Alex van den Brandhof

Kansrekening

een introductie

Epsilon Uitgaven Utrecht

Voorwoord

In moderne wetenschappen is kansrekening een onmisbaar vak. Het verdient daarom een vaste plaats in het wiskundeonderwijs. Kansrekening kent uitdagende problemen uit het leven van alledag en speelt een grote rol in de maatschappij met al haar risico's en onzekerheden. Toepassingen zijn er in onder andere de natuurkunde, biologie, geneeskunde, economie, sociologie en psychologie. Maar ook in de zuivere wiskunde is kansrekening een gevestigde tak.

Kansrekening – een introductie is geschreven voor middelbare scholieren met Wiskunde B in het vakkenpakket, maar kan ook worden gebruikt in bijvoorbeeld het hoger beroepsonderwijs. Het boek gaat minder ver dan de meeste inleidingen in de kansrekening voor wiskundestudenten aan een universiteit, maar het overstijgt de kansrekening die in de wiskundemethodes voor het vwo wordt aangeboden. Raakvlakken met de analyse en de meetkunde worden duidelijk zichtbaar gemaakt. Het boek bevat een uitgebreide collectie uitdagende opgaven, met daarbij ook een aantal (soms bewerkte) vwo-eindexamensommen Wiskunde B1 uit de periode 2001-2009.

Dit boek legt zich toe op *kansrekening*. Kennis van *combinatoriek* wordt verondersteld; hoofdstuk 1 bevat combinatorische opgaven, echter zonder verklaring van de theorie. De eindtermen van het domein *Kansrekening en Statistiek* in het programma Wiskunde D op het vwo worden wat betreft de kansrekening ruimschoots gedekt. Schattings- en toetsingstheorie, onderdelen uit de statistiek, komen in dit boek niet aan bod. In de literatuurlijst worden enkele toegankelijke teksten over deze onderwerpen vermeld.

Af en toe wordt een beroep gedaan op de grafische rekenmachine. De uitleg over het gebruik van de grafische rekenmachine is beknopt gehouden en – zonder reclame voor Texas Instruments te willen maken – gebaseerd op de TI-84 Plus. Grafische rekenmachines van andere merken hebben vergelijkbare mogelijkheden. Op internet is een schat aan interessant aanvullend materiaal te vinden. Interessante links, artikelen en softwarepakketten over kansrekening zijn te vinden op `www.kansrekening.nl`. Verder is *WolframAlpha* een aanrader: `www.wolframalpha.com`. Deze 'antwoordmachine' is bedacht door Stephen Wolfram, de ontwerper van het wetenschappelijke softwareprogramma *Mathematica*. Geef je in de 'vraagbalk' bijvoorbeeld `binomial distribution n=30 p=0.2` in, dan krijg je allerhande informatie over deze kansverdeling, zoals de kansdichtheid, de verwachtingswaarde, de variantie en de standaardafwijking, een kanshistogram en een grafiek van de verdelingsfunctie.

Ik dank Jan van de Craats en Henk Tijms voor hun nuttige adviezen en voor het feit dat ik diverse opgaven uit hun publicaties heb mogen overnemen, waarbij ik in het bijzonder het boek *Understanding Probability* van Henk Tijms wil noemen.

Amsterdam, januari 2012
Alex van den Brandhof

Inhoud

1 Voorkennis

In dit hoofdstuk staan opgaven uit de *combinatoriek*. De theorie hiervan wordt bekend verondersteld; we volstaan hier met het geven van een overzicht van telmodellen.

Veel combinatorische problemen kunnen geherformuleerd worden in termen van het doen van *trekkingen uit een doos met al dan niet genummerde ballen*. Als je wilt weten hoeveel uitkomsten er zijn bij drie worpen met een dobbelsteen, kun je doen alsof je drie keer trekt uit een doos met zes genummerde ballen. Als je geïnteresseerd bent in het aantal manieren waarop twintig mensen op verschillende dagen jarig kunnen zijn, kun je doen alsof je twintig maal trekt uit een doos met 365 genummerde ballen.

Met deze voorbeelden is het duidelijk dat we nog wel moeten aangeven of een getrokken bal vóór de volgende trekking in de doos wordt teruggelegd of dat dit niet het geval is; we spreken van *trekkingen met* en *trekkingen zonder teruglegging*. Verder moeten we afspreken of we al dan niet letten op de volgorde waarin de ballen worden getrokken; we onderscheiden *trekkingen met* en *trekkingen zonder* (in achtneming van de) *volgorde*. Aldus zijn er vier soorten trekkingen; in elk van de gevallen geven we hieronder de telregels, zonder bewijs. Daarbij geven we het aantal trekkingen aan met r, en stellen we het aantal ballen in de doos gelijk aan n ($r \leq n$ bij trekking zonder teruglegging).

Trekking met teruglegging en met volgorde
Het aantal manieren om r ballen te trekken uit een doos met n ballen, waarbij na elke trekking de getrokken bal wordt teruggelegd en waarbij de volgorde van belang is, is gelijk aan

$$\text{het aantal } \textit{herhalingsvariaties} \text{ van } r \text{ uit } n = n^r$$

Voorbeeld 1 Bulle maakt een driekeuzetoets die uit tien vragen bestaat. Op hoeveel manieren kan hij het antwoordformulier invullen?

