

## 4.2 Verschil kwalitatieve variabelen

### Inleiding

Kun je zonder meer zeggen dat mannen langer zijn dan vrouwen? Dat er meer meisjes wiskunde A kiezen dan jongens? Dat Belgen meer vreemde talen spreken dan Nederlanders? Bij de verschillende soorten statistische variabelen horen verschillende meetniveaus, die ken je al. Als je verschillen tussen statistische variabelen in kaart wil brengen, hangt de manier waarop je dit kunt doen af van het meetniveau van deze variabelen.



Figuur 1

### Je leert in dit onderwerp

- het verschil tussen twee kwalitatieve variabelen te beschrijven en te interpreteren;
- te werken met een formulekaart om deze verschillen te beschrijven.

### Voorkennis

- soorten statistische variabelen herkennen;
- de begrippen onderzoek, steekproef, populatie en representatief, simulatie;
- meetniveaus onderscheiden bij antwoordmogelijkheden op vragen.

### Verkennen

#### Opgave V1

Langzamerhand kom je in de buurt van echt statistisch onderzoek.

Probeer te beschrijven hoe je zou onderzoeken of meisjes naar verhouding structureel vaker wiskunde A kiezen en jongens naar verhouding wiskunde B.

#### Uitleg 1

Er is een sportdag op school. De leerlingen uit klas 1 en klas 2 mogen kiezen: binnensport of buitensport. De sportkeuze is een nominale kwalitatieve variabele. Van de leerlingen uit klas 1 kiest 21% binnensport. Van de leerlingen uit klas 2 kiest 43% binnensport.

Hoe verschillend kiezen de leerlingen uit klas 1 en klas 2?

Het verschil tussen deze percentages is  $43 - 21 = 22\%$ .

Om antwoord te geven op de vraag gebruik je vuistregels zoals deze:

- Als het verschil groter of gelijk is aan 40% is de conclusie:  
Het verschil is groot.
- Als het verschil tussen 20% en 40% is, is de conclusie:  
Het verschil is middelmatig.
- Als het verschil kleiner of gelijk is aan 20% is de conclusie:  
Het verschil is gering.

Volgens deze vuistregels is de conclusie:

Het verschil tussen klas 1 en klas 2 is 'middelmatig' als je kijkt naar de (percentages) sportkeuze.

Leer dergelijke vuistregels niet uit het hoofd, je krijgt ze gegeven als je ze moet gebruiken.

### Opgave 1

Gebruik de gegevens uit **Uitleg 1**. In klas 3 kiest 48% voor binnensport.

- a Als je op de percentages binnensportkiezers let, wat kun je dan zeggen over het verschil tussen klas 2 en klas 3?
- b Als je op de percentages binnensportkiezers let, wat kun je dan zeggen over het verschil tussen klas 1 en klas 3?

### Opgave 2

Om te onderzoeken of vitamine C helpt tegen verkoudheid heeft een onderzoeker 139 personen vitamine C toegediend en 140 personen een placebo (fopmiddel). De personen in kwestie wisten niet of ze vitamine C of de placebo toegediend kregen.

De onderzoeker telde het aantal personen dat verkouden werd en maakte deze tabel.

	verkouden	niet verkouden
vitamine C	19	122
placebo	29	109
totaal	48	231

Tabel 1

- a Van welke soort is de statistische variabele?
- b Hoeveel procent van de verkouden personen slikte vitamine C?
- c Hoeveel procent van de personen die niet verkouden zijn, slikte vitamine C?
- d Wat is het verschil in percentage verkouden personen van de vitamine C-slikkers en de placebo-slikkers?
- e Is het verschil tussen wel of niet verkouden worden bij wel of niet vitamine C slikken gering, middelmatig of groot?

### Uitleg 2

Regelmatig heb je aantallen eerst omgerekend naar percentages. Er is ook een manier om direct uit aantallen conclusies te trekken.

