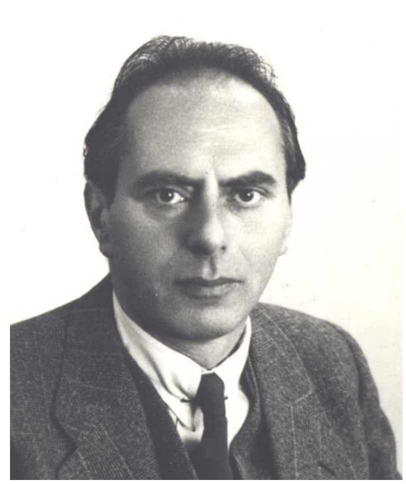
Statistisch Priesterschap

Peter Grünwald

Gebaseerd op een lezing gehouden 12-12 2006, ter ere van het 60-jarig bestaan van het CWI

1. David van Dantzig (1900-1959)



De titel van deze presentatie verwijst naar een artikel van David van Dantzig, een van de oprichters van het CWI. Deze buitengewoon getalenteerde en gedreven man had zich vóór de tweede wereldoorlog vooral met de zuivere wiskunde beziggehouden. Na de oorlog ontwikkelde hij zich tot, in de woorden van J. Hemelrijk, de "Godfather van de kansrekening en statistiek in Nederland". De professionele beoefening van statistiek in Nederland begint met Van Dantzig.

In 1957 publiceerde Van Dantzig het artikel "Statistical Priesthood" (*Statistica Neerlandica 11*,1-16). Het artikel is feitelijk een bespreking van het boek *The Foundations of Statistics* van Leonard Savage. In dit boek presenteert Savage een aantal nieuwe ideeën op het gebied van de statistiek, ideeën die we tegenwoordig "Bayesiaans" zouden noemen. Van Dantzig veegt de vloer aan met het boek. De geschiedenis heeft hem geen gelijk gegeven: Savage's boek wordt tegenwoordig gezien als een van de belangrijkste boeken over de statistiek die ooit verschenen zijn. Met veel aspekten van Van Dantzig's recensie ben ik het dan ook niet eens. Maar Van Dantzig stipt één punt aan waar ik het wel van harte mee eens ben; een punt dat tegenwoordig, 50 jaar later, vaak nog steeds niet begrepen wordt. Dit is het feit dat de uitspraak "deze gebeurtenis heeft *X* procent kans" vaak *betekenisloos* is. In veel alledaagse situaties kan men eigenlijk niet spreken van "kansen", hoewel de meeste mensen (inclusief wiskundigen) dit vaak wel doen. Deze voordracht gaat over dit soort "kansloze situaties," die ik bespreek aan de hand van drie voorbeelden:

1.1 Kansloze Situaties:

-Het 3-Gevangenen Probleem. Een wiskundige puzzel die laat duidelijk laat zien dat een eenduidige "kans" soms niet bestaat (Sectie 2).

-Het 3-Deuren Probleem. Een, veel bekendere, wiskundige puzzel die laat zien dat onze intuitie hierover vaak verkeerd is (Sectie 3).

-Het 1-Gevangene Probleem. Dit is geen wiskundig maar een maatschappelijk probleem dat thans in Nederland speelt. Het laat zien dat onze pogingen om over "kansen" te praten als die er niet zijn, desastreuze gevolgen kunnen hebben! (Sectie 4)

Voordat ik begin, zal ik even uw geheugen opfrissen met betrekking tot kansrekening.

1.2 De Dobbelsteen – Conditionele Kansen

Stel ik gooi met een eerlijke dobbelsteen; ik zie de uitkomst (een getal tussen 1 en 6) maar u ziet de uitkomst niet. Ik vertel u ofwel "*de uitkomst is even*" ofwel "*de uitkomst is oneven*". Stel dat ik u vertel "de uitkomst is even." Wat is volgens u dan de kans dat er "4" is gegooid? U zegt: er zijn nog drie mogelijkheden over. Die hebben allemaal gelijke kans, dus: **de kans op "4" is nu 1/3.**

Dat is een correcte redenering. Eerst was de kans 1/6. U past deze kans aan omdat u nieuwe informatie heeft; dit heet **conditioneren**. De kans is nu 1/3 geworden. We zeggen: "de *conditionele* kans op "X=4", *gegeven* dat "X is even", is 1/3". (in het Nederlands wordt meestal trouwens van 'voorwaardelijke kansen' in plaats van 'conditionele kansen' gesproken. Ik houd het bij 'conditioneel' omdat ik het woord 'conditioneren' veel ga gebruiken)

2. Het Drie Gevangenen Probleem (1959)

Dit probleem duikt voor het eerst op in de literatuur in 1959, in de column van Martin Gardner in de *Scientific American*. Het gaat als volgt. Er zijn drie gevangenen, **A**, **B** en **C.** Twee van hen worden willekeurig uitgekozen en zullen worden terechtgesteld. De gevangenen weten dit. Bijvoorbeeld, A wordt met kans 2/3 geëxecuteerd, dus hij overleeft met kans 1/3. Rita, de cipier, komt langs. A vraagt haar of zij misschien kan zeggen of B of C wordt terechtgesteld. De cipier zegt: *B.* Wat is nu de nieuwe kans dat A wordt terechtgesteld?