Uitwerking Voor elk van de tien vragen zijn er drie alternatieven (A, B en C). We gebruiken het model met teruglegging, want nadat vraag 1 met (bijvoorbeeld) alternatief A wordt beantwoord, kan vraag 2 opnieuw met alternatief A worden beantwoord. Verder is de volgorde van belang, want (bijvoorbeeld) de rijtjes AAAACCCCBB en BBACACACAC zijn, ondanks het even vaak voorkomen van de alternatieven A, B en C, verschillend. Het aantal manieren om het antwoordformulier in te vullen, is dus gelijk aan het aantal herhalingsvariaties van tien uit drie: $3^{10} = 59049$. □

Voorbeeld 2 Een pincode bestaat uit vier cijfers. Hoeveel verschillende pincodes zijn er?

Uitwerking Voor elk van de vier cijfers zijn er tien mogelijkheden (0 tot en met 9). Het aantal pincodes is dus gelijk aan het aantal herhalingsvariaties van vier uit tien: $10^4 = 10000$. □

2 Kansbegrip

Kansrekening vindt zijn oorsprong in de zeventiende eeuw, toen gokkers in rokerige salons de wetten van het toeval mathematisch probeerden te doorgronden. In 1654 ontstond een briefwisseling tussen de Franse wiskundige Pierre de Fermat (1601-1665) en Blaise Pascal (1623-1662), over kansproblemen die aan hen door gokkers werden voorgelegd. De correspondentie is bewaard gebleven en tegenwoordig beschikbaar (ook in Engelse vertaling) op het internet, zie de literatuurlijst achterin.

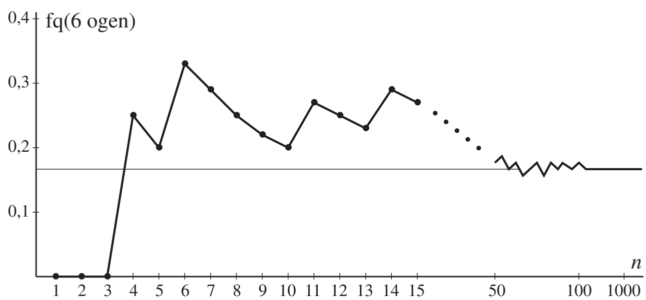
Na de briefwisseling volgde een reeks publicaties over kansrekening. In 1657 schreef de Nederlander Christiaan Huygens (1629-1695) het eerste boek over dit onderwerp, getiteld *De ratiociniis in ludo aleae* ('Over berekeningen bij het dobbelspel'). Drie jaar later verscheen van zijn hand *Van Rekeningh in Spelen van Geluck*; in 1998 verscheen hiervan een vertaling in de Epsilonreeks, zie opnieuw de literatuurlijst.

2.1 Kans als empirische relatieve frequentie

Bij een toevalsexperiment, zoals het werpen met een dobbelsteen, is de uitkomst onvoorspelbaar. Toch blijkt een dergelijk proces aan een zekere wetmatigheid te voldoen. Stel dat we duizend keer met een dobbelsteen werpen en daarbij letten op het aantal keer dat er 6 ogen wordt gegooid. Na elke worp noteren we het *frequentiequotiënt*

$$\text{fq}(6 \text{ ogen}) = \frac{n(6 \text{ ogen})}{n}.$$

Hierbij is $n(6 \text{ ogen})$ het aantal keer dat er 6 ogen wordt geworpen in de eerste n worpen. Na elke worp zetten we de waarde van het frequentiequotiënt uit in een grafiek, die er bijvoorbeeld zo uit kan zien:



De opeenvolgende frequentiequotiënten vertonen in het begin, als n nog klein is, vrij sterke schommelingen, maar naarmate n groter wordt, vertonen de frequentiequotiënten steeds minder variatie. Het lijkt alsof het frequentiequotiënt tot een *limietwaarde* nadert. Dit experimenteel gevonden verschijnsel staat bekend als de *empirische wet van grote aantallen*. (Er bestaan ook diverse *wiskundige wetten van grote aantallen*, zie hoofdstuk 8.) Intuïtief zouden we de gevonden limietwaarde willen laten corresponderen met de *kans* dat er 6 ogen wordt gegooid bij het werpen met een dobbelsteen.

OPGAVEN

1 Paul werpt vijfhonderd maal met een viervlaksdobbelsteen. De ogentallen 1, 2, 3 en 4 komen achtereenvolgens 117, 132, 145 en 106 keer voor.

a. Bereken fq(2 ogen).
b. Bereken fq(aantal ogen is oneven).

2 In het jaar 2001 werden volgens het Centraal Bureau voor de Statistiek 103.800 jongens en 98.800 meisjes in Nederland geboren. Wat is op grond van deze gegevens volgens de empirische wet van grote aantallen de kans dat bij een geboorte (in Nederland) een meisje ter wereld komt?

3 Bij een worp met een punaise kan de punaise op twee manieren terechtkomen: met de punt naar boven (⊥) of met de punt naar beneden (⋌). Ralph werpt duizend keer met een punaise; daarbij belandt de punaise 355 keer met de punt naar boven.

a. Hoe groot schat je op grond van dit gegeven de kans dat een punaise met de punt naar boven valt?

Thea werpt ook duizend keer met een punaise; de punaise belandt 298 keer met de punt naar boven.

b. Hoe groot schat je de kans dat een punaise met de punt naar boven valt, als je alle worpen (van Ralph en Thea tezamen) in ogenschouw neemt?

