Interpretaties van kansen, de *Dutch Book* stelling en het driedeurenprobleem

Mirte Dekkers en **Klaas Landsman**

mdekkers@math.ru.nl landsman@math.ru.nl

Institute for Mathematics, Astrophysics, and Particle Physics (IMAPP)
en
Genootschap voor Meetkunde en Kwantumtheorie (GQT-cluster)

Radboud Universiteit Nijmegen
Toernooiveld 1, 6525 ED Nijmegen

ALLEREERSTE CONCEPT D.D. 22 JANUARI 2007

Inhoudsopgave

In de nu lopende experimenteerfase voor wiskunde D staat 160 slu voor het domein *Statistiek en Kans*. Het is een goede mogelijkheid om het bestaande materiaal van wiskunde A1 en A12 (200 slu) in wiskunde D versneld in 120 slu te behandelen en 40 slu te besteden aan een keuzeonderwerp.

De tekst die nu volgt is bedoeld als aanzet voor een dergelijk keuzeonderwerp van 40 slu. De tekst kan met behulp van de in de bibliografie opgenomen literatuur worden aangevuld met een toepassing op de rechtszaak tegen Lucia de B. De docent kan hiertoe zelf tekst toevoegen of het als onderwerp voor een werkstukje aan de leerlingen overlaten. Intussen werken wij ook zelf aan een hoofdstuk hierover.

Het gaat hier om een ruwe eerste versie van de tekst. Commentaar (per e-mail aan `landsman@math.ru.nl`) is van harte welkom!

1

Uitkomsten en gebeurtenissen

In het onderwijs over kansrekening ben je in aanraking gekomen met het uitrekenen van kansen op verschillende gebeurtenissen. We gaan nu de algemene wiskundige regels die schuil gaan achter die berekeningen beter bekijken. In dit hoofdstuk voeren we de abstracte begrippen uitkomsten, gebeurtenissen, uitkomstruimte en gebeurtenisruimte in. Deze begrippen hebben we nodig om de wiskundige regels van de kanstheorie te formuleren. Dat gebeurt in het volgende hoofdstuk. We definiëren de begrippen aan de hand van twee voorbeelden van kansexperimentsen waar je zeker vertrouwd mee bent: het gooien met een dobbelsteen en een enquête.

De uitkomstruimte van een kansexperiment.
Elk kansexperiment heeft een aantal mogelijke ***uitkomsten***. Een opsomming van alle mogelijke uitkomsten noemen we de ***uitkomstruimte*** van het kansexperiment. We zetten de mogelijke uitkomsten tussen accolades $\{\ldots\}$; in de wiskunde is dit de gebruikelijke notatie voor een verzameling.
Een aantal vertrouwde voorbeelden van uitkomstruimten zijn de volgende:

Voorbeeld 1:
De uitkomstruimte van een worp met een dobbelsteen is $\{1, 2, 3, 4, 5, 6\}$.

Voorbeeld 2:
De uitkomstruimte van een vraag aan een willekeurige proefpersoon die slechts postitief of negatief beantwoord kan worden is $\{ja, nee\}$; we kunnen dit ook noteren als $\{+, -\}$. De uitkomstruimte voor twee van dergelijke vragen A en B is $\{ja/ja, ja/nee, nee/ja, nee/nee\}$ ofwel $\{++, +-, -+, --\}$.

Opgave 1.1

- **a)** Geef de uitkomstruimte bij twee worpen met één dobbelsteen.
- **b)** Geef de uitkomstruimte bij één worp met twee dobbelstenen.
- **c)** Onder welke aanname is het antwoord op vraag a en b hetzelfde?

Opgave 1.2

- **a)** Hoeveel uitkomsten zijn er mogelijk bij drie ja/nee vragen? Geef de bijbehorende uitkomstruimte.
- **b)** Hoeveel mogelijkheden zijn er bij vier ja/nee vragen?
- **c)** Stel we hebben een enquête met n ja/nee vragen. Druk het aantal mogelijke uitkomsten uit in n.

Opgave 1.3

Bedenk zelf een kansexperiment en geef de bijbehorende uitkomstruimte.[1]

Gebeurtenissen.
Meestal willen we niet alleen de kans op één mogelijke uitkomst weten, maar zijn we ook benieuwd naar de kans op een combinatie van verschillende uitkomsten. Bij het werpen met een dobbelsteen zijn we niet alleen geïnteresseerd in de kans op '6 gooien', maar ook in de kans op '5 of 6 gooien' en de kans op 'een even aantal ogen'. Een combinatie van één of meerdere uitkomsten noemen we een ***gebeurtenis***.

Voorbeeld:
We zagen al dat $\{++, +-, -+, --\}$ de uitkomstruimte is bij een enquête met twee vragen A en B. Een mogelijke gebeurtenis bij dit kansexperiment is dat 'de eerste vraag met ja beantwoord wordt' ofwel '$A = +$'. Deze gebeurtenis correspondeert met de uitkomsten $\{++, +-\}$. Een andere gebeurtenis is '$A \neq B$' en komt overeen met de uitkomsten $\{+-, -+\}$. Ook de situatie dat vraag A en B beide positief beantwoord worden is een gebeurtenis. Deze correspondeert slechts met één enkele uitkomst, namelijk $\{++\}$. Let op: het is dus ook mogelijk dat een gebeurtenis slechts één mogelijke uitkomst bevat.

Opgave 1.4

a) *Over de enquête met drie vragen.* Welke combinatie van uitkomsten correspondeert met de gebeurtenissen:
(a) '$A = +$'
(b) '$A = +\ en\ B = -$'
(c) '$A \neq B$'

b) *Over het werpen van een dobbelsteen.* Welke combinatie van uitkomsten correspondeert met de gebeurtenissen:
(a) 'het geworpen aantal ogen is 4.'
(b) 'het geworpen aantal ogen is even.'
(c) 'het geworpen aantal ogen is niet 3 of 5.'

Wat je aan deze twee voorbeelden heel duidelijk kunt zien is dit:

Een gebeurtenis is niets anders dan een gedeelte van de uitkomstruimte. [2]

We hebben gezien dat de uitkomstruimte van een kansexperiment bestaat uit een opsomming van alle mogelijke uitkomsten. Ook van alle mogelijke gebeurtenissen met betrekking tot een bepaalde uitkomstruimte kunnen we een opsomming maken. Een opsomming van **álle** mogelijke gebeurtenissen noemen we de ***gebeurtenisruimte***.
Het verschil tussen uitkomsten en gebeurtenissen, en daarmee ook tussen de uitkomstruimte en de gebeurtenisruimte van een kansexperiment, is erg belangrijk. Hoe sneller je er aan went, hoe beter.

Opgave 1.5

Over de enquête met twee vragen. We hebben al gezien dat de uitkomstruimte gelijk is aan $\{++, +-, -+, --\}$. Wat is de bijbehorende gebeurtenisruimte?

Let op!

We hebben al gezien dat één uitkomst op zichzelf ook een gebeurtenis is.[3]

Ten tweede is het een nuttige afspraak in de wiskunde dat elke gebeurtenisruimte ook de 'onmogelijke' gebeurtenis bevat. Bij deze gebeurtenis wordt geen enkele mogelijke uitkomst gerealiseerd. De onmogelijke gebeurtenis omvat dus geen enkele uitkomst, wat je als volgt kunt noteren: $\{\}$. In de wiskunde

1. In dit project beperken we ons tot eindige uitkomstruimten. Er bestaan natuurlijk ook kansexperimentsen waarbij de uitkomstruimte oneindig is.
2. Als je de uitkomstruimte als verzameling Z ziet, is een gebeurtenis dus niets anders dan een gedeelte van de verzameling Z. In de wiskunde wordt dit ook wel een deelverzameling van Z genoemd. Deze terminologie was in de jaren zestig heel gebruikelijk op de middelbare school, maar is sindsdien vrijwel uitgestorven.
3. Soms wordt een uitkomst daarom een ***elementaire gebeurtenis*** genoemd.

is het echter gebruikelijk dit te noteren met $\emptyset$.[4]

Ten slotte hoort ook de 'zekere' gebeurtenis, die de combinatie is van alle mogelijke uitkomsten, bij de gebeurtenisruimte. We noemen de zekere gebeurtenis Z. In het voorbeeld van de opgave is dus $Z = \{++, +-, -+, --\}$. Je ziet dus dat Z hetzelfde is als de totale uitkomstruimte.

