Kanstheorie

2de bachelor wiskunde
Vrije Universiteit Brussel

U. Einmahl

Academiejaar 2016/2017

Inhoudsopgave

Hoofdstuk 1

Kansruimten

1.1 Toevallige experimenten

Een **toevallig experiment** is een experiment waar men de uitkomst niet volledig kan voorspellen.

Voorbeelden

1. Het werpen van een muntstuk met de mogelijke uitkomsten:
   (K) kop, (M) munt.

2. Het werpen van een dobbelsteen met de mogelijke uitkomsten (aantal ogen):
   1, 2, 3, 4, 5, 6.

3. Het installeren van een gloeilamp. Men weet niet hoe lang deze zal houden. De leeftijd T van de gloeilamp hangt af van het toeval: $T \in]0, \infty[$.

4. Het volgende experiment waar we een bol in een vacuüm laten vallen en we willen weten hoe lang het duurt tot dat deze de grond bereikt is **geen** toevallig experiment, omdat we uit de natuurkunde precies weten hoe lang dit duurt. (We veronderstellen hier natuurlijk dat we de (initiële) afstand van de bol tot de grond kennen.)

Gegeven een toevallig experiment, noteren we de verzameling van de mogelijke uitkomsten altijd met Ω. We noemen deze de **uitkomstenverzameling** (of sample space) van het experiment. Verder kiezen we voor elk experiment een klasse van deelverzamelingen van Ω die we met $\mathcal{F}$ noteren. De verzamelingen in $\mathcal{F}$ heten de **gebeurtenissen**. Als Ω ten hoogste aftelbaar is, kunnen we $\mathcal{F} = 2^\Omega$ (al de deelverzamelingen van Ω) stellen, maar in het algemeen is dit niet mogelijk.

Voorbeelden:

1. Het werpen van twee muntstukken: Dan is $\Omega = \{(K,K),(K,M),(M,K),(M,M)\}$. De gebeurtenis "1 keer kop, 1 keer munt"wordt voorgesteld door de verzameling $A = \{(K,M),(M,K)\}$.

2. Het werpen van een dobbelsteen: Dan is het evident dat $\Omega = \{1,\ldots,6\}$ en de verzameling $A = \{2,4,6\}$ is de gebeurtenis dat het aantal ogen even is.

Een **kans** is een afbeelding van $\mathcal{F}$ naar $[0,1]$ die met elke gebeurtenis A zijn kans $\mathbb{P}(A)$ associeert. De vraag is nu hoe men dit op een wiskundige manier kan definiëren.

Uitgaande van de intuïtie, is de volgende definitie via "relatieve frequenties"voor de hand liggend:

Veronderstel dat men een bepaald experiment kan herhalen zodat de uitkomsten "onafhankelijk van elkaar" zijn. Stel n_A = het aantal van de eerste n experimenten met een uitkomst in A. Dan kan men verwachten dat de relatieve frequenties n_A/n convergeren en de limiet gelijk aan $\mathbb{P}(A)$, de kans op A is. Hoewel dit klopt, kunnen we daarop geen theorie baseren omdat we niet precies weten wat "onafhankelijk van elkaar" betekent.

Er bestaat echter een axiomatische definitie van het begrip"kans"zodat we exact kunnen definiëren wat onafhankelijkheid betekent. In dit kader kunnen we dan bepaalde limietstellingen bewijzen. Een van deze stellingen, de sterke wet van de grote getallen, impliceert dan dat voor elke gebeurtenis $A \in \mathcal{F}$ de relatieve frequenties n_A/n naar $\mathbb{P}(A)$ convergeren.

1.2 De axioma's van Kolmogorov

Zij Ω een niet-lege verzameling.

Definitie 1.1 *Een klasse $\mathcal{F}$ van deelverzamelingen van Ω heet een σ-algebra op Ω als $\mathcal{F}$ voldoet aan de volgende drie voorwaarden:*
(i) $\Omega \in \mathcal{F}$.
(ii) $A \in \mathcal{F} \Rightarrow A^c := \Omega \setminus A \in \mathcal{F}$.
(iii) $A_n \in \mathcal{F}, n \geq 1 \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Eigenschappen: Als $\mathcal{F}$ een σ-algebra is, geldt:

1. $\varnothing \in \mathcal{F}$. (Dit is triviaal wegens $\varnothing = \Omega^c$.)

2. $A_1, \ldots, A_m \in \mathcal{F} \Rightarrow \bigcup_{j=1}^{m} A_j \in \mathcal{F}$.
(Stel $A_j = \varnothing$, $j \geq m+1$ en gebruik (iii).)

3. $A_1, \ldots, A_m \in \mathcal{F} \Rightarrow \bigcap_{j=1}^{m} A_j \in \mathcal{F}$.
(Dit is evident wegens $\bigcap_{j=1}^{m} A_j = (\bigcup_{j=1}^{m} A_j^c)^c$, (ii) en (2).)

4. $A_n \in \mathcal{F}, n \geq 1 \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$.
(Analoog argument als in (3).)

5. $A, B \in \mathcal{F} \Rightarrow A \setminus B \in \mathcal{F}$ en $A \triangle B \in \mathcal{F}$,
waar $A \triangle B = (A \setminus B) \cup (B \setminus A)$ het symmetrische verschil van de twee gebeurtenissen A, B is. (Dit betekent dat precies één van de twee gebeurtenissen optreedt.)
Om (5) te bewijzen noteren we dat $A \setminus B = A \cap B^c$ wegens (ii) en (3) in $\mathcal{F}$ zit. Eigenschap (2) impliceert dan ook $A \triangle B \in \mathcal{F}$.

Opmerking: Uit het bovenstaande blijkt dat verzamelingen die men na een aftelbaar aantal operaties zoals $\cap, \cup, \triangle, ^c$ uit $\mathcal{F}$ kan verkrijgen, nog tot $\mathcal{F}$ behoren. Dit is in het algemeen fout, als men meer dan aftelbaar veel operaties gebruikt.

Definitie 1.2 *Zij $\mathcal{F}$ een σ-algebra op Ω.*

(i) *Een afbeelding $\mathbb{P} : \mathcal{F} \to [0,1]$ wordt een* **kans** *(of kansmaat) genoemd indien*

(a) $\mathbb{P}(\Omega) = 1$

(b) $A_n \in \mathcal{F}, n \geq 1$ *(paarsgewijs) disjunct (i.e., $A_n \cap A_m = \varnothing, n \neq m$)* $\Rightarrow \mathbb{P}(\dot\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$. *($\sigma$-additiviteit)*

(ii) *Een kansruimte is een drietal $(\Omega, \mathcal{F}, \mathbb{P})$ bestaande uit een verzameling Ω, een σ-algebra $\mathcal{F}$ op Ω en een kans $\mathbb{P}$.*

Enkele eigenschappen van kansmaten:

1. $\mathbb{P}(\varnothing) = 0$.
(Stel $A_n = \varnothing, n \geq 1$. Dan geldt: $\mathbb{P}(A_1) = \mathbb{P}(\dot\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$, wat natuurlijk impliceert: $0 = \sum_{n=2}^{\infty} \mathbb{P}(A_n) = \mathbb{P}(\dot\bigcup_{n=2}^{\infty} A_n) = \mathbb{P}(\varnothing)$.)

2. (*additiviteit*) $A_1, \dots, A_m \in \mathcal{F}$ disjunct, $\Rightarrow \mathbb{P}(\dot\bigcup_{i=1}^{m} A_i) = \sum_{i=1}^{m} \mathbb{P}(A_i)$.
(Beschouw de rij $A_1, \dots, A_m, \varnothing, \varnothing, \dots$ en gebruik (1) in verband met de σ-additiviteit van $\mathbb{P}$.)

3. $A \in \mathcal{F} \Rightarrow \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
($1 = \mathbb{P}(\Omega) = \mathbb{P}(A \dot\cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c)$.)

4. $A, B \in \mathcal{F}$ en $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$.
(Gebruik het feit dat $B = A \dot\cup (B \smallsetminus A)$ en de additiviteit van $\mathbb{P}$.)

Verder hebben we:

5. $A, B \in \mathcal{F} \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

6. Als $A_1, \dots, A_m \in \mathcal{F}$, waar $m \geq 2$, dan geldt er: $\mathbb{P}(\cup_{i=1}^{m} A_i) \leq \sum_{i=1}^{m} \mathbb{P}(A_i)$.

7. Als $A_1, \dots, A_m \in \mathcal{F}$, waar $m \geq 3$, dan geldt ook:

$$\mathbb{P}(\bigcup_{i=1}^{m} A_i) = \sum_{r=1}^{m} (-1)^{r+1} \sum_{1 \leq i_1 < \ldots < i_r \leq m} \mathbb{P}(\bigcap_{j=1}^{r} A_{i_j}).$$

("Formule van het in- en uitsluiten")

We bewijzen (5). Eigenschappen (6) en (7) volgen dan via volledige inductie (oefening). Gezien $A \cup B = A \dot\cup (B \smallsetminus A)$, waar de twee laatste gebeurtenissen disjunct zijn, geldt $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \smallsetminus A)$. Verder is B gelijk aan de unie van de twee disjuncte gebeurtenissen $B \smallsetminus A$ en $A \cap B$ en dus geldt: $\mathbb{P}(B) = \mathbb{P}(B \smallsetminus A) + \mathbb{P}(A \cap B)$ of $\mathbb{P}(B \smallsetminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$, wat in verband met de bovenstaande formule voor $\mathbb{P}(A \cup B)$ eigenschap (5) impliceert.

Definitie 1.3 *Een rij van verzamelingen $A_n, n \geq 1$ heet* **stijgend** *(Notatie: $A_n \nearrow$) indien geldt: $A_1 \subset A_2 \subset A_3 \subset \dots$ Analoog noemen we en rij $A_n, n \geq 1$* **dalend** *indien $A_1 \supset A_2 \supset A_3 \supset \dots$ (Notatie: $A_n \searrow$) In beide gevallen zeggen we dat de rij* **monotoon** *is.*
Verder definiëren we:

$$\lim_{n \to \infty} A_n = \begin{cases} \bigcup_{n=1}^{\infty} A_n & \text{als } A_n \nearrow \\ \bigcap_{n=1}^{\infty} A_n & \text{als } A_n \searrow \end{cases}$$

Stelling 1.1 (Kansmaten zijn continu van beneden en boven)
Zij $\mathbb{P}$ een kans op een σ-algebra $\mathcal{F}$ en zij $A_n, n \geq 1$ een monotone rij in $\mathcal{F}$. Dan geldt als $n \to \infty$,

$$\mathbb{P}(A_n) \to \mathbb{P}(\lim_{n\to\infty} A_n).$$

Bewijs. (i) ($\mathbb{P}$ is continu van beneden) Dus zij $A_n \nearrow$ en stel $A = \bigcup_{n=1}^{\infty} A_n$. Dan zijn de gebeurtenissen $B_1 := A_1, B_n := A_n \setminus A_{n-1}, n \geq 2$ disjunct en bovendien geldt:

$$\dot{\bigcup_{i=1}^{\infty}} B_i = A, \ \dot{\bigcup_{i=1}^{n}} B_i = A_n, n \geq 1.$$

Dus volgt uit de σ-additiviteit van $\mathbb{P}$,

$$\mathbb{P}(A) = \mathbb{P}(\dot{\bigcup_{i=1}^{\infty}} B_i) = \sum_{i=1}^{\infty} \mathbb{P}(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{P}(B_i) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

(ii) ($\mathbb{P}$ is continu van boven.) Zij $C_n \searrow$ en stel $C = \bigcap_{n=1}^{\infty} C_n$. Dan geldt natuurlijk $A_n = C_n^c \nearrow$ en $\lim_n A_n = \bigcup_n C_n^c = (\bigcap_n C_n)^c = C^c$. Dus volgt uit deel (i) dat $1 - \mathbb{P}(C_n) = \mathbb{P}(C_n^c) \to \mathbb{P}(C^c) = 1 - \mathbb{P}(C)$ en bijgevolg $\mathbb{P}(C_n) \to \mathbb{P}(C)$ als $n \to \infty$. □

Als eerste toepassing van stelling 1.1, noteren we dat kansmaten σ-subadditief zijn, dwz

$$A_n \in \mathcal{F}, n \geq 1 \Rightarrow \mathbb{P}(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n). \tag{1.1}$$

Dit volgt direct via eigenschap 6 van kansmaten, omdat $\bigcup_{m=1}^{n} A_n \nearrow \bigcup_{n=1}^{\infty} A_n$ als $n \to \infty$.
Zoals voor rijen reëele getallen, bestaan ook een **limes inferior** en in **limes superior** voor rijen van gebeurtenissen. Deze zijn zoals volgt gedefinieerd:

Definitie 1.4 *Zij $A_n, n \geq 1$ een rij in $\mathcal{F}$. Dan stellen we*

$$\liminf_{n\geq 1} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k = \{\omega \in \Omega : w \in A_n \ \textit{uiteindelijk}\}$$

en

$$\limsup_{n\geq 1} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \{\omega \in \Omega : \omega \in A_n \ \textit{oneindig vaak}\}.$$

Het is evident dat $\liminf_{n\geq 1} A_n$ en $\limsup_{n\geq 1} A_n$ weer in $\mathcal{F}$ zitten. Verder geldt er

$$\left(\liminf_{n\geq 1} A_n\right)^c = \limsup_{n\geq 1} A_n^c$$

en analoog

$$\left(\limsup_{n\geq 1} A_n\right)^c = \liminf_{n\geq 1} A_n^c.$$

Uit bovenstaande definitie volgt verder dat $\bigcap_{k=n}^{\infty} A_k \nearrow \liminf_{n\geq 1} A_n$ en $\bigcup_{k=n}^{\infty} A_k \searrow \limsup_{n\geq 1} A_n$, hetgeen via stelling 1.1 impliceert dat

$$\mathbb{P}(\liminf_{n\geq 1} A_n) = \lim_{n\to\infty} \mathbb{P}(\bigcap_{k=n}^{\infty} A_k)$$

en

$$\mathbb{P}(\limsup_{n\geq 1} A_n) = \lim_{n\to\infty} \mathbb{P}(\bigcup_{k=n}^{\infty} A_k).$$

Het volgende bijna triviale lemma zal in het vervolg zeer nuttig zijn.

Lemma 1.1 (Borel-Cantelli, deel 1)
Zij A_n een rij gebeurtenissen zodanig dat $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$. Dan geldt er: $\mathbb{P}(\limsup_{n\geq 1} A_n) = 0$.

Bewijs Wegens de σ-subadditiviteit van $\mathbb{P}$ hebben we:

$$\mathbb{P}(\limsup_{n\geq 1} A_n) = \lim_{n\to\infty} \mathbb{P}(\bigcup_{k=n}^{\infty} A_k) \leq \limsup_{n\to\infty} \sum_{k=n}^{\infty} \mathbb{P}(A_k).$$

Deze laatste lim sup is gelijk aan 0 omdat $\sum_{k=n}^{\infty} \mathbb{P}(A_k)$ naar 0 convergeert als $n \to \infty$ aangezien de reeks $\sum_{n=1}^{\infty} \mathbb{P}(A_n)$ convergent is. □

1.3 Eindige kansruimten

Veronderstel dat Ω eindig is met N elementen $\omega_1, \ldots, \omega_N$. Zij $p_i \in]0,1], 1 \leq i \leq N$ zodanig dat $\sum_i p_i = 1$. Dan is

$$\mathbb{P}(A) := \sum_{i:\omega_i \in A} p_i = \sum_{i=1}^{N} p_i I_A(\omega_i),\ A \subset \Omega$$

een kans op $\mathcal{F} = 2^{\Omega}$.

We definiëren voor elke verzameling $A \subset \Omega : I_A(\omega) = \begin{cases} 1 & \text{als } \omega \in A \\ 0 & \text{als } \omega \in A^c \end{cases}$ en we noemen I_A de **indicatorfunctie** van A.

Als al de uitkomsten $\omega_1, \ldots, \omega_N$ dezelfde kans hebben, verkrijgen we een **uniforme** kansruimte. In dit geval geldt: $1 = \mathbb{P}(\Omega) = \sum_{i=1}^{N} \mathbb{P}(\{\omega_i\}) = N\mathbb{P}(\{\omega_1\})$ en bijgevolg $\mathbb{P}(\{\omega_1\}) = \mathbb{P}(\{\omega_i\}) = p_i = 1/N, 2 \leq i \leq N$ en het is evident dat

$$\mathbb{P}(A) = \#A/\#\Omega,$$

waar $\#B$ het aantal elementen van de verzameling B betekent. We kunnen dan al de kansen via de combinatoriek berekenen.