De in de vorige opgaven gebruikte methode om kansen toe te kennen aan 'gebeurtenissen' wordt in de praktijk veel gebruikt. Vanuit wiskundig oogpunt treden er echter problemen op. In het voorbeeld van het werpen met een dobbelsteen zullen bij een tweede serie van duizend worpen de frequentiequotiënten (met name in het begin) hoogstwaarschijnlijk andere waarden hebben. Een toevalsexperiment kan in de praktijk niet oneindig vaak worden herhaald. En zelfs als dat wel zou kunnen, dan is het niet duidelijk dat het frequentiequotiënt altijd convergeert. Ten slotte is het probleem van het bepalen van een kans met behulp van frequentiequotiënten dat deze methode niet van toepassing is op eenmalige gebeurtenissen.

2.2 Een wiskundig model

Als oplossing van de aan het eind van de vorige paragraaf genoemde bezwaren kan een *wiskundig model* van een toevalsexperiment worden gemaakt. Het betreft hier een *vertaalfase* van praktische gegevens naar modeleigenschappen. Hierbij kan het gebeuren dat verschillende personen verschillende uitgangspunten wensen te kiezen. Zodra echter een model gekozen is – en vaak ligt er één model zeer voor de hand – liggen de waarden van allerlei kansen vast.

Beschouw wederom het experiment waarbij met een dobbelsteen wordt geworpen. De verzameling van de *mogelijke uitkomsten* noemen we de *uitkomstenruimte* en noteren we met Ω; hier geldt dus $\Omega = \{1,2,3,4,5,6\}$. Een *gebeurtenis* is een uitspraak over de uitkomst en dus te identificeren met een *deelverzameling* van Ω, namelijk met de verzameling van die uitkomsten die de betreffende eigenschap hebben. De gebeurtenis 'aantal ogen is even' correspondeert met de deelverzameling $\{2,4,6\}$ van Ω.

Als een kansexperiment bestaat uit het drie keer werpen met een munt, dan is de uitkomstenruimte $\Omega = \{MMM, MMK, MKM, KMM, MKK, KMK, KKM, KKK\}$. Hierbij staat MKK voor: eerste worp munt, tweede worp kop, derde worp kop. De gebeurtenis 'twee keer kop' correspondeert met de deelverzameling $\{MKK, KMK, KKM\}$ van Ω.

Bij het dobbelsteen- en het muntexperiment ligt het op grond van symmetrie-overwegingen voor de hand om aan elke mogelijke uitkomst dezelfde kans *toe te kennen*. Dus de gebeurtenis '6 ogen' bij het werpen met een dobbelsteen heeft kans $\frac{1}{6}$ en de kans op MKK bij het drie keer werpen met een munt is gelijk aan $\frac{1}{8}$. Noteren we de gebeurtenis '6 ogen' met A (dus $A = \{6\}$), dan noteren we de *kans* op deze gebeurtenis als $\mathbf{P}(A)$, of, informeler, als $\mathbf{P}(6\text{ ogen})$. De letter P komt van het Latijnse woord *probabilitas* (= waarschijnlijkheid). Er geldt dus:

$$\mathbf{P}(A) = \tfrac{1}{6}.$$

En als we met B de gebeurtenis 'aantal ogen is even' noteren (dus $B = \{2,4,6\}$), dan geldt:

$$\mathbf{P}(B) = \tfrac{3}{6} = \tfrac{1}{2}.$$

En als we bij het muntexperiment met C de gebeurtenis 'twee keer kop' noteren (dus $C = \{MKK, KMK, KKM\}$), dan geldt:

$$\mathbf{P}(C) = \tfrac{3}{8}.$$

We hebben hierbij gebruik gemaakt van de *kansdefinitie van Laplace*.

Kansdefinitie van Laplace
Bij een kansexperiment met een eindig aantal uitkomsten die alle even waarschijnlijk zijn, is de kans op een gebeurtenis A gelijk aan

$$\mathbf{P}(A) = \frac{\#A}{\#\Omega}$$

Om $\#A$ en $\#\Omega$ te berekenen, moet vooraf worden vastgesteld welk *telmodel* er gebruikt moet worden, zoals uit de volgende voorbeelden blijkt.

Voorbeeld 1 Een pincode bestaat uit vier cijfers. Wat is de kans dat een bankpas een pincode heeft die uit vier verschillende, oneven cijfers bestaat?

Uitwerking De uitkomstenruimte Ω is de verzameling van alle rijtjes van vier cijfers. Er geldt $\#\Omega = 10^4 = 10000$. Verder is A de verzameling van alle rijtjes van vier

28 Zes meisjes en drie jongens gaan naar het theater. Ze komen op een rij van negen stoelen te zitten. Iedereen neemt willekeurig op een stoel plaats. Bereken de kans dat elke jongen tussen twee meisjes in zit.

29 Tien stoelen staan in een rij. Vijf echtparen nemen willekeurig plaats. Bereken de kans dat iedereen naast zijn/haar eigen echtgeno(o)t(e) zit.

30 Uit een volledig spel kaarten worden na elkaar twee kaarten getrokken. Bereken de kans dat de tweede kaart hoger is dan de eerste kaart. Hanteer, ongeacht de kleur van de kaarten, de volgende ordening: $2 < 3 < \cdots < 10 <$ boer $<$ vrouw $<$ heer $<$ aas.
Aanwijzing: gebruik het feit dat

$$\mathbf{P}(2^{\text{de}} \text{ kaart is hoger}) + \mathbf{P}(2^{\text{de}} \text{ is lager}) + \mathbf{P}(2^{\text{de}} \text{ is gelijk}) = 1,$$

en maak gebruik van de symmetrie-eigenschap $\mathbf{P}(2^{\text{de}} \text{ is hoger}) = \mathbf{P}(2^{\text{de}} \text{ is lager})$.

