Statistiek in de rechtszaal

Het proces van Lucia de B

Willem R. van Zwet

HOVO Leiden, november 2011

V=27 verplegenden

D=1029 diensten van 8 uur (d.w.z. 343 dagen)

D_L=142 diensten van Lucia

I=8 incidenten (plotseling en onverwacht overlijden of reanimatie)

I_L : incidenten tijdens Lucia's diensten.

Het feit dat $I_L = 8$ lijkt er op te wijzen dat de kans op incidenten toeneemt als Lucia aanwezig is. Dit kan men statistisch onderzoeken. Als ook bewezen zou zijn dat de 8 incidenten (pogingen tot) moord zijn, dan zou het gegeven $I_L = 8$ Lucia zeer verdacht maken. Zonder bewijs van 8 moorden zegt het echter niets. Het feit dat $I_L = 8$ kan ook niet bijdragen tot het bewijs dat de incidenten (pogingen) tot moord zijn.

H: hypothese dat kans op incident niet afhangt van Lucia's aanwezigheid (nul-hypothese).

A: alternatieve hypothese dat de kans op een incident groter is als Lucia aanwezig is.

Verwerp H als $I_L \geq k$, waarbij de kritieke waarde k de kleinste waarde is waarvoor $P(I_L \geq k \mid H) \leq \alpha$.

Onbetrouwbaarheid α klein (bijv. $\alpha = 0{,}01$ of $0{,}001$).

Men doet dit vaak door de overschrijdingskans

$$p = P(I_L \geq 8 \mid H)$$

van waargenomen waarde van I_L (in ons geval 8) te berekenen en H te verwerpen als $p \leq \alpha$.

Voorlopig nemen wij onder H ook aan dat de kans op een incident tijdens alle diensten dezelfde is.

k	$P(I_L = k \mid H)$	$P(I_L \geq k \mid H)$
0	0,3035037	1
1	0,3917957	0,6964959
2	0,2194679	0,3047002
3	0,0696723	0,0852323
4	0,0137095	0,0155600
5	0,0017121	0,0018506
6	0,0001325	0,0001384
7	0,0000058	0,0000059
8	0,0000001	0,0000001 = p

Getuige deskundige Elffers vermeldt de overschrijdings-kans $p = P(I_L \geq 8 \mid H)$ en vindt net als hierboven:

$$P(I_L \geq 8 \mid H) = (142/1029) \times (141/1028) \times \ldots \times (135/1022)$$
$$= 1.1 \times 10^{-7} = 0{,}00000011$$

correctie: $\times 27 = 3 \times 10^{-6} = 0{,}000003$ **(en ook foutje)**

Elffers: "Dit is geen toeval! De rest is aan de rechter!" Daar is niets mis mee, maar het was niet verstandig. De rechter legde dit vermoedelijk uit als: de kans dat Lucia onschuldig is aan moorden, is 0,000003.

Later veel pogingen om deze overschrijdingskans onderuit te halen in plaats van de betekenis ervan.

Rechter combineert voor ieder incident

- **Verklaringen van getuigen**
- **Dagboek**
- **Toxicologische bevindingen**
- **Het statistische argument.**

en concludeert dat Lucia 8 (pogingen tot) moord heeft gepleegd en veroordeelt haar tot levenslang.

Rechter onderscheidt niet de twee vragen:

- **Waren de incidenten (pogingen tot) moord ?**
- **Zo ja, is Lucia de dader?**

Rechter redeneert circulair als hij de statistiek gebruikt om de moorden aan te tonen die nodig zijn om de statistiek zin te geven.

Hoger beroep bij Gerechtshof Den Haag)

Meester, Van Lambalgen:

“De analyse van Elffers heeft weinig betekenis”

- Onze aandacht werd getroffen doordat er iets heel onwaarschijnlijks gebeurde op een bepaalde afdeling van een bepaald ziekenhuis. Maar, er zijn vele afdelingen in vele ziekenhuizen. Correctie door vermenigvuldigen met N=27 verpleegsters is onzin. Vermenigvuldig met aantal verpleegsters in Den Haag, Nederland, de wereld, …??