Als je de gebeurtenissen $\emptyset$ en Z in de vorige opgave was vergeten moet je die dus nog snel even toevoegen! De gebeurtenisruimte bestaat immers uit álle mogelijke gebeurtenissen.

We geven een gebeurtenis meestal aan met een letter als G. De kans op deze gebeurtenis wordt dan genoteerd met $P(G)$. In het vervolg kom je dikwijls paren van gebeurtenissen tegen, die we dan G_1 en G_2 of A en B noemen. Het volgende begrip speelt daarbij een centrale rol.

Definitie 1.1 *Twee gebeurtenissen G_1 en G_2 sluiten elkaar uit als ze niet tegelijkertijd kunnen plaatsvinden. Voor de bijbehordende twee gedeelten van de uitkomstruimte betekent dit dat ze geen enkele overlap hebben (met andere woorden, niet één dezelfde uitkomst bevatten).*

In het voorbeeld van de enquête met twee vragen sluiten de gebeurtenissen $G_1 = \{++, +-\}$ en $G_2 = \{--\}$ elkaar uit, terwijl G_1 en $G_3 = \{++\}$ dat niet doen. Immers als iemand beide vragen van de enquête met 'ja' beantwoordt vindt zowel gebeurtenis G_1 als G_3 plaats.

Opgave 1.6

a) Welke paren gebeurtenissen in opgave 5.4 sluiten elkaar uit (en welke dus niet)?

b) We hebben gezien dat ook Z een gebeurtenis is. Bestaat er een gebeurtenis G zodat Z en G elkaar uitsluiten?

Nu weet je genoeg om over te gaan tot de wiskundige regels voor kansen!

4. Dit symbool staat voor de lege verzameling. Dat is een verzameling zonder elementen.

2

Kansverdelingen en kansfuncties

Nu je weet wat een uitkomstruimte en een gebeurtenisruimte zijn, voeren we de begrippen kansverdeling en kansfunctie in. Je zult zien dat een kansfunctie in de gevallen waar je mee vertrouwd bent, zoals het gooien met een dobbelsteen, aan bepaalde wiskundige regels voldoet. Laten we de voorbeelden los en formuleren we deze regels in meer algemene zin dan krijgen we de zogenaamde axioma's van de klassieke kansrekening.

De kansverdeling op een uitkomstruimte.
Bij een toevalsproces met bijbehorende uitkomstruimte horen kansen op elk van deze uitkomsten. Een overzichtje van alle mogelijke uitkomsten van een toevalsproces met de bijbehorende kans noemen we een *kansverdeling*. We zullen in het volgende hoofdstuk bekijken welke betekenis deze kansen precies hebben. Maar ook zonder een dergelijke studie is iedereen het er wel over eens dat de kansverdeling voor het werpen met een 'eerlijke' dobbelsteen er als volgt uit ziet:

Uitkomst	*Kans*
1	$P(1) = \frac{1}{6}$
2	$P(2) = \frac{1}{6}$
3	$P(3) = \frac{1}{6}$
4	$P(4) = \frac{1}{6}$
5	$P(5) = \frac{1}{6}$
6	$P(6) = \frac{1}{6}$

Tabel 2.1: Kansen bij het dobbelen

Opgave 2.1

Maak een soortgelijke kansverdeling voor de uitkomstruimte van een enquête onder tien personen met drie vragen op grond van Tabel 2.2 op de volgende bladzijde (of een eigen tabel).

Regels voor de kansverdeling.
In het bovenstaande voorbeeld (en talloze andere uit je schoolboeken over statistiek en kansrekening) zie je dat een kansverdeling aan twee eenvoudige regels voldoet:

1. $0 \leq P(U) \leq 1$, waarbij U een willekeurige uitkomst is;
2. Als $\{U_1, U_2, \ldots, U_n\}$ de uitkomstruimte is, dan geldt $P(U_1) + P(U_2) + \cdots + P(U_n) = 1$.

A	B	C
+	-	-
-	+	+
-	-	+
+	-	-
+	+	-
+	+	+
+	-	-
-	-	+
-	+	+
-	+	-

Tabel 2.2: Resultaten enquête onder tien personen met drie vragen

De kansfunctie op een gebeurtenisruimte.
Met behulp van een kansverdeling op een uitkomstruimte kunnen we ook de kans $P(G)$ op een willekeurige gebeurtenis G uitrekenen (zoals je je uit het vorige hoofdstuk herinnert is G niets anders dan een gedeelte van de uitkomstruimte). Het volgende zal je overbekend voorkomen en je beschouwt het misschien als een belediging van je verstand dat we hier over beginnen, maar het is onze bedoeling dat je even stil staat bij manipulaties die zo voor de hand liggen dat je er al lang niet meer over nadenkt.

Kijk bijvoorbeeld eens naar de volgende gebeurtenissen die bij het werpen met een dobbelsteen op kunnen treden.

- A: het geworpen aantal ogen is 4.
- B: het geworpen aantal ogen is 1 of 2.
- C: het geworpen aantal ogen is even.
- D: het geworpen aantal ogen is oneven.
- E: het geworpen aantal ogen is zes.
- F: het geworpen aantal ogen is minder dan zes.
- G: het geworpen aantal ogen is niet 3 of 5.

Als je het vorige hoofdstuk goed hebt begrepen, weet je nu dat de gebeurtenis A correspondeert met het deel $\{4\}$ van de uitkomstruimte en de gebeurtenis D met het deel $\{1, 3, 5\}$, enzovoort. In plaats van $P(A)$ kun je dus ook $P(\{4\})$ schrijven,[1] en uit Tabel 2.1 lees je af dat $P(\{4\}) = 1/6$.

Opgave 2.2

a) Schrijf $P(B)$ tot en met $P(G)$ in bovenstaande notatie.
b) Bereken de kans op de gebeurtenissen B tot en met G volgens je kennis van school.

Regels voor de kansfunctie
Om $P(B)$ te berekenen heb je bewust of onbewust gebruik gemaakt van de eigenschap

$$P(\{1, 2\}) = P(\{1\}) + P(\{2\}) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}. \tag{2.1}$$

Ook als het ingewikkelder wordt, blijken alle berekeningen in de kansrekening van de middelbare school te berusten op drie eigenschappen van de functie P die bij elke gebeurtenis de bijbehorende kans geeft.

Definitie 2.1 *Een functie met als domein een gebeurtenisruimte van een bepaald toevalsproces die aan de volgende drie eisen voldoet heet een **kansfunctie**:*

1. $0 \leq P(G) \leq 1$, *waarbij G een willekeurige gebeurtenis is;*
2. $P(Z) = 1$, *waarbij Z de hele uitkomstruimte (ofwel de zekere gebeurtenis) is;*
3. $P(A \text{ of } B) = P(A) + P(B)$, *waarbij A en B gebeurtenissen zijn die elkaar uitsluiten (zie vorige hoofdstuk).*

1. Je bent waarschijnlijk de notatie $P(X = 4)$ gewend waarbij X een toevalsvariabele is die het aantal ogen van de dobbelsteen representeert

Opgave 2.3

In plaats van symbolen kun je ook woorden gebruiken om de drie eigenschappen van een kansfunctie te formuleren. Eigenschap 1 wordt dan bijvoorbeeld: "De waarde van een kansfunctie is voor een willekeurige gebeurtenis is altijd groter of gelijk aan 0 en kleiner of gelijk aan 1". Formuleer op deze manier ook eigenschap 2 en 3 in woorden in plaats van symbolen.

Misschien zien de eigenschappen er ingewikkeld uit. Als je ze echter toepast op het voorbeeld van het werpen met een dobbelsteen, zijn ze volkomen vanzelfsprekend. De derde eigenschap bijvoorbeeld zegt dat je de kans op 1 of 2 gooien kunt uitrekenen door de kans op 1 en de kans op 2 bij elkaar op te tellen: dat is precies (2.1).