31 Een volledig spel kaarten wordt goed geschud; vervolgens worden de kaarten in de dan ontstane volgorde één voor één omgekeerd.

a. Bereken de kans dat de veertiende kaart die wordt omgekeerd, een aas is.
b. Bereken in drie decimalen nauwkeurig de kans dat de eerste aas die wordt omgekeerd, de veertiende kaart is.

32 *(Verjaardagsprobleem)* Ga bij deze opgave uit van jaren van 365 dagen.

a. Bereken in drie decimalen nauwkeurig de kans dat twee willekeurige personen op verschillende dagen jarig zijn.
b. Bereken in drie decimalen nauwkeurig de kans dat vijf willekeurige personen op verschillende dagen jarig zijn.
c. Bereken in drie decimalen nauwkeurig de kans dat in een groep van twintig personen er ten minste twee op dezelfde dag jarig zijn.
d. Uit hoeveel personen moet een groep minimaal bestaan, opdat de kans dat er ten minste twee op dezelfde dag jarig zijn, groter dan 50% is?

33 Ga ook bij deze opgave uit van jaren van 365 dagen. Bij de kassa van een theater staat een rij personen. De eerste persoon in de rij die tegelijk jarig is met één van de personen vóór hem of haar in de rij, krijgt een gratis kaartje.

a. Carl staat als vijfde in de rij. Bereken in drie decimalen nauwkeurig de kans dat hij het gratis kaartje bemachtigt.

Noteer met p_n de kans dat je het gratis kaartje bemachtigt indien je de n-de persoon in de rij bent.

b. Geef een uitdrukking voor p_n.
c. Voor welke n is de waarde van p_n maximaal? Hoe groot is die optimale kans op het bemachtigen van het gratis kaartje?

34 Vier violisten, twee altisten en twee cellisten willen op een vaste weekavond (dus geen zaterdag- of zondagavond) het octet van Mendelssohn gaan repeteren. Nu hebben die acht muzikanten, onafhankelijk van elkaar, elk één weekavond waarop ze absoluut niet kunnen. Bereken in drie decimalen nauwkeurig de kans dat ze toch een gemeenschappelijke wekelijkse repetitieavond kunnen vinden.

oppervlakte van dat deelgebied. Maak bij de hierna volgende opgaven steeds een wiskundig model (denk goed na over de keuze voor Ω) alvorens de gevraagde kans te bepalen.

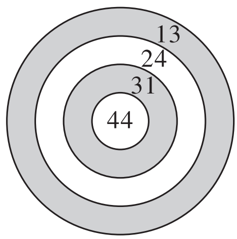
OPGAVEN

1 Hiernaast zie je een cirkelvormige putdeksel. De cirkel die met een witte rand is aangegeven, heeft een diameter die drie keer zo klein is als die van de gehele putdeksel.

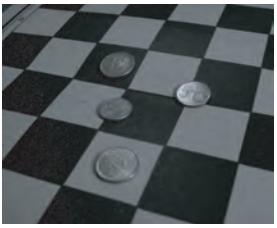
a. Bereken de kans dat de eerste druppel van een regenbui die op de putdeksel terechtkomt, binnen de witte cirkel belandt.
b. Bereken in drie decimalen nauwkeurig de kans dat van de eerste twintig druppels op de putdeksel, er acht binnen de witte cirkel belanden.

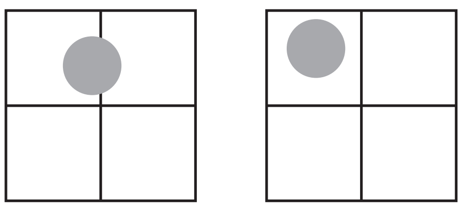
2 Op een rechthoek van 20 cm bij 50 cm wordt willekeurig een punt gekozen. Bereken de kans dat het punt minder dan 5 cm van een hoekpunt van de rechthoek ligt.

3 Een schietschijf bestaat uit vier concentrische cirkels, met stralen van 1, 2, 3 en 4. In de figuur zie je de punten die gescoord kunnen worden.
Aart werpt drie pijlen, Ben vier en Cor vijf. Bereken voor alledrie de kans op een totaalscore van 101 punten. Geef je antwoorden in drie decimalen nauwkeurig.



4 Euromunten van 10, 20 en 50 cent hebben een diameter van respectievelijk 19,75, 22,25 en 24,25 millimeter. Wanneer je een munt op een schaakbord met hokjes van 40 bij 40 millimeter laat vallen, is het mogelijk dat de munt helemaal binnen een hokje valt, zoals in de rechtertekening hieronder.





a. Bereken voor elk van de drie muntsoorten (10, 20 en 50 cent) in drie decimalen nauwkeurig de kans dat de munt helemaal binnen een hokje valt.
Aanwijzing: bezie de ligging van het middelpunt van de munt binnen een hokje.
b. Bereken de diameter van een munt die met kans 0,5 binnen een hokje valt.

