



## De Chi-squaretoets

Sil Aarts · E. Wouters

Published online: 1 September 2017  
© Bohn Stafleu van Loghum, onderdeel van Springer Media B.V. 2017

**Samenvatting** Bij toetsende statistiek wordt, met behulp van een statistische berekening, een hypothese getoetst. Er wordt bijvoorbeeld bepaald of het verschil dat we zien tussen twee groepen op toeval berust of niet. Of er wordt bepaald of de samenhang die we zien tussen twee variabelen wel of niet toevallig is. De keuze welke toetst van toepassing is op de onderzoeksresultaten is onder meer afhankelijk van het meetniveau van de variabelen. De Chi-squaretoets is van toepassing bij nominale, dichotome en ordinale variabelen en wordt in dit artikel nader toegelicht. De tabellen in dit artikel lijken op die in het SPSS-programma, maar zijn wel esthetisch aangepast.

**Trefwoorden** Chi-squaretoets · kruistabel · toetsende statistiek

### Inleiding

In het vorige artikel over statistisch toetsen is in vogelvlucht een overzicht gegeven van de verschillende toetsen die gebruikt worden bij het analyseren van onderzoeksdata [1]. In het huidige artikel zal dieper worden ingegaan op een van die toetsen: de Chi-squaretoets.

Met de Chi-squaretoets (ook wel de kruistabel-analyse of Chi-kwadraattoets genoemd) kan berekend worden of een gevonden verschil statistisch significant is. Bij een Chi-squaretoets kan er gerekend worden

met nominale, dichotome of ordinale variabelen (zie kader 'Soorten variabelen'). Als er sprake is van ratio- en intervalvariabelen is het verstandiger om te kiezen voor een andere statistische toets, zoals een T-toets of een ANOVA (meer hierover in het volgende artikel). Met behulp van een Chi-squaretoets kunnen er dus vragen vanuit de praktijk worden beantwoord, die te vertalen zijn naar nominale of ordinale variabelen.

Om het verschil tussen de variabelen te verduidelijken, gaan we uit van het volgende voorbeeld. Een student voert een afstudeeronderzoek uit. Ze wil nagaan hoe tevreden patiënten zijn over een nieuwe therapie die gericht is op het voorkomen van valinciden-ten. Daarbij wil zij ook kijken of mannen en vrouwen verschillen in hoe tevreden zij zijn over deze therapie. Er wordt een vragenlijst ontwikkeld. Deze wordt verstuurd naar verschillende podotherapeuten in het land, met de vraag of zij deze onder hun patiënten willen verspreiden. Uiteindelijk ontvangt de student 161 ingevulde vragenlijsten retour. Een van de vragen luidde: 'Hoe tevreden bent u over de podotherapie die u ontvangt?' De drie antwoordmogelijkheden waren:

1. 'ontevreden',
2. 'tevreden',
3. 'heel tevreden'.

Zoals te zien is in tab. 1 (uitgevoerd met behulp van het statistiek programma SPSS [2] [alle SPSS-tabellen zijn esthetisch aangepast]), hangt de Chi-squaretoets nauw samen met de kruistabel. In deze tabel wordt de samenhang tussen geslacht en tevredenheid getoond. Deze tabel wordt ook wel een 2×3-tabel genoemd; twee mogelijkheden in de rijen (man *vs.* vrouw) en drie mogelijkheden in de kolommen ('ontevreden'/'tevreden'/'heel tevreden'). Een vak waarin cijfers zijn opgenomen, wordt ook wel een 'cel' genoemd. In deze tabel worden de werkelijk verkregen waarden getoond (in SPSS ook wel *observed*

---

In deze rubriek dragen de auteurs een steentje bij aan het vergroten van de kennis over wetenschappelijk onderzoek en de toepasbaarheid ervan in de podotherapeutische praktijk.