We kunnen hier op twee manieren naar kijken. In eerste instantie lijkt het alsof de cipier geen nieuwe informatie geeft over A's overlevingskans geeft. A wist toch al dat B of C zou worden terechtgesteld. Dus volgens deze redenering blijft de kans dat A overleeft 1/3.

De tweede manier om ernaar te kijken is te gaan conditioneren. Het blijkt dat met conditioneren, de kans ineens omhoog springt naar 1/2! Er waren nl. eerst *drie mogelijkheden* met gelijke kans: A overleeft, B overleeft, C overleeft. Nadat de cipier zegt "B wordt terechtgesteld," zijn er nog *twee mogelijkheden* over: A overleeft, C overleeft. Dus de kans dat A overleeft is nu 1/2 !

Op dit moment richtte ik mij tot het publiek en vroeg of iedereen die dacht dat conditioneren (antwoord 1/2) hier correct was, zijn hand wilde opsteken. Zo'n 10% van de zaal stak zijn hand op. Helaas is conditioneren hier verkeerd! Om dit in te zien, hoeven we alleen maar te bedenken dat als de cipier C in plaats van B had geantwoord, we met conditioneren ook op 1/2 waren gekomen. Dus: A stelt een vraag, en *wat het antwoord op die vraag ook is*, nadat hij het antwoord heeft gehoord gaat de kans dat hij het overleeft omhoog. Dat kan niet goed zijn! Conditioneren geeft blijkbaar niet altijd het juiste antwoord.

Ik vroeg toen aan het publiek wie er dacht dat het eerste antwoord (kans blijft 1/3) correct was. Weer stak zo'n 10% van het publiek zijn hand op. Bij nader inzien is dat ook niet helemaal goed, alhoewel het hier subtieler ligt.

Het juiste antwoord is: *je kunt niet meer zeggen wat de kans is***,** tenzij je extra aannames doet over de psyche van de cipier. Om dit in te zien kan ik het beste eerst even teruggaan naar het dobbelsteenverhaal.

2.1 Dobbelstenen en Correcte Kansuitspraken

Net als zojuist gooi ik met een eerlijke dobbelsteen die u niet ziet. Ik vertel u ofwel "de uitkomst lag *tussen 1 en 3*" ofwel "de uitkomst lag *tussen 4 en 6*". Ik vraag u wat de kans op '4' is. U bepaalt het antwoord door te conditioneren. Als ik zeg "tussen 1 en 3," dan zegt u: "de kans op '4' is 0". Als ik zeg "tussen 4 en 6," dan zegt u: "de kans op '4' is 1/3". Dit is geheel correct. Maar *waarom* is het eigenlijk correct? Wat voor uitspraak doet iemand eigenlijk over de wereld om hem heen als hij of zij zegt "de kans op 4 is 1/3"? Om dat te zien doen we een gedachtenexperiment: we gaan hetzelfde spelletje 6000 keer herhalen . Dan zal ik ongeveer 3000 keer zeggen "tussen 4 en 6". In ongeveer 1000 van die 3000 gevallen (1/3 dus) is de uitkomst '4'.

Dus, en dit is de clou van het het verhaal: de uitkomst is '4' *in ongeveer 1 op de 3 van de gevallen waarin u* **zegt** *"de kans op '4' is 1/3."* We kunnen dan zeggen dat uw kansuitspraak "correct" is: als we de situatie waarin u de uitspraak doet voldoende vaak herhalen, en de onderliggende aannames (eerlijke dobbelsteen) kloppen, dan zal de waargenomen frequentie van een uitkomst ongeveer gelijk zijn aan de kans die u heeft vermeld.[1] In dit voorbeeld "werkt" conditioneren dus, omdat het tot correcte kansuitspraken leidt. Ik geef nu echter een soortgelijk voorbeeld waarin conditioneren *niet* werkt.