12 Bij de Amerikaanse roulette is het speelveld verdeeld in 38 sectoren, die genummerd zijn van 1 tot en met 36 en verder met 0 en 00. Van de nummers 1 tot en met 36 is de helft rood, de andere helft zwart. De nummers 0 en 00 zijn groen. Per spelletje komt het balletje willekeurig op één van de 38 sectoren terecht.
Voor \$ 1 mag je inzetten op één van de nummers $1,\ldots,36$. Je krijgt \$ 36 uitbetaald indien het balletje op jouw nummer belandt, anders krijg je niets.

a. Bereken in dit geval de verwachtingswaarde van je winst.

Je mag ook voor \$ 1 inzetten op één van de kleuren rood en zwart. Je ontvangt dan \$ 2 indien het balletje op een nummer van jouw kleur belandt.

b. Bereken in dit geval de verwachtingswaarde van je winst.

Willem zet in op de kleur rood. Als hij dan wint, stopt hij. Anders speelt hij nóg twee spelletjes en zet daarbij steeds in op rood (daarna stopt hij sowieso). De stochast W is Willems (uiteindelijke) winst.

c. Bereken $\mathbf{P}(W > 0)$ in drie decimalen nauwkeurig.
d. Bereken $\mathbf{E}(W)$ in drie decimalen nauwkeurig.

Bij het berekenen van de verwachtingswaarde van een discreet verdeelde stochast X sommeer je over de elementen van de waardenverzameling van X. Als de waardenverzameling eindig is, levert de definitie van verwachtingswaarde geen problemen op. Bij oneindige waardenverzamelingen is de verwachtingswaarde een som van oneindig veel getallen; in dat geval zeggen we dat de verwachtingswaarde *bestaat* indien $\mathbf{E}(X)$ *eindig* is.

OPGAVEN

13 Zie opgave 2. Bereken $\mathbf{E}(Y)$.

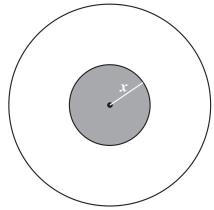
14 *(St. Petersburgparadox)* Bij een spel wordt met een munt geworpen, net zolang totdat 'kop' verschijnt. De stochast X is het aantal benodigde worpen. De uitbetaling in euro is gelijk aan $Y = 2^X$.

a. Bepaal de kansdichtheid van X. Toon aan dat de som van de kansen inderdaad 1 is met behulp van een van de reeksen uit de appendix.
b. Bereken $\mathbf{E}(X)$ met behulp van een van de reeksen uit de appendix.
c. Laat zien dat $\mathbf{E}(Y)$ niet bestaat. (Zou jij bereid zijn € 1000 als inzet te betalen?)
d. Bereken in drie decimalen nauwkeurig de kans dat je winst maakt bij een inzet van € 1000.

15 Gegeven is een discreet verdeelde stochast X met $W_X = \{1,2,3,\ldots\}$.

Bewijs dat $\mathbf{E}(X) = \sum_{n=1}^{\infty} \mathbf{P}(X \geq n)$.

16 Zie opgave 2 en 13. Bereken nogmaals $\mathbf{E}(Y)$, nu met behulp van de bij opgave 15 bewezen formule.



Uitwerking Zie de figuur hiernaast. De gehele schietschijf is de uitkomstenruimte Ω; het grijze cirkeltje met straal x noemen we A.
Er geldt: $\mathbf{P}(X \leq x) = \dfrac{\text{opp}(A)}{\text{opp}(\Omega)} = \dfrac{\pi x^2}{\pi} = x^2$, dus de verdelingsfunctie van X is:

$$F(x) = \begin{cases} 0, & \text{voor } x < 0; \\ x^2, & \text{voor } 0 \leq x \leq 1; \\ 1, & \text{voor } x > 1. \end{cases}$$

Omdat voor de kansdichtheid f geldt dat $f(x) = F'(x)$, heeft f het volgende functievoorschrift:

$$f(x) = \begin{cases} 2x, & \text{voor } 0 < x < 1; \\ 0, & \text{anders.} \end{cases} \qquad \square$$

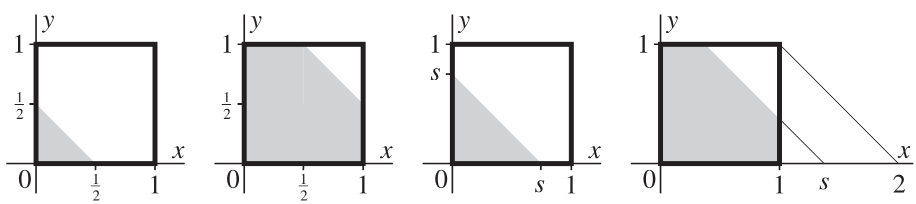
Opmerking: de reden dat we schrijven $f(x) = 2x$ voor $0 < x < 1$, en niet voor $0 \leq x \leq 1$, is gelegen in het feit dat F niet differentieerbaar *is in de punten $x = 0$ en $x = 1$. In dergelijke punten doet het er niet toe welke waarde $f(x)$ heeft. Het is echter gebruikelijk om voor die punten $f(x)$ gelijk aan nul te stellen. In feite mag je voor eindig veel waarden x de waarde van $f(x)$ vrij kiezen, de integraal van f verandert daardoor niet. De functie f ligt dus nooit helemaal vast. We kiezen in de regel voor de meest voor de hand liggende (dat wil zeggen: continue) keuze van f, en spreken dan van* de *kansdichtheid van X in plaats van* een *kansdichtheid.*

Voorbeeld 4 Een punt A wordt willekeurig gekozen binnen het eenheidsvierkant. De stochast S is de som van de x- en de y-coördinaat van A.

a. Bereken $\mathbf{P}(S \leq \frac{1}{2})$.
b. Bereken $\mathbf{P}(S \leq 1\frac{1}{2})$.
c. Bepaal de verdelingsfunctie van S en schets de grafiek van deze verdelingsfunctie.
   Aanwijzing: maak bij het bepalen van $F_S(s)$ onderscheid tussen de gevallen $0 \leq s \leq 1$ en $1 \leq s \leq 2$.
d. Bepaal de kansdichtheid van S en schets de grafiek van deze kansdichtheid.