- De redenering is te simpel. Bijv. verschillende kans op incident overdag en ‘s nachts, of voor bepaalde verpleegsters? Of krijgt Lucia juist de ernstige gevallen?

Het Hof vond (terecht) dat het met andere ziekenhuizen niets te maken had en sloot alle alternatieve verklaringen uit !!! Zonder zich nog verder over kansen uit te laten of zelfs maar het woord “statistiek” te gebruiken - maar wel steeds op de aanwezigheid van Lucia te wijzen – gaf het Hof Lucia weer levenslang. De Hoge Raad bevestigde dit deel van het vonnis.

Statistiek zonder kansen!

De Vos : Bayes analyse (NRC-Handelsblad)

H*: Hypothese dat Lucia onschuldig is;

A*: Hypothese Lucia schuldig is.

$P(A^*) = (1/4)\times 10^{-5}$ **: a priori kans Lucia schuldig**

$\Longrightarrow$

$P(I_L=8) = P(I_L=8 \mid H^*) \times P(H^*) + P(I_L=8 \mid A^*) \times P(A^*)$

$= 1{,}1\times 10^{-7}\times(1- (1/4)\times 10^{-5}) + 1 \times (1/4)\times 10^{-5} \approx (1/4)\times 10^{-5}$

$\Longrightarrow$

$P(H^* \mid I_L=8) = P(H^* \,\&\, I_L=8) / P(I_L=8) \approx$

$\{1{,}1 \times 10^{-7}\} / \{(1/4) \times 10^{-5}\} = 0{,}044$ **: a posteriori kans.**

H* niet verwerpen.

Maar $P(I_L=8 \mid K^*) = 1$ **en Elffers'** $P(I_L=8 \mid H) = 1{,}1 \times 10^{-7}$

betekenen alleen iets als 8 moorden zijn bewezen!

Voorwaardelijke Bayes analyse

We gaan er van uit dat de rechter het feit M dat de 8 patiënten zijn vermoord bewezen acht. Anders is er geen probleem. Dus is de a priori kans dat een verpleegster moorden pleegt niet relevant, want we weten dat een van de 27 verpleegsters zo'n moordenaar is. De a priori kans dat dit Lucia is, is nu de voorwaardelijke kans gegeven M en het lijkt redelijk om

a priori $P(A^* \mid M) = 1/27$ te kiezen. Dan volgt

a posteriori $P(H^* \mid M \,\&\, I_L{=}8) = 26 \times 1{,}1 \times 10^{-7}$

(vergelijk Elffers' $27 \times 1{,}1 \times 10^{-7}$ en Stigler over het O.J. Simpson proces.)

Vanaf 2004 discussie binnen de Vereniging voor Statistiek. Om mijn mening gevraagd, schreef ik in juni 2004 in een kort artikel ondermeer dat je drie vragen moet onderscheiden.

(1) Zijn er inderdaad 8 patienten vermoord?

Geen rol maar een voorwaarde voor statistiek.

(2) Zo ja, wie heeft het gedaan?

Hier kan de statistiek een bijdrage leveren mits (1) overtuigend bevestigend is beantwoord.

(3) Kloppen de gegevens? Waar komen die verdachte gevallen vandaan?

De ervaring leert dat dit vaak fout gaat.

Vraag 1: **Zijn er patienten vermoord?** Aanvankeljk werd in alle gevallen **“natuurlijke dood”** gerapporteerd. Later werd bij 1 patiëntje digoxine-vergiftiging bewezen geacht. Twee van de belangrijkste experts ter wereld achten deze conclusie ongerechtvaardigd. Dit blijkt al uit een tijdschrift artikel uit 1989 dat steeds wordt aangehaald, maar waarvan niemand het belang leek in te zien. Bij een tweede patiëntje werd vergiftiging met chloral-hydraat bewezen geacht. Lucia’s gedrag werd steeds verdacht geacht, maar beide gevallen werden door Derksen grondig onderuit gehaald.

