Onverwachte kansen

F. van der Blij
OW & OC, RU Utrecht

Samenvatting

Aan de hand van enkele meetkundige- en getallenvoorbeelden wordt aangetoond dat kansrekening zelfs in heel eenvoudige gevallen verradelijker is dan vaak gedacht wordt.
Zinvol over kans, verwachtingswaarde etc. spreken kan pas als eerst de verzamelingen goed gedefinieerd zijn, vastgelegd is hoe deze zijn opgebouwd en welke waarschijnlijkheidsmaat er opgelegd wordt.

Summary

Probability within a geometrical context is rather suspect. This is illustrated by a couple of examples, one of them being the following: What is the probability of a triangle having one obtuse angle?
But even in discrete cases one should be careful: A box with an equal number of black and white balls is worth f 5,– and a box with an unequal number of black and white balls is worth f 3,–. What is the expected value of a box?
It may be worth f 4,– or f 3,86 or f 3,67, depending on the "probability measure".

Soms wordt kansrekening als een triviaal vak gezien. Ieder begrijpt toch dat je een kans $\frac{1}{6}$ hebt om met een dobbelsteen 6 te gooien, net zoals kans $\frac{1}{2}$ om met een gulden kop te gooien. En wat er verder volgt is eenvoudige rekenarij. Soms lijken opgaven moeilijk, omdat de echte wiskundige helderheid tussen zoveel tekst verstopt zit.
En dan zijn er alle grappenmakers, die je met opzet in de war brengen. Met een dobbelsteen gooi je 6 of geen 6. Dus heb je 50% kans om een zes te gooien! Wie daar inloopt is een domoor.
En wat denkt u van het verhaal over de plaatsingskansen van sollicitanten? Een extreem goede sollicitant heeft een kans van 95%, een middelmatige van 30% blijkt in een bepaalde situatie. Een moderne eerlijkecijferaar-ethicus zegt: soit, de extreem goede heeft ongeveer driekeer zoveel kans als de modale, dat vind ik aanvaardbaar. Maar zijn collega, opkomend voor de "underdog" zegt: de "modale" heeft een kans van 70% om niet aangenomen te worden, terwijl de extreem goede maar een kans van 5% heeft om niet aangenomen te worden. De verhouding van de kansen is dus als 1 : 14 en dat is maatschappelijk onaanvaardbaar!

Meetkunde en kans

Als wiskundedocent zult u over het bovenstaande de schouders ophalen; duidelijke γ-praat, ze rekenen niet beter. Toch bent u misschien ook wel eens geschrokken van die verhalen over de willekeurige driehoek. Op het gevaar af oude koeien uit de sloot te halen, merken we nog maar eens op dat een willekeurige driehoek òf scherphoekig òf rechthoekig òf stomphoekig is. Dus kans $\frac{1}{3}$ op een stomphoekige? Nee hoor, de eerste hoek kan scherp of stomp zijn, dus al kans $\frac{1}{2}$ op stomp. Als de eerste scherp is, kan de tweede stomp zijn, kans $\frac{1}{4}$. Als de eerste en de tweede scherp zijn kan de derde stomp zijn, kans $\frac{1}{8}$.
Alles tezamen is de kans voor een stomphoekige driehoek $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}$. Voor een scherphoekige driehoek moeten alle hoeken scherp zijn, dus kans $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$. De rechthoekigen zijn we voor het gemak maar even vergeten!
Natuurlijk is dit een onzinverhaal. Laten we de driehoek met hoeken α, β en γ voorstellen door een punt met coördinaten (α,β). (Er geldt immers $\gamma = 180° - \alpha - \beta$).
In figuur 1 zien we dat bij alle punten in de drie gearceerde driehoeken een tripel (α,β,γ) hoort, waarvan één hoek stomp is. Terwijl bij de punten van de niet gearceerde driehoek scherpe hoeken α, β en γ horen. De kans op een stomphoekige driehoek is dus $\frac{3}{4} = 0{,}75$.

