

Statistiek en Rechtspraak

prima combinatie of riskant?

BWI-werkstuk geschreven door:

Margaret van Valkengoed
Bedrijfswiskunde & Informatica
Vrije Universiteit, Amsterdam

Voorwoord

Een van de laatste onderdelen van de studie Bedrijfswiskunde en Informatica is het schrijven van het BWI-werkstuk naar aanleiding van een literatuuronderzoek. Het is de opdracht om een onderwerp te behandelen wat zich op de scheidslijn bevindt tussen bedrijfsleven en wiskunde of bedrijfsleven en informatica. De rechtspraak en zo de bedrijven en organisaties die daarbij betrokken zijn, is misschien niet de meest voor de hand liggende “bedrijfstuk”. Maar wel één die ontegenzeggelijk een belangrijke rol speelt in onze maatschappij. Bij de bestudering van literatuur en gesprekken met buitenstaanders heb ik gemerkt dat de scheidslijn tussen de statistiek en de rechtspraak een onderwerp is wat zeer tot de verbeelding spreekt en waarin een hoop ontwikkeling gaande is. Mede hierdoor ligt er hier voor u een werk waar hopelijk voor u interessante zaken naar voren komen.

Ik wil dan ook mijn begeleider, dr. Sandjai Bhulai bedanken voor het aanreiken van dit onderwerp, waardoor het laatste loodje zo ontzettend veel lichter heeft kunnen wegen.

Margaret van Valkengoed
September 2004

Inhoudsopgave

1. INLEIDING	5
1.1 MOTIVATIE.....	5
1.2 CENTRALE VRAAG	5
1.3 OPBOUW	5
2. RECHTSPRAAK	6
2.1 STRAFRECHT.....	6
2.2 DE RECHTER	6
2.3 WETTIGE BEWIJSMIDDELEN.....	6
2.4 DE ROL VAN EEN DESKUNDIGE.....	7
2.5 DE ZITTING	7
3. STATISTIEK	8
3.1 WAT IS STATISTIEK?	8
3.2 KLASSIEKE VS BAYESIAANSE STATISTIEK	8
3.2.1 <i>De klassieke statistiek</i>	8
3.2.2 <i>De Bayesiaanse statistiek</i>	9
3.2.3 <i>Klassiek vs Bayes</i>	9
4. FORENSISCHE STATISTIEK	11
4.1 BEWIJSKRACHT	11
4.2 TOEPASSINGSGEBIEDEN.....	11
4.3 DE BAYESIAANSE STROMING BINNEN DE FORENSISCHE STATISTIEK.....	12
4.4 MOEILIKHEDEN VAN FORENSISCHE STATISTIEK	14
4.4.1 <i>Presentatie</i>	14
4.4.2 <i>Klassieke denkfouten bij interpreteren bewijs</i>	14
4.4.3 <i>Het Bayesiaanse model: mensen en cijfers</i>	15
4.4.4 <i>Fout positieve conclusies bij DNA-bewijs</i>	16
4.4.5 <i>Onderzoek binnen de forensische statistiek</i>	18
5. KANSANALYSE ALS BEWIJSMIDDEL IN RECHTSZAKEN	19
5.1 DE ZAAK LUCIA DE B.	19
5.1.1 <i>De feiten</i>	19
5.1.2 <i>De klassieke berekening van Elffers</i>	19
5.1.3 <i>De Bayesiaanse berekening van De Vos</i>	20
5.1.4 <i>De verdeelde meningen in de zaak Lucia de B.</i>	21
5.1.5 <i>De uitspraak</i>	22
5.2 DE ZAAK SALLY CLARK	23
5.2.1 <i>De feiten</i>	23
5.2.2 <i>De berekening</i>	23
5.2.3 <i>De meningen achteraf</i>	23
5.2.4 <i>De uitspraak</i>	24
6. CONCLUSIES	25
Literatuurlijst	2
5	
Bijlage A: Verslag van de gedane uitspraak m.b.t. de rechtszaak tegen de verpleegster Lucia de B.	de de 26

Bijlage B: Letter from the President to the Lord Chancellor regarding the use of statistical evidence in court cases.....27

1. Inleiding

1.1 Motivatie

Statistiek wordt toegepast op vele terreinen. Maar recentelijk zorgde een kansanalyse van een statisticus dienende als bewijs in een rechtszaak voor grote opschudding. Deze zaak is vervolgens breed uitgemeten door verschillende statistici die het niet eens waren over de aanpak van de betreffende statisticus en bijbehorende conclusie.

Dit heeft velen, vooral vele niet-statistici, zich doen laten afvragen: statistiek en rechtspraak, is dat wel een combinatie? Hierbij komt er ook een ethisch aspect om de hoek kijken. Aangezien de getrokken statistische conclusies een rol spelen in een uiteindelijke conclusie die enorme invloed heeft op de betreffende persoon.

1.2 Centrale vraag

Er zijn meer gebieden in de rechtspraak waar statistiek een rol speelt dan alleen in kansanalyses. Dit werkstuk is bedoeld om meer inzicht te verschaffen in welke rol statistiek speelt en kan spelen bij rechtspraak en welke problemen de toepassing van statistiek op dit gebied tegenkomt. Centraal zal staan de vraag:

Hoe is statistiek van toepassing in rechtspraak?

Door middel van een studie van verschillende publicaties in vakbladen en op het internet zal hier een antwoord op worden gezocht.

1.3 Opbouw

Het werkstuk begint met een overzicht van rechtspraak en statistiek in hoofdstuk 2 en 3, waarin het verschil tussen de klassieke en Bayesiaanse school in de statistiek naar voren komt. Vervolgens is hoofdstuk 4 gewijd aan de forensische statistiek: zijn verschillende gebieden, het Bayesiaanse model en moeilijkheden in toepassing. In hoofdstuk 5 zal de hiervoor genoemde rechtszaak uitgebreid besproken worden met de verschillende meningen van een aantal statistici waarna in hoofdstuk 6 de conclusies worden getrokken.

2. Rechtspraak

2.1 Strafrecht

In de rechtspraak zijn er verschillende soorten recht van toepassing. Strafrecht is de vorm van recht waar een persoon mee te maken krijgt wanneer hij ervan wordt verdacht een overtreding of een misdrijf te hebben gepleegd. In dat geval vraagt het Openbaar Ministerie de rechter om een uitspraak te doen. Is de verdachte schuldig of niet? Welke straf is passend als de verdachte schuldig is?

Om een strafzaak voor de rechter te brengen is informatie nodig. In het strafrecht heeft de politie als taak deze informatie te verzamelen. De politie onderzoekt wat er precies gebeurd is en verzamelt bewijzen om aan te kunnen tonen dat de verdachte een strafbaar feit heeft gepleegd. De resultaten van het onderzoek zet de politie op papier. Dit is het proces-verbaal. De officier van justitie vertegenwoordigt het Openbaar Ministerie. Hij leest het proces-verbaal en besluit dan of de zaak voor de rechter komt of niet.

2.2 De Rechter

De rechter treedt op als onafhankelijke buitenstaander. Hij doet een uitspraak, waaraan beide partijen zich moeten houden.

Alle rechters zijn onafhankelijk, maar wel gebonden aan het recht. De rechter moet zich houden aan de zogenaamde bronnen van het recht. Deze zijn de Grondwet en andere wetten, eerdere uitspraken (jurisprudentie), internationale verdragen, gewoonten en gebruiken.

2.3 Wettige bewijsmiddelen

In het wetboek van strafvordering staan de criteria voor wettige bewijsmiddelen geformuleerd. Zo zijn er verschillende vormen van wettige bewijsmiddelen:

- Eigen waarneming van de rechter
- Verklaring van de verdachte
- Verklaring van de getuige
- Verklaring van een deskundige
- Schriftelijke bescheiden

In Nederland kan één bewijsmiddel alleen, nooit voldoende bewijs zijn. Er zijn altijd diverse bewijzen nodig om tot een overtuigend bewijs te komen. In de meeste zaken is dit niet problematisch. De verdachte is bijvoorbeeld op heterdaad betrapt, er zijn meerdere getuigen, etc. In die zaken, waar het bewijsmateriaal lastiger ligt, kan het zijn dat er deskundigen aan te pas moeten komen. Vaak betreft het hier deskundigen die behoren tot de dienst Nationale Recherche Informatie (NRI) of onderzoekers van het Nederlands Forensisch Instituut (NFI). Zij rapporteren bijvoorbeeld over vingerafdrukken, DNA-profielen etc.

2.4 De rol van een deskundige

De verklaring van de deskundige vormt een bewijsmiddel. Het wetboek van strafvordering omschrijft een verklaring van een deskundige als volgt: *“Onder verklaring van een deskundige wordt verstaan zijn bij het onderzoek op de terechtzitting medegedeeld gevoelen betreffende hetgeen zijne wetenschap hem leert omtrent datgene wat aan zijn oordeel onderworpen is”*. Een deskundige heeft dus tijdens een rechtszaak als taak zijn mening kenbaar te maken betreffende hetgeen hem gevraagd wordt. Dit zal nooit de schuld of onschuld betreffen van de verdachte.

