

Het dilemma van het bewijs

J.Sittig¹

Het volgende is een bewerking van gedeelten uit een publicatie van 25 jaar geleden: J. Sittig, Rechtszekerheid bij onzekerheid, Statistica Neerlandica 1969, p.213-225.

Bij het scheiden van goede en slechte produkten maakt men gebruik van een statistisch keuringsrecept waarbij de waargenomen waarde van een produkteigenschap wordt vergeleken met een kritieke waarde ervan.

Hierbij zijn twee foute beoordelingen onvermijdelijk: een goed produkt wordt soms afgekeurd, en een slecht produkt wordt soms goedgekeurd. In beide gevallen ontstaat schade. Men kiest de kritieke waarde van de toetsingsgrootte zodanig, dat de totale schade per eenheid van tijd geminimaliseerd wordt.

Dezelfde gedachtengang kan worden toegepast op de "keuring" van een van een strafbaar feit verdachte: de kritieke waarde van de toetsingsgrootte is dan het geheel van eisen die aan de wetmatigeheid van het bewijs worden gesteld.

Kwantificeerbaarheid en gevoeligheid worden besproken.

1. Probleemstelling

De opsporingsmethoden die de politie gebruikt om bewijzen tegen verdachten te verzamelen zijn in Nederland een onderwerp van discussie. Wat in dit opzicht kan en wat niet kan interesseert niet alleen de wetgever maar ook de burger.

Of bepaalde opsporingsmethoden - denk bijvoorbeeld aan het af luisteren met behulp van elektronische apparatuur - toelaatbaar zijn of niet, wordt bepaald door de bestaande wettelijke voorschriften en de jurisprudentie; die zijn op hun beurt de neerslag van de bestaande (of vroegere) opvattingen over goed en kwaad. Als deze opvattingen veranderen, zullen - met de nodige vertraging - nieuwe regels ontstaan die aansluiten bij de veranderde situatie. In het

¹J.Sittig, P.C.Boutenssingel 12, 2902 BH Capelle a/d IJssel
tel. 010-4507099 b.g.g. 071-613193

volgende zal worden gepoogd factoren aan te wijzen, die tot verandering van de opvattingen, en dus ook van de wetten, leiden, en te bepalen in welke richting deze factoren werken. Het uitgangspunt van de schrijver is niet juridisch-ethisch, maar besliskundig-statistisch.

2. Het dilemma

Wij beschouwen een minder emotioneel geladen situatie.

Stel dat in Nederland het gouden tientje weer zou worden ingevoerd. Door de ontwaarding van het geld wordt het nu het "gouden honderdje".

Stel verder dat valsemunten namaak in circulatie brengen, waarbij de valse munten op het oog niet van de echte zijn te onderscheiden; maar zij zijn gemiddeld lichter.

Door onnauwkeurigheden in het fabricageproces, door slijtage en door meetfouten zijn de waargenomen gewichten van echte zowel als van valse munten stochastische grootheden met verdelingen die op het gemiddelde na aan elkaar gelijk zijn.

De twee verdelingen overlappen elkaar: sommige echte munten blijken na weging lichter dan sommige valse.

De procedure is nu de volgende: Kies een kritieke waarde $x(t)$; als het resultaat van de weging gelijk is aan of groter is dan $x(t)$ beschouwt men de munt als echt, in het omgekeerde geval als onecht.

Klaarblijkelijk zijn er vier mogelijkheden:

	Beoordeling: echt	Beoordeling: vals
munt is echt	A	B
munt is vals	C	D

Het dilemma is nu, dat - hoe men $x(t)$ ook kiest - de ongewenste schadelijke situaties B en C zullen blijven optreden; door de keuze van $x(t)$ kan men alleen de frequentie waarmee B en C optreden beïnvloeden: een grotere $x(t)$ laat de frequentie van de gevallen B toenemen en die van C afnemen, bij een kleinere $x(t)$ is het andersom.