2.2 Dobbelstenen en Correcte Kansuitspraken – Vers 2

Stel, we spellen het bovenstaande spel, maar ik vertel u nu ofwel "de uitkomst lag *tussen 1 en 4* " (in plaats van: tussen 1 en 3, zoals net) ofwel "de uitkomst lag tussen 4 en 6". Ik vraag u nu de kans op '4' te bepalen. U bepaalt deze kans wederom door te conditioneren. Dus als ik zeg "tussen 4 en 6" zegt u nog steeds: de kans op "4" is 1/3. Stel nu dat we deze variatie van het spel vaak herhalen. Het verschil met de vorige situatie is dat ik thans, steeds als de uitkomst 4 is, een ***keuze*** heb in wat ik u ga vertellen. Ik kan het bijvoorbeeld zo doen: als de uitkomst 4 is zeg ik altijd "tussen 1 en 4" en *nooit* "tussen 4 en 6." Als we het spelletje 6000 keer spelen, zal ik dan ongeveer 2000 keer "tussen 4 en 6" zeggen. Elk van die keren zult u zeggen: "de kans op 4 is 1/3", maar in werkelijkheid zal van al die keren de uitkomst geen enkele keer 4 zijn. Uw uitspraak "kans op 4 is 1/3" is dan dus *niet* correct; in dit geval had u moeten zeggen '0', dus conditioneren geeft het verkeerde antwoord. Maar het echte probleem ligt niet bij conditioneren. Er bestaat in feite geen enkele methode om uw kansen aan te passen die wel altijd het juiste antwoord geeft. Immers, ik zou zelf ook een eerlijk muntje kunnen gooien als de daadwerkelijke uitkomst 4 is. Bij kop zeg ik "tussen 1 en 4", en bij "munt" zeg ik "tussen 4 en 6." In dat geval zal ik ongeveer 2500 keer "tussen 1 en 4" zeggen, en ongeveer 500 van die keren zal de uitkomst "4" zijn. De "correcte" kans is dan dus 1/5. Ik kan ook met een valse munt gooien om te bepalen wat ik ga zeggen, en op deze manier kan ik iedere frequentie tussen 0 en 1/3 bereiken. Ik zou de keuze ook af kunnen laten hangen van de kleur van de eerstvolgende auto die ik buiten langs zie rijden, of wat dan ook.

---

[1] Er zijn veel toepassingen van kansrekening waarbij het subtieler ligt, maar daar zal ik hier verder niet op ingaan.

Het is duidelijk: als u niet weet op wat voor manier ik de keuze maak wat ik u ga vertellen telkens als de daadwerkelijke uitkomst 4 is, dan is het voor u onmogelijk om een correcte kans te bepalen, met wat voor methode dan ook. We zouden ook kunnen zeggen: een “correcte” kans bestaat nu niet meer, tenzij we bereid zijn extra aannames te doen. Een uitspraak als “de kans is X”, voor wat voor X dan ook, is potentieel geheel verkeerd, of zelfs betekenisloos, tenzij X een interval in plaats van een getal is, of tenzij er bij verteld wordt wat voor aannames gedaan worden over de keuzes die ik maak.

2.3 Terug naar de Drie Gevangenen – Keuze en Overlap

Ook het 3-gevangenen probleem wordt veroorzaakt doordat de cipier (soms) een **keuze** heeft in wat zij gaat vertellen. Als A overleeft (B en C worden geëxecuteerd), kan de cipier kiezen of zij B of C zegt. De echte kansen zijn wederom niet te bepalen als men niet weet wat voor strategie de cipier dan volgt. Als de cipier een eerlijk muntje gooit om te kiezen of zij B of C zegt, dan blijft A’s overlevingskans 1/3, wat ze ook zegt. Vandaar dat sommigen wel iets voelen voor het antwoord ‘de kans blijft 1/3’: daarin zit de impliciete aanname dat de keuze van de cipier onafhankelijk is van de daadwerkelijke uitkomst. Toch is het niet goed om zomaar te zeggen ‘1/3’; wel goed is om te zeggen: ‘*als* ik ervan uitga dat de keuze van de cipier onafhankelijk is van de daadwerkelijke uitkomst, *dan* geeft zij A geen informatie en blijft de kans 1/3’.

Kort samengevat: u kunt uw kansen alleen aanpassen aan nieuwe informatie als er geen mogelijke *overlap* zit in de informatie die u kunt krijgen, zodat de persoon of machine van wie u de informatie krijgt, nooit een keuze kan maken. In dat geval dient u de kansen aan te passen door te conditioneren. Als er wel mogelijke overlap is, dan zijn de kansen in feite niet meer goed gedefinieerd, en bestaat er geen enkele correcte methode om uw kans aan te passen.