2.5 De zitting

De strafrechter onderzoekt op de rechtszitting wat er precies is gebeurd. Hij luistert naar wat de officier van justitie en de verdachte te vertellen hebben. Daarnaast hoort hij getuigen en deskundigen. Als de rechter vindt dat hij genoeg informatie heeft, doet hij uitspraak.

Ingewikkelde zaken komen voor drie rechters. Dit is de meervoudige kamer. Eén van de drie rechters is de voorzitter. Hij heeft de leiding. De meervoudige kamer doet na twee weken uitspraak, terwijl de enkelvoudige kamer meestal meteen uitspraak doet. De beslissingen van de meervoudige kamer komt tot stand na uitvoerig overleg in de raadkamer. Van de discussies in de raadkamer mag niets naar buiten komen.

3. Statistiek

3.1 Wat is statistiek?

Historisch gezien komt de term statistiek van het beschrijven van de staten: de toestand van de staat neergelegd in archieven van belastinggegevens en volkstellingen. Maar de statistiek is natuurlijk al lang veel meer dan dat. In het woordenboek tref ik aan: de wetenschap, methode en technieken van het verzamelen, bewerken en interpreteren van kwantitatieve gegevens [Kramers, 1996]. De statistiek geeft methoden en handvaten om uit verzamelde data conclusies te trekken. Door het toepassen van data analyse en statistische toetsing worden deskundigen op vele gebieden geholpen bij het interpreteren van verkregen data.

De Vos ziet het zo: “Statistiek is een rare wetenschap. Zij houdt zich bezig met de kunst precies geformuleerde conclusies te trekken uit onvoldoende gegevens. Dat is een hachelijke onderneming, maar omdat het overgrote deel van “wetenschap” hierop neerkomt erg nuttig”.

3.2 Klassieke vs Bayesiaanse statistiek

Samengevat zijn er twee verschillende scholen binnen de statistiek, genoemd de klassieke en de Bayesiaanse. Ik zal in het kort een idee geven van de denkwijze en methode van de verschillende scholen, waarna ik een aantal verschillen zal bespreken.

3.2.1 De klassieke statistiek

De klassieke school is ontwikkeld tussen 1930 en 1960, onder aanleiding van Ronald Fisher. De klassieke theorie berust zich op de frequentistische definitie van kans. Kansbegrip is gedefinieerd als frequentie in herhalingsexperimenten.

In de klassieke statistiek spreekt men zich uit over het al dan niet of wel statistisch significant zijn van de uitkomst van het experiment. Oftewel: is er met de gevonden uitkomst statistisch gezien iets te concluderen aangaande een bepaald ‘fenomeen’. Om zich daar over uit te kunnen spreken wordt er getoetst. Het toetsen op significantie vormt een belangrijk deel van de klassieke statistiek. Om te kunnen toetsen wordt er eerst een kansmodel geformuleerd, waarbij verschillende aannames worden gedaan ten opzichte van de toetsingsgrootte X . Onder dat model kan er een nulhypothese opgesteld worden, waarin de verwachte uitkomst voor de toetsingsgrootte verwoord wordt.

Wanneer de gevonden uitkomst (nogal sterk) afwijkt van de waarde onder de nulhypothese, dan kan dit nog aan toeval gewijd worden. Maar als de kans op een ‘extremere’ uitkomst (de overschrijdingskans) onder de nulhypothese wel heel erg klein is, kleiner dan een gegeven α , dan is er genoeg rede om de alternatieve hypothese voor waarheid aan te nemen en de nulhypothese te verwerpen. α geeft de kritieke grens aan van wanneer een uitkomst een indicatie is voor de onjuistheid van een hypothese en niet meer toeval. Hiermee is het ook de kans dat de nulhypothese onterecht

verworpen wordt, de zogenaamde 1^{ste} fout. Daarmee is alpha het onbetrouwbaarheidsniveau van de toets. De kans op de 2^{de} fout van een toets, de fout waarbij de nulhypothese onterecht niet verworpen wordt, probeert men te minimaliseren. Duidelijk hebben de twee fouten invloed op elkaar. Door een zeer kleine alpha te kiezen, wordt de nulhypothese minder snel verworpen en stijgt hiermee de kans op de 2^{de} fout. Gebruikelijke waarden voor alpha zijn daarom niet zeer klein en liggen tussen de 1% en 10%.

3.2.2 De Bayesiaanse statistiek

Deze school heeft zijn beginsel in de geloofswaarde van een bepaalde hypothese. Deze geloofswaarde bestaat er initieel, voordat er van data kennis is genomen. Met de introductie van data, kijkt men hoe deze de geloofswaarde van de hypothese beïnvloedt om vervolgens tot een uiteindelijke geloofswaarde te komen van de hypothese gegeven de data. Men spreekt zich in de Bayesiaanse statistiek dan ook uit over de waarschijnlijkheid van een bepaalde hypothese.

De methode die hiervoor wordt gebruikt, berust op onderstaande regel aangaande voorwaardelijke kansen, geïntroduceerd door de Thomas Bayes een achttiende-eeuwse dominee.

$$P[A | B] = P[A] \cdot P[B | A] / P[B]$$

Met behulp van deze regel kan worden aangegeven hoe men de waarschijnlijkheid van een bepaalde hypothese moet wijzigen in het licht van nieuwe data. Hierbij wordt de initiële waarschijnlijkheid van de hypothese de a priori kans genoemd ($P[A]$). Deze geeft de kans op juistheid van de hypothese voordat de nieuwe data bekend is. De waarschijnlijkheid van de data, aangenomen dat de hypothese juist is, geeft de likelihood functie ($P[B | A]$). Door deze te vermenigvuldigen met de a priori kans en vervolgens te delen door de kans op de data, wordt de posteriori kans verkregen ($P[A | B]$). De posteriori kans geeft de waarschijnlijkheid van de hypothese gegeven de nieuwe data. Als data kan er ook subjectieve data gebruikt worden. Data die voortkomt uit iemands inschatting, gevoel of ervaring voortkomt.

Door introductie van data wordt er dus een posteriori kans verkregen. Bij introductie van nieuwe data, kan deze posteriori kans dienen als a priori kans in de nieuwe berekening. Verschillende data kan zo in een berekening verwerkt worden en uitmonden in één gezamenlijke posteriori kans.

3.2.3 Klassiek vs Bayes

Verschillend zijn de klassieke en de Bayesiaanse school in de manier waarop ze zich uitspreken over kans. Bij de klassieke toetsing berekent men de overschrijdingskans van een bepaald onderzoeksresultaat onder voorwaarde dat de gestelde hypothese juist is. Als er daartoe reden wordt gezien, verworpt men de nulhypothese. Aan de hypothesen zelf worden geen kansen toegekend. Men vraagt zich niet af hoe waarschijnlijk een bepaalde hypothese is. In de Bayesiaanse statistiek daarentegen, richt men zich op de waarschijnlijkheid van de hypothese. Er zijn situaties waarvoor een

uitspraak over de waarschijnlijkheid van een hypothese zich uitermate goed leent in de context. In het hoofdstuk forensische statistiek komt er een voorbeeld naar voren. In dit zicht kan de Bayesiaanse statistiek een antwoord zijn op een wens die de klassieke statistiek niet kan vervullen.

Een probleem echter waar de Bayesianen (de aanhangers van de Bayesiaanse statistiek) mee te maken kunnen krijgen is de bepaling van de a priori kans. Als deze bepaald wordt op basis van subjectieve informatie, waarvoor de Bayesiaanse statistiek de ruimte geeft, rijst het gevaar dat deze afwijkt van de werkelijkheid. Er is onzekerheid betreft de juistheid van de gekozen a priori verdeling. Om de invloed van deze mogelijke onjuiste a priori verdeling te beperken kan een robuustheids onderzoek uitgevoerd worden waarmee gekeken wordt naar de invloed van verandering van de a priori verdeling op de posteriori verdeling. Door deze invloed klein te houden, zal de posteriori verdeling grotendeels alleen beïnvloed worden door de likelihood functie. Er is zelfs een ontwikkeling gaande waarbij er een zogenaamde niet informatieve a priori verdeling kiest. Bij vergelijking van waarschijnlijkheid voor twee verschillende hypothesen, hebben beide hypothesen a priori evenveel kans. De beweging binnen de Bayesiaanse statistiek om de rol van de a priori verdeling zo klein mogelijk te houden, of zelfs volledig non-informatief te maken, zal hen dichterbij de opvattingen van de klassieke statistici brengen. Bayesiaanse en frequentistische standpunten naderen elkaar na decennia van opgewonden betogen. [Van Zwet, 2004]

De introductie van subjectieve data, wat bij de klassieke statistiek niet kan, geeft kansen en gevaren. Kansen, daar waar weinig objectieve informatie aanwezig is en de subjectieve informatie aanwezig zeer relevant. Ook biedt de Bayesiaanse school de mogelijkheid om de gevoeligheid van een posteriori kans te tonen van een a priori kans. Hiermee geeft het een handvat voor de gevoeligheid van een 'klassieke conclusie' ten opzichte van mogelijke waarden een subjectieve a priori kans.

