

Lucia de B. en de statistiek

[Ronald Meester]



Inleiding

De rechtszaken tegen de verpleegster Lucia de B., die veroordeeld werd voor de vermeende moord op een aantal van haar patiënten, hebben de afgelopen jaren heel wat stof doen opwaaien. Een van de bijzondere aspecten in deze casus was het gebruik van statistiek en kansrekening in de bewijsvoering. In eerste instantie werd Lucia de B. een verdachte doordat zich tijdens haar diensten onevenredig veel incidenten^[1] leken te hebben voorgedaan, en de vraag drong zich op of het optreden van zoveel incidenten aan het 'toeval' kon worden toegeschreven. De officier van justitie vroeg een statisticus, Henk Elffers, om op deze vraag een antwoord te geven, en het antwoord van Elffers was ondubbelzinnig: als alle incidenten gewoon toevallig zouden plaatsvinden, zo zei hij, dan zou de kans dat Lucia de B. minstens zoveel incidenten zou meemaken als feitelijk was gebeurd, kleiner zijn dan 1 op de 342 miljoen. Een onwaarschijnlijk klein getal, en dit getal inspireerde Henk Elffers om te concluderen: 'Als statisticus heb ik de vraag: "Kan dat toeval zijn?", beantwoord met een ongeclausuleerd "Nee".'^[2]

De berekening, en vooral ook de interpretatie van Elffers, werden vrijwel onmiddellijk door vele wiskundigen, onder wie ikzelf, scherp bekritiseerd. In hoger beroep heb ik daarop als getuige-deskundige betoogd dat het getal dat Elffers leverde, weinig of geen betekenis heeft, en dat men in het algemeen buitengewoon voorzichtig dient te zijn met dit soort berekeningen. De rechters waren tijdens dit verhoor *not amused* over wat ik te vertellen had. Het werd mij snel duidelijk dat zij de getuigenis van Elffers wensten te aanvaarden, omdat hij een eenduidig getal

leverde, namelijk de al eerder genoemde 1 op de 342 miljoen. Mijn opvatting dat een eenduidig getal als dit helemaal niet zinvol te leveren was, en dat er sowieso ernstige problemen waren in Elffers' methode, werd niet gewaardeerd.

Elffers' methode

Laten we nu eerst eens kijken wat Elffers eigenlijk gedaan had. De data die Elffers gebruikte, waren als volgt. De verdachte had in de periode waar het om ging, in twee ziekenhuizen gewerkt: in het Juliana Kinderziekenhuis (JKZ) en in het Rode Kruisziekenhuis (RKZ) alwaar ze op twee verschillende afdelingen diensten had gedraaid. In onderstaande tabel vinden we de gegevens over de diensten in de bestudeerde periode.

ziekenhuis (en nummer afdeling)	JKZ	RKZ-41	RKZ-42
totaal aantal diensten	1029	336	339
aantal diensten Lucia de B.	142	1	58
totaal aantal incidenten	8	5	14
aantal incidenten in diensten Lucia de B.	8	1	5

Op het eerste gezicht lijkt dit inderdaad wel wat te veel van het goede. Zo vonden bijvoorbeeld alle 8 incidenten in het JKZ plaats tijdens een dienst van Lucia de B. Onder de aannames dat:

1. er een vaste kans p bestaat dat gedurende een dienst een incident optreedt, en een kans
2. $1 - p$ dat er geen incident optreedt, en
3. incidenten onafhankelijk van elkaar optreden,

berekende Elffers eerst de *conditionele kans* dat alle incidenten in het JKZ gedurende Lucia's diensten plaatsvonden, *gegeven* dat er in totaal 8 incidenten plaatsvonden. Deze kans is eenvoudig uit te rekenen. Stel in het algemeen dat het totaal aantal diensten n is, en dat Lucia de B. er daarvan r voor haar rekening neemt. Laat nu A_k de gebeurtenis zijn dat er precies k incidenten plaatsvinden, en laat L_x de gebeurtenis zijn dat Lucia de B. tijdens haar diensten x incidenten ziet. We berekenen dan:

$$P(L_x | A_k) = \frac{P(L_x \cap A_k)}{P(A_k)} = \frac{\binom{r}{x} p^x (1-p)^{r-x} \cdot \binom{n-r}{k-x} p^{k-x} (1-p)^{n-r-k+x}}{\binom{n}{k} p^k (1-p)^{n-k}} = \frac{\binom{r}{x} \binom{n-r}{k-x}}{\binom{n}{k}}$$

