

Bayesiaanse statistiek in de rechtspraak



$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

BWI werkstuk

*Jenneke IJzerman
September 2004*

Vrije Universiteit Amsterdam

Voorwoord

Het werkstuk dat voor u ligt is het uiteindelijke verslag van een veldonderzoek. Deze zoektocht is een belangrijk en verplicht onderdeel van de studie Bedrijfswiskunde en informatica. Deze opleiding combineert de disciplines van economie, wiskunde en informatica.

Het onderwerp van het werkstuk moet gerelateerd zijn aan tenminste twee van de drie disciplines, en het is de bedoeling de bestaande literatuur daarover te onderzoeken. Dit gebeurt aan de hand van een vraagstelling. De student mag zelf een onderwerp kiezen, maar kan ook een onderwerp aangeboden door de wetenschappelijke staf nemen.

Met dit in het achterhoofd ben ik op zoek gegaan naar een geschikt onderwerp waarbij ik graag een combinatie maak van de theorie en de praktijk. Ik heb hiervoor zelf wat onderwerpen bedacht over nieuwe ontwikkelingen in de statistiek. Na overleg met de heer Jongbloed bleek het onderwerp "Bayesiaanse statistiek in de rechtspraak" een geschikt onderwerp. Dit onderwerp is de laatste tijd wat meer in het nieuws door de rechtszaak van verpleegster Lucy de B.

Hierbij wil ik van de gelegenheid gebruik maken om de heer A. Korzec te bedanken voor het toezenden van zijn proefschrift.

Ook wil ik graag de heer G. Jongbloed bedanken voor het begeleiden van dit werkstuk.

Jenneke IJzerman
2004

Samenvatting

De vraagstelling die ik in dit werkstuk probeer te beantwoorden is:

Hebben Bayesiaanse methoden een toegevoegde waarde in de rechtspraak?

Specifieker gezegd wil ik achterhalen wat de voor- en nadelen van het gebruik van Bayesiaanse statistiek in de rechtspraak zijn.

In de statistiek zijn twee fundamenteel verschillende stromingen, de Bayesiaanse en de klassieke. Het grootste verschil tussen deze twee stromingen, is dat de klassieke statistiek een kans definieert als de frequentie van een bepaalde gebeurtenis bij onafhankelijke herhalingen van een experiment, terwijl de Bayesiaanse een kans interpreteert als de subjectieve graad van geloof in een bepaalde gebeurtenis.

Centraal in de Bayesiaanse statistiek staat de regel van Bayes. De mate waarin wij iets geloven is niet eens en voor altijd gegeven. De mate van geloof wordt beïnvloed door binnenkomende informatie. Dit mechanisme van updaten van waarschijnlijkheden nadat nieuwe informatie is verkregen is de kern van de regel van Bayes.

Twee methodes in de Bayesiaanse statistiek zijn de Markov Chain Monte Carlo methode en de Bayesian belief netwerken. Deze methoden zijn nog in ontwikkeling, maar door de komst van snellere computers worden deze methodes steeds beter toepasbaar in de praktijk

Het vakgebied van statistiek in de rechtspraak heet forensische statistiek. De Bayesiaanse statistiek wordt al jaren toegepast in het DNA bewijs. Een nieuwere ontwikkeling is om bepalen of een verdachte recidief is. Een voorbeeld hiervan is de Bayesian Alcoholism Test, die gebruikt kan worden om aan te tonen of een dronken automobilist een alcoholist is.

Enkele voordelen van Bayesiaanse statistiek in de rechtspraak zijn:

- geeft een numerieke definitie van bewijskracht;
- geeft een definitie voor de taak van de deskundigen en de rechter;
- ondersteunt het combineren van bewijsmiddelen;
- is toepasbaar bij besluiten onder onzekerheid en bij incidentele gevallen.

Echter, behalve voordelen zijn er ook nadelen:

- onopgelost probleem bij het formuleren van bewijzen;
- denkfouten bij het waarderen van bewijs;
- geen threshold waarde aanwezig voor veroordeling;
- geen richtlijnen voor bepalen van subjectieve a-priori kansen.

Bayesiaanse statistiek is geschikter dan de klassieke benadering in de rechtspraak en Bayesiaanse methoden leveren duidelijk een meerwaarde. Maar er zal nog veel onderzoek verricht moeten worden voordat Bayes werkelijk gemeengoed in de Nederlandse rechtspraak is.

Inhoudsopgave

VOORWOORD	3
SAMENVATTING	5
INHOUDSOPGAVE	7
INLEIDING	9
WAT IS RECHTSPRAAK	9
STRUCTUUR VAN HET VERSLAG	9
1 BAYESIAANSE STATISTIEK	11
1.1 HISTORIE	11
1.2 WAT IS BAYESIAANSE STATISTIEK?	11
1.3 BAYESIAANSE VERSUS KLASSIEKE STATISTIEK	12
1.3.1 <i>Betekenis van kans</i>	12
1.3.2 <i>Conditionele kansen</i>	12
1.3.3 <i>Confidence interval versus credible interval</i>	13
1.3.4 <i>p-waarde</i>	13
2 BAYESIAANSE METHODEN	15
2.1 BAYESIAANSE GEVOLGTREKKING	15
2.2 MARKOV CHAIN MONTE CARLO	16
2.3 MACHINE LEARNING	17
2.3.1 <i>Bayesian Belief Network</i>	17
3 FORENSISCHE STATISTIEK	19
3.1 GESCHIEDENIS STATISTIEK IN RECHTSPRAAK	19
3.2 BAYESIAANS MODEL	20
3.3 VOORBEELDEN	20
4 DNA	23
4.1 GEBRUIK BAYES BIJ DNA	23
4.2 DISCUSSIE	23
4.2.1 <i>Formulieren van hypothesen</i>	23
4.2.2 <i>DNA databank</i>	25
4.2.3 <i>Denkfouten</i>	25
4.2.4 <i>Vonnis</i>	27
4.3 CONCLUSIE	27
5 ALCOHOL	29
5.1 GEBRUIK BAYES IN DIAGNOSE	29
5.2 BAYESIAN ALCOHOLISM TEST	31
5.3 DISCUSSIE	34
5.3.1 <i>Ontkenning</i>	34
5.3.2 <i>Systeemkeuze</i>	34
5.3.3 <i>Bronnen</i>	34
5.3.4 <i>Validatie</i>	35
5.3.5 <i>Mannen</i>	36
5.4 CONCLUSIE	37
6 AFSLUITING	39
6.1 STERKE PUNTEN	39
6.2 ZWAKKE PUNTEN	39
6.3 SLOTCONCLUSIE	40
APPENDIX A REKENREGELS KANSREKENING	41
AXIOMA'S KOLMOGOROV	41
CENTRALE LIMIETSTELLING	41
WET VAN GROTE AANTALLEN	42

APPENDIX B BEWIJZEN	43
PREVALENTIE	43
POSITIEVE PREDICTIEVE WAARDE	44
APPENDIX C CAGE VRAGEN	45
GEBRUIKTE AFKORTINGEN	47
LITERATUURLIJST	49
REFERENTIES.....	49
AANVULLENDE LITERATUUR.....	51

Inleiding

...Er waren eens een Bayesiaan en een frequentist die ter dood veroordeeld waren. Toen de rechter hen vroeg wat hun laatste wensen waren, antwoordde de Bayesiaan dat hij graag nog eenmaal de frequentist een les wilde geven. De rechter willigde zijn verzoek in en wende zich vervolgens tot de frequentist om te vragen wat zijn laatste wens was. Waarop de frequentist vlot antwoordde dat hij de les nogmaals wilde horen en nogmaals en nogmaals en nogmaals...

Deze grap geeft een inleiding op de fundamentele verschillen tussen de Bayesiaanse en klassieke benadering in de statistiek. De Bayesiaanse statistiek is nog steeds een ondergeschoven kindje. Tijdens mijn onderzoek heb ik een beter begrip gekregen van deze stroming. Ik hoop daarom dat dit werkstuk voor de lezer ook opheldering zal geven over de Bayesiaanse stroming en de toepasbaarheid ervan binnen het rechtssysteem. In dit werkstuk zal geprobeerd worden een antwoord te geven op de volgende onderzoeksvraag:

Hebben Bayesiaanse methoden een toegevoegde waarde in de rechtspraak?

Aan de hand van een diepgaand literatuuronderzoek zullen conclusies getrokken worden over hoe de Bayesiaanse statistiek kan worden toegepast binnen het rechtssysteem. Hierin zullen Bayesiaanse methoden en enkele toepassingen bekeken worden.

Wat is rechtspraak

Een persoon die ervan wordt verdacht een overtreding of een misdrijf te hebben gepleegd, krijgt te maken met het strafrecht. In een strafzaak staan de verdachte en het Openbaar Ministerie (OM) tegenover elkaar. Het OM treedt hierbij op namens de samenleving.

Alle strafbare feiten staan in wetten. Voorbeelden zijn het Wetboek van Strafrecht, de Wet Economische Delicten, de Opiumwet en de Wegenverkeerswet.

Om een strafzaak voor de rechter te brengen is informatie nodig. In het strafrecht heeft de politie als taak deze informatie te verzamelen. Dit levert een proces-verbaal op.

De officier van justitie leest het proces-verbaal en besluit dan of de zaak voor de rechter komt of niet. In het strafrecht spelen diverse deelnemers een rol, zoals bijvoorbeeld de verdachte, de aanklager en de deskundigen. Tijdens de zitting behandelt de rechter de zaak. Hij beslist of de verdachte een straf krijgt opgelegd. Als de verdachte of de officier van justitie het oneens zijn met de uitspraak, is hoger beroep mogelijk.

Structuur van het verslag

Het eerste hoofdstuk gaat over de Bayesiaanse statistiek. Hoe deze stroming ontstaan is en wat de verschillen zijn met de klassieke stroming. In hoofdstuk 2 staat een korte uitleg over veel gebruikte Bayesiaanse methoden. Om vervolgens in hoofdstuk 3 over te gaan op een inleidend stuk over het gebruik van statistiek in de rechtsspraak, genaamd forensische statistiek. Hierbij heb ik gebruik gemaakt van artikelen van Sjerps (2004). Daarna is er een hoofdstuk (4) gewijd aan het gebruik van de Bayesiaanse statistiek op het gebied van DNA bewijsvoering. In hoofdstuk 5 ga ik in op een praktijk onderzoek naar een nieuwe test die een gerechtelijke toepassing kan hebben. Deze test kan bepalen of een dronken automobilist recidief is. De conclusies zijn beschreven in hoofdstuk 6.

De appendices bevatten enkele rekenregels uit de kansrekening en een uitleg over de CAGE test. Verder bevat dit werkstuk ook nog een lijst met gebruikte afkortingen en een literatuurlijst.

1 Bayesiaanse statistiek

Er zijn twee soorten statistiek, en deze twee verschillen fundamenteel. De ene vorm, de klassieke, is ontwikkeld tussen 1930 en 1960, met als grote voorman Ronald Fisher. Deze theorie vult vrijwel alle leerboeken en ligt ten grondslag aan 80% van de wetenschappelijke publicaties over statistiek. De andere statistiek heet de Bayesiaanse statistiek.

1.1 Historie

De Engelsman Thomas Bayes (1702-1761) was een presbyteriaans¹ predikant en wiskundige. Hij bestudeerde de waarschijnlijkheidsrekening. De Bayesiaanse statistiek is gebaseerd op zijn postuum gepubliceerde essay "Essay towards solving a problem in the Doctrine of chances"(1763). Bayes leerde hoe je over kansen kunt leren uit data. In het artikel beschrijft Bayes (een gedeelte van) een stelling die later de regel van Bayes genoemd zal worden.



Laplace herontdekte de stelling veertien jaar later en gebruikte het om problemen op te lossen, bijvoorbeeld om het gewicht van Saturnus te bepalen.

Pas na 1950 werd de regel belangrijker door onder andere het werk van Leonard Savage(1954).

1.2 Wat is Bayesiaanse statistiek?

In de Bayesiaanse statistiek draait het om het filosofische principe, dat de wiskundige theorie van waarschijnlijkheid op de subjectieve **graad van geloof** die men hecht aan een gebeurtenis toepast.

De regel van Bayes (soms ook wel de omkeerformule genoemd, omdat het de "omgekeerde" voorwaardelijke kans berekent.) is

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)} \quad \text{als } P(A) > 0, P(B) > 0,$$

oftewel, "de kans op A, als B gegeven is, is gelijk aan de kans op B, als A gegeven is maal de kans op A, gedeeld door de kans op B".

Er zijn enkele termen die gewoonlijk gebruikt worden door Bayesiaanse statistici. Eén daarvan is de **a-priori** waarschijnlijkheid, $P(A)$. Dit is de kans dat een gebeurtenis is opgetreden voordat er ook maar enige data geobserveerd is. Een tweede term is de **a-posteriori** waarschijnlijkheid, $P(A|B)$. Deze refereert naar de kans dat een gebeurtenis is opgetreden nadat met geobserveerde data rekening is gehouden. Een derde erg gebruikelijke term is de **likelijkheid**, $P(B|A)$, deze wordt gebruikt om de conditionele kans te beschrijven van de data gegeven een bepaald model.

De regel van Bayes wordt daarom ook wel geschreven als:

$$a \text{ posteriori} \propto a \text{ priori} * \text{likelihood}$$

Waarbij \propto betekent 'is evenredig aan'².

¹ een presbyteriaan is een lid van de Schots/Britse Presbyterian Church. Deze kerk is de evenknie van de Nederlandse Gereformeerde Kerk.

² Waar is de kans $P(B)$ gebleven? De kans $P(B)$ kan weggelaten worden omdat $P(B) = \int P(A)P(B | A) dA$.

Oftewel, $P(B)$ kan bepaald worden uit de de a-priori verdeling en de likelijkheid en daardoor kan deze kans door schaling weggelaten worden.