Gevaren rijzen daar waar onjuiste data wordt gebruikt, en de posteriori verdeling voor waar wordt aangenomen. [Bijnen, 1976]

4. Forensische statistiek

De toepassing van statistiek binnen de rechtspraak, de zogenaamde forensische statistiek, beslaat een zeer breed gebied. Statistiek en kansrekening liggen vaak ten grondslag aan de verklaringen van deskundigen en spelen daarom een belangrijke rol in het strafrecht. [Sjerps]. Het betreft hier deskundigen op verschillende gebieden, zoals bijvoorbeeld biologie, pathologie fysica en psychologie. Ondanks deze belangrijke rol is er tot op heden in de literatuur nog relatief weinig aandacht geschonken aan de toepassing van statistiek in het strafrecht. Er is dan ook nog volop onderzoek en ontwikkeling gaande in dit gebied.

4.1 Bewijskracht

Ondanks de verscheidenheid van deskundigen die op zichzelf ook weer verschillende statistische technieken toepassen, zijn er een aantal aspecten die de verschillende gebieden in de forensische statistiek gemeen hebben. Zo is er meestal een verdachte in beeld, en gaat het erom om iets te kunnen zeggen over de bewijskracht van een bepaald bewijsmiddel. In hoeverre kan er met dat specifieke bewijsmiddel meegeholpen worden aan de overtuiging van de schuld van de verdachte. Want onschuld hoeft niet bewezen te worden, het is de schuld die bewezen geacht moet worden. Er wordt een spoor die men associeert met het delict vergeleken met een bepaald kenmerk van de verdachte. Bij al deze vergelijkingen, en de gevonden overeenkomsten of verschillen, moet de grootte van bewijskracht, dus sterkte van de overtuigingskracht van het bewijsmiddel bepaald worden.

4.2 Toepassingsgebieden

Er zijn een aantal belangrijke toepassingsgebieden te onderscheiden, waar statistische overwegingen een belangrijke rol spelen.

- **DNA-onderzoek**

Wanneer men sporen als bloed sperma of speeksel van de vermoedelijke dader aantreft kan men proberen hieruit een DNA-profiel te verkrijgen. Vervolgens kan dit DNA-profiel vergeleken worden met het profiel van de verdachte. Als deze overeenkomt, kan de deskundige met behulp van kansrekening zich uitspreken over de match-kans van dit profiel. Deze kans spreekt zich uit over de kans dat een willekeurige persoon dit profiel zal hebben. Als er uit de sporen een volledig profiel verkregen kan worden, ontleent het grote bewijskracht aan de extreem lage match-kans. Bijvoorbeeld wanneer er uit de match-kans volgt dat er vermoedelijk maar 1 op de 5 miljoen mensen dit DNA-profiel heeft.

Lastiger wordt het wanneer er geen compleet profiel bepaald kan worden. Bijvoorbeeld wanneer er te kleine hoeveelheden van materiaal wordt teruggevonden of wanneer het materiaal al half vergaan is. Daarmee neemt de match-kans sterk toe en neemt de bewijskracht sterk af.

Ook vraagt men zich af hoeveel fouten er eigenlijk worden gemaakt in DNA-onderzoek. Daarom wordt er ook uitgebreid onderzoek gedaan naar de

invloed van de kans op deze zogeheten fout-positieve uitslagen. Oftewel een onterecht geconcludeerde match van aangetroffen DNA met dat van de verdachte. Hierover later meer.

- Sporenonderzoek

Bij een delict kunnen allerlei andere sporen gevonden worden, die vergeleken kunnen worden met kenmerken of spullen van de verdachte. Als we het hebben over lichamelijke kenmerken, denken we hierbij aan vingerafdrukken, een geurtest die een speurhond naar de verdachte leidt, een test voor spraak herkenning, afdruk van een gebit etc.

Maar ook kunnen er schoenen vergeleken worden met gevonden voetsporen, of kogelresten met een aangetroffen pistool.

- Gegevensverzameling

Daar waar conclusies getrokken moeten worden uit gegevens, is duidelijk ook van belang hoe er aan die gegevens is gekomen en onder welke omstandigheden. In deze situatie waar de gegevens nog verzameld moeten worden, speelt bijvoorbeeld representatieve monsternamen een rol. Dit kan zijn inzake milieudelicten, waar onderzocht moet worden of de concentratie van een bepaalde stof de wettelijke limiet overschrijdt. In hoeverre is dan om te gaan met kansen op een zogenaamde hot-spot in een partij, waar een piek in concentratie bevindt. Hoe kan je op die manier een monster nemen, zodanig dat er een representatieve waarde uit komt.

- Kansanalyse

Een heel andere kant van statistiek binnen de rechtspraak is wanneer een statisticus tijdens een rechtszaak als deskundige wordt opgeroepen om een kansanalyse te doen. Hoe groot is de kans op een bepaalde waarneming als er alleen toeval in het spel is. En wat is de kans dat ditzelfde een willekeurig persoon zou overkomen. Bekende gevallen hierin zijn zogenaamde dienststroosterzaken: een aantal incidenten vindt plaats wanneer een bepaalde werknemer dienst heeft. Kan er vanuit de data een uitspraak gedaan worden over een mogelijke associatie tussen de incidenten en de dienstdoende werknemer? Kan er statistisch gezien nog gesproken worden van toeval? Hiervan zullen we later nog een voorbeeld zien vanuit de praktijk.

4.3 De Bayesiaanse stroming binnen de forensische statistiek

Een invloedrijke stroming binnen de forensische statistiek is gebaseerd op Bayesiaanse ideeën. Sjerps vindt de Bayesiaanse benadering beter passen dan de klassieke benadering. “In het rechtssysteem, waar de vraag die de rechter moet beantwoorden uiteindelijk is of de verdachte schuldig dan wel onschuldig is, past een Bayesiaanse benadering mijns inziens beter dan de klassieke benadering” [Sjerps, 2004]. Voordat we verder gaan, laten we eerst kijken naar hoe het Bayesiaanse gedachtegoed vertaald is naar een model wat toepasbaar is in de forensische statistiek.

De hypothesen waar het in de rechtszaak meestal om zal draaien, is die van de aanklager inzake schuld van de verdachte (H_p) tegenover die van de verdediging inzake onschuld van de verdachte (H_d). Met behulp van de regel van Bayes de volgende formule aangaande de kansverhouding tussen de

twee hypothesen en hoe die verandert door introductie van een feit aangevoerd als bewijsmiddel (E):

$$\frac{P[H_p|E]}{P[H_d|E]} = \frac{P[H_p]}{P[H_d]} \cdot \frac{P[E|H_p]}{P[E|H_d]} \quad \text{of in woorden:}$$

a posteriori kansverhouding = *a priori kansverhouding* · LR.

De a priori kansverhouding, is een weergave van de initiële kansverhouding tussen de twee verschillende hypothesen. De is voorbehouden aan de rechter om te bepalen alvorens kennis te nemen van bewijsmiddelen. De 'likelikhoud ratio' (LR) oftewel aannemelijkheids ratio geeft een kansverhouding weer op de aannemelijkheid van het geïntroduceerde bewijs, onder aanname van de verschillende twee hypothesen. De a posteriori kansverhouding, gegeven door het produkt van de a priori kansverhouding en de LR, geeft vervolgens de kansverhouding aan van twee hypothesen gegeven het bewijsmiddel. Uit deze kansverhouding zou de overtuiging moeten spreken van een grotere waarschijnlijkheid voor een van de twee hypothesen, gegeven het bewijsmiddel om aan een uitspraak te kunnen meewerken.

Door dit model toe te passen kan er een numerieke waarde worden toegekend aan de bewijskracht van een bepaald bewijsmiddel. Immers geeft de LR de verhouding aan tussen hoeveel waarschijnlijker de verklaring van de aanklager wordt ten opzichte van die van de verdediging onder invloed van het geïntroduceerde bewijsmiddel. Op deze manier biedt het een deskundige een handvat om de bewijskracht te formuleren. Volgens deze Bayesiaanse stroming is het dan ook de taak van de deskundige om een LR te rapporteren en moet het schatten van de a priori verdeling overgelaten worden aan de rechter.

Verder heeft het Bayesiaanse model een speciale eigenschap die goed van pas kan komen in de praktijk van een proces. De gevonden a posteriori kansverhouding kan namelijk weer als a priori kansverhouding gebruikt kan worden in verdere berekeningen. Op deze manier kunnen verschillende bewijsmiddelen gecombineerd worden en uiteindelijk resulteren in één a posteriori kansverhouding.

In praktijk zou het er dan als volgt uit moeten zien:

De aanklager heeft een bepaalde hypothese (bijvoorbeeld de verdachte heeft de moord gepleegd), en de verdediging heeft een alternatieve hypothese (de verdachte heeft de moord niet gepleegd).