In het concrete geval waarnaar we hier kijken is $n = 1029$, $r = 142$, $k = 8$, en tenslotte $x = 8$. De (conditionele) kans is dan gelijk aan $\frac{\binom{142}{8} \binom{1029-142}{8-8}}{\binom{1029}{8}} \approx 1,1 \times 10^{-7}$.

De kansverdeling die we hebben gevonden, staat bekend als de *hypergeometrische verdeling*.

Elffers bracht hiermee het probleem terug tot een berekening die je ook in termen van vazen en ballen kunt formuleren. Merk op dat de uitkomst van de berekening helemaal niet afhangt van de onbekende p . Volgens Elffers is dat maar goed ook, omdat deze onbekende p volgens hem niet te bepalen of te schatten is.

Bovenstaande berekening kun je voor elke afdeling doen waar Lucia de B. werkte. Bij het RKZ wordt dan de kans berekend dat de verdachte *minstens* 1, respectievelijk minstens 5 incidenten zou zien tijdens haar diensten.^[3]

Echter, nog steeds volgens Elffers, deze berekening is niet helemaal eerlijk voor de verdachte, omdat de berekening wordt uitgevoerd juist omdat de verdachte zoveel incidenten had meegemaakt. Vergelijk dit maar met de winnaar van een loterij: als je achteraf gaat uitrekenen wat de kans is dat juist deze persoon wint, dan zul je zien dat die kans heel klein is, maar natuurlijk maakt dit feit de winnaar niet tot een verdacht persoon. Het is daarom beter, aldus Elffers, om in het JKZ, het ziekenhuis waar de verdachte opviel, de kans uit te rekenen dat *één van de verpleegkundigen* alle 8 incidenten meemaakt. Ruwweg (maar niet helemaal) komt dat neer op het vermenigvuldigen van de gevonden kans met het aantal verpleegkundigen, namelijk 27. Voor



het RKZ is deze correctie niet nodig, omdat de verdachte toen al geïdentificeerd was in het JKZ.

Tenslotte komt Elffers tot zijn getal van 1 op de 342 miljoen door de uitkomsten van deze drie berekeningen met elkaar te vermenigvuldigen; *met* correctie in het JKZ, en *zonder* correctie in het RKZ.

Kritiek op de gang van zaken

Mijn kritiek kan onderverdeeld worden in een aantal verschillende categorieën. In deze bijdrage maak ik het volgende onderscheid.

1. De manier waarop de gegevens tot stand kwamen.
2. De keuze van het model van Elffers.
3. De berekeningen van Elffers.
4. De interpretatie van de cijfers van Elffers door het hof.

In deze zaak is weinig goed gegaan vanuit statistisch perspectief, en ik zal nu deze vier punten één voor één langslopen, en kort weergeven wat er mis is gegaan.

1. De manier waarop de gegevens tot stand kwamen

In het rapport van de Commissie Evaluatie Afsloten Strafzaken onder leiding van J.W.M. Grimbergen is duidelijk te lezen wat er op dit punt is misgegaan. Uit het rapport komt overduidelijk naar voren dat Lucia de B. in een veel te vroeg stadium als enige verdachte werd aangemerkt. Dit had tot gevolg dat incidenten waarvan al duidelijk was dat Lucia de B. er niets mee te maken had, als niet-relevant terzijde werden geschoven, en dus in bovengenoemde data helemaal niet voorkomen.