fig.1 Keuren van gouden munten

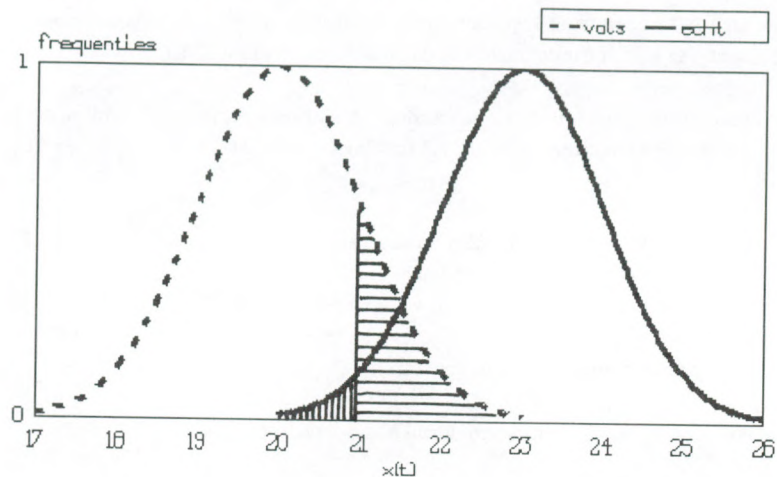
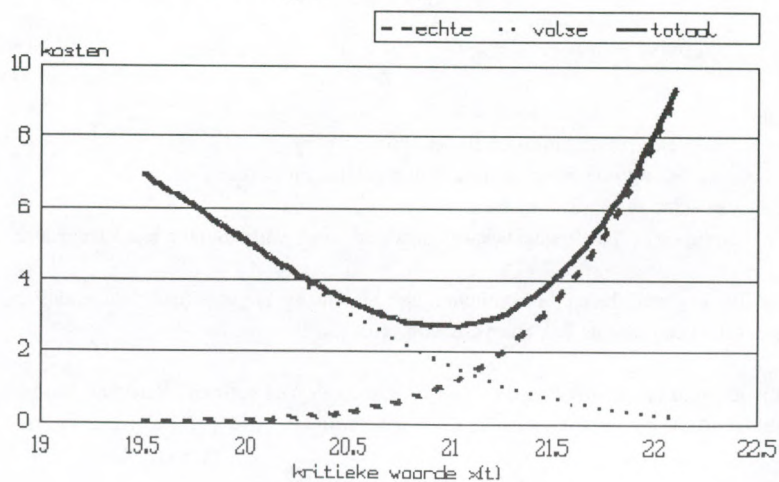


fig.2 Kosten bij keuring



3. Het optimum

Beschouwt men B en C als voor het geldverkeer schadelijke situaties dan ligt het voor de hand $x(t)$ zodanig te kiezen dat de totale schade zo klein mogelijk wordt.

Telkens wanneer B optreedt, ontstaat een schade, zeg $s(e)$: een echte munt wordt als vals behandeld. Wanneer C optreedt ontstaat schade $s(v)$: een valse munt wordt als echt beschouwd.

Kunnen wij de schade in beide gevallen uitdrukken in dezelfde optelbare grootheid, bijvoorbeeld geld, dan zullen wij nu als opdracht kiezen: bepaal $x(t)$ zodanig dat, gegeven de twee frequentieverdelingen (fig.1), de som van de kosten (bijvoorbeeld per jaar) wegens optreden van B en wegens optreden van C lager is dan bij enige andere waarde van $x(t)$. Dit is het maatschappelijke optimum van $x(t)$ (zie fig.2).

Deze optimale waarde van de toetsingsgrootheid kan worden berekend.

In fig.1 is $x(e)$ het waargenomen gewicht van een echte munt, $x(v)$ dat van een valse. Alle munten waarvan het waargenomen gewicht kleiner is dan de kritieke waarde $x(t)$ worden als vals behandeld, alle overige als echt. Munten in de gearceerde gebieden worden dus fout beoordeeld; zij hebben schade tot gevolg.

De wiskundige behandeling leidt tot een eenvoudige voorwaarde voor de optimale waarde van $x(t)$, nl.

$$y(e)/y(v) = s(v)n(v) / s(e)n(e) \quad (1)$$

Hierbij zijn:

$y(e)$ en $y(v)$ de ordinaten in het optimale punt,

$s(e)$ en $s(v)$ de per munt optredende schade bij verkeerde

beoordeling van echte resp. valse munten,

$n(e)$ en $n(v)$ de frequenties waarmee echte resp. valse munten worden gekeurd.

Zoals men uit (1) ziet, hangt het optimum niet af van de afzonderlijke frequenties en schadebedragen, maar slechts van hun verhoudingen.

Als nu bijvoorbeeld ceteris paribus $s(v)$, dus de schade wegens het niet ontdekken van een valse munt toeneemt, volgt uit de formule dat ook de ordinaat van $x(e)$ moet toenemen, met

andere woorden dat het nieuwe optimum optreedt bij een grotere $x(t)$: de toetsing is strenger geworden, er wordt meer afgekeurd.

Hetzelfde geldt blijkens (1) voor de toeneming van $n(v)$, het aantal valse munten in circulatie.

4. Bewijs in het strafrecht

Er is een grote mate van overeenkomst tussen onze gouden honderdjes en de van een strafbaar feit verdachte. Veroordeling enerzijds, sepot of vrijspraak wegens gebrek aan bewijs (of onvoldoende bewijs) anderzijds zijn de mogelijke uitslagen van de procedure. De variabelen hebben overeenkomstige betekenis als in het voorbeeld. Maar wat betekent nu de verplaatsing van de "kritieke waarde" $x(t)$?

$x(t)$ moeten wij nu opvatten als het pakket van eisen aan de wettigheid van de bewijsvoering gesteld. Een verhoging van $x(t)$ komt nu neer op een verlaging van de drempel die ter veroordeling moet worden overschreden, dus een relaxatie van de voorwaarden die aan het toegelaten bewijs worden gesteld.