Aan het begin van het proces wordt de rechter geacht een bepaalde kansverhouding in te schatten ten aanzien van deze twee hypothesen. Naar aanleiding van de verklaring van een deskundige ten aanzien van de LR betreffende een bewijsmiddel komt hij door deze te vermenigvuldigen met zijn a priori kansverhouding op een nieuwe kansverhouding. Als het bewijsmiddel in het voordeel spreekt van de verdachte zal de LR kleiner zijn dan 1, en als het in het nadeel spreekt van de verdachte zal de LR groter zijn dan 1. Door dit te herhalen voor alle bewijsmiddelen die worden aangevoerd tijdens het proces (telkens de resulterende posteriori kansverhouding nemende als nieuwe a priori kansverhouding) komt de rechter op een uiteindelijke kansverhouding gegeven door de resulterende posteriori kansverhouding. Aan de hand van deze kansverhouding zou de rechter dan kunnen beslissen welke van de twee hypothesen aan te hangen.

Dit lijkt een vrij simpele toepassing. Maar welke moeilijkheden deze methodiek onder andere met zich meebrengt zullen we zien in de volgende paragraaf.

4.4 Moeilijkheden van forensische statistiek

Een moeilijkheid die niet alleen in de forensische statistiek optreedt, maar op verschillende terreinen waar statistiek wordt toegepast voor moeilijkheden zorgt is de interpretatie van de gevonden verbanden. Al in de periode direct na de oorlog heeft Van Dantzig, hoogleraar in de kansrekening en de mathematische statistiek, het over de principiële onmogelijkheid om met behulp van kansrekening iets over de wereld om ons heen te zeggen. “Wij trachten dat te doen door een wiskundig model aan te nemen en daarbinnen wiskunde te bedrijven. Het probleem zit echter in het in- en uitschakelen van het formalisme: hoe kies ik mijn model en hoe kom ik er weer van af, dat wil zeggen, hoe vertaal ik mijn wiskundige resultaten terug naar praktische conclusies.”

4.4.1 Presentatie

Wat betreft het vertalen van die resultaten naar praktische conclusies gaat het nog steeds wel eens mis. Deels door gebrekkige presentatie, waardoor het gemakkelijk verkeerd te interpreteren is voor een buitenstaander zoals een rechter of jurylid.

Een belangrijk punt in presentatie is bijvoorbeeld om diegene waaraan de vraag beantwoordt wordt in te lichten onder welke voorwaarde(n) de statistische conclusie geldt. Zoals Van Zwet schrijft: “De klant (diegene die de vraag stelt) realiseert zich soms onvoldoende dat statistische conclusies slechts betekenis hebben binnen een bepaald begrippenkader en heeft daardoor vaak een aanzienlijk rooskleuriger beeld van wat er statistisch is aangetoond.”

Daarnaast is het belangrijk is om redelijk aannemelijk te kunnen maken dat er geen alternatieve verklaring voor het gevonden statistische verband is. Andere verklaring voor het statistisch verband, dan het beweerde inzake schuld, moeten uitgesloten kunnen worden.

4.4.2 Klassieke denkfouten bij interpreteren bewijs

Bij de interpretatie van statistische conclusies zijn er een aantal denkfouten die in praktijk regelmatig worden gemaakt. Ik zal er een tweetal behandelen.

- Prosecutor’s fallacy

Deze denkfout is het gemakkelijkst te laten zijn met behulp van een voorbeeld. Stel, het DNA-profiel verworven uit een aangetroffen bloedvlek vermoedelijk van de dader wordt vergeleken met het DNA-profiel van de verdachte. Het komt overeen en de deskundige verklaart dat de kans dat een willekeurige persoon dit profiel heeft 1 op een miljoen. Vanuit deze verklaring, wetende dat het profiel overeenkomt, redeneert men vervolgens: de kans dat de bloedvlek niet van de verdachte is 1 op de miljoen.

De fout die gemaakt wordt is te zeggen dat de kans op een overeenkomstig DNA-profiel gegeven dat het bloed niet van de verdachte is, gelijk is aan de kans dat het bloed niet van de verdachte is gegeven de overeenkomst. Deze fout is menselijk en al langer bekend in de psychologie. Maar zeker waar het DNA-profiel een van de weinige bewijsmiddelen vormt, en de match-kans uitzonderlijk klein is, kan deze denkfout immense consequenties hebben.

- Defence fallacy

De defence fallacy is eigenlijk een reactie op de prosecutor's fallacy. Kijkend naar de hiervoor beschreven zaak zal de verdediging kunnen zeggen: Als de kans 1 op de miljoen is dat een willekeurig persoon dit profiel heeft en er wonen nog zo'n 16 miljoen andere mensen in Nederland, dan zijn er naar verwachting 17 personen met hetzelfde DNA-profiel als de verdachte.

Hiervan uitgaande is de kans dat het gevonden bloed van de verdachte gelijk aan 1/17 is, dus ongeveer 6 procent.

In principe is deze redenering juist, maar alleen wanneer voor al deze 17 personen geldt dat het net zo waarschijnlijk is voor hen als voor de verdachte om op de plaats van delict geweest te zijn. Stel dat een baby van 6 maanden hetzelfde DNA-profiel heeft, is deze dan ook een van de mogelijke verdachten ervan uitgaande dat het bloed tot de verdachte behoort? Bij een zedendelict waar het DNA-profiel uit sperma wordt verworven is al helemaal geen vraag meer.

Wat er misgaat is dat er bewust of onbewust bepaalde aannames worden gedaan. Bijvoorbeeld dat de dader een Nederlander is, dat alle Nederlanders (inclusief baby's, bejaarden en gehandicapten) met even grote waarschijnlijkheid de dader zijn enz.

Belangrijk is dus om bij het formuleren van conclusies dit op zo'n manier op zo'n manier te doen dat denkfouten en misverstanden worden voorkomen.

4.4.3 Het Bayesiaanse model: mensen en cijfers

Bij de Bayesiaanse stroming wordt er verwacht van betrokkenen, rechter en/of jury om een gevoel voor cijfers te hebben. Zij moeten immers een a priori kansverdeling kunnen bepalen om de gepresenteerde LR te kunnen verwerken tot een posteriori kansverdeling. Onderzoek daarentegen heeft meerdere malen uitgewezen dat mensen grote moeite hebben met het cijfermatig denken en het schatten. Zo is er een voorbeeld waarbij een Duitse krant aan haar lezers vroeg:

“ Wat betekent 40%?”

met als mogelijke antwoorden:

- een kwart
- 4 uit 10
- elke 40^{ste}

Ongeveer 33% van de mensen koos een onjuist alternatief.

Als voorbeeld ter illustratie van hoe lastig het kan zijn introduceer ik een zaak die gespeeld heeft in Engeland. Het gaat om het geval van verdachte A, verdacht van verkrachting en het enige bewijsmateriaal is DNA-bewijs. De match kans van dit profiel is 1 op 200.000.000, dus uitermate klein. Het slachtoffer heeft een glimp opgevangen van de dader en beschrijft hem als tussen de 20 en 25 jaar oud. De verdachte is echter 37 jaar oud en in een confrontatie herkent het slachtoffer hem niet als de dader. De verdachte heeft een alibi voor de avond waarop het delict heeft plaatsgevonden, ondersteund door zijn vriendin.

De verdediging roept Professor Donnelly op als deskundige die het Bayesiaanse model aanwijst als een logische aanpak in deze zaak. Hij deelt het bewijs, behalve het DNA-bewijs, op in de volgende secties:

- Wat is de kans dat de verkrachter een willekeurige man uit de buurt is, zoals A?
- Wat is de kans dat A niet overeenkomt met de beschrijving van de ooggetuige als hij onschuldig was?
- Hoe overtuigend is het alibi van A als hij schuldig was? En als hij onschuldig was?

Vervolgens na een uitleg hoe het Bayesiaanse model toe te passen, introduceert Donnelly waarden voor de verschillende kansen. Hij maakt uitdrukkelijk duidelijk aan de jury dat zij zelf hier hun invulling aan moeten geven om op die manier op hun eigen kansverhouding te komen.

Als ik in de positie van jurylid was geweest had ik het een hachelijk situatie gevonden. Het Bayesiaanse model laat het inschatten van de a priori kansverhouding over aan de rechter en/of jury. Immers het gebied waarover die kansverhouding spreekt, de initiële ingeschatte verhouding tussen schuld en onschuld behoort niet tot de expertise van de deskundigen.

De expertise aangaande die inschatting behoort tot de rechter. Maar in hoeverre is die rechter/jury, instaat zijn ontstane inschattingen om te zetten in getalswaarden? Zeer onwaarschijnlijk, kan vele orden van grootte hebben. En deze verschillende orden van grootte kunnen grote invloed uitoefenen op de posteriori kans.