Bij Bayesiaanse kansen wordt uitgegaan van reeds bekende kansen, a-priori kansen, op basis van eerder onderzoek. Als de waarde hiervan niet bekend is, kan hiervoor ook het oordeel van een expert (ervaringsdeskundige) worden gevraagd. Hij zal dan een bepaalde waarschijnlijkheid toekennen aan gebeurtenissen, bijvoorbeeld door te schatten dat een voorval voor 0.7 (70%) plausibel wordt geacht.

1.3 Bayesiaanse versus Klassieke statistiek

Een van de grootste fundamentele tegenstellingen in de moderne wetenschap is die tussen de klassieke en Bayesiaanse statistiek. Deze controversie is van filosofische aard.

Of men nu in het klassieke dan wel in het Bayesiaans kader werkt, de rekenregels voor kansen zijn dezelfde. De regel van Bayes volgt meteen uit de definitie van de conditionele kans.³

Tussen de klassieke en Bayesiaanse statistiek zijn de volgende vier verschillen aan te merken (Bullard, 2001):

1.3.1 Betekenis van kans

In de klassieke, frequentistische, statistiek wordt er van uitgegaan dat parameters (zoals bijvoorbeeld de kans op een bepaalde ziekte) door de natuur gegeven constanten zijn, die we zo goed mogelijk proberen te schatten met behulp van steekproeven (bijvoorbeeld in een klinische studie). Een kans is dus de waarschijnlijkheid op een bepaalde gebeurtenis. Deze kans wordt bepaald aan de hand van een voldoende aantal keer uitvoeren van een experiment, en het **frequentie quotiënt** van een bepaalde gebeurtenis in de verzameling wordt dan bepaald. Hoe meer experimenten worden uitgevoerd, hoe dichter deze kans komt bij de theoretische kans. Dus door het herhalen van een willekeurige trekking kunnen we de verdeling in de onderliggende grote populatie kennen.

Een andere manier om kans te interpreteren is als **de subjectieve graad van geloof** die men hecht aan de gebeurtenis dat de uitkomst in een bepaalde verzameling valt. Die benadering wordt Bayesiaans genoemd. Deze kans is op veel meer gebeurtenissen van toepassing dan de klassieke waarschijnlijkheid toestaat: de niet herhaalbare gebeurtenissen, en alles waar men onzeker over is. De subjectieve benadering van kans laat het toe dat verschillende mensen een verschillende schatting van de kans hebben omdat zij verschillende informatie, of verschillende ervaringen kunnen hebben.

Een klassieke statisticus zou aan de gebeurtenis van het gooien van kop bij het gooien van een muntje een kans 1/2 kunnen geven. Dit kan hij alleen doen als hij weet dat dat de relatieve frequentie is. Een Bayesiaan zou een kans 1/2 (of een ander cijfer) aan een persoonlijk geloof in de hypothese dat er een miljard jaar geleden leven op Mars was kunnen geven, zonder ook maar iets over de relatieve frequentie te beweren.

1.3.2 Conditionele kansen

De discussie ontstaat nu als je in de regel van Bayes A ziet als “waarheid” en B als “data”. De klassieke kant gaat uit van het bestaan van een vaste, zij het onbekende, waarheid die bepaald kan worden uit waarnemingen met een toevalskarakter. In de klassieke gevolgtrekking worden de significantie testen uitgevoerd in de veronderstelling dat een hypothese (de nulhypothese) waar is. De klassieke onderzoeken dus de waarschijnlijkheid van de data gegeven een model, $P(\text{data} \mid \text{model}=\text{hypothese})$.

³ De definitie van de conditionele kans zegt het volgende; $P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A)$.

De Bayesiaan hangt het kansbegrip op aan het feit dat de mogelijke waarheden onzeker zijn. En hij zal de waarschijnlijkheid van een model gegeven de gegevens willen onderzoeken, $P(\text{model}=\text{waarheid} \mid \text{data})$.

Het probleem is nu dat je aan de waarheid een kans geeft voordat je de data gezien hebt. Dit is in feite het kern principe in de Bayesiaanse statistiek; alleen wat *werkelijk* geobserveerd wordt (én natuurlijk de prior) is relevant voor het bepalen van de kans dat een bepaald model waar is. Waardes die meer extreem zijn dan de geobserveerde waarde zijn bijvoorbeeld irrelevant.

In de Bayesiaanse statistiek wordt aan de parameter wel een kansverdeling toegeschreven. Klassieke statistici hebben dus problemen met het idee dat een a-priori kans verklaring voor “waarheid” nodig is. Vanuit hun oogpunt is “waarheid” niet stochastisch.

In sommige statistische problemen is de waarde van de a-priori kans nagenoeg niet van belang, omdat deze kans weinig informatie geeft. Het is ook mogelijk dat de likelihood erg informatief is omdat er veel data aanwezig is. In deze gevallen komen de resultaten van de Bayesiaanse en klassieke statistiek dicht bij elkaar.

1.3.3 Confidence interval versus credible interval

In de statistiek wordt een intervallschatting aan de hand van steekproefgegevens gebruikt om een interval van mogelijke waarden van een bepaalde modelparameter te berekenen.

Voorbeelden van modelparameters zijn bijvoorbeeld gemiddelde, mediaan en variantie.

In de klassieke statistiek heet dit interval een confidence interval (betrouwbaarheidsinterval).

De verkregen puntschatting uit de steekproef is zelden of nooit precies gelijk aan de

modelparameter die hij schat. Een betrouwbaarheidsinterval geeft dus een

onzekerheidsmarge rond de steekproefuitkomst. De onder- en bovengrenzen van dit interval

zijn de betrouwbaarheidsgrenzen. De waarschijnlijkheidsverklaring die genoemd wordt bij

een betrouwbaarheidsinterval verwijst naar de willekeur van de herhaaldelijk uitgevoerde

steekproeven. De experimenten produceren een interval die de parameter 95% van de tijd

bevat (of 90%, enz.). Daarom hoort bij een 95% betrouwbaarheidsinterval *niet* de

interpretatie dat de parameter met 95% kans binnen dat interval ligt. Dit omdat de parameter

constant is en dus wel of niet in het interval ligt. Zodra de steekproef is genomen en een

bepaald betrouwbaarheidsinterval is geconstrueerd, is er geen willekeur meer, en dus ook

geen waarschijnlijkheid.

In de Bayesiaanse statistiek kan een interval ook worden gebruikt om een modelparameter

te schatten. Op basis van de a-posteriori verdeling kan een Bayesiaans

betrouwbaarheidsinterval gegeven worden dat de intuïtieve interpretatie heeft dat de kans

dat de parameter tussen twee waarden ligt, 95% is. Dit interval wordt een “credible interval”

(geloofwaardigheidsinterval of waarschijnlijkheidsinterval) genoemd. De

waarschijnlijkheidsverklaring bij een credible interval is dus de waarschijnlijkheid dat de

parameter in het interval ligt.

1.3.4 p-waarde

De p-waarde of overschrijdingskans (van een gegeven steekproefuitkomst) is de kans dat onder de verdeling, gegeven door de nulhypothese, de waarde van de toetsingsgrootheid in

een statistische toets wordt overschreden. Dit is de klassieke interpretatie. De p-waarde

geeft aan hoe extreem de gevonden waarde voor de toetsingsgrootheid in de verdeling

onder de nulhypothese is. Hoe kleiner de p-waarde, hoe extremer de uitkomst.

Wanneer twee hypothesen vergeleken worden, resulteren de klassieke methodes in het

verwerpen of niet-verwerpen van de nulhypothese met een bepaalde graad van vertrouwen

(meestal 95%).

De Bayesiaanse methodes geven als resultaat een uitkomst dat een bepaalde hypothese waarschijnlijker is dan de andere. Hierbij wordt geen gebruik gemaakt van niet geobserveerde (extreme) data.

KLASSIEK	BAYESIAANS
Waarschijnlijkheid wordt geïnterpreteerd als de relatieve frequentie van een bepaalde gebeurtenis over een lange reeks herhaalde vergelijkbare experimenten. De kans ligt objectief in de wereld, niet bij de waarnemer.	Waarschijnlijkheid is geïnterpreteerd als een meting van iemands geloof van (on)zekerheid over een gebeurtenis. De kans ligt in de geest van de waarnemer en kan verschillend zijn tussen mensen die verschillend informatie of verschillende ervaringen uit het verleden hebben.
Gevolgtrekking is uitgevoerd door het evalueren van de kans van de geobserveerde data, of de uiterste waarden van data, gegeven een hypothese model.	Gevolgtrekking wordt uitgevoerd door het evalueren van de kans van een hypothese model gegeven geobserveerde data.
Een 95% betrouwbaarheidsinterval is het resultaat van een procedure die 95% kans had voor het genereren van een interval die de geschatte parameter bevat.	De kans is 95% dat de geschatte parameter in het 95% geloofwaardigheids(credible) interval ligt.
De P-waarde in een significantie test is de kans voor het krijgen van een resultaat tenminste zo extreem als degene die geobserveerd is, gegeven dat de nul hypothese waar is.	Men kan de kans evalueren van een bepaald model of set van modellen gegeven geobserveerde data. Niet geobserveerde data (meer extreme waardes) zijn niet relevant.

Tabel 1: Verschillen Bayesiaanse en klassieke statistiek

2 Bayesiaanse methoden

De Bayesiaanse methoden zijn in de afgelopen tien jaar in populariteit gegroeid door de vondst van slimme methoden om de a-posteriori verdeling te simuleren. Methoden als MCMC en Bayesian Belief networks worden gebruikt om de Bayesiaanse gevolgtrekking te ondersteunen.

Ook heeft de komst van snellere computers een belangrijke rol in de ontwikkeling van de methoden gespeeld. Dit omdat de Bayesiaanse aanpak nogal rekenintensief kan zijn.

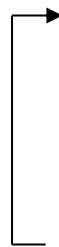
Een probleem met deze methoden is dat er nog weinig standaard software beschikbaar is voor complexe modellen.⁴

2.1 Bayesiaanse gevolgtrekking

Het herhaaldelijk uitvoeren van een model-opzet (bepalen van hypothesen en verzamelen van data), de berekening met Bayes, het bepalen van karakteristieken en een conclusie trekken over het model wordt Bayesiaanse gevolgtrekking genoemd.

De gebruikte model parameters in de Bayesiaanse statistiek zijn schattingen. Maar over de parameters kunnen wel probabilistische verklaringen gegeven worden aan de hand van hun karakteristieken. Voorbeelden van karakteristieken zijn gemiddelde, variantie, mediaan en waarschijnlijkheidsinterval.

Nog specifiekier gezien bevat de Bayesiaanse gevolgtrekking de volgende zich herhalende stappen (Traat, 2002):

- 
- De informatie gegeven door de data wordt gecombineerd met de a-priori kennis van de parameter(s) (de a-priori verdeling). Door middel van de regel van Bayes wordt de a-posteriori verdeling van de parameters verkregen.
 - De informatie van de parameters wordt gebaseerd op de verschillende karakteristieken van de a-posteriori verdeling.
 - Als er nieuwe data beschikbaar komen, kan de a-posteriori verdeling als a-priori verdeling genomen worden en een nieuwe a-posteriori verdeling bepaald worden. (output wordt nu input)

In de praktijk is het vaak niet mogelijk om een a-posteriori verdeling op analytische wijze te vinden. Een manier om de Bayesiaanse gevolgtrekking te benaderen is op stochastische wijze:

De exacte verdeling van de a-posteriori is onbekend, maar er kunnen wel waardes uit gesimuleerd worden. Dit door bijvoorbeeld de empirische verdeling met gesimuleerde waardes te nemen als schatting voor de a-posteriori verdeling. De karakteristieken van de empirische verdeling benaderen dan de karakteristieken van de a-posteriori verdeling. Een voorbeeld van zo'n stochastische methode is de Markov Chain Monte Carlo (MCMC) methode.

De stochastische methoden worden ondersteund door de volgende twee resultaten uit de waarschijnlijkheidstheorie (zie appendix B):

- De wet van de grote aantallen: hoe meer gesimuleerde waardes, hoe beter de benadering.
- Centrale Limietstelling: geeft een maat voor de benaderingsfout.

⁴ Computerprogramma's als Splus, SPSS en R (gratis te downloaden van www.R-project.org) kunnen gebruikt worden om statistische analyses uit te voeren. BUGS (Bayesian inference Using Gibbs Sampling) is een voor algemeen doel geschikt systeem voor statistische analyse in grafische modellen. www.mrc-bsu.cam.ac.uk/bugs

2.2 Markov Chain Monte Carlo

Bootstrappen is het uit een standaard kansverdeling pseudo random getallen simuleren. De basis stap is om getallen van de uniforme verdeling $U(0,1)$ te genereren. De uitkomsten van andere verdelingen worden vervolgens hieruit verkregen door geschikte transformaties/manipulaties uit te voeren.

Zodra deze onafhankelijke replicaties uit een gegeven verdeling gebruikt worden om een probleem op te lossen heet het de Monte Carlo methode. Het is interessant om te vermelden dat Monte Carlo een puur frequentistische methode is (O'Hagan, 1987, MacKay, 1999).

Historisch gezien werd de naam Monte Carlo gebruikt als codenaam tijdens WOII voor top-secret berekeningen die werden uitgevoerd om de stroom van neutronen in een atoombom te voorspellen. Aangezien de wegen van de neutronen varieerden en de bouw van de atoombom een grote gok was, werden de berekeningen de codenaam Monte Carlo gegeven omdat in de gokstad Monte Carlo een simpele random nummer generator genaamd roulette stond.

(Bron: Sabri Pillana, <http://www.geocities.com/CollegePark/Quad/2435/history.html>)

De truc is dat de ingewikkelde afleiding van de a-posteriori verdeling vervangen wordt door simulatie van deze verdeling, en dit met behulp van de theorie van de Markovketens.⁵

De theorie van de Markovketens is ontwikkeld door Andrei A. Markov(1906, 1971) aan het begin van de twintigste eeuw. Hij bestudeerde hoe letters in woorden verschenen. Gebruik maken van het gedicht 'Jevgeni Onegin' van Poeshkin ontwikkelde hij een waarschijnlijkheidsmodel dat het volgende element (klinker of medeklinker) alleen afhangt van zijn onmiddellijke voorganger.