S. Aarts, PhD MD (✉) · E. Wouters, PhD MD  
Health Innovations & Technology, Fontys Paramedische Hogeschool, Eindhoven, Nederland  
s.aarts@fontys.nl



**Soorten variabelen**

- Nominale variabelen zijn variabelen die slechts een naam hebben, een 'benoeming'. Voorbeelden hiervan zijn: land van herkomst (Europa/Azië/Amerika) of kleur ogen (groene ogen/ blauwe ogen/bruine ogen).
- Dichotome variabelen zijn variabelen met een naam, een benoeming, met slechts twee waarden. Voorbeelden hiervan zijn: ja *vs.* nee, tevreden *vs.* niet tevredenheid, diabetes *vs.* niet diabetes.
- Ordinale variabelen beschrijven variabelen met een naam, een benoeming, die een natuurlijke rangorde hebben. Voorbeelden hiervan zijn: opleidingsniveau (mavo/havo/vwo) of rangen in het leger. Bij dit soort variabelen is er sprake van een volgorde: vwo duidt op een hogere opleiding dan havo, maar het verschil tussen vwo en havo is mogelijk niet hetzelfde als het verschil tussen mavo en havo.
- Interval- en ratiovariabelen zijn variabelen met numerieke waarden. Deze worden ook wel numerieke of continue variabelen genoemd. Bij intervalvariabelen zijn de intervallen tussen de waarden gelijk, maar is er geen absoluut nulpunt (bijv. graden Celsius). Zo is het verschil tussen 25° en 30° graden en 15° en 20° graden in beide gevallen 5° graden, maar kan er niet worden geconcludeerd dat 10° graden twee keer zo koud is als 20° graden. Bij ratiovariabelen is er ook sprake van gelijke intervallen, maar hier heeft het nulpunt een absolute waarde (bijv. leeftijd in jaren). Zo is het verschil in leeftijd tussen kinderen van 5 en 10 jaar hetzelfde als tussen kinderen van 10 en 15 jaar. Echter, nu kan er worden gezegd dat iemand van 10 jaar twee keer zo oud is als iemand van 5 jaar en dat iemand van 15 jaar drie keer zo oud is als iemand van 5 jaar.

*count* genoemd) die met de vragenlijst zijn verzameld. In totaal hebben 161 mensen deelgenomen; 79 mannen en 82 vrouwen. Een grote meerderheid van zowel de mannen als de vrouwen is zeer tevreden over de therapie.

**Tabel 1** Kruistabel uit SPSS van de relatie tussen geslacht en tevredenheid.

		Tevredenheid			Totaal
		1	2	3	
Geslacht	1	7	7	65	79
	2	11	8	63	82
Totaal		18	15	128	161

Geslacht: 1 = man, 2 = vrouw  
 Tevredenheid: 1 = ontevreden, 2 = tevreden, 3 = zeer tevreden

Zoals te lezen is in het voorgaande artikel over toetsende statistiek [1], wordt een onderzoeksvraag beantwoord door het toetsen van een hypothese. Op basis van eerdere kennis- en praktijkervaring wordt de *nulhypothese*,  $H_0$ , opgesteld. Dat wil zeggen: de veronderstelling dat er geen verschil (tussen de groepen) of samenhang (tussen de variabelen) bestaat. In dit geval houdt de  $H_0$  dus in dat er géén statistisch significant verschil bestaat tussen mannen en vrouwen in hun tevredenheid. Het doel van een statistische toets is nagaan of de nulhypothese verworpen wordt. Als de  $H_0$  verworpen wordt, neemt men de *alternatieve hypothese*,  $H_1$ , aan. Deze hypothese benoemt dat er juist wél een verschil is tussen mannen en vrouwen in hun tevredenheid.

De volgende hypothesen worden geformuleerd:

- $H_0$  Er is geen verschil tussen mannen en vrouwen in hun tevredenheid over de therapie die zij ontvangen.
- $H_1$  Er is wel een verschil tussen mannen en vrouwen in hun tevredenheid over de therapie die zij ontvangen.