De Bayesiaanse theorie leent zich goed voor de gevraagde uitspraak: schuldig of onschuldig en ook al laat het zien hoe je bewijsmiddelen kan interpreteren en zo bewijskracht toekennen. Toch kan het in de praktijk lastig zijn om door 'niet statistici' toegepast te worden. Ook voor deskundigen die zich niet in de getallen begeven, als handschriftdeskundigen of wapendeskundigen, kan het lastig zijn om zich uit te moeten drukken in een LR, een kansverhouding. Iets, wat voor toepassing van het Bayesiaanse model vereist is.

4.4.4 Fout positieve conclusies bij DNA-bewijs

Over de jaren heen is de nauwkeurigheid van het vaststellen van een DNA-profiel ontzettend verbeterd wat zeer lage match-kansen oplevert. Maar waar nog steeds weinig helderheid over is, is de kans op fout positieve conclusies. Dit zijn gevallen waarbij de conclusie luidt dat het DNA-profiel overeenkomt met dat van de verdachte, terwijl dat in de werkelijkheid niet zo is.

Er zijn verschillende oorzaken te bedenken waardoor deze fout positieve conclusies worden gemaakt:

- Onnauwkeurige verzameling van materiaal
- Onnauwkeurige uitvoering van test
- Verwisseling van materiaal

Er zijn verschillende voorbeelden bekend waarbij een fout positieve conclusie is getrokken. Maar hoeveel gevallen zijn er onbekend? Er is bezorgdheid over de invloed van dit soort fouten op de bewijskracht van DNA-bewijs. De waterdichtheid die het eens bezat (er zijn zelfs deskundigen geweest die hebben verklaard dat fout positieve conclusies niet voorkomen bij DNA-tests) behoort tot de verleden tijd. Hier in Nederland, waar een enkel bewijsmiddel niet genoeg is om iemand te veroordelen, vormt dat misschien nog niet zo'n groot risico. Er moet immers een tweede bewijsmiddel worden aangedragen om overtuigend te kunnen zijn.

In landen echter als Engeland en de Verenigde Staten, lijkt men niet schuw te zijn van het aanvoeren van DNA-bewijs als enige bewijs. Interessant is dan ook een recente studie die de invloed van de kans op fout positieve conclusies op de bewijswaarde van DNA heeft [Thompson et.al., 2003]. Hierin wordt een Bayesiaans model gebruikt waarin match-kans en a priori kansverhouding van schuld van de verdachte wordt gecombineerd met verschillende waarden voor de kans op fout positieve conclusies. We zien in hun model het geval waar de a priori kansverhouding een zeer kleine kans op schuld geeft (bijvoorbeeld als er geen ander bewijs is) en de match-kans 1 op de 100 miljoen, dat de posteriori kansverhouding wijzigt van 1 op de miljoen naar 1 op 10 als de fout positief kans wijzigt van 0 naar 1 op 10.000. Dit betekent een vermindering in kans van de orde van grootte van 5, wanneer er een relatief kleine kans op fout positieve conclusies in het spel komt. Naarmate die kans stijgt, hoe kleiner de posteriori kans. Volgens deze studie kunnen posteriori kansen toch nog heel klein uitpakken, ook al is de initiële match kans indrukwekkend. De schrijvers van dit stuk begrijpen dan ook niet hoe er zo weinig aandacht besteed kan worden aan een numerieke waarde geven aan de kans op fout positieve conclusies, aangezien ze zoveel invloed hebben. Een reden die er wordt genoemd is rechters uit ervaring wel zouden weten in te schatten. Maar, maar brengen zij hier tegenin: "To rely on jurors' common sense to produce accurate estimates when experts cannot agree seems unduly optimistic" [Thompson et. al., 2003]. Zij zouden dan ook liever zien dat er meer moeite werd gedaan om tot een inschatting te kunnen komen.

Een manier om tot een inschatting te komen van het percentage foute positief conclusies is door zogenaamde 'proficiency tests' van DNA laboratoria, een soort bekwaamheidstests. Wanneer er dan onder dezelfde omstandigheden gewerkt zou worden als normaal, en de test blind is (dus zonder dat de onderzoeker zich bewust is van een test) zou hiermee een indicatie gegeven kunnen worden van het percentage fout positieve conclusies. Maar in 1996 verklaarde in de Verenigde Staten de National Research Council dat de mogelijkheid van hertesten, de beste manier is om fouten tegen te gaan. Aan de fouten die al gemaakt zijn bij het verzamelen van het materiaal wordt hier dan wel voorbij gegaan. Net als de gevallen waar te weinig materiaal aangetroffen is, om genoeg materiaal over te hebben voor een tweede test.

4.4.5 Onderzoek binnen de forensische statistiek

Uit het voorgaande blijkt dat er nog een hoop interessante discussies gaande zijn, om statistiek op een goede manier van dienst te kunnen laten zijn binnen de rechtspraak. Het wordt op een breed terrein ingezet, maar voor verschillende problemen moeten nog een oplossing gevonden worden.

Het Nederlands Forensisch Instituut is gespecialiseerd in het verzamelen, bewerken en interpreteren van allerhande bewijs waarbij forensische statistiek van toepassing is. Daarbij komen het problemen tegen waarnaar verder onderzoek gedaan moet worden. Uit de volgende activiteiten blijkt waar men verder onderzoek naar doet op het moment:

- Steekproef opzet bij verdovende middelen, haren, textiel en milieu-onderzoek
- Ontwikkelen van een onzekerheidsanalyse t.b.v. de verkeersreconstructie van botssnelheden
- Classificatie en individualisatie van oorafdrukken, kras- en indruksporen, gezichten en spectra
- Gebruik van databases in combinatie met DNA-bewijs
- Combineren van conclusies (met onzekerheden) in deelonderzoeken tot een samenvattende eindconclusie
- Interpretatie en conclusies bij DNA-identificatie
- Onderzoek naar het gebruik van waarschijnlijkheidsschalen bij het formuleren van conclusies
- Gebruik van Bayesiaanse theorieën en methoden in juridisch bewijs
- Onzekerheden die ontstaan bij de interpretatie en analyse van foto's en videobeelden

[www.forensischinstituut.nl]

5. Kansanalyse als bewijsmiddel in rechtszaken

Een enkele keer wordt een statisticus gevraagd als zijnde deskundige op te treden in de rechtszaal in Nederland. In het buitenland is men al meer gewend aan dit fenomeen. Het gaat dan meestal om zaken waarin aan de statisticus wordt gevraagd een oordeel te vellen over de ‘mogelijke toevalligheid’ van het voorgevallene. Wat voor rol de statistiek kan spelen en hoe, zal ik laten zien aan de hand van de veelbesproken dienstroosterzaak van de verpleegster Lucia de B (LdB). De meningen over de gemaakte berekeningen verschilden nogal onder een aantal statistici, waarmee het een interessante zaak is om nader te bekijken.

Daarna zal ik nog kort ingaan op een zaak die in Engeland heeft gespeeld.

5.1 De zaak Lucia de B.

5.1.1 De feiten

In een ziekenhuis vinden op de kinderafdeling 8 incidenten plaats, waarbij reanimatie van het patiëntje nodig is, in een periode van 343 dagen. Dit is opvallend, aangezien het met de patiëntjes kort daarvoor goed leek te gaan en zij zich geen van allen in een terminale toestand bevonden. Dan blijkt dat tijdens alle incidenten een bepaalde verpleegster, mevrouw V, de dienst had. In alle overige diensten, van de 26 andere verpleegsters, vinden in de betreffende periode geen incidenten plaats.

Gedurende de periode rijst er een verdenking betreffende de verpleegster, dat zij de hand in de incidenten heeft. Bij het achtste incident stapt de ziekenhuisleiding naar de politie die een onderzoek instelt. Er worden nog meer incidenten achterhaald in andere ziekenhuizen waar de verdachte hiervoor heeft gewerkt. En na het vinden van aanwijzingen van moord uit toxologisch onderzoek van de stoffelijke resten van de slachtoffers, besluit de officier van justitie de betreffende verpleegster te vervolgen. Een van de bewijsmiddelen die het hierbij wil opvoeren is dat het geen toeval kan zijn dat de verpleegster dienst deed bij alle incidenten. Het Openbaar Ministerie vraagt daarom een statisticus als deskundige op te treden om die vraag te beantwoorden naar zijn kunnen: “Kan dat toeval zijn?”

5.1.2 De klassieke berekening van Elffers

De opgeroepen statistisch deskundige door het OM is Henk Elffers. Hij behoort tot de klassieke school van statistici. Mijn benadering is me af te vragen wat het eigenlijk zou betekenen als geldt: ‘het is toeval dat mevrouw V zoveel incidenten meemaakt.’ “[Elffers]

Zich die vraag stellende, en conditionerend op de gegeven data komt hij tot kansmodel wat het ‘toevallig voorkomen van een incident’ beschrijft met behulp van een urnmodel. Hierbij staat elke bal in de urn voor een dienst (in totaal zijn er 1029 diensten in de 343 dagen) waarbij er 8 zwarte ballen zijn die staan voor de diensten met incident. De overige ballen zijn wit. De nulhypothese is dat dit model geldt, en dat het voorkomen van incidenten in

de diensten waarin V aanwezig was vergeleken kan worden met het trekken (zonder teruglegging) van ballen uit deze urn, zodat er geen associatie is tussen het dienst hebben van V en het voorkomen van een incident. Hij stelt hierbij de alternatieve hypothese op het er wel zijn van een associatie. Onder dit model is het aantal zwarte ballen bij een trekking van 142 ballen (het aantal diensten van V) hypergeometrisch verdeeld.