Voorbeelden (1) In een doos zitten 10 ballen waarop nummers van 1 t/m 10 staan. We trekken 2 ballen en noteren de nummers. Wat is de kans dat het verschil tussen de twee getrokken nummers tenminste 2 is, als we
(a) de eerste getrokken bal weer terugleggen "sampling with replacement"),
(b) niet terugleggen "sampling without replacement")?

Oplossing. (a) Stel $\Omega_a = \{1, \ldots, 10\}^2$ (= alle paren (i,j) met $i,j \in \{1,\ldots,10\}$.)
Dan geldt natuurlijk $\#\Omega_a = 100$ en als we $A = \{(i,j) \in \Omega_a : |i-j| \geq 2\}$ stellen, geldt $A^c = \{(i,i) : 1 \leq i \leq 10\} \cup \{(1,2),(2,1),(2,3),\ldots,(9,8),(9,10),(10,9)\}$ en dus heeft A^c 10+18 elementen wat natuurlijk impliceert dat $\mathbb{P}(A^c) = 0,28$ en bijgevolg $\mathbb{P}(A) = 0,72$.
(b) Stel $\Omega_b = \{(i,j) \in \{1,\ldots,10\}^2 : i \neq j\}$. Dan heeft Ω_b 90 elementen en verder geldt als we B gelijk aan $\{(i,j) \in \Omega_b : |i-j| \geq 2\}$ stellen, $\#B^c = 18$ en dus $\mathbb{P}(B) = 72/90 = 0,8$.

(2) (*Het verjaardagenprobleem*) Wat is de kans dat in een groep van m personen allen een verschillende verjaardag hebben? Onderstel dat alle dagen van het jaar even waarschijnlijk zijn (niemand jarig op 29 februari is), en dat er geen verband tussen de verjaardagen van deze personen bestaat. (Als er een tweeling is, mag maar één van hun meedoen).

Oplossing. Stel $\Omega_m = \{1, \ldots, 365\}^m$. (Als $m = 3$ betekent bv $(25, 119, 33)$ dat de eerste persoon jarig is op 25 januari, de tweede op 29 april en de derde op 2 februari.) Onze voorwaarden zijn zo dat we van een uniforme kansruimte kunnen uitgaan. De gebeurtenis "m verschillende verjaardagen" wordt voorgesteld door $A_m = \{(i_1, \ldots, i_m) \in \Omega_m : i_j \neq i_k, j \neq k\}$. Dus is de gezochte kans gelijk aan

$$\#A_m/\#\Omega_m = 365 \cdot \ldots \cdot (365 - m + 1)/365^m = \prod_{j=1}^{m-1}(1 - j/365).$$

Deze kans is kleiner dan 1/2 zodra $m \geq 23$.

(3) (*Het ontmoetingsprobleem.*) N paren gaan naar een feest waar gedanst wordt. De gastheer vraagt zowel de mannen als de vrouwen nummers te trekken. Elke vrouw danst met de man die een identiek nummer heeft. Wat is de kans dat (a) geen enkel paar samen blijft en (b) precies k paren samen blijven?

Oplossing. Stel $\Omega =$ alle permutaties van $1, \ldots, N$, waar de permutatie $(j_1, \ldots, j_N)$ betekent dat de vrouw die het nummer i heeft getrokken met de man van de vrouw danst die het nummer j_i heeft getrokken.
(a) Stel verder $A_i = \{$de vrouw met het nummer i danst met haar man$\}$, $1 \leq i \leq N$.
Dan geldt: $\mathbb{P}(\bigcup_{j=1}^N A_j) = \mathbb{P}(\text{tenminste een paar blijft samen}) = 1 - \mathbb{P}(\text{geen paar blijft samen})$.
We veronderstellen dat elke permutatie even waarschijnlijk is. Dan is het evident dat

$$\mathbb{P}(A_i) = \#\{(j_1, \ldots, j_N) \in \Omega : j_i = i\}/N! = (N-1)!/N! = 1/N, \ 1 \leq i \leq N$$

en algemeen voor $i_1 < \ldots < i_r$ en $r \geq 2 : \mathbb{P}(\bigcap_{k=1}^r A_{i_k}) = (N - r)!/N!$, hetgeen via de formule van het in- en uitsluiten impliceert:

$$\mathbb{P}(\bigcup_{i=1}^N A_i) = \sum_{r=1}^N (-1)^{r+1} \binom{N}{r} \frac{(N-r)!}{N!} = \sum_{r=1}^N (-1)^{r+1} \frac{1}{r!}.$$

We kunnen nu concluderen dat

$$\mathbb{P}(\text{geen paar blijft samen}) = \sum_{r=0}^N (-1)^r \frac{1}{r!}.$$

De laatste som kunnen we voor $N \geq 5$ door $e^{-1} = \sum_{r=0}^\infty (-1)^r/r! = 0,36775...$ benaderen. Dus is de bovenstaande kans min of meer onafhankelijk van N (als $N \geq 5$).
(b) Uit deel (a) weten we dat het aantal permutaties $(j_1, \ldots, j_m)$ van $(1, \ldots, m)$ zodanig dat $j_i \neq i, 1 \leq i \leq m$ gelijk is aan

$$m! \sum_{r=0}^m (-1)^r \frac{1}{r!} \ \ (m \geq 2)$$

en het is evident dat voor $1 \leq k < N$ geldt:

$$\begin{aligned} \mathbb{P}(k \text{ paren blijven samen}) &= \#\{A \subset \{1, \ldots, N\} : \#A = k\} \times \\ &\qquad \#\{\text{permutaties van } (1, \ldots, N-k) \\ &\qquad\qquad \text{zodat } j_i \neq i, 1 \leq i \leq N-k\}/N! \\ &= \binom{N}{k} \frac{1}{N!}(N-k)! \sum_{r=0}^{N-k} (-1)^r \frac{1}{r!} = \frac{1}{k!} \sum_{r=0}^{N-k} (-1)^r \frac{1}{r!}. \end{aligned}$$

Deze formule klopt ook als $k = N$. (In dit geval is de kans gelijk aan $1/N!$) Een benadering van de bovenstaande kans als k klein en N groot is, is $e^{-1}/k!$.

1.4 Voorwaardelijke kans

Om de definitie te motiveren bekijken we eerst een eenvoudig

Voorbeeld. Werp een dobbelsteen. Dan is de kans dat we een "2" verkrijgen gelijk aan 1/6. Veronderstel nu dat we al over de informatie beschikken dat het aantal ogen even is. In dit geval kunnen we de originele uitkomstenverzameling $\Omega = \{1, \ldots, 6\}$ vervangen door $\{2, 4, 6\}$ en de voorwaardelijke (of conditionele) kans op "2", gegeven dat het aantal ogen even is, is gelijk aan 1/3.

Om de juiste definitie te vinden, veronderstellen we weer dat men een gegeven toevalsexperiment kan herhalen zodat al de uitkomsten onafhankelijk van elkaar zijn. Om de kans op en gebeurtenis B onder de voorwaarde dat een andere gebeurtenis A is opgetreden, te berekenen, beschouwen we de gewijzigde relatieve frequenties $n_{A\cap B}/n_A$, waar weer n_C het aantal van de eerste n experimenten met een uitkomst in C is. Als $\mathbb{P}(A) > 0$ is, convergeren deze frequenties naar $\mathbb{P}(A\cap B)/\mathbb{P}(A)$ en het is dus voor de hand liggend dat we de voorwaardelijke kans op B gegeven A zo gaan definiëren.

Definitie 1.5 *Zij $(\Omega, \mathcal{F}, \mathbb{P})$ een kansruimte en $A \in \mathcal{F}$ een gebeurtenis met $\mathbb{P}(A) > 0$. Dan definiëren we voor elke gebeurtenis $B \in \mathcal{F}$ de* **voorwaardelijke kans** *op B, gegeven (dat de gebeurtenis) A (is opgetreden) door $\mathbb{P}(A \cap B)/\mathbb{P}(A)$. Notatie: $\mathbb{P}(B|A)$.*

De volgende stelling is soms handig om voorwaardelijke kansen te berekenen.

Stelling 1.2 *Zij $(\Omega, \mathcal{F}, \mathbb{P})$ een kansruimte, $A \in \mathcal{F}$ een gebeurtenis zodanig dat $\mathbb{P}(A) > 0$. Dan is $\mathbb{P}(\cdot|A) : \mathcal{F} \to [0, 1]$ een kansmaat.*

Bewijs. We noteren eerst dat $\mathbb{P}(\Omega|A) = \mathbb{P}(\Omega \cap A)/\mathbb{P}(A) = \mathbb{P}(A)/\mathbb{P}(A) = 1$. Verder is $B_n \cap A, n \geq 1$ een rij disjuncte gebeurtenissen als de gebeurtenissen $B_n, n \geq 1$ disjunct zijn. Dus geldt:

$$\mathbb{P}(\dot{\bigcup_{n=1}^{\infty}} B_n|A) = \mathbb{P}(\dot{\bigcup_{n=1}^{\infty}}(B_n \cap A))/\mathbb{P}(A) = \sum_{n=1}^{\infty} \mathbb{P}(B_n \cap A)/\mathbb{P}(A) = \sum_{n=1}^{\infty} \mathbb{P}(B_n|A)$$

en het is duidelijk dat $\mathbb{P}(\cdot|A)$ een kans is. □

Voorbeeld. Veronderstel dat er even veel meisjes als jongens geboren worden. Wat is dan de voorwaardelijke kans dat in een gezin met twee kinderen er zowel een jongen als een meisje is, als
(a) het oudste kind een jongen is,
(b) er tenminste één van de twee een jongen is?
Oplossing. Stel $\Omega = \{(J, J), (J, M), (M, J), (M, M)\}$ waar bv (M, J) betekent dat het oudste kind een meisje is en het jongste een jongen. Stel $B = \{(J, M), (M, J)\}$ de gebeurtenis dat er zowel een jongen als een meisje is.
In (a) hebben we de informatie $A = \{(J, J), (J, M)\}$ en bijgevolg: $\mathbb{P}(B|A) = 1/2$.
In (b) stellen we $A = \{(J, M), (M, J), (J, J)\}$ en dus geldt $\mathbb{P}(B|A) = 2/3$.

In het vorige voorbeeld hebben we de voorwaardelijke kans via de "gewone" kansen berekend. Er zijn echter meer toepassingen waar men voorwaardelijke kansen gebruikt om "gewone" kansen te berekenen. De twee volgende stellingen zijn daarvoor heel handig.

Stelling 1.3 (Kettingregel) *Zij $(\Omega, \mathcal{F}, \mathbb{P})$ een kansruimte en $A_1, \ldots, A_n \in \mathcal{F}$ $(n \geq 2)$ gebeurtenissen zodanig dat $\mathbb{P}(A_1 \cap \ldots \cap A_{n-1}) > 0$. Dan geldt:*

$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1) \cdot \ldots \cdot \mathbb{P}(A_n|A_1 \cap \ldots \cap A_{n-1}).$$

Bewijs. via inductie.

Voorbeeld. In een doos zitten 10 ballen, waarvan één blauw is. De andere ballen zijn rood. We vragen een groep van 10 personen dat iedereen en bal trekt (zonder terugleggen). De persoon die de blauwe bal trekt wint een prijs. Wat is de beste strategie? Is het een goed idee als eerste te trekken, of is het beter een beetje te wachten?

Oplossing. We bewijzen via inductie dat iedereen de gelijke kans heeft de blauwe bal te trekken. (Dus ook de laatste persoon.)
Stel $A_i = \{\text{persoon } i \text{ trekt de blauwe bal}\}$, $1 \leq i \leq 10$. Dan is het evident dat $\mathbb{P}(A_1) = 1/10$.
Veronderstel nu dat $\mathbb{P}(A_1) = \ldots = \mathbb{P}(A_{m-1}) = 1/10$, waar $2 \leq m < n$.
Dan volgt wegens $A_m \subset A_j^c, j \neq m, \mathbb{P}(A_m) = \mathbb{P}(A_m \cap \bigcap_{j=1}^{m-1} A_j^c)$ en dus

$$\mathbb{P}(A_m) = \mathbb{P}(\bigcap_{j=1}^{m-1} A_j^c)\mathbb{P}(A_m|\bigcap_{j=1}^{m-1} A_j^c) = (1 - \mathbb{P}(\bigcup_{j=1}^{m-1} A_j))\frac{1}{10 - m + 1}.$$

Vermits de gebeurtenissen $A_1, \ldots, A_{m-1}$ disjunct zijn, volgt dat $\mathbb{P}(\bigcup_{j=1}^{m-1} A_j) = (m-1)/10$ en we zien dat $\mathbb{P}(A_m) = 1/10$.

Stelling 1.4 (Wet van de totale kans) *Zij $(\Omega, \mathcal{F}, \mathbb{P})$ een kansruimte en zij $A_i \in \mathcal{F}, i \in I$ een partitie van Ω. Veronderstel dat $\mathbb{P}(A_i) > 0, i \in I$. Dan is I ten hoogste aftelbaar en voor elke gebeurtenis $B \in \mathcal{F}$ geldt:*

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

Bewijs. We tonen eerst aan dat I ten hoogste aftelbaar is.
Stel $I_m := \{i \in I : \mathbb{P}(A_i) \geq 1/m\}, m \geq 1$. Dan geldt $\#I_m \leq m$ omdat we anders indices $i_1, \ldots, i_{m+1}$ konden vinden zodat $\mathbb{P}(A_{i_j}) \geq 1/m, 1 \leq j \leq m+1$. Maar dan was $\mathbb{P}(\Omega) \geq \mathbb{P}(\dot{\bigcup}_{j=1}^{m+1} A_{i_j}) = \sum_{j=1}^{m+1} \mathbb{P}(A_{i_j}) \geq 1 + 1/m$, wat natuurlijk niet kan. Dus is $I = \bigcup_{m=1}^{\infty} I_m$ als een aftelbare unie van eindige verzamelingen ten hoogste aftelbaar.
Het is evident dat

$$B = \dot{\bigcup_{i \in I}}(B \cap A_i),$$

waar de gebeurtenissen $B \cap A_i, i \in I$ disjunct zijn. Gezien I ten hoogste aftelbaar is, volgt:

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B \cap A_i) = \sum_{i \in I} \mathbb{P}(B|A_i)\mathbb{P}(A_i)$$

en de stelling is bewezen. □

Voorbeeld. Ann werpt 3 muntstukken en Jan werpt 2 muntstukken. Ann krijgt 10 Euro van Jan als haar aantal "kop"groter is dan het aantal "kop"van Jan. Anders wint Jan 10 Euro. Is dit een goed spel voor Ann?

Oplossing. Zij B de gebeurtenis dat Ann wint en stel voor $i = 0, 1, 2$,

$$A_i = \{\text{Jan heeft } i\text{-keer kop geworpen}\}$$

Dan geldt:

$$\mathbb{P}(B) = \mathbb{P}(B|A_0)\mathbb{P}(A_0) + \mathbb{P}(B|A_1)\mathbb{P}(A_1) + \mathbb{P}(B|A_2)\mathbb{P}(A_2),$$

waar $\mathbb{P}(A_0) = \mathbb{P}(A_2) = 1/4, \mathbb{P}(A_1) = 1/2$. Zij verder K_i de gebeurtenis dat Ann i-keer kop heeft geworpen, $0 \le i \le 2$. Dan geldt $\mathbb{P}(K_0) = \mathbb{P}(K_3) = 1/8, \mathbb{P}(K_1) = \mathbb{P}(K_2) = 3/8$ en dus $\mathbb{P}(B|A_0) = \mathbb{P}(K_1) + \mathbb{P}(K_2) + \mathbb{P}(K_3) = 7/8$. Analoog volgt dat $\mathbb{P}(B|A_1) = 1/2$ en $\mathbb{P}(B|A_2) = 1/8$ en we zien dat $\mathbb{P}(B) = 1/2$. Dus is het een fair spel.

De volgende stelling is een onmiddellijk gevolg van de wet van de totale kans.

