

# **Inleiding Kansrekening en Statistiek**



# **Inleiding Kansrekening en Statistiek**

S.J. de Lange

© VSSD

**Eerste druk 1989**

**Tweede druk 1991 - 2007**

Uitgegeven door de VSSD

Leeghwaterstraat 42, 2628 CA Delft, The Netherlands

tel. +31 15 27 82124, telefax +31 15 27 87585, e-mail: [hlf@vssd.nl](mailto:hlf@vssd.nl)

internet: <http://www.vssd.nl/hlf>

URL met informatie over dit boek: **<http://www.vssd.nl/hlf/a014.htm>**

*All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photo-copying, recording, or otherwise, without the prior written permission of the publisher.*

Printed in The Netherlands.

Gedrukte versie ISBN-13 978-90-6562-095-8

Elektronische versie ISBN-13 978-90-6562-180-1

NUR 916

*Keywords:* kansrekening, statistiek

# Voorwoord

Dit boek is geschreven voor technische studenten die in één semester enige vaardigheid moeten krijgen in het hanteren van de begrippen en methoden van de kansrekening.

De zeer beperkte tijd die in de diverse studieprogramma's voor dit vak beschikbaar is heeft tot gevolg dat enige onderwerpen, die zelfs op dit niveau eigenlijk niet mogen ontbreken, niet of nauwelijks worden genoemd. Zo ontbreken voortbrengende en karakteristieke functies geheel en is de combinatoriek te vluchtig behandeld.

Daar statistiek een belangrijk toepassingsgebied is van de kansrekening worden ook de statistische standaardtechnieken behandeld.

De leer der verzamelingen is bekend verondersteld evenals de analyse die in het eerste jaar aan de TU Delft wordt onderwezen (in het bijzonder het berekenen van meervoudige integralen). De kansrekening zelf wordt echter, zij het beknopt, van de grond af aan opgebouwd.

Vooraf in de latere hoofdstukken wordt een beroep op de intuïtie en de goedgegeloofigheid van de lezer gedaan. Het boek is dan ook niet bestemd voor (aspirant-)wiskundigen. Achterin is een aantal statistische tabellen opgenomen. Voor het toepassen van de besproken technieken zijn deze tabellen voldoende.

Vanzelfsprekend zijn op- en aanmerkingen, vooral voorstellen ter verbetering, welkom.

De vraagstukken zijn voor een deel overgenomen uit *Vraagstukken over waarschijnlijkheidsrekening* van dr. P.J.A. Kanters. Ik dank hem voor zijn toestemming daartoe. Mijn dank gaat voorts uit naar ir. Th.C.A. Mensch en dr. J.A.M. v.d. Weide voor hun bijdragen bij de totstandkoming van de inhoud. Zeer veel dank ben ik verschuldigd aan mevr. F.A. Zuidervaart-Murray en de medewerkers van de VSSD voor de prettige samenwerking bij de productie van dit boek.

Delft, december 1988

S.J. de Lange

## BIJ DE TWEDE OPLAGE

In deze nieuwe oplage zijn de wijzigingen beperkt tot enkele noodzakelijke correcties.

Delft, december 1991

S.J. de Lange



# Inhoud

<b>VOORWOORD</b>	5
<b>NOTATIE, LITERATUUR</b>	10
<b>1. KANSREKENING</b>	11
1.1. Uitkomstenruimte en gebeurtenissen	11
1.2. Axioma's van de kansrekening	13
1.3. Kansruimten	16
1.4. Combinatoriek	18
1.5. Enige voorbeelden	20
1.6. Conditionele kans	21
1.7. Onafhankelijkheid	25
<b>2. STOCHASTISCHE VARIABELEN</b>	29
2.1. Kansfunctie	29
2.2. Verdelingsfunctie en kansdichtheid	31
2.3. Verwachting en variantie	35
2.4. Momenten en andere kentallen	38
2.5. Functies van een stochastische variabele	40
<b>3. VEEL VOORKOMENDE VERDELINGEN</b>	43
3.1. Bernoulli- en binomiale verdeling	43
3.2. Geometrische en hypergeometrische verdeling	45
3.3. Poisson-verdeling	48
3.4. Uniforme en exponentiële verdeling	49
3.5. Normale verdeling	52
<b>4. SIMULTANE VERDELINGEN</b>	57
4.1. Twee-dimensionale verdelingen	57
4.2. Verwachting, variantie, covariantie en correlatiecoëfficiënt	64
4.3. Onafhankelijke stochastische variabelen; voorwaardelijke verdelingen	69
4.4. Stochastische vectoren in $n$ dimensies	77
4.5. De ongelijkheid van Chebychev; Wet van de grote aantallen	81
4.6. De Centrale Limietstelling; benaderingen	83
4.7. De verdeling van functies van twee of meer stochastische variabelen	85
4.8. Convolutie	94
4.9. De negatief-binomiale verdeling	97

<b>5. STATISTIEK</b>	99
5.1. Inleiding	99
5.2. Steekproef en populatie	99
5.3. Gemiddelde en variantie van een steekproef	101
5.4. Andere steekproeffuncties: $\chi^2$ , t en F	102
5.5. Het rekenwerk	108
<b>6. SCHATTEN</b>	111
6.1. Inleiding	111
6.2. Puntchatting	111
6.3. Constructie van schatters	114
6.4. Betrouwbaarheidsintervallen	119
<b>7. TOETSEN VAN HYPOTHESEN</b>	127
7.1. Inleiding	127
7.2. Parametrische toetsen	129
7.3. Enige standaardtoetsen	132
7.3.1. Toetsen voor de verwachting bij bekende variantie	132
7.3.2. Toetsen voor de verwachting bij onbekende variantie	133
7.3.3. Toetsen voor de variantie bij bekende verwachting	133
7.3.4. Toetsen voor de variantie bij onbekende verwachting	134
7.4. Twee steekproeven	134
7.4.1. Het verschil van de verwachtingen bij bekende varianties	135
7.4.2. Het verschil van de verwachtingen bij onbekende maar gelijke varianties	135
7.4.3. Het quotiënt van de varianties bij bekende verwachtingen	135
7.4.4. Het quotiënt van de varianties bij onbekende verwachtingen	136
7.5. Twee andere toetsen	138
7.5.1. Voor de parameter van een exponentiële verdeling	138
7.5.2. Voor de parameter $p$ van een alternatief verdeelde populatie	138
<b>8. VERDELINGSVRIJE TOETSEN</b>	141
8.1. Inleiding	141
8.2. De $\chi^2$ -toets voor aanpassing	141
8.3. De tekentoets	144
8.4. De toets van Wilcoxon	145
<b>9. GEORDENDE STEEKPROEVEN</b>	149
<b>OPGAVEN</b>	155
<b>ANTWOORDEN</b>	181
<b>TREFWOORDENLIJST</b>	190



<b>APPENDIX: FORMULES EN TABELLEN</b>	193
Klein repertorium	194
Overzicht verdelingen	198
Cumulatieve binomiale verdeling	199
Cumulatieve Poissonverdeling	205
Linker-kritieke waarden $K_{1-\alpha}(n)$ van de tekentoets	207
Linker-kritieke waarden $W_{1-\alpha}(n_1, n_2)$ van de toets van Wilcoxon	208
F-verdeling	210
Chi-kwadraat-verdeling	214
Standaard-normale verdeling	215
Student-verdeling	216

## Notatie

Stochastische variabelen zijn aangeduid door onderstreping.

$\underline{x} \sim \dots$  : de stochastische variabele  $\underline{x}$  heeft de  $\dots$  verdeling.