Een markovketen convergeert in veel gevallen naar een stationaire verdeling die een equilibrium (evenwichts)verdeling genoemd wordt.

De MCMC techniek houdt in dat de verdeling waarin we geïnteresseerd zijn (de a-posteriori verdeling van de modelparameters) wordt gepresenteerd als een unieke stationaire verdeling van een Markovketen. Het simuleren van de verdeling houdt in dat we een Markov keten simuleren totdat een evenwicht bereikt is en vervolgens komen de gesimuleerde waarden van onze verdeling. En de gevolgtrekking wordt vervolgens gebaseerd op de gesimuleerde waarden.

De meeste Bayesiaanse analyses maken gebruik van één van de twee volgende MCMC algoritmen:

- Metropolis-Hastings algoritme

De afhankelijkheid van de huidige waarde in Markovketen op de vorige waarde wordt bepaald door twee stappen

- Proposal: stel een bepaalde nieuwe waarde voor

- Acceptance: de voorgestelde waarde wordt alleen goedgekeurd onder een bepaalde voorwaarde, die voldoet aan de eis naar de stationaire verdeling te gaan.

Het artikel van Metropolis e.a.(1953) introduceerde deze methode.

- Gibbs sampling (Geman en Geman, 1984)

Is een componentsgewijze Metropolis Hastings algoritme. De afhankelijkheid van de huidige waarde in de Markovketen op de vorige waarde wordt bepaald door de voorwaardelijke verdelingen.

⁵ Een Markovketen is een random proces waar iedere waarde in het traject alleen afhankelijk is van zijn direct voorgaande waarde

De MCMC techniek vergemakkelijkt de Bayesiaanse gevolgtrekking door een middel te verstrekken om een reeks willekeurige steekproeven van de a-posteriori verdeling te produceren. Hoewel ze niet gebonden zijn aan een Bayesiaanse context, kunnen MCMC-methoden, in sommige situaties waar een klassieke benadering problemen oplevert, uitkomst bieden.

2.3 Machine learning

Machine learning is een gebied in de kunstmatige intelligentie waarin technieken ontwikkeld worden om computers te laten "leren". Leren wordt hierbij gedefinieerd als het automatisch verbeteren door ervaringen. Specifieker gezegd is machine learning een methode om computerprogramma's te maken door de analyse van data.

Het Bayesiaanse perspectief op machine learning zorgt voor een probabilistisch verband tussen oorzaak en gevolg.

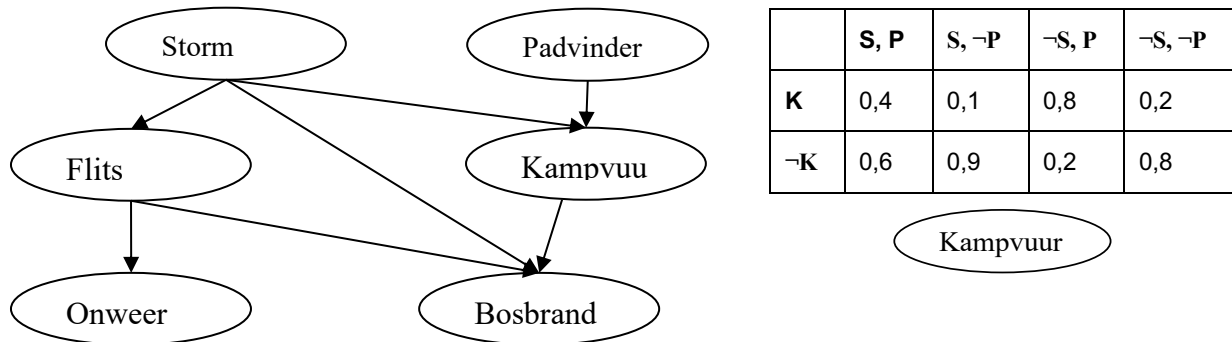
Bayesiaanse methoden hebben in het vakgebied van de machine learning twee rollen (Mitchell, 1997):

- ze zorgen voor praktische leeralgoritmes: zoals Bayesian Belief Networks
- ze zorgen voor een bruikbaar conceptueel framework

2.3.1 Bayesian Belief Network

Een Bayesian Belief network, ook wel Bayes Net genoemd, is een gerichte acyclische graaf van knopen die variabelen vertegenwoordigen en pijlen die afhankelijkheids relaties tussen de variabelen vertegenwoordigen (Heckerman, 1996). Als er een pijl van knoop A naar een andere knoop B is, dan zeggen we dat A een ouder van B is. Een knoop kan ieder soort variabele vertegenwoordigen: een waargenomen meting, een parameter, een latente variabele, of een hypothese.

Onderstaand figuur geeft de grafische structuur van een Belief netwerk weer.



Figuur 1: Simpel Belief netwerk met ernaast de kanstabel behorende bij de knoop Kampvuur. Kampvuur is afgekort tot K, Storm tot S en Padvinders tot P.

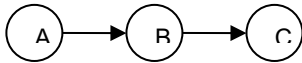
Een Bayesian belief netwerk beschrijft de gezamenlijke kansverdeling (joint probability distribution) over een set van variabelen. In een Belief netwerk wordt de kennis over relaties tussen variabelen op een probabilistische wijze gerepresenteerd. De grafische structuur van het netwerk geeft de rechtstreekse verbanden tussen de variabelen aan. De kanstabellen kwantificeren deze verbanden. En met de regels van de kansrekening kan met deze kennis worden gemanipuleerd en geredeneerd.

Dit type netwerk is dus handig als er een conditionele afhankelijkheid tussen subsets van variabelen is. Het netwerk laat het combineren van a-priori kennis over (on)afhankelijkheden

tussen variabelen met geobserveerde trainingsdata toe. Met behulp van een Belief netwerk kan op vele manieren geredeneerd worden, ongeacht de richting van de pijlen. Een van de belangrijkste toepassingen is het maken van een diagnose.

Er geldt dus dat bij iedere knoop een kanstabel hoort. In deze kanstabel wordt de conditionele kans op de knoop gegeven zijn directe ouders gespecificeerd. Hierbij wordt gebruik gemaakt van de aanname dat een knoop alleen afhankelijk is van zijn directe ouders. En dus onafhankelijk is van al zijn andere voorouders, gegeven de kennis van zijn ouders.

Neem bijvoorbeeld als voorbeeld het onderstaande zeer simpele netwerk met de variabelen A, B en C.



De aanname specificereert dan dat $P(C | A, B) = P(C | B)$.

Om de kanstabellen samen te stellen zijn verschillende aanpakken mogelijk:

- aan de hand van expertkennis wordt de tabel handmatig gevuld
- aan de hand van data wordt de tabel 'geleerd'. Dit komt neer op het bepalen van frequenties in de data.
- een combinatie van beide methoden.

Welke aanpak gekozen wordt hangt af in hoeverre er expertkennis en data beschikbaar is.

Het Bayesian Belief netwerk is een actief onderzoeksgebied, er wordt bijvoorbeeld gekeken naar effectievere gevolgtrekkingsmethoden. Monte Carlo methoden voor Bayesiaanse netwerken zijn bijvoorbeeld ontwikkeld door Neal (1996).

3 Forensische statistiek

Forensische wetenschappen krijgen de laatste tijd veel aandacht in de media zoals bij de zaak Lucy de B.⁶ Verder laten populaire programma's zoals 'Peter R. de Vries, misdaadverslaggever' (SBS6), 'NET5 Detectives' (NET5) en 'Crime scene investigation' (RTL4) zien wat forensisch onderzoek kan betekenen bij de bewijsvoering in een strafzaak. De term forensische statistiek zegt eigenlijk al wat het inhoudt. De exacte definitie van forensisch is volgens de Van Dale:

fo·ren·sisch (bn.)

1 [jur.] betr. hebbend op justitie

Statistiek en kansrekening zijn vaak een onderdeel van de verklaringen van (wetenschappelijke) deskundigen, en de forensische statistiek is een groeiend vakgebied. Het gaat hierbij om nogal uiteenlopende deskundigheidsgebieden, zoals bijvoorbeeld pathologie, biologie, fysica, chemie en psychologie. Maar bij de toepassing in strafrecht hebben deze gebieden een aantal aspecten gemeen. Er is meestal een verdachte aanwezig en een spoor van het plaats delict wordt vergeleken met een bepaald kenmerk van de verdachte.

Het gaat hierbij om de bewijskracht van de waargenomen verschillen en overeenkomsten.

3.1 Geschiedenis statistiek in rechtspraak

Tot circa twintig jaar geleden was er geen algemene en bruikbare beschrijving van het begrip bewijskracht. Ieder deskundigheidsgebied gebruikte zijn eigen methode om de bewijskracht van de onderzoeksresultaten weer te geven. Door de opkomst van de Bayesiaanse denkrichting in de forensische statistiek werd dit probleem opgelost. De regel van Bayes wordt nu gebruikt om een numerieke uitdrukking voor de bewijskracht van allerlei bewijsmateriaal te geven.

De juridische context veroorzaakt dat een deskundige bij het toetsen van hypothesen op de klassieke manier in een ander type fout geïnteresseerd is dan gebruikelijk. De enige nuttige nulhypothese in een gerechtelijke context is dat het bewijsmateriaal en het referentiemateriaal een gemeenschappelijke bron hebben (Rudin en Inman, 2003). Voor een voorbeeld zie de voetnoot⁷.

Gewoonlijk is een statisticus geïnteresseerd in een fout van de eerste soort. Dit is de kans om de nulhypothese te verwerpen terwijl deze waar is. Deze fout is dus geassocieerd met de kans dat de verdachte ten onrechte wordt vrijgesproken. Voor de deskundige is dit echter niet de belangrijkste fout. Van groter belang is het om de fout van de tweede soort te controleren. Dit is de conclusie dat de nulhypothese niet verworpen moet worden, terwijl dat wel zou moeten. Deze conclusie kan leiden tot een onterechte veroordeling.

De Bayesiaanse statistiek is in dit verband een natuurlijkere keuze. Dit omdat in de regel van Bayes de verhouding van de nulhypothese en alternatieve hypothese meegenomen wordt in de vorm van de likelihood.

⁶ Voor meer informatie over deze zaak verwijst ik naar <http://www.nos.nl/nieuws/achtergronden/luciadeb.html>, Anders rekenen met Lucy de B., Trouw, De verdieping, 13 maart 2004 en Een schijn van kans, NRC, Wetenschap en onderwijs, 13 maart 2004

⁷ Als de vraag is of twee kogels door hetzelfde wapen gevuld zijn, kunnen de volgende hypothesen gebruikt worden. *Nulhypothese: Deze kogels zijn door hetzelfde wapen geschoten.*

Alternatieve hypothese: Deze kogels zijn door verschillende wapens geschoten.

Hierbij kunnen deze kogels twee bewijsmaterialen zijn, of is één kogel bewijsmateriaal en de andere een referentie kogel.

3.2 Bayesiaans model

In de forensische statistiek wordt een model gebruikt dat is gebaseerd op de regel van Bayes voor twee alternatieven.

Dit model kan gebruikt worden om de bewijskracht uit te drukken.

$$\frac{P(H_p | E)}{P(H_d | E)} = \frac{P(H_p)}{P(H_d)} \cdot \frac{P(E | H_p)}{P(E | H_d)}$$

In een rechtszaak dragen zowel de aanklager als de verdediging een hypothese aan. De regel van Bayes wordt dus toegepast op twee hypothesen H_p en H_d en het bewijsmateriaal E .⁸

De kansverhouding aan de linkerkant ontstaat uit de oude kansverhouding door deze te vermenigvuldigen met de likelihood ratio (meest rechter quotiënt).

$$\text{a-posteriori} = \text{a-priori} \cdot \text{likelihood ratio}$$

Bovenstaande formule geeft dus weer dat de kansverhouding van de twee hypothesen verandert door de introductie van het bewijsmateriaal.

Deze likelihood ratio term wordt ook wel de Bayes factor genoemd.

Dit Bayesiaanse model kan op de volgende manier toegepast worden in de rechtszaal. De rechter heeft aan het begin van het proces een a-priori inschatting van de kansverhouding. Vervolgens worden tijdens het proces bewijsmiddelen naar voren gebracht. Hierbij heeft de aanklager een bepaalde hypothese (bijvoorbeeld: de verdachte heeft de bom geplaatst) en de verdediging heeft een alternatieve hypothese (bijvoorbeeld: iemand anders heeft de bom geplaatst). Het is nu de taak van een deskundige om een schatting te geven voor de likelihood ratio (LR) van het bewijsmateriaal voor de hypothesen H_p en H_d .

Hierbij is de LR groter dan 1 als het bewijs in het nadeel van de verdachte is, dus als

$$P(E | H_p) > P(E | H_d),$$

en kleiner dan 1 als het bewijs in voordeel van de verdachte is, met dus

$$P(E | H_p) < P(E | H_d).$$

Bij elk bewijsmiddel past de rechter de waarde van de a-posteriori kansverhouding aan met behulp van de bovenstaande regel van Bayes. Bij het aandragen van een volgend bewijsmiddel zal deze aangepaste a-posteriori kansverhouding de a-priori kansverhouding worden.

Aan het eind van het proces als alle bewijsmiddelen aangedragen zijn, zal de rechter zijn beslissing op basis van de uiteindelijk geschatte a-posteriori kansverhouding baseren. De rechter zal de verdachte veroordelen als de hypothese van de aanklager, gezien al het bewijs, veel waarschijnlijker is dan die van de verdediging. Dus als de a-posteriori kansverhouding tussen deze twee hypothesen groot is. Wat de exacte waarde van deze a-posteriori uitkomst moet zijn om een verdachte te kunnen veroordelen is niet vastgelegd.