Om te onderzoeken of vitamine C helpt tegen verkoudheid heeft een onderzoeker 139 personen vitamine C toegediend en 140 personen een placebo (fopmiddel). De personen in kwestie wisten niet of ze vitamine C of de placebo toegediend kregen.

	verkouden	niet verkouden
vitamine C	19 ( <i>a</i> )	122 ( <i>b</i> )
placebo	29 ( <i>c</i> )	109 ( <i>d</i> )

Tabel 2

De onderzoeker telde het aantal personen dat verkouden werd en maakte deze kruistabel.

Met deze aantallen kun je de groepen vergelijken met behulp van het getal *phi* ofwel  $\varphi$  (spreek uit: 'fi'):

- Noem de waarden in de vier cellen van de kruistabel *a*, *b*, *c* en *d* op de manier die je in de tabel ziet.
- Vul deze waarden in de formule voor  $\varphi$  in:

$$\varphi = \frac{ad-bc}{\sqrt{(a+b) \cdot (a+c) \cdot (b+d) \cdot (c+d)}}$$

- Je vindt:  $\varphi \approx -0,100$

Conclusies trek je met de volgende vuistregels:

- Als  $\varphi < -0,4$  of  $\varphi > 0,4$  is de conclusie: het verschil is groot.
- Als  $-0,4 \leq \varphi < 0,2$  of  $0,2 < \varphi \leq 0,4$  is de conclusie: het verschil is middelmatig.
- Als  $-0,2 \leq \varphi \leq 0,2$ , is de conclusie: het verschil is gering.

De vuistregels staan ook op de **Formulekaart**.

Op grond van deze vuistregels trek je een conclusie.

### Opgave 3

Gebruik de gegevens uit **Uitleg 2**.

- Bereken  $\varphi$  nu zelf.
- Welke conclusie trek je volgens de vuistregels op de formulekaart?

### Opgave 4

Gebruik de gegevens uit **Uitleg 2**.

- Bereken  $\varphi$  bij de volgende tabel.

aantal	jongens	meisje
houdt van Harry Potter	30	13
houdt niet van Harry Potter	12	40

Tabel 3

- Wat kun je zeggen over de grootte van het verschil tussen jongens en meisjes, als je kijkt of ze houden van Harry Potter?

### Uitleg 3

De leerlingen in 4 havo zijn naar een toneelvoorstelling geweest. Daarna werd hun gevraagd of ze de voorstelling boeiend vonden, hun *voorstellingsbeleving*.

In de linker tabel staan de resultaten uitgesplitst naar profiel.

Om goed te kunnen vergelijken bij groepen die niet even groot zijn, moeten eerst de absolute aantallen worden omgezet naar percentages. In de rechter tabel staan onder  $p$  de percentages en onder  $cp$  de cumulatieve percentages.

	CM	EM	NG	NT
1 = niet boeiend	5	8	6	17
2 = gaat wel	12	12	18	13
3 = boeiend	9	18	15	8
4 = erg boeiend	9	10	6	2

	EM			NG			V <sub>cp</sub>
	aantal	p	cp	aantal	p	cp	
1 = niet boeiend	8	16,7	16,7	6	13,3	13,3	3,3
2 = gaat wel	12	25,0	41,7	18	40,0	53,3	11,7
3 = boeiend	18	37,5	79,2	15	33,3	86,7	7,5
4 = erg boeiend	10	20,8	100,0	6	13,3	100,0	0,0
	48			45			

Figuur 2

De percentages onder  $p$  van de EM-leerlingen en de NG-leerlingen verschillen: procentueel hebben veel meer NG-leerlingen 'gaat wel' geantwoord dan dat dit bij de EM-leerlingen het geval was. En procentueel hebben veel minder NG-leerlingen 'erg boeiend' geantwoord dan dat dit bij de EM-leerlingen het geval was. Toch hebben procentueel ook minder NG-leerlingen 'niet-boeiend' geantwoord. Een conclusie trekken is nog niet eenvoudig.

Daarom is voor deze situatie afgesproken dat het maximale verschil van de cumulatieve percentages  $\max V_{cp}$  wordt gebruikt als maat voor het verschil tussen twee groepen.