Het is duidelijk dat dit een onacceptabele gang van zaken is. Voordat het statistisch onderzoek überhaupt zou kunnen beginnen, is er al een fout begaan die niet meer gecorrigeerd kan worden: de gegevens waarop de statisticus zich moet baseren, zijn eenzijdig vastgesteld, en zijn dus in feite incorrect. Als alle incidenten waar Lucia de B. op voorhand geen invloed op kan hebben weggelaten worden uit de berekeningen, dan is het duidelijk dat dit sterke nadelen van haar is. Als bijvoorbeeld op het JKZ niet 8 maar 16 incidenten zouden hebben plaatsgevonden (ik noem maar wat), dan zou Lucia de B. niet alle, maar slechts de helft van alle incidenten hebben meegemaakt. Dit zou van grote invloed zijn op de uitkomst van de berekeningen. Ik wijs hier met grote nadruk op het feit dat deze rechtszaak in die zin bijzonder was, dat helemaal niet vaststond dat er überhaupt een moord was gepleegd. Dat gegeven maakt dat we buitengewoon voorzichtig moeten zijn met de data.

2. De keuze van het model van Elffers

Elffers heeft een wel heel eenvoudig model gekozen, en de vraag is of dit eenvoudige model de realiteit wel voldoende kan beschrijven.

Allereerst valt op dat hij de kans op een incident per dienst gelijk neemt voor alle diensten. Je kunt je afvragen of dat reëel is. Wellicht is er een verschil tussen bijvoorbeeld dag- en nachtdiensten, of is er verschil in de beleving van patiënten bij verschillende verpleegkundigen. Deze nuance is in het model van Elffers niet mee te nemen.

Ten tweede beperkt het model van Elffers zich tot de afdelingen waar de verdachte werkte. Deze keuze is niet goed of fout, maar het is wel een keuze. De vraagstelling was of de verdachte een 'te groot' aantal incidenten mee had gemaakt, en in deze vraag komt het woord 'afdeling' niet voor. Als je andere afdelingen of zelfs andere ziekenhuizen mee zou nemen, dan zou het model misschien te groot en te gecompliceerd worden, maar dat wil niet zeggen dat andere afdelingen of ziekenhuizen daardoor irrelevant zijn. Immers, ook als Lucia de B. iets overkomen is dat een heel kleine kans heeft, dan nog zal een dergelijke gebeurtenis 'ergens in Nederland' ook wel eens voorkomen. Het is een beetje te vergelijken met de winnaar van de loterij waar ik hierboven al over schreef. Net zoals het onredelijk is om uit te rekenen wat de kans is dat de feitelijke winnaar wint, is het ook onredelijk om zonder verder commentaar de berekeningen te beperken tot juist die afdeling waar iets bijzonders is gebeurd.

Maar er is nog iets belangrijkers. Door te conditioneren op het aantal incidenten dat plaatsgevonden heeft, valt de onbekende p weg uit zijn berekeningen, en zoals ik al schreef, is dat volgens Elffers maar goed ook, want over die onbekende p is volgens hem niet veel te zeggen. Welnu, of dat waar is, is op zijn minst onduidelijk, maar zelfs als hij hierin gelijk heeft, speelt die p op de achtergrond een belangrijke rol. Laat ik dit proberen duidelijk te maken met het volgende gedachtenexperiment.

Stel je even voor dat in een lange periode voordat Lucia op de afdeling kwam werken, er gemiddeld twee maal zoveel incidenten plaatsvonden als gedurende de tijd waarin Lucia werkzaam was op de afdeling. Dit zou natuurlijk relevante informatie zijn, want het aantal sterfgevallen gedurende de tijd dat Lucia er werkte, zou dan lager zijn dan normaal, en dan zou het toch wel enigszins merkwaardig zijn dat Lucia de B. als verdachte aangemerkt zou worden (ik roep in herinnering dat we helemaal niet weten of er überhaupt een moord gepleegd is). Ook deze overwegingen kunnen in het model van Elffers niet meegenomen worden, en dus is het heel goed mogelijk dat belangrijke informatie gemist wordt. In het eerder genoemde rapport van de commissie Grimbergen wordt op dit punt dan ook terecht forse kritiek geleverd op de gang van zaken.

