

Bionieuws 10

Miquel Bulnes

Wetenschappers zouden zich harder moeten verzetten tegen flauwekul die verkocht wordt als wetenschap: misleidende reclames, slecht gedocumenteerde wetenschapskaternen, speculaties in overheidsrapporten. Van alle misbruikte takken van wetenschap is statistiek toch wel de voornaamste en hier wil ik even gaan zeuren over de ergste en hardnekkigste vorm van misbruik, de verkiezingspeilingen. Het zou goed zijn als af en toe een echte statisticus (en niet een hobbyist zoals ik) zich hier harder tegen zou uitspreken.

Peilingen, voor alle duidelijkheid, zijn afgeleide parameters. Om dat uit te leggen een voorbeeld:

Stel je voor, je wilt uitzoeken hoe lang Nederlandse vijftienjarigen zijn. Wat is dan betrouwbaarder: (1) driehonderd vijftienjarigen meten, of (2) driehonderd veertienjarigen bellen en vragen hoe lang ze over een jaar zullen zijn?

Peilingen zijn vergelijkbaar met optie 2, waardes waarvan je áánneemt dat ze voorspellend zijn voor wat je echt wilt weten, maar die – per definitie – een onzekerheidsmarge kennen.

Af en toe geven de onderzoeksbureaus zelf die onzekerheidsmarges op; als je hen moet geloven zijn die minimaal. Je leest uitspraken in de trant van: *“Er is een steekproef gehouden onder vierduizend kiezers en daarmee is de maximale afwijking één procent.”* Maurice de Hond en consorten doen zich hier dommer voor dan ze zijn. Ze vergelijken de peiling namelijk niet met uitslagen maar met grotere peilingen.

Dank je de koekoek.

Dat is net zo’n onzin als roepen: *“Als er vandaag verkiezingen worden gehouden, haalt het CDA 38 zetels.”* Er worden “vandaag” helemaal geen verkiezingen gehouden, dus er valt achteraf niets te meten. Tijdens de zomerstop ga je ook niet voorspellen: *“Als Ajax-PSV vandaag wordt gespeeld, wint PSV met drie – nul.”*

Een voorspelling moet je doen voor de realiteit en moet je spiegelen aan deze realiteit— in dit geval dus een verkiezingsuitslag. Wetenschap begint bij het stellen van de juiste vraag, en de enige legitieme vraag is: *“Als vandaag het kabinet valt en er worden*

verkiezingen uitgeschreven, wat is dan de voorspellende waarde van de peilingen van vandaag?”

Hieraan kunnen we gaan rekenen. We hebben namelijk het geluk dat de afgelopen drie verkiezingen voortkwamen uit kabinetten die vielen, en we beschikken over peilingen óp of net ná die datum. Voor het vaststellen van waarde van een voorspelling is dat alles wat je nodig hebt: de testuitslag en de werkelijke waarde (“gouden standaard”).

Met een betrekkelijk eenvoudige formule¹ is op basis van de verschillen in peiling en uitkomst de “standaardafwijking” uit te rekenen: de maat voor de onzekerheid van een steekproef of voorspelling. Vul de getallen in voor de vier grootste partijen en je komt uit op een standaardafwijking van 32%. Dat betekent dat als er op een dag dat het kabinet viel door een onderzoeksbureau een voorspelling werd gedaan, er een kans was van 68,2%² dat de uitslag van de verkiezingen tussen de 32% onder en 32% boven deze voorspelling uitkwam.

Dus, 38 zetels voor het CDA in de peilingen betekent niet:

“Als er vandaag verkiezingen waren zou het CDA uitkomen op 38 zetels.”

Maar:

“Als er vandaag verkiezingen worden uitgeschreven, is er een kans van 68.2% dat het CDA tussen de 26 en de 50 zetels haalt.”

Dát is wetenschap. Of het nieuws is, is een tweede.

	Peiling 17 april 2002 (NIPO)	Uitslag 15 mei 2002	Peiling 17 oktober 2003 (De Hond)	Uitslag 22 januari 2003	Peiling 3 juli 2006 (NIPO)	Uitslag 22 november 2006
CDA	29	43	51	44	30	41
PvdA	31	23	28	42	43	33
VVD	25	24	28	28	32	22
SP	7	9	16	9	20	25

¹ $\sqrt{(\sum(x_{\text{voorspeld}} - x_{\text{uitkomst}})^2 / n)}$. Hij ziet er moeilijker uit dan hij is

²68,2% is een statistische constante die hoort bij de standaardafwijking.