Evenals in het voorbeeld van de munten verhoging van $x(t)$ tot meer afkeuringen zal leiden, en wel meer terechte, maar ook meer onterechte afkeuringen, zo zal een minder stringente bewijsmethode tot meer veroordelingen van schuldigen, maar ook tot meer veroordelingen van onschuldigen leiden. Dat is jammer, maar als men zich tot doel heeft gesteld, de totale maatschappelijke schade wegens niet vervolgen of vrijspreken van de schuldige enerzijds, en het veroordelen van de onschuldige anderzijds te minimaliseren, zal men toch in richting moeten zoeken die hier is aangeduid.

Volgens de bovenvermelde formule betekent dit dat bij stijgen van $n(v)$ (het aantal delicten) en/of bij stijgen van $s(v)$ (de zwaarte van de delicten) het nieuwe maatschappelijke optimum moet worden gezocht bij een relaxatie van de vereiste bewijsvoering.

Dit resultaat komt overeen met wat velen op gevoelsgronden menen; maar: de *gustibus non est disputandum*. Door gebruik te maken van een besliskundig model zet men in ieder geval een stap in de richting naar een meer geargumenteerde gedachtenwisseling.

In het midden latende of het mogelijk is de schadesoorten in een geschikte eenheid zoals geld te meten (de schrijver is in dit opzicht optimistisch) kunnen wij in ieder geval ervan uitgaan dat het slechts de verhouding van twee schadesoorten en de verhouding van twee frequenties is die de optimale ligging van $x(t)$ bepalen.

5. Kwantificeerbaarheid en gevoeligheid

Het zoeken van optimale oplossingen in de maatschappij - evenals in het bedrijf, waar het al lang wordt toegepast - is *mogelijk* als voor de variabelen die de ligging van het optimum bepalen, getallenwaarden kunnen worden gevonden (kwantificeerbaarheid). Het is *zinvol* als niet-optimale (kleinere of grotere) waarden van de beslissingsvariabele tot totale kosten leiden die aanzienlijk hoger zijn dan in het optimale punt (gevoeligheid).

Wat de kwantificeerbaarheid in het geval van de optimale $x(t)$ bij het bewijs in strafzaken betreft, is reeds in de vorige paragraaf gesteld, dat zelfs als rechtstreekse kwantificering moeilijk uitvoerbaar is, de voorwaarden voor de optimale $x(t)$ zodanig zijn dat slechts de verhouding van twee schadesoorten $s(e)$ en $s(v)$ een rol speelt alsmede de verhouding van de frequenties waarmede procedures tegen schuldige en onschuldige verdachten aanhangig worden gemaakt. Men hoeft dus niet de twee vragen te beantwoorden: Welke schade loopt de maatschappij als de van een bepaald delict verdachte schuldig is, maar niet gestraft wordt? en: Welke schade ontstaat als een onschuldige, van hetzelfde delict verdachte, veroordeeld wordt? De vraag waarop een antwoord moet worden gegeven is slechts: Hoeveel maal groter is de schade in het tweede geval dan in het eerste?

Hierbij komt nog, dat de gevoeligheid van de optimale $x(t)$ voor verandering in de "invloedgrootheden" betrekkelijk gering is. In het voorbeeld dat ten grondslag ligt aan fig. 1 en 2 is de verhouding van de twee schadesoorten gelijk aan 2.5, de verhouding van de twee frequenties gelijk aan 2. De verdelingen zijn normaal, met gemiddelden van 20 en 23 g, en met standaardafwijkingen van 1 g. Hierbij behoort dan een optimale $x(t)$ van 20.96 g. Wanneer men de schadeverhouding abusievelijk op respectievelijk de helft en het dubbele van de genoemde waarde stelt, leidt de berekening tot (schijn)optimale waarden van $x(t)$ die respectievelijk 1.1% kleiner en 1.1% groter uitvallen dan de juiste 20.96 g.

Het andere aspect van gevoeligheid is, hoe sterk de te minimaliseren grootheid (in ons geval de schade) toeneemt als men niet de optimale oplossing kiest maar een andere. Stel dat men in ons voorbeeld, waarin de optimale $x(t)$ gelijk is aan 20.96 g met waarden voor $x(t)$ zou werken die respectievelijk 1% groter en kleiner zijn dan de genoemde 20.96 g. De optredende totale kosten zouden dan met respectievelijk 6.5% en 5.6% toenemen.

De twee aspecten van gevoeligheid samenvattende komen wij tot de conclusie dat overschatting met 100% (van de het moeilijkst te kwantificeren invloedsvariabele) tot vergroting van de maatschappelijke schade met ongeveer 6% leidt.

Indien deze conclusie op het eerste gezicht verrassend is, dient men te bedenken dat de verandering van de totale schade bij een niet-optimale waarde van de beslissingsvariabele het saldo is van een positieve en een negatieve verandering in de twee schadesoorten.

6. Slot

De schrijver is van mening dat toepassing van voor andere doeleinden ontwikkelde kwantitatieve methoden op nieuwe gebieden niet alleen mogelijk is, maar aanbeveling verdient, al ware het alleen maar omdat gebruik van statistische en besliskundige modellen de discussie boven het peil van welles-nietes kan verheffen.