De vuistregels zijn nu:

- Als  $\max V_{cp} \leq 20\%$  is het verschil 'gering'.
- Als  $20\% < \max V_{cp} \leq 40\%$  is het verschil 'middelmatig'.
- Als  $\max V_{cp} > 40\%$  is het verschil 'groot'.

Je vindt deze afspraken op de **Formulekaart**.

### Opgave 5

Gebruik de gegevens uit **Uitleg 3**.

- Waarom is de variabele *voorstellingsbeleving* een ordinale kwalitatieve variabele?
- Vergelijk nu de cumulatieve percentages. Wat betekent het dat de cumulatieve percentages van de EM-leerlingen zowel bij 'gaat wel' als bij 'boeiend' lager zijn dan die van de NG-leerlingen?
- In de kolom  $V_{cp}$  zie je het verschil van de cumulatieve percentages. Laat zien dat daarbij niet wordt gelet op welk cumulatieve percentage groter is.
- Waarom is de  $V_{cp}$  bij de hoogste waarde van de variabele altijd 0?
- Hoe groot is het maximale verschil  $\max V_{cp}$  van de cumulatieve percentages?
- Schrijf de conclusie op.

### Opgave 6

Vergelijk de *voorstellingsbeleving* van de NG-leerlingen en de NT-leerlingen.

- Bepaal bij elke waarde van deze variabele het verschil tussen de cumulatieve relatieve percentages. Neem het absolute verschil, laat dus mintekens weg.
- Bepaal nu  $\max V_{cp}$ .
- Als je dezelfde criteria hanteert als in het voorbeeld, is er dan een gering, middelmatig of groot verschil tussen NG- en NT-leerlingen?

## Theorie en voorbeelden

### Om te onthouden

In statistisch onderzoek worden vaak twee groepen onderzocht en met elkaar vergeleken.

Als er vragen worden gesteld met twee mogelijke antwoorden (kwalitatieve, nominale variabele), kunnen percentages of aantallen voor deze vergelijking worden gebruikt.

- **Percentages vergelijken.**

Gebruik de vuistregels:

- Als het verschil in percentage groter of gelijk is aan 40% is de conclusie: Het verschil is groot.
- Als het verschil in percentage tussen 20% en 40% is, is de conclusie: Het verschil is middelmatig.
- Als het verschil in percentage kleiner of gelijk is aan 20% is de conclusie: Het verschil is gering.

- **Aantallen vergelijken:**

Maak een  $2 \times 2$  **kruistabel**.

Bereken  $\phi = \varphi$ :

$$\phi = \frac{ad-bc}{\sqrt{(a+b) \cdot (a+c) \cdot (b+d) \cdot (c+d)}}$$

Conclusies trek je met behulp van de vuistregels op de **Formulekaart**.

Als er vragen worden gesteld met meerdere mogelijke antwoorden die kunnen worden geordend (kwalitatieve, ordinale variabele), kunnen **cumulatieve percentages** worden gebruikt. Dan wordt het maximale verschil van de cumulatieve percentages,  $\max V_{cp}$ , berekend.

De bijbehorende vuistregels staan op de **Formulekaart**.

### Voorbeeld 1

Bekijk in de tabel de keuze voor wiskunde A voor jongens en meisjes in 4 havo.

	jongens	meisjes
wiskunde A	27	32
geen wiskunde A	33	25

**Tabel 4**

Is er een gering, een middelmatig of groot verschil tussen jongens en meisjes op deze school voor wat betreft de keuze voor wiskunde A? Gebruik percentages en de vuistregels:

- Als het verschil in percentage groter of gelijk is aan 40% is de conclusie:  
Het verschil is groot.
- Als het verschil in percentage tussen 20% en 40% is, is de conclusie:  
Het verschil is middelmatig.
- Als het verschil in percentage kleiner of gelijk is aan 20% is de conclusie:  
Het verschil is gering.

Antwoord

Er is sprake van een nominale variabele (variabele is aantal jongens en aantal meisjes dat wel of geen wiskunde A kiest).

Je kijkt nu naar het verschil in procenten tussen jongens en meisjes.

	percentage jongens	percentage meisjes
wiskunde A	45%	56,1%

**Tabel 5**

Reken de aantallen om naar percentages. Het totaal aantal jongens is 60, het totaal aantal meisjes is 57.