Het uitvoeren van Fishers exacte toets leert dat bij het aantreffen van 8 incidenten de nulhypothese wordt verworpen met een overschrijdingskans van $1.1 \cdot 10^{-7}$. Wat betekent dat Elffers een associatie concludeert tussen de incidenten en V.

Er is nu alleen een probleem en dat is dat de data zogenaamd ‘dubbel’ gebruikt is. Dit, doordat de data eerst gebruikt is om de verdenking op haar te laten vallen en daarna om aan te tonen het aantal incidenten dat zij meemaakt niet een toevalsmodel te rijmen is. Een methode om hier mee om te gaan in de klassieke statistiek is door het doen van een post-hoc correctie. Door de kans uit te rekenen dat bij minstens een van de verdachten (de 27 verpleegsters van de afdeling) zoiets extreems optreedt als bij LdB, wordt deze fout gecorrigeerd.

Hij past een zelfde analyse toe op de gegevens van de andere incidenten. De overschrijdingskansen onder H_0 van de combinatie van de drie datasets gecombineerd, resulteert in een overschrijdingskans van $2.92 \cdot 10^{-9}$.

Zijn antwoord op de vraag luidt dan ook: “Nee, het is geen toeval”, aangevuld met “de rest is aan u”.

5.1.3 De Bayesiaanse berekening van De Vos

Zoals eerder al beschreven is het met de Bayesiaanse methode van statistische analyse mogelijk om een uitspraak te doen over de kans op een bepaalde hypothese. De rechter wil weten wat de feiten te maken hebben met de schuldvraag. “Dit kan alleen via Bayes”, zegt De Vos.

De hypothesen die hier bekeken worden zijn die van de aanklager inzake schuld van de verdachte V (H_p) tegenover die van de verdediging inzake onschuld van de verdachte V (H_d), gebruikende het eerder geïntroduceerde model:

$$\frac{P[H_p|E]}{P[H_d|E]} = \frac{P[H_p]}{P[H_d]} \cdot \frac{P[E|H_p]}{P[E|H_d]}$$

Hierbij gaat De Vos uit van $P[E|H_d] \approx 1.1 \cdot 10^{-7}$, de kans op de 8 moorden gegeven dat LdB onschuldig is. Dezelfde kans als die Elffers berekende. Maar dan moet er aan de andere coëfficiënten nog een waarde worden verbonden.

Om aan een a priori kansverdeling te komen redeneert De Vos het volgende: er zijn 40.000 verpleegkundigen in Nederland. Stel dat er eens per 10 jaar een verplegende moorden pleegt, dan is de kans dat een willekeurige verpleegster moorden pleegt 1 op de 400.000. Wat geeft

$P[\text{Hp}] = (1/4) \cdot 10^{-5}$ een a priori kansverhouding 1/400000.

Voor $P[\text{E}|\text{Hp}]$, de kans op de 8 moorden gegeven dat LdB schuldig is, stelt De Vos op 1/2. Hij houdt er daarbij rekening mee dat door het starten van het onderzoek, haar de kans ontnomen is om nog meer moorden te plegen. Door een 50% kans te nemen, denkt hij er hooguit een factor 2 naast te kunnen zitten.

Verder rekenend, leidt dit vervolgens tot:

$$P[\text{Hd} | \text{E}] \approx 0.088$$

Dus de kans dat LdB onschuldig gegeven het feit van de 8 moorden komt volgens deze Bayesiaanse methode op 0,088.

5.1.4 De verdeelde meningen in de zaak Lucia de B.

Als we het voorgaande bekijken valt op dat de bij beide methodes van berekening, ook al wordt de conclusie anders geformuleerd, er hetzelfde beeld wordt gevormd van de kans op toeval/onschuld. Waar de klassieke stroming Hd verwerpt ten gunste van Hp, komt de Bayesiaanse stroming uit op een grote LR. Beiden concluderen dus min of meer dat het bewijsmateriaal veel waarschijnlijker is als Hp waar is dan als Hd waar is, en in die zien een sterke aanwijzing vormt voor Hp ten opzichte van Hd. [Sjerps] Alhoewel Sjerps de voorkeur van De Vos deelt voor de Bayesiaanse methode heeft zij ook kritiek op de manier van aanpak van De Vos. Door een a priori kansverhouding tussen Hp en Hd te formuleren, begeeft De Vos zich op een terrein wat volgens Sjerps gereserveerd is voor de rechter en niet voor de deskundige. Immers, de deskundige heeft alleen speciale kennis over het ene bewijsmiddel E. Als gevolg hiervan zou de deskundige dus ook geen uitspraken moeten doen over de a posteriori kansverhouding; aldus Sjerps. Van Zwet is het ermee eens dat de rechter en deskundige strict gescheiden taken hebben, maar vindt dat Sjerps in deze de rechter een onmogelijke opdracht geeft. Zoals De Vos het probleem stelt moet er een getalswaarde bepaald worden voor de kans dat LdB onschuldig is, zonder in acht te nemen dat er moorden gepleegd zijn. Van Zwet het niet eens met de invulling in deze van De Vos. Zoals reeds uitgelegd, schat De Vos zijn a priori kansverdeling in door alle verpleegkundigen in Nederland in het vizier te nemen en een veronderstelling te doen aangaande hun 'moordlustigheid'. Van Zwet brengt daar tegenin dat als je weet dat:

- Er in het ziekenhuis patiënten vermoord zijn
- Er N verdachten zijn voor de gepleegde moorden
- Je er vanuit gaat dat alle moorden door een verdachte zijn gepleegd

het niet meer interessant is om te weten dat de Nederlandse verpleegster met kans $(1/4) \cdot 10^{-5}$ gaat moorden. De moorden zijn namelijk gepleegd, en de groep van mogelijke verdachten is gedefinieerd (de 27 verpleegsters van de afdeling). Zonder nadere informatie heet de moordenaar dan met kans 1/N LdB. (Van Zwet)

Volgens van Zwet is het dan ook van essentieel belang dat er helderheid wordt verschaft door de rechtbank inzake deze veronderstellingen:

- Zijn er moorden gepleegd
- Zo ja, hoeveel moorden en op wie
- Wie vormen de groep van mogelijke daders

Zonder expliciet aan deze voorwaarden te hebben voldaan is volgens Van Zwet reden te twifelen aan de validiteit van de door Elffers gehanteerde methode. En zal een a priori veronderstelling in de Bayesiaanse analyse discutabel blijven.

Meester, de statisticus die werd opgeroepen door de verdediging als deskundige, heeft een aantal punten van kritiek op de aanpak van Elffers en denkt dat het getal waar Elffers op komt daarom weinig betekenis heeft. Ten eerste heeft hij er grote moeite mee dat de hypothese inzake LdB is opgesteld nadat de data al bekend was. Hij vergelijkt het toeval van de casus LdB met het toeval wat je zou kunnen aantreffen bij het opstellen van een model van een mogelijke lotto winnaar. Als je dit zou doen nadat de trekking heeft plaatsgevonden, dan is volgens het model de kans dat die bepaalde persoon de winnaar wordt bijzonder klein. Maar vervolgens wanneer dit dan toch gebeurt, zullen we toch niet zeggen dat de loterij niet eerlijk was. In analogie hiermee beperkt Elffers zich tot de afdeling waar iets opvallends heeft afgespeeld, zonder er rekening mee te houden dat dit van tevoren uitermate onwaarschijnlijk was. Daarbij noemt hij ook nog een aantal zaken die eventueel van invloed zouden kunnen zijn op het 'toeval', zoals een grotere kans van overlijden in de nacht waar LdB veel nachtdiensten draaide, of dat in het bijzonder LdB erg zieke mensen onder haar hoede nam. Maar hij gaat hierbij voorbij aan het feit dat het niet gaat om plotseling overlijden van ernstig zieke mensen, maar om incidenten waarbij reden is om uit te gaan van poging tot moord. En poging tot moord laat zich allerminst leiden door toeval. Het is dan ook eigenlijk verwonderlijk dat Elffers zich hier niet over uitsprekt. Met zijn uitspraak naar de rechter toe: "En de rest is aan U", schiet hij daarmee toch eigenlijk tekort in zijn taak als deskundige. Waarbij het toch ook een deel van zijn taak is om uit te leggen op welke manier zijn berekening uitgelegd moet worden in 'normaal Nederlands'.