Opgave 2.4

a) Ga precies na in je antwoord van de vorige opgave waar je gebruik hebt gemaakt van de bovenstaande drie eigenschappen.
b) Welke gedeelten van de uitkomstruimte horen bij de gebeurtenissen A, B en 'A of B'? Bereken $P(A \text{ of } B)$. Is dit gelijk aan $P(A) + P(B)$? Waarom wel/niet?
c) Beantwoord de vorige vraag ook voor de gebeurtenissen B en C.

Een kansfunctie op de gebeurtenisruimte van een toevalsproces ligt geheel vast door de kansverdeling van de uitkomstruimte van het bewuste toevalsproces. Kennen we de kansverdeling, dan kunnen we met bovenstaande eigenschappen immers de kans op elke gebeurtenis (ofwel elk gedeelte van de uitkomstruimte) uitrekenen. Dit is precies wat je gedaan hebt bij het uitrekenen van de kansen op gebeurtenissen A tot en met G. Omgekeerd wordt de kansverdeling op de uitkomstruimte bepaald door de kansfunctie op de gebeurtenisruimte.

Voorwaardelijke kansen
Er is nog een vierde regel voor de kansfunctie, die de zogenaamde voorwaardelijke kans definieert. Als B een gebeurtenis is met kans $P(B) \neq 0$ en A een willekeurige andere gebeurtenis, dan is de ***voorwaardelijke kans*** op A gegeven B gelijk aan

$$P(A|B) = \frac{P(A \text{ en } B)}{P(B)}. \tag{2.2}$$

Je kunt formule (2.2) ook schrijven als

$$P(A \text{ en } B) = P(A|B)P(B). \tag{2.3}$$

Je hoeft bij (2.3) niet meer aan te nemen dat $P(B) \neq 0$: als $P(B) = 0$ dan staan er gewoon $0 = 0$.
Het is duidelijk dat $0 \leq P(A \text{ en } B) \leq P(B) \leq 1$, want als A én B plaatsvinden, vindt zeker ook A plaats. Daarmee volgt de eigenschap

$$0 \leq P(A|B) \leq 1. \tag{2.4}$$

Opgave 2.5

a) Ga na dat de regel (2.2) klopt bij het dobbelen: leid af dat de kans op twee ogen gegeven het feit dat het aantal gegooide ogen even is, gelijk is aan 1/3. Bedenk zelf nog een paar voorbeelden.
b) Bewijs dat $0 \leq P(A|B) \leq 1$.
c) Bewijs de ***regel van Bayes***:

$$P(A|B) = P(B|A)\frac{P(A)}{P(B)}. \tag{2.5}$$

De eerste drie eigenschappen in de definitie van een kansfunctie lijken volkomen vanzelfsprekend. De definitie (2.2) van een voorwaardelijke kans is weliswaar in overeenstemming met je intuïtie uit opgave a) van zojuist, maar deze vierde eigenschap van de kansfunctie ligt toch minder voor de hand. Over de regel van Bayes zullen we later nog meer zeggen: het lijkt een wiskundige identiteit, maar als je ziet hoe de regel in de praktijk wordt gebruikt, zul je begrijpen dat nadere discussie noodzakelijk is. We gaan daarom nu nader op de materie in en geven een afleiding van de vier eigenschappen van een kansfunctie.

3

De frequentie-interpretatie van kansen

Hoewel we de regels nu kennen, is het niet zo duidelijk wat een kans eigenlijk precies betekent. In dit hoofdstuk bekijken we een veelgebruikte manier om een betekenis te geven aan het kansbegrip, namelijk de frequentie-interpretatie.

In het vorige hoofdstuk heb je de vier regels leren kennen waar de kansrekening op is gebaseerd. Bij het formuleren van deze regels hebben we voor het gemak aangenomen dat het gegeven kansexperiment een eindig aantal uitkomsten heeft. Zelfs voor dat relatief eenvoudige geval werden de regels pas in 1933 opgesteld door de beroemde Russische wiskundige Andrei Kolmogorov (1903–1987). Hij gaf ook meteen de regels voor willekeurige kansexperimenten, maar daar gaan we hier niet op in. Daarmee axiomatiseerde hij de kansrekening, net zoals Euclides lang geleden de meetkunde had geaxiomatiseerd. Zo werden kansen in ieder geval voor een wiskundige teruggebracht tot 'iets dat aan de axioma's voldoet'.

De axioma's van de kansrekening
Zoals je hebt gezien zijn er slechts 4 axioma's waarop de hele wiskundige kansrekening gebaseerd is. Alle andere stellingen en resultaten van de kansrekening kunnen uit deze axioma's worden afgeleid. Voor het gemak lopen we nog een keer kort door de axioma's heen.

De eerste drie axioma's hebben betrekking op absolute of onvoorwaardelijke kansen. Het vierde axioma heeft betrekking op voorwaardelijke kansen. Absolute kansen worden gegeven door een *kansfunctie* P. Het domein van een kansfunctie is een lijst van gebeurtenissen, waarbij een gebeurtenis per definitie een combinatie van uitkomsten van een gegeven kansexperiment is. Definitie 2.1 en vergelijking (2.2) geven dan de axioma's van de kansrekening:

1. $0 \leq P(G) \leq 1$ voor een willekeurige gebeurtenis G;
2. $P(Z) = 1$, waarbij Z de hele uitkomstruimte is;
3. $P(A \text{ of } B) = P(A) + P(B)$, waarbij A en B gebeurtenissen zijn die elkaar uitsluiten;
4. Voor de voorwaardelijke kans $P(A|B)$ op A gegeven B geldt (mits $P(B) \neq 0$)

$$P(A|B) = \frac{P(A \text{ en } B)}{P(B)}. \tag{3.1}$$

De interpretatie(s) van de kansrekening
In de voorgaande paragraaf hebben we de axioma's geschetst waaraan een kansfunctie P moet voldoen. Als je deze axioma's eenvoudig als rekenregels ziet (en zo hebben we ze ook in het vorige hoofdstuk gemotiveerd), dan vatten ze in feite samen wat je in concrete berekeningen met kansen altijd al zonder nadenken deed. Om tot een dieper begrip van de kansrekening te komen is het echter noodzakelijk om de axioma's beter te motiveren. Dit zullen we nu gaan doen door ons af te vragen welke betekenis we eigenlijk kunnen geven aan een kans $P(A)$.

Op deze vraag blijkt geen eenduidig antwoord te zijn: sinds de achttiende eeuw zijn er verschillende interpretaties van het kansbegrip ontstaan. Hoewel deze interpretaties alle dezelfde axioma's accepteren, geven ze een andere betekenis aan het kansbegrip en kennen ze soms zelfs een andere kans toe aan een bepaalde gebeurtenis. We bekijken de twee belangrijkste interpretaties. In dit hoofdstuk behandelen we de frequentie-interpretatie, in het volgende de zogenaamde mentale interpretatie.

Objectief: de frequentie-interpretatie van kansen
In deze opvatting is de kans op een gebeurtenis A een objectieve numerieke waarde behorende bij A. Het woord 'objectief' drukt uit dat deze waarde niet afhankelijk is van de invloed of kennis van mensen, maar als het ware 'in de natuur' ligt.

In de eenvoudigste versie van de frequentie-interpretatie verrichten we een gegeven kansexperiment n keer en is de kans op A gelijk aan de relatieve frequentie van de gebeurtenis A. Om precies te zijn wordt $P(A)$ bepaald door het aantal keer $n(A)$ dat de gebeurtenis A in de serie van n experimenten optreedt te delen door het totale aantal experimenten n:

$$P(A) := \frac{n(A)}{n}. \tag{3.2}$$

Evenzo geldt

$$P(A|B) := \frac{n(A \text{ en } B)}{n(B)}. \tag{3.3}$$

Opgave 3.1

Laat zien dat zo aan de axioma's 1 t/m 4 op de vorige bladzijde voldaan is.