In de overige gevallen berust het bewijs op het dagboek (compulsie), leugenachtigheid en het statistische argument van Elffers **!!** Dit laatste is circulair redeneren.

Vraag 2: Zo ja, wie heeft het gedaan ?

Als zou zijn aangetoond dat acht patienten vergiftigd zijn, dan zou de statistische analyse van Elffers het redelijk aannemelijk maken dat Lucia de dader zou zijn.

Als alleen twee moorden zouden zijn aangetoond en de rest m.b.v. schakelbewijs, dan valt het statistische argument om. Tweemaal in de buurt zijn bij moord en zesmaal als er niets aan de hand is, zegt helemaal niets.

Nu het bewijs van moord ook in die twee gevallen faalt, blijft eigenlijk alleen het dagboek en de leugenachtigheid over, wat niet voldoende is voor een veroordeling.

Wat blijft is de vraag hoe die I_L=8 tot stand kwam.

Vraag 3: Kloppen de gegevens? Waar komen die verdachte gevallen vandaan?

In 2004 eindigde ik mijn betoog met:

“Tot slot nog één opmerking die door geen der betrokkenen wordt gemaakt. Toen men in het ziekenhuis verdenking tegen LdB had opgevat, ging men alle gevallen van onverwacht overlijden of reanimatie na. Bij zo’n onderzoek ligt het levensgrote gevaar op de loer dat de onderzoekers al overtuigd zijn van wat zij moeten vinden. Als er dus een onverklaarbaar overlijden wordt gevonden waarbij LdB aanwezig was, dan wordt dit verder onderzocht en vindt zo’n geval al gauw een plaats onder de moorden. Als daarentegen LdB niet aanwezig was, dan zal ongetwijfeld de neiging bestaan om dit geval als natuurlijke dood te classificeren. Dit is een overbekend verschijnsel in de statistiek en ik hoop maar dat iemand ook echt met de onderzoekers gepraat heeft om uit te vinden of zoiets mogelijkerwijs kan zijn gebeurd.”

Dit vermoeden is bewaarheid. De gegevens zijn op een onverantwoorde manier verzameld met vooringenomenheid tegen Lucia. De definitie van "incident" is onduidelijk en lijkt sterk gerelateerd aan "Lucia aanwezig" (De vijf incidenten van agent Spaans). Dan is het niet verbazend dat de twee vaak samengaan. Elffers had dus gelijk: het was geen toeval maar er was een alternatieve verklaring die het Hof uitsloot !!

Deze manier van werken is niet onbegrijpelijk als het gaat om een politie-onderzoek naar een mogelijke dader, maar de zo verzamelde gegevens zijn niet bruikbaar voor een objectieve statistische analyse.

Dit proces is een droeve zaak hoewel het openbaar ministerie dit soort missers steeds als "incidenten" afdoet. De getuigen deskundigen waren niet allemaal competent en zouden beter gekozen moeten worden.

Zaken als statistiek (en moleculaire biologie: DNA) zijn voor de rechterlijke macht kennelijk een struikelblok. De Juridische Faculteiten moeten hier in de opleiding iets aan doen.

C.P. Snow (1959) "The two cultures"

Tweede wet thermodynamica en Shakespeare.

De Noo, Metta (2010)

Er werd mij verteld over Lucia de B. *De klok, de klepel en de klokkenluider*. Aspect.

Derksen, Ton (2006)

Lucia de B. *Reconstructie van een gerechtelijke dwaling.* Veen magazines.

De Berk, Lucia (2010)

Lucia de B. *Levenslang en tbs.* Arbeiderspers