Uitwerking In de figuren hieronder is Ω steeds met een dikke rand aangegeven.

a. In het eerste plaatje bestaat het grijze gebied uit alle punten (binnen Ω) waarvoor geldt dat $S \leq \frac{1}{2}$; uit deze figuur blijkt onmiddellijk dat $\mathbf{P}(S \leq \frac{1}{2}) = \frac{1}{8}$.
b. In het tweede plaatje bestaat het grijze gebied uit alle punten (binnen Ω) waarvoor geldt dat $S \leq 1\frac{1}{2}$; uit deze figuur blijkt onmiddellijk dat $\mathbf{P}(S \leq 1\frac{1}{2}) = \frac{7}{8}$.
c. In het derde plaatje bestaat het grijze gebied uit alle punten (binnen Ω) waarvoor geldt dat $S \leq s$ voor $0 \leq s \leq 1$; uit deze figuur blijkt dat $\mathbf{P}(S \leq s) = \frac{1}{2}s^2$.
   In het vierde plaatje bestaat het grijze gebied uit alle punten (binnen Ω) waarvoor geldt dat $S \leq s$ voor $1 \leq s \leq 2$; uit deze figuur blijkt dat $\mathbf{P}(S \leq s) = 1 - \frac{1}{2}(2 - s)^2$.



Voorbeeld 9 De stochast X is normaal verdeeld met $\mu = 10$ en $\sigma = 1{,}8$.

a. Bereken a in twee decimalen nauwkeurig, indien gegeven is dat $\mathbf{P}(X < a) = 0{,}55$.
b. Bereken b in twee decimalen nauwkeurig, indien gegeven is dat $\mathbf{P}(X > b) = 0{,}27$.

Uitwerking

a. $a \approx 10{,}23$ (via `invnorm(.55,10,1.8)`).
b. $b \approx 11{,}10$ (via `invnorm(.73,10,1.8)`; let op: gegeven is dat de oppervlakte *rechts* van b gelijk is aan 0,27; dat wil zeggen dat de oppervlakte *links* van b gelijk is aan $1 - 0{,}27 = 0{,}73$). □

OPGAVEN

44 Een fabrikant maakt gloeilampen waarvan de brandtijd normaal verdeeld is met $\mu = 100$ uur en $\sigma = 7$ uur. Hoe lang branden de lampen die tot de 5% lampen met de langstlevende levensduur horen minimaal?

45 Een fabrikant maakt moeren waarvan de diameter normaal verdeeld is met $\mu = 20$ mm en $\sigma = 0{,}5$ mm. De 4% van de moeren waarvan de diameter het meest afwijkt van het gemiddelde, wordt vernietigd. Wat weet je van de diameter van deze moeren?

46 Van een grote partij forellen is de lengte normaal verdeeld met $\mu = 27{,}5$ cm en $\sigma = 3{,}1$ cm. De forellen worden in drie klassen verdeeld: ‘klein’, ‘normaal’ en ‘groot’. Elke klasse bevat evenveel forellen. Bereken de grenzen van de drie klassen.

In sommige gevallen is de waarde van een van de parameters μ en σ onbekend. In zo’n geval kun je aan de hand van de gegevens een *vergelijking* opstellen en die *grafisch-numeriek* oplossen. Hoe dat gaat, zie je in het volgende voorbeeld.

Voorbeeld 10 De stochast X is normaal verdeeld met $\mu = 20$. Bereken σ in één decimaal nauwkeurig, indien gegeven is dat $\mathbf{P}(16 < X < 22) = 0{,}4$.

Uitwerking Voer in `y1=normalcdf(16,22,20,x)` en `y2=.4`. Plot de grafieken (kies een geschikt *window*!) en gebruik de optie ‘intersect’. Je vindt $\sigma \approx 5{,}6$. □

Voorbeeld 11 De stochast X is normaal verdeeld met $\sigma = 4$. Bereken μ in één decimaal nauwkeurig, indien gegeven is dat $\mathbf{P}(X < 18) = 0{,}55$.

Uitwerking Voer in `y1=normalcdf(-E99,18,x,4)` en `y2=.55`. Plot de grafieken en gebruik de optie ‘intersect’. Je vindt $\mu \approx 17{,}5$. Je kunt ook `y1=invnorm(.55,x,4)` en `y2=18` invoeren; dat geeft uiteraard hetzelfde resultaat. □

OPGAVEN

47 Een machine vult flessen wijn van 75 cl. De inhoud van de flessen is normaal verdeeld met een standaardafwijking van 2 cl. De machine kan op elk gewenst gemiddelde worden afgesteld, zonder dat daarbij de standaardafwijking verandert. Op welk gemiddelde moet de machine worden afgesteld, opdat niet meer dan 8% van de flessen minder dan 75 cl bevat? Rond je antwoord af op twee decimalen.