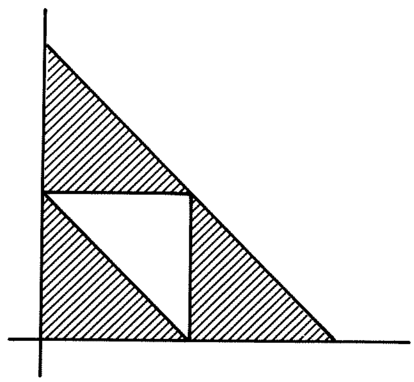

fig. 1

Maar we kunnen ook naar de zijden van de driehoek kijken. Natuurlijk mogen we zo verkleinen of vergroten dat de grootste zijde 1 is. De beide anderen zijn dan a en b met $0 \leq a \leq 1$, $0 \leq b \leq 1$. Zo'n driehoek stellen we weer voor door een punt met coördinaten (a,b). Bij de punten in de driehoek ABC vinden we tweetallen (a,b), zodat a, b, 1 de zijden van een driehoek met grootste zijde 1 zijn.
Die in het gearceerde gebied van fig. 2 behoren bij stomphoekige driehoeken. De oppervlakte van het gearceerde sectortje is $\frac{\pi}{4} - \frac{1}{2}$. De kans op een stomphoekige driehoek lijkt nu dus $\frac{\pi}{2} - 1 = 0{,}57$.

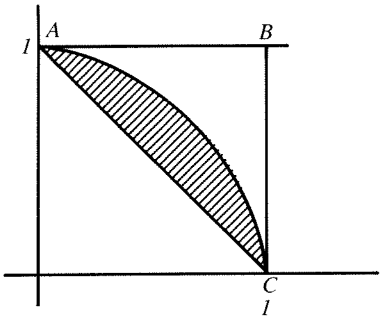

fig. 2

Hoe langer we praten hoe kleiner de kans op een stomphoekige driehoek lijkt te worden. Verzint u zelf even een vervolg op dit verhaal?
Vermoedelijk zult u mij nu voor de voeten werpen dat iedereen toch weet dat *meetkundige* kansen gevaarlijk zijn. En wellicht weet u van de klassieke discussies over de gemiddelde lengte van een koorde in een cirkel met straal 1. U kunt redeneren dat een koorde op afstand x van het middelpunt een lengte $2\sqrt{1-x^2}$ heeft. De gemiddelde lengte zal dus

$$\int_0^1 2\sqrt{1-x^2}\,dx = 2\int_0^1 \sqrt{1-x^2}\,dx = \frac{\pi}{2} = 1{,}57.$$

zijn.

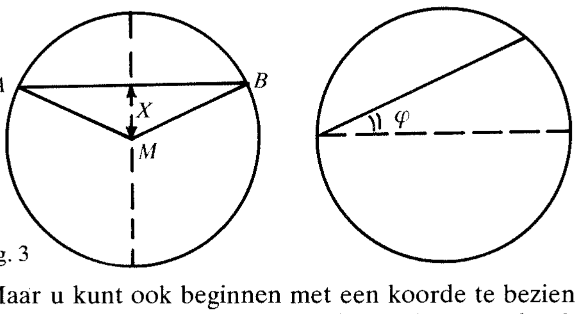

fig. 3

Maar u kunt ook beginnen met een koorde te bezien die een hoek φ met de middellijn maakt. Deze heeft een lengte 2 cos φ. De gemiddelde lengte is dus

$$\frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cdot 2\cos\varphi\, d\varphi = \frac{4}{\pi} = 1{,}27.$$

Wil de ware koorde zich even melden?