3.3 Voorbeelden

De regel van Bayes met de LR, soms ook diagnostische waarde genoemd, wordt succesvol toegepast in gebieden waar veel getallen voor handen zijn. In een groot aantal rechtszaken speelt statistiek een grote maar onopvallende rol. Bij technisch bewijs worden statistiek en kansrekening regelmatig gebruikt om de conclusie te onderbouwen.

⁸ De H staat voor hypothese, p voor prosecution (aanklager), d voor defense (verdediging) en E voor evidence (bewijsmiddel).

De regel van Bayes wordt onder andere toegepast bij DNA onderzoek: de vergelijking van het spoor met een dader of een met een database. Of in DNA verwantschap analyse om bijvoorbeeld vaderschap aan te tonen. Toch wordt DNA onderzoek in een rechtszaak nog niet altijd toegestaan en gewaardeerd. Het probleem is dat de verdachte niet altijd verplicht is om DNA af te staan voor een test.

Tijdens de spraakmakende rechtszaak van Lucy de B. stond statistiek veel in de belangstelling. De wetenschappers waren het totaal oneens over de waarde van het statistiek bewijs. De berekeningen van prof. dr. H. Elffers werden aangevallen door Meester en Lambalgen en vervolgens verscheen ook nog een groot stuk van A. de Vos. De opspraak in deze zaak wordt veroorzaakt door de tegenstelling in de klassieke en Bayesiaanse stroming.

Maar ook bij de diagnose of een verdachte in herhaling kan vallen wordt Bayesiaanse statistiek toegepast. Dit gebeurt onder andere voor automobilisten die in beroep gaan tegen het intrekken van hun rijbewijs. Hierbij wordt Bayes toegepast om te bepalen of iemand alcoholist is of niet.

En er zijn nog veel meer voorbeelden te noemen waar de statistiek een rol speelt in de vergaring of waardering van het strafrechtelijke bewijsmateriaal. Steekproeven in drugszaken, geuridentificatieproef door honden, snelheidsbepaling van auto's, de Oslo confrontatie (line-up van verdachten) en epidemiologische aspecten van de toxicologie (vergiftenleer).

4 DNA

De toepassing van DNA-technieken in het strafprocesrecht stamt uit de jaren tachtig. In 1989 gaf een Nederlandse rechter voor het eerst toestemming voor DNA-onderzoek en wel op uitdrukkelijk verzoek van een verdachte van serieverkrachtingen die hadden plaatsgevonden in de omgeving van het World Trade Centre. Dit illustreert de noodzaak van een wettelijke regeling.⁹

4.1 Gebruik Bayes bij DNA

Wanneer er bloed, speeksel, huid of haar (die allen cellen bevatten) achtergelaten wordt op een crime-scene kan hiervan een DNA profiel gemaakt worden. Dit profiel geeft aan hoe het DNA (een afkorting van het Engelse woord voor **deoxyribosenucleïnezuur**) van een bepaald persoon is opgebouwd. Het DNA profiel kan vergeleken worden met het DNA profiel van een verdachte, of met een database met DNA profielen. In het eerste geval kan DNA bewijs extra bewijs tegen de verdachte opleveren. In het tweede geval kan een match in de database een mogelijke verdachte aanwijzen die tot dan toe nog niet verdacht was.



Figuur 2: DNA keten

Het is nu de taak van de deskundige om een match kans geven. Deze kans is gerelateerd aan de LR: de ratio bewijsmateriaal (bijv. 'celmateriaal en verdachte hebben DNA profiel X') voor de hypothese (bijv. 'de niet-verwante persoon'). De likelihood ratio staat gelijk aan 1/match kans van profiel X. De bewijskracht van een match is groter dan één op een miljoen. Het is duidelijk dat deze kans nog niet zegt of een verdachte al dan niet de eigenaar van het celmateriaal is, oftewel, of de verdachte schuldig is. Een DNA-profielmatch is niet voldoende voor een veroordeling. Er moet ook ander bewijsmateriaal of een bekentenis zijn.

Het is nu aan de rechter om een schatting te doen voor de a-priori kansverhouding om de kans te bepalen dat het celmateriaal van de verdachte is of niet. Hij maakt dus bijvoorbeeld een schatting voor hoeveel potentiële daders er naast de verdachte zijn waarvoor het even waarschijnlijk is dat ze het celmateriaal achter gelaten hebben. Vervolgens kan met beide kansen bepaald worden wat de kans is dat het celmateriaal van de verdachte is.

4.2 Discussie

4.2.1 Formuleren van hypothesen

Een nog niet opgelost probleem in het gebruik van DNA als bewijs betreft het formuleren van de hypothesen. Soms is het rapporteren van een LR nogal problematisch. Dit wordt hieronder uitgelegd aan de hand van een bepaald forensisch probleem, namelijk het "two-stain problem" (Evet, 1987).

⁹ Pas in 1994 trad de (nog) bestaande regeling 'DNA-onderzoek in strafzaken' in werking, die weliswaar duidelijkheid verschafte, maar ook tot veel nieuwe vraagpunten aanleiding gaf. In de afgelopen jaren zijn de technische en juridische mogelijkheden van DNA-onderzoek bij de waarheidsvinding in strafzaken fors uitgebreid. Sinds 1 november 2001 kan in "het belang van het onderzoek", bij delicten waar 4 jaren of meer gevangenisstraf op staat, van verdachten, zonodig onder dwang, celmateriaal worden afgenomen voor DNA-onderzoek.

Beschouw een misdaad die gedaan wordt door twee mensen, en ieder van hen heeft een bloedvlek achtergelaten op de plek van de misdaad. Eén op een kussen (A), een ander op een vel papier (B). De DNA profielen van deze vlekken worden bepaald. Stel dat Jan Smit wordt gearresteerd om hele andere redenen, die niets te maken hebben met bovenstaande misdaad, maar dat blijkt dat het DNA profiel van Smit overeenkomt met A. Hij wordt nu dus als verdachte gezien. Om nu te bekijken hoe sterk het DNA bewijs tegen Smit werkelijk is, moet een likelihood ratio bepaald worden. De expert moet hiervoor twee concurrerende hypothesen opzetten. Hiervoor zijn de volgende sets hypothesen mogelijk:

Hypothese set 1

H_p : de verdachte Smit is donor van 1 van de 2 vlekken

H_d : de verdachte Smit is niet donor van 1 van de 2 vlekken

Deze set van hypothesen wordt vaak gebruikt (Evetts 1987, Aitken 1995, Stoney 1991).

Hypothese set 2

H_p : de verdachte Smit is donor van vlek A

H_d : de verdachte Smit is niet donor van 1 van de 2 vlekken (vlek A en B)

Het verschil tussen de eerste en de tweede set is dat er in de eerste set geen onderscheid gemaakt wordt tussen de vlekken op het kussen en het vel papier. Terwijl in de tweede set het onderscheid bewust gemaakt is.

Hypothese set 3

H_p : de verdachte Smit is donor van vlek A

H_d : de verdachte Smit is niet donor van vlek A

Deze hypothesen negeren compleet de vlek op het vel papier. Gegeven het feit dat Smit matched met de kussenvlek, is deze derde set misschien wel de meest natuurlijke.

Hierbij is het verschil dat hypothese set 1 opgesteld kan worden zonder dat de DNA profielen van de verdachte en de vlekken bekend zijn. En bij de hypothesen 2 en 3 alleen als de DNA profielen wel bekend zijn.

Op het eerste gezicht is het niet duidelijk waarom het een probleem is dat er verschillende sets van hypothesen mogelijk zijn. Maar het probleem wordt duidelijk als de LR's van de verschillende sets van hypothesen berekend worden. De LR zijn verschillend. Maar de LR is juist datgene wat als meting gebruikt wordt voor de bewijskracht tegen de verdachte Jan Smit.

Dit wordt nog verwarrender wanneer je realiseert dat na het bewijs, dat is, na het uitvinden dat de misdaad profielen A en B zijn en dat Smit profiel A heeft, alle drie de hypothesen equivalent aan elkaar zijn. Maar welke hypothese set is nu de juiste?

De LR verschilt voor de hypothesen, maar alle zijn ze wel gebaseerd op het feit dat de vlekken van 2 verschillende personen zijn (oftewel, het DNA profiel is al bepaald).

Mag een deskundige bij het formuleren van zijn hypothesen gebruik maken van het bewijs dat al voorhanden is? Dit is niet zoals je het wilt, de deskundigen moeten juist onafhankelijk hun waarde geven van de LR.

Het is zo dat de a-posteriori kans niet alleen afhangt van het nieuwe bewijs, maar ook van de a-priori kans. Deze moeten bepaald worden door de rechtbank, en niet door de expert. Maar een mogelijke moeilijkheid hierbij is dat het switchen tussen verschillende conditioneel gelijke hypothesen een verschillende LR kan geven. We noemen een set van hypothesen conditioneel gelijk als ze gelijk zijn ná het DNA bewijs.

Natuurlijk kan het wel zo zijn dat er goede redenen zijn om een bepaalde set van hypothesen te prefereren voor een andere set.

4.2.2 DNA databank

Een verkregen DNA profiel kan niet alleen gebruikt worden als er een verdachte is, maar ook om een verdachte op te zoeken in een DNA databank.

Stel dat het profiel vergeleken wordt met profielen in een databank ter grootte 3000. Nu blijkt dat één van de personen hetzelfde DNA profiel heeft.. Het spoor is nu dus vergeleken met 3000 personen in plaats van 1. De bewijskracht van deze DNA match is nu groter geworden, maar dit hoeft niet te betekenen dat kans dat de verdachte de dader is, groter is. Er zou bijvoorbeeld nog helemaal geen ander bewijs tegen de verdachte aanwezig kunnen zijn. De LR van deze verdachte kan met betrekking tot DNA profiel dan wel groot zijn, maar zijn totale a-posteriori kans niet omdat zijn a-priori kans klein was.

De reden dat de bewijskracht groter is, is dat er 2999 niet matchende personen uitgesloten kunnen worden als dader.

4.2.3 Denkfouten

Het gebruik van de Bayesiaanse methode in de rechtspraak leverde ook enkele denkfouten bij het waarderen van bewijs (Broeders, 2003). Deze denkfouten zijn al langer bekend in de psychologie.

Prosecutor's fallacy

Het gaat hierbij om een denkfout die de stelling van het openbaar ministerie sterker doet voorkomen dan ze in feite is.

Er is namelijk onderscheid tussen vragen over de waarschijnlijkheid van het bewijs (deskundige), en over vragen die zich richten op de waarschijnlijkheid van de hypothese (dat de verdachte al dan niet de dader is) en waarvan de beantwoording is voorbehouden aan de rechter. Dit komt ook voor in Nederland.

De prosecutor's fallacy is een denkfout die regelmatig voorkomt in strafzaken. Een openbare aanklager heeft wat bewijs verzameld (bijvoorbeeld een DNA match) en laat een expert getuigen dat de kans voor het vinden van dit bewijs als de verdachte onschuldig was zeer klein was. De denkfout is nu gemaakt als men nu doorgaat met het beweren dat de kans dat de verdachte onschuldig is vergelijkbaar klein is.

Een voorbeeld:

Neem aan dat een verkrachting is gepleegd in een stad. 20.000 mannen in deze stad hebben hun DNA vergeleken met een sample van de misdaad. Eén van deze mannen heeft een overeenkomend DNA profiel, en tijdens zijn rechtszaak wordt getuigd dat de kans dat twee DNA profielen matchen slechts 1 op 10.000 is. Dat wil zeggen dat de kans dat het profiel overeenkomt terwijl het bloed toch niet van de verdachte is, 1 op de 10.000 bedraagt.

Dit betekent niet dat de kans dat de verdachte onschuldig is gelijk is aan 1/10.000. Er kan dus niet gezegd worden, het profiel komt overeen, dus de kans dat het bloed niet van de verdachte is bedraagt 1 op de 10.000.¹⁰

De fout in de argumentatie van de aanklager wordt veroorzaakt door de verwarring van twee conditionele kansen.

$P(\text{overeenkomst} | \text{bloed niet van verdachte}) \neq P(\text{bloed niet van verdachte} | \text{overeenkomst})$
Deze kansen zijn niet hetzelfde. De eerste voorwaardelijke kans is de kans die de expert geeft, en de tweede kans wordt gebruikt door de rechter om te bepalen of een verdachte schuldig is.

¹⁰ Omdat er 20.000 mannen getest waren, zijn er 20.000 mogelijkheden geweest om een match toevallig te vinden. De kans dat er tenminste 1 DNA match is, is $1 - (1 - 1/10.000)^{20.000}$, ongeveer 86%. En dit is duidelijk meer dan 1/10.000. De kans dat exact 1 van de 20.000 mannen een match heeft is $20.000 * (1/10.000) * (1 - 1/10.000)^{19.999}$, of ongeveer 27%, wat nog steeds behoorlijk hoog is.

Deze denkfout is vooral kwalijk als de kans dat de verdachte de donor zou zijn berekend kan worden aan de hand van het DNA bewijs alleen. Wat natuurlijk onlogisch impliceert dat elk ander bewijs in de zaak (zelfs als de verdachte een dode vrouw is, of als de dader gefilmd is bij de actie) helemaal geen verschil maakt.

Dus de gevolgen kunnen groot zijn in de gevallen waarin de match kans groot is en er vrijwel geen overig bewijs is. De denkfout ligt in het feit dat de a-priori kans van schuld niet in overweging is genomen. Als deze kans klein is, dan is het enige effect van de gepresenteerde bewijs dat deze kans iets stijgt, maar niet perse dramatisch.

