Zucht, nee hè ... niet nog een paradox!

Onder de wiskundige problemen die garant staan voor een hoop verwarring en verhitte discussies [1]) is het 'drie-deurenprobleem' wel het bekendste, en vermoedelijk kent iedere lezer van dit blad het wel. Zelfs in de kolommen van NRC Handelsblad woedde een jaar of twee terug via ingezonden brieven een pittige tweestrijd tussen de voorstanders van het *maakt niet uit* standpunt en degenen die het *altijd wisselen* propageerden. De ingenieursoplossing in deze situatie is overigens de volgende. Een ingenieur die kennis heeft genomen van de standpunten van beide kampen, maar zelf ook niet precies weet hoe het zit, kiest voor de 'wisselstrategie', en wel onder het motto: 'baat het niet, het schaadt ook niet'.

Minder bekend is het volgende probleem. Het valt waarschijnlijk ook onder de klassiekers, al kennen ondergetekenden de herkomst niet. De situatie is als volgt. Iemand houdt je twee identiek ogende enveloppen voor, die elk een cheque bevatten. Je hebt inmiddels gehoord dat het bedrag op de ene cheque twee keer zo hoog is als dat op de andere. Van de envelop die je kiest mag je het op de cheque erin vermelde bedrag innen....., of de andere envelop kiezen en dan het bedrag dat daarin is vermeld innen. De vraag is nu wat je moet doen, als je zoveel mogelijk geld wilt opstrijken.

Oplossing 1. Je kiest willekeurig een envelop (met kans 50%), opent hem en ziet het vermelde bedrag, zeg 100 gulden. Je weet nu dat de andere envelop ofwel 50 gulden waard is, ofwel 200 gulden, beide situaties doen zich met kans 50% voor. Als je dus besluit het bedrag dat de andere envelop waard is te innen, is je verwachte uitbetaling 125 gulden (precies het rekenkundig gemiddelde van 50

[1]) Ook wij zijn het zelfs een avond in een café jaren geleden over talloze glazen bier uren oneens geweest: vage herinnering van één van de auteurs, de andere herinnert zich niets meer.

en 200 gulden). Conclusie: het is dus beter om van envelop te wisselen.

Bovenstaande is een mooi verhaal, maar gaat helaas toch mank. Met elementaire kansrekening valt dit aan te tonen (goede oefening), maar een eenvoudiger en overtuigend tegenargument is het volgende. Kennelijk wissel je in oplossing 1 altijd (!) van envelop, ongeacht welke waarde deze blijkt te hebben. Dan had je net zo goed meteen de tweede envelop kunnen kiezen. Doe je dat, dan zou je volgens dezelfde redenering ook altijd wisselen. En dan kom je weer bij de eerste uit. Kortom, deze strategie levert je niets meer op dan lukraak kiezen en bij je keuze blijven.

Het lijkt zelfs zo te zijn, dat er geen strategie is die meer oplevert dan willekeurig een envelop kiezen. En toch is die er! Het geheim hierachter is het inschakelen van een orakel. Het orakel is niet helderziend en speelt ook niet vals (kan bijvoorbeeld niet door de enveloppen heen kijken om erachter te komen welke bedragen er op de cheques vermeld staan). Het orakel hoeft zelfs geen mens te zijn, maar kan ook een dobbelsteen of een loterij zijn. Het belangrijkste is dat het orakel een getal produceert. Hoe het dat doet, doet er in eerste instantie niet toe, als het maar onafhankelijk van de keuze van de envelop is. Abstraheren we de situatie, dan zeggen we te maken te hebben met twee bedragen, h en l, met $h > l > 0$. Het orakel produceert een (stochastisch) getal X. De gunstige strategie is nu als volgt.

Oplossing 2. Kies een envelop en vergelijk het vermelde bedrag met X. Is dit bedrag groter dan of gelijk aan X, dan voel je je een tevreden mens en je gaat onverwijld naar de bank. Is het vermelde bedrag lager dan X, dan krijg je het gevoel dat het orakel het er beter af heeft gebracht dan jij en probeer je je onvrede te compenseren door voor de andere envelop te kiezen (wie weet word je er beter van?).

Omdat de lezer wiskundige is, introduceren we nog wat notatie. De hoogte van het eerste gekozen bedrag geven we aan met K, de hoogte van de uitbetaling met U. Het is duidelijk dat $P(K = h)$ gelijk aan $\frac{1}{2}$ is – we schaffen ook gelijk het uitdrukken in procenten af – maar wat is voor deze strategie nu $P(U = h)$? De gebeurtenis $\{U = h\}$ valt op te splitsen in de gebeurtenissen $\{K = h, X \leq h\}$ en $\{K = l, X > l\}$. De kans op de eerste van deze twee is, wegens de veronderstelde onafhankelijkheid, $\frac{1}{2}P(X \leq h)$ en de kans op de tweede is $\frac{1}{2}P(X > l)$. Bij elkaar vinden we dus

$$P(U = h) = \tfrac{1}{2}(P(X > l) + P(X \leq h)).$$

Op analoge wijze vinden we

$$P(U = l) = \tfrac{1}{2}(P(X \leq l) + P(X > h)).$$

De strategie is gunstig als $P(U = h) > P(U = l)$ en dit doet zich voor zo gauw $P(X \leq h) > P(X \leq l)$.

Merk nog even op dat we ook (bijna) hebben bewezen dat deze strategie nooit slechter is dan lukraak kiezen, omdat altijd geldt dat $P(X \leq h) \geq P(X \leq l)$. We concluderen dat de strategie echt beter is dan lukraak kiezen onder de vermelde conditie $P(X \leq h) > P(X \leq l)$, en bovendien valt er meer winst te halen naar mate de twee laatste kansen verder uit elkaar liggen. Ook valt eenvoudig na te rekenen dat voor deze oplossing geldt dat $P(U = h) = \frac{1}{2} + \frac{1}{2}P(l < X \leq h)$, waaruit we de 'winst' van oplossing 2 boven het lukraak kiezen kunnen aflezen. Nu doet zich het probleem voor dat we niet kunnen optimaliseren als h en l niet bekend zijn (anders zitten we ook gelijk in een geheel andere situatie). Een manier om er voor te zorgen dat we echt winst kunnen behalen is om X te kiezen uit een verdeling op $\mathbb{R}^+$ met strikt stijgende verdelingsfunctie, dan is $P(l < X \leq h)$ altijd positief. Een voorbeeld hiervan is de exponentiële verdeling.

Echt leuk wordt het pas met een helderziend en je goed gezind orakel. Dit produceert altijd een getal $X \in (l, h]$. Je kunt nu eenvoudig narekenen, dat $P(U = h) = 1$. Helaas bestaat zo'n orakel alleen in sprookjes.

Tot slot keren we nog even terug naar het drie-deurenprobleem en de afloop van de discussie in NRC Handelsblad. De brievenserie werd gestopt na een brief van een gezaghebbend mathematisch statisticus die het zat was [2]). Hij merkte op dat kansrekening

[2]) vage herinnering van een van de auteurs

en statistiek veel meer is dan gezwets over dit soort tegenintuïtieve probleempjes en in zijn volle breedte een belangrijke maatschappelijke bijdrage levert op veel gebieden. We zijn het daar mee eens. Kansrekening en Mathematische Statistiek hebben zich ontwikkeld tot volwassen disciplines binnen de wiskunde en kennen veel toepassingen erbuiten.

Bert van Es en Peter Spreij
vanes@science.uva.nl en spreij@science.uva.nl