Stelling 1.5 (De regel van Bayes)
Zij $(\Omega, \mathcal{F}, \mathbb{P})$ een kansruimte en zij $A_i \in \mathcal{F}, i \in I$ een partitie van Ω, waar $\mathbb{P}(A_i) > 0, i \in I$. Dan geldt voor $i_0 \in I$:

$$\mathbb{P}(A_{i_0}|B) = \frac{\mathbb{P}(B|A_{i_0})\mathbb{P}(A_{i_0})}{\sum_{i\in I}\mathbb{P}(B|A_i)\mathbb{P}(A_i)}$$

Voorbeeld (Betrouwbare medische testen) Een medisch test voor een bepaalde ziekte is positief in 99% van de gevallen waar de patiënt deze ziekte heeft. Maar hij geeft ook een vals "positief" resultaat in 1% van de gevallen waar de patiënt gezond is. Veronderstel dat 0,5% van de bevolking aan deze ziekte lijdt. Wat is de kans dat een patiënt ziek is als de test positief is?

Oplossing. Zij B de gebeurtenis dat de test positief is en A de gebeurtenis dat de patiënt ziek is. Dan geldt: $\mathbb{P}(B|A) = 0,99, \mathbb{P}(B|A^c) = 0,01, \mathbb{P}(A) = 0,005$ wat impliceert dat

$$\mathbb{P}(A|B) = \frac{(0,99)\cdot(0,005)}{(0,99)\cdot(0,005) + (0,01)\cdot(0,995)} = \frac{495}{495+995} = \frac{99}{298} < \frac{1}{3}.$$

Dus is in dit geval de kans nog relatief klein dat de patiënt aan deze ziekte lijdt, hoewel het natuurlijk een dramatisch verschil in vergelijking met de originele kans van 0,005 is.

1.5 Onafhankelijke gebeurtenissen

Zij $(\Omega, \mathcal{F}, \mathbb{P})$ een kansruimte en A, B gebeurtenissen met $\mathbb{P}(A), \mathbb{P}(B) > 0$. We zeggen dat A onafhankelijk van B is als de voorwaardelijke kans op A, gegeven B gelijk aan de originele kans op A is. Dus, de informatie dat B gebeurd is vertelt ons niets over de kans op A die wij daarom niet kunnen wijzigen. Als we de definitie van "voorwaardelijke kans" gebruiken, zien we onmiddellijk dat $\mathbb{P}(A|B) = \mathbb{P}(A)$ impliceert dat $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Maar dit betekent ook dat $\mathbb{P}(B|A) = \mathbb{P}(B)$. Dus is "$A$ onafhankelijk van B" hetzelfde als "B onafhankelijk van A" en we zeggen gewoon dat de twee gebeurtenissen A en B onafhankelijk zijn. Formeel definiëren we ook voor gebeurtenissen waar de kans gelijk aan 0 mag zijn:

Definitie 1.6
Twee gebeurtenissen $A, B \in \mathcal{F}$ heten **onafhankelijk** *indien* $\boxed{\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)}$.

Opmerkingen

1. Als A een gebeurtenis met $\mathbb{P}(A) \in \{0, 1\}$ en $B \in \mathcal{F}$ een willekeurige gebeurtenis is, dan zijn A en B onafhankelijk.

2. Zij $A, B \in \mathcal{F}$ zodanig dat $A \subset B$ en $0 < \mathbb{P}(A) \leq \mathbb{P}(B) < 1$. Dan zijn A en B afhankelijk (= niet onafhankelijk). Dit is evident omdat in dit geval $\mathbb{P}(B|A) = 1$.

3. Zij $A, B \in \mathcal{F}$ disjuncte niet-triviale gebeurtenissen (i.e met kans $\in]0, 1[$). Dan zijn A, B weer afhankelijk. (Als we weten dat A gebeurd is, weten we zeker dat B niet gebeurd is. Dus kunnen deze twee gebeurtenissen niet onafhankelijk zijn.)

4. We concluderen dat als A, B niet-triviale, onafhankelijke gebeurtenissen zijn, ze niet disjunct mogen zijn en geen van de twee mag de andere omvatten.

Voorbeeld. Werp twee keer een dobbelsteen en beschouw de gebeurtenissen

$$A = \{\text{som van de ogen} = 7\}, B = \{\text{som van de ogen} = 6\}, C = \{\text{een drie de eerste keer}\}.$$

Als we veronderstellen dat al de uitkomsten $(i, j) \in \{1, \ldots, 6\}^2$ even waarschijnlijk zijn, volgt dat $\mathbb{P}(A) = 1/6$, $\mathbb{P}(B) = 5/36$, $\mathbb{P}(C) = 1/6$. Gezien $\mathbb{P}(A \cap C) = \mathbb{P}\{(3, 4)\} = 1/36$ volgt dat A en C onafhankelijk zijn. Maar B en C zijn afhankelijk.

Stelling 1.6 *Zij A, B onafhankelijke gebeurtenissen. Dan geldt:*
(i) *A en B^c zijn onafhankelijk.*
(ii) *A^c en B zijn onafhankelijk.*
(iii) *A^c en B^c zijn onafhankelijk.*

Bewijs. Het is voldoende (i) te bewijzen. ((ii) is equivalent met (i), en (iii) volgt na twee opeenvolgende toepassingen van (i).)
Gezien A de unie van de twee disjuncte gebeurtenissen $A \cap B$ en $A \cap B^c$ is, volgt onmiddellijk:

$$\mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B),$$

hetgeen wegens de onafhankelijkheid van A en B gelijk is aan

$$\mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)(1 - \mathbb{P}(B)) = \mathbb{P}(A)\mathbb{P}(B^c),$$

en de stelling is bewezen. □
We bekijken nu gebeurtenissen $A_i, i \in I$, waar de indexverzameling I niet noodzakelijk aftelbaar is.

Definitie 1.7 *Zij $A_i \in \mathcal{F}, i \in I$ gebeurtenissen, waar I een verzameling is.*
(i) *$A_i, i \in I$ heet* **onafhankelijk** *indien voor elke eindige deelverzameling $\{i_1, \ldots, i_r\} \subset I, r \geq 2$ geldt:*

$$\mathbb{P}(\bigcap_{j=1}^{r} A_{i_j}) = \prod_{j=1}^{r} \mathbb{P}(A_{i_j}).$$

(ii) *$A_i, i \in I$ heet* **paarsgewijs onafhankelijk** *indien*

$$\mathbb{P}(A_{i_1} \cap A_{i_2}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}), i_1 \neq i_2, i_1, i_2 \in I.$$

Opmerking Het is triviaal dat geldt "(i) ⇒ (ii)". Het volgende elementaire voorbeeld toont aan dat de omgekeerde implicatie in het algemeen niet klopt.

Voorbeeld Werp twee keer een dobbelsteen en bekijk de gebeurtenissen

$$A = \{\text{som van de ogen} = 7\}, B = \{\text{een drie de eerste keer}\}, C = \{\text{een vier de tweede keer}\}$$

We hebben al eerder bewezen dat A en B onafhankelijk zijn. Hetzelfde argument toont aan dat A en C onafhankelijk zijn. Het is ook evident dat B en C onafhankelijk zijn. Dus zijn deze gebeurtenissen paarsgewijs onafhankelijk. Ze zijn **niet** onafhankelijk omdat $A \supset B \cap C$ wat impliceert dat A en $B \cap C$ afhankelijk zijn. Maar als drie gebeurtenissen A, B, C onafhankelijk zijn, dan moeten ook de twee gebeurtenissen $A, B \cap C$ onafhankelijk zijn.

Stelling 1.7 *Zij $A_i, i \in I$ een klasse onafhankelijke gebeurtenissen en $I_j, j \in J$ disjuncte deelverzamelingen van I. Als we voor elke j een gebeurtenis B_j uit de klasse $\mathcal{B}_j$ - de gebeurtenissen die uit $A_i, i \in I_j$ verkrijgbaar zijn (waar men ten hoogste aftelbaar veel operaties zoals $\cap$, $\cup$, $\triangle$,^{c} gebruikt)- kiest, dan zijn de gebeurtenissen $B_j, j \in J$ onafhankelijk.*

Bewijs via onafhankelijkheid van σ-algebra's (zie maattheorie, bachelor 3). In 2.1 zullen we reeds de precieze wiskundige definitie van $\mathcal{B}_j$ geven.

Natuurlijk impliceert stelling 1.7 stelling 1.6.

Voorbeelden

1. Als $A_n, n \geq 1$ onafhankelijke gebeurtenissen zijn, dan zijn ook de gebeurtenissen A_1, $A_2 \cap A_5$, $A_3 \triangle A_{10}$, $\cup_{n=5}^{\infty} A_{2n+1}$ onafhankelijk.

2. A, B, C onafhankelijk $\Rightarrow A, B \cup C$ onafhankelijk. (*Oefening*: Geef een direct bewijs.)

Definitie 1.8 *Een experiment met twee mogelijke uitkomsten waarvan men één "succes"(en de andere "mislukking") noemt, heet een Bernoulli(p)-experiment als de kans op succes p is.*

Probleem Veronderstel dat men een bepaald Bernoulli(p)-experiment n keer kan uitvoeren zodat de gebeurtenissen $A_i = \{\text{i-de experiment succes}\}, 1 \leq i \leq n$ onafhankelijk zijn. Wat is dan de kans dat er precies k successen zijn ($0 \leq k \leq n$)?
Oplossing Zij B_k de gebeurtenis dat er precies k successen zijn. Dan is het evident dat

$$\mathbb{P}(B_0) = \mathbb{P}(\bigcap_{i=1}^{n} A_i^c) = \prod_{i=1}^{n} \mathbb{P}(A_i^c) = (1-p)^n.$$

Analoog volgt dat

$$\mathbb{P}(B_n) = p^n.$$

Als $1 \leq k \leq n-1$, dan geldt

$$B_k = \bigcupdot_{\substack{K \subset \{1,\ldots,n\} \\ \#K = k}} (\bigcap_{i \in K} A_i \cap \bigcap_{i \in K^c} A_i^c).$$

Uit stelling 1.7 volgt dat de n gebeurtenissen $A_i, i \in K, A_i^c, i \notin K$ onafhankelijk zijn en dus geldt

$$\mathbb{P}(\bigcap_{i \in K} A_i \cap \bigcap_{i \in K^c} A_i^c) = p^k (1-p)^{n-k} \text{ als } \#K = k.$$

Vermits de gebeurtenissen $\bigcap_{i\in K} A_i \cap \bigcap_{i\in K^c} A_i^c,\ K \subset \{1,\dots,n\}$ disjunct zijn, volgt dat

$$\mathbb{P}(B_k) = \#\{K \subset \{1,\dots,n\} : \#K = k\} p^k (1-p)^{n-k} = \begin{pmatrix} n \\ k \end{pmatrix} p^k (1-p)^{n-k}, 1 \le k \le n,$$

wat ook juist is als $k \in \{0, n\}$.

Voorbeeld Werp een dobbelsteen 5 keer. Wat is de kans op (a) minstens 2 keer "5", (b) precies 3 keer een even getal?

Oplossing (a) In dit geval gaat het om een Bernoulli(1/6)-experiment. Dus is de kans in (a) gelijk aan

$$1 - \mathbb{P}\{\text{geen } 5\} - \mathbb{P}\{\text{één } 5\} = 1 - (5/6)^5 - 5(1/6)(5/6)^4 = 0,1962..$$

(b) Nu is de kans op succes gelijk aan 1/2. Dus: $\mathbb{P}\{3 \text{ successen}\} = \begin{pmatrix} 5 \\ 3 \end{pmatrix} \frac{1}{32} = 0,3125.$

We bewijzen nog een "omkering" van het Borel-Cantelli lemma voor onafhankelijke gebeurtenissen.

Lemma 1.2 (Borel-Cantelli lemma, deel 2) *Zij $A_n, n \ge 1$ een rij onafhankelijke gebeurtenissen. Veronderstel dat $\sum_{n=1}^\infty \mathbb{P}(A_n) = \infty$. Dan is $\mathbb{P}(\limsup_{n\ge1} A_n) = 1$.*

Bewijs Zij $1 \le n \le N$. Dan kunnen we uit de onafhankelijkheid van de gebeurtenissen $A_k, n \le k \le N$ en de triviale ongelijkheid $1 - t \le \exp(-t), t \ge 0$ concluderen:

$$\mathbb{P}(\bigcup_{k=n}^N A_k) = 1 - \mathbb{P}(\bigcap_{k=n}^N A_k^c) = 1 - \prod_{k=n}^N (1 - \mathbb{P}(A_k)) \ge 1 - \exp(-\sum_{k=n}^N \mathbb{P}(A_k)).$$

Gezien $\sum_{n=1}^\infty \mathbb{P}(A_n) = \infty$, volgt voor elke $n \ge 1$: $\exp(-\sum_{k=n}^N \mathbb{P}(A_n)) \to 0$ als $N \to \infty$.
Dus als we B_n gelijk aan $\bigcup_{k=n}^\infty A_k$ stellen, geldt er: $\mathbb{P}(B_n) = \lim_{N\to\infty} \mathbb{P}(\bigcup_{k=n}^N A_k) = 1\ \forall n \ge 1$. Aangezien $B_n \searrow \limsup_n A_n$ en $\mathbb{P}$ continu van boven is, zien we dat $\mathbb{P}(\limsup_n A_n) = \lim_{n\to\infty} \mathbb{P}(B_n) = 1$. □

Opmerking Omdat er maar twee mogelijkheden zijn, de reeks $\sum_{n=1}^\infty \mathbb{P}(A_n)$ is convergent of divergent, kan de kans $\mathbb{P}(\limsup_n A_n)$ alleen 0 of 1 zijn als de gebeurtenissen A_n onafhankelijk zijn. Dit is een speciaal geval van de 0-1 wet van Kolmogorov (→ maattheorie). Bekijken we de complementen van A_n zien we dat dan ook $\mathbb{P}(\liminf_{n\ge1} A_n) \in \{0,1\}$. Dus hebben we in dit geval ook een 0-1-wet voor $\liminf_{n\ge1} A_n$.

Hoofdstuk 2

Toevalsvariabelen en verdelingen

2.1 Meetbare afbeeldingen

Stelling 2.1 *Zij Ω een niet-lege verzameling en $\mathcal{A}$ een klasse van delen van Ω. Er bestaat een kleinste σ-algebra op Ω die $\mathcal{A}$ bevat. We noemen deze de σ-algebra voortgebracht door de klasse $\mathcal{A}$. Notatie: $\sigma(\mathcal{A})$.*

Bewijs Zij $K_{\mathcal{A}}$ de klasse van al de σ-algebra's op Ω die $\mathcal{A}$ bevatten. (Deze is niet leeg omdat geldt $2^\Omega \in K_{\mathcal{A}}$.) Stel $\mathcal{F}_0 = \bigcap_{\mathcal{G} \in K_{\mathcal{A}}} \mathcal{G}$. Het is evident dat $\mathcal{F}_0$ als een doorsnede van σ-algebra's op Ω weer een σ-algebra op Ω is. Verder volgt uit de definitie van $K_{\mathcal{A}}$ dat $\mathcal{F}_0$ de klasse $\mathcal{A}$ bevat. Gezien elke σ-algebra $\mathcal{G}$ met deze eigenschap in $K_{\mathcal{A}}$ zit, volgt dat $\mathcal{F}_0 \subset \mathcal{G}$ en dus is $\mathcal{F}_0$ de kleinste σ-algebra die $\mathcal{A}$ bevat. □

Opmerking. We kunnen nu de klassen $\mathcal{B}_j$ in stelling 1.7 formeel definiëren als

$$\mathcal{B}_j = \sigma(\{A_i : i \in I_j\}), j \in J.$$

Definitie 2.1 *De k-dimensionale Borel-σ-algebra is de σ-algebra op $\mathbb{R}^k$ voortgebracht door de open verzamelingen van $\mathbb{R}^k$. Notatie: $\mathcal{R}^k$ en $\mathcal{R} = \mathcal{R}^1$.*

We noemen een koppel $(\Omega, \mathcal{F})$, waar Ω een niet-lege verzameling en $\mathcal{F}$ een σ-algebra op Ω is, een **meetbare ruimte**.