Hieronder een ingewikkeld voorbeeld van hoe een klein verschil in formulering al fout is (Brenner, 2004).

correcte verklaring/bewering versus prosecutor's fallacy	
Correcte verklaring	Prosecutor's fallacy
De kans is 1/7000 dat een(bijzonder)persoon anders dan de verdachte een vlek als de werkelijke vlek zou achterlaten.	De kans is 1/7000 dan iemand (om het even wie) anders dan de verdachte de vlek achtergelaten heeft.

Tabel 2: Goede versus foute formulering

Defense attorney's fallacy

Ook de verdediging kan zich bewust of onbewust van onjuiste redeneringen bedienen bij het gebruik van voorwaardelijke kansen. Zo zou de advocaat de impact van 10.000 tot 1 matching odds uit het eerder genoemde voorbeeld kunnen vergoelijken door te zeggen:

Er zijn in de stad naast de verdachte nog 100.000 andere mannen. Hiervan zullen waarschijnlijk 10 personen hetzelfde DNA profiel hebben als de verdachte. De kans dat het bloed gevonden op de plek van de misdaad van de verdachte is, is slechts 1 op de 11, dus ongeveer 9%. Dus de kans dat de verdachte onschuldig is, is 91%.

Dit is de juiste redenering als al het andere bewijs suggereert dat iedere man in de stad dezelfde toegang tot de crime-scene had als de verdachte.

De fout die hier gemaakt wordt, is dat in feite stilzwijgend wordt uitgegaan van een gelijke a-priori kans voor alle personen met hetzelfde DNA profiel, inclusief de verdachte. In veel gevallen zal de verdachte echter op grond van andere en van het DNA materiaal onafhankelijke evidentie de status van verdachte hebben bereikt. Er wordt dus aangenomen dat ieder persoon a-priori met even grote waarschijnlijkheid de dader kan zijn.

Hierbij wordt dus al aangenomen dat de dader in deze stad woont en dat iedereen met even grote kans de dader kan zijn (inclusief bejaarden, baby's en gehandicapten e.d.)

Base rate fallacy

In het maken van probabilistische gevolgtrekkingen moeten waarnemers algemene, breed gebaseerde informatie over populatie karakteristieken in beschouwing nemen. En meer specifiek de prior kans dat een bepaalde gebeurtenis voorkomt. De neiging om zulke informatie te negeren of te weinig in account te nemen heet base rate fallacy.

Voorbeeld

Stel er is een haar gevonden in de pet die de dader verloren heeft tijdens zijn ontvluchting. Tijdens zijn overval sprak de dader Nederlands. Er is nog geen verdachte, maar het DNA profiel kan wel vergeleken worden met verschillende populaties. Er wordt nu bijvoorbeeld gekeken naar de frequentie van het DNA profiel in de gehele Nederlandse bevolkingsgroep (A) en deze blijkt 1 op 1 miljoen te zijn. En de frequentie bij de gehele Belgische bevolkingsgroep (B) bijvoorbeeld 1 op 1 miljard. De likelihood ratio voor de hypothese dat de dader uit A komt en niet uit B is dan 1000.

Nu is de denkfout dat men geneigd is te denken dat dit inhoudt dat de kans dat de dader uit A komt veel groter is dan uit B. Maar de kans dat een dader uit A komt hangt behalve van de grootte van A, ook af van andere factoren.

Het is dus voor een forensisch deskundige belangrijk te bedenken waarover hij een uitspraak moet doen. Maar hij moet ook zijn conclusie op de juiste wijze formuleren om zo denkfouten te voorkomen. Hierbij moet wel gezegd worden dat de perfecte oplossing om deze denkfouten te voorkomen, nog niet gevonden is.

4.2.4 Vonnis

De deskundige mag dus uitdrukkelijk niet de rol van de rechter spelen. Hij zal geen inschattingen mogen doen voor de a-priori en a-posteriori kansverhoudingen. Deze opvatting van de rol van de deskundige past ook bij de wettelijke bepalingen. In 344 Sv lid 1 sub 4 staat dat deskundigen verslag moeten uitbrengen "behelzende hun gevoelens betreffende hetgeen hunne wetenschap hen leert omtrent datgene wat aan hun oordeel onderworpen is".

Vanuit het gezichtspunt van een bepaalde rechter zijn er twee betekenisvolle categorieën van a-posteriori: klein genoeg om "reasonable doubt" te hebben, of groot genoeg om geen significante twijfel te hebben. Dus er moet een minimum threshold waarde van a-posteriori zijn voor het veroordelen door elke rechter. De grootste waarde van alle rechters voor deze threshold waarde is een waarde die aan het geweten van alle rechters voldoet. Deze waarde zou wettelijk vastgelegd kunnen worden en vervolgens als threshold waarde in alle rechtszaken toegepast kunnen worden.

4.3 Conclusie

De grote bewijskracht van een DNA profiel wordt veroorzaakt door de extreem lage matchkans. Maar DNA als forensische techniek is geen wondermiddel. De aanwezigheid van celmateriaal op een plaats delict hoeft niet direct in verband te staan met dat de donor van het materiaal ook bij het onderzochte feit is betrokken. Omdat een spoor van celmateriaal door iedereen constant wordt achtergelaten. Daarom zal de bewijskracht van DNA gecombineerd moeten worden met andere bewijzen in een rechtszaak.

Hierbij moet wel goed over de formulatie van de hypothesen nagedacht worden. Verder zal bij het bepalen van het DNA bewijs er opgelet moeten worden dat men niet in de valkuil van de standaard denkfouten valt. Het is dus van belang voor een expert om een juiste formulatie van zijn conclusies te geven.

Er zal in de toekomst nog onderzoek verricht moeten worden naar het formuleren van hypothesen en naar de threshold waarde.

5 Alcohol

Mijn begeleider, de heer Jongbloed, wees mij op het proefschrift van de heer A. Korzec: "Confirming alcoholism in drivers under influence". In dit proefschrift wordt verslag gelegd van een onderzoek naar een nieuwe test die objectief vast kan stellen of je alcoholist bent of niet. Het bijbehorende computerprogramma is in 5 jaar tijd ontwikkeld. Deze test zou dan gebruikt kunnen worden door het Centraal Bureau Rijvaardigheidsbewijzen om te bepalen welke automobilisten hun rijbewijs moeten inleveren. Het computerprogramma is gebaseerd op de wiskundige formules van Bayes.

Een aanhouding door de politie resulteert bij een alcoholpromillage van 1.8 of hoger voor de automobilist niet alleen in een boete, maar ook in een bestuursrechterlijke procedure¹¹. De rechter besluit of de automobilist te verslaafd is om nog langer auto te rijden, en psychiaters geven de rechtbank hierbij advies. Voor de bewijsvoering in een proces stelt een psychiater in opdracht van het CBR vast of de beklaagde éénmalig ver over de schreef gegaan is en dus een gelegheidsdrinker is of een alcoholverslaving heeft.

Er bestaan voor alcoholisme twee soorten aandoeningen:

- Alcohol use disorder (AUD). Hierbij gaat het om de sociale en geestelijke gevolgen van het misbruik van alcohol. Deze aandoening wordt door een psychiater gediagnosticeerd.
- Hazardous and harmful alcohol use (HHAU). Hierbij gaat het om de lichamelijke gezondheid, dus de lichamelijke gevolgen van het alcoholmisbruik.

Naar schatting kunnen mannen 21-35 glazen en vrouwen 14-21 glazen verdeeld over één week ook op de lange duur zonder lichamelijke en/of geestelijk schade verdragen.

bron : NHG-Standaard (Nederlands Huisartsen Genootschap <http://nhg.artsennet.nl>)

5.1 Gebruik Bayes in diagnose

De regel van Bayes wordt op diagnostisch gebied regelmatig toegepast. Deze regel wordt toegepast als een methode om de klinische waarde van diagnostische tests te onderzoeken. (van Aelst, 2003)

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}$$

In diagnostische termen betreft de Bayes theorie de waarde van een diagnostische test B voor de aanwezigheid van een ziekte A. Alcohol is giftig en alcoholisme (A) veroorzaakt een verandering in vele lichamelijke functies die gebruikt kunnen worden als diagnostische tests (B). De resulterende waarde van Bayes kan geformuleerd worden als de kans dat de ziekte A aanwezig is als de diagnostische test B positief is. De kans P(B) dat de test positief is, dus boven een bepaalde gekozen cut-off waarde, kan verkregen worden door het testen van grote steekproeven van de doel populatie.

¹¹ Artikel 426 over ordeverstoring, het hinderen van verkeer en het bedreigen van andermans veiligheid onder invloed van alcohol

Artikelen 130-134 van de Wegenverkeerswet 1994 over de vorderingsprocedure en de vereiste dat het subject zich heeft onthouden van alcohol misbruik in de laatste 12 maanden

In een diagnostische context maakt men dus gebruik van conditionele kansen.¹² Het verband tussen het resultaat van een diagnostische test en de werkelijke situatie kan dus weergegeven worden in termen van de volgende twee conditionele kansen.

Definities.

- De **sensitiviteit** van een diagnostische test is de voorwaardelijke kans dat de test positief is als de ziekte aanwezig is:

$$P(\text{test} + | \text{ziek})$$

De waarde hiervan kan verkregen worden door het testen van grote aantallen proefpersonen met de ziekte. De waarde van de sensitiviteit is dan de verhouding van de positieve test resultaten gedeeld door alle test resultaten als de ziekte aanwezig is.

- De **specificiteit** van een diagnostische test is de kans dat de test negatief is wanneer de ziekte afwezig is:

$$P(\text{test} - | \text{niet ziek})$$

De specificiteit wordt verkregen door de verhouding te nemen van de negatieve test resultaten gedeeld door alle test resultaten als de ziekte afwezig is.

Gewoonlijk zijn deze waardes bij de arts bekend. Maar voor een diagnose zijn echter de omgekeerde conditionele kansen van belang. De werkelijke waarde van een test wordt dus het beste omschreven met onderstaande kansen.

Definities.

- De **positieve predictieve waarde** van een test is de kans dat de ziekte aanwezig is als de test positief is:

$$P(\text{ziek} | \text{test} +)$$

- De **negatieve predictieve waarde** van een test is de kans dat de ziekte afwezig is als de test negatief is.

$$P(\text{niet ziek} | \text{test} -)$$

Verder is de klinische waarde van een test afhankelijk van de prevalentie van een specifieke ziekte in een specifieke populatie. Deze waarde is dus afhankelijk van tot welke klasse van leeftijd, sekse, rokerstatus enzovoort de patiënt behoort.

Definitie.

- De **prevalentie** is de kans dat een willekeurig persoon uit de doel populatie ziek is:

$$P(\text{ziek})$$

De waarde van de prevalentie wordt verkregen door het aantal personen waarbij de ziekte aanwezig is te delen door het totaal aantal personen in de steekproef. Meestal is de enige onbekende waarde in de Bayes formule de kans dat de ziekte aanwezig is in de doel populatie. Er moet dus diagnostisch onderzoek gedaan worden naar de prevalentie.

Er zijn twee mogelijke strategieën hiervoor:

- op het individuele niveau door elk subject in de populatie te diagnosticeren met een specifieke diagnostische procedure. Als deze ene procedure niet perfect werkt, kunnen verschillende procedures parallel of achtereenvolgens gecombineerd worden. De resulterende prevalentie ratio is het percentage van de positief gediagnosticeerde personen in de populatie.

- op het populatie niveau door aan te nemen dat een gegeven variabele aanwezig is in een bepaald percentage van de ziekte. De benadering van de prevalentie wordt dan bepaald met de volgende formule die afgeleid is van de Bayes theorie:

¹² Zie appendix A voor enkele rekenregels uit de kansrekening

$$prevalentie = \frac{\text{proportie van positieve testen} - (1 - \text{specificiteit})}{(\text{sensitiviteit} + \text{specificiteit} - 1)}$$

De theorie van Bayes kan dus vertaald worden in diagnostische termen als positieve predictieve waarde, sensitiviteit, specificiteit en prevalentie. De regel van Bayes impliceert dan:

$$\text{positieve predictieve waarde} = \frac{\text{prevalentie} * \text{sensitiviteit}}{\text{prevalentie} * \text{sensitiviteit} + (1 - \text{prevalentie}) * (1 - \text{specificiteit})}$$

Deze beide formules worden afgeleid in appendix B.

De positieve predictieve waarde van een test voor alcoholisme verandert significant als het wordt toegepast in populaties met verschillende prevalentie. Het is daarom van belang een goede waarde voor de prevalentie te nemen.

Op internet is een calculator te vinden die uit verkregen resultaten bovenstaande waardes kan berekenen. Deze calculator is te vinden op www.intmed.mcw.edu/clincalc/bayes.html.

5.2 Bayesian Alcoholism Test

Artsen maken in hun diagnose van alcoholisme gebruik van testen. In het klinisch redeneren wordt de uitslag van een test geïnterpreteerd als het stijgen of dalen van de kans van een positieve diagnose. Klinische argumenten zijn dus subjectieve kansen gebaseerd op klinische ervaring en het lezen van research. Maar de conventionele testen falen in het nauwkeurig en objectief bevestigen van HHAU. Daarom wil de schrijver van het proefschrift een diagnostisch instrument bouwen dat gebaseerd wordt op dit klinische redeneren, zodat dit instrument als een bevestigingstest voor alcoholisme kan dienen.

Dit leverde een test op genaamd de Bayesian Alcoholism Test (BAT). De BAT is een expert systeem waarin informatie uit bloedtests, lichamelijk onderzoek en een gesprek met de mogelijke drankmisbruiker wordt ingevoerd. Het computerprogramma analyseert de uitkomsten van deze reeks testresultaten en weegt deze mee voor de eindiagnose. Deze eindiagnose is een kans: hoe waarschijnlijk is het dat de onderzochte persoon alcoholist is. Verder geeft het systeem ook de kans op leverziektes en diabetes.¹³

De benodigde data voor de BAT zijn dus:

- prevalentie van aandoeningen
- prevalentie van klinische tekens
- conditionele kansen tussen ziektes en testen.