$a \approx b$ :  $a$  is ongeveer gelijk aan  $b$ .

$P(\lambda)$ : Poisson-verdeling met parameter  $\lambda$ .

$\text{Exp}(\lambda)$ : Exponentiële verdeling met parameter  $\lambda$ , waarbij is gekozen voor die variant waarbij de verwachting  $1/\lambda$  is.

$N(a;b)$ : Normale verdeling met verwachting  $a$  en variantie  $b$ .

$B(n;p)$ : Binomiale verdeling met parameters  $n$  en  $p$ .

$NB(r;p)$ : Negatief-binomiale verdeling met parameters  $r$  en  $p$ .

## Literatuur

Er zijn veel inleidende boeken over kansrekening en statistiek. Het merendeel is in het Engels. Uit die overvloed worden er slechts enkele genoemd.

Feller, W., *An Introduction to Probability Theory and its Applications*, Vol. I and Vol. II, John Wiley, New York.

Genugten, B.B. van der, *Inleiding tot de Waarschijnlijkheidsrekening en de Mathematische Statistiek*, deel I, Stenfert Kroese, Leiden.

Hogg, R.V. and A.T. Craig, *Introduction to Mathematical Statistics*, Collier MacMillan International Editions, London.

Kanters, dr. P.J.A., *Vraagstukken over waarschijnlijkheidsrekening*, DUM, Delft.

Larson, H.J., *Introduction to Probability Theory and Statistical Inference*, John Wiley, New York.

Meelen, A.J., J. van Soest en J.M.G. Vermeulen, *Aanvulling op Elementaire Statistiek*, DUM, Delft.

Mood, A.M., F.A. Graybill and D.C. Boes, *Introduction to the Theory of Statistics*, McGraw-Hill, New York.

Roes, P.B.M. en H.J.L. van Oorschot, *Kansrekening en Statistiek*, DUM, Delft.

Soest, ir. J. van, *Elementaire Statistiek*, DUM, Delft.

Stam, dr. A.J., *Inleiding tot de Waarschijnlijkheidsrekening*, Technische Uitgeverij H. Stam, Haarlem.

# 1

## Kansrekening

### 1.1. Uitkomstenruimte en gebeurtenissen

In den beginne was er het experiment. Het resultaat dat optreedt na uitvoering van een experiment, de uitkomst, kan afhankelijk zijn van het toeval. Als dat het geval is schieten de gebruikelijke deterministische methoden tekort en moet men de kansrekening gebruiken om numerieke uitspraken te formuleren over de gevolgen van de uitvoering van het experiment.

#### *Definities*

- Een *experiment* is een handeling met één of meer mogelijke resultaten (uitkomsten).
- De *uitkomstenruimte*,  $\Omega$ , is de verzameling van alle mogelijke uitkomsten van het experiment.  
Een element van  $\Omega$  duiden we aan met  $\omega$ .
- Een *gebeurtenis*  $A$  is een deelverzameling van de uitkomstenruimte  $\Omega$ . De gebeurtenis  $A$  treedt op als het experiment eindigt in een van de uitkomsten die tot  $A$  behoren.
- Het *complement*  $\bar{A}$  van de gebeurtenis  $A$  is de gebeurtenis dat  $A$  niet optreedt.
- Een *elementaire gebeurtenis*  $\{\omega\}$  is een deelverzameling van  $\Omega$  die slechts één element bevat.
- De *zekere gebeurtenis* is die deelverzameling van  $\Omega$  die alle elementen van  $\Omega$  bevat, m.a.w.  $\Omega$  zelf.
- Een *onmogelijke gebeurtenis*,  $\Phi$ , is een deelverzameling van  $\Omega$  die geen enkel element bevat, m.a.w. een lege verzameling.
- Twee gebeurtenissen  $A$  en  $B$  heten *disjunct* of elkaar uitsluitend als  $A \cap B = \Phi$ .

Het aantal elementen van  $\Omega$  kan eindig zijn of oneindig. In het laatste geval kan het aantal elementen aftelbaar zijn of meer dan aftelbaar. Is het aantal elementen eindig of aftelbaar dan noemt men  $\Omega$  discreet.

#### **Voorbeelden**

- a. Gooi een dobbelsteen éénmaal. Mogelijke uitkomsten: 1, 2, 3, 4, 5, 6;

$\Omega = \{1,2,3,4,5,6\}$ . De gebeurtenis A: ‘de worp is even’ bevat de uitkomsten 2, 4 en 6;  $A = \{2,4,6\}$ . De gebeurtenis B: ‘de worp is oneven’ bevat de uitkomsten 1, 3 en 5;  $B = \{1,3,5\}$ .

$$A \cap B = \Phi; A \cup B = \Omega; \bar{A} = B.$$

- b. Gooi een dobbelsteen net zo lang totdat de eerste 6 valt. De uitkomst van het experiment is het aantal malen dat de dobbelsteen moet worden gegooid.

$$\Omega = \{1,2,3,\dots\}.$$

De uitkomstenruimte heeft oneindig veel elementen maar is aftelbaar.

- c. Schiet op een schietschijf met een straal van 25 cm. De uitkomst van het experiment is de afstand in centimeters tussen het middelpunt van de schijf en het punt waar het schot de schijf heeft getroffen. Het is mogelijk dat de schijf niet wordt geraakt; ongeacht hoever het schot ernaast gaat, noemen we zo’n uitkomst ‘mis’. De elementen van  $\Omega$  hoeven dus niet gelijksoortig te zijn.

$$\Omega = \{\text{mis}\} \cup \{x \mid 0 \leq x \leq 25\}.$$

De uitkomstenruimte heeft meer dan aftelbaar veel elementen.

Zij A de gebeurtenis dat de schijf op minder dan 10 cm van het middelpunt wordt geraakt:

$$A = \{x \mid 0 \leq x < 10\}.$$

Zij B de gebeurtenis dat de schijf *niet* op minder dan 10 cm van het middelpunt wordt geraakt, m.a.w.  $B = \bar{A}$ :

$$B = \{\text{mis}\} \cup \{x \mid 10 \leq x \leq 25\}.$$

- d. Gooi een rode en een groene dobbelsteen. Mogelijke uitkomsten: (1,1),(1,2), ..., (1,6),(2,1),..., (6,6) waarbij steeds het eerste getal het aantal ogen van de rode dobbelsteen aangeeft en het tweede getal dat van de groene.

$$\Omega = \{(i,j) \mid i = 1, 2, \dots, 6; j = 1, 2, \dots, 6\}.$$

De uitkomstenruimte bevat 36 elementen. Elk element bestaat uit een tweetal (i,j).

Zij A de gebeurtenis: ‘het totale aantal ogen bij de worp is niet groter dan 4’:

$$A = \{(1,1), (1,2), (1,3), (2,1), (2,2), (3,1)\}.$$

- e. Laat de waterhoogte in cm ten opzichte van N.A.P., op een bepaald moment, op tien plaatsen langs de kust opmeten en registreer de resultaten:

$$\Omega = \{(x_1, x_2, \dots, x_{10}) \mid a_i \leq x_i; i = 1, 2, \dots, 10\}.$$

Er is geen bovengrens aangegeven voor de  $x_i$  omdat er geen zekerheid is dat de waterhoogte een vooraf door ons gestelde grens niet zal overschrijden. De ondergrens kan echter gelijkgesteld worden aan het niveau van de zeebodem ter plaatse.