Definitie 2.2 *Veronderstel dat $(\Omega, \mathcal{F}), (\Omega', \mathcal{F}')$ meetbare ruimten zijn.*

(i) *Een afbeelding $T : \Omega \to \Omega'$ heet $\mathcal{F}, \mathcal{F}'$-meetbaar indien*

$$T^{-1}A' := \{\omega \in \Omega : T(\omega) \in A'\} \in \mathcal{F} \quad \forall A' \in \mathcal{F}'$$

(ii) *Een functie $f : \Omega \to \mathbb{R}$ (een afbeelding $f = (f_1, \dots, f_k) : \Omega \to \mathbb{R}^k$) heet $\mathcal{F}$-meetbaar indien deze $\mathcal{F}, \mathcal{R}$-meetbaar ($\mathcal{F}, \mathcal{R}^k$-meetbaar) is.*

Stelling 2.2

(i) *Zij $\mathcal{F}' = \sigma(\mathcal{A}')$. Dan geldt: $T^{-1}A' \in \mathcal{F} \;\; \forall A' \in \mathcal{A}' \Rightarrow T$ is $\mathcal{F}, \mathcal{F}'$-meetbaar.*

(ii) *Zij $T_1 : \Omega_1 \to \Omega_2$ $\mathcal{F}_1, \mathcal{F}_2$-meetbaar en $T_2 : \Omega_2 \to \Omega_3$ $\mathcal{F}_2, \mathcal{F}_3$-meetbaar. Dan geldt: $T_2 \circ T_1 : \Omega_1 \to \Omega_3$ is $\mathcal{F}_1, \mathcal{F}_3$-meetbaar.*

Bewijs (i) Stel $\mathcal{S}' := \{A' \in \mathcal{F}' : T^{-1}A' \in \mathcal{F}\}$. Dan is $\mathcal{S}'$ een σ-algebra (oefening) en $\mathcal{S}' \supset \mathcal{A}'$, hetgeen impliceert dat $\mathcal{S}' \supset \mathcal{F}'$. Dus is T $\mathcal{F},\mathcal{F}'$-meetbaar.
(ii) $(T_2 \circ T_1)^{-1}(A_3) = T_1^{-1}(\underbrace{T_2^{-1}(A_3)}_{\in \mathcal{F}_2}) \in \mathcal{F}_1,\ \forall A_3 \in \mathcal{F}_3$. □

Stelling 2.3

(i) $f : \mathbb{R}^{k_1} \to \mathbb{R}^{k_2}$ *continu* $\Rightarrow f$ $\mathcal{R}^{k_1}$*-meetbaar.*

(ii) $f = (f_1, \dots, f_k) : \Omega \to \mathbb{R}^k$ $\mathcal{F}$*-meetbaar* $\Longleftrightarrow f_i : \Omega \to \mathbb{R}$ $\mathcal{F}$*-meetbaar,* $1 \le i \le k$.

(iii) $f_i : \Omega \to \mathbb{R}$ $\mathcal{F}$*-meetbaar,* $1 \le i \le k$, $g : \mathbb{R}^k \to \mathbb{R}$ *continu* $\Rightarrow \omega \to g(f_1(\omega), \dots, f_k(\omega))$ *is* $\mathcal{F}$*-meetbaar.*

Bewijs (i) Gezien $\mathcal{R}^{k_2}$ de σ-algebra voortgebracht door de open verzamelingen in $\mathbb{R}^{k_2}$ is, is het voldoende aan te tonen dat

$$f^{-1}G \in \mathcal{R}^{k_1}\ \forall G \subset \mathbb{R}^{k_2} \text{ open.}$$

Omdat f continu is, is $f^{-1}G$ open in $\mathbb{R}^{k_1}$ en dus een k_1-dimensionale Borelverzameling is.
(ii) $\boxed{\Rightarrow}$ We hebben $f_i = \pi_i \circ f$, waar de projectie $\pi_i(x_1, \dots, x_k) = x_i$ continu en bijgevolg $\mathcal{R}^k$-meetbaar is. Dus is f_i als een compositie van meetbare afbeeldingen zelf meetbaar. $(i = 1, \dots, k)$
$\boxed{\Leftarrow}$ We kunnen elke open verzameling in $\mathbb{R}^k$ opschrijven als een aftelbare unie van verzamelingen uit de klasse

$$\mathcal{A} = \{\prod_{i=1}^{k}]a_i, b_i[: -\infty < a_i \le b_i < \infty, 1 \le i \le k\},$$

wat natuurlijk impliceert dat $\mathcal{R}^k = \sigma(\mathcal{A})$. Gezien

$$f^{-1}(\prod_{i=1}^{k}]a_i, b_i[) = \bigcap_{i=1}^{k} \underbrace{f_i^{-1}(]a_i, b_i[)}_{\in \mathcal{F}} \in \mathcal{F}$$

volgt uit stelling 2.2.i dat f $\mathcal{F}$-meetbaar is.
(iii) Volgt onmiddellijk uit (i), (ii) en deel (ii) van stelling 2.2. □

Voorbeeld Neem twee $\mathcal{F}$-meetbare functies $f_1, f_2 : \Omega \to \mathbb{R}$ $\mathcal{F}$. Dan zijn functies zoals $f_1^2, |f_1|, f_1 + f_2, f_1 \cdot f_2, \dots$ weer $\mathcal{F}$-meetbaar.

2.2 Toevalsvariabelen

Als $(\Omega, \mathcal{F}, \mathbb{P})$ een kansruimte is, noemen we de $\mathcal{F}$-meetbare functies toevalsvariabelen en duiden ze meestal door hoofdletters zoals $X, Y, Z, \dots$ aan. We noteren de gebeurtenissen $X^{-1}A$ met $\{X \in A\}, A \in \mathcal{R}$.

Stelling 2.4 *Zij* $X : \Omega \to \mathbb{R}$ *een toevalsvariabele. Dan is* $\mu(A) := \mathbb{P}\{X \in A\}, A \in \mathcal{R}$ *een kansmaat. We noemen deze de verdeling van X. Notatie:* $\mathbb{P}_X$

Bewijs. Gezien X $\mathcal{F}$-meetbaar is, is $\mu(A)$ altijd gedefinieerd voor $A \in \mathcal{R}$. Verder is het evident dat $\{X \in \mathbb{R}\} = \Omega$ en bijgevolg $\mu(\mathbb{R}) = 1$. Als $A_n, n \geq 1$ disjuncte Borel-verzamelingen zijn, dan zijn de verzamelingen $\{X \in A_n\} = X^{-1}(A_n), n \geq 1$ disjuncte gebeurtenissen in $\mathcal{F}$ en het volgt dat

$$\mu(\dot{\bigcup_{n=1}^{\infty}} A_n) = \mathbb{P}(X^{-1}(\dot{\bigcup_{n=1}^{\infty}} A_n)) = \mathbb{P}(\dot{\bigcup_{n=1}^{\infty}} X^{-1}(A_n)) = \sum_{n=1}^{\infty} \mathbb{P}(X \in A_n) = \sum_{n=1}^{\infty} \mu(A_n).$$

Dus is μ ook σ-additief en bijgevolg een kansmaat. □

Definitie 2.3 *Zij $X : \Omega \to \mathbb{R}$ een toevalsvariabele. Dan heet de functie*

$$F(x) := \mathbb{P}\{X \leq x\} = \mathbb{P}_X(]-\infty, x]), x \in \mathbb{R}$$

de **verdelingsfunctie** *van X.*

Stelling 2.5 *Zij $F(x), x \in \mathbb{R}$ de verdelingsfunctie van een toevalsvariabele $X : \Omega \to \mathbb{R}$. Dan geldt:*

(i) *$x \to F(x)$ is monotoon niet-dalend*

(ii) *F is rechts continu*

(iii) $\lim_{x\to\infty} F(x) = 1, \lim_{x\to-\infty} F(x) = 0.$

(iv) *F is continu in $a \Leftrightarrow \mathbb{P}\{X = a\} = 0$.*

Bewijs (i) $x_1 \leq x_2 \Rightarrow \{X \leq x_1\} \subset \{X \leq x_2\}$. Gezien $\mathbb{P}$ monotoon is (eigenschap (4) van kansmaten in 1.2) volgt dat $F(x_1) = \mathbb{P}\{X \leq x_1\} \leq \mathbb{P}\{X \leq x_2\} = F(x_2)$.
(ii) Het is voldoende te bewijzen dat geldt: $x_n \searrow x \Rightarrow F(x_n) \to F(x)$. In dit geval is $\{X \leq x_n\}$ een dalende rij van gebeurtenissen met limiet gelijk aan $\{X \leq x\}$. Gezien $\mathbb{P}$ continu van boven is volgt dat $F(x_n) = \mathbb{P}\{X \leq x_n\} \to \mathbb{P}\{X \leq x\} = F(x)$.
(iii) Analoog bewijs.
(iv) We noteren eerst dat $x_n < a, x_n \nearrow a$ impliceert dat $F(x_n) \to \mathbb{P}\{X < a\}$. (Dit volgt omdat nu $\{X \leq x_n\}$ een stijgende rij van gebeurtenissen is met limiet $\{X < a\}$.) Dus geldt: $F(a-) = F(a) - \mathbb{P}\{X = a\}$. Gezien F rechts continu is en dus $F(a+) = F(a)$ is het nu evident dat F continu in a is als en slechts als $\mathbb{P}\{X = a\} = 0$. □

Opmerkingen

1. Deel (iv) van stelling 2.5 impliceert dat het aantal discontinuïteitspunten van F ten hoogste aftelbaar is (zie bewijs van stelling 1.4).

2. Uit een van de hoofdstellingen van de maattheorie (de uitbreidingsstelling, zie maattheorie, bachelor 3) volgt dat voor elke functie F die aan de drie voorwaarden (i) - (iii) voldoet, er een kansmaat μ op $\mathcal{R}$ bestaat zodat geldt: $F(x) = \mu(]-\infty, x]), x \in \mathbb{R}$.

3. Bovendien kan men bewijzen dat deze kansmaat **uniek** is. Dus als we weten dat twee toevalsvariabelen dezelfde verdelingsfunctie hebben, kunnen we concluderen dat hun verdelingen overeenstemmen.

2.3 Discrete toevalsvariabelen

Toevalsvariabelen die ten hoogste aftelbaar veel waarden kunnen aannemen noemen we **discrete** toevalsvariabelen. Toevalsvariabelen waar maar eindig veel waarden mogelijk zijn heten **elementaire** toevalsvariabelen.

Algemeen is het voldoende voor discrete toevalsvariabelen de **kansfunctie** te bepalen, i.e. $p(x) = \mathbb{P}\{X = x\}, x \in \mathbb{R}$. Per definitie geldt dan $\mathbb{P}\{X \in S\} = 1$, waar $S := \{x : p(x) > 0\}$ ten hoogste aftelbaar is en we kunnen kansen $\mathbb{P}\{X \in A\}$ onmiddellijk via de formule

$$\mathbb{P}\{X \in A\} = \sum_{x \in S \cap A} p(x)$$

berekenen. (De verdelingsfunctie is in dit geval een trapfunctie met discontinuïteitspunten in S en het geldt $F(x) - F(x-) = p(x), x \in S$.)

Voorbeelden

1. Zij $A \in \mathcal{F}$ een gebeurtenis met $\mathbb{P}(A) = p$. Dan heeft de toevalsvariabele(!) I_A (ga na dat deze $\mathcal{F}$-meetbaar is) een **Bernoulli(p)-verdeling**. In dit geval geldt $S = \{0, 1\}$ en $p(0) = (1 - p), p(1) = p$.

2. Zij $A_1, \ldots, A_n$ onafhankelijke gebeurtenissen met $\mathbb{P}(A_i) = p, 1 \leq i \leq n$. Stel $X = \sum_{i=1}^n I_{A_i}$. Vermits $I_{A_i}, 1 \leq i \leq n$ toevalsvariabelen zijn, volgt uit stelling 2.3.iii dat ook X een toevalsvariabele is. Het is evident dat $S = \{0, \ldots, n\}$. Gezien we A_i als de gebeurtenis "succes" bij het i-de van n onafhankelijke Bernoulli(p)-experimenten kunnen beschouwen, volgt zoals in hoofdstuk 1.5 dat voor de kansfunctie van deze toevalsvariabele geldt:
$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}, 0 \leq k \leq n.$$
De verdeling van deze toevalsvariabele heet de **binomiaal(n,p)-verdeling**.
(Natuurlijk is de Bernoulli(p)-verdeling het speciale geval daarvan, waar $n = 1$.)

3. Werp een muntstuk waar de kans op "kop" gelijk aan p is, totdat je de eerste keer kop krijgt. Stel $X = \#$ worpen. Dan zijn natuurlijk de mogelijke waarden: $1, 2, 3, \ldots$ Als we onderstellen dat de worpen onafhankelijk van elkaar zijn, volgt dat
$$p(1) = \mathbb{P}\{X = 1\} = \mathbb{P}\{\text{de eerste worp: kop}\} = p$$
en verder geldt er voor $k \geq 2$:
$$p(k) = \mathbb{P}\{X = k\} = \mathbb{P}\{(k-1) \text{ keer munt en dan kop}\} = (1-p)^{k-1} p.$$
Deze verdeling heet de **geometrische** verdeling met parameter p.

4. Als we in het vorige voorbeeld het muntstuk werpen totdat we de r-de keer "kop" hebben gekregen, heeft de toevalsvariabele $X = \#$ worpen een **negatief-binomiaal(r,p)-verdeling**. We hebben in dit geval: $S = \{r, r+1, \ldots\}$ en de kansfunctie is gegeven door
$$p(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, k = r, r+1, \ldots$$
(Als $r = 1$ verkrijgen we weer de geometrisch(p)-verdeling.)

5. In een vaas zitten M blauwe ballen en N gele ballen. Wij trekken n ballen (a) met terugleggen, (b) zonder terugleggen ($n \le N + M$).
   Stel X = # getrokken blauwe ballen. In geval (a) heeft X een binomiaal(n,M/(M+N))-verdeling. In geval (b) geldt:

$$\mathbb{P}\{X = k\} = \frac{\binom{M}{k}\binom{N}{n-k}}{\binom{M+N}{n}}, \max(0, n-N) \le k \le \min(n, M).$$

   We noemen deze verdeling de **hypergeometrische** verdeling.
   Als $\min(N, M)$ groot is en n in vergelijking relatief klein, kunnen we echter de kansfunctie van de hypergeometrische verdeling door de binomiale kansfunctie benaderen.

6. Zij voor elke $n \ge 1$ X_n binomiaal(n, p_n)-verdeeld en veronderstel dat $np_n \to \lambda \in]0, \infty[$ als $n \to \infty$. Dan geldt voor $k \ge 0$ en $n \ge k$:

$$\mathbb{P}\{X_n = k\} = \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{1}{k!} \prod_{j=1}^{k} (1 - \frac{j-1}{n})(np_n)^k (1-p_n)^n (1-p_n)^{-k}.$$

   Als $n \to \infty$ (en k vast blijft) volgt dat $\prod_{j=1}^k (1 - \frac{j-1}{n}) \to 1$, $(np_n)^k \to \lambda^k$, $(1 - \frac{np_n}{n})^n \to e^{-\lambda}$ (omdat $np_n \to \lambda$) en natuurlijk $(1-p_n)^{-k} \to 1$ (wegens $p_n \to 0$). Dus geldt:

$$\mathbb{P}\{X_n = k\} \to \frac{1}{k!} \lambda^k e^{-\lambda} =: p(k), k = 0, 1, \ldots$$

   Wegens $\sum_{k \ge 0} p(k) = 1$ volgt dat $p(k), k \ge 0$ een kansfunctie is en we noemen de daardoor bepaalde verdeling de **Poisson-verdeling** met parameter λ. Als we nog een keer het ontmoetingsprobleem (zie hoofdstuk 1.3) bekijken zien we dat de toevalsvariabele X = het aantal paren die samen blijven, in benadering een Poisson(1)-toevalsvariabele is. Maar we kunnen dit niet via het bovenstaande argument bewijzen omdat we wel hebben dat $X = \sum_i I_{A_i}$, waar A_i de gebeurtenis is dat het i-de paar samen blijft, $1 \le i \le N$, maar deze gebeurtenissen zijn **niet** onafhankelijk. Dus is X geen binomiaalverdeelde toevalsvariabele, maar convergeert nog steeds naar een Poissonverdeling als $N \to \infty$.

2.4 Absoluut continue verdelingen

We bekijken nu toevalsvariabelen die meer dan aftelbaar veel waarden aannemen. Velen van deze hebben een absoluut continue verdeling.