Merk op dat deze interpretatie alleen betrekking heeft op kansen op *herhaaldelijke* gebeurtenissen. Dat wil zeggen: gebeurtenissen waarvan we herhaaldelijk kunnen vaststellen of ze wel of niet plaatsvinden. Een voorbeeld van een herhaaldelijke gebeurtenis is: het regent op zondag. Een voorbeeld van een éénmalige gebeurtenis is: het regent op zondag 21 januari 2007.

Deze methode om kansen aan gebeurtenissen toe te kennen wordt in de praktijk zeer veel gebruikt. De vraag is echter of zo wel een objectieve eigenschap van de gebeurtenis A bepaald wordt. Stel dat we een eerlijke munt hebben. In het extreme geval dat $n = 1$ kunnen we slechts de mogelijkheden $P(\text{kop}) = 0$ of $P(\text{kop}) = 1$ tegenkomen. Maar ook als we vaker gooien en n dus groter kiezen, is het niet waarschijnlijk dat we zo ooit $P(\text{kop}) = 1/2$ vinden.

Om dit probleem op te lossen, moeten we n 'groot' kiezen. Wiskundig kunnen we de enigzins vage omschrijving 'groot' opvangen door de limiet van het aantal experimenten n naar oneindig te nemen. We krijgen dan dus

$$P(A) := \lim_{n\to\infty} \frac{n(A)}{n}; \tag{3.4}$$

$$P(A|B) := \lim_{n\to\infty} \frac{n(A \text{ en } B)}{n(B)}, \tag{3.5}$$

als deze limieten bestaan.

Opgave 3.2

Laat zien dat ook nu aan de axioma's 1 t/m 4 voldaan is.

Helaas treden ook nu problemen op. Ten eerste, hoe kunnen we een kansexperiment in de praktijk oneindig veel keren herhalen? En zelfs als we dit zouden kunnen, waarom zou dan de limiet $\lim_{n\to\infty}$ bestaan? En als deze limiet bestaat, hangt die niet af van de volgorde waarin de experimenten worden gedaan?

Opgave 3.3

Stel dat we willen bewijzen dat de kans dat een natuurlijk getal (zoals 1, 2, 3, enzovoort) even is, gelijk is aan $P(\text{even}) = 1/2$.

a) We zien het opschrijven van de natuurlijke getallen in de juiste volgorde 1, 2, 3,… als een kansexperiment. Toon aan dat de reeks kansen $n(\text{even})/n$ gelijk is aan

$$0, 1/2, 1/3, 1/2, 2/5, 1/2, 3/7, \ldots,$$

en dat de limiet voor $n \to \infty$ van deze reeks inderdaad 1/2 is.

b) Schrijf de natuurlijke getallen nu op in de ongebruikelijke volgorde

$$1, 3, 5, 2, 7, 9, 11, 4, 13, 15, 17, 5, 19, 21, 23, 6, 25, 27, 29, \ldots$$

Beargumenteer dat je zo inderdaad alle natuurlijke getallen krijgt en laat zien dat de limiet van reeks kansen $n(\text{even})/n$ voor $\lim_{n\to\infty}$ nu $1/4$ is (in plaats van 1/2).

Er moeten dus nog extra eisen aan de oneindige reeks herhalingen van het kansexperiment worden gesteld om tot een redelijke interpretatie te komen. Een dergelijke eis werd in 1931 geformuleerd door Richard von Mises (1883–1953), maar het wordt wel heel ingewikkeld op die manier. Bovendien is de relatie tussen kansen zoals bepaald door (3.4) en (3.5) en kansen zoals gemeten in een eindig aantal experimenten allerminst duidelijk.

Afgezien van dergelijke praktische en wiskundige bezwaren, is het grootste probleem met de frequentie-interpretatie dat deze niet van toepassing is op éénmalige gebeurtenissen. Toch willen we daar vaak een kans aan toekennen. Mede om die reden is de interpretatie in het volgende hoofdstuk bedacht.

4

De mentale interpretatie van kansen

Lijnrecht tegenover de objectieve frequentie-interpretatie van kansen staat de subjectieve mentale interpretatie. Deze associeert de kans $P(A)$ op een gebeurtenis A met de mate van geloof in A op basis van beschikbare achtergrondinformatie. Om een mentale kans in een getal om te zetten wordt mate van geloof vertaald in bereidheid om bepaalde weddenschappen aan te gaan. Uiteindelijk leidt dit ertoe dat de mentale interpretatie aan de axioma's van de kansrekening voldoet.

De mentale interpretatie van kansen gaat terug op Thomas Bayes (1702–1761), vandaar dat men ook vaak spreekt van de ***Bayesiaanse*** interpretatie. De kans op gebeurtenis A is in deze interpretatie afhankelijk van je kennis over de gebeurtenis A.
Een mooi voorbeeld van een mentale kans is het volgende: bij de huidige (beperkte) kennis over de banen van planetoïden is er een kans van 0.0014 dat er de komende 100 jaar één op de aarde inslaat. Onlangs is een nieuwe telescoop in gebruik genomen, waarmee de banen van planetoïden veel beter in kaart kunnen worden gebracht. De hoop is met deze nieuwe informatie de inslagkans tot 0 terug te brengen. Dit is precies in lijn met het mentale kansbegrip. De kans van 0.0014 representeert volgens deze interpretatie immers de mate van geloof die wetenschappers op basis van hun voorkennis hebben in de gebeurtenis (in dit geval de inslag van een planetoïde).
Als de voorkennis met betrekking tot deze gebeurtenis verandert, zal ook de kans die ze aan de gebeurtenis toeschrijven veranderen. De voorkennis met betrekking tot een bepaalde éénmalige gebeurtenis kan niet alleen met de tijd veranderen, zoals in het bovenstaande voorbeeld, maar kan op een gegeven moment ook van persoon tot persoon verschillen. Het zoontje van de tweede auteur gelooft bijvoorbeeld in Sinterklaas (i.e. de kans dat Sinterklaas echt bestaat is volgens Julius Landsman gelijk aan 1), terwijl die kans volgens zijn vader 0 is.

Van mentale kansen naar getallen
Wat in het bovenstaande verhaal nog ontbreekt is een argument dat een kans als *mate van geloof* vertaalt naar een kans als een *getal* tussen 0 en 1. In de praktijk wordt deze vertaling natuurlijk gemaakt: we identificeren een hoge mate van geloof met een kans van bijna één en wanneer er geen enkel geloof in een bepaalde bewering aanwezig is associeren we dit met een kans gelijk aan nul. Maar wat nu als we 'enigzins', 'redelijk veel' of 'een beetje' in het optreden van een gebeurtenis geloven? Hoe kunnen we een een getal toekennen aan deze intuïtieve uitspraken?

Een niet-wiskundige zou een 'mate van geloof' misschien helemaal niet in een getal uit willen drukken. Je zou bijvoorbeeld kunnen onderzoeken wat de emoties zijn van iemand die een bepaald geloof heeft, in de hoop dat een sterker geloof (bijvoorbeeld in de terugkeer van de Heiland) ook sterkere emoties met zich mee brengt. Maar daar kom je niet ver mee: juist iemand die iets heel sterk gelooft denkt daar al nauwelijks meer over na en toont dus ook weinig emoties. Maar precies hetzelfde geldt ook voor iemand die juist helemaal niet gelooft in de tweede komst van Jezus Christus op aarde. Bovendien zijn wij wiskundigen en zijn we nu eenmaal gewend kansen en vele andere zaken in getallen uit te drukken.

Mentale kansen en weddenschappen

Een belangrijke doorbraak in het probleem mate van geloof te vertalen naar wiskundige kansen werd tussen 1925 en 1930 bereikt in het werk van Frank Ramsey (1903-1930) en Bruno de Finetti (1906 - 1985). Zij stelden voor om je 'mate van geloof' in een (éénmalige) gebeurtenis te vertalen in de bereidheid om een ***weddenschap*** aan te gaan over deze gebeurtenis. Deze bereidheid zal natuurlijk afhangen van de uitbetaling van de weddenschap, dat wil zeggen van je winst of verlies als de gebeurtenis waarover een weddenschap wordt afgesloten plaatsvindt. We zullen een theorie opstellen waarbij de uitbetaling van een weddenschap afhangt van een zogenaamde ***wed-ratio***. Het is de bedoeling dat deze wed-ratio een maat zal zijn voor het geloof van een persoon in een bepaalde gebeurtenis.