Naar mening van Van Zwet moet het voorbeeld van Meester inzake de loterij aangepast worden, om het analoog te maken met de situatie zoals beschreven in het ziekenhuis. Stel dat er ontdekt wordt dat er geknoeid is met de random generator van de loterij en de groep verdachten wordt gevormd door 5 hoogleraren die deze geleverd hebben. En dan blijkt dat een van de hoogleraren woont op het adres waard de prijs op was gevallen. Toeval? Of een reden om de hoogleraar op te pakken. Er is immers een misdrijf vastgesteld: knoeien met de generator. En oppakken gebeurt dus op andere gronden dan op basis van de uitslag van de loterij.

Hier wordt eens te meer duidelijk hoe belangrijk bepaalde aannames zijn voor een model, om een zinnige conclusie te kunnen trekken. De voorzichtigheid die Meester dan ook maant inzake het geven van een statistisch bewijs, is dan ook begrijpelijk.

5.1.5 De uitspraak

Lucia de B. is schuldig bevonden van 3 moorden en 4 pogingen tot moord. In de uitspraak van de rechter kwam uitdrukkelijk naar voren, dat aan de kansanalyse geen waarde is gehecht om tot de slotconclusie te komen. Aangezien de statistici het onderling ook niet eens konden worden over de uitkomst, waardoor elk antwoord discutabel werd, is dit niet verwonderlijk. Als de statistici het onderling al niet eens zijn, hoe moet een buitenstaander dan omgaan met de conclusie

De rechter zag zich overtuigd van haar schuld in het overige bewijs.

Een verhandeling van de veroordeling is te vinden in Bijlage A.

5.2 De zaak Sally Clark

Een zaak die in Engeland voor grote opschudding zorgde, was die van Sally Clark. Ook hier werd een kansanalyse gebruikt als bewijsmiddel, maar betreffende deskundige was geen statisticus.

5.2.1 De feiten

Het echtpaar Clark krijgt binnen twee jaar twee kindjes, die beiden sterven op de leeftijd van een aantal weken. In het eerste geval is er alleen maar medelijden. Maar bij het tweede geval rijst er argwaan. Is het niet te toevallig, dat wiegendoed van een kindje dit gezin twee keer overkomt? Er kunnen geen duidelijke sporen gevonden worden die op een onnatuurlijke dood wijzen. Ook wijst het gedrag van de moeder niet op enige vorm van verwaarlozing of bedreiging. Toch wordt de moeder aangeklaagd.

Rechtspraak in Engeland wordt gedaan met behulp van een jury.

5.2.2 De berekening

In het proces tegen Sally Clark wordt een medicus als deskundige opgeroepen. Hem wordt de vraag voorgelegd, hoe hij de kans inschat van hetgeen geclaimd is gebeurd te zijn, twee maal wiegendoed in dit gezin.

Hij berekent, rekening houdend met de situatie van de ouders (rijk en niet-rokend) dat de kans op wiegendoed zeer klein is: 1 op 8500. Dan, is volgens hem, de kans op tweemaal wiegendoed gelijk aan het kwadraat van de kans op 1 maal wiegendoed wat geeft 1 op 73 miljoen.

5.2.3 De meningen achteraf

Over de berekening van de betrokken deskundige is een algehele verwondering te lezen in vele meningen. Door het produkt te nemen van de kans op 1 maal wiegendoed, om de kans op tweemaal wiegendoed te krijgen, gaat de deskundige uit van het feit dat wiegendoed onafhankelijk plaatsvindt. Maar dit is allerminst bewezen. In de studies die erna zijn gedaan, blijkt dat er juist een associatie lijkt te bestaan, waarmee eens wiegendoed, een grotere kans op tweemaal wiegendoed betekent. Over een genetische afwijking is geen duidelijkheid. Maar het zou erfelijk kunnen zijn. Dit in acht nemend is het nemen van het produkt onnavolgbaar.

Een ander argument wat zich laat horen, is dat men alleen heeft gekeken naar de onwaarschijnlijkheid van het plaatsvinden van tweemaal wiegendoed, zonder te kijken naar de onwaarschijnlijkheid van een dubbele moord door de moeder. Want ook al is het eerste zeer onwaarschijnlijk (maar meerdere gevallen hebben zich al vrij snel gemeld), het gaat er toch om welke van de twee verklaringen het minst onwaarschijnlijk is. Daarmee wordt het een geheel andere afweging. Immers, is kijkend naar de statistieken, de kans dat een welvarende moeder, met partner zich aan haar kinderen vergrijpt ook zeer klein. Maar daar is niet naar gekeken.

Als er nu duidelijk was aangetoond dat de kindjes zijn vermoord, dan sprak men al van een hele andere zaak. Maar dat was niet het geval.

5.2.4 De uitspraak

Sally Clark werd door de jury schuldig bevonden. Niet kan worden uitgesloten dat men hier de prosecutor's fallacy heeft begaan.

In Bijlage B valt een brief te lezen opgesteld door de voorzitter van het statistisch genootschap in Engeland over het optreden van de deskundige in de rechtszaal. In deze brief roept hij de rechter op om statistici te betrekken bij zaken als deze, waar om een kansanalyse wordt gevraagd. Om op die manier de misstanden die nu zijn gedaan te voorkomen.

6. Conclusies

Terugkomend op de vraag waar we mee begonnen zijn:

Hoe is statistiek van toepassing in rechtspraak?

Kunnen we concluderen dat statistiek op vele manieren van toepassing is in rechtspraak. Als ondersteuning van forensisch bewijs is het niet meer weg te denken. Hoe zou DNA-bewijs anders gewaardeerd moeten worden? Een groep mensen zou zelfs een verdere integratie van statistiek in de rechtspraak zien door middel van het gebruik van het Bayesiaanse model.

Er zijn echter wel wat problemen waar men mee te maken krijgt bij de vertaling van de ‘normale wereld’ van de leek naar de ‘wereld van getallen’ van de wiskundige en vice versa. Klassieke denkfouten en verkeerde interpretatie, kunnen hierdoor een juiste gang van zaken verstoren. Het negeren van onzekerheden waar mee rekening gehouden zou moeten worden (zoals bij waardering van DNA-bewijs), kan een statistische conclusie een onschendbaarheid geven die in de situatie niet gepast is.

Desalniettemin geeft de toepassing van statistiek de kans om dit soort onzekerheden, die inherent zijn aan het vakgebied van recht, in te schatten en daardoor de invloed ervan juist inzichtelijker te maken.

Het is belangrijk om meer onderzoek te doen in de forensische statistiek, zoals naar de kans op fout-positieve conclusies in DNA-bewijsvoering. De techniek geeft ons de mogelijkheid om vele zaken tot in groot detail vast te kunnen stellen, zoals een DNA-profiel. Maar ik denk dat het gevaarlijk is om vanuit dit gedachtegoed verder te redeneren dat verklaringen betreft bijvoorbeeld een zeer lage DNA-matchings kans zekerheid verschaft. Gelukkig is zich hierover zich een bewustzijn aan het ontwikkelen en buigt zich men over de kwestie hoe DNA-bewijs te presenteren.

Wat dat betreft ben ik het zeer eens met Meester dat er niet al te lichtzinnig met getallen omgegaan moet worden, aangezien de conclusies worden meegenomen in een zeer ingrijpende beslissing aangaande schuld of onschuld van iemand. Naar mijn inzien zal kansanalyse dan ook alleen als ondersteunend bewijs mogen dienen, daar in die zaken dat ander bewijs al aanwezig is. Uitleg van Van Zwet heeft immers ook al laten zien, dat waar overig bewijs van het veronderstelde (moord) mist, er eigenlijk geen zinnige kansanalyse gedaan kan worden. De manier waarop statistiek een rol heeft gehad in de veroordeling van Sally Clark laat in mijn ogen zien dat deze gemaande voorzichtigheid niet (altijd) overbodig is.

Dus terugkomend op de ondertitel: prima combinatie, alleen riskant bij onzorgvuldigheid in uitvoering of in presentatie van feiten.

Tot slot wil ik nog even kwijt dat ik blij ben dat we hier in Nederland niet werken met een rechter/jury stelsel. Ervan uitgaande dat de grote ervaring van de rechter hem beter in staat stelt om te gaan met verklaringen in getallen van deskundigen dan de gemiddelde persoon. Alhoewel het misschien een idee is voor rechters om het boek “How to know when numbers deceive you” van Gerd Gigerenzer te lezen.