## 1.2. Axioma's van de kansrekening

Om de kansrekening goed te kunnen bedrijven hebben we een kansruimte nodig. Een kansruimte bestaat uit drie ingrediënten:

- een uitkomstenruimte  $\Omega$ ;
- een  $\sigma$ -algebra  $\mathbf{A}$  van verzamelingen in  $\Omega$ ;
- een kansmaat  $P$  die aan elk element  $A$  van  $\mathbf{A}$  een getal,  $P(A)$ , toekent.\*

De uitkomstenruimte kwamen we al tegen in de vorige paragraaf. Het zou mooi zijn als we verder konden werken met de klasse van *alle* deelverzamelingen van  $\Omega$ . Dat kan echter alleen als  $\Omega$  discreet is. Als het aantal elementen van  $\Omega$  meer dan aftelbaar is kunnen er problemen ontstaan zodat we ons enigszins moeten beperken, en wel tot een  $\sigma$ -algebra.

### Definitie

Een  $\sigma$ -algebra  $\mathbf{A}$  is een klasse van deelverzamelingen van  $\Omega$  waarvoor het volgende geldt:

- $\Omega \in \mathbf{A}$ ;
- Als  $A \in \mathbf{A}$  dan ook  $\bar{A} \in \mathbf{A}$ ;
- Als  $A_i \in \mathbf{A}$  voor  $i = 1, 2, 3, \dots$  dan ook  $\bigcup_{i=1}^{\infty} A_i \in \mathbf{A}$ .

Het is gemakkelijk aan te tonen dat ook  $\Phi$  en  $\bigcap_{i=1}^{\infty} A_i$  tot  $\mathbf{A}$  behoren. Kortom  $\mathbf{A}$  is gesloten onder aftelbare verzamelingstheoretische operaties. Deelverzamelingen van  $\Omega$  die niet tot  $\mathbf{A}$  behoren zijn geen gebeurtenissen.

De kansmaat  $P$  is een functie van  $\mathbf{A}$  naar  $[0,1]$  met de volgende eigenschappen:

- $P(\Omega) = 1$ ;
- $P(A) \geq 0$  voor alle  $A \in \mathbf{A}$ ;
- $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$  voor iedere rij disjuncte gebeurtenissen  $A_1, A_2, A_3, \dots$  (d.w.z.  $A_i \cap A_j = \Phi$  als  $i \neq j$ ).

Bij elkaar krijgt men zo de kansruimte  $(\Omega, \mathbf{A}, P)$ , ook een kans-triple genaamd. Er is een zeer grote vrijheid bij het bepalen van de functiewaarden van  $P$ . Het is echter

---

\* Men komt ook andere schrijfwijzen tegen zoals  $P\{A\}$ ,  $\Pr(A)$ ,  $\Pr\{A\}$  of  $P[A]$ .

gebruikelijk, en ook verstandig, om zodanige waarden te kiezen dat aansluiting bij de praktijk wordt verkregen.

De drie eisen waaraan de kansmaat  $P$  moet voldoen worden de *axioma's* van de kansrekening genoemd. Uitgaande van deze axioma's kan een heel bouwwerk van stellingen worden opgetrokken.

**Stelling 1.1.**  $P(\Phi) = 0$ .

Bewijs:  $A, \Phi, \Phi, \Phi, \dots$  is een rij disjuncte gebeurtenissen, dus kan axioma 3 worden toegepast.

$$P(A \cup \Phi \cup \Phi \cup \dots) = P(A) + P(\Phi) + P(\Phi) + \dots \geq P(A) + P(\Phi)$$

dus

$$P(A) \geq P(A) + P(\Phi) \quad \text{ofwel } P(\Phi) \leq 0.$$

Uit axioma 2 volgt echter dat  $P(\Phi) \geq 0$  zodat  $P(\Phi) = 0$ . □

**N.B.:** Het omgekeerde geldt niet: als  $P(A) = 0$  volgt daar niet uit dat  $A = \Phi$ .

**Stelling 1.2.**  $P(A \cup B) = P(A) + P(B)$  als  $A \cap B = \Phi$ .

Bewijs: Als  $A \cap B = \Phi$  dan is  $A, B, \Phi, \Phi, \Phi, \dots$  een rij disjuncte gebeurtenissen zodat axioma 3 weer kan worden gebruikt.

$$P(A \cup B \cup \Phi \cup \Phi \cup \Phi \cup \dots) = P(A) + P(B) + P(\Phi) + P(\Phi) + P(\Phi) + \dots$$

$$P(A \cup B) = P(A) + P(B) + 0 + 0 + 0 + \dots$$

zodat  $P(A \cup B) = P(A) + P(B)$ . □

Deze stelling laat zich eenvoudig uitbreiden tot de volgende:

**Stelling 1.3.** Als  $A_1, A_2, \dots, A_n$  disjuncte gebeurtenissen zijn dan geldt

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i). \quad \square$$

**Stelling 1.4.**  $P(\bar{A}) = 1 - P(A)$ .

Bewijs:  $A \cap \bar{A} = \Phi$  dus  $P(A \cup \bar{A}) = P(A) + P(\bar{A})$ ;  
 $A \cup \bar{A} = \Omega$  dus  $1 = P(A) + P(\bar{A})$ , waaruit het gestelde volgt. □

**Stelling 1.5.**  $0 \leq P(A) \leq 1$ .

Bewijs: Volgt direct uit axioma 2 en de voorgaande stelling. □

**Stelling 1.6.**  $P(A \cap B) = P(B) - P(\bar{A} \cap B)$ .

Bewijs:  $B = B \cap \Omega = B \cap (A \cup \bar{A}) = (B \cap A) \cup (B \cap \bar{A})$ .

$B \cap A$  en  $B \cap \bar{A}$  zijn disjunct, dus

$$P(B) = P\{(B \cap A) \cup (B \cap \bar{A})\} = P(B \cap A) + P(B \cap \bar{A}) \quad \text{ofwel}$$

$$P(A \cap B) = P(B) - P(\bar{A} \cap B). \quad \square$$

**Stelling 1.7.** Als  $A \subset B$  dan geldt  $P(A) \leq P(B)$ .

Bewijs: Uit  $A \subset B$  volgt  $A \cap B = A$ . Dus geldt

$$P(B) = P(B \cap A) + P(B \cap \bar{A}) = P(A) + P(B \cap \bar{A}).$$

Volgens axioma 2 is  $P(B \cap \bar{A}) \geq 0$  zodat

$$P(B) \geq P(A). \quad \square$$

**Stelling 1.8.**  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ .

Bewijs:  $A \cup B = (A \cap \bar{B}) \cup (A \cap B) \cup (\bar{A} \cap B)$ .

Het rechterlid is een vereniging van drie disjuncte gebeurtenissen. Dus

$$\begin{aligned} P(A \cup B) &= P(A \cap \bar{B}) + P(A \cap B) + P(\bar{A} \cap B) = \\ &= P(A) - P(A \cap B) + P(A \cap B) + P(B) - P(A \cap B) = \\ &= P(A) + P(B) - P(A \cap B). \quad \square \end{aligned}$$

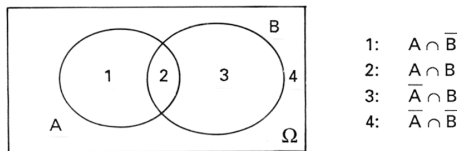
Deze stelling, de *algemene somregel*, kan worden uitgebreid tot elk eindig aantal gebeurtenissen. De algemene vorm, bekend als de *regel van inclusie/exclusie*, is

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n) &= \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_i \cap A_j) + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) - \dots \\ &\quad - (-1)^n P(A_1 \cap A_2 \cap \dots \cap A_n). \end{aligned}$$

In het geval dat  $n = 3$  wordt het:

$$\begin{aligned} P(A \cup B \cup C) &= \\ &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C). \end{aligned}$$

Vaak is een Venn-diagram een goed hulpmiddel bij het bewijzen van stellingen. In een Venn-diagram duidt men grafisch de verschillende gebeurtenissen aan. Ter illustratie:



Figuur 1.1.