Getallen en kansen

Meetkundige verhalen bij kansrekening zijn dus op zijn zachtst gezegd suspect te noemen. Maar als je met discrete zaken zoals natuurlijke getallen werkt gaat alles toch goed? Zoals de helft van alle getallen is even. Immers van 1–10 is 50% even van 1–25 is 48% even, van 1–100 is 50% even, van 1–35 is 49,9% even. En als je de limiet voor N naar oneindig neemt voor het percentage even getallen onder 1 – N vind je heus 50%.
Helaas zijn er primitieve mensen die de getallen op een andere manier schrijven, zoals woordenboekschrijvers: acht, achtendertig, achtennegentig.
Het lijkt dan hopeloos iets over het percentage even getallen te schrijven. Hoe zit het eigenlijk met een alfabetische lijst van *alle* natuurlijke getallen? Maar zelfs een vrij volledig lijkende lijst van de natuurlijke getallen: 1, 3, 2, 5, 7, 4, 9, 11, 6, 13, 15, 8, 17, 19, 10, ... (twee oneven gevolgd door één even) geeft voor de limiet van het percentage even getallen op een rij van N getallen ons nu 33%. En u begrijpt weer dat u eruit kunt krijgen wat u wilt, door de rij waarin u alle natuurlijke getallen laat optreden, maar geschikt te kiezen. Gelukkig werken we op school alleen maar met kansen in gevallen van eindig veel objecten. Dan komen deze rariteiten niet voor. Helaas, ik moet u teleurstellen. Op een internationale conferentie, nog niet zo heel lang geleden, waren deskundigen ruzie aan het maken over een kans-probleem dat ik voor u in een eenvoudiger vorm giet.

Een doosje met evenveel witte als zwarte knikkers is *f* 5,– waard, een doosje waarin de aantallen witte en zwarte knikkers niet gelijk zijn is *f* 3,– waard. Wat is de verwachte waarde van een doosje?
Eerste oplossing: In ieder doosje zit zowel een witte als een zwarte. Daarbij komen er nog twee, dit kan zijn WW, WZ, ZW of ZZ met kansen $\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}$ en waarden *f* 3,–, *f* 5,–, *f* 5,– en *f* 3,– resp. De verwachte waarde is dus $\frac{1}{2} \times$ *f* 3,– + $\frac{1}{2} \times$ *f* 5,– = *f* 4,–.
Tweede oplossing: Van de zestien mogelijke combinaties WWWW, WWWZ enz. sluiten we er twee uit, namelijk WWWW en ZZZZ. Van de veertien overigen zijn er zes met twee zwarten en twee witten, vier met drie zwarten en één witte, vier met drie witten en één zwarte. De verwachtingswaarde is dus $\frac{6}{14} \times$ *f* 5,– + $\frac{8}{14} \times$ *f* 3,– = *f* 3,86.
Welke van de twee oplossingen is goed? Of zijn ze allebei fout, maar wat is dan de goede oplossing?
Het hangt er maar vanaf hoe je het gegeven leest. Als de fabrikant van de doosjes deze willekeurig gevuld heeft met vier knikkers en daarna de "misdrukken", d.w.z. de doosjes met enkel witten of enkel zwarten verwijderd heeft, is de tweede berekening goed. Heeft de fabrikant echter in ieder doosje eerst een witte en een zwarte knikker gedaan en er daarna willekeurig nog twee bijgevoegd, dan is de eerste oplossing de juiste. Kunt u nog andere vullingsstrategieën bedenken, die weer andere verwachtingswaarden geven? Als ik mij goed herinner had Laplace een fabrikant op het oog, die met gelijke kansen doosjes met een overwicht aan zwarten, doosjes met een overwicht aan witten en eerlijke doosjes produceerde.
Dan is de verwachtingswaarde $\frac{2}{3} \times$ *f* 3,– + $\frac{1}{3} \times$ *f* 5,– = = *f* 3,67.
Wilt u een moraal van dit verhaal? Kansrekening is zelfs in de eenvoudigste gevallen verradelijker dan u denkt! Niet alleen bij meetkundige situaties kunt u pas zinvol over kans, verwachtingswaarde etc. spreken als u eerst goed gedefinieerd hebt over welke verzamelingen u spreekt, hoe die zijn opgebouwd, en om het geleerd te zeggen, welke waarschijnlijkheidsmaat u erop legt.