Voorbeeld 1 Om te beoordelen of een dobbelsteen zuiver is, besluit je er 600 keer mee te gooien. Als het aantal keren 6 ogen ten minste 20 afwijkt ten opzichte van wat je zou mogen verwachten bij een zuivere dobbelsteen, dan besluit je te concluderen dat de dobbelsteen onzuiver is.

a. Geef met behulp van de ongelijkheid van Chebyshev een bovengrens voor de kans dat je *ten onrechte* besluit dat de dobbelsteen onzuiver is. Geef je antwoord in drie decimalen nauwkeurig.
b. Bereken in drie decimalen nauwkeurig de kans dat je ten onrechte besluit dat de dobbelsteen onzuiver is.

Uitwerking De stochast X is het aantal keer 6 ogen bij 600 worpen met de dobbelsteen. Als de dobbelsteen zuiver is, dan is X binomiaal verdeeld met $n = 600$ en $p = \frac{1}{6}$.

a. Er geldt: $\mu = np = 100$ en $\sigma^2 = np(1-p) = \frac{250}{3}$. Met de ongelijkheid van Chebyshev vinden we dat

$$\mathbf{P}(|X-100| \geq 20) \leq \frac{\left(\frac{250}{3}\right)}{400} = \frac{5}{24} \approx 0{,}208.$$

b. De kans dat de dobbelsteen als onzuiver wordt bestempeld, terwijl hij in feite zuiver is, is gelijk aan $\mathbf{P}(X \leq 80) + \mathbf{P}(X \geq 120) \approx 0{,}01447 + 0{,}01801 \approx 0{,}032$. □

In het voorbeeld is de bovengrens die met behulp van de ongelijkheid van Chebyshev werd gevonden, zeer grof. Omdat de ongelijkheid van Chebyshev in het algemeen geen scherpe grens geeft, is het praktisch belang ervan gering. De ongelijkheid is echter wél heel nuttig bij het bewijs van de *zwakke wet van grote aantallen*.

Zwakke wet van grote aantallen

Als $X_1, X_2, X_3, \ldots$ onafhankelijke stochasten zijn, elk met verwachting μ, dan geldt voor elke $\varepsilon > 0$:

$$\lim_{n\to\infty} \mathbf{P}(|\overline{X}_n - \mu| < \varepsilon) = 1$$

OPGAVEN

5 Veronderstel dat $X_1, X_2, X_3, \ldots$ onafhankelijke stochasten zijn, elk met verwachting μ en variantie σ^2. Hoe volgt in dat geval de zwakke wet van grote aantallen uit de ongelijkheid van Chebyshev?

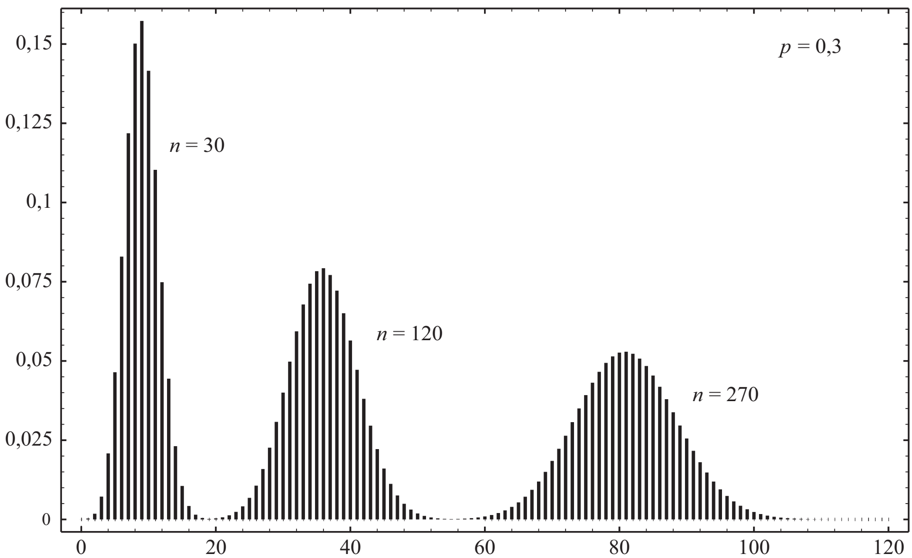
6 Je werpt n keer met een munt waarvan de kans op ‘kop’ gelijk is aan p en de kans op ‘munt’ gelijk is aan $1-p$. De stochast $\overline{X}_n$ is het gemiddeld aantal keren ‘kop’.

a. Bepaal $\mathbf{E}(\overline{X}_n)$ en $\mathrm{Var}(\overline{X}_n)$.
b. Voor welke waarde van p is $p(1-p)$ maximaal? Wat is dat maximum?
c. Bewijs dat $\mathbf{P}(|\overline{X}_n - p| \geq \varepsilon) \leq \dfrac{1}{4n\varepsilon^2}$.

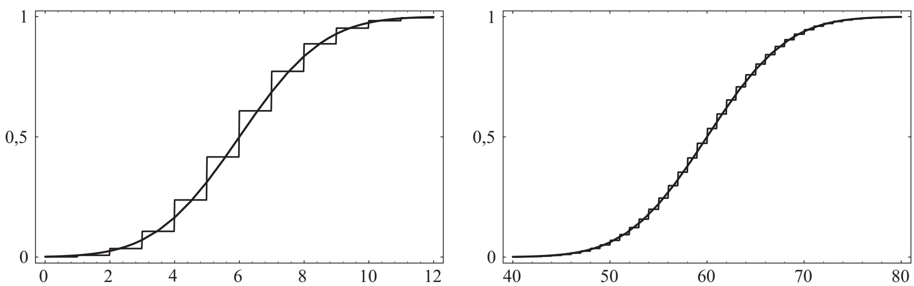
De Centrale Limietstelling wordt wel eens aangemerkt als een van de markantste inzichten die binnen de wiskunde zijn bereikt. De praktische betekenis ervan kan nauwelijks worden overschat. Het bewijs van deze opmerkelijke stelling is stof voor hogerejaarsstudenten wiskunde. Een speciaal geval van deze stelling kunnen we echter wel bewijzen; zie hiervoor de laatste opgave van dit hoofdstuk.