Voortaan zal, indien dat geen verwarring wekt, het doorsnijdingsteken ‘ $\cap$ ’ worden weggelaten en  $AB$  de betekenis  $A \cap B$  hebben. Dit is vergelijkbaar met het weggelaten van het maalteken in sommige gevallen:  $ab = a \cdot b = a \times b$ . Het verenigingsteken ‘ $\cup$ ’ zal echter (evenals het plusteken) steeds worden geplaatst.

### 1.3. Kansruimten

Om de kansruimte bruikbaar te maken moeten de functiewaarden van  $P$  worden vastgesteld. Iedere keuze die niet strijdig is met de axioma's is toelaatbaar. We willen echter resultaten krijgen die in praktische situaties betekenis hebben. Bij een experiment met een eindig aantal mogelijke uitkomsten kan dat worden bereikt door frequentie-quotiënten van de gebeurtenissen te bepalen. Men herhaalt het experiment  $n$  maal en telt het aantal malen  $n(A)$  dat daarbij de gebeurtenis  $A$  optreedt. Het frequentie-quotiënt  $f_n(A)$  is dan gelijk aan

$$f_n(A) = \frac{n(A)}{n}.$$

Naarmate  $n$  groter wordt zal  $f_n(A)$  minder fluctueren en in de buurt van een getal  $c$  blijven. Door nu  $P(A) = c$  te stellen verkrijgt men een bruikbare waarde voor de kans op de gebeurtenis  $A$ . Deze methode kan echter niet altijd worden toegepast.

Vaak is het mogelijk om bij een experiment een eindige uitkomstenruimte te bepalen, zodanig dat daarmee een symmetrische kansruimte kan worden geconstrueerd. In een symmetrische kansruimte heeft elk van de  $n$  elementen van  $\Omega$  een even grote kans  $1/n$ . De kans op een gebeurtenis  $A$  is dan gelijk aan het aantal uitkomsten van het experiment die tot de gebeurtenis  $A$  leiden, vermenigvuldigd met  $1/n$ .

Iets anders geformuleerd verkrijgt men zo de klassieke kansdefinitie van Laplace:

Als bij een experiment alle mogelijke uitkomsten gelijkwaardig zijn dan geldt voor de kans dat de gebeurtenis  $A$  optreedt:

$$P(A) = \frac{\text{aantal uitkomsten dat gunstig is voor } A}{\text{totaal aantal mogelijke uitkomsten}}.$$

#### Voorbeelden

- Gooi een zuivere dobbelsteen:  $\Omega = \{1,2,3,4,5,6\}$ . Omdat de dobbelsteen zuiver is zijn de uitkomsten gelijkwaardig en is  $P(1) = P(2) = \dots = P(6) = \frac{1}{6}$ .
- Gooi twee zuivere dobbelstenen. Als we alleen geïnteresseerd zijn in de som van de ogen aantallen is het mogelijk om te werken met  $\Omega' = \{2,3,\dots,12\}$ . Hier zijn echter de elf uitkomsten niet gelijkwaardig. Een symmetrische kansruimte is mogelijk indien we onderscheid maken tussen de twee dobbelstenen zodat elke uitkomst door een getallenpaar wordt aangegeven. De uitkomsten (1,3) en (3,1) zijn verschillend omdat de drie in het eerste geval bij de tweede steen valt en in het laatste geval bij de eerste steen. Zo ontstaat een uitkomstenruimte met 36 elementen:

$$\Omega = \{(1,1),(1,2), \dots, (1,6),(2,1), \dots, (6,6)\},$$

waarin elk element een kans  $1/36$  heeft.



Iets meer moeite kost de constructie van een kansruimte als  $\Omega$  oneindig veel elementen bevat, zoals uit de volgende voorbeelden blijkt.

- c. Gooi een zuivere dobbelsteen net zo lang totdat de eerste 4 gevallen is. De uitkomst van het experiment is het aantal benodigde worpen, dus

$$\Omega = \{1, 2, 3, \dots\}.$$

In voorbeeld b had net zo goed met een dobbelsteen twee keer achter elkaar gegooid kunnen worden. Dan blijkt bij zes van de 36 uitkomsten de eerste worp al een 4 op te leveren. Er zijn 5 uitkomsten waarbij pas bij de tweede worp een 4 verschijnt.

Een symmetrische kansruimte voor het drie maal werpen met een zuivere dobbelsteen heeft een uitkomstenruimte met 216 elementen van de vorm  $(i, j, k)$ . Daarbij zijn er 36 met  $i = 4$ ; 30 met  $i \neq 4$  en  $j = 4$ ; 25 met  $i \neq 4, j \neq 4$  en  $k = 4$ . Dus

$$p_1 = P\{\text{de eerste 4 valt bij worp 1}\} = \frac{36}{216};$$

$$p_2 = P\{\text{de eerste 4 valt bij worp 2}\} = \frac{30}{216};$$

$$p_3 = P\{\text{de eerste 4 valt bij worp 3}\} = \frac{25}{216}.$$

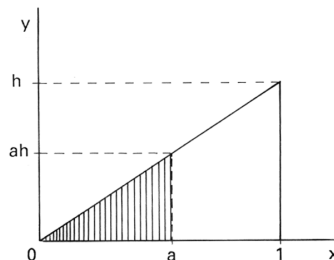
Het valt op dat  $p_2 = \frac{5}{6}p_1$  en  $p_3 = (\frac{5}{6})^2 \cdot p_1$ . Het vermoeden is dus dat bij het gooien net zolang totdat er een 4 valt zal gelden:

$P\{\text{de eerste 4 valt bij worp } n+1\} = \frac{5}{6}P\{\text{de eerste 4 valt bij worp } n\}, n=1, 2, 3, \dots$   
waaruit volgt:

$$P\{\text{de eerste 4 valt bij worp } n\} = \frac{5^{n-1}}{6^n}, n=1, 2, 3, \dots$$

Later zal blijken dat dit vermoeden gegrond is.

Merk op dat  $p_n = \{1 - (p_1 + p_2 + \dots + p_{n-1})\}/6$ .



Figuur 1.2.

- d. Prik blindelings in een driehoekig stuk karton als aangegeven in figuur 1.2. De uitkomst van het experiment is de  $x$ -coördinaat van het geprikte punt.

$$\Omega = \{x | 0 \leq x \leq 1\}.$$

‘Blindelings’ geeft aan dat elk punt in de driehoek een even grote kans heeft om gekozen te worden.

Uit het blindelings prikken volgt dat oppervlakten van gelijke grootte gelijke kans hebben om het geprikte punt te bevatten. Daaruit volgt weer dat de kans om in een deelgebied  $M$  van de driehoek te prikken gelijk is aan

$$\frac{\text{oppervlakte van gebied } M}{\text{oppervlakte van de hele driehoek}}.$$

Zij  $A$  de gebeurtenis dat  $0 \leq x \leq a$  dan is dus

$$P(A) = \frac{1/2 \cdot a \cdot ah}{1/2 \cdot 1 \cdot h} = a^2.$$

Zij  $B$  de gebeurtenis dat  $x = a$  dan volgt

$$P(B) = \frac{0}{1/2 \cdot 1 \cdot h} = 0$$

omdat het oppervlak dat gunstig is voor de gebeurtenis  $B$  (een lijn) de grootte nul heeft. Het is nu dus niet zinvol om, zoals bij een discrete uitkomstenruimte, voor elk element de kans aan te geven.