Wed-ratio's

We gaan het bovenstaande nu precies maken met behulp van een voorbeeld. Omdat de volgende twee rollen nog vaak voor zullen komen geven we ze een naam: Beauty en Nerd. Stel Beauty wil van Nerd weten in welke mate hij gelooft dat een bepaalde éénmalige gebeurtenis G zal plaatsvinden. Als voorbeeld zullen we voor G de gebeurtenis 'Nederland wint het EK 2008' nemen.
Omdat Beauty wil weten in welke mate Nerd gelooft dat Nederland het EK zal winnen, stelt ze hem voor hierover te wedden. Nerd mag daarbij de *wed-ratio* r bepalen en daarna zal Beauty een getal I kiezen dat zowel positief als negatief mag zijn. Dat het teken van I aan Nerd niet bekend is op het moment dat hij r mag kiezen zorgt ervoor dat Nerd r niet hoger of lager kan kiezen dan zijn werkelijke mate van geloof. We komen hier zo meteen op terug. De wed-ratio r bepaalt samen met het getal I het bedrag dat Nerd zal winnen of verliezen, zodra de uitslag van de weddenschap bekend is. In de onderstaande tabel zie je hoe deze bedragen afhangen van r en I. [1]

Gebeurtenis	*Uitbetaling*
Nederland wint het EK	$I(1-r)$
Nederland wint het EK niet	$-Ir$

Tabel 4.1: Uitbetaling van de weddenschap

Je kunt dit ook zo zien:

1. Nerd kiest de wed-ratio r voor een gebeurtenis G.
2. Beauty kiest de inzet I voor G (inclusief het teken).
3. Nerd 'betaalt' Beauty rI (als $I < 0$ en $r > 0$ *ontvangt* Nerd dus geld!).
4.
   - Als de gebeurtenis G plaatsvindt, 'betaalt' Beauty aan Nerd I.
   - Als de gebeurtenis G niet optreedt, betaalt Beauty niets aan Nerd.

In het vervolg zullen we bewijzen dat een verstandige keuze van r altijd tussen 0 en 1 moet liggen. Met een verstandige keuze bedoelen we een keuze van r die voorkómt dat Nerd in een situatie verzeild raakt waarbij hij met zekerheid geld zal verliezen, onafhankelijk van de eindstand van het EK. Een dergelijke situatie wordt ook wel een ***Dutch Book*** genoemd; meer hierover in het volgende hoofdstuk.

Nemen we even aan dat Nerd daadwerkelijk zo verstandig is geweest r tussen 0 en 1 te kiezen, dan kunnen we het volgende opmerken: Als Beauty $I > 0$ kiest dan zal Nerd geld *winnen* als Nederland het EK wint en geld *verliezen* als Nederland niet het EK wint. Als Beauty $I < 0$ kiest dan zal Nerd juist geldt *verliezen* als Nederland het EK wint en geld *winnen* als Nederland het EK niet wint. Eigenlijk weet Nerd dus op het moment dat hij r kiest nog niet of hij juist vóór of tegen de Duitse overwinning gaat wedden. Dit wordt veroorzaakt door het feit dat hij op het moment dat hij r kiest niet weet of I positief of negatief is. Deze onwetendheid over het teken van I zorgt ervoor dat hij r niet hoger of lager kan kiezen dan zijn werkelijke mate van geloof.

Stel je immers voor dat Nerd op het moment dat hij r mag kiezen al weet dat I een positief getal is. Het is dan, ongeacht de mate van geloof die hij heeft, voordelig voor hem om r zo klein mogelijk te kiezen. Kiest hij bijvoorbeeld $r = -100$ dan zal hij verzekerd zijn van winst ongeacht de uitslag van het EK. [2] Anderzijds, als hij weet dat I een negatief getal is dan is het in zijn voordeel r zo groot mogelijk te kiezen. Om dit te

1. Met de uitbetaling van de weddenschap bedoelen we hier en in de volgende hoofdstukken altijd de winst of het verlies van degene die de wed-ratio kiest. In ons voorbeeld is de uitbetaling dus de winst of het verlies van Nerd.
2. Merk op: Omdat het het teken van I in dit geval bekend is voordat r wordt bepaald gaat het bewijs dat een verstandige keuze van r tussen 0 en 1 moet liggen niet meer op.

voorkomen moet Nerd de waarde r kiezen voordat Beauty het teken van I bepaalt.

Opgave 4.1

Stel dat Nerd denkt dat de gebeurtenis G een kans $P(G)$ heeft om op te treden. Toon aan dat de verwachte winst van Nerd uitgedrukt in $P(G), r$ en I volgens hemzelf gelijk is aan[3]

$$P(G)I(1-r) + (1-P(G))(-Ir) = (P(G)-r)I. \tag{4.1}$$

Hieruit volgt dat Nerds wed-ratio r gelijk moet zijn aan de kans $P(G)$ die Nerd aan G toekent. Zodra $P(G) \neq r$ kan Beauty het teken van immers I zo kiezen dat de weddenschap gemiddeld in het nadeel van Nerd zal uitvallen: als $P(G) > r$ kiest ze $I < 0$ en als $P(G) < r$ kiest ze $I > 0$. Conclusie:

De wed-ratio r die Nerd kiest is gelijk aan de (subjectieve) kans $P(G)$ die Nerd aan G toekent.

Voorwaardelijke weddenschappen
Ook voorwaardelijke kansen vallen onder dit schema. We moeten onze theorie over wed-ratio's daartoe enigszins uitbreiden.
Een *voorwaardelijke weddenschap* op A gegeven B is een weddenschap op A die alleen doorgang zal vinden als B plaatsvindt. Vindt B niet plaats, dan wordt de weddenschap afgeblazen.

We geven de wed-ratio die Nerd zal kiezen wanneer hij een voorwaardelijke weddenschap op A gegeven B aangaat opnieuw aan met r. We kunnen de uitkomst van een voorwaardelijke weddenschap op A gegeven B met wed-ratio r dan als volgt in een tabel weergeven:

Gebeurtenis	*Uitbetaling*
A en B	$I(1-r)$
(niet-A) en B	$-Ir$
niet-B	0

Opgave 4.2

Laat op soortgelijke wijze als in de vorige opgave zien dat r gelijk moet zijn aan $P(A|B)$, de kans die Nerd aan A toekent als hij weet dat B plaatsvindt.

Dat ook de mentale interpretatie van kansen voldoet aan de axioma's is een gevolg van de *Dutch Book* stelling, die in het volgende hoofdstuk wordt behandeld. In het bijzonder geldt formule (3.1) voor de voorwaardelijke kans, en daarmee de regel van Bayes (2.5):

$$P(A|B) = P(B|A)\frac{P(A)}{P(B)}. \tag{4.2}$$

Deze regel wordt door volgelingen van Bayes en de mentale interpretatie echter op een opmerkelijke manier gebruikt. De kansen $P(A)$, $P(B)$, en $P(B|A)$ worden geschat of berekend vóórdat het bekend is of B al dan niet plaatsvindt. Dit zou dus eigenlijk ook voor $P(A|B)$ moeten gelden, in welk geval (4.2) een wiskundige identiteit is.

De voorwaardelijke kans $P(A|B)$ in (4.2) wordt door Bayesianen echter geïnterpreteerd als de kans dat A optreedt *nadat bekend is dat B geldt*. Op deze manier actualiseren Bayesianen hun kansen in het licht van nieuwe kennis. Het rechterlid van (4.2) slaat dus op de tijd vóórdat B bewezen wel of niet plaatsvindt, terwijl het linkerlid van (4.2) wordt gerekend op een tijdstip dat bekend is dat B plaatsvond. Om deze reden is de regel van Bayes *zoals die in de praktijk wordt gebruikt* allerminst een identiteit, en is de wereld der kansrekenaars verdeeld in Bayesianen en anti-Bayesianen. De laatsten veroordelen het gebruik van (4.2) op de beschreven manier scherp. Wij kiezen in dit debat de kant van de Bayesianen, en zullen later in het voorbeeld van het Driedeurenprobleem zien dat hun regel dit probleem oplost.