De data in de BAT is op twee manieren verkregen. Ten eerste is er een literatuuronderzoek naar epidemiologische data¹⁴ gedaan. De data is gebaseerd op mannen in West-Europa of de Verenigde Staten. Als hiervan geen gegevens aanwezig waren, zijn de gegevens genomen van gemengde data van mannen en vrouwen, of van klinische populaties. Dit leverde ongeveer 40% van de benodigde kansen.

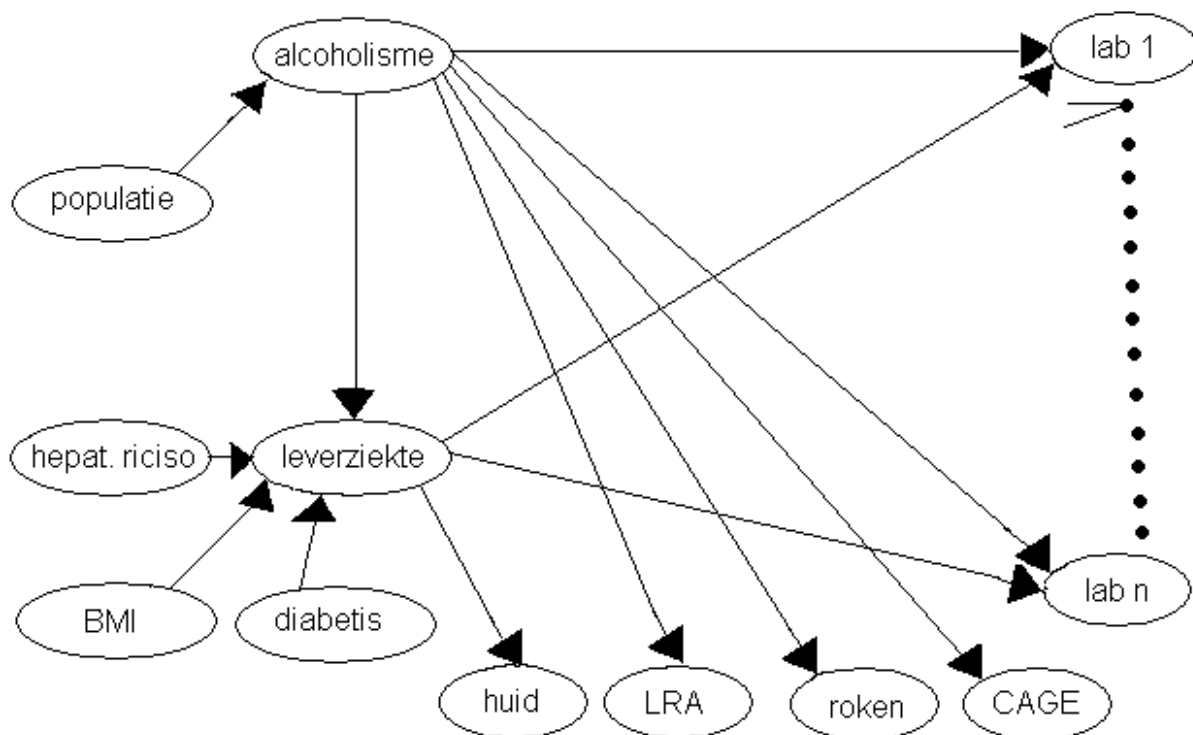
¹³ De reden dat deze ziektes ook meegenomen zijn is omdat deze dezelfde symptomen hebben. Er wordt rekening gehouden met dat bepaalde afwijkende bloedwaardes niet alleen veroorzaakt kunnen worden door alcoholisme.

¹⁴ Een voorbeeld van epidemiologische data is: van alle mannen die bovenmatig veel enzym X in hun bloed hebben, is tien procent alcoholist.

Ten tweede is de hulp van experts ingeroepen. Het consulten van vakmensen is gedaan wanneer de literatuur inconsistent of niet beschikbaar was. Dit gebeurde dus in 60% van de gevallen.

Het gebruikte systeem is een expert systeem. Een expert systeem is een computer programma dat bestaande algemene kennis over een domein codificeert op zo'n manier dat het invoeren van data van een bepaalde patiënt tot een geldige diagnostische kans leidt voor het hebben van een ziekte. Het systeem heeft als framework een Bayesiaans belief netwerk. Dit netwerk bevat grafische structuren waarin de knopen de ziektes, syndromen, patho-fysiologische entiteiten, symptomen en diagnostische testen representeren en waar een pijl van knoop n naar knoop m aangeeft dat de kans van m afhangt van de kans van n. De statistische software HUGIN 5.7 is gebruikt om dit netwerk te bouwen. (hugin expert A/S, Aalborg, Denemarken)

Hieronder staat een schematische weergave van het gebruikte Bayesiaanse netwerk.



Figuur 3: Aan de linkerkant staan de a-priori kansen voor de ziektes en toestanden. Deze worden gecombineerd met de biochemische (rechterkant) en de klinische (onderkant) bevindingen. Een pijl die van een ziekte naar een symptoom of biochemische test gaat, houdt in dat het symptoom of de test afhankelijk is van de ziekte of toestand.

Het systeem is onderverdeeld in drie componenten waarover gegevens ingevuld kunnen worden:

- klinische tekens: zoals het aantal sigaretten dat gerookt wordt, de CAGE vragen¹⁵ en het LRA (level of response to alcohol)
- biochemische testen: Laboratorium gegevens zoals bijvoorbeeld CDT en ALT. In bovenstaand schema weergegeven als lab 1 tot en met lab n.
- andere gegevens: het al dan niet hebben van diabetes, het hepatitis risico en de BMI¹⁶ van de persoon.

¹⁵ zie Appendix C

¹⁶ BMI is de afkorting van Body Mass Index, een meting tussen het gewicht en de lengte

Het Bayesiaanse netwerk werkt met conditionele kanstabellen die de conditionele kansverdeling weergeven van een ziekte die verschillende symptomen en biochemische abnormaliteiten veroorzaakt. Hieronder staat een voorbeeld van een data tabel in het systeem. Het is de conditionele kanstabel voor de knoop ALT, één van de gebruikte biochemische testen. In de tabel staan de kansen voor waarden van alanine aminotransferase afhankelijk van de aanwezigheid van alcohol gebruik en van leverziekte.

Alcoholism	harmful use >80g/d			
liverdisease	fatty liver	hepatitis	liver cirrhosis	no liverdisea
highly elevated >100	0.15	0.3	0.3	0.1
slightly elev. 50-100	0.35	0.4	0.3	0.2
not elevated <50	0.5	0.3	0.4	0.7

Alcoholism	hazardous use 40-80			
liverdisease	fatty liver	hepatitis	liver cirrhosis	no liverdisea
highly elevated >100	0.05	0.25	0.25	0.05
slightly elev. 50-100	0.15	0.35	0.25	0.05
not elevated <50	0.8	0.4	0.5	0.9

Alcoholism	no HHAU			
liverdisease	fatty liver	hepatitis	liver cirrhosis	no liverdisea
highly elevated >100	0.05	0.2	0.1	0
slightly elev. 50-100	0.15	0.3	0.3	0.025
not elevated <50	0.8	0.5	0.6	0.975

Figuur 4: Conditionele kanstabel voor ALT

Na het invullen van (een deel van) de gegevens uit de testen geeft het BAT als output de kans op HHAU en enkele andere ziektes. De output bestaat uit continue waarden (tussen 0 en 100 procent). Het is nu vervolgens de taak van de arts om deze continue kansen om te zetten in een binaire beslissing. Dus aan de hand van de verkregen kans een antwoord geven op de vraag of de persoon al dan niet alcoholverslaafd is.

Om te kijken of het systeem een werkelijke verbetering is, wordt in het proefschrift de diagnostische nauwkeurigheid van het door computer geassisteerd diagnostisch systeem BAT vergeleken met conventionele testen van alcoholisme. Dit is op de volgende manieren onderzocht. Eerst heeft er een controle plaatsgevonden of het systeem in staat is om onderscheid te maken tussen 'clear alcoholics' (alcoholisten) en 'social drinkers' (sociale drinkers). Vervolgens is de diagnostische onderscheidmogelijkheid van de BAT in de populatie van zware drinkers bekeken. De populatie van zware drinkers is het spectrum tussen alcoholisten en sociale drinkers.

Er is dus gebruikt gemaakt van volgende drie studie groepen:

- een controle groep van niet alcoholisten ter grootte 47
- een groep van behandeling zoekende alcoholisten van 67 personen
- een groep van niet behandeling zoekende alcoholisten met daarin 68 mannen

Er werd hierbij gekeken naar de valse positieve en de valse negatieve gevallen. Een vals positief resultaat wordt gedefinieerd als de gebeurtenis waarbij de test positief is maar de werkelijke situatie negatief is. Analoog wordt een vals negatief resultaat verkregen wanneer de test negatief is terwijl de werkelijke situatie positief is. Beide resultaten zijn fout en in het algemeen wenst men testen die beide fouten zo klein mogelijk maken.

5.3 Discussie

5.3.1 Ontkenning

Jaarlijks komen 10000 mensen bij de forensisch psychiater terecht voor een alcoholismetest. Na een jaar mag de alcoholist proberen te bewijzen dat hij van de drank af is. De beklaagden zullen natuurlijk altijd proberen hun verslaving te ontkennen, aangezien ze niet graag hun rijbewijs verliezen. Daardoor kan een klassieke anamnese niet afgenomen worden want dit levert een ernstige onderdiagnose op. Dus als de patiënt onwillig is alcoholisme te onthullen of zich niet bewust is van alcohol gerelateerde problemen dat is er geen accurate test om de diagnose objectief te bevestigen.

De keuze om het onderzoek toe te spitsen op HHAU in plaats van AUD is daarom een verstandige keuze. Dit omdat het betere resultaten op zal leveren want lichamelijke gevolgen zijn objectiever te meten dan geestelijke. Verder is het zo dat HHAU erg geschikt is om om te zetten in data. Ook is het zo dat AUD ook te bepalen is als iemand op de dag van het onderzoek niet heeft gedronken.

5.3.2 Systeemkeuze

Natuurlijk zou een arts exact dezelfde testen en vragen kunnen gebruiken in zijn onderzoek als de testen die gebruikt worden in het BAT model. Dit omdat elke testuitslag op een bepaalde kans, of waarschijnlijkheid duidt dat de persoon in kwestie een alcoholprobleem heeft. Maar het probleem is vervolgens dat de arts al die informatie moeilijk kan verwerken in een objectief oordeel. Bij het combineren en uitsluiten van die waarschijnlijkheden kun je vrij makkelijk iets over het hoofd zien.

Expert systemen zijn bruikbaar wanneer er grote aantallen diagnostische tests voor een gegeven ziekte zijn en wanneer de relatie tussen de ziekte die gediagnosticeerd wordt en de resultaten van de testen voor de ziekte van een probabilistische aard is. De keuze voor een expert systeem om alcoholisme te diagnosticeren is zeker verstandig.

Verder is het zo dat de Bayesiaanse berekeningen zich uitstekend lenen om opgenomen te worden in een computerprogramma.

Hierbij moet wel opgemerkt worden dat de kennis van een arts nog wel steeds nodig zal blijven om uit te sluiten dat de testresultaten niet vertekend worden door een zeer zeldzame ziekte, die niet in het model meegenomen is. En ook om de verkregen kans uit de test te interpreteren.

5.3.3 Bronnen

De gebruikte waardes in de kanstabellen in het systeem zijn door een uitgebreid onderzoek verkregen. De objectieve kansen zijn verkregen uit epidemiologische data, dus objectieve informatie, óf van experts, dus subjectieve schattingen door deskundigen.

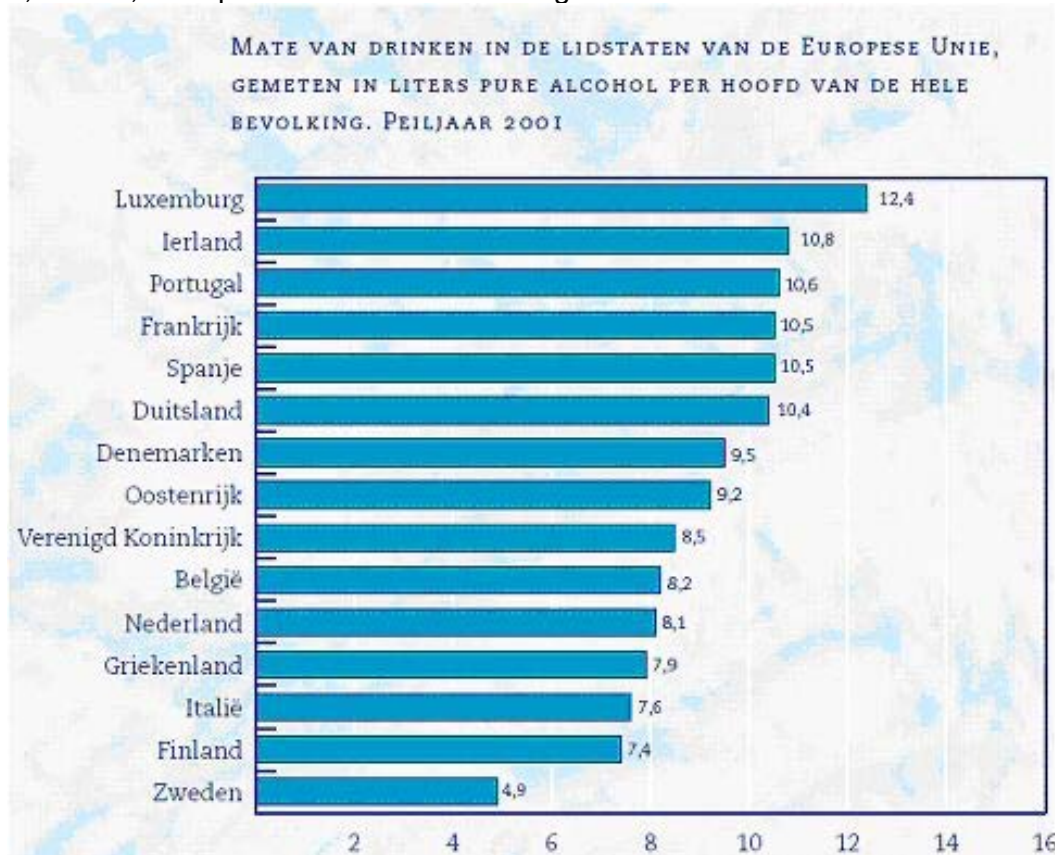
Het probleem bij deze kansen is dat veel gebruikte studies tekortkomingen hadden in hun onderzoeksmethode of niet overtuigende gegevens produceerden. Ook kwamen waardes voor dezelfde conditionele kans in verschillende onderzoeken niet met elkaar overeen.