3. Het Drie Deuren (Quizmaster) Probleem (+/- 1970)

U doet mee aan een televisiequiz. In de studio zijn drie deuren. Achter één deur staat een auto, achter beide andere deuren een geit. Het spel gaat als volgt: U gaat voor een van de deuren staan. Monty Hall, de quizmaster, opent een van de twee andere deuren, en laat zien dat er een geit achter zit. U mag nu blijven staan, of wisselen naar de de deur die nog dicht is. Vervolgens doet Monty de deur open waar u uiteindelijk voor bent gaan staan. Uw prijs is wat er achter die deur zit. U hoopt natuurlijk dat het de auto is. De vraag is nu: *is* *het verstandig om te wisselen nadat Monty Hall een deur met een geit erachter heeft opengemaakt?* We gaan er hierbij vanuit dat u zelf de eerste deur willekeurig hebt gekozen, en dat Monty (die weet waar de auto is) vervolgens hoe dan ook een deur met een geit zal openen.

Het blijkt dat veranderen van deur *zeer verstandig* is. Grofweg gezegd verhoogt u hiermee uw winstkansen van 1/3 tot 2/3. Vrijwel alle mensen, inclusief de meeste wiskundigen, denken echter in eerste instantie dat het niets uitmaakt of u van deur wisselt of niet! Voor beide dichte deuren geldt immers dat de kans dat de prijs erachter zit, gelijk is?

Een variant van het 3-deuren spel werd daadwerkelijk gespeeld in de quiz *Let's Make a Deal!* die in de jaren '60 en '70 heel populair was in de VS, en die, inderdaad, gepresenteerd werd door ene Monty Hall. Het schijnt dat hetzelfde spel ook op de Nederlandse televisie te zien was, in een van de spelshows van Willem Ruys in de jaren '80. In 1990 beweerde Marilyn Vos Savant (volgens aanhoudende geruchten de "vrouw met het hoogste IQ ter wereld") in haar column in het tijdschrift *Parade* dat het veel beter was om van deur te verwisselen. Meer dan 10000 lezers reageerden op deze column met de opmerking dat Vos Savant natuurlijk ongelijk had. Onder deze briefschrijvers waren meer dan honderd wiskundehoogleraren. Ook Paul Erdös, een van de grootste wiskundigen van de 20^{e} eeuw, weigerde te geloven dat Vos Savant gelijk had. Er zijn hele boeken over het probleem geschreven (zie bijvoorbeeld G. von Randow, *Das Ziegenproblem,* Rowohlt 1992). Toch: hoewel er wel wat aan te merken valt op de analyse van Vos Savant, is haar conclusie volkomen correct: in dit spel is het beter om van deur te verwisselen.

Het grappige is dat, wiskundig gezien, het 3-gevangenen probleem en het 3-deuren probleem equivalent zijn. Net als bij het 3-gevangenen probleem zijn er in het 3-deuren probleem drie mogelijkheden: A, B en C. (A betekent hier dat de auto achter deur A zit; B betekent dat de auto achter deur B zit, en C dat de auto achter C zit). Stel u gaat voor deur A staan. Net als Rita geeft Monty u dan als informatie dat ofwel B niet het geval is, ofwel C niet het geval is ('B is niet het geval' betekende eerst 'B overleeft niet', en nu 'geen auto achter deur B'). Stel dat Monty B zegt. Wanneer u denkt dat het nu niets uitmaakt of u voor A blijft staan of naar C gaat, komt dat omdat u (misschien onbewust) conditioneert: er waren eerst drie mogelijkheden met gelijke kans, A, B en C. Daarvan zijn er twee over, A en C. Die hebben dus nog steeds gelijke kans. Die kans is dus 1/2, en het maakt dus niet uit. We hebben al gezien dat deze redenering verkeerd is. *Het gekke is dat vrijwel iedereen – wiskundige of niet – bij het 3-gevangenen probleem de juiste intuitie heeft (nl.: conditioneren werkt niet) terwijl bij het 3-deuren probleem vrijwel iedereen de verkeerde intuitie heeft (nl. conditioneren werkt wel).* Dit voorbeeld laat nog eens duidelijk zien hoezeer onze intuitie bepaald wordt door de manier waarop we een en hetzelfde wiskundige probleem "verpakken." Het probleem kan zodanig verpakt worden dat bijna iedereen de verkeerde intuitie heeft.