Definitie 2.4 *Een kansmaat μ op $\mathcal{R}$ heet absoluut continu als de verdelingsfunctie $x \to F(x)$ continu is en er een Borel-meetbare functie $f : \mathbb{R} \to [0, \infty[$ bestaat zodanig dat*

$$F(x) = \int_{-\infty}^{x} f(t) dt, x \in \mathbb{R}.$$

We noemen f de dichtheidsfunctie van μ.

Opmerkingen

1. De integraal in de definitie van F is de Lebesgue-integraal die voor elke positieve Borel-meetbare functie bestaat (met een waarde in $[0,\infty]$). Deze wordt formeel in de maattheorie (Bachelor 3) gedefinieerd. In deze cursus bekijken we bijna uitsluitend functies f die ook Riemann-integreerbaar zijn. Men kan tonen dat als $f \geq 0$ en de Riemann-integraal bestaat, deze gelijk aan de Lebesgue-integraal is en dus is de Riemann-integraal (nog) voldoende.

2. Vermits verdelingen door hun verdelingsfuncties bepaald zijn, geldt natuurlijk voor elke Borelverzameling $A \in \mathcal{R}$: $\mu(A) = \int_A f(t)dt$. Verder volgt als X een toevalsvariabele met een absoluut continue verdeling is: $\mathbb{P}\{X = x\} = 0 \ \ \forall x \in \mathbb{R}$.

Met behulp van de volgende stelling kunnen we vaak aantonen dat een gegeven verdeling absoluut continu is. Het niet-triviale bewijs daarvan wordt hier niet gegeven.

Stelling 2.6 *Als μ een kansmaat op $\mathcal{R}$ met een continue verdelingsfunctie F is, die behoudens in hoogstens aftelbaar veel punten een afgeleide F' bezit, dan is μ absoluut continu en we kunnen de dichtheidsfunctie definiëren door*

$$f(t) = \begin{cases} F'(t) & \textit{als } F \textit{ differentieerbaar in } t \textit{ is} \\ 0 & \textit{elders} \end{cases}$$

Opmerkingen De voorwaarde "de verdelingsfunctie F is continu" is belangrijk omdat elke verdelingsfunctie van een elementaire toevalsvariabele aan de tweede voorwaarde voldoet. (In dit geval is $F'(t) = 0$ voor alle punten t waar F continu is.) Er bestaan ook kansmaten met continue verdelingsfuncties die **niet** absoluut continu zijn. Deze spelen echter geen belangrijke rol in de toepassingen van de kanstheorie. Verder bestaan er absoluut continue verdelingen waar de verdelingsfuncties geen afgeleide bezitten voor meer dan aftelbaar veel punten (en stelling 2.6 niet van toepassing is). De volgende stelling toont dat men in dit geval ten minste voor de continuïteitspunten van f deze door differentiëren van de verdelingsfunctie kan verkrijgen.

Stelling 2.7 *Zij X een toevalsvariabele met een absoluut continue verdeling. Dan is de verdelingsfunctie van X differentieerbaar in alle punten waar de dichtheidsfunctie f continu is.*

Bewijs $\frac{1}{h}(F(x+h) - F(x)) = \begin{cases} \frac{1}{h}\mathbb{P}\{x < X \leq x+h\} & \text{als } h > 0 \\ -\frac{1}{h}\mathbb{P}\{x+h < X \leq x\} & \text{als } h < 0 \end{cases} = \frac{1}{h}\int_x^{x+h} f(t)dt \to f(x)$ als $h \to 0$ en f continu in x is. □

Zoals in het discrete geval bekijken we weer enkele speciale verdelingen.

Voorbeelden van absoluut continue verdelingen

(1) Zij $a < b$. Dan is de **uniform(a,b)-verdeling** de absoluut continue verdeling met dichtheidsfunctie

$$f(t) = \begin{cases} \frac{1}{b-a} & \text{als } t \in]a,b[\\ 0 & \text{als } t \notin]a,b[\end{cases}$$

Deze verdeling is een goed model voor een toevalsvariabele waar men weet dat waarden tussen a en b mogelijk zijn en voor deelintervallen van $]a, b[$ die dezelfde lengte hebben de kans dat X in zo'n interval zit identiek is. Random getallen op de computer (hoewel deze een resultaat van een deterministische algoritme zijn) beschouwt men als de waarden van een uniform(0,1)-toevalsvariabele. Deze zijn handig omdat men toevalsvariabelen met een willekeurige andere verdeling door een transformatie van uniform(0,1)-toevalsvariabelen kan verkrijgen.

Stelling 2.8 *Zij $F : \mathbb{R} \to [0, 1]$ een verdelingsfunctie. Stel*
$\phi(u) := \inf\{x : F(x) \geq u\}, 0 < u < 1.$
Als $U : \Omega \to]0, 1[$ uniform(0,1)-verdeeld is, heeft de toevalsvariabele $X := \phi \circ U$ de door F bepaalde verdeling.

Bewijs. We noteren eerst dat $\phi(u) \in \mathbb{R}, u \in]0, 1[$. Bovendien geldt:

$$\{x : F(x) \geq u\} = [\phi(u), \infty[$$

("$\subset$" is triviaal. "$\supset$" Zij $x \geq \phi(u)$ Uit de definitie van $\phi(u)$ als een infimum volgt dat er een rij $x_n \searrow \phi(u) \leq x$ bestaat zodanig dat $F(x_n) \geq u$ wat impliceert dat $F(\phi(u)) \geq u$ omdat F rechts continu is. Maar F is ook monotoon. Dus $F(x) \geq F(\phi(u)) \geq u$.)
We concluderen dat

$$\{\omega : X(\omega) \leq x\} = \{\omega : \phi(U(\omega) \leq x\} = \{\omega : U(\omega) \leq F(x)\}$$

Dit toont dat X $\mathcal{F}$-meetbaar is (gebruik stelling 2.2.i en het feit dat $\mathcal{R} = \sigma(\{]-\infty, x] : x \in \mathbb{R}\})$ en verder dat $\mathbb{P}\{X \leq x\} = \mathbb{P}\{U \leq F(x)\} = F(x), x \in \mathbb{R}$. □

Opmerkingen. Als er een open interval $I =]a, b[$ bestaat (eindig of oneindig) zodat $F|_I : I \to]0, 1[$ een 1-1 -afbeelding is, dan is ϕ gelijk aan de inverse afbeelding $(F|_I)^{-1} :]0, 1[\to I$. Vandaar dat we ϕ ook de *gegeneraliseerde inverse afbeelding* van F noemen.
Verder is het niet moeilijk te zien dat ϕ altijd links continu is (als gevolg van het feit dat F rechts continu is).

Voorbeeld Zij F de verdelingsfunctie van een Bernoulli(p)-toevalsvariabele, dus :
$F(x) = 0, x < 0, F(x) = 1 - p, 0 \leq x < 1, F(x) = 1, x \geq 1$. Dan is $\phi = I_{]1-p,1[}$.

(2) **De gamma-verdelingen.**
Herinnering (analyse): de gamma-functie is gedefinieerd door:

$$\Gamma(\alpha) := \int_0^\infty e^{-y} y^{\alpha-1} dy, \alpha > 0.$$

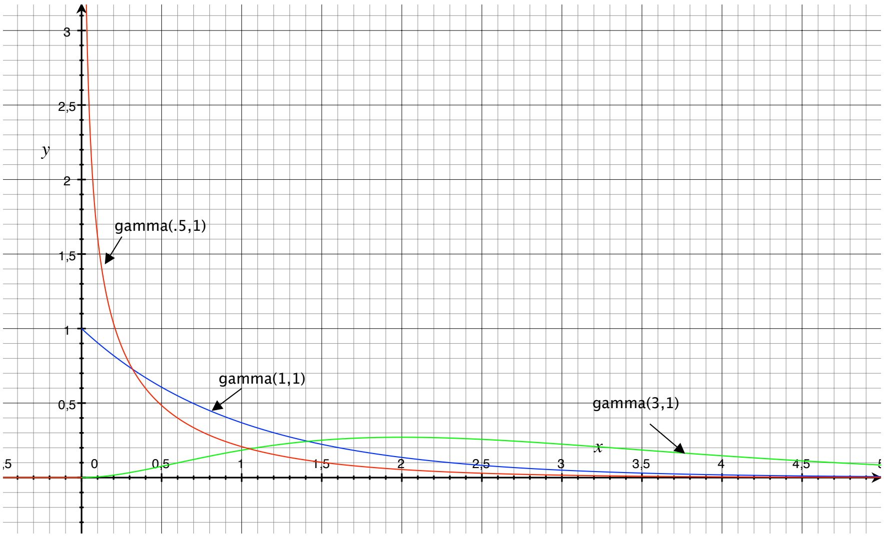
Verder geldt (partiële integratie):

$$\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1), \alpha > 1,$$

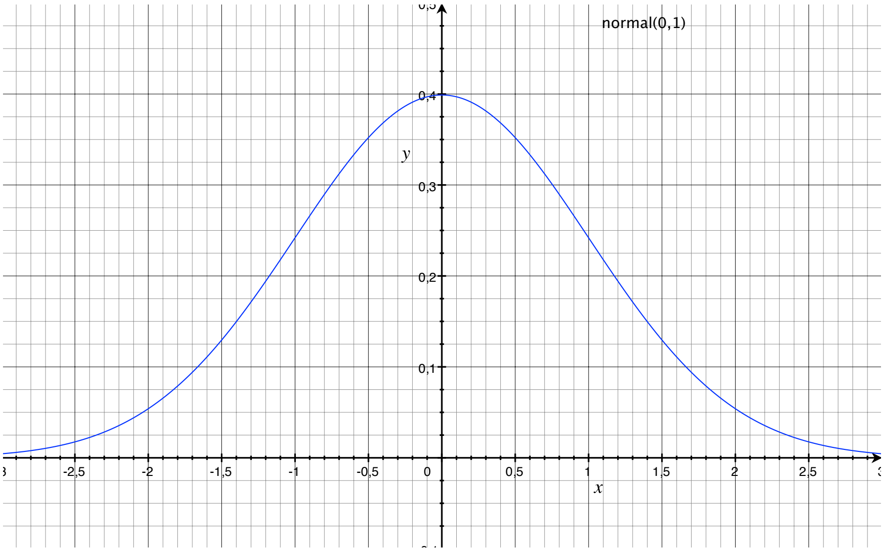
en bijgevolg: $\Gamma(k) = (k - 1)!, k = 1, 2, \ldots$
Als $\alpha, \beta > 0$ stellen we

$$f_{\alpha,\beta}(x) = \begin{cases} x^{\alpha-1} e^{-x/\beta} \beta^{-\alpha} / \Gamma(\alpha) & \text{als } x \geq 0 \\ 0 & \text{als } x < 0 \end{cases}$$

Gezien $\int_0^\infty f_{\alpha,\beta}(x)dx = 1$ (narekenen, substitutie), kunnen we een absoluut continue verdeling met $f_{\alpha,\beta}$ als dichtheidsfunctie definiëren. We noemen deze de gamma(α, β)-verdeling.
In het speciale geval $\alpha = 1$ spreken we ook van een **exponentiële** verdeling met parameter β. (in dit geval is de dichtheidsfunctie $f_{1,\beta}(x) = e^{-x/\beta}/\beta, x > 0$.)



(3) **Normaalverdelingen.**
Stel $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}, x \in \mathbb{R}$. Ga na dat $\int_{-\infty}^\infty \varphi(x) = 1$.
(Hint:$(\int_{-\infty}^\infty e^{-x^2/2}dx)^2 = \int_{-\infty}^\infty \int_{-\infty}^\infty e^{-(x^2+y^2)/2}dxdy$ waar de dubbelintegraal gemakkelijk via een transformatie naar poolcoördinaten te berekenen is.)



Dus is φ de dichtheidsfunctie van een absoluut continue verdeling die we de **standaard-normaal**-verdeling noemen.
We noteren de verdelingsfunctie van de standaard-normaal-verdeling met Φ, dus

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2}dt, x \in \mathbb{R},$$

waarvoor geen eenvoudige formule bestaat zodat we meestal een tabel of een computerprogramma moeten gebruiken om $\Phi(x)$ te bepalen.

Definitie 2.5 *Een toevalsvariabele heeft een normaal*(μ, σ^2)*-verdeling indien deze absoluut continu met dichtheidsfunctie* $x \to \varphi((x-\mu)/\sigma)/\sigma$ *is.*

Lemma 2.1 *Zij* X *een toevalsvariabele met normaal*(μ, σ^2)*-verdeling.*
Dan heeft $Z = (X-\mu)/\sigma$ *een standaard-normaalverdeling.*

Bewijs Via de substitutie $z = (t-\mu)/\sigma$ volgt:
$\mathbb{P}\{Z \le x\} = \mathbb{P}\{X \le \mu + \sigma x\} = \int_{-\infty}^{\mu+x\sigma} \sigma^{-1}\varphi((t-\mu)/\sigma)dt = \int_{-\infty}^{x} \varphi(z)dz = \Phi(x), x \in \mathbb{R}.\square$

Als we kansen voor algemene normaalverdeelde toevalsvariabelen moeten berekenen, kunnen we deze altijd via lemma 2.1 herschrijven als kansen betreffende standaard-normaalverdeelde toevalsvariabelen.

De volgende stelling is belangrijk voor de statistiek.

Stelling 2.9 *Zij* Z *een standaard-normaalverdeelde toevalsvariabele.*
Dan heeft Z^2 *een gamma(1/2,2)-verdeling.*

Bewijs Voor de verdelingsfunctie F van Z^2 geldt:
$F(x) = \mathbb{P}\{|Z| \le \sqrt{x}\} = \Phi(\sqrt{x}) - \Phi(-\sqrt{x}), x \ge 0$ en (triviaal) $F(x) = 0, x < 0$.
Dus is F continu en verder bestaat $F'(x), x \ne 0$ en we kunnen via stelling 2.6 concluderen dat Z^2 een dichtheidsfunctie f heeft gegeven door
$f(x) = (\Phi'(\sqrt{x}) - \Phi'(-\sqrt{x})/(2\sqrt{x}) = x^{-1/2}e^{-x/2}/\sqrt{2\pi}, x > 0$ en $f(x) = 0, x \le 0$. Dus geldt: $f(x) = cf_{1/2,2}(x), x \in \mathbb{R}$, waar $f_{1/2,2}$ de gamma(1/2,2)-dichtheidsfunctie en $c > 0$ een constante is. Gezien het om twee dichtheidsfuncties gaat (waar dus de integraal over $\mathbb{R}$ gelijk aan 1 is) volgt dat $c = 1$. (Dit impliceert trouwens ook dat $\Gamma(1/2) = \sqrt{\pi}$.) $\square$

De vorige stelling toont aan dat als X een standaard-normaalverdeelde toevalsvariabele is en $g(x) = x^2, x \in \mathbb{R}$, de toevalsvariabele $g \circ X$ een absoluut continue verdeling heeft en specificeert de dichtheidsfunctie.

Transformatie van dichtheidsfuncties

Een generalisatie daarvan is de volgende stelling waar we "m-1-afbeeldingen" bekijken. (Natuurlijk kunnen we zo'n stelling niet voor willekeurige afbeeldingen bewijzen omdat als $g(\mathbb{R})$ eindig of aftelbaar is, $g \circ X$ een discrete verdeling heeft.)

Stelling 2.10 *Zij* $X : \Omega \to \mathbb{R}$ *een toevalsvariabele met absoluut continue verdeling en dichtheidsfunctie* f_X, $g : \mathbb{R} \to \mathbb{R}$ *Borel-meetbaar en* $I_i, 1 \le i \le m$ *disjuncte open intervallen (eindig of oneindig) zodanig dat*

(i) $\sum_{i=1}^m \mathbb{P}\{X \in I_i\} = 1$

(ii) $g(I_i) =]c, d[, 1 \le i \le m,$ *waar* $c, d \in [-\infty, \infty]$.

(iii) $g_i := g|_{I_i} : I_i \to]c, d[$ *is differentieerbaar met een continue afgeleide* $g_i'(a) \ne 0, a \in I_i$.