Verder zijn de gegevens ook verzameld onder buitenlandse populaties.

Dit houdt dus in dat als er nieuwe onderzoeken gedaan worden en hieruit nieuwe data beschikbaar komen voor de conditionele kansen, de BAT nog betere resultaten zal kunnen geven.

Het bepalen van de juiste prevalentie blijkt ook een zeer lastige bezigheid te zijn. De prevalentie van alcoholisme in de Nederlandse DUI (drivers under influence) populatie ligt tussen de 8% en 82%. Dit grote bereik wordt veroorzaakt door het gebruik van verschillende diagnostische criteria voor alcoholisme en ook door keus van een beoordelingsprocedure. Ook hierbij geldt dat de waarde deels gebaseerd is op (oude) onderzoeken onder

buitenlandse populaties. Dit is naar mijn mening niet verstandig, aangezien landen onderling verschillen in alcohol gebruik en er de laatste 15 jaar veel veranderingen geweest zijn in alcohol gebruik en in het rijden onder invloed. De alcoholconsumptie in Europa varieert van 4,9 tot 12,4 liter per hoofd van de bevolking in 2001.



Figuur 5: gemiddeld aantal liter drank per persoon per land in 2001 (Bron: productschap dranken, commissie gedistilleerd)

De exacte waarde van de prevalentie is dus onbekend. Er is uiteindelijk gekozen voor de waarde van 50%.¹⁷

5.3.4 Validatie

De output waarden zijn ook onderworpen aan verschillende statistische testen.¹⁸ En deze waarden zijn vergeleken met andere bestaande alcoholisme testen. Met behulp van bestaande testen worden maximaal 60% van de alcoholisten geïdentificeerd. Bestuurders die dagelijks tussen de 4 en 8 glazen alcohol drinken, hebben met de nieuwe test een kans van 50 % dat ze het stempel alcoholist krijgen en dus hun rijbewijs verliezen

Alcoholgebruik en gezondheid hebben veel met elkaar te maken. Verantwoord alcoholgebruik wil zeggen: het aantal glazen dat u zonder bezwaar voor uw gezondheid kunt drinken. Hiervoor gelden de volgende drie richtlijnen.

1. Drink *gemiddeld* niet meer dan twee glazen alcohol per dag. Wanneer u gemiddeld niet meer dan twee glazen per dag drinkt, hoeft u zich geen zorgen te maken voor uw gezondheid.

¹⁷ Deze waarde is gevonden met de klinische diagnostische procedure (CDP). Wat opvalt is dat deze waarde gelijk is aan de nationale prevalentie waarde gevonden door de divisie Vorderingen (CBR). Maar dit komt omdat dat ze dezelfde methode gebruiken

¹⁸ Confidence intervals met Wilson's methode, LR's met de score methode, differentiatie a.d.h.v. entropie.

2. Drink maximaal drie glazen *per keer*.

De *gemiddelde* hoeveelheid alcohol is iets anders dan de hoeveelheid alcohol *per keer*. Wist u dat veel alcohol ineens niet gezond is? Het is aan te raden niet meer dan drie glazen per gelegenheid te drinken, zeker niet elke week. Overkomt dit u wel eens, dan kunt u uw lichaam de tijd geven om te herstellen door twee aaneengesloten dagen helemaal geen alcohol te drinken.

3. Drink ten minste twee dagen in de week geen alcohol.

Alcohol drinken kan heel makkelijk een gewoonte worden. Juist omdat het vaak naar aanleiding van iets gebeurt: om iets te vieren of om even ergens niet aan te hoeven denken. Of uit gewoonte, bijvoorbeeld elke dag vóór het avondeten. Als u ten minste twee dagen per week niets drinkt, voorkomt u dat u ongemerkt meer gaat drinken.

Bron: www.drinktest.nl

De pakkans loopt op naarmate de bestuurders dagelijks meer alcohol drinken. Voor meer dan 15 glazen is de pakkans 95%. De nieuwe methode is daarmee beter dan de bestaande tests.

De gebruikte test personen zijn naar mijn mening niet representatief voor rijders onder invloed. De groep van behandeling zoekende alcoholisten is verworven door mensen te selecteren in verslaving behandelingsfaciliteiten. Maar dit zijn personen die hun verslaving niet ontkennen aangezien ze al hulp gezocht hebben. Terwijl automobilisten die aangehouden zijn naar alle waarschijnlijkheid hun probleem zullen ontkennen aangezien ze hun rijbewijs niet willen verliezen.

Verder is de groep van niet behandeling zoekende alcoholisten door middel van advertenties in een wijnmagazine en door het bezoeken van wijnproef conventies verkregen. Deze zware drinkers zijn wijndrinkers. Naar mijn mening zijn wijndrinkers niet representatief voor mogelijke alcoholisten. Alcoholisten grijpen eerder naar sterkere drank zoals bijvoorbeeld wodka (reukloos).

De validiteit van een test heeft slechts betekenis voor de populatie waarvoor de test is gevalideerd.

5.3.5 Mannen

Een duidelijk minpunt naar mijn mening is dat de BAT alleen werkt voor mannen. Dit komt doordat Korzec tijdens zijn onderzoek geen epidemiologische gegevens heeft verzameld van vrouwen. Hij relativeert zijn keuze hiervoor door te zeggen dat maar 1 op de 10 rijders onder invloed, vrouw zijn.

Als men 20 jaren geleden over alcoholisme sprak, dan dacht men automatisch aan mannelijke drinkers. Ook in studie en onderzoek over alcoholisme bleef de vrouw afwezig of tenminste onzichtbaar en conclusies van wetenschappelijk onderzoek in verband met alcoholisme betroffen meestal mannen.

(bron: Gezondheid NV)

Korzec kon dus weinig anders dan zijn test alleen baseren op gegevens van mannen, omdat gegevens van vrouwen niet beschikbaar zijn.

En hierbij komt ook nog eens dat vrouwen heel anders op alcohol reageren dan mannen.

Vrouwen krijgen bij hetzelfde alcoholgebruik een hoger alcoholpromillage in het bloed dan mannen. Lichaamsvet voorkomt dat alcohol wordt verdund. Omdat vrouwen over het algemeen lichter zijn en relatief meer lichaamsvet hebben, hebben zij bij het gebruik van dezelfde hoeveelheid alcohol een hogere concentratie alcohol in het bloed.

(BRON: CBR)

De BAT zal dus uitdrukkelijk niet voor vrouwen gebruikt mogen worden. In de toekomst zou de BAT herschreven moeten worden om ook de Bayesiaanse test op vrouwelijke dronken automobilisten toe te kunnen passen.

5.4 Conclusie

De BAT kan gebruikt worden door psychiaters om de diagnose van alcoholmisbruik te laten stellen. Zeker als het om twijfelgevallen gaat. Het programma moet psychiaters in de toekomst meer houvast bieden als zij bij de rechter een advies moeten geven over het wel of niet intrekken van een rijbewijs.

Het programma kan voor een objectievere diagnose van de patiënt zorgen. Gewoonlijk geeft een arts een diagnose in het belang van de patiënt, maar nu is het doel in een wettelijke context en is de verkeersveiligheid het referentie punt. Ook wordt iedere patiënt op dezelfde wijze gediagnosticeerd.

De BAT is, na validatie bij de juiste populatie, op vele gebieden in te zetten, zoals bij forensisch onderzoek, in de verzekeringswereld en in bepaalde beroepsgroepen (tijdens sollicitaties bijvoorbeeld).

Bijkomende voordelen van de test zijn dat bijvoorbeeld wijnproevers zich regelmatig ermee laten testen, om te kijken of hun "beroep"¹⁹ nog geen schadelijke gevolgen voor hun lichaam heeft.

¹⁹ De èchte wijnproever spuugt de wijn uit!

6 Afsluiting

Statistiek en kansrekening zijn al jaren geaccepteerd in de rechtszaal. Het gebruik van statistiek is een essentieel onderdeel van verschillende soorten bewijsmiddelen. En zelfs Bayesiaanse statistiek neemt al heel lang een positie in bij de rechtspraak, het wordt toegepast in vele soorten bewijsmiddelen en in het bepalen van het oordeel de rechter.

6.1 Sterke punten

Het gebruik van statistiek in de rechtszaal is niet verrassend, aangezien statistiek en de wet in essentie overeenkomen. Ondanks dat beide afkomstig zijn uit verschillende achtergronden, behandelen beide gebieden het interpreteren van bewijsmateriaal, het analyseren van gegevens, het testen van hypothesen en het nemen van besluiten onder onzekerheid.

In de rechtspraak zal de Bayesiaanse aanpak van meer nut zijn dan een klassieke aanpak. Dit omdat de subjectieve kans per definitie relevant is voor incidentele gevallen. In dit werkstuk is naar voren gekomen dat een voorbeeld hiervan het vergelijkend onderzoek (bijvoorbeeld DNA) is. De belangstelling gaat hierbij uit naar de relatie tussen het spoor en een bepaald individu. Als gevolg hiervan sluit het Bayesiaanse subjectieve kansbegrip veel beter aan bij de vraag dan het klassieke kansbegrip. Het is erg gekunsteld om in de forensische context een kans te zien als een frequentie of als één of andere abstracte kansmaat.

Verder worden er in een zaak meerdere technische bewijsmiddelen gevonden (DNA, vingerafdruk, schoensporen). Het is dan niet alleen de vraag wat de bewijskracht is van deze afzonderlijke bewijzen, maar ook wat de bewijskracht is van de combinatie van de bewijsmiddelen. Het Bayesiaanse denkraam biedt hiervoor een uitstekende richtlijn.

De Bayesiaanse stroming levert ook een definitie op van de taak van de deskundige. Zijn taak is het bepalen van de LR. Ook wordt de taak van de rechter ermee gedefinieerd: het bepalen van de a-priori en a-posteriori kansen. Verder geeft de regel van Bayes ook nog een numerieke definitie van bewijskracht.

Expertsystemen die gebruik maken van Bayesiaanse statistiek, zoals Bayesian belief netwerken, zijn ook in de rechtspraak van nut. Deze systemen zouden de taak van de deskundigen of de rechter kunnen ondersteunen.

6.2 Zwakke punten

De Bayesiaanse statistiek zal wel met enige voorzichtigheid toegepast moeten worden. Een feit is namelijk dat leken, en veel mensen zijn leken als het op statistieken aankomt, zich snel laten imponeren door cijfers, percentages, risicometingen en dergelijke. Juristen zijn geen uitzondering op die regel. En dat is natuurlijk geen goede zaak. Als zij niet het verschil kunnen zien tussen goed en slecht statistisch onderzoek, gepast gebruik en misbruik van statistiek, dan lopen zij het risico onzin te verkondigen en veel kwaad te doen. Het is daarom van belang dat juristen enig inzicht hebben in het gebruik van statistische methoden.

Verder blijft het een probleem om de subjectieve kansen te bepalen. Hiervoor zijn tot nu toe geen richtlijnen vastgelegd. Het zou natuurlijk optimaal zijn als er technieken waren waarmee iedereen dezelfde kans zou geven, als men over dezelfde relevante informatie beschikt voor de waarheid van een onzeker voorstel. Hoe de a-priori kansverdeling gekozen moet worden

staat nergens beschreven. Dit zou dus in kunnen houden dat rechters verschillende schalingen aanhouden. Hierbij geldt natuurlijk wel dat hoe meer bewijs er is, hoe minder de waarde van deze startkeuze van belang is. Ook in het interpreteren van de uiteindelijk verkregen a-posteriori waarde ligt een probleem. Er is in Nederland geen wet die in termen van kansen op een terechte beslissing definieert wat "wettig en overtuigend bewijs" is.

Waar de experts rekening mee moeten houden is dat toegepast statistisch onderzoek niet beter kan zijn dan de kwaliteit van de gegevens waar dat onderzoek op gebaseerd is. Er moet bijvoorbeeld rekening gehouden worden met het feit dat statistieken van verschillende landen meestal niet vergelijkbaar zijn. Hier geldt het GIGO-principe: 'Garbage in, Garbage out'. Als de kwaliteit van de data twijfelachtig is, zal de expert dus duidelijk bij zijn conclusie de twijfelachtigheid hiervan moeten melden. Het is dan vervolgens de taak van de rechter om te bepalen of hij deze informatie gebruikt.

De statistiek bewijst of verklaart niets. Statistiek geeft in het beste geval een beschrijving aan de hand van getallen van een situatie die misschien wel, maar misschien ook niet, om een verklaring vraagt. Een hoge correlatie kan louter toevallig zijn of het resultaat van een buiten het onderzoek gebleven factor. Daarom zal de uitspraak in een rechtszaak, en de uitkomst van de berekening met de regel van Bayes nooit gebaseerd mogen worden op één bewijsstuk.

6.3 Slotconclusie

Aanhangers van de Bayesiaanse richting hebben de neiging zich te gedragen als bekeerlingen die het licht hebben gezien. Zij zingen zelfs ketterse liederen, met als meest bekende "There's no Theorem like Bayes' Theorem"²⁰. Tot nu toe wordt de Bayesiaanse statistiek in veel minder gevallen toegepast als de klassieke statistiek. En dat is jammer. Want juist in het rechtssysteem is het gebruik van de regel van Bayes in de gevolgtrekking bijzonder handig.