3.1. Een Complicatie

"Maar wacht nou eens even!" zult u nu misschien zeggen: de les van het 3-gevangenen probleem was nou juist dat, nadat de cipier B of C heeft gezegd, de kans op A eigenlijk niet meer gedefinieerd is, of in ieder geval, niet bepaald kan worden. Waarom kan men dan wèl zo'n stellige kansuitspraak doen in het equivalente drie-deuren probleem? We zeiden immers:

"de kans om te winnen als men van deur verwisselt, is gelijk is aan 2/3." *(Uitspraak I)*

Dit lijkt hetzelfde als de uitspraak:

“de kans dat de auto achter C zit nadat de quizmaster deur B heeft geopend, is 2/3; de kans dat de auto achter A zit, is nu dus 1/3.” *(Uitspraak II)*

Maar in het 3-gevangenen probleem hadden we juist gezegd dat, nadat Rita ‘B’ heeft gezegd, de uitspraak ‘de kans op A blijft 1/3’ *verkeerd* is, omdat de echte kans van Rita’s keuzes afhangt, en dus niet meer te bepalen is. Hoe zit dat nou?

Het antwoord is subtiel: stel dat we *voordat* Monty Hall de eerste deur open doet, al bepalen dat we de volgende strategie gaan gebruiken: we kiezen eerst een willekeurige deur uit, zodat elke deur met kans 1/3 gekozen wordt. Vanaf dat moment noemen we de deur waar we voor staan deur A. Van de twee andere deuren moet er eentje verder naar links staan dan de andere. Die noemen we deur B; en de overgebleven deur noemen we deur C. Nu zal Monty dus of deur B of deur C openmaken en een geit laten zien. Welke deur Monty ook open maakt, wij gaan van deur veranderen en kiezen de overgebleven deur.

Het is eenvoudig in te zien dat, als we het spel 3000 keer herhalen, en we volgen steeds bovenstaande strategie, we ongeveer 2000 van de 3000 keer zullen winnen. Dus de kans dat we *winnen* is inderdaad 2/3: uitspraak I, het antwoord van Vos Savant is correct. Maar toch is uitspraak II *niet* correct. Deze uitspraak impliceert namelijk dat, wanneer we het spel herhaaldelijk spelen, *ongeveer 2/3* **van de keren dat Monty deur B opendoet***, de auto achter deur C zit.* En dit hoeft helemaal niet waar te zijn: Monty zou bijvoorbeeld altijd de meest rechts gelegen deur waarachter nog een geit zit open kunnen doen (dus als de auto achter A zit, en Monty dus kan kiezen tussen B en C, zal hij altijd deur C kiezen). In dat geval zal *iedere* keer dat Monty deur B opendoet, de auto achter deur C zitten. Het verschil tussen uitspraak I en uitspraak II is dat uitspraak I alleen iets zegt over het aantal keren dat we winnen op het *totaal* aantal keer dat we spelen; bij de strategie van altijd wisselen hangt de vraag of deze uitspraak correct is niet af van de door de quizmaster gemaakte keuze. Uitspraak II zegt ook iets over het aantal keren dat er iets gebeurt *in de gevallen waarin quizmaster deur B opent.* Dit hangt wel af van de door de quizmaster gemaakte keuze, en we kunnen er niet een bepaalde vaste kans aan toekennen.

Terzijde I. We hebben betoogd dat conditioneren in het drie deuren probleem niet werkt omdat de quizmaster (soms) een *keuze* heeft in wat hij ons vertelt. Een van de gangbare verklaringen waarom het antwoord ‘de kans is 1/2’ verkeerd is, is echter juist dat we in 2/3 van de gevallen in eerste instantie voor een geit staan en daarom de quizmaster *dwingen* om de enige andere deur met een geit erachter open te maken, en daarmee aan ons te signaleren dat de auto zich achter de overgebleven deur bevindt. Volgens deze redenering lijkt conditioneren juist niet te werken doordat de quizmaster in 2/3 van de gevallen *geen* keuze heeft. Dit lijkt in tegenspraak met de eerdere overweging. Bij nadere beschouwing verdwijnt de tegenspraak echter. Het is eenvoudig aan te tonen dat conditioneren altijd werkt (in de zin dat frequenties, bij veelvuldige herhaling van een experiment, ongeveer gelijk zijn aan de kansen) als degene die de informatie verstrekt *nooit* een keus heeft wat hij/zij gaat vertellen. Het is ook eenvoudig aan te tonen dat conditioneren altijd werkt als degene die de informatie verschaft *altijd* de keus heeft uit 2 berichten, en *altijd* met een eerlijke munt gooit om te bepalen wat hij gaat vertellen. In het 3-deuren probleem heeft Monty *soms* een keus tussen deur B en C (als de auto achter deur A staat), en *soms* geen keus (als de auto achter B of C staat). Deze situatie van soms wel, soms geen keuze is precies het geval waarin conditioneren niet werkt - zelfs als Monty “eerlijk” is en, als de auto achter A staat, met een eerlijke munt gooit om tussen B en C te kiezen.