Dan is de verdeling van $Y = g \circ X$ *ook absoluut continu met dichtheidsfunctie*

$$f_Y(y) = \begin{cases} \sum_{i=1}^m \frac{1}{|g'(h_i(y))|} f_X(h_i(y)) & \text{als } y \in]c, d[\\ 0 & \text{als } y \notin]c, d[\end{cases}$$

waarbij $h_i = g_i^{-1} :]c, d[\to I_i$ *de inverse functie van* g_i *is.* $(1 \le i \le m)$

Bewijs Zij $x \in]c,d[$. Dan volgt:

$$\begin{aligned}
\mathbb{P}\{Y \leq x\} &= \sum_{i=1}^{m} \mathbb{P}\{g \circ X \leq x, X \in I_i =]a_i, b_i[\} \\
&= \sum_{i: g_i' > 0} \mathbb{P}\{X \leq h_i(x), X \in]a_i, b_i[\} + \sum_{i: g_i' < 0} \mathbb{P}\{X \geq h_i(x), X \in]a_i, b_i[\} \\
&= \sum_{i: g_i' > 0} \int_{a_i}^{h_i(x)} f_X(t)dt + \sum_{i: g_i' < 0} \int_{h_i(x)}^{b_i} f_X(t)dt \\
&= \sum_{i=1}^{m} \int_c^x |h_i'(y)| f_X(h_i(y))dy \quad (\text{substitutie: } t = h_i(y), 1 \leq i \leq m) \\
&= \int_c^x \sum_{i=1}^{m} \frac{1}{|g'(h_i(y))|} f_X(h_i(y))dy
\end{aligned}$$

Dus heeft Y dezelfde verdelingsfunctie als de absoluut continue kansmaat met de bovenstaande dichtheidsfunctie. Dit impliceert dat deze dan ook de verdeling van Y moet zijn (zie opmerking (3), hoofdstuk 2.2). $\square$

Voorbeelden

1. Als X een toevalsvariabele met $\mathbb{P}\{X \in]a,b[\} = 1$, $g :]a,b[\to]c,d[$ een 1-1-afbeelding met continue afgeleide $g'(t) \neq 0, t \in]a,b[$ is en $h :]c,d[\to]a,b[$ de inverse afbeelding van g is, dan geldt voor de dichtheidsfunctie f_Y van $Y = g \circ X$:

$$f_Y(y) = \begin{cases} \frac{1}{|g'(h(y))|} f_X(h(y)) & y \in]c,d[\\ 0 & \text{elders.} \end{cases}$$

   Bijvoorbeeld, als X uniform(0,1)-verdeeld is volgt dat $Y = -\log(X)$ een exponentiële verdeling met parameter 1 heeft. (Stel $g(x) = -\log(x), x \in]0,1[$.)

2. Zij X een toevalsvariabele met absoluut continue verdeling en dichtheidsfunctie f_X. Dan heeft de toevalsvariabele X^2 altijd een absoluut continue verdeling met dichtheidsfunctie $g = f_{X^2}$ die voldoet aan

$$g(y) = \frac{1}{2\sqrt{y}}(f_X(\sqrt{y}) + f_X(-\sqrt{y})), \ y > 0.$$

   Om dit via stelling 2.11 te bewijzen, stellen we

$$g(x) = x^2, x \in \mathbb{R}, I_1 =]0,\infty[, I_2 =]-\infty, 0[\text{ en }]c,d[=]0,\infty[.$$

   Dan is $h_1(y) = \sqrt{y}, h_2(y) = -\sqrt{y}, y > 0$ en we verkrijgen onmiddellijk de bovenstaande formule voor g.

   Als X standaard-normaalverdeeld is, kunnen we ook via bovenstaande formule tonen dat X^2 een gamma(1/2,2)-verdeling heeft.

   We kunnen verder concluderen dat als X een toevalsvariabele met een uniform(0,1)-verdeling is, dwz $f_X = I_{]0,1[}$, we hebben:

$$g(y) = \begin{cases} \frac{1}{2\sqrt{y}} & \text{als } 0 < y < 1 \\ 0 & \text{elders} \end{cases}$$

   (Hier hebben we het triviale feit gebruikt dat $I_{]0,1[}(-\sqrt{y}) = 0, y > 0$.)

Hoofdstuk 3

Toevalsvectoren en verdelingen

3.1 Gezamenlijke en marginale verdelingsfuncties

Zij $(\Omega, \mathcal{F}, \mathbb{P})$ een kansruimte. We noemen dan de $\mathcal{F}$-meetbare afbeeldingen $\vec{X} : \Omega \to \mathbb{R}^k$ de (k-dimensionale) toevalsvectoren. Uit stelling 2.3 volgt dat $\vec{X} = (X_1, \ldots, X_k) : \Omega \to \mathbb{R}^k$ toevalsvector is als en slechts als $X_i : \Omega \to \mathbb{R}$ toevalsvariabelen zijn, $1 \le i \le k$.
Dus als X_1, X_2 toevalsvariabelen zijn en we kansen zoals $\mathbb{P}\{X_1 > X_2\}, \mathbb{P}\{X_1 + X_2 \le x\}$ moeten berekenen, kunnen we deze als kansen betreffende een 2-dimensionale toevalsvector beschouwen, dwz als $\mathbb{P}\{(X_1, X_2) \in A\}$ voor een tweedimensionale Borelverzameling.

Definitie 3.1 *Als $X_1, \ldots, X_k$ toevalsvariabelen zijn, noemen we de door*

$$\mathbb{P}\{(X_1, \ldots, X_k) \in A\}, A \in \mathcal{R}^k$$

gedefinieerde kansmaat de **gezamenlijke verdeling** *van $X_1, \ldots, X_k$ of de* **verdeling van de toevalsvector** $\vec{X} = (X_1, \ldots, X_k)^t$*. Notatie: $\mathbb{P}_{\vec{X}}$.*

Dat we op deze manier een kansmaat verkrijgen, volgt precies zoals in stelling 2.4. We noemen dan verder de (1-dimensionale) verdelingen van de toevalsvariabelen $X_1, \ldots, X_k$ de **marginale** verdelingen.

Als we de gezamenlijke verdeling kennen, kunnen we de marginale verdeling onmiddellijk bepalen (bijvoorbeeld geldt: $\mathbb{P}\{X_1 \in B\} = \mathbb{P}\{(X_1, \ldots, X_k) \in B \times \mathbb{R}^{k-1}\}$), maar in het algemeen kunnen we niet de gezamenlijke verdeling bepalen, als we maar informatie over de marginale verdelingen hebben.

Definitie 3.2 *Als $X_1, \ldots, X_k$ toevalsvariabelen zijn, noemen we de functie*

$$F_{\vec{X}}(\underline{x}) := \mathbb{P}\{X_1 \le x_1, \ldots, X_k \le x_k\}, \underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \in \mathbb{R}^k$$

de (gezamenlijke) verdelingsfunctie van de toevalsvariabelen $X_1, \ldots, X_k$.

Zoals in het 1-dimensionale geval kan men via de maattheorie bewijzen dat er een eenduidig verband tussen k-dimensionale verdelingsfuncties en k-dimensionale verdelingen bestaat. In het bijzonder geldt:

Stelling 3.1 *Als $X_1, \ldots, X_k$ en $Y_1, \ldots, Y_k$ toevalsvariabelen zijn zodat de twee gezamenlijke verdelingsfuncties overeenstemmen, i.e. $F_{\vec{X}}(\underline{x}) = F_{\vec{Y}}(\underline{x})$, $\underline{x} \in \mathbb{R}^k$, hebben we: $\mathbb{P}_{\vec{X}} = \mathbb{P}_{\vec{Y}}$.*

In principe kunnen we dus kansen betreffende toevalsvariabelen $X_1, \ldots, X_k$ via hun gezamenlijke verdelingsfunctie berekenen. Maar dit kan vrijwel ingewikkeld worden.

Voorbeeld Stel X, Y toevalsvariabelen met gezamenlijke verdelingsfunctie $F(x,y), x, y \in \mathbb{R}$. Bepaal: (a) $p_1 = \mathbb{P}\{1 < X \le 2, 3 < Y \le 4\}$, (b) $p_2 = \mathbb{P}\{X > 1, Y > 2\}$.

Oplossing

(a) $F(2,4) = \mathbb{P}\{X \le 2, Y \le 4\} = \mathbb{P}\{1 < X \le 2, 3 < Y \le 4\}$
$+\mathbb{P}\{X \le 1, 3 < Y \le 4\}$ $+\mathbb{P}\{X \le 2, Y \le 3\} = p_1 + F(1,4) - F(1,3) + F(2,3)$
Dus geldt: $p_1 = F(2,4) - F(1,4) + F(1,3) - F(2,3)$.

(b) $p_2 = 1 - \mathbb{P}(\{X \le 1\} \cup \{Y \le 2\}) = 1 - \mathbb{P}\{X \le 1\} - \mathbb{P}\{Y \le 2\} + F(1,2)$.
Gezien $\mathbb{P}\{X \le 1\} = \lim_{y \to \infty} F(1,y) =: F(1,\infty)$ en analoog $\mathbb{P}\{Y \le 2\} = \lim_{x \to \infty} F(x,2) =: F(\infty,2)$ volgt dat $p_2 = 1 - F(1,\infty) - F(\infty,2) + F(1,2)$.

In het vorige voorbeeld hebben we de volgende stelling gebruikt.

Stelling 3.2 *Als $X_1, \ldots, X_k$ toevalsvariabelen met gezamenlijke verdelingsfunctie $F(x_1, \ldots, x_k)$ zijn, geldt voor de marginale verdelingsfuncties*

$$F_{X_i}(x_i) = \mathbb{P}\{X_i \le x_i\} = \lim_{\substack{x_j \to \infty \\ j \neq i}} F(x_1, \ldots, x_k), x_i \in \mathbb{R}, 1 \le i \le k.$$

Bewijs Zij $A_n = \{X_j \le x_j^{(n)}, j \neq i, X_i \le x_i\}$, waar $x_j^{(n)} \nearrow \infty$. Dan is het evident dat $A_n \nearrow \{X_i \le x_i\}$ en dus geldt $\mathbb{P}(A_n) \to F_{X_i}(x_i)$ als $n \to \infty$. □

3.2 Discrete verdelingen

We bekijken hoofdzakelijk het geval $k = 2$. De uitbreiding naar hogere dimensies is meestal evident.
Als X, Y discrete toevalsvariabelen zijn, dan bestaan er ten hoogste aftelbaar veel paren (x,y) zodanig dat de **gezamenlijke kansfunctie** $p(x,y) := \mathbb{P}\{X = x, Y = y\} > 0$ is en we kunnen de kansen via de volgende formule berekenen:

$$\mathbb{P}\{(X,Y) \in A\} = \sum_{(x,y) \in S \cap A} p(x,y), A \subset \mathbb{R}^2$$

waar $S : \{(x,y) \in \mathbb{R}^2 : p(x,y) > 0\}$. (De som over de lege verzameling is gedefinieerd als 0.)

De marginale kansfuncties van X en Y kunnen we via de volgende stelling bepalen.

Stelling 3.3 *Als X, Y discrete toevalsvariabelen met kansfunctie $p(x,y)$ zijn, geldt: (a) $\mathbb{P}\{X = x\} = \sum_y p(x,y), x \in \mathbb{R}$ en (b) $\mathbb{P}\{Y = y\} = \sum_x p(x,y), y \in \mathbb{R}$.*

Bewijs (a) Zij $S_x := \{y : p(x,y) > 0\}$. Dan is S_x ten hoogste aftelbaar en het geldt:

$$\begin{aligned} \mathbb{P}\{X = x\} &= \mathbb{P}((X,Y) \in \{x\} \times S_x) + \mathbb{P}((X,Y) \in \{x\} \times S_x^c) \\ &= \sum_{y \in S_x} p(x,y) + 0 = \sum_y p(x,y) \end{aligned}$$

Het bewijs van (b) verloopt analoog. □

Voorbeelden

1. *In een doos zitten 3 ballen waarop de nummers 1,2,3 staan. We trekken 2 ballen (a) met teruglegging en (b) zonder teruglegging. Stel X = het nummer op de eerste bal, Y = het nummer op de tweede bal. Bepaal de gezamenlijke kansfunctie van X en Y en de twee marginale kansfuncties p_X, p_Y.*

   Oplossing In geval (a) geldt $p(x,y) = 1/9, (x,y) \in \{1,2,3\}^2$.
   Dus $p_X(x) = 1/3, x = 1,2,3$ en $p_Y(y) = 1/3, y = 1,2,3$.

   In geval (b) geldt $p(x,y) = 1/6, x \neq y$, maar niettemin zijn p_X, p_Y zoals in (a).
   Dit toont dat we in het algemeen gezamenlijke verdelingen niet via de marginale verdelingen kunnen bepalen.

2. *Beschouw twee discrete toevalsvariabelen X, Y met gezamenlijke kansfunctie*

$$p(x,y) = \frac{\lambda^y}{x!(y-x)!} e^{-2\lambda}, x = 0,1,2,\ldots, y = x, x+1, x+2, \ldots$$

   Bepaal: (a) de marginale verdelingen, (b) $\mathbb{P}\{X = Y\}$.
   Oplossing (a) Een directe toepassing van stelling 3.3 levert:

$$p_X(x) = e^{-2\lambda} \sum_{y=x}^{\infty} \frac{\lambda^y}{x!(y-x)!} = e^{-2\lambda} \frac{\lambda^x}{x!} \sum_{y=x}^{\infty} \frac{\lambda^{y-x}}{(y-x)!} = e^{-\lambda} \frac{\lambda^x}{x!}, x = 0,1,\ldots$$

   waaruit blijkt dat X Poisson(λ)-verdeeld is.
   Analoog volgt:

$$p_Y(y) = \sum_{x=0}^{y} \frac{\lambda^y}{x!(y-x)!} e^{-2\lambda} = \frac{\lambda^y}{y!} e^{-2\lambda} \sum_{x=0}^{y} \binom{y}{x} = \frac{(2\lambda)^y}{y!} e^{-2\lambda}, y = 0,1,2,\ldots,$$

   Dus is Y Poisson-verdeeld met parameter 2λ.
   (b) $\mathbb{P}\{X = Y\} = \sum_{x=0}^{\infty} p(x,x) = e^{-2\lambda} \sum_{x=0}^{\infty} \lambda^x / x! = e^{-\lambda}$.

3. **De multinomiale verdeling**
   Veronderstel dat we een experiment dat r mogelijke uitkomsten heeft ($r \geq 2$) n keer kunnen uitvoeren zodat de uitkomsten van de n experimenten onafhankelijk van elkaar zijn. Bepaal de verdeling van $(X_1, \ldots, X_r)$, waar X_i = het aantal experimenten met uitkomst i. ($i = 1, \ldots, r$)

   Oplossing We hebben $X_1 + \ldots X_r = n$ en verder

$$\mathbb{P}\{X_1 = n_1, \ldots, X_r = n_r\} = \binom{n}{n_1, \ldots, n_r} p_1^{n_1} \cdot \ldots \cdot p_r^{n_r},$$

   als $n_1, \ldots, n_r = 0,1,2\ldots$ zodat $n_1 + \ldots + n_r = n$ en p_i = kans op uitkomst $i, 1 \leq i \leq r$.
   ($\Rightarrow$ $p_1 + \ldots p_r = 1$)
   De marginale verdelingen van de toevalsvariabelen X_i zijn dan binomiaal(n, p_i), $i = 1, \ldots, r$.
   Als we bijvoorbeeld een dobbelsteen n keer gooien en X_i aangeeft hoe vaak we "i" verkrijgen, $1 \leq i \leq 6$, dan heeft $(X_1, \ldots, X_6)$ een multinomiale verdeling, waar $p_i = 1/6, 1 \leq i \leq 6$ (onder de voorwaarde dat de n worpen onafhankelijk zijn.)

3.3 Toevalsvariabelen met een gezamenlijk continue verdeling

Zoals in het vorige hoofdstuk zullen we ons hoofdzakelijk beperken tot het geval $k = 2$.

Definitie 3.3 *Twee toevalsvariabelen* $X, Y : \Omega \to \mathbb{R}$ *hebben een* **gezamenlijk continue verdeling** *indien er een Borel-meetbare functie* $f : \mathbb{R}^2 \to [0, \infty[$ *bestaat zodanig dat*

$$\mathbb{P}\{(X, Y) \in A\} = \iint_A f(x, y)dxdy$$

voor alle Borelverzamelingen $A \in \mathcal{R}^2$. *We noemen* f *de dichtheidsfunctie van* (X, Y).