In dit geval hadden we ook een  $\sigma$ -algebra van toegelaten gebeurtenissen moeten kiezen. De  $\sigma$ -algebra van alle intervallen is ruim genoeg voor praktisch gebruik en beperkt genoeg om ons te behoeden voor problemen.

## 1.4. Combinatoriek

Symmetrische kansruimten worden al gauw erg groot. Het tellen van het aantal gunstige (en mogelijke) uitkomsten moet dan systematisch gebeuren. De mogelijkheid daartoe biedt de combinatoriek. Alle resultaten van de combinatoriek volgen uit de volgende twee grondregels.

### Productregel

Als handeling  $A$  op  $n_1$  manieren kan worden uitgevoerd en handeling  $B$  op  $n_2$  manieren, dan zijn er  $n_1 \cdot n_2$  manieren om ‘handeling  $A$  en handeling  $B$ ’ uit te voeren, mits de volgorde van  $A$  en  $B$  vastligt of er niet toe doet.

### Somregel

Als handeling  $A$  op  $n_1$  manieren kan worden uitgevoerd en handeling  $B$  op  $n_2$  manieren, dan zijn er  $n_1 + n_2$  manieren om ‘handeling  $A$  of handeling  $B$ ’ uit te voeren, mits het niet mogelijk is met één handeling zowel  $A$  als  $B$  uit te voeren.

Reeds bekende en zeer bruikbare resultaten zijn:

- Het aantal permutaties van  $n$  voorwerpen is  $n!$
- Het aantal permutaties van  $r$  voorwerpen uit  $n$  voorwerpen is  $\frac{n!}{(n-r)!}$  voor  $n \geq r$ .
- Het aantal combinaties van  $r$  voorwerpen uit  $n$  voorwerpen is  $\binom{n}{r}$  voor  $n \geq r$ .

Het aantal combinaties van  $r$  uit  $n$  is ook het aantal manieren waarop  $n$  voorwerpen over twee personen, A en B, verdeeld kunnen worden, zó dat A  $r$  voorwerpen krijgt en B  $n - r$  voorwerpen, waarbij A en B niet op de volgorde letten.

Dit kan worden uitgebreid tot het verdelen van  $n$  voorwerpen over  $k$  personen waarbij persoon  $i$   $r_i$  voorwerpen ontvangt,  $i = 1, 2, \dots, k$ , en  $\sum_{i=1}^k r_i = n$  en niet op de volgorde wordt gelet. Kies eerst  $r_1$  voorwerpen uit voor persoon 1; dit kan op  $\binom{n}{r_1}$  manieren. Kies uit de  $n - r_1$  resterende de  $r_2$  voorwerpen voor persoon 2; dat kan op  $\binom{n-r_1}{r_2}$  manieren, enz. Het totaal aantal manieren wordt dan

$$\binom{n}{r_1} \cdot \binom{n-r_1}{r_2} \dots \binom{n-r_1-r_2-\dots-r_{k-1}}{r_k}$$

hetgeen na uitschrijven van de binominaalcoëfficiënten

$$\frac{n!}{r_1! r_2! \dots r_k!}$$

oplevert.

Men krijgt hetzelfde resultaat voor het aantal permutaties van  $n$  voorwerpen als er daarbij  $r_i$  van soort  $i$  zijn, voor  $i = 1, 2, \dots, k$ , terwijl  $\sum_{i=1}^k r_i = n$ . Hierbij worden de voorwerpen van één soort niet van elkaar onderscheiden.

In combinatorische problemen komt men vaak de term ‘trekken met teruglegging’ of ‘trekken zonder teruglegging’ tegen. Bij trekken met teruglegging wordt elk getrokken voorwerp weer toegevoegd aan de verzameling waar men uit trekt, voordat een volgende trekking wordt gedaan. Het herhaald gooien met een dobbelsteen kan men beschouwen als trekken met teruglegging uit de verzameling  $\{1, 2, 3, 4, 5, 6\}$ .

Bij trekken zonder teruglegging blijven de getrokken voorwerpen terzijde liggen. De verzameling waaruit men trekt wordt dus steeds kleiner. Als er wordt getrokken uit een vaas met  $n$  ballen dan kan men bij trekken zonder teruglegging hoogstens  $n$  maal een bal trekken. Bij trekken met teruglegging is er niet zo’n grens.

De toevoeging van de term ‘aselect’, ‘blindelings’ of ‘willekeurig’ duidt aan dat alle vergelijkbare uitkomsten even waarschijnlijk zijn. Dat betekent dus dat men een symmetrische kansruimte kan gebruiken.

## 1.5. Enige voorbeelden

- a. In een doos zitten 7 bouten en 3 moeren. We pakken blindelings twee keer iets uit de doos zonder de getrokken voorwerpen terug te leggen. Hoe groot is de kans op de gebeurtenis A dat we een bout en een moer trekken.

*Volgorde vast:*

Om een symmetrische kansruimte te krijgen nummeren we de voorwerpen  $1, 2, \dots, 10$  (1-7 de bouten, 8-10 de moeren). Als uitkomstenruimte kiezen we de paren  $(i, j)$  met  $i \neq j$  en  $i = 1, 2, \dots, 10; j = 1, 2, \dots, 10$ , waarbij  $i(j)$  het resultaat van de eerste (tweede) keer pakken weergeeft. Er zijn  $\frac{10!}{(10-2)!} = 90$  permutaties van 2 uit 10 mogelijk, dus  $\Omega$  heeft 90 elementen, elk met kans  $1/90$ . Hoeveel van deze uitkomsten zijn er gunstig voor A? Dat zijn al de uitkomsten waarbij  $i \leq 7$  en  $j \geq 8$  is of waarbij  $i \geq 8$  en  $j \leq 7$  is. Volgens de productregel, die mag worden toegepast omdat de volgorde vastligt zijn er  $7 \cdot 3$  manieren om een  $i \leq 7$  en een  $j \geq 8$  te kiezen terwijl er  $3 \cdot 7$  manieren zijn om een  $i \geq 8$  en een  $j \leq 7$  te kiezen. De somregel mag worden toegepast zodat er  $21 + 21 = 42$  gunstige mogelijkheden zijn en dus volgt:

$$P(A) = \frac{42}{90} = \frac{7}{15}.$$

*Volgorde doet er niet toe:*

In dit geval was het echter helemaal niet nodig om de volgorde waarin de voorwerpen gepakt worden in de beschouwingen te betrekken. Alleen wat uiteindelijk naast de doos ligt is van belang, niet of een gekozen voorwerp als eerste of als tweede uit de doos kwam. Zo redenerend komt men tot een uitkomstenruimte met  $\binom{10}{2} = 45$  elementen  $(i, j); i = 1, 2, \dots, 9; j = i+1, i+2, \dots, 10$ .

Het aantal gunstige uitkomsten volgt weer m.b.v. de productregel, die nu mag worden toegepast omdat de volgorde er niet toe doet, en is gelijk aan  $7 \cdot 3 = 21$ . Ook op deze wijze verkrijgt men

$$P(A) = \frac{21}{45} = \frac{7}{15}.$$

- b. Kies een willekeurig getal van drie cijfers. Hoe groot is de kans op de gebeurtenis B dat minstens een van de drie cijfers een 6 is?