Het verschil in procenten tussen jongens en meisjes die wiskunde A kiezen, is  $56,1 - 45 = 11,1\%$ .

Dit verschil in procenten is  $\leq 20\%$ . Je spreekt dus van een gering verschil.

### Opgave 7

Gebruik de gegevens uit **Voorbeeld 1**.

- Bereken zelf de percentages die in de tabel staan.
- Bereken het procentuele verschil tussen jongens en meisjes die geen wiskunde A kiezen. Wat is dan je conclusie?

### Opgave 8

Bekijk de tabel met de keuze voor Frans voor jongens en meisjes in 4 havo.

	jongens	meisjes
Frans	22	40
geen Frans	38	17

**Tabel 6**

Is er een gering, een middelmatig of groot verschil tussen jongens en meisjes op deze school voor wat betreft de keuze voor Frans? Gebruik de percentages en de vuistregels uit het voorbeeld.

### Voorbeeld 2

Er wordt veel onderzoek gedaan naar bijwerkingen van medicijnen. Onderzoek onder mannen en vrouwen naar de bijwerkingen gaf de volgende resultaten:

	bijwerking	geen bijwerking
mannen	86	14
vrouwen	61	39

Tabel 7

Bereken  $\phi$  en trek een conclusie met behulp van de vuistregels op de [Formulekaart](#).

Antwoord

$\phi$  berekenen geeft  $\phi \approx 0,28$ .

Volgens de vuistregels voor vergelijken van groepen is dus sprake van een middelmatig verschil.

### Opgave 9

Gebruik de gegevens uit [Voorbeeld 2](#).

Bereken zelf  $\phi$  en controleer de conclusie.

### Opgave 10

Dit is het resultaat van een statistisch onderzoek.

	slagen	zakken
gemotiveerd	70	10
ongemotiveerd	10	30

Tabel 8

- Bereken  $\phi$ .
- Wat is je conclusie over de mate van verschil tussen gemotiveerde en ongemotiveerde leerlingen?
- Trek ook een conclusie door percentages te gebruiken.

### Voorbeeld 3

Hotelgasten is gevraagd het dinerbuffet te beoordelen. In de tabel staan de resultaten uitgesplitst naar werelddeel van herkomst.

	Europa	Afrika	Azië	Noord-Amerika	Zuid-Amerika	Oceanië
1 = zeer goed	12	4	6	3	4	0
2 = goed	15	6	22	6	8	4
3 = matig	10	2	15	4	2	2
4 = slecht	3	0	4	4	0	2

Tabel 9

Op grond van deze tabel lijkt het alsof de Europeanen het dinerbuffet hoger hebben gewaardeerd dan de Aziaten, maar die groepen zijn niet even groot.

Is er een gering, een middelmatig of groot verschil tussen de Europeanen en de Aziaten en hun oordeel over het dinerbuffet? Gebruik percentages en de vuistregels op de [Formulekaart](#).

Antwoord

Om goed te kunnen vergelijken bij groepen die niet even groot zijn, moeten eerst de absolute aantallen worden omgezet naar percentages.

Bekijk in de tabel onder  $p$  de percentages, onder  $cp$  de cumulatieve percentages en onder  $V_{cp}$  het verschil van  $cp$ .

	Europa			Azië			$V_{cp}$
	aantal	$p$	$cp$	aantal	$p$	$cp$	
1 = zeer goed	12	30	30	6	12,8	12,8	17,2
2 = goed	15	37,5	67,5	22	46,8	59,6	7,9
3 = matig	10	25	92,5	15	31,9	91,5	1
4 = slecht	3	7,5	100	4	8,5	100	0
	40			47			

Tabel 10

Er geldt  $\max V_{cp} \leq 20\%$ , dus het verschil is gering.

### Opgave 11

Gebruik de gegevens uit [Voorbeeld 3](#).

Is er een gering, een middelmatig of groot verschil tussen de Afrikanen en de Oceaniërs en hun oordeel over het dinerbuffet?

