

Uitslagen voorspellen

In de tijd voor Tweede Kamerverkiezingen worden allerlei onderzoeken gedaan naar kiezersgedrag.

Media publiceren vrijwel elke dag voorspellingen gebaseerd op onderzoek. Zo ging het ook voor de verkiezingen in juni 2010. Op 3 juni publiceerde de krant Tubantia de persoonlijke voorspellingen van elf lijsttrekkers over de te verwachten zetelverdeling voor de elf partijen. Zie tabel 1. Deze tabel staat vergroot op de uitwerkbijlage.

tabel 1

	PVV	SP	GroenLinks	Trots op NL	PvdA	CDA	D66	VVD	P.v.d.Dieren	SGP	ChristenUnie
	G. Wilders	E. Roemer	F. Halsema	R. Verdonk	J. Cohen	J.P. Balkenende	A. Pechtold	M. Rutte	M. Thieme	K. v.d. Staaij	A. Rouvoet
CDA	29	27	29	28	27	34	26	29	24	28	28
PvdA	29	30	33	26	35	28	28	29	29	27	32
SP	10	18	11	14	9	17	13	11	21	12	10
VVD	29	29	31	27	34	32	30	34	31	34	32
PVV	25	15	11	14	16	12	15	17	12	17	14
GroenLinks	8	10	13	9	9	9	12	10	9	10	10
ChristenUnie	8	7	6	6	7	5	6	6	6	7	10
D66	8	10	12	10	9	10	15	10	12	10	10
P.v.d.Dieren	1	2	2	3	2	1	3	2	4	2	2
SGP	2	2	2	3	2	2	2	2	2	3	2
Trots op NL	1	0	0	10	0	0	0	0	0	0	0
Totaal	150	150	150	150	150	150	150	150	150	150	150

In tabel 1 valt onder andere op dat de voorspellingen van Wilders en Thieme behoorlijk van elkaar verschillen, terwijl de voorspellingen van Rutte en Van der Staaij tamelijk dicht bij elkaar liggen.

Om voorspellingen met elkaar te kunnen vergelijken, gebruiken we het begrip **afstand**. Om de afstand tussen twee voorspellingen te berekenen, tellen we alle verschillen tussen de voorspelde zetelaantallen bij elkaar op. Zo is de afstand tussen de voorspellingen van Roemer (lijsttrekker SP) en Halsema (lijsttrekker GroenLinks) 24, want de som van de positieve verschillen tussen hun voorspellingen is:

$$(29 - 27) + (33 - 30) + (18 - 11) + (31 - 29) + (15 - 11) + (13 - 10) + (7 - 6) + (12 - 10) + (2 - 2) + (2 - 2) + (0 - 0) = 24$$

- 3p 1 Onderzoek of de afstand tussen de voorspellingen van Wilders en Thieme meer dan twee maal zo groot is als de afstand tussen de voorspellingen van Roemer en Halsema.

Na afloop van de verkiezingen kun je de voorspellingen van ieder van de lijsttrekkers met de werkelijke uitslag vergelijken. Dat doen we hier op twee verschillende manieren. Bij de eerste methode berekenen we de **afstand tussen de voorspelling en de werkelijke uitslag**. Die werkelijke uitslag van de verkiezingen op 9 juni 2010 staat in tabel 2.

tabel 2

partij	CDA	PvdA	SP	VVD	PVV	GL	CU	D66	PvdD	SGP	TON
werkelijk aantal zetels	21	30	15	31	24	10	5	10	2	2	0

De voorspelling van Roemer blijkt de kleinste afstand, namelijk 22, tot de werkelijke uitslag op te leveren.

De afstand tussen de voorspelling van Wilders en de werkelijke uitslag blijkt exact gelijk te zijn aan de afstand tussen de voorspelling van Van der Staaij en de werkelijke uitslag.

2p **2** Bereken deze afstand.

Een andere methode om voorspellingen te vergelijken met de werkelijke uitslag is om te kijken naar het totaal **aantal juist voorspelde zetels**. Als een partij bijvoorbeeld 8 zetels haalt terwijl er 5 voorspeld zijn, dan krijgt de voorspeller daar 5 punten voor. En als er 8 zetels behaald worden terwijl er 10 voorspeld zijn, dan krijgt de voorspeller 8 punten.

Op deze manier is het aantal juist voorspelde zetels van Roemer:

$$21 + 30 + 15 + 29 + 15 + 10 + 5 + 10 + 2 + 2 = 139$$

Als je het aantal juist voorspelde zetels van Wilders vergelijkt met het aantal juist voorspelde zetels van Van der Staaij, blijkt ook nu weer dat deze aantallen aan elkaar gelijk zijn.

2p **3** Bereken het aantal juist voorspelde zetels bij deze twee lijsttrekkers.

Dat deze aantallen aan elkaar gelijk zijn, is niet toevallig als je kijkt naar het aantal juist voorspelde zetels en de afstand tussen de voorspelling en de werkelijke uitslag. Tussen deze afstand (de eerste methode) en het aantal juist voorspelde zetels (de tweede methode) bestaat een verband. Bij de afstand let je op de verschillen (altijd positief) en bij de tweede methode tel je het aantal goed voorspelde zetels. Het verband heeft de volgende vorm:

$$\text{aantal juist voorspelde zetels} = a \cdot \text{afstand} + b$$

4p **4** Bereken de waarden van a en b in bovenstaand verband.