Terzijde II. We hebben laten zien dat, als Rita of Monty ons vertellen dat B *niet* het geval is, de conclusie ‘de kans op A blijft 1/3’ inderdaad niet correct is. Toch suggereert bovenstaande analyse dat de uitspraak “de kans op A blijft 1/3” op de een of andere manier wel iets “beter” is dan de uitspraak “de kans op A wordt 1/2”. Dit kan men volgens mij inderdaad hard maken, maar dit vereist een niet-standaard wiskundige definitie van het begrip ‘kansverdeling’. De gebruikelijke definitie wordt dan een speciaal geval. Dit is een onderzoeksvraag waar ik thans aan werk. Een eerste aanzet in deze richting is gegeven in het paper *When Ignorance is Bliss (Proceedings UAI 2004)* dat ik samen met Joe Halpern (Cornell) heb geschreven. Het vervolgpaper belooft spannend en hoogst controversieel te worden!

4. Het ÉÉN Gevangene Probleem (2001-heden)

In 2004 werd verpleegkundige **Lucia de B.** in hoger beroep veroordeeld tot levenslang voor *7 moorden en 3 pogingen tot moord*. Zij zit al 5 jaar in de gevangenis, maar heeft nooit schuld bekend.Ton Derksen, auteur van het *boek Lucia de B., Reconstructie van een gerechtelijke dwaling* heeft de zaak aangekaart bij de Commissie Posthumus-II (evaluatie afgesloten strafzaken). Deze commissie heeft drie wijze mannen aangesteld, die sinds october 2006 aan het onderzoeken zijn of de zaak heropend moet worden. Als ik het goed heb begrepen wordt hun advies in de lente van 2007 verwacht.

In deze zaak heeft statistiek een cruciale rol gespeeld.[2] Er zijn in de statistische analyse een aantal fouten gemaakt. Zoals we zullen zien is een van de belangrijkste fouten dat er van een ‘kans’ wordt gesproken in een situatie waarin dat eigenlijk niet kan – de kans is dus betekenisloos, net als in het 3-gevangenen probleem.

---

[2] *Officieel* heeft statistiek ‘in de vorm van kansberekeningen’ geen rol gespeeld in het hoger beroep. Maar wie het arrest bekijkt ziet dat de (foutieve) statistische redeneringen ervan af druipen. Dat blijkt al uit het citaat uit het arrest op deze pagina. Het boek van Derksen bevat hier nog vele andere voorbeelden van. Ik geef er hier slechts een: bij een van de “vermoorde” patienten is aan zes medische experts gevraagd of het om een natuurlijke dood ging. Vijf van de zes dachten van wel. De *enige* expert die dacht dat het niet om een natuurlijke dood ging, is dezelfde arts die oorspronkelijk een natuurlijke doodverklaring had afgegeven. Maar aan die niet-natuurlijke dood dacht hij pas vier jaar later, zoals hij zelf verklaart, nadat “in de media aandacht werd besteed aan onverklaarbare sterfgevallen in de diverse Haagse ziekenhuizen". Het hof volgt deze laatste expert, die zich duidelijk heeft laten leiden door de statistische redenering dat zoveel onverklaarbare sterfgevallen ‘geen toeval kunnen zijn.’

Lucia werkte op de Medium Care Unit van het Juliana Kinderziekenhuis (JKZ) in Den Haag. Er waren veel meer “incidenten” (plotselinge sterfgevallen en reanimaties) wanneer Lucia wel dienst had, dan wanneer Lucia geen dienst had. Dat kon toch geen toeval zijn!?!?

Het hof vroeg een statisticus, Henk Elffers, om de gegevens nader te analyseren. Elffers concludeerde dat de kans dat een verpleegkundige bij toeval een dergelijk incidenten-patroon zou meemaken, kleiner is dan 1 op 342.000.000. Hij trok hieruit de conclusie dat het geen toeval kon zijn. Ik moet er meteen bijzeggen dat hij benadrukte dat dit *niet* betekent dat Lucia een moordenaar is. Ter illustratie geeft Elffers een aantal mogelijke alternatieve verklaringen, zoals ‘het zou bijvoorbeeld kunnen dat Lucia vaker nachtdiensten draait, en dat ’s nachts meer patienten sterven’. Het hof schrijft echter in zijn arrest (11.13):

Er is geen enkele aannemelijke verklaring gevonden voor het feit dat de verdachte in die korte periode bij zoveel overlijdensgevallen en levensbedreigende incidenten betrokken was.