## Verwerken

### Opgave 12

Het verschil bij een onderzoek onder 200 patiënten naar het gebruik van twee medicijnen is 14%. Is er een gering, een middelmatig of groot verschil?

- A. gering verschil
- B. middelmatig verschil
- C. groot verschil

### Opgave 13

Er is veel onderzoek gedaan naar de effecten van roken. Gekeken is bijvoorbeeld naar aandoeningen aan hart- en bloedvaten bij rokers en niet-rokers met het volgende resultaat:

	aandoening	geen aandoening
roker	86	14
niet-rokers	61	39

Tabel 11

- a Gaat het hier om kwalitatieve of kwantitatieve variabelen?
- b Gaat het om nominale of ordinale variabelen?
- c Bereken het verschil in procenten tussen rokers en niet-rokers als er wordt gelet op wel of geen aandoeningen aan hart- en bloedvaten.
- d Is het verschil tussen roken en niet roken gering, middelmatig of groot volgens de vuistregels?
- e Welke kritische vragen heb je bij dit onderzoek?
- f Zegt dit onderzoek iets over de oorzaak van de aandoeningen aan hart- en bloedvaten?

### Opgave 14

Dit is het resultaat van een onderzoek door de Rijksoverheid naar de uitstroom van studenten uit het voltijd HBO per studierichting. De getallen zijn percentages.

studierichting	werk en/of uitkering	geen werk en/of uitkering
Onderwijs	87	13
Landbouw en natuur	76	24
Techniek	77	23
Gezondheidszorg	86	15
Economie	77	23
Gedrag en maatschappij	83	17
Taal en Cultuur	62	38

Tabel 12

- Welke studierichting heeft het laagste percentage uitstroom van studenten met werk en/of uitkering?
- Voor welke studierichtingen geldt: er is een middelmatig verschil met uitstroom van studenten van de richting Taal en Cultuur, als gelet wordt op het percentage werk en/of uitkering?

### Opgave 15

Er is veel onderzoek naar bijwerkingen van medicijnen. Onderzoek onder gebruikers van aspirines gaf de volgende  $2 \times 2$  tabel.

	hartfalen	geen hartfalen
aspirine gebruikers	139	10898
geen aspirine gebruikers	239	10795

Tabel 13

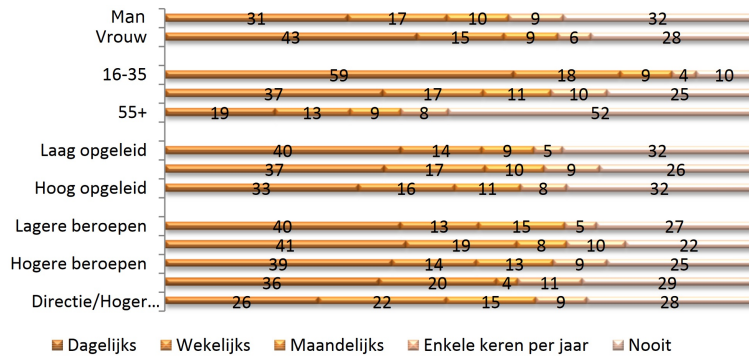
Je wilt weten welk verschil er is tussen gebruikers van aspirine en niet-gebruikers. De grootheid  $\varphi$  is een maat voor dit verschil.

- Bereken  $\varphi$ .
- Welke conclusie trek je over de samenhang tussen gebruik van aspirine en hartfalen? Gebruik de **Formulekaart**.
- Reken de gegevens in de tabel om naar procenten. Welke conclusie kun je nu trekken?



### Opgave 16

In een rapport van Universiteit Twente staat het volgende overzicht over het bezoek aan sociale netwerksites.



Figuur 3

Bekijk of er een verschil is tussen mannen en vrouwen. Gebruik de cijfers uit de figuur.

- a Vul de percentages voor mannen en vrouwen in de tabel in. *cp* is de afkorting voor cumulatief percentage.

	mannen	cp mannen	vrouwen	cp vrouwen
nooit				
enkele keren per jaar				
maandelijks				
wekelijks				
dagelijks				

Tabel 14

- b Gebruik de vuistregels op de **Formulekaart**. Is er een gering, middelmatig of groot verschil tussen mannen en vrouwen? Licht je antwoord toe.

