

Dialecten vergelijken

Taalkundigen doen veel onderzoek naar de dialecten in Nederland en Vlaanderen.

Onderzoeker M. Spruit onderzocht in 2008 in hoeverre dialecten op elkaar lijken. De mate waarin twee dialecten verschillen, wilde hij uitdrukken in een getal. Daarom vergeleek hij steeds twee dialecten op een aantal kenmerken en telde hij vervolgens de verschillen. Elk verschil tussen deze twee dialecten leverde een punt op. Het totale aantal punten is de **Hammingafstand** tussen deze twee dialecten.

Bijvoorbeeld: in Lunteren kan men zeggen: “Jan herinnert **zich** dat verhaal wel”, maar ook: “Jan herinnert **z’n eigen** dat verhaal wel”. In Veldhoven zegt men altijd: “Jan herinnert **zich** dat verhaal wel”. In geen van beide dialecten gebruikt men hier “**hem**” of “**zichzelf**” of “**hemzelf**”, iets dat in andere dialecten wel voorkomt.

Het vergelijken van deze vijf kenmerken levert dus in totaal 1 punt op voor de Hammingafstand. Dat is in tabel 1 weergegeven.

tabel 1

	Lunteren	Veldhoven	punten (voor Hammingafstand)
zich	+	+	0
hem	–	–	0
z’n eigen	+	–	1
zichzelf	–	–	0
hemzelf	–	–	0

Stel men vergelijkt dialect X met het dialect van Lunteren. En stel dat vergelijken van de vijf kenmerken uit tabel 1 in totaal 3 punten oplevert voor de Hammingafstand. In dialect X wordt ook “zich” gebruikt.

- 4p 10 Schrijf alle mogelijkheden voor deze vijf kenmerken voor dialect X op. Gebruik hiervoor de tabel op de uitwerkbijlage.

De onderzoeker vergeleek niet vijf, maar 507 kenmerken. Het resultaat is een tabel waarin per tweetal dialecten de Hammingafstand te zien is. In tabel 2 zie je hier een gedeelte van.

Het getal 66 in deze tabel voor het tweetal Lunteren-Bellingwolde (of Bellingwolde-Lunteren) betekent dat bij deze twee dialecten 66 van de 507 kenmerken verschillen: de Hammingafstand is 66.

tabel 2

Dialect	Lunteren	Bellingwolde	Hollum	Doel	Sint-Truiden	Veldhoven
Lunteren	–	66	52	122	77	47
Bellingwolde	66	–	56	134	81	51
Hollum	52	56	–	116	63	59
Doel	122	134	116	–	115	111
Sint-Truiden	77	81	63	115	–	72
Veldhoven	47	51	59	111	72	–

In tabel 2 heeft de onderzoeker dus 15 Hammingafstanden berekend. In totaal stonden er echter geen 6 dialecten, maar 267 dialecten in de tabel. Bij elk tweetal heeft de onderzoeker de Hammingafstand berekend.

3p 11

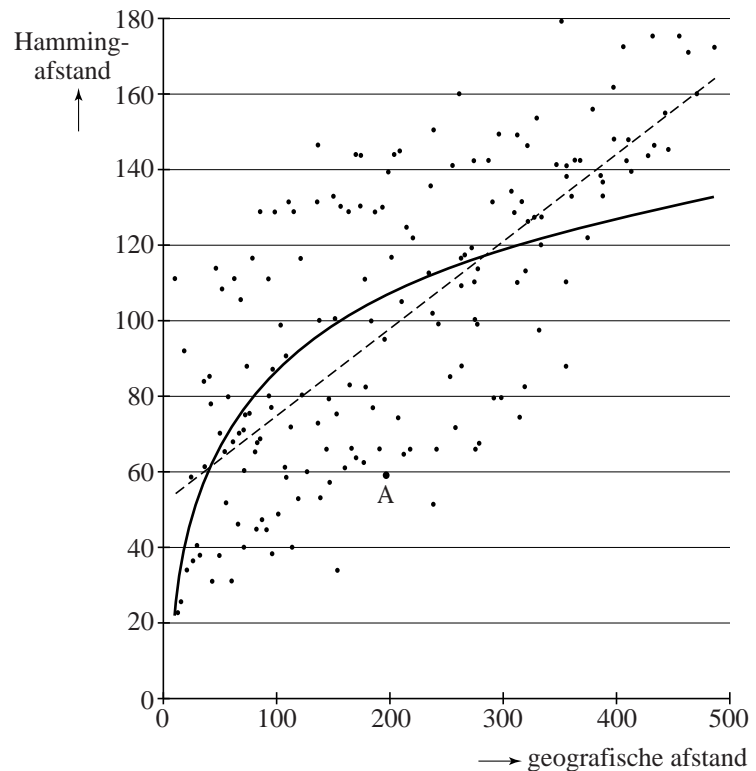
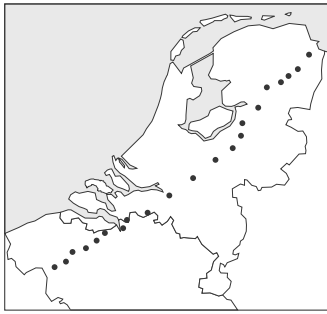
Bereken hoeveel Hammingafstanden de onderzoeker in totaal heeft berekend.

De onderzoeker zocht naar een verband tussen de geografische afstand en de Hammingafstand van dialecten. In het kaartje in de figuur zie je een aantal dialecten met stippen aangegeven. In het assenstelsel is voor elk tweetal dialecten de Hammingafstand (in punten) uitgezet tegen de geografische afstand (in km).

In het assenstelsel kun je zien dat bij punt A de afstand tussen twee plaatsen gelijk is aan 200 km en de Hammingafstand ongeveer gelijk is aan 58.

De onderzoeker heeft op twee manieren geprobeerd het verband tussen de geografische afstand en de Hammingafstand met een wiskundig verband te benaderen. Beide manieren, een lineair verband en een logaritmisch verband, zijn weergegeven in het assenstelsel.

figuur



De onderzoeker heeft in het assenstelsel dus ook een grafiek voor een logaritmisch verband getekend. De formule voor dit logaritmische verband is:

$$H = -45,88 + 66,44 \log(x)$$

Hierin is H de Hammingafstand en x de geografische afstand in km.

- Op de getekende rechte lijn liggen de punten $(10, 55)$ en $(400, 145)$.
- 5p 12 Stel met behulp van deze twee punten een formule op voor het lineaire verband in het assenstelsel en bereken met behulp van de formules bij welke geografische afstanden de Hammingafstanden volgens het lineaire en het logaritmische verband gelijk zijn. Rond je antwoord af op gehele kilometers.

De grafiek van het logaritmische verband in het assenstelsel gaat bijvoorbeeld door de punten $(50, 67)$, $(100, 87)$, $(200, 107)$ en $(400, 127)$. Hieraan kun je zien dat volgens het logaritmisch verband bij een verdubbeling van de geografische afstand de Hammingafstand steeds met 20 toeneemt. Met behulp van de formule $H = -45,88 + 66,44 \log(x)$ kun je aantonen dat dit altijd zo is.

- 3p 13 Toon met behulp van de rekenregels van de logaritmen aan dat $-45,88 + 66,44 \log(2x)$ ongeveer gelijk is aan $-45,88 + 66,44 \log(x) + 20$.