3. Neem hier alvast het volgende aan: als Nerd's mate van geloof in een Nederlandse overwinning gelijk is aan $P(G)$, dan is Nerd's mate van geloof in een Nederlandse nederlaag gelijk aan $1 - P(G)$. Dit volgt uit de axioma's van de kansrekening. Je zult in het volgende hoofdstuk bewijzen dat onze manier om de mate van geloof te kwantificeren deze axioma's respecteert.

5

De Dutch Book stelling

We hebben in het vorige hoofdstuk een numerieke waarde aan de mate van geloof toegekend en geprobeerd aannemelijk te maken dat deze waarde een indicatie vormt voor een bepaalde mate van geloof. Eerder merkten we op dat een verstandig persoon zijn wed-ratio r altijd zo zal kiezen dat deze tussen 0 en 1 ligt. De eis dat een kans altijd een waarde tussen 0 en 1 moet hebben is één van de axioma's van de kansrekening. In dit hoofdstuk bewijzen we dat de zojuist ingevoerde wed-ratio's r de axioma's van de kansrekening respecteren (de Dutch Book stelling).

Dutch Books en coherentie
Om te bewijzen dat de wed-ratio's voldoen aan de axioma's van de kansrekening doen we een voor de hand liggende maar belangrijke aanname. We gaan er vanuit een persoon nooit tegen zijn eigen belang handelt. In ons geval betekent dat dat Nerd zijn wed-ratio nooit zo zal kiezen dat hij, ongeacht de uitkomst van de weddenschap, met zekerheid zal verliezen.
Een weddenschap hoeft zich niet per se te beperken tot één gebeurtenis G, maar kan ook over een aantal gebeurtenissen $G_1, G_2, \ldots, G_n$ gaan. Omdat Nerd in elke gebeurtenis natuurlijk een ander mate van geloof kan hebben mag bij hij bij elke gebeurtenis G_i een wed-ratio r_i kiezen. Beauty kiest bij elke G_i en r_i een getal I_i. We krijgen zo twee rijtjes getallen $r_1, r_2, \ldots, r_n$ en $I_1, I_2, \ldots, I_n$, die samen de uitbetaling van de weddenschap bepalen zodra vastgesteld is welke van de gebeurtenissen G_i hebben plaatsgevonden en welke niet.

Definitie:
Een *Dutch Book* is een weddenschap op een serie gebeurtenissen waarbij Beauty de getallen $I_1, I_2, \ldots, I_n$ zo kiest dat Nerd *ongeacht de uitkomst* geld zal verliezen.

Definitie:
Als Beauty bij de door Nerd gekozen wedratio's $r_1, r_2, \ldots, r_n$ geen mogelijkheid heeft door een slimme keuze van $I_1, I_2, \ldots, I_n$ een Dutch book af te dwingen, dan worden de wed-ratio's $r_1, r_2, \ldots, r_n$ *coherent* genoemd.

We kunnen de aanname dat niemand een weddenschap zal aangaan waarbij hij zeker is van verlies nu als volgt herformuleren: iedereen zal ervoor zorgen dat zijn wed-ratio's coherent zijn.

Kansexperimenten
We nemen nu aan dat de gebeurtenissen $G_1, G_2, \ldots$ samen de gebeurtenisruimte van een kansexperiment vormen.[1] We hebben in het vorige hoofdstuk gezien dat het toekennen van wedratio's aan gebeurtenissen dan opgevat kan worden als een functie R op de onderhavige gebeurtenisruimte. Omdat we nog niet weten dat R een kansfunctie is schrijven we R in plaats van P. We willen dit nu wel gaan bewijzen. Voor het gemak

1. Dit kunnen we altijd voor elkaar krijgen. Stel dat Nerd maar op één gebeurtenis G wedt. Dan nemen we $G_1 = G$ en $G_2 =$ niet-G.

herhalen we de axioma's in Definitie 2.1 en vergelijking (2.2):

1. $0 \leq R(G) \leq 1$;
2. $R(Z) = 1$, waarbij Z de zekere gebeurtenis is;
3. $R(A \text{ of } B) = R(A) + R(B)$, waarbij A en B gebeurtenissen zijn die elkaar uitsluiten;
4. $R(A|B) = \frac{R(A \text{ en } B)}{R(B)}$.

We concentreren ons eerst op de axioma's 1-3 en komen later terug op de betekenis van voorwaardelijke kansen.

De *Dutch Book* stelling
Het cruciale resultaat van de mentale interpretatie van kansen is als volgt (*Dutch Book* stelling):

> *Coherente wedratio's voldoen aan de eerste drie axioma's van de kansrekening.*

Bewijs:
Het bewijs zit als volgt in elkaar: we tonen voor elk axioma aan dat als Nerd zijn wed-ratio's zo kiest dat ze *niet* aan dit axioma voldoen, Beauty een mogelijkheid heeft om een Dutch Book af te dwingen. Met andere woorden: als Nerds wed-ratio's één van de axioma's schenden en Beauty handelt optimaal, dan zal hij met zekerheid geld verliezen. We gebruiken de uitbetaling U als variabele om de winst (of het verlies) van Nerd met betrekking tot de weddenschap aan te geven. Als $U < 0$ zal Nerd dus geld verliezen.

- $0 \leq R(G) \leq 1$.

  Stel $R(G) < 0$. Beauty hoeft alleen maar een $I < 0$ te kiezen om Nerd tot een Dutch Book te dwingen. Stel zij kiest $I = -10$. Er zijn twee mogelijkheden:
  1. G vindt plaats; dan geldt $U = -10(1 - R(G))$. Omdat $(1 - R(G)) > 0$, geldt $U < 0$.
  2. G vindt niet plaats; dan geldt $U = 10R(G)$. Omdat $R(G) < 0$, geldt eveneens $U < 0$.

  Stel $R(G) > 1$. Nu hoeft Beauty slechts $I > 0$ te kiezen om Nerd tot een Dutch Book te dwingen. Stel zij kiest $I = 10$. Wederom zijn er twee mogelijkheden:
  1. G vindt plaats; dan geldt $U = 10(1 - R(G))$. Omdat $(1 - R(G)) < 0$, geldt $U < 0$.
  2. G vindt niet plaats; dan geldt $U = -10R(G)$. Omdat $R(G) > 1$ geldt $U < 0$.

  Om aan de coherentie aanname te voldoen en daarmee de mogelijkheid tot een Dutch Book uit te sluiten moet dus gelden: $0 \leq R(G) \leq 1$.

- $R(Z) = 1$.
  Omdat we zeker weten dat Z zal plaatsvinden, geldt $U = (1 - R(Z))I$.
  Stel $R(Z) < 1$. Dan is $(1 - R(Z)) > 0$. Beauty hoeft slechts $I < 0$ te kiezen. Dan geldt $U < 0$ en dwingt ze een Dutch Book af.

  Stel $R(Z) > 1$. Dan is $(1 - R(Z)) < 0$. Nu hoeft Beauty slechts $I > 0$ te kiezen om er voor te zorgen dat $U < 0$ en zodoende Nerd tot een Dutch Book te dwingen.
  De coherentie aanname zorgt er dus voor dat $R(Z) = 1$.

- $R(A \text{ of } B) = R(A) + R(B)$.

Opgave 5.1

Bewijs deze eigenschap zelf.

Uit bovenstaande kunnen we concluderen dat coherente wed-ratio's aan de eerste drie axioma's van de kansrekening voldoen.
Einde bewijs

Voorwaardelijke kansen
We gaan nu axioma 4 (oftewel formule (2.2) voor voorwaardelijke kansen) bewijzen. Lees eerst even het stukje in het vorige hoofdstuk over voorwaardelijke weddenschappen opnieuw.

Om te laten zien dat Nerd zijn wedratio $R(A|B)$ gelijk moet kiezen aan $\frac{R(A \text{ of } B)}{R(B)}$ moeten we onze aanname van coherentie een beetje versterken. Er bestaan namelijk geen voorwaardelijke weddenschappen waarbij Nerd met zekerheid geld zal verliezen. Als B niet plaatsvindt zal de hele weddenschap immers worden afgeblazen en zal Nerd dus ook geen geld verliezen.