### Opgave 17

Een marketingmanager heeft de volgende resultaten van een onderzoek naar de werking van een medicijn ontvangen. Welke conclusie zal hij eruit trekken?

aantal	positieve reactie	geen reactie
medicijn niet gebruikt	161	739
medicijn wel gebruikt	1796	1306

Tabel 15

## Toepassen

### Opgave 18: Intelligentie van hondenrassen

Voor een onderzoek naar de intelligentie van hondenrassen is aan hondenbezitters gevraagd hoelang het duurde voor de hond naar zijn naam luisterde vanaf het moment dat ze hem als pup kregen. De resultaten van labradors en beagles staan in de tabel.

aantal	labrador	beagle
luistert naar naam binnen een week	55	39
luistert naar naam na meer dan een week	32	27

Tabel 16

- a Op welke manier zou je hier een conclusie uit kunnen trekken?
- b Trek aan de hand van de **bijbehorende vuistregels** en de tabel een conclusie over het verschil in intelligentie tussen een labrador en een beagle.

## Testen

### Opgave 19

Dit is het resultaat van een statistisch onderzoek.

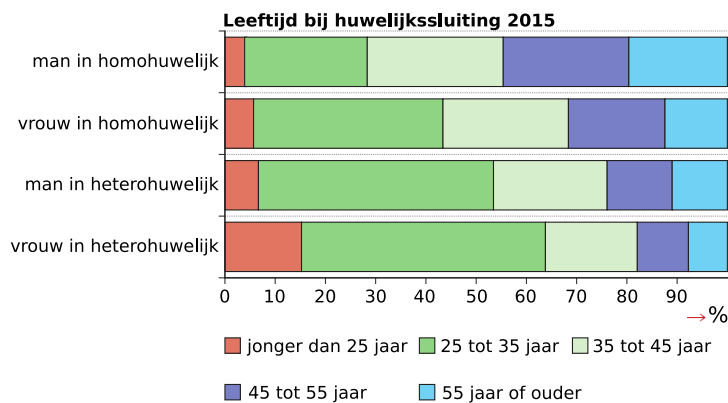
	huid verbrandt snel	huid verbrandt niet snel
rood haar	55	8
bruin haar	87	102

Tabel 17

- a Trek op grond van deze gegevens een conclusie over het verschil tussen roodharigen en bruinharigen met betrekking tot het verbranden van de huid door de zon. Gebruik de vuistregels van de **Formulekaart**.
- b Trek op grond van percentages een conclusie over het verschil tussen roodharigen en bruinharigen met betrekking tot het verbranden van de huid door de zon. Gebruik de vuistregels:
  - Als het verschil in percentage groter of gelijk is aan 40% is de conclusie: Het verschil is groot.
  - Als het verschil in percentage tussen 20% en 40% is, is de conclusie: Het verschil is middelmatig.
  - Als het verschil in percentage kleiner of gelijk is aan 20% is de conclusie: Het verschil is gering.

### Opgave 20

Bekijk de leeftijd bij huwelijkssluiting in 2015, gegroepeerd naar soort huwelijk.



Figuur 4

Maak een cumulatieve tabel en trek op grond van deze tabel de conclusies van de soort “Er is een middelmatig verschil tussen het aantal ...” voor dezelfde leeftijdscategorieën.



© 2021

Deze paragraaf is een onderdeel van het Math4All wiskundemateriaal.

Math4All stelt het op prijs als onvolkomenheden in het materiaal worden gemeld en ideeën voor verbeteringen in de content of dienstverlening kenbaar worden gemaakt.

Email: [f.spijkers@math4all.nl](mailto:f.spijkers@math4all.nl)

Met de Math4All maatwerkdienst kunnen complete readers worden samengesteld en toetsen worden gegenereerd. Docenten kunnen bij [a.f.otten@xs4all.nl](mailto:a.f.otten@xs4all.nl) een gratis inlog voor de maatwerkdienst aanvragen.

---