Literatuurlijst

- [1] Elffers H., “*Geacht hof, het was geen toeval. De rest is aan u*”, Stator, jaargang 5, nr. 2, blz. 4-11, juni 2004.
- [2] Meester R., *Voorzichtig met statistiek in de rechtzaal!*, Stator, jaargang 5, nr. 2, blz. 12-15, juni 2004.
- [3] Sjerps M., *Statistiek als bewijsmiddel in het strafrecht*, Stator, jaargang 5, nr. 2, blz. 16-20, juni 2004.
- [4] Van Zwet W., *Niets nieuws onder de zon*, Stator, jaargang 5, nr. 2, blz. 21-29, juni 2004.
- [5] Sjerps M., *Forensische statistiek*, Nieuw archief voor wiskunde, vijfde serie, deel 5, nummer 2, blz. 106-111, juni 2004.
- [6] Bijnen E.J., *Bayesiaanse statistiek*, 1976.
- [7] Thompson W.C., Taroni F., Aitken C.G.G., *How probability of a false positive affects the value of DNA evidence*, *Forensic Science*, Vol. 48, No.1, januari 2003.
- [8] Brenner C.H., *Forensic mathematics of DNA matching*, <http://dna-view.com/profile.htm>
- [9] *Kramers Handwoordenboek*, 1996.
- [10] Merckelbach H., *Rekenen met Gigerenzer (boekbespreking)*, *Skepter* 15 (4), december 2002.
- [11] Thomas R., *Seeking truth with statistics*, <http://plus.maths.org>
- [12] Kugh, H.E., *Statistics the essentials for research*, 1970.
- [13] Joyce H., *Beyond reasonable doubt*, <http://plus.maths.org>

Bijlage A: Verslag van de gedane uitspraak m.b.t. de rechtszaak tegen de verpleegster Lucia de B.

Inhoudsindicatie

De rechtbank achtte vier moorden en drie pogingen tot moord bewezen. De in het Juliana Kinderziekenhuis (JKZ) opgenomen slachtoffers waren jonge kinderen in de leeftijd van 0-6 jaar oud. De slachtoffers in het Rode Kruis Ziekenhuis en het Leyenburg Ziekenhuis waren drie bejaarde patiënten. Het hof acht, vooral op basis van toxicologisch bewijs, bewezen dat L.d.B een bijna 6 maanden oud meisje door middel van een injectie een dodelijke dosis digoxine (een hartversterkend middel) heeft toegediend en heeft gepoogd een 6-jarige jongen van het leven te beroven door hem een overdosis chloralhydraat (een slaapmiddel) toe te dienen. De overige levensdelicten zijn soortgelijk. Het overlijden of het aan een reanimatie ten grondslag liggende incident was steeds plotseling en onverwacht en medisch onverklaarbaar. Alle in aanmerking komende natuurlijke doodsoorzaken konden worden uitgesloten. Het overlijden of de reanimatie heeft steeds plaatsgevonden terwijl L.d.B dienst had. L.d.B. heeft volgens het hof met het gebruik van het woord "compulsie" in haar dagboek aantekeningen gedaan op haar vreemde dwangmatige drang om patiënten om het leven (trachten) te brengen. Er is geen statistisch bewijs in de vorm van toevalsberekeningen gebruikt. Wel is voor het bewijs van belang geacht dat de in het JKZ gepleegde delicten een betrekkelijk korte periode bestrijken en de meeste delicten voornamelijk op een gewone verpleegafdeling hebben plaatsgevonden en hiervoor geen verklaring is gevonden. Door het hof zijn voorts de handelingen van L.d.B en de daaraan gegeven uitleg bij alle levensdelicten als buitengewoon suspect aangemerkt. Het hof acht bewezen dat de overige overlijdensgevallen en reanimaties steeds kunnen worden verklaard door een door L.d.B. veroorzaakt misdrijf. De gedragsdeskundigen van het Pieter Baan Centrum zijn er niet in geslaagd zijn om diepgaand in de persoon van L.d.B. door te dringen. Uit de onderzoeksgegevens blijkt geen verband tussen de bij L.d.B. geconstateerde persoonlijkheidsstoornis en de levensdelicten. Nu L.d.B. geweigerd heeft zich door andere gedragsdeskundigen dan die van het Pieter Baan Centrum te laten onderzoeken wordt L.d.B. voor deze levensdelicten volledig toerekeningsvatbaar geacht. In navolging van de rechtbank en overeenkomstig de eis van het openbaar ministerie legt het hof aan L.d.B. een levenslange gevangenisstraf op. Alleen een zodanige straf kan leiden tot adequate vergelding van de door L.d.B. begane misdrijven, gekenmerkt door een in Nederland tot nu toe voor schier onmogelijk gehouden omvang en uitzonderlijke ernst en tot effening van de schade die L.d.B. door die misdrijven aan de nabestaanden en de samenleving heeft toegebracht. Een levenslange gevangenisstraf biedt, gezien de mogelijkheid dat L.d.B. na eventuele gratie onbehandeld weer in de samenleving terugkeert, onvoldoende waarborgen. Door deze straf te combineren met een TBS met dwangverpleging wil het hof gewaarborgd zien dat het recidiverisico door delictsbespreking en behandeling tot een voor de samenleving aanvaardbaar niveau is teruggebracht. [www.rechtspraak.nl]

Bijlage B: Letter from the President to the Lord Chancellor regarding the use of statistical evidence in court cases.

The Lord Chancellor,
Lord Chancellor's Department,
Selborne House,
54-60 Victoria Street, London SW1E 6QW

23 January 2002

Dear Lord Chancellor,

I am writing to you on behalf of the Royal Statistical Society to express the Society's concern about some aspects of the presentation of statistical evidence in criminal trials.

You will be aware of the considerable public attention aroused by the recent conviction, confirmed on appeal, of Sally Clark for the murder of her two infants. One focus of the public attention was the statistical evidence given by a medical expert witness, who drew on a published study to obtain an estimate of the frequency of sudden infant death syndrome (SIDS, or "cot death") in families having some of the characteristics of the defendant's family. The witness went on to square this estimate to obtain a value of 1 in 73 million for the frequency of two cases of SIDS in such a family. This figure had an immediate and dramatic impact on all media reports of the trial, and it is difficult to believe that it did not also influence jurors.

The calculation leading to 1 in 73 million is invalid. It would only be valid if SIDS cases arose independently within families, an assumption that would need to be justified empirically. Not only was no such empirical justification provided in the case, but there are very strong reasons for supposing that the assumption is false. There may well be unknown genetic or environmental factors that predispose families to SIDS, so that a second case within the family becomes much more likely than would be a case in another, apparently similar, family.

A separate concern is that the characteristics used to classify the Clark family were chosen on the basis of the same data as was used to evaluate the frequency for that classification. This double use of data is well recognised by statisticians as perilous, since it can lead to subtle yet important biases.

For these reasons, the 1 in 73 million figure cannot be regarded as statistically valid. The Court of Appeal recognised flaws in its calculation, but seemed to accept it as establishing "... a very broad point, namely the rarity of double SIDS" [AC judgment, para 138]. However, not only is the error in the 1 in 73 million figure likely to be very large, it is almost certainly in one particular direction - against the defendant. Moreover, following from the 1 in 73 million figure at the original trial, the expert used a figure of about 700,000 UK births per year to conclude that "... by chance that happening will occur every 100 years". This conclusion is fallacious, not only because of the invalidity of the 1 in 73 million figure, but also because the 1 in 73 million figure relates only to families having some characteristics matching that of the defendant. This error seems not to have been recognised by the Appeal Court, who cited it without critical comment [AC judgment para 115].

Leaving aside the matter of validity, figures such as the 1 in 73 million are very easily misinterpreted. Some press reports at the time stated that this was the chance that the deaths of Sally Clark's two children were accidental. This (mis-)interpretation is a serious error of logic known as the Prosecutor's Fallacy¹. The jury needs to weigh up two competing explanations for the babies' deaths: SIDS or murder. The fact that two deaths by SIDS is quite unlikely is, taken alone, of little value. Two deaths by murder may well be even more unlikely. What matters is the relative likelihood of the deaths under each explanation, not just how unlikely they are under one explanation.

The Prosecutor's Fallacy has been well recognised in the context of DNA profile evidence. Its commission at trial has led to successful appeals (R v. Deen, 1993; R v. Doheny/Adams 1996). In the latter judgment, the Court of Appeal put in place guidelines for the presentation of DNA evidence. However, we are concerned that the seriousness of the problem more generally has not been sufficiently recognised. In particular, we are concerned that the Appeal Court did not consider it necessary to examine the expert statistical evidence, but were content with written submissions.

The case of R v. Sally Clark is one example of a medical expert witness making a serious statistical error. Although the Court of Appeal judgment implied a view that the error was unlikely to have had a profound effect on the outcome of the case, it would be better that the error had not occurred at all. Although many scientists have some familiarity with statistical methods, statistics remains a specialised area. The Society urges you to take steps to ensure that statistical evidence is presented only by appropriately qualified statistical experts, as would be the case for any other form of expert evidence.

Without suggesting that there are simple or uniform answers, the Society would be pleased to be involved in further discussions on the use and presentation of statistical evidence in courts, and to give advice on the validation of the expertise of witnesses.

Yours sincerely

Professor Peter Green,
President, Royal Statistical Society.

Cc: David McIntosh, President, Law Society.

¹ Balding DJ & Donnelly P, Criminal Law Review, October 1994
[http://www.rss.org.uk/archive/evidence/sc_letter.html]