$\Omega = \{100, 101, \dots, 999\}$  en bevat 900 elementen. De kansruimte is symmetrisch. Het bepalen van het aantal gunstige uitkomsten kan door inspectie van de uitkomsten gebeuren. Handiger is het echter om  $P(B)$  te bepalen via  $P(\bar{B})$ , en dan gebruik te maken van  $P(B) = 1 - P(\bar{B})$ .  $\bar{B}$  is de gebeurtenis dat het gekozen getal geen enkele 6 bevat. Hoeveel van zulke getallen zijn er in  $\Omega$ ?

Voor het eerste cijfer zijn er 8 mogelijkheden (0 en 6 vallen uit). Voor het tweede zowel als voor het derde cijfer zijn er 9 mogelijkheden (alleen de 6 mag

niet). Met behulp van de productregel krijgen we dus  $8 \cdot 9 \cdot 9$  uitkomsten waar geen 6 in voorkomt.

$$\text{Dus } P(\bar{B}) = \frac{8 \cdot 9 \cdot 9}{900} = \frac{18}{25}$$

$$\text{en } P(B) = 1 - \frac{18}{25} = \frac{7}{25}.$$

- c. De zes deelnemers aan een ronde-tafel conferentie kunnen op  $5!$  verschillende manieren worden geplaatst. Omdat de tafel rond is maakt het niet uit op welke stoel de eerste deelnemer gaat zitten. De vijf anderen kunnen dan op  $5!$  manieren over de resterende vijf stoelen worden verdeeld.
- d. Iemand heeft een binair getal van vier cijfers opgeschreven, d.w.z. een rij ter lengte vier met op elke plaats een 0 of een 1. U mag raden welk getal het is. Er zijn  $2 \times 2 \times 2 \times 2 = 16$  mogelijkheden en dus is de kans dat u het getal goed raadt  $1/16$ . De kans dat u minstens 3 van de 4 symbolen goed raadt is  $5/16$  (ga dit na). Als u nu als extra informatie krijgt dat het getal uit 2 nullen en 2 enen bestaat zijn er maar 6 getallen mogelijk. De kans dat u de juiste volgorde van 2 nullen en twee enen raadt is  $1/6$ . De kans dat u, zodoende, precies 3 van de 4 symbolen goed raadt is 0 (waarom?). U kunt echter toch de kans dat u minstens 3 van de vier symbolen goed raadt gelijk aan  $\frac{1}{2}$  maken en wel door in een willekeurige volgorde 1 nul en 3 enen (of 3 nullen en 1 een) op te schrijven. U hebt dan met kans  $\frac{1}{2}$  één symbool goed en eveneens met kans  $\frac{1}{2}$  drie symbolen goed.

## 1.6. Conditionele kans

Laat een experiment bestaan uit het tweemaal gooien met een zuivere dobbelsteen en de belangstelling uitgaan naar het totaal aantal ogen dat daarbij wordt geworpen.

$$\Omega = \{(i,j) \mid i = 1,2,\dots,6; j = 1,2,\dots,6\}.$$

Zij A de gebeurtenis dat  $i + j = 10$ , dan is  $P(A) = 3/36$ .

Zij B de gebeurtenis dat de eerste worp een drie is.

Met een drie bij de eerste worp is het onmogelijk om nog een totaal van tien te halen. Met andere woorden de kans op de gebeurtenis A, als gegeven is dat de gebeurtenis B is opgetreden is gelijk aan nul. Men spreekt in zo'n geval van de *conditionele* (of *voorwaardelijke*) kans op A gegeven B.

In het algemeen kan men, zodra bekend is dat B is opgetreden, een aangepaste kansmaat  $P_B$  bij het experiment construeren waarbij alle elementen van  $\bar{B}$  de kans nul krijgen omdat zij niet meer als uitkomst van het experiment kunnen optreden. Het experiment kan alleen nog maar eindigen in één van de uitkomsten die bevat zijn in B. Als de gebeurtenis A optreedt eindigt het experiment in een van de uit-

komsten die bevat zijn in  $A$ . Dus  $A$  treedt op bij gegeven  $B$  als het experiment eindigt in een van de uitkomsten die behoren tot  $AB$ . Evenzo volgt dat  $\bar{A}$  optreedt bij gegeven  $B$  als het experiment eindigt in een van de uitkomsten behorend tot  $\bar{A}B$ .

Het is duidelijk dat  $P_B(A) + P_B(\bar{A}) = 1$  moet zijn en dat

$$P_B(A) : P_B(\bar{A}) = P_B(AB) : P_B(\bar{A}B).$$

Hieruit volgt

$$P_B(A) : 1 = P(AB) : P(B), \text{ zodat } P_B(A) = \frac{P(AB)}{P(B)}.$$

Natuurlijk moet  $P(B) \neq 0$  zijn. De gebruikelijke notatie voor de conditionele kans  $P_B(A)$  is  $P(A|B)$ . Het is gemakkelijk aan te tonen dat de conditionele kansmaat aan de drie axioma's voldoet, m.a.w. dat voor  $P(B) \neq 0$  geldt:

1.  $P(\Omega|B) = 1$ ;
2.  $P(A|B) \geq 0$  voor alle  $A \in \mathcal{A}$ ;
3.  $P(\bigcup_i A_i|B) = \sum_i P(A_i|B)$  mits  $A_i A_j = \Phi$  voor  $i \neq j$ .

Dat betekent dat elke stelling die geldt voor gewone (of absolute) kansen ook geldt voor conditionele kansen.

### Voorbeeld

Men trekt aselect en zonder teruglegging tweemaal een bal uit een vaas die 4 witte en 6 rode ballen bevat. Zij  $W_i$  de gebeurtenis dat de  $i^e$  getrokken bal wit is,  $i = 1$  of  $2$ .

$$P(W_1) = \frac{4}{10} = \frac{2}{5}; \quad P(W_1 W_2) = \frac{4 \cdot 3}{10 \cdot 9} = \frac{2}{15}; \quad P(\bar{W}_1 W_2) = \frac{6 \cdot 4}{10 \cdot 9} = \frac{4}{15};$$

$$P(W_2|W_1) = \frac{P(W_1 W_2)}{P(W_1)} = \frac{2/15}{2/5} = \frac{1}{3};$$

$$P(W_2) = P(W_1 W_2) + P(\bar{W}_1 W_2) = \frac{2}{15} + \frac{4}{15} = \frac{2}{5};$$

$$P(W_1|W_2) = \frac{P(W_1 W_2)}{P(W_2)} = \frac{2/15}{2/5} = \frac{1}{3}.$$

### Voorbeeld

Men gooit tweemaal met een zuivere dobbelsteen. Hoe groot is de conditionele kans op 'tweemaal even' gegeven dat minstens één van de twee worpen een even resultaat geeft.

Zij  $A$  de gebeurtenis 'tweemaal even';  $P(A) = \frac{3 \cdot 3}{6 \cdot 6} = 1/4$ .