Soortgelijke formuleringen komen een aantal keer voor in het arrest, en spelen een belangrijke rol in de bewijsvoering. Het hof gebruikt dus wel degelijk statistiek – hoewel Elffers’ getal niet genoemd wordt, wordt zijn statistische conclusie ‘het kan geen toeval zijn’ wel degelijk overgenomen. Helaas blijft er bij nadere analyse weinig over van deze conclusie.

Elffers deed een “nulhypothese toets” met significantieniveau 1/10000. Dit is een standaard statistische methode. In grote lijnen werkt het als volgt: we formuleren eerst een zogenoemde ‘nulhypothese’ en een ‘alternatieve’ hypothese. In dit geval was de nulhypothese ‘Lucia heeft dezelfde kans om een incident mee te maken dan andere verpleegkundigen’. De alternatieve hypothese is ‘Lucia heeft een hogere kans om een incident mee te maken als andere verpleegkundigen’. We kijken nu wat de kans[3] is op de daadwerkelijk geobserveerde gegevens als de nulhypothese waar zou zijn. Als die zgn. *overschrijdingskans* kleiner is dan het gekozen significantieniveau (in dit geval, 1 op 10000), dan verwerpen we de nulhypothese. Stel bijv. dat er 10 incidenten waren in de tijd dat Lucia op de afdeling werkte, en dat Lucia bij 8 incidenten aanwezig was. Dan berekenen we wat de kans is dat Lucia 8 *of meer* van die 10 incidenten meemaakt onder de aanname dat Lucia een evengrote kans heeft op een incident als andere verpleegkundigen. Elffers vond dat de kans dat Lucia evenveel of meer incidenten meemaakte, dan zij daadwerkelijk meemaakte, kleiner was dan 1 op 342 miljoen. Dat is veel kleiner dan 1 op 10000. Daarom verwerpt hij de hypothese “Lucia heeft dezelfde kans op incidenten als andere verpleegkundigen,” en hij concludeert hieruit “wat er gebeurd is, is geen toeval.”

4.1 “De” kans bestaat niet

Hoe werkt nulhypothese toetsen nou precies? De methode zit zo in elkaar, dat, als een

[3] We moeten hierbij heel voorzichtig zijn. We mogen niet zomaar de kans op de gegevens berekenen, want *elke* verzameling gegevens heeft uiteindelijk een hele kleine kans. Als we 10 keer met een eerlijke dobbelsteen gooien, dan heeft de uiteindelijke reeks die we gooien een kans van (1/6) tot de macht 10, vele malen kleiner dan 1 op 10000. Dit geldt altijd, welke reeks we ook gooien. We mogen hieruit natuurlijk niet concluderen dat de dobbelsteen vals is! Bij een nulhypothese toets bepalen we daarom niet de kans op de gegevens zelf , maar een zgn. *overschrijdingskans.* Dit is de kans op een speciaal gekozen *eigenschap* van de gegevens, waarbij die eigenschap aan bepaalde voorwaarden moet voldoen. Bij de dobbelsteen kunnen we bijv. kijken naar het gemiddeld aantal ogen. Als we waarnemen dat dat 4.5 in plaats van de verwachte 3.5 is, en de kans op een aantal ogen van 4.5 *of hoger* is kleiner dan 1 op 10000, dan kunnen we wel degelijk concluderen dat de dobbelsteen vermoedelijk niet eerlijk is. In het geval Lucia kijken we naar de (overschrijdings-) kans op *evenveel of meer incidenten* dan Lucia heeft meegemaakt.

statisticus hem herhaaldelijk (en correct) zou toepassen, dan zou gelden dat de statisticus gemiddeld maximaal 1 op 10000 keer *zegt* “dat kan geen toeval zijn” terwijl het wèl toeval is. Hij doet zo’n verkeerde uitspraak dus maximaal 1 op de 10000 keer.

Helaas is de nulhypothese toets in het geval Lucia niet correct toegepast, en *kan* hij ook helemaal niet correct toegepast worden. Dat zien we meteen als we ons gaan afvragen *wat “herhaaldelijk toepassen” hier zou moeten betekenen.*

Doen we de uitspraak ‘wel/geen toeval’

- telkens als een verpleegkundige *in het Juliana Kinderziekenhuis* zoiets meemaakt als Lucia?
- Of telkens als een verpleegkundige *ergens in Nederland/in Europa/op de wereld* zoiets meemaakt?
- Of telkens als er een *rechtzaak* is waarin het OM een nulhypothese toets gebruikt?