Stelling 3.4 *Als* X, Y *toevalsvariabelen met een gezamenlijk continue verdeling zijn, dan zijn de (marginale) verdelingen van* X *en* Y *absoluut continu met dichtheidsfuncties* f_X, f_Y *gegeven door*

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy, x \in \mathbb{R} \text{ en } f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx, y \in \mathbb{R}.$$

Bewijs Uit de definitie van "gezamenlijk continu" volgt:

$$\mathbb{P}\{X \leq t\} = \mathbb{P}\{(X, Y) \in]-\infty, t] \times \mathbb{R}\} = \int_{-\infty}^{t} \int_{-\infty}^{\infty} f(x, y)dydx = \int_{-\infty}^{t} f_X(x)dx, t \in \mathbb{R}.$$

Dus is de verdelingsfunctie van X gelijk aan de verdelingsfunctie van de absoluut continue verdeling met dichtheidsfunctie f_X en dit impliceert dat X deze verdeling heeft (zie opmerking (3), hoofdstuk 2.2). Analoog volgt dat Y absoluut continu met dichtheidsfunctie f_Y is. □

Opmerking Als X, Y absoluut continue verdelingen hebben, impliceert dit in het algemeen niet dat X, Y gezamenlijk continu zijn. Om dit in te zien, bekijken we een standaardnormaalverdeelde toevalsvariabele X en we stellen $Y = X$.
Dan hebben X, Y absoluut continue verdelingen, maar ze zien niet gezamenlijk continu omdat $\boxed{\mathbb{P}\{(X, Y) \in D\} = 1,}$ waar $D = \{(x, x) : x \in \mathbb{R}\}$.
Anderzijds geldt voor elke Borel-meetbare functie $f : \mathbb{R}^2 \to [0, \infty[$: $\boxed{\iint_D f(x, y)dxdy = 0.}$

Voorbeelden

1. Stel X, Y toevalsvariabelen met een gezamenlijk continue verdeling, waar $f(x, y) = e^{-x-y}, x, y > 0$ en $f(x, y) = 0$ elders. Bepaal:
   (a) $\mathbb{P}\{X > 2, Y > 3\}$ en
   (b) $\mathbb{P}\{X > 2Y\}$.
   (c) Toon dat X/Y een absoluut continue verdeling heeft en bepaal de dichtheidsfunctie van X/Y.

   Oplossing De kans in (a) is gelijk aan $\int_2^{\infty} \int_3^{\infty} e^{-y}dye^{-x}dx = e^{-5}$.
   De kans in (b) is gelijk aan $1 - F(2)$, waar F de verdelingsfunctie van X/Y is zodat

we de oplossing van (b) kunnen verkrijgen via (c).
Het is evident dat voor $t > 0$ geldt:

$$\begin{aligned} F(t) &= \mathbb{P}\{X/Y \le t\} = \mathbb{P}\{Y \ge X/t\} \\ &= \int_0^\infty \int_{x/t}^\infty e^{-y} dy e^{-x} dx \\ &= \int_0^\infty e^{-x(t+1)/t} dx = \frac{t}{t+1} \end{aligned}$$

We zien nu dat $F(t) = \begin{cases} 1 - \frac{1}{t+1} & \text{als } t > 0 \\ 0 & \text{als } t \le 0 \end{cases}$ continu is en dat verder de afgeleide overal behalve in $t = 0$ bestaat. Uit stelling 2.6 volgt dat X/Y een absoluut continue verdeling heeft.

Bovendien geldt voor de dichtheidsfunctie f van X/Y: $f(t) = \begin{cases} (t+1)^{-2} & \text{als } t > 0 \\ 0 & \text{als } t \le 0 \end{cases}$

Ten slotte volgt uit de formule voor F dat de kans in (b) gelijk aan $1/3$ is.

2. Onderstel dat X, Y toevalsvariabelen met een gezamenlijk continue verdeling zijn, waar

$$f(x,y) = \frac{1}{2\pi} e^{-y^2 - xy - x^2/2}, (x,y) \in \mathbb{R}^2.$$

Bepaal de marginale dichtheidsfuncties f_X en f_Y.

Oplossing Dit is een directe toepassing van stelling 3.4.

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi} \int_{-\infty}^\infty e^{-y^2 - xy - x^2/2} dy \\ &= \frac{1}{2\pi} e^{-x^2/2} \int_{-\infty}^\infty e^{-(y+x/2)^2} e^{x^2/4} dy \\ &= \frac{1}{2\pi} e^{-x^2/4} \int_{-\infty}^\infty e^{-u^2} du \text{ (substitutie)} \\ &= c e^{-x^2/4}, x \in \mathbb{R}. \end{aligned}$$

Dus geldt $f_X(x) = \bar{c} g(x), x \in \mathbb{R}$ waar g de dichtheidsfunctie van een normaal(0,2)-toevalsvariabele is, en het volgt dat $\bar{c} = 1$ (of $c = 1/\sqrt{4\pi}$). Dus heeft X een normaal(0,2)-verdeling.

Analoog volgt:

$$\begin{aligned} f_Y(y) &= \frac{1}{2\pi} \int_{-\infty}^\infty e^{-y^2 - xy - x^2/2} dx \\ &= \frac{1}{2\pi} e^{-y^2/2} \int_{-\infty}^\infty e^{-(y+x)^2/2} dx \\ &= \frac{1}{2\pi} e^{-y^2/2} \int_{-\infty}^\infty e^{-u^2/2} du \text{ (substitutie)} \\ &= \frac{1}{\sqrt{2\pi}} e^{-y^2/2}, y \in \mathbb{R}. \end{aligned}$$

Dus is Y standaard-normaalverdeeld. (Later in hoofdstuk 3.5 zullen we zien dat (X,Y) een twee-dimensionale normaalverdeling heeft en we kunnen de marginale verdelingen dan op een meer efficiënte manier bepalen.)

Als het laatste resultaat in 3.3 vermelden we nog een k-dimensionale versie van stelling 2.10.

Stelling 3.5 *Zij $X_1, \dots, X_k : \Omega \to \mathbb{R}$ toevalsvariabelen met een gezamenlijk continue verdeling en dichtheidsfunctie f, $g = \begin{bmatrix} g_1 \\ \vdots \\ g_k \end{bmatrix} : \mathbb{R}^k \to \mathbb{R}^k$ een Borel-meetbare afbeelding en zij $I_i, 1 \le i \le m$ disjuncte open verzamelingen in $\mathbb{R}^k$ zodanig dat*

(i) $\sum_{i=1}^m \mathbb{P}\{(X_1, \dots, X_k) \in I_i\} = 1$
(ii) $g(I_i) = I_0, 1 \le i \le m$, *waar I_0 een open verzameling is.*
(iii) $g|_{I_i} : I_i \to I_0$ *is 1-1 met inverse afbeelding $h_i : I_0 \to I_i$, $1 \le i \le m$.*
(iv) *g heeft continue partiële afgeleiden op $\bigcup_{i=1}^m I_i$ zodanig dat*

$$J_g(\underline{a}) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1}(\underline{a}) & \dots & \frac{\partial g_1}{\partial x_k}(\underline{a}) \\ \vdots & & \vdots \\ \frac{\partial g_k}{\partial x_1}(\underline{a}) & \dots & \frac{\partial g_k}{\partial x_k}(\underline{a}) \end{vmatrix} \neq 0, \ \underline{a} \in \bigcup_{i=1}^m I_i$$

Dan zijn de toevalsvariabelen $Y_1 = g_1(X_1, \dots, X_k), \dots, Y_k = g_k(X_1, \dots, X_k)$ gezamenlijk continu met dichtheidsfunctie

$$f_{\vec{Y}}(\underline{y}) = \begin{cases} \sum_{i=1}^m \frac{1}{|J_g(h_i(\underline{y}))|} f_{\vec{X}}(h_i(\underline{y})) & \text{als } \underline{y} \in I_0 \\ 0 & \text{als } \underline{y} \notin I_0 \end{cases}$$

Gevolg 3.1 *Zij A een (k,k)-matrix met $\det(A) \neq 0$ en zij $\underline{\mu}$ een vector in $\mathbb{R}^k$. Als $X_1, \dots, X_k$ toevalsvariabelen met een gezamenlijk continue verdeling en dichtheidsfunctie $f_{\vec{X}}$ zijn, volgt dat ook $\vec{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_k \end{bmatrix} = A \begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix} + \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_k \end{bmatrix}$ gezamenlijk continu is met dichtheidsfunctie*

$$\underline{y} \to f_{\vec{Y}}(\underline{y}) = \frac{1}{|\det(A)|} f_{\vec{X}}(A^{-1} \cdot (\underline{y} - \underline{\mu})), \underline{y} \in \mathbb{R}^k.$$

Bewijs Stel $(g_1(\underline{x}), \dots, g_k(\underline{x}))^t = A\underline{x} + \underline{\mu}$, $\underline{x} \in \mathbb{R}^k$ en gebruik stelling 3.5. □

Dit impliceert bijvoorbeeld dat als X_1, X_2 gezamenlijk continu zijn, ook

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} X_1 + X_2 \\ X_1 - X_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

gezamenlijk continu is met dichtheidsfunctie

$$f_{\vec{Y}}(y_1, y_2) = \frac{1}{2} f_{\vec{X}}\left(\frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2}\right), (y_1, y_2) \in \mathbb{R}^2.$$

3.4 Onafhankelijke toevalsvariabelen

Definitie 3.4 *Als $X_1, \dots, X_k$ toevalsvariabelen zijn, zeggen we dat deze* **onafhankelijk** *zijn indien voor alle Borelverzamelingen $A_1, \dots, A_k \in \mathcal{R}$ geldt:*

$$\mathbb{P}\{X_1 \in A_1, \dots, X_k \in A_k\} = \prod_{j=1}^k \mathbb{P}\{X_j \in A_j\}.$$

Opmerkingen

1. Een equivalente definitie is dat voor elke keuze $B_i \in \mathcal{R}, 1 \le i \le k$ de gebeurtenissen $\{X_1 \in B_1\}, \ldots, \{X_k \in B_k\}$ onafhankelijk zijn.
(Het is triviaal dat dit onafhankelijkheid zoals in de bovenstaande definitie impliceert. Voor de omgekeerde implicatie moeten we aantonen dat
$$\mathbb{P}(\bigcap_{i \in I}\{X_i \in B_i\}) = \prod_{i \in I} \mathbb{P}\{X_i \in B_i\}, I \subset \{1, \ldots, k\}.$$
Maar dit volgt onmiddellijk als we in definitie 3.4 $A_i = B_i, i \in I$ en $A_i = \mathbb{R}, i \notin I$ stellen.)

2. Uit bovenstaande definitie volgt ook direct dat als $f_i : \mathbb{R} \to \mathbb{R}$ Borel-meetbaar zijn en $X_i, 1 \le i \le k$ onafhankelijk zijn, dat dit dan ook het geval voor de toevalsvariabelen $Y_i = f_i \circ X_i, 1 \le i \le k$ is.

De volgende stelling geeft een eenvoudige methode (ten minste in het discrete en continue geval) om na te gaan of toevalsvariabelen onafhankelijk zijn. Daarbij is $F(x_1, \ldots, x_k)$ de gezamenlijke verdelingsfunctie van $X_1, \ldots, X_k$ en F_{X_j} de (marginale) verdelingsfunctie van X_j, $1 \le j \le k$. Analoog is $p_{X_1,\ldots,X_k}$ ($f_{X_1,\ldots,X_k}$) de gezamenlijke kansfunctie (dichtheidsfunctie) en p_{X_j} (f_{X_j}) de kansfunctie (dichtheidsfunctie) van $X_j, 1 \le j \le k$ in het discrete (continue) geval.

Stelling 3.6 *Zij $X_1, \ldots, X_k$ toevalsvariabelen. Dan zijn equivalent:*
(i) $X_1, \ldots, X_k$ *zijn onafhankelijk.*
(ii) $F_{X_1,\ldots,X_k}(x_1, \ldots, x_k) = \prod_{j=1}^k F_{X_j}(x_j), \forall \underline{x} \in \mathbb{R}^k$.
(iii) $\begin{cases} p_{X_1,\ldots,X_k}(x_1, \ldots, x_k) = \prod_{j=1}^k p_{X_j}(x_j), \forall \underline{x} \in \mathbb{R}^k & \textit{(discreet geval)} \\ f_{X_1,\ldots,X_k}(x_1, \ldots, x_k) = \prod_{j=1}^k f_{X_j}(x_j), \forall \underline{x} \in \mathbb{R}^k & \textit{(continu geval)} \end{cases}$

Bewijs De implicatie "(i) $\Rightarrow$ (ii)" is evident. (Stel in definitie 3.4 $A_j =]-\infty, x_j], 1 \le j \le k$.)
Voor de implicatie"(ii) $\Rightarrow$ (i)"hebben we de maattheorie nodig. ($\to$ bachelor 3)
We bekijken nu het discrete geval. Als we in definitie 3.4 $A_j = \{x_j\}, 1 \le j \le k$ stellen, zien we dat (i) in dit geval (iii) impliceert. Om de implicatie "(iii) $\Rightarrow$ (i)"te bewijzen, stellen we $S_j = \{x_j \in \mathbb{R} : \mathbb{P}\{X_j = x_j\} > 0\}, 1 \le j \le k$. Dan zijn deze verzamelingen ten hoogste aftelbaar en het volgt voor $A_j \subset \mathbb{R}, 1 \le j \le k$ dat

$$\begin{aligned} \mathbb{P}\{X_1 \in A_1, \ldots, X_k \in A_k\} &= \sum_{x_1 \in A_1 \cap S_1} \cdots \sum_{x_k \in A_k \cap S_k} p_{X_1,\ldots,X_k}(x_1, \ldots, x_k) \\ &= \sum_{x_1 \in A_1 \cap S_1} \cdots \sum_{x_k \in A_k \cap S_k} \prod_{j=1}^k p_{X_j}(x_j) \\ &= \prod_{j=1}^k \sum_{x_j \in A_j \cap S_j} p_{X_j}(x_j) = \prod_{j=1}^k \mathbb{P}\{X_j \in A_j\} \end{aligned}$$

en we zien dat de toevalsvariabelen $X_1, \ldots, X_k$ onafhankelijk zijn.
In het continue geval is de implicatie "(iii) $\Rightarrow$ (i)" evident (oef.), terwijl we weer de maattheorie nodig hebben om "(i) $\Rightarrow$ (iii)" te bewijzen. □

Voorbeelden (1) Als we nog een keer de twee voorbeelden in 3.3 (na stelling 3.4) bekijken, zien we dat de toevalsvariabelen in het eerste voorbeeld onafhankelijk zijn, terwijl de toevalsvariabelen uit het tweede voorbeeld afhankelijk (= **niet** onafhankelijk) zijn.

(2) Onderstel dat X_1, X_2 onafhankelijke toevalsvariabelen zijn zodanig dat X_i normaal$(0, \sigma_i^2)$-verdeeld is, $i = 1, 2$. Stel $Y_1 = X_1 + X_2$, $Y_2 = -\sigma_2^2 X_1 + \sigma_1^2 X_2$. We beweren dat dan deze twee toevalsvariabelen ook onafhankelijk zijn.

Bewijs Vermits X_1 en X_2 onafhankelijk zijn, volgt dat ze gezamenlijk continu zijn met dichtheidsfunctie

$$f_{\vec{X}}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-x_1^2/(2\sigma_1^2) - x_2^2/(2\sigma_2^2)}$$

Stel $A = \begin{bmatrix} 1 & 1 \\ -\sigma_2^2 & \sigma_1^2 \end{bmatrix}$. Dan geldt $\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = A \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ en gevolg 3.1 impliceert dat de toevalsvariabelen Y_1, Y_2 gezamenlijk continu zijn met dichtheidsfunctie

$$f_{\vec{Y}}(y_1, y_2) = \frac{1}{|\det(A)|} f_{\vec{X}}(A^{-1}(y_1, y_2)^t)$$

Na een beetje rekenwerk kunnen we dit herschrijven als

$$f_{\vec{Y}}(y_1, y_2) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} e^{-y_1^2/2(\sigma_1^2+\sigma_2^2)} \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}\sigma_1\sigma_2} e^{-y_2^2/2(\sigma_1^2+\sigma_2^2)\sigma_1^2\sigma_2^2} =: f_1(y_1) f_2(y_2)$$

Gezien zowel f_1 als f_2 dichtheidsfuncties van normaalverdelingen zijn, volgt onmiddellijk dat de marginale dichtheidsfunctie van Y_i gelijk aan f_i is, $i = 1, 2$ en uit stelling 3.6 volgt dat Y_1 en Y_2 inderdaad onafhankelijk zijn. $\square$

Opmerkingen

1. We noteren dat ons bewijs ook toont dat $Y_1 = X_1 + X_2$ normaal$(0, \sigma_1^2 + \sigma_2^2)$ verdeeld is als de toevalsvariabelen X_i normaal $(0, \sigma_i^2)$-verdeeld en onafhankelijk zijn.