Definitie:
We noemen Nerds wedratio's *strikt coherent* als Beauty haar getallen I niet zo kan kiezen dat alleen zij kans heeft op een positieve uitkomst.

Omdat we hebben aangenomen dat Nerd altijd in zijn eigen belang handelt is het redelijk te veronderstellen dat hij zijn wed-ratio's strikt coherent zal kiezen. Hij heeft er immers geen belang bij een weddenschap af te sluiten waarbij hij geen enkele kans maakt op een positieve uitkomst. Het volgende resultaat breidt de *Dutch Book* stelling uit tot voorwaardelijke kansen.

Stelling:
Strikt coherente wed-ratio's voldoen aan axioma 4.

Bewijs:
Wederom voeren we een vereenvoudigde notatie in:
$r_1 := R(A \text{ en } B)$;
$r_2 := R(B)$;
$r_3 := R(A|B)$.
Dat wil zeggen r_1 en r_2 zijn de wed-ratio's die Nerd kiest voor respectievelijk de gebeurtenissen A en B en B. Voor de voorwaardelijke weddenschap op A gegeven B kiest hij wed-ratio r_3. We kunnen het vierde axioma met deze notatie als volgt uitdrukken: $r_3 = \frac{r_1}{r_2}$.

Stel $r_3 < \frac{r_1}{r_2}$.
Beauty denkt wederom goed na, voert een aantal berekeningen uit, en besluit de volgende drie weddenschappen met Nerd af te sluiten:

1. Ze sluit een weddenschap af over de gebeurtenis A en B waarbij ze $I = 10$ kiest.
2. Ze sluit een weddenschap af over de gebeurtenis B waarbij ze $I = -\frac{10r_1}{r_2}$ kiest.[2]
3. Ze sluit een voorwaardelijke weddenschap af over de gebeurtenis A gegeven B waarbij ze $I = -10$ kiest.

We geven de uitbetaling van de drie afzonderlijke weddenschappen en de totale uitbetaling weer in een tabel.

Gebeurtenis	*Uitbetaling 1*	*Uitbetaling 2*	*Uitbetaling 3*	*Totaal*
A en B	$10(1-r_1)$	$-\frac{10r_1}{r_2}(1-r_2) = -10(\frac{r_1}{r_2}-r_1)$	$-10(1-r_3)$	$10(r_3-\frac{r_1}{r_2})$
niet-A en B	$-10r_1$	$-\frac{10r_1}{r_2}(1-r_2) = -10(\frac{r_1}{r_2}-r_1)$	$10r_3$	$10(r_3-\frac{r_1}{r_2})$
A en niet-B	$-10r_1$	$r_2\frac{10r_1}{r_2} = 10r_1$	0	0
niet-A en niet-B	$-10r_1$	$r_2\frac{10r_1}{r_2} = 10r_1$	0	0

Omdat $r_3 < \frac{r_1}{r_2}$ geldt in alle gevallen $U \leq 0$. Nerd heeft zijn wedratio's dus niet strikt coherent gekozen.
In het geval dat $r_3 > \frac{r_1}{r_2}$ geldt een zelfde soort argument. Beauty sluit dezelfde weddenschappen af maar nu met tegengestelde tekens van de getallen I. Ook dan zal blijken dat in alle gevallen $U \leq 0$, ofwel Nerds wedratio's zijn niet strikt coherent.

Uit bovenstaande kunnen we concluderen dat strikt coherente wedratio's voldoen aan het vierde axioma van de kansrekening.
Einde bewijs

2. Herinner je dat Beauty I pas hoeft te kiezen nadat Nerd zijn wed-ratio's gekozen heeft. Ze kan dus I op deze manier laten afhangen van Nerds wed-ratio's.

6

Het driedeurenprobleem

Een bekend problemen uit de kansrekening (waarbij je intuïtie je in de steek laat) is het Driedeurenprobleem. Stel je bent beland in de finale van een spelshow en staat voor 3 gesloten deuren. Achter één van deze deuren staat de hoofdprijs: een prachtige splinternieuwe auto. Achter ieder van de andere twee deuren bevindt zich een geit. De quizmaster vraagt je één van de drie deuren te kiezen. Nadat jij je keuze gemaakt hebt, opent de quizmaster één van de overgebleven twee deuren. Achter de deur die de quizmaster opent zit een geit. De quizmaster heeft bewust een deur met een geit geopend, hij weet immers waar de hoofdprijs zich bevindt en deze moet nog verborgen blijven. Nu geeft de quizmaster je de keuze om alsnog van deur te wisselen. Wat moet je doen om de meeste kans te maken op de hoofdprijs: bij je oorspronkelijke keuze blijven of wisselen van deur?

Je inuïtie zegt waarschijnlijk dat het niet uitmaakt of je van deur wisselt of niet: er zijn nog twee gesloten deuren en op het eerste gezicht lijkt het alsof voor elk van beide deuren de kans $\frac{1}{2}$ is dat de hoofdprijs zich achter deze deur bevindt. Deze redenering blijkt echter niet juist te zijn. Het maakt wel degelijk uit of je wisselt van deur of niet. Als je wisselt van deur verdubbel je je kans op de hoofdprijs: deze stijgt van $\frac{1}{3}$ naar $\frac{2}{3}$. We merken hier al op dat het bij deze uitkomst van groot belang is dat de quizmaster weet waar de hoofdprijs zich bevindt en bewust een andere deur opent. We komen hier later nog op terug.

Geschiedenis van het probleem
Er zijn in de wiskunde weinig problemen waarover zoveel onenigheid is geweest als over het Driedeurenprobleem. In Amerika staat dit probleem bekend als het *Monty Hall problem*, vernoemd naar de presentator van de spelshow *Let's make a deal* waarin de hierboven beschreven situatie zich daadwerkelijk voordeed. Het probleem kreeg grote bekendheid toen het in 1990 verscheen in het Amerikaanse tijdschrift *Parade Magazine*. Marilyn Savant Vos schreef erover in haar column 'Ask Marilyn', waarin ze wekelijks wiskundevragen van lezers beantwoorde. Alhoewel Marilyn volgens het *Guiniss Book of Records* de intelligentste vrouw ter wereld was, twijfelden velen aan haar bewering dat je je kans op de hoofdprijs verhoogt van $\frac{1}{3}$ naar $\frac{2}{3}$ door van deur te wisselen. Zelfs na haar uitleg bleven de brieven van mensen, waaronder voorname wetenschappers en wiskundigen, die het tegendeel beweerden, binnenstromen. Enkele reacties waren:

> "Ik maak me grote zorgen over het gebrek aan wiskundig inzicht bij het grote publiek. Help alstublieft door uw fout toe te geven."

> "Ongelooflijk dat u uw fout nog steeds niet inziet nadat u door zeker drie wiskundigen bent verbeterd."

> "U hebt het volledig mis... Hoeveel woedende wiskundigen zijn ervoor nodig om u te overtuigen."

Zowel in populair wetenschappelijke als serieuze wiskundige tijdschriften verschenen regelmatig artikelen over het Monty Hall probleem.
Ook in Nederland zorgde het probleem voor veel opwinding toen er in 1995 een artikel over verscheen in het NRC Handelsblad. De krant kreeg massa's brieven binnen en eindigde uiteindelijk de discussie met de woorden:

> "Stop, stop, stop met brieven sturen. Het onbegrip tussen het gezond verstand en de wiskundigen is kennelijk onoverbrugbaar."

Opgave 6.1

> Kijk op het internet onder Driedeurenprobleem, Monty Hall Problem of Ziegenproblem (de Duitse naam). Geef een samenvatting van wat je gevonden hebt.