In een gezamenlijke publicatie met Van Lambalgen, Gill en Collins^[4] geef ik met een berekening aan wat er zou kunnen gebeuren als je een model maakt waarin deze overweging wel een plaats krijgt. Kern van dat model is, dat het aantal incidenten dat een bepaalde verpleegkundige meemaakt, een Poisson-verdeling krijgt met een parameter μ die eventueel van verpleegkundige tot verpleegkundige kan variëren. We krijgen dan dat het aantal incidenten X van een verpleegkundige die r diensten draait, de kansverdeling:

$$P(X = k) = e^{-\mu r} \cdot \frac{(\mu r)^k}{k!} \quad (k = 0, 1, \dots)$$

heeft.^[5] De hypothese dat Lucia de B. geen relatie heeft tot de incidenten, zou dan geformuleerd kunnen worden door te zeggen dat elke verpleegkundige dezelfde μ heeft; deze hypothese geven we aan met H_d . De officier van justitie zou dan de hypothese kunnen formuleren dat de μ_L die bij Lucia de B. hoort, anders is dan de 'normale' μ van de andere verpleegkundigen. Deze hypothese geven we aan met H_p . Als we nu E schrijven voor de gebeurtenis dat verpleegkundige i precies k_i incidenten meemaakt, Lucia de B. nummer j krijgt, en het aantal verpleegkundigen I is, dan geldt dat:

$$P(E | H_d) = \prod_{i=1}^I e^{-\mu r_i} \frac{(\mu r_i)^{k_i}}{k_i!}$$

en

$$P(E | H_p) = \frac{e^{-\mu_L r_j} (\mu_L r_j)^{k_j}}{k_j!} \cdot \prod_{i=1, i \neq j}^I e^{-\mu r_i} \frac{(\mu r_i)^{k_i}}{k_i!}$$

De zogenoemde 'likelihood ratio' wordt dan gegeven door:

$$LR = \frac{P(E | H_p)}{P(E | H_d)} = e^{-(\mu - \mu_L) r_j} \left(\frac{\mu_L}{\mu} \right)^{k_j}$$

Hoe hoger dit getal, hoe relatief meer waarschijnlijk de gebeurtenissen zijn onder de aanname H_p van de officier van justitie, dan onder de aanname H_d van de verdediging. De formule is niet altijd even makkelijk te interpreteren, maar als nu bijvoorbeeld uit data voorafgaand aan de bestudeerde periode, zou blijken dat $\mu > \mu_L$, dan zien we dat dit getal kleiner (en dus ontlastender voor Lucia de B.) wordt naarmate het aantal incidenten k_j dat zij meemaakt, groter wordt. Zo *tegenintuïtief* kan het kennelijk zijn.

3. De berekeningen van Elffers

Hoeveel kritiek er op de keuze van het model ook te geven is, zodra het model eenmaal vastgesteld is, zouden berekeningen binnen dat model eenduidig moeten kunnen zijn. Maar zelfs dat is niet het geval bij de berekeningen van Elffers. Ik wil hier twee punten in Elffers' berekening nader belichten. Elffers zegt zich in zijn berekening alleen te willen baseren op dienstrooster-data, dat wil zeggen, de diensten van verdachte en andere verpleegkundigen, en de opgetreden incidenten. Het enige dat telt, is het aantal diensten dat door de verdachte gedraaid wordt, samen met het totaal aantal gedraaide diensten. Het aantal verpleegkundigen op de afdeling zou dus totaal irrelevant moeten zijn. Immers, de vraag die Elffers gesteld werd is of er reden is om te denken dat zich tijdens de diensten van de verdachte iets opvallends voorgedaan heeft. De enige gegevens die daarvoor nodig zijn, zijn gegevens omtrent de incidenten die zich tijdens, of juist buiten de diensten van de verdachte hebben afgespeeld; het aantal verpleegkundigen waarover de resterende diensten verdeeld worden is totaal irrelevant. We voeren nu een klein gedachtenexperiment uit. Als de diensten buiten de verdachte niet onder 26 resterende verpleegkundigen verdeeld zouden zijn maar onder 99 (ik noem maar een willekeurig getal), dan zou er voor Lucia de B. eigenlijk niets mogen veranderen. Maar als er 100 verpleegkundigen zouden zijn, dan zou de correctie van Elffers een vermenigvuldiging met 100, en niet met 27 moeten zijn. Natuurlijk, in de realiteit waren er 27 verpleegkundigen, en geen 100, maar dit gedachtenexperiment maakt duidelijk dat de feitelijkheid van de 27 verpleegkundigen geen enkele rol kan en mag spelen in de feitelijke berekening. Echter, in de berekening van Elffers speelt dit getal *wel* een rol.