We gaan er vanuit dat de variabelen 'geslacht' en 'tevredenheid' onafhankelijk zijn; de student heeft geen reden om aan te nemen dat mannen meer tevreden zijn over de nieuwe therapie dan vrouwen of omgekeerd. Als deze veronderstelling juist is, zouden de geobserveerde waarden overeen moeten komen met de zogenoemde te verwachten waarden (ook wel *expected count* genoemd). De verwachte waarden zijn de waarden die we verwachten, gegeven het feit dat de twee variabelen 'geslacht' en 'tevredenheid' volledig onafhankelijk zijn.

De verwachte waarden worden als volgt berekend: het totaal van de rij wordt vermenigvuldigd met het totaal van de kolom, en deze uitkomst wordt gedeeld door  $n$  (het aantal deelnemers). Voor mannen die ontevreden zijn, geldt bijvoorbeeld:  $(18 \times 79)/161 = 8,8$ .

Als ook de te verwachten waarden in de tabel opgenomen worden, is de tabel ineens een stuk complexer geworden: zie tab. 2.

Als er geen verschil zou zijn tussen mannen en vrouwen in hun tevredenheid (in dat geval zijn de twee variabelen dus volledig onafhankelijk en is er geen statistisch significant verschil tussen mannen en vrouwen in hun tevredenheid), dan zouden, op basis van het totaal aantal deelnemers (er hebben meer vrouwen dan mannen deelgenomen aan het onder-

**Tabel 2** Kruistabel uit SPSS van de relatie tussen geslacht en tevredenheid, inclusief de verwachte waarden.

			Tevredenheid			Totaal
			1	2	3	
Geslacht	1	Count	7	7	65	79
		Expected count	8,8	7,4	62,8	79,0
	2	Count	11	8	63	82
		Expected count	9,2	7,6	65,2	82,0
Totaal		Count	18	15	128	161
		Expected count	18,0	15,0	128,0	161,0

*Geslacht 1 = man, 2 = vrouw. Tevredenheid 1 = ontevreden, 2 = tevreden, 3 = zeer tevreden. Count geobserveerde waarden. Expected count verwachte waarden (met decimalen weergegeven)*

**Tabel 3** Chi-squaretoets uit SPSS voor de relatie tussen geslacht en tevredenheid.

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	0,931	2	<b>0,628</b>
Likelihood Ratio	0,938	2	0,625
Linear-by-Linear Association	0,909	1	0,340
N of Valid Cases	161		

0 cells (0,0 %) have expected count less than 5. The minimum expected count is 7,36

zoek), de drie categorieën volstrekt evenredig verdeeld zijn. Als we in dit specifieke geval de geobserveerde waarden vergelijken met de verwachte waarden, dan verschillen deze, op het oog, weinig.

Let op: als ook percentages getoond worden, kan er direct worden afgelezen hoeveel procent van de mannen tevreden is en hoeveel procent van de vrouwen. De tabel wordt dan uiteraard nog groter en complexer. Het is niet moeilijk voor te stellen wat er met de tabel gebeurt als de nominale of ordinale variabelen worden vervangen door ratio- of intervalvariabelen (bijv. lengte in cm); de tabel zou dan erg groot en enorm complex worden. Er geldt dus: hoe groter de tabel wordt, hoe moeilijker leesbaar deze is en hoe moeilijker de uitkomsten te interpreteren zijn.

### Hoe ziet de Chi-squaretoets eruit?

De kruistabellen uit ons voorbeeld (tab. 1 en 2) laten de relatie zien tussen geslacht en tevredenheid. Is die relatie nu statistisch significant of niet? Hoewel de waarden, op het oog, niet veel van elkaar lijken te verschillen, kunnen we deze vraag alleen beantwoorden als we een statistische toets uitvoeren. Dat kan met de Chi-squaretoets, want deze vergelijkt alle geobserveerde waarden met de verwachte waarden.