Dit omdat het in rechtszaken altijd gaat over mogelijke toestanden van de werkelijkheid, niet over zich herhalende experimenten. Ook moeten er beslissingen onder onzekerheid genomen worden. Met behulp van de regel van Bayes krijgt de oordeelsvorming van de rechter een rationele onderbouwing.

De rechter zou bij wijze van spreken een rekenmachine voor zijn neus kunnen neerzetten en met behulp van de door de deskundigen geleverde LR's een uitspraak over de verdachte kunnen doen. Maar één ding is zeker, voordat het zover is zal er eerst verdergaand onderzoek verricht moeten worden.

Naar mijn mening kunnen Bayesiaanse methoden een meerwaarde leveren in de rechtspraak.

²⁰ Gezongen op de wijs van "there's no business like show business", Bron: A. de Vos, *De godsdienstoorlog der statistici*, <http://staff.feweb.vu.nl/avos/default.htm>

Appendix A Rekenregels kansrekening

Voor meer informatie over deze regels en stellingen uit de kansrekening verwijst ik naar het dictaat "Inleiding in de waarschijnlijkheidsrekening", van de Vrije Universiteit Amsterdam.

Axioma's Kolmogorov

In de kansrekening hebben we te maken met een willekeurige niet-lege verzameling Ω (de uitkomstenruimte), en een set deelverzamelingen hiervan genaamd de gebeurtenissen.

Op de set gebeurtenissen is een kans P gedefinieerd. De kans moet voldoen aan de volgende voorwaarden van Kolmogorov:

1 Een kans is niet negatief

$$P(A) \geq 0$$

2 De kans op de verzameling met alle mogelijke uitkomsten is 1

$$P(\Omega) = 1$$

3 Als gebeurtenissen niet tegelijkertijd kunnen optreden, is de kans dat één van deze gebeurtenissen optreedt, gelijk aan de som van de kansen op de afzonderlijke gebeurtenissen.

$$P(\cup A_i) = \sum P(A_i) \text{ met } (A_i) \text{ rij disjuncte gebeurtenissen.}$$

Uit bovenstaande axioma's zijn de volgende gevolgen afleidbaar:

- De kans op de lege verzameling is nul.

$$P(\emptyset) = 0$$

- De kans dat de uitkomst in A of B ligt is

$$P(A \text{ of } B) = P(A \cup B) = P(A) + P(B) - P(A \text{ en } B)$$

- De kans dat de uitkomst niet in A ligt is

$$P(A^c) = 1 - P(A)$$

- De kans dat de uitkomst in A en B ligt is

$$P(A \text{ en } B) = P(B \cap A) = P(A)P(B | A)$$

- De conditionele kans op A gegeven AB wordt gegeven door

$$P(A | B) = \frac{P(A \text{ en } B)}{P(B)}$$

En heeft als betekenis: Als je al weet dat B waar is, wat is dan nog de kans dat A waar is?

Centrale limietstelling

De stelling geeft aan dat de som van een aantal onafhankelijke en gelijk verdeelde stochastische variabelen een normale verdeling benadert als het aantal voldoende groot wordt gekozen. De variabelen zelf hoeven daarvoor geen normale verdeling te hebben.

Zij X_1, X_2, \dots een rij van onafhankelijke, gelijkverdeelde stochastische variabelen met verwachting μ en variantie σ^2 .

Dan geldt voor de bijbehorende partieelsom $S_n = X_1 + \dots + X_n$:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} U \quad [n \rightarrow \infty]$$

met U standaard normaal verdeeld.

Wet van grote aantallen

In de kansrekening zeggen de verschillende wetten van grote aantallen dat het gemiddelde van een opeenvolging van willekeurige variabelen (met een gemeenschappelijke verdeling) naar hun gezamenlijke verwachting convergeert, als het aantal variabelen naar oneindig gaat.

Zwakke wet van grote aantallen

Zij X_1, X_2, \dots een rij van paarsgewijs onafhankelijke, gelijkverdeelde stochastische variabelen met verwachting μ . Dan convergeert voor $n \rightarrow \infty$ het gemiddelde

$\overline{X}_n = (X_1 + \dots + X_n)/n$ in waarschijnlijkheid naar μ :

$$\lim_{n \rightarrow \infty} P(|\overline{X}_n - \mu| \geq a) = 0 \quad [a > 0]$$

Sterke wet van grote aantallen

Zij X_1, X_2, \dots een rij van onafhankelijke, gelijkverdeelde stochastische variabelen met verwachting μ . Dan convergeert voor $n \rightarrow \infty$ het gemiddelde $\overline{X}_n = (X_1 + \dots + X_n)/n$ met kans 1 naar μ :

$$P(\lim_{n \rightarrow \infty} \overline{X}_n = \mu) = 1$$

Appendix B Bewijzen

Voor de duidelijkheid worden hieronder nogmaals de verschillende symbolen herhaald.

A	ziek
B	test positief
$P(A)$	prevalentie
$P(B)$	proportie positieve testen
$P(B^c A^c)$	specificiteit
$P(B A)$	sensitiviteit
$P(A B)$	positieve predictieve waarde

Door deze waarden in te vullen in de gegeven formules, kan nu met behulp van verschillende rekenregels uit de kansrekening, gecontroleerd worden of de gegeven formules kloppen. Hierbij wordt aangenomen dat $P(A) > 0$ en $P(B) > 0$.

Prevalentie

BEWERING

$$\text{prevalentie} = \frac{\text{proportie van positieve testen} - (1 - \text{specificiteit})}{(\text{sensitiviteit} + \text{specificiteit} - 1)}$$

BEWIJS

$$P(A) = \frac{P(B) - (1 - P(B^c | A^c))}{P(B | A) + P(B^c | A^c) - 1} \Leftrightarrow$$

$$P(A) = \frac{P(B) - P(B | A^c)}{P(B | A) - P(B | A^c)} \Leftrightarrow$$

$$P(A)P(B | A) - P(A)P(B | A^c) = P(B) - P(B | A^c) \Leftrightarrow$$

$$P(B) = P(A)P(B | A) - P(A)P(B | A^c) + P(B | A^c) \Leftrightarrow$$

$$P(B) = P(A)P(B | A) + (1 - P(A))P(B | A^c) \Leftrightarrow$$

$$P(B) = P(B \cap A) + P(B \cap A^c) \Leftrightarrow$$

$$P(B) = P(B) \quad \square$$

Positieve predictieve waarde

BEWERING

$$\text{pos. pred. waarde} = \frac{\text{prevalentie} * \text{sensitiviteit}}{\text{prevalentie} * \text{sensitiviteit} + (1 - \text{prevalentie}) * (1 - \text{specificiteit})}$$

BEWIJS

$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + (1 - P(A))(1 - P(B^C|A^C))} \Leftrightarrow$$

$$P(A|B) = \frac{P(B)P(A|B)}{P(A)P(B|A) + P(A^C)P(B|A^C)} \Leftrightarrow$$

$$1 = \frac{P(B)}{P(A)P(B|A) + P(A^C)P(B|A^C)} \Leftrightarrow$$

$$P(A)P(B|A) + P(A^C)P(B|A^C) = P(B) \Leftrightarrow$$

$$P(B \cap A) + P(B \cap A^C) = P(B) \Leftrightarrow$$

$$P(B) = P(B)$$

□

Appendix C CAGE vragen

Van de bestaande vragenlijsten voor het diagnosticeren van alcoholisme wordt de CAGE-vragenlijst het meest toegepast. De oorsprong en de vragen van de CAGE test zijn beschreven door Ewing (1984).

De CAGE (acroniem van 'Cut down', 'Annoyed', 'Guilty' en 'Eye-opener') bestaat uit de volgende vier vragen:

1. Heeft u wel eens het gevoel gehad te moeten minderen met drinken ('Cut down')?
2. Raakt u wel eens geïrriteerd door aanmerkingen van anderen op uw alcoholgebruik ('Annoyed')?
3. Heeft u zich wel eens vervelend of schuldig gevoeld over uw drinken ('Guilty')?
4. Heeft u wel eens direct na het opstaan gedronken om de zenuwen de baas te worden of om van een kater af te komen ('Eye-opener')?

Bij twee of meer bevestigende antwoorden is de CAGE-test positief. Een positieve test bevestigt problematisch alcoholgebruik. Een negatieve test sluit dit echter niet uit. Bij een beginnend overmatig alcoholgebruik is de CAGE-test nog wel eens negatief.
(BRON: NHG-Standaard Problematisch alcoholgebruik)

Gebruikte afkortingen

ALT	ALanine aminoTransferase
AUD	Alcohol Use Disorder
BAT	Bayesian Alcoholism Test
CBR	Centraal Bureau Rijvaardigheidsbewijzen
CDP	Clinical Diagnostic Procedure
CDT	Carbohydrate-Deficient Transferrin
DNA	Deoxyribo Nucleic Acid
HHAU	Hazardous and Harmful Alcohol Use
LR	Likelihood Ratio
LRA	Level of Response to Alcohol
MCMC	Markov Chain Monte Carlo
OM	Openbaar Ministerie

Literatuurlijst

Referenties

- *Aelst, Stefan Van, Goetghebeur, Els
Medische statistiek
Dictaat van het vak Medische statistiek, Ghent University, H2, 2003
<http://allserv.ugent.be/~svaelst/courses/medstat.html>

- *Aitken, C.G.G.
Statistics and the evaluation of evidence for forensic scientists
Wiley, Chichester, 1995
<http://allserv.ugent.be/~svaelst/courses/medstat.html>

- *Bayes, Thomas
Essay towards solving a problem in the Doctrine of chances
Philosophical Transactions of the Royal Society of London, Vol. 53, 370-418, 1763
<http://www.stat.ucla.edu/history/essay.pdf>

- *Brenner, Charles
Forensic Mathematics, 24 augustus 2004
<http://dna-view.com/>

- *Broeders, A.P.A.
Op zoek naar de bron - Over de grondslagen van de criminalistiek en de waardering van het forensisch bewijs
Academisch proefschrift, Kluwer, Deventer, 2003

- *Bullard, Floyd
A brief introduction to bayesian statistics
NCTM, 1999
<http://courses.ncssm.edu/math/TALKS/PDFS/BullardNCTM2001.pdf>

- *Ewing, J.A.
Detecting alcoholism: The CAGE questionnaire
Journal of the American Medical Association, 252(14): 1905–1907, 1984

- *Geman, S. and Geman, D.
Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images
IEEE Transactions on Pattern Analysis and Machine Intelligence, nr. 6, 721-741, 1984

- *Harn K. Van, Holewijn, P.J.
Inleiding in de waarschijnlijkheidsrekening
dictaat, Vrije Universiteit Amsterdam, 1999

- *Heckerman, David
A Tutorial on Learning with Bayesian Networks
Learning in Graphical Models, MIT Press, Cambridge, 1999
http://research.microsoft.com/research/pubs/view.aspx?msr_tr_id=MSR-TR-95-06

- *Korzec, A
Confirming alcoholism in drivers under influence

Academisch proefschrift, Uitgeverij Prometheus, Amsterdam, 2004

*MacKay, D. J. C

Introduction to Monte Carlo methods

Learning in Graphical Models, MIT Press, Cambridge, 1999

*Markov, A. A.

Rasprostranenie zakona bol'shikh chisel na velichiny, zavisyaschie drug ot druga
Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete, 2-ya
seriya, tom 15, 135-156, 1906

*Markov, A. A.

*Extension of the limit theorems of probability theory to a sum of variables connected
in a chain*

reprinted in Appendix B of: R. Howard. Dynamic Probabilistic Systems, volume 1:
Markov Chains. John Wiley and Sons, 1971

*Metropolis, N., Rosenbluth, A., Rosenbluth, M.N., Teller, A., and Teller, E

Equations of state calculations by fast computing machines

Journal of Chemical Physics, nr. 21, 1087-1092, 1953

*Mitchell, Tom M.

Machine Learning, Bayesian learning 154-200

McGraw-Hill, Singapore, 1997

*Neal, R. M.

Bayesian Learning for Neural Networks

Springer-Verlag, New York, 1996

*NHG

Standaard Problematisch alcoholgebruik

NHG 2003

<http://nhg.artsennet.nl/upload/104/standaarden/M10/std.htm#noot2>

*O'Hagan, A.

Monte Carlo is fundamentally unsound

The Statistician, nr. 36, 247-249, 1987

*Rudin, N., Inman, K.

Articulating Hypothesis – The null hypothesis and beyond

News of the California association of criminalists, 1st Quarter, 7-8, 2003

*Savage, Leonard J.

The foundations of statistics

Wiley, New York, 1954

*Sjerps, Marjan

Forensische statistiek

Nieuw archief voor de wiskunde, nr.5, 4-9, 2004

<http://www.kennislink.nl/web/show/?id=111885>

*Sjerps, Marjan

Statistiek in de rechtszaal

Vakpagina Wiskunde, Kennislink, 28 mei 2004

<http://www.kennislink.nl/web/show/?id=111865>

*Stoney, D.A., Aitken, C.G.G.
The use of statistics in forensic science
Ellis Horwood Limited, Chichester, 1991

*Traat, Imbi
Bayesian inference with Markov chain Monte Carlo, Najaar 2002
<http://www.ms.ut.ee/ained/SoomeMCMC/MCMCkava>.

Aanvullende literatuur

*Bayesian songbook
<http://www.biostat.umn.edu/%7Ebrad/cabaret.html>

*Centraal Bureau Rijvaardigheidsbewijzen
<http://www.cbr.nl>

*de Rechtspraak
<http://www.rechtspraak.nl>

*Van Dale - online woordenboek
<http://www.vandale.nl>

*Wikipedia - De vrije encyclopedie
<http://nl.wikipedia.org/wiki/>
<http://en.wikipedia.org/wiki/>