Een tweede gedachtenexperiment geeft aan dat ook de spreiding van de diensten over de ziekenhuizen van belang is. Als de verdachte al haar diensten op één afdeling zou hebben gedraaid (in plaats van op drie), waarbij het totaal aantal diensten, het aantal door de verdachte gedraaide diensten, het totaal aantal incidenten en het aantal incidenten tijdens een dienst van de verdachte gelijk zouden zijn aan die van de reële casus, dan zou er een totaal andere kans uit komen, een veel grotere kans wel te verstaan. Dus het feit dat de verdachte haar diensten spreidt over meerdere afdelingen, is nadelig voor haar, en

het lijkt duidelijk dat er iets mis moet zijn. De achtergrond van dit probleem is dat de vermenigvuldiging die Elffers toepast, onjuist is. Je mag de getallen van de drie afdelingen niet met elkaar vermenigvuldigen, en dit is als volgt eenvoudig in te zien.

Stel dat een bepaalde verpleegkundige op 10 afdelingen werkt, en dat ze op elke afdeling een 'normaal' aantal incidenten meemaakt. Laten we, om precies te zijn, zeggen dat de kans dat ze minimaal zoveel incidenten meemaakt als feitelijk gebeurt, op elke afdeling gelijk is aan $\frac{1}{2}$. Volgens Elffers zou dan de kans dat ze op de 10 afdelingen bij elkaar zoveel incidenten meemaakt als feitelijk gebeurt is, gelijk moeten zijn aan $(\frac{1}{2})^{10}$. Dit is een zo kleine kans dat dit de verpleegkundige verdacht maakt... Onzin, natuurlijk! Maar het is wel wat Elffers in zijn berekeningen doet.

4. De interpretatie van de cijfers van Elffers door het hof

In haar vonnis van 24 maart 2003 nam de rechtbank het getal van Elffers over, maar uit de motivering bij het vonnis blijkt herhaaldelijk dat de rechters de betekenis van het getal 1 op de 342 miljoen niet begrepen hebben. We geven een drietal voorbeelden.

Onder punt 7 van de bewijsoverwegingen schrijft de rechtbank het volgende:

In zijn rapport van 29 mei 2002 komt dr. H. Elffers tot de conclusie dat de kans dat een verpleegkundige bij toeval net zoveel incidenten (...) zou meemaken als verdachte 1 op de 342 miljoen bedraagt.

Onder punt 8 vinden we:

Uit zijn rapport van 29 mei 2002 blijkt verder dat dr. H. Elffers tevens de navolgende deelkansen heeft berekend:

- a. *De kans dat bij toeval een van de 27 verpleegkundigen bij 142 uit 1029 diensten betrokken zou zijn bij 8 incidenten (...) is kleiner dan 1 op 300.000.*
- b. *De kans dat verdachte bij toeval (...).*

Onder punt 11 vinden we tenslotte:

De rechtbank is van oordeel dat uit de door dr. H. Elffers uitgevoerde waarschijnlijkheidsberekeningen, zoals neergelegd in zijn rapport van 29 mei 2002, volgt dat het uitermate onwaarschijnlijk moet worden geacht dat de verdachte de in de tenlastelegging genoemde incidenten bij toeval heeft meegemaakt. Deze berekeningen geven derhalve aan dat het in hoge mate waarschijnlijk is dat er een verband bestaat tussen het werkzaam zijn van de verdachte en het zich voordoen van bedoelde incidenten.