2. Via inductie volgt dan ook dat als Y_i onafhankelijk en normaal(μ_i, σ_i^2)-verdeeld zijn, $1 \leq i \leq k$, dat dan $\sum_{i=1}^k Y_i$ normaal$(\sum_{i=1}^k \mu_i, \sum_{i=1}^k \sigma_i^2)$-verdeeld is. (Gebruik ook het feit dat in dit geval $X_i = Y_i - \mu_i$ normaal$(0, \sigma_i^2)$-verdeeld is, $1 \leq i \leq k$.)

3.5 De k-dimensionale normaalverdeling

Als $X_1, \ldots, X_k$ onafhankelijke standaard-normaalverdeelde toevalsvariabelen zijn, kunnen we hun gezamenlijke dichtheidsfunctie die gelijk aan het product van de marginale dichtheidsfuncties is, opschrijven als

$$f_{\vec{X}}(\underline{x}) = (2\pi)^{-k/2} \exp(-\|\underline{x}\|^2/2) = (2\pi)^{-k/2} \exp(-\langle \underline{x}, \underline{x} \rangle/2),$$

waar $\langle \underline{x}, \underline{y} \rangle = \sum_{i=1}^k x_i y_i$ het scalaire product in $\mathbb{R}^k$ is.

Zij verder A een (k, k)-matrix met maximaalrang en $\underline{\mu} = (\mu_1, \ldots, \mu_k)^t$ een vector in $\mathbb{R}^k$. Stel

$$(Y_1, \ldots, Y_k)^t = (\mu_1, \ldots, \mu_k)^t + A(X_1, \ldots, X_k)^t \tag{3.1}$$

Dan volgt uit stelling 3.5 dat $(Y_1, \ldots, Y_k)$ gezamenlijk continu is met dichtheidsfunctie

$$f_{\vec{Y}}(\underline{y}) = \frac{1}{\sqrt{2\pi}^k \sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}\langle \underline{y} - \underline{\mu}, \Sigma^{-1}(\underline{y} - \underline{\mu}) \rangle\right), \ \underline{y} \in \mathbb{R}^k, \tag{3.2}$$

waar $\Sigma = A \cdot A^t$ een positief definiete, symmetrische (k, k)-matrix is.

Definitie 3.5 *Zij Σ een positief definiete, symmetrische matrix en zij $\underline{\mu}$ een vector in $\mathbb{R}^k$. We noemen dan de k-dimensionale verdeling die door bovenstaande dichtheidsfunctie (3.2) bepaald is, de* **k-dimensionale normaalverdeling** *met verwachting $\underline{\mu}$ en covariantiematrix Σ.*

De reden voor deze naam wordt in hoofdstuk 4 gegeven.

Via de lineaire algebra kan men tonen dat voor elke positief definiete, symmetrische matrix Σ een unieke positief definiete, symmetrische matrix A bestaat zodanig dat $A^2 = \Sigma$. Dit betekent dat we elke k-dimensionale normale verdeling met Σ positief definiet via (3.1) als de verdeling van een affiene transformatie van k onafhankelijke standaard normaal verdeelde toevalsvariabelen kunnen verkrijgen. Daarvan bestaat ook een omkering (zie Lemma 3.1 beneden).

De volgende stelling toont algemeen dat toevalsvectoren die we via een affiene transformatie (met maximaalrang) van normaalverdeelde toevalsvectoren verkrijgen, ook normaalverdeeld zijn.

Stelling 3.7 *Zij $(Y_1, \ldots, Y_k)$ normaal($\underline{\mu}, \Sigma$)-verdeeld, waar Σ positief definiet is. Als B een (k,k)-matrix is met $\det(B) \neq 0$ en $\underline{c} \in \mathbb{R}^k$, dan is*

$$\vec{X} = B\vec{Y} + \underline{c}$$

normaal($B\underline{\mu} + \underline{c}, B\Sigma B^t$)-verdeeld.

Bewijs oefening.

Het volgende lemma is een k-dimensionale versie van Lemma 2.1 en het volgt direct uit stelling 3.7.

Lemma 3.1 *Zij $(Y_1, \ldots, Y_k)$ normaal($\underline{\mu}, \Sigma$)-verdeeld, waar Σ positief definiet is. Stel*

$$(Z_1, \ldots, Z_k)^t = A^{-1}(Y_1 - \mu_1, \ldots, Y_k - \mu_k)^t,$$

waar A de symmetrische positief definitie (k,k)-matrix is zodanig dat $A^2 = \Sigma$.
Dan zijn de toevalsvariabelen $Z_1, \ldots, Z_k$ onafhankelijk en standaard-normaalverdeeld.

Een belangrijke eigenschap van de k-dimensionale normaalverdeling is dat als een toevalsvector zo'n verdeling heeft, ook de marginale verdelingen normaal zijn.

Lemma 3.2 *Zij $(Y_1, \ldots, Y_k)$ normaal($\underline{\mu}, \Sigma$)-verdeeld, waar Σ positief definiet is. Dan zijn de toevalsvariabelen Y_i normaal($\mu_i, \Sigma_{i,i}$) verdeeld, $1 \le i \le k$.*

Bewijs Zonder beperking van de algemeenheid kunnen we veronderstellen dat

$$(Y_1, \ldots, Y_k)^t = A(Z_1, \ldots, Z_k)^t + (\mu_1, \ldots, \mu_k)^t,$$

waar A de symmetrische, positief definitie matrix met $A^2 = \Sigma$ en $Z_1, \ldots, Z_k$ onafhankelijk en standaard-normaalverdeeld zijn.
Dan kunnen we via opmerking 2 op het einde van hoofdstuk 3.4 concluderen dat $Y_i = \sum_{j=1}^k A_{i,j} Z_j + \mu_i$ normaal(μ_i, σ_i^2)-verdeeld is met $\sigma_i^2 = \sum_{j=1}^k A_{i,j}^2 = \sum_{j=1}^k A_{i,j} A_{j,i} = (A^2)_{i,i} = \Sigma_{i,i}, 1 \le i \le k$, waarmee het lemma bewezen is. □

We bekijken nog een keer het voorbeeld uit hoofdstuk 3.3, waar (X, Y) een toevalsvector was met dichtheidsfunctie $(x, y) \to f(x, y) = e^{-y^2-xy-x^2/2}/(2\pi)$. Dit is de normaal$(0, \Sigma)$-verdeling met $\Sigma = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$ en we zien nu onmiddellijk dat X ~normaal$(0, 2)$ en Y ~normaal$(0, 1)$.

De volgende stelling geeft een zeer eenvoudige karakterisatie van onafhankelijkheid voor toevalsvariabelen $Y_1, \ldots, Y_k$ met een gezamenlijke k-dimensionale normaalverdeling.

Stelling 3.8 *Zij $(Y_1, \ldots, Y_k)$ normaal$(\underline{\mu}, \Sigma)$-verdeeld, waar Σ positief definiet is. Dan zijn equivalent:*
(i) *$Y_1, \ldots, Y_k$ zijn onafhankelijk.*
(ii) *Σ is een diagonaalmatrix.*

Bewijs $\boxed{\text{(i)} \Rightarrow \text{(ii)}}$ Als de toevalsvariabelen $Y_1, \ldots, Y_k$ onafhankelijk zijn, volgt uit lemma 3.2 dat hun gezamenlijke dichtheidsfunctie $f_{\vec{Y}}(\underline{x}), \underline{x} \in \mathbb{R}^k$ gelijk is aan

$$\prod_{i=1}^{k} \frac{1}{\sqrt{2\pi\Sigma_{i,i}}} \exp(-(x_i - \mu_i)^2/(2\Sigma_{i,i})) = \frac{1}{\sqrt{2\pi}^k \sqrt{\det(\Gamma)}} \exp\left(-\frac{1}{2}\langle \underline{x} - \underline{\mu}, \Gamma^{-1}(\underline{x} - \underline{\mu})\rangle\right), \underline{x} \in \mathbb{R}^k,$$

waar Γ de diagonaalmatrix met $\Gamma_{i,i} = \Sigma_{i,i}, 1 \le i \le k$ is.
Anderzijds weten we ook dat

$$f_{\vec{Y}}(\underline{x}) = \frac{1}{\sqrt{2\pi}^k \sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}\langle \underline{x} - \underline{\mu}, \Sigma^{-1}(\underline{x} - \underline{\mu})\rangle\right), \underline{x} \in \mathbb{R}^k.$$

Kiezen we $\underline{x} = \underline{\mu}$, volgt dat $\det(\Gamma) = \det(\Sigma)$, en we kunnen concluderen dat

$$\langle \underline{y}, \Gamma^{-1}\underline{y}\rangle = \langle \underline{y}, \Sigma^{-1}\underline{y}\rangle, \forall \underline{y} \in \mathbb{R}^k.$$

Dit impliceert dat $\Gamma^{-1} = \Sigma^{-1}$ en dan natuurlijk $\Gamma = \Sigma$. Dus is Σ een diagonaalmatrix.

$\boxed{\text{(ii)} \Rightarrow \text{(i)}}$ Als Σ diagonaal is, volgt direct uit de definitie van de k-dimensionale dichtheidsfunctie van $\vec{Y}$ dat

$$f_{\vec{Y}}(\underline{x}) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi\Sigma_{i,i}}} \exp(-(x_i - \mu_i)^2/(2\Sigma_{i,i})) = \prod_{i=1}^{k} f_{Y_i}(x_i), \underline{x} \in \mathbb{R}^k.$$

Stelling 3.6(iii) impliceert dan dat de toevalsvariabelen $Y_1, \ldots, Y_k$ onafhankelijk zijn. $\square$

De volgende stelling is belangrijk voor de statistiek.

Stelling 3.9 *Zij $n \ge 2$. Stel $Y_1, \ldots, Y_n$ onafhankelijke normaal(μ, σ^2)-verdeelde toevalsvariabelen. Dan zijn $\overline{Y} = \frac{1}{n}\sum_{i=1}^n Y_i$ en $S^2 = \frac{1}{n-1}\sum_{i=1}^n (Y_i - \overline{Y})^2$ onafhankelijke toevalsvariabelen.*

Bewijs Zonder beperking van de algemeenheid kunnen we veronderstellen dat

$$Y_j = \sigma Z_j + \mu, 1 \le j \le n$$

waar $Z_j, 1 \le j \le n$ onafhankelijke standaard-normaalverdeelde toevalsvariabelen zijn.
Dan geldt $\overline{Y} = \mu + \sigma\overline{Z}$ en $S^2 = \sigma^2 S_z^2$, waar $S_z^2 = \frac{1}{n-1}\sum_{i=1}^n (Z_i - \overline{Z})^2$, waaruit blijkt dat het

voldoende is de stelling voor standaard-normaalverdeelde toevalsvariabelen te bewijzen.

Gezien $\|\underline{x}\| = (\sum_{i=1}^n x_i^2)^{1/2}, \underline{x} \in \mathbb{R}^n$ de Euklidnorm is, geldt voor elke orthogonale (n,n)-matrix A (dus, A is inverteerbaar met $A^{-1} = A^t$):

$$\|A\underline{z}\| = \|\underline{z}\|, \underline{z} \in \mathbb{R}^n. \tag{3.3}$$

Zij nu A een orthogonale (n,n)-matrix waar de eerste rij gelijk is aan $(1/\sqrt{n}, \ldots, 1/\sqrt{n})$. Zo'n matrix bestaat. (oef., lin. algebra. *Hint*: kies een orthonormale basis van $\mathbb{R}^n$ die de bovenstaande vector bevat.) Stel:

$$\begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} = A \begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix}$$

Wegens onafhankelijkheid is $(Z_1, \ldots, Z_n)^t$ normaal$(0, I)$-verdeeld, waar I de n-dimensionale identiteitsmatrix is. Stelling 3.7 impliceert dan dat $(\xi_1, \ldots, \xi_n)^t$ normaal$(0, AIA^t)$-verdeeld en dus normaal$(0, I)$-verdeeld is. Gezien I diagonaal is, volgt via stelling 3.8 dat de toevalsvariabelen $\xi_1, \ldots, \xi_n$ onafhankelijk en standaard-normaalverdeeld zijn.
Verder impliceert betrekking (3.3) dat

$$\sum_{i=1}^n \xi_i^2 = \sum_{i=1}^n Z_i^2 \tag{3.4}$$

We kunnen concluderen dat

$$\frac{1}{n-1}\sum_{i=2}^n \xi_i^2 = \frac{1}{n-1}(-\xi_1^2 + \sum_{i=1}^n Z_i^2) = \frac{1}{n-1}(-n\overline{Z}^2 + \sum_{i=1}^n Z_i^2) = \frac{1}{n-1}\sum_{i=1}^n (Z_i - \overline{Z})^2 = S_z^2.$$

Uit de onafhankelijkheid van de toevalsvariabelen $\xi_1, \ldots, \xi_n$ volgt dan:

$$\overline{Z} = \xi_1/\sqrt{n} \text{ en } S_z^2 = \frac{1}{n-1}\sum_{i=2}^n \xi_i^2 \text{ zijn onafhankelijk,} \tag{3.5}$$

waarmee de stelling bewezen is. □

3.6 Sommen van onafhankelijke toevalsvariabelen

Probleem Gegeven X, Y onafhankelijke toevalsvariabelen waarvan de verdelingen gekend zijn. Vind de verdeling van $X + Y$.

We bekijken eerst het discrete geval.

Stelling 3.10 *Als X, Y onafhankelijke, discrete toevalsvariabelen zijn, dan is $X + Y$ ook discreet en als p_X, p_Y, p_{X+Y} de kansfuncties van $X, Y, X + Y$ zijn, geldt:*

$$p_{X+Y}(z) = \sum_x p_X(x)p_Y(z-x) = \sum_y p_X(z-y)p_Y(y).$$

Bewijs Het is triviaal dat $X + Y$ discreet is. Verder is $S_1 = \{x : p_X(x) > 0\}$ ten hoogste aftelbaar en het volgt dat

$$\begin{aligned} p_{X+Y}(z) &= \mathbb{P}(\{X + Y = z\} \cap \bigcup_{x \in S_1} \{X = x\}) \\ &= \mathbb{P}(\bigcup_{x \in S_1} \{X = x, Y = z - x\}) = \sum_{x \in S_1} p_X(x) p_Y(z - x), \end{aligned}$$

waar we in de laatste stap de onafhankelijkheid van X en Y hebben gebruikt.
Een modificatie van het bovenstaande bewijs toont dat ook de tweede formule voor p_{X+Y} correct is. □

Voorbeeld Als X, Y onafhankelijke Poisson-variabelen met parameters λ_1 en λ_2 zijn, is $Z = X + Y$ Poisson-verdeeld met parameter $\lambda_1 + \lambda_2$.

Bewijs Uit de vorige stelling volgt onmiddellijk voor $j = 0, 1, 2, \dots$

$$p_Z(j) = \sum_{i=0}^{\infty} \mathbb{P}\{X = i\} \mathbb{P}\{Y = j - i\} = \sum_{i=0}^{j} \frac{\lambda_1^i}{i!} \frac{\lambda_2^{j-i}}{(j - i)!} e^{-\lambda_1 - \lambda_2},$$

waar de laatste som wegens de binomiale stelling gelijk aan $(\lambda_1 + \lambda_2)^j e^{-\lambda_1 - \lambda_2}/j!$ is. □

In het continue geval bestaat er een analoog resultaat waar het natuurlijk niet meer volledig triviaal is dat een som van onafhankelijke toevalsvariabelen met absoluut continue verdelingen ook zo'n verdeling heeft.

Stelling 3.11 *Stel X, Y onafhankelijke toevalsvariabelen met absoluut continue verdelingen. Duid de dichtheidsfuncties van X en Y met f_