

Johan, Florence en Samuel – Leugens over een citaat

In deze tijd van de invoering van het nieuwe zorgstelsel realiseert menigeen zich weer eens hoe lastig het soms kan zijn om met onzekerheden om te gaan en om kansen in te schatten. Zelfs binnen een heel precies gedefinieerd wiskundig model voor onzekerheid en toeval is het vaak moeilijk de juiste weg te vinden. Berucht is het drie dozenprobleem dat de Franse wiskundige Jules Henri Poincaré (1854–1912) in 1896 formuleerde en dat enkele jaren geleden nog verhitte discussies in de NRC opleverde.

Je zou staande kunnen houden dat de meeste mensen een betere intuïtie hebben voor het gedrag van bijvoorbeeld meetkundige objecten dan voor toevalsvariabelen. Zo is het axiomastelsel van Euclides 22 eeuwen eerder ontwikkeld dan de axiomatische opbouw van de kansrekening door Andrei Nikolaevich Kolmogorov (1903–1987) in 1933.

Een bekende raadspensionaris van Holland was een van de eersten die kansrekening en statistiek toepasten. Johan de Witt (1625–1672) verbeterde de bepaling van het rentepercentage bij de toen gebruikelijke methode van staatsfinanciering door middel van lijfrenten op een wiskundige manier. Hij publiceerde zijn verhandeling *Waardijje van Lyf-renten naer Proportie van Los-renten* een jaar voor zijn gewelddadige dood op 20 augustus 1672.

Als een van de eerste toegepast statisticae wordt Florence Nightingale (1820–1910) beschouwd. Als feministe vocht zij voor haar recht wiskunde te studeren. Zij meende dat sociologische verschijnselen statistisch onderzocht moesten worden. Tijdens de Krim-oorlog van 1853-1856 overtuigde Florence door middel van effectieve grafische voorstelling van de verzamelde gegevens over het overlijden van oorlogsgewonden de verantwoordelijke militairen en politici van de maatregelen die genomen moesten worden op hygiënisch gebied in de lazaretten om onnodige doden te voorkomen. Erkenning van haar werk blijkt uit haar benoeming in 1858 tot Fellow of the Royal Statistical Society en in 1874 tot erelid van de American Statistical Association.

In bovenstaande voorbeelden, maar ook in de etymologie van het woord statistiek, zit een verwantschap tussen statistiek en *staat*. Oorspronkelijk is statistiek dan ook een begrip uit de politieke wetenschappen waarbij het gaat om de gegevens van een staat of samenleving. In dit licht is het niet verwonderlijk dat Samuel Langhorne Clemens (1835–1910) onder zijn schrijversnaam Mark Twain een politicus ten tonele voert als hij beweert: ‘The remark attributed to Disraeli would often apply with justice and force: *There are three kinds of lies: lies, damned lies, and statistics*’. Maar Benjamin Disraeli (1804–1881), eerste minister van Groot Brittannië in 1868 en van 1874 tot 1880, heeft zoiets zeer waarschijnlijk nooit gezegd. Wel is er uit zijn tijd een citaat van de econoom en politicus Leonard Henry Courtney (1832-1918): ‘After all, facts are facts, and although we may quote one to another with a chuckle the words of the Wise Statesman, *Lies - damn lies - and statistics*, still there are some easy figures the simplest must understand, and the astutest cannot wriggle out of.’

Je zou kunnen zeggen dat hier niet sprake is van een opklimmende reeks zoals bij Twain, maar van een tegenstelling: je hebt leugens en statistieken. Het is in ieder geval duidelijk dat allen hier met statistics statistieken bedoelen en zeker niet de wetenschap statistiek. Toch word je niet alleen als toegepast statisticus, maar ook als wiskundig statisticus vaak geplaagd met dit ‘citaat’, meestal in situaties waarin onoordeelkundig gebruik van statistische technieken tot duidelijke onzin heeft geleid. Het volledige citaat roept er in feite toe op om de productie van statistieken aan echte statistici over te laten en niet aan charlatans. Mark Twain schrijft immers in zijn posthuum verschenen Autobiography (1925): ‘Figures often beguile me, particularly when I have the arranging of them myself; in which case the remark attributed to Disraeli would often apply with justice and force: *There are three kinds of lies: lies, damned lies, and statistics*.’

Chris Klaassen
chrisk@science.uva.nl