Ik citeer zo uitgebreid uit de bewijsoverwegingen om te laten zien dat de rechtbank systematisch formuleringen hanteert zoals *kans (...) bij toeval*. Elffers beoogt de kans te berekenen dat iets gebeurt als we aannemen dat incidenten gewoon per toeval en zonder samenhang plaatsvinden. De rechtbank meent echter ten onrechte dat Elffers de kans op toeval uitrekent. Deze fout wordt vaak gemaakt, en staat bekend als de *prosecutor's fallacy*. Kort gezegd komt de fout hierop neer dat $P(A | B)$ en $P(B | A)$ verwisseld worden, terwijl dit doorgaans toch echt twee geheel verschillende kansen zijn. Dit is dus niet alleen maar een kwestie van verkeerde woorden, zoals Elffers ter zitting van het hof wilde doen geloven. De interpretatie van de rechtbank is fundamenteel onjuist, en suggereert een veel sterkere bewijskracht van de getallen dan deze werkelijk hebben. Bovendien is de conclusie die de rechtbank onder punt 11 trekt ('deze berekeningen geven derhalve aan...') ook nog eens fout, zelfs als men de eerdergenoemde onjuiste interpretatie van de getallen voor lief neemt.

Conclusie

Is de conclusie dat we dan maar helemaal moeten afzien van het gebruik van kansrekening en statistiek in het strafrecht, gezien de mate van subjectiviteit die hierbij een rol speelt?

Nee, dat is niet de conclusie die ik trek, en dat brengt me ook bij de rol die de statisticus als expert zou moeten spelen in een proces. Een statisticus staat in een proces ten dienste van rechter, officier van justitie en/of verdediging. Zijn taak is om ervoor te zorgen dat gegevens die zich lenen voor statistische analyse, op de juiste wijze geïnterpreteerd worden. Dat betekent niet dat er slechts op één manier mee om te gaan is, dat heb ik betoogd. Juiste interpretatie betekent aan de ene kant dat de rechters inzicht moeten krijgen in de betekenis van gegevens. Is het gek dat bij een zuivere munt 65 van de 100 keer 'kop' gegoooid wordt? Ja, dat is gek, en het is de taak van de statisticus om dat uit te leggen. Maar ook, en misschien wel vooral, is het belangrijk dat rechters gaan begrijpen dat getallen een beperkte betekenis hebben. Een getal is nooit alleen maar een getal, maar altijd het gevolg van keuzes. Uiteindelijk moet de rechter de getallen zelf

beoordelen, en ik kan me zo voorstellen dat hij of zij daar de hulp van een goede statisticus heel goed kan gebruiken. De statisticus helpt, kort gezegd, bij het interpreteren en begrijpen van de data.

Toen ik, tijdens de zitting in hoger beroep, de rechter duidelijk trachtte te maken hoe statistiek werkt, en ik vertelde over het maken van berekeningen onder bepaalde hypothesen, werd ik onderbroken:

‘Mijnheer Meester, hier in de rechtszaal houden we ons niet bezig met hypothetische situaties, maar met de realiteit.’

Dat bedoel ik nou.

Noten

- [1] Wat een incident precies is, is voor mijn bijdrage hier niet van belang. Het komt ongeveer neer op de noodzaak tot reanimatie, ongeacht de afloop ervan.
- [2] In: *STAtOR* 5(2), juni 2004, p. 11.
- [3] Deze ‘minstens’ was in het JKZ niet nodig, omdat daarbij alle incidenten al tijdens Lucia de B.’s diensten plaatsvonden.
- [4] In: *Law, Probability & Risk* 2007.
- [5] De keuze van een Poisson-verdeling ligt enigszins voor de hand, vanwege de aanname van onafhankelijkheid tussen verschillende diensten.

Over de auteur

Ronald Meester is hoogleraar kansrekening aan de Vrije Universiteit Amsterdam. Reacties naar diens e-mailadres (rmeester@few.vu.nl).