Een gevolg van de Centrale Limietstelling is dat voor grote waarden van n de verdelingsfuncties van de som S_n en het gemiddelde $\overline{X}_n$ goed benaderd kunnen worden door de verdelingsfuncties van de normale verdeling met parameters $n\mu$ en $\sigma\sqrt{n}$ respectievelijk μ en $\sigma/\sqrt{n}$.

Hieronder zie je drie kanshistogrammen. Ze horen bij de binomiale verdeling met $p = 0{,}3$; van links naar rechts is de waarde van n achtereenvolgens 30, 120 en 270. De normale verdeling wordt steeds beter herkenbaar.



In onderstaande figuur zie je de grafieken van de verdelingsfuncties van de binomiale verdeling met $p = 0{,}3$. In de linkersituatie is $n = 20$ en in de rechtersituatie is $n = 200$. Erdoorheen is de verdelingsfunctie van de bijbehorende normale verdeling getekend; in de linkersituatie zijn de parameters $\mu = 20 \cdot 0{,}3 = 6$ en $\sigma = \sqrt{20 \cdot 0{,}3 \cdot 0{,}7} \approx 2{,}05$ en in de rechtersituatie geldt $\mu = 200 \cdot 0{,}3 = 60$ en $\sigma = \sqrt{200 \cdot 0{,}3 \cdot 0{,}7} \approx 3{,}55$.



44 De stochast X is homogeen verdeeld op het interval $[1,2]$ en de stochast Y is gedefinieerd als $Y = 1/X$.

a. Bepaal de waardenverzameling van Y.
b. Bepaal de verdelingsfunctie van Y.
c. Bepaal de kansdichtheid van Y.
d. Bereken $\mathbf{E}(Y)$ met behulp van de kansdichtheid van Y.
e. Bereken $\mathbf{E}(Y)$ nogmaals, nu met behulp van de substitutiemethode.

45 Bij de Europese roulette is het speelveld verdeeld in 37 sectoren, die genummerd zijn van 0 tot en met 36. Van de nummers 1 tot en met 36 is de helft rood, de andere helft zwart. Het nummer 0 is groen. Per spelletje komt het balletje willekeurig op één van de 37 sectoren terecht.

a. Voor € 1 mag je inzetten op één van de kleuren rood en zwart. Je ontvangt € 2 indien het balletje op een nummer van jouw kleur belandt, anders ben je je inzet kwijt. Bereken de verwachtingswaarde van je winst W.
b. Bereken de verwachtingswaarde van het aantal spelletjes X dat nodig is totdat het balletje op nummer 0 belandt.

Een gokker gaat met jou de weddenschap aan dat bij de eerste n spelletjes het balletje ten minste één keer op nummer 0 belandt.

c. Stel $n = \mathbf{E}(X)$. Bereken de kans dat de gokker deze weddenschap wint.
d. Bereken voor welke waarden van n deze weddenschap voor jou gunstig is, dat wil zeggen: de kans dat je deze weddenschap wint is groter dan $\frac{1}{2}$.

46 Een punt A wordt blindelings gekozen in het eenheidsinterval. Dit punt verdeelt het eenheidsinterval in twee stukken die vervolgens als de twee (aangrenzende) zijden van een rechthoek worden genomen. De stochast X is de afstand van 0 tot A en de stochast Y is de oppervlakte van de rechthoek.

a. Bepaal de verdelingsfunctie en de kansdichtheid van X.
b. Druk Y uit in X en bepaal de waardenverzameling van Y.
c. Bereken $\mathbf{P}(Y \leq \frac{1}{8})$.
d. Bepaal de verdelingsfunctie van Y en schets de grafiek van deze verdelingsfunctie.
e. Bepaal de kansdichtheid van Y en schets de grafiek van deze kansdichtheid.
f. Bereken $\mathbf{E}(Y)$, $\mathrm{Var}(Y)$ en σ_Y met behulp van de kansdichtheid van Y. Geef je antwoorden in drie decimalen nauwkeurig.
g. Bereken $\mathbf{E}(Y)$, $\mathrm{Var}(Y)$ en σ_Y nogmaals, nu exact, met de substitutieregel.

47 Een punt A wordt blindelings gekozen in het eenheidsinterval. Dit punt verdeelt het eenheidsinterval in twee stukken. De stochast X is de afstand van 0 tot A. Verder definiëren we de stochast $Z = X/(1-X)$; in woorden: Z is het quotiënt van het linkerstuk en het rechterstuk.

a. Bepaal de waardenverzameling van Z.
b. Bepaal de verdelingsfunctie van Z en schets de grafiek van deze verdelingsfunctie.
c. Bepaal de kansdichtheid van Z en schets de grafiek van deze kansdichtheid.
d. Bereken $\mathbf{P}(2 < Z < 4)$.
e. Toon aan dat $\mathbf{E}(Z)$ niet bestaat.

Index