Tab. 3 zit er uitgebreid uit, maar toch zijn er slechts twee begrippen voor nodig om deze tabel te begrijpen: de *Pearson Chi-square*, de naam van de toets, en de *Asymp. Sig.*, een afkorting voor significantie. Om de uitkomst, de *p*-waarde, van de statistische toets te vinden, kijken we dan ook naar de cel met de uitkomst van deze twee (rood in SPSS, hier: vet). Hier zien we een *p*-waarde van 0,628.

### Wat betekenen de uitkomsten van deze toets?

De *p*-waarde is, in ons voorbeeld, de kans dat het gevonden verschil tussen mannen en vrouwen in hun tevredenheid berust op toeval. De kans, uitgaande van het feit dat er geen verschil is tussen mannen en vrouwen in de tevredenheid over de nieuwe therapie, is 62,8 % ( $0,628 \times 100$ ). Dit is een vrij grote kans. Als de kans zo groot is, is er weinig 'evidence' voor de  $H_1$ . (In de wetenschap wordt niet gesproken over 'waar' vs. 'niet waar', maar over 'evidence'. Zie ons artikel *Evidence-based practice in de dagelijkse praktijk* [3].) In dat geval wordt  $H_1$  verworpen en blijft de  $H_0$  gelden.

De conclusie luidt dan ook dat er geen statistisch significant verschil is in de tevredenheid tussen mannen en vrouwen bij een significantieniveau van 5 %. De uitspraak 'er is een significant verschil', is helaas onvolledig, aangezien het woord 'significant' slechts 'belangrijk' betekent. Een statistisch significant verband wil niet per se wijzen op een belangrijk verband (en vice versa) [4]. Een verband kan bijvoorbeeld ook statistisch significant zijn, maar klinisch irrelevant (meer hierover in een volgend artikel).

### Voorwaarden

Aan tab. 3 zijn nog een paar dingen opvallend en van belang. Onder de tabel staan de volgende gegevens vermeld: '0 cells (0,0 %) have expected count less than 5. The minimum expected count is 7,36.' Wat betekent dit nu eigenlijk? Alvorens een Chi-squaretoets kan worden uitgevoerd, dient aan een aantal belangrijke voorwaarden te worden voldaan. Ten eerste dient geen enkele verwachte celfrequentie (aantal per cel) kleiner te zijn dan 1 (dus 0). Tevens mag niet meer dan 20 % van de cellen een verwachte waarde hebben van minder dan 5. Zoals te zien in tab. 3, wordt in de uitvoer aangegeven of er voldaan wordt aan deze twee voor-

waarden. Wordt er niet voldaan aan de voorwaarden, dan kan een Chi-square niet worden uitgevoerd. Er kan dan worden uitgeweken naar andere statistische toetsen (zoals de Fisher's exact test).

### Vooruitblik

Het volgende artikel zal uitvoerig ingaan op T-toetsen en variantieanalyse (ANOVA).

Dit artikel is mede gebaseerd op:

- Field A. *Discovering statistics using SPSS*. Deel 3. 3e druk. London: Sage Publications Ltd.; 2009.
- Howell DC. *Statistical methods for psychology*. International edition. 8e druk. Boston, Massachusetts, VS: Cengage Learning, Inc.; 2012.

### Literatuur

1. Wouters E. Wat is toetsende statistiek? *PodoSophia*. 2017;25(2):72–5.
2. IBM. *SPSS statistics for Windows*. Version 21. Armonk: IBM; 2016.
3. Aarts S, Zaalen Y van, Wouters E. Evidence-based practice in de dagelijkse praktijk. *PodoSophia*. 2015;3:30–1.
4. Aarts S, Akker M van den, Winkens B. The importance of effect sizes. *Eur J Gen Pract*. 2014;20(1):61–4.

**Sil Aarts**, docent/onderzoeker

**E. Wouters**, lector health innovations and technology

Hier staat een advertentie.



Houten 2017