun constructieve feedback. Ook mijn dank aan Brenda Mathijssen voor haar bijdrage aan deel 2 van deze serie. Ontzettend leuk om ons college zo in geschreven vorm terug te zien en onze visies op kwalitatief onderzoek nader uit te kunnen wisselen. Ten slotte dank aan de studenten in de cursus ‘Conducting research on religion, health and wellbeing’, aan wie ik de uitleg van de thema’s in deze serie de afgelopen jaren heb mogen scherpen.

Dr. A. Visser is universitair docent geestelijke verzorging aan de Rijksuniversiteit Groningen. E-mail: a.visser-nieraeth@rug.nl.



## Literatuur

Ferguson, C.J. (2009). An effect size primer: a guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40 (5), 532-538.

King, D.A., Lyness, J.M., Duberstein, P.R., He, H., Tu, X.M. & Seaburn, D.B. (2007). Religious involvement and depressive symptoms in primary care elders. *Psychological Medicine*, 37 (12), 1807-1815.

Schnittker, J. (2001). When is faith enough? The effects of religious involvement on depression. *Journal for the Scientific Study of Religion*, 3, 393-411.

Visser, A., De Jager Meezenbroek, E.C. & Garssen, B. (2018). Does spirituality reduce the impact of somatic symptoms on distress in cancer patients? Cross-sectional and longitudinal findings. *Social Science and Medicine*, 214 (augustus), 57-66.

## Noten

1. Dit artikel is gratis te downloaden via <https://research.rug.nl/en/publications/does-spirituality-reduce-the-impact-of-somatic-symptoms-on-distre>.
2. 'Distress' wordt door de American Psychological Association beschreven als: 'Een negatieve stressreactie die gepaard gaat met negatieve gevoelens en fysiologische reactiviteit: een type stress die mensen ervaren als ze overweldigd worden door eisen, verliezen of waargenomen bedreigingen. Het lokt negatieve fysieke en psychologische aanpassingsstrategieën uit die een groot gezondheidsrisico kunnen vormen' (<https://dictionary.apa.org/distress>, 2 augustus 2021). Helaas bestaat er geen adequaat Nederlands woord voor, daarom gebruik ik hier de onvertaalde term.
3. In de (godsdienst)psychologie zijn correlaties vaak zwak tot matig. Als een correlatie sterk of zeer sterk is, dan is dat meestal een reden om goed na te gaan welke conclusies de onderzoekers trekken. De kans is dan vrij groot dat iemand twee dezelfde of zeer vergelijkbare eigenschappen heeft gemeten. Conclusies over mogelijke 'invloeden' van de eigenschappen op elkaar zijn dan vaak niet gerechtvaardigd.
4. De werkelijke betekenis van de  $p$ -waarde is iets complexer. De  $p$ -waarde geeft de waarschijnlijkheid weer dat de nulhypothese verworpen wordt, onder de aanname dat de nulhypothese waar is. Statistische toetsen vergelijken de informatie die bij de deelnemers

(de steekproef) verzameld is met een veronderstelling (de nulhypothese). Bij toetsen over samenhang is de nulhypothese meestal dat er geen verband is tussen de onderzochte eigenschappen. Deze hypothese wordt verworpen als de verzamelde gegevens te ver afwijken van de getallen die je op basis van de veronderstelling zou verwachten. Het verband tussen de eigenschappen is dan statistisch significant.

5. Bij eigenschappen op een ratiomeetniveau is er sprake van een absoluut nulpunt (bijvoorbeeld: een lengte van 0 is de afwezigheid van lengte). Bij eigenschappen op een intervalmeetniveau is er geen absoluut nulpunt (bijvoorbeeld: een temperatuur van 0°C is niet de afwezigheid van temperatuur).
6. Het is bij ordinale variabelen niet helemaal duidelijk hoeveel meer er is van een eigenschap als iemand hoger scoort, en er hadden ook andere getallen gekozen kunnen worden om de antwoorden mee te coderen, maar er is wel sprake van een logische volgorde in de scores. Een voorbeeld hiervan is tevredenheid, gecodeerd als 1 = erg ontevreden, 2 = enigszins ontevreden, 3 = niet tevreden/niet ontevreden, 4 = enigszins tevreden, en 5 = erg tevreden.
7. Ik ga hier uit van meervoudige regressieanalyse, omdat deze het meest gebruikt wordt. Hierbij is er sprake van één afhankelijke variabele en meerdere onafhankelijke variabelen. Het is echter ook mogelijk om enkelvoudige regressieanalyse uit te voeren. Dit lijkt op een correlatie, behalve dat er door de onderzoeker een keuze wordt gemaakt voor wat oorzaak en gevolg is. In enkelvoudige regressieanalyse kan er dus een sterkere uitspraak worden gedaan over de bijdrage van een eigenschap aan de score op een andere eigenschap dan bij correlaties. Er is echter nog steeds geen zekerheid over causaliteit.
8. Ik ga hier uit van lineaire regressieanalyse. In logistische regressieanalyse is de afhankelijke variabele op een dichotome schaal gemeten (bijvoorbeeld man of vrouw, wel of geen depressieve stoornis, wel of geen ziekte). De relatie volgt dan een  $s$ -vorm. In een logistische regressieanalyse wordt berekend of de onafhankelijke variabele(n) de kans vergroot dat iemand in de ene of de andere categorie valt. Het is ook mogelijk om in een lineaire regressieanalyse non-lineaire verbanden tussen de onafhankelijke en afhankelijke variabele te onderzoeken, maar daar zal ik hier verder niet op ingaan.