Zij  $B$  de gebeurtenis 'minstens éénmaal even' dan is  $\bar{B}$  'tweemaal oneven';

$$P(\bar{B}) = \frac{3 \cdot 3}{6 \cdot 6} = \frac{1}{4} \text{ zodat } P(B) = \frac{3}{4}. \text{ Gevraagd wordt } P(A|B) = \frac{P(AB)}{P(B)}.$$

$$\text{Omdat } A \subset B \text{ volgt } AB = A, \text{ dus } P(A|B) = \frac{P(A)}{P(B)} = \frac{1/4}{3/4} = \frac{1}{3}. \quad \square$$

**Stelling 1.9.**  $P(AB) = P(A|B) \cdot P(B)$ , als  $P(B) \neq 0$ ;  
 $= 0$ , als  $P(B) = 0$ .

Bewijs: De eerste regel volgt onmiddellijk uit  $P(A|B) = \frac{P(AB)}{P(B)}$ ; de tweede uit  $0 \leq P(AB) \leq P(B)$ . □

**Stelling 1.10.**  $P(A_1 A_2 \dots A_n) =$   
 $= P(A_n | A_1 A_2 \dots A_{n-1}) \cdot P(A_{n-1} | A_1 A_2 \dots A_{n-2}) \dots P(A_2 | A_1) \cdot P(A_1)$ ,  
mits  $P(A_1 A_2 \dots A_{n-1}) \neq 0$ .

Dit heet de *algemene productregel*.

Bewijs:  $P(A_2 | A_1) \cdot P(A_1) = P(A_1 A_2)$  want uit  $P(A_1 A_2 \dots A_{n-1}) \neq 0$  volgt  $P(A_1) \neq 0$  omdat  $A_1 \supset A_1 A_2 \dots A_{n-1}$ .

Pas nu weer de vorige stelling toe waarbij  $A_1 A_2$  optreedt als conditie B dan volgt  $P(A_3 | A_1 A_2) \cdot P(A_1 A_2) = P(A_1 A_2 A_3)$ , enz. □

### Definitie

Een *partitie*  $\mathcal{B}$  van  $\Omega$  is een klasse van deelverzamelingen  $B_i$ ,  $i = 1, 2, \dots, n$ , van  $\Omega$  zodanig dat:

$$\bigcup_i B_i = \Omega;$$

$$B_i B_j = \Phi \text{ als } i \neq j;$$

$$P(B_i) \neq 0.$$

**Stelling 1.11.** Zij  $\mathcal{B}$  een partitie van  $\Omega$  dan geldt:

$$P(A) = \sum_{i=1}^n P(A|B_i) P(B_i).$$

Dit heet de stelling van de *totale waarschijnlijkheid*.

Bewijs:  $P(A|B_i) \cdot P(B_i) = P(AB_i)$  want  $P(B_i) \neq 0$ .

$AB_i$  en  $AB_j$  zijn disjunct want  $B_i$  en  $B_j$  zijn disjunct, dus:

$$\begin{aligned} \sum_{i=1}^n P(A|B_i) \cdot P(B_i) &= \sum_{i=1}^n P(AB_i) = P\left(\bigcup_{i=1}^n AB_i\right) = P\left(A \cap \bigcup_{i=1}^n B_i\right) = \\ &= P(A \cap \Omega) = P(A). \end{aligned} \quad \square$$

**Stelling 1.12.** Zij  $B$  een partitie van  $\Omega$  en  $P(A) \neq 0$  dan geldt:

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_{i=1}^n P(A|B_i)P(B_i)} .$$

Dit is de stelling van *Bayes*.

Bewijs: Volgens de stelling van de totale waarschijnlijkheid is

$$\sum_{i=1}^n P(A|B_i)P(B_i) = P(A).$$

Volgens de algemene productregel is  $P(A|B_k)P(B_k) = P(AB_k)$ . Dit ingevuld levert

$$P(B_k|A) = \frac{P(AB_k)}{P(A)}$$

hetgeen juist is omdat  $P(A) \neq 0$ . □

**Voorbeeld**

Een kast heeft drie laden. Lade A bevat 2 gouden munten, lade B een gouden en een zilveren terwijl lade C twee zilveren munten bevat. Men kiest aselekt een lade en vervolgens aselekt een munt uit die lade.

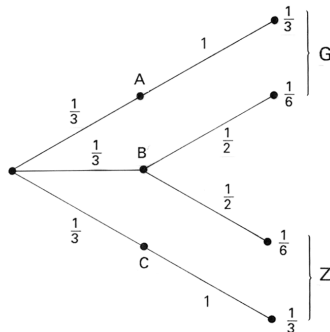
Zij A de gebeurtenis dat lade A wordt gekozen, idem voor B en C. Zij G de gebeurtenis dat een gouden munt wordt gekozen.

$$P(G) = P(G|A)P(A) + P(G|B)P(B) + P(G|C)P(C) = 1 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} = \frac{1}{2} .$$

$$P(B|G) = \frac{P(G|B)P(B)}{P(G)} = \frac{1/2 \cdot 1/3}{1/2} = \frac{1}{3} .$$

Zo volgt ook  $P(A|G) = \frac{2}{3}$  en  $P(C|G) = 0$ .

Men kan dit illustreren met een kansboom waarin alle mogelijkheden om het experiment uit te voeren grafisch worden weergegeven . Voor het bovenstaande voorbeeld krijgt men zo:



Figuur 1.3.

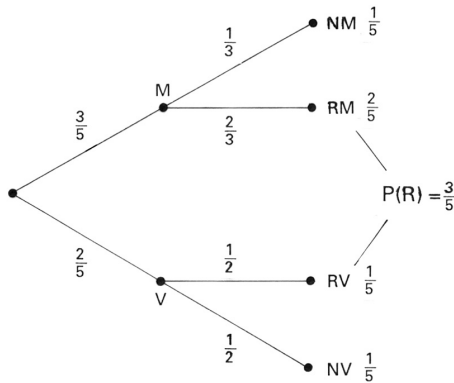


Hieruit valt af te lezen dat  $P(G) = \frac{1}{2}$  en  $P(AG) = \frac{1}{3}$ .

### Voorbeeld

In een trein zitten 60 mannen en 40 vrouwen. Van de mannen rookt  $\frac{2}{3}$ , van de vrouwen de helft. De conducteur kiest aselekt een reiziger uit waarvan hij het plaatsbewijs gaat controleren. Zij  $M(V)$  de gebeurtenis ‘het is een man(vrouw)’. Zij  $R(N)$  de gebeurtenis ‘de gekozen persoon rookt (rookt niet)’.

$$P(V|R) = \frac{P(R|V) \cdot P(V)}{P(R|V)P(V) + P(R|M)P(M)} = \frac{1/2 \cdot 40/100}{1/2 \cdot 40/100 + 2/3 \cdot 60/100} = \frac{1/5}{3/5} = \frac{1}{3}.$$



$$P(V|R) = \frac{P(VR)}{P(R)} = \frac{1/5}{3/5} = \frac{1}{3}$$

Figuur 1.4.

## 1.7. Onafhankelijkheid

### Definities

- Twee gebeurtenissen A en B zijn (stochastisch) onafhankelijk als geldt:

$$P(AB) = P(A) \cdot P(B).$$

- De gebeurtenissen  $A_1, A_2, \dots, A_n$  zijn (onderling) onafhankelijk als geldt:

$$P(A_i A_j) = P(A_i) \cdot P(A_j) \text{ voor alle } i \neq j$$

en

$$P(A_i A_j A_k) = P(A_i) \cdot P(A_j) \cdot P(A_k) \text{ voor alle } i \neq j \neq k$$

en

⋮

en

$$P\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P(A_i).$$

– De gebeurtenissen  $A_1, A_2, \dots, A_n$  zijn paarsgewijs onafhankelijk als geldt:

$$P(A_i A_j) = P(A_i) \cdot P(A_j) \text{ voor alle } i \neq j.$$

Het is nodig om er met nadruk op te wijzen dat onafhankelijkheid en disjunctie van twee gebeurtenissen geheel verschillende begrippen zijn. Zodra de uitkomstenruimte van een experiment is bepaald kan worden nagegaan of  $A$  en  $B$  al dan niet disjunct zijn. Over onafhankelijkheid van  $A$  en  $B$  kan pas iets worden gezegd als de kansmaat  $P$  gekozen is. Door de keuze van  $P$  kan men desgewenst  $A$  en  $B$  onafhankelijk maken.