Dit is volstrekt onduidelijk. En als we Elffers’ berekening proberen aan te passen aan de drie gevallen hierboven, komen we in alle drie de gevallen op volledig verschillende getallen uit. Met andere woorden: zonder een precieze *context* aan te geven, is de uitspraak “het kan geen toeval zijn want de kans is 1 op 342 miljoen” simpelweg *betekenisloos*. Net als in het 3-gevangenen probleem is dit een ‘kansloze situatie’ waarin we niet, of in ieder geval niet zonder meer, van kansen kunnen spreken. In het hoger beroep is de rechter hierop expliciet geattendeerd door de hoogleraren M. Van Lambalgen (logica) en R. Meester (kansrekening), die optraden als deskundigen van de verdediging. Maar de rechter wilde hier niet aan, en bleef maar vragen ‘als u het niet met Elffers eens bent, wat is volgens u de kans dan wèl?’

Normaalgesproken worden nulhypothese toetsen toegepast in situaties waarbij de nul- en alternatieve hypothese van te voren geformuleerd worden, en getest worden op nieuwe, onafhankelijk verkregen gegevens. Er wordt bijvoorbeeld een speciaal experiment opgezet om die gegevens te verkrijgen. Als men dit zorgvuldig doet, dan kan men garanderen dat *gemiddeld van alle keren dat iemand, in wat voor context dan ook, een nulhypothesetoets* **correct** *uitvoert, de nulhypothese maar 1 op de 10000 keer onterecht verworpen zal worden.* De toets kan dus gaan over een nieuw geneesmiddel, de levensduur van gloeilampen of wat dan ook; de 1 op 10000 garantie kan gegeven worden. Maar in het geval van Lucia wordt de nulhypothese getoetst aan dezelfde data waardoor hij gesuggereerd is. Dan kan de 1 op 10000 garantie alleen gegeven worden als bekend is in wat voor context de toets uitgevoerd wordt, en als die context niet bekend is, is de uitkomst van de toets feitelijk betekenisloos.[4]

Het voorgaande suggereert dat de statistische analyse, hoewel die een zeer grote impact heeft gehad, eigenlijk niet zoveel zegt. Wanneer we andere relevante gegevens (beschikbaar ten tijde van de rechtszaak maar genegeerd door het hof) bekijken, dan krijgen we de indruk dat ‘de statistiek’, zo die überhaupt al iets kan zeggen, Lucia eerder vrijpleit dan verdacht maakt. Het blijkt nl. dat in de drie jaar dat Lucia op de medium care unit van het JKZ werkte, er daar **zes** sterfgevallen waren. *In de drie voordat ze er werkte, waren er* ***zeven****.* Voor een nadere analyse van wat dit betekent, verwijs ik naar Derksen’s boek. Derksen maakt ook

[4] Elffers realiseert zich wel dat er een probleem is, en past daarom een ‘post-hoc correctie’ op zijn hypothesetoets toe – hij vermenigvuldigt de overschrijdingskans met het aantal verpleegkundigen op Lucia’s afdeling. Maar met deze correctie kan nog steeds niet gegarandeerd worden dat maar 1 op de 10000 gevallen onterecht gezegd wordt ‘het is geen toeval’, omdat nog steeds onduidelijk is wat de context is: 1 op *welke* 10000 gevallen?

aannemelijk dat de gegevens waarop Elffers zijn analyse baseerde niet betrouwbaar zijn (dit valt Elffers natuurlijk niet aan te rekenen!). Verder werd er ook nog een rekenfout gemaakt (vermenigvuldigen van p-waarden). Hiermee blijft er niets, maar dan ook niets van de oorspronkelijke statistiek over.

Gelukkig beginnen statistici zich te roeren! Professor *Richard Gill*, hoogleraar statistiek in Leiden, winnaar van de ***Van Dantzig*** –prijs, lid van de Koninklijke Nederlandse Academie der Wetenschappen en oud MC/CWI'er, stuurde samen met mij een brief naar de Commissie Posthumus - zie het interview in Vrij Nederland van 6 december 2006 en het bericht in Nature van 19 Januari 2007. Professor W. van Zwet, emeritus hoogleraar statistiek in Leiden en "grand old man" van de Nederlandse statistiek, ook winnaar van de ***Van Dantzig*** -prijs, ook KNAW lid, en ook oud MC/CWI'er, heeft sinds 2004 lezingen over de zaak gegeven, inmiddels al op vier werelddelen. Ook Elffers is overigens oud MC/CWI'er.

En met deze verwijzing naar Van Dantzig, om wie het allemaal begonnen was, wil ik graag afsluiten. Ik dank u voor uw aandacht!