We gaan de juiste oplossing van het probleem nu op twee manieren afleiden: eerst met behulp van de frequentie-interpretatie en vervolgens door middel van de mentale interpretatie van kansen. De toepasbaarheid van de frequentie-interpretatie lijkt twijfelachtig omdat het hier om een unieke gebeurtenis lijkt te gaan. Als de quiz echter vaak op de tv wordt vertoond en de deelnemer al die afleveringen heeft gezien, komen we toch in het domein van deze interpretatie. De mentale interpretatie is zonder meer uitstekend geschikt om dit probleem op te lossen. Gelukkig leiden beide interpretaties tot hetzelfde antwoord.

De frequentie-interpretatie
Stel we spelen de finale van de spelshow 300 keer achter elkaar. Omdat je geen idee hebt achter welke deur de hoofdprijs zich bevindt, zal je ongeveer 200 keer een deur kiezen met een geit. Heb je in eerste instantie een geit gekozen, dan zal wisselen van deur je de hoofdprijs opleveren. In ongeveer 200 van de 300 gevallen zal je dus winnen als je van deur wisselt. De frequentietheorie zegt dan dat de kans op winst bij wisselen ongeveer $2/3$ is. Als je daarentegen niet wisselt, win je alleen als je keuze reeds goed was. Dat zal in ongeveer 100 gevallen zo zijn. De kans op winst als je bij je keuze blijft ligt dus rond $1/3$.

Deze oplossing bevat toch nog een aantal onderliggende aannamen. De belangrijkste is dat de deelnemers toevallig kiezen en dat tegelijk de quizmaster de auto toevallig achter een deur zet.

Opgave 6.2

> Stel dat de deelnemers in opeenvolgende afleveringen de deuren $1, 2, 3, 1, 2, 3, 1, 2, 3$ enzovoort kiezen en dat de quizmaster (zonder deze keuzes van tevoren te kennen) de auto achtereenvolgens achter precies deze deuren zet. Laat zien dat in de frequentie-benadering de winstkans bij niet wisselen nu 1 is en bij wisselen 0.

Toch is het in het scenario van deze opgave zo dat de deelnemers iedere deur met kans $1/3$ kiezen en de quizmaster de auto ook even vaak achter iedere deur zet!

Opgave 6.3

> Formuleer precieze aannamen over het keuzegedrag van de deelnemers en de manier waarop de quizmaster de auto in achtereenvolgende afleveringen neerzet, die garanderen dat in de limiet van n naar oneindig de juiste kansen volgen.

De mentale interpretatie
Zoals al opgemerkt lijkt de mentale benadering in de onderhavige situatie beter van toepassing dan de frequentie-interpretatie. We definiëren de volgende gebeurtenissen:

- $A_1 :=$ de auto bevindt zich achter deur 1.
- $A_2 :=$ de auto bevindt zich achter deur 2.
- $A_3 :=$ de auto bevindt zich achter deur 3.
- $Q_1 :=$ de quizmaster opent deur 1.
- $Q_2 :=$ de quizmaster opent deur 2.
- $Q_3 :=$ de quizmaster opent deur 3.

Stel je hebt in eerste instantie deur 1 gekozen. We gaan nu kijken wat de informatie uit de probleembeschrijving ons vertelt over de kansfunctie op bovenstaande gebeurtenissen.[1] Omdat we nu de mentale interpretatie toepassen en kansen dus associëren met de mate van geloof, is het belangrijk op te merken dat we de kansfunctie beschouwen die *de deelnemer* hanteert. We kunnen nu uit de gegeven informatie de volgende drie conclusies trekken:

1. Omdat de deelnemer *a priori* geen idee heeft achter welke deur de hoofdprijs zich bevindt geldt:

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}. \tag{6.1}$$

2. Nadat hij deur 1 heeft gekozen, heeft de deelnemer geen idee welke van de twee overgebleven deuren de quizmaster open gaat maken:

$$P(Q_2) = P(Q_3) = \frac{1}{2}. \tag{6.2}$$

3. Omdat de deelnemer weet dat de quizmaster nooit de deur met de hoofdprijs openmaakt en ook niet de deur die hij gekozen hebt (in dit geval deur 1), geeft informatie over de plaats waar de hoofdprijs zich bevindt zekerheid over welke deur de quizmaster opent. Dit kunnen we met behulp van voorwaardelijke kansen als volgt formuleren:

$$P(Q_2|A_3) = P(Q_3|A_2) = 1. \tag{6.3}$$

4. Als de deelnemer aanneemt dat de auto achter deur 1 staat, weet hij niet welke deur de quizmaster zal openen. Dit geeft

$$P(Q_2|A_1) = P(Q_3|A_1) = \frac{1}{2}. \tag{6.4}$$

Stel de quizmaster opent deur 2. Wat is nu de kans dat de deelnemer de hoofdprijs wint als hij wisselt van deur? Hij wint als de hoofdprijs zich achter deur 3 bevindt. We willen dus weten wat de kans hierop is, *gegeven* dat de quizmaster deur 2 heeft geopend. Met andere woorden, wat is $P(A_3|Q_2)$? Met behulp van de regel van Bayes (2.5) vinden we:

$$P(A_3|Q_2) = \frac{P(Q_2|A_3)P(A_3)}{P(Q_2)} \tag{6.5}$$

Invullen van de bovenstaande gevonden waarden voor de kansfunctie geeft:

$$P(A_3|Q_2) = \frac{1 \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}. \tag{6.6}$$

Als de quizmaster deur 2 opent heeft de deelnemer dus een kans van $\frac{2}{3}$ op de hoofdprijs als hij van deur wisselt. Op eenzelfde manier vinden we $P(A_2|Q_3) = \frac{2}{3}$. Dus ook als de quizmaster deur 3 opent is het verstandiger om van deur te wisselen.

Met hetzelfde argument kun je de kans op de hoofdprijs berekenen als je *niet* van deur wisselt. Heeft de deelnemer deur 1 gekozen en opent de quizmaster deur 2 dan is deze kans gelijk aan:

$$P(A_1|Q_2) = \frac{P(Q_2|A_1)P(A_1)}{P(Q_2)} = \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}. \tag{6.7}$$

Conclusie: wisselen van deur vergroot altijd de kans op de hoofdprijs.

Wat gebeurt er als de quizmaster niet weet waar de hoofdprijs zich bevindt?
We hebben al eerder opgemerkt dat het van groot belang is dat de quizmaster weet waar de hoofdprijs zich bevindt en dus *bewust* een deur opent met een geit daarachter.

Opgave 6.4

Werk beide interpretaties uit in de situatie dat de deelnemer weet dat de quizmaster *toevalling* een deur met een geit opent (hij had dus ook de deur met de auto kunnen openen). Laat zien dat wisselen en niet wisselen dan dezelfde winstkans geven.

1. We maken ons niet druk over hoe het domein van de kansfunctie er precies uitziet. We eisen alleen dat bovenstaande gebeurtenissen er in zitten.

7
Literatuur

Interpretatie kansrekening:

D. Gillies, *Philosophical Theories of Probability* (Cambridge Univ. Press, 2000).
J. von Plato, *Creating Modern Probability* (Routlegde, 1994).
C. Howson & P. Urbach *Scientific Reasoning, The Bayesian Approach* (Open Court, 1989).
A. Hájek, *Interpretations of probability*,
    `http://www.science.uva.nl/ seop/entries/probability-interpret/`

Driedeurenprobleem:

Monty Hall paradox, `http://www.mathematik.uni-bielefeld.de/~sillke/PUZZLES/monty-hall`
The Monty Hall dilemma revisited, `http://129.3.20.41/eps/exp/papers/9906/9906001.html`
Monty Hall problem, `http://en.wikipedia.org/wiki/`
Hommeles over drie deuren, `http://www.kennislink.nl/web/show?id=159743`

Lucia de B.:

R. Meester et al., On the (ab)use of statistics in the legal case against the nurse Lucia de B.,
    `http://www.cs.vu.nl/~rmeester/preprints/`
Marjan Sjerps, Forensische statistiek, *Nieuw Archief voor Wiskunde* 5/5 nr 2 juni 2004
Marjan Sjerps, Statistiek in de rechtszaal, `http://www.kennislink.nl/web/show?id=111865`
A.F. de Vos, Rekenen aan de zaak Lucy de B.,
    `http://www.portill.nl/articles/Borgers/devos.041104.pdf`