**Stelling 1.13.** Als  $A$  en  $B$  disjuncte gebeurtenissen zijn dan zijn zij alleen dan onafhankelijk als minstens één van de twee de kans nul heeft.

Bewijs:  $AB = \Phi$  dus  $P(AB) = 0$  volgt uit de disjunctie. Voor onafhankelijkheid moet gelden  $P(AB) = P(A) \cdot P(B)$ , dus hier  $P(A) \cdot P(B) = 0$ .  $\square$

**Stelling 1.14.** Als  $A$  en  $B$  onafhankelijke gebeurtenissen zijn en  $0 < P(A) < 1$  dan geldt dat:

$$P(B|A) = P(B|\bar{A}) = P(B).$$

Bewijs: 
$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{P(A) \cdot P(B)}{P(A)} = P(B).$$

$$\bar{A}B \cup AB = B, \text{ dus } P(\bar{A}B) + P(AB) = P(B) \text{ ofwel}$$

$$P(\bar{A}B) = P(B) - P(A) \cdot P(B) = P(B) \cdot P(\bar{A}),$$

zodat ook  $P(B|\bar{A}) = P(B)$ .  $\square$

Dit resultaat verklaart de naam onafhankelijkheid. Het wel of niet optreden van de gebeurtenis  $A$  heeft geen invloed op de kans dat gebeurtenis  $B$  optreedt.

Bij het construeren van een kansmaat zorgt men er dus voor dat gebeurtenissen waarvan men aanneemt dat ze elkaars kans van optreden niet beïnvloeden, onafhankelijk zijn.

Laat een experiment  $E$  bestaan uit het uitvoeren van  $n$  onafhankelijke proeven  $E_i$ ,  $i = 1, 2, \dots, n$ . Bij elke proef  $E_i$  bepaalt men de uitkomstenruimte  $\Omega_i$  met de bijbehorende kansmaat  $P_i$ . Een geschikte uitkomstenruimte  $\Omega$  voor  $E$  is de productruimte  $\Omega_1 \times \Omega_2 \times \dots \times \Omega_n$ . Elk element  $\omega$  van  $\Omega$  is van de vorm  $(\omega_1, \omega_2, \dots, \omega_n)$  met  $\omega_i \in \Omega_i$ . Door nu te stellen

$$P(\{\omega\}) = P_1(\{\omega_1\}) \cdot P_2(\{\omega_2\}) \cdot \dots \cdot P_n(\{\omega_n\})$$

ontstaat een kansmaat voor  $\Omega$  waarin de gewenste onafhankelijkheid aanwezig is.

### Voorbeeld

Werp tweemaal met een zuivere munt en eenmaal met een zuivere dobbelsteen.

$$\Omega_1 = \{k, m\}, \Omega_2 = \{k, m\}, \Omega_3 = \{1, 2, 3, 4, 5, 6\};$$

$$P_1(\{k\}) = P_1(\{m\}) = 1/2; P_2(\{k\}) = P_2(\{m\}) = 1/2; P_3(\{i\}) = 1/6 \text{ voor } i = 1, 2, \dots, 6;$$

$$P(\{(k, m, 5)\}) = P_1(\{k\}) \cdot P_2(\{m\}) \cdot P_3(\{5\}) = 1/2 \cdot 1/2 \cdot 1/6 = 1/24.$$

Zij A de gebeurtenis dat  $1 \times$  kruis,  $1 \times$  munt en hoogstens 4 met de dobbelsteen wordt geworpen. Noem 'hoogstens 4'  $A_3$  dan volgt dat

$$A = \bigcup_{i=1}^4 \{(m, k, i)\} \cup \bigcup_{j=1}^4 \{(k, m, j)\}$$

$$\text{zodat } P(A) = \sum_{i=1}^4 P(\{(m, k, i)\}) + \sum_{j=1}^4 P(\{(k, m, j)\}) =$$

$$= P_1(m) \cdot P_2(k) \cdot P_3(A_3) + P_1(k) \cdot P_2(m) \cdot P_3(A_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}.$$

Dezelfde constructie wordt ook toegepast bij een aftelbaar aantal onafhankelijke deel-experimenten maar is daar formeel iets gecompliceerder.

### Definitie

Een aftelbare rij gebeurtenissen is onderling onafhankelijk als iedere eindige greep uit deze rij uit onafhankelijke gebeurtenissen bestaat.

Naar analogie van deze definitie kan men hier denken aan een experiment E dat bestaat uit de aftelbare rij onafhankelijke deel-experimenten  $E_1, E_2, \dots$ . Als er voor elke eindige greep uit de rij deel-experimenten een uitkomstenruimte en kansmaat op de bovenstaande wijze worden geconstrueerd dan gaat het wel goed.

### Voorbeeld

Gooi met een zuivere dobbelsteen tot de eerste zes valt. Vooraf kunnen we geen grens aangeven waar het aantal benodigde worpen onder zal blijven. Beschouw het eenmaal werpen van de dobbelsteen als deel-experiment  $E_i$  met uitkomstenruimte  $\Omega_i = \{s, m\}$  waarbij s staat voor succes (= 6 valt) en m voor mislukking (1, 2, 3, 4 of 5 valt). Neem als  $P_i$ :  $P_i(m) = 5/6$  en  $P_i(s) = 1/6$ .

Voor het berekenen van de kans op de gebeurtenis  $A_n$  dat bij de  $n^e$  worp ( $n = 1, 2, \dots$ ) de eerste zes valt hebben we alleen de eindige greep  $E_1, E_2, \dots, E_n$  nodig.

$$A_n = \{ \underbrace{(m, m, m, \dots, m, s)}_{(n \pm 1)x} \} \text{ zodat}$$

$$P(A_n) = P_1(m) \cdot P_2(m) \dots P_{n-1}(m) \cdot P_n(s) = \left(\frac{5}{6}\right)^{n-1} \cdot \frac{1}{6}, \quad n = 1, 2, \dots$$

Het gaat ook goed als men de kans op de gebeurtenis  $B_n$  ( $n = 1, 2, \dots$ ), dat de eerste zes valt bij of na de  $n^e$  worp, wil berekenen. Voor het optreden van  $B_n$  is nodig en voldoende dat de eerste  $n - 1$  worpen alle een mislukking opleveren. Men vindt zo  $P(B_n) = (5/6)^{n-1}$ ,  $n = 1, 2, \dots$ . De kans dat het aantal worpen tot en met de eerste zes even is kan als volgt worden bepaald.

$$\begin{aligned} P(A_2 \cup A_4 \cup \dots \cup A_{2m}) &= \sum_{i=1}^m P(A_{2i}) = \frac{1}{6} \sum_{i=1}^m \left(\frac{5}{6}\right)^{2i-1} = \\ &= \frac{1}{5} \sum_{i=1}^m \left(\frac{25}{36}\right)^i = \frac{1}{5} \frac{25}{36} \frac{1 - (25/36)^m}{1 - 25/36} = \frac{5}{11} \left\{1 - \left(\frac{25}{36}\right)^m\right\}. \end{aligned}$$

Maar voor de gebeurtenis ‘even aantal worpen nodig’ is het niet nodig  $m$  te begrenzen, zodat door het onbegrensd laten groeien van  $m$  volgt:

$$P(\text{even aantal worpen}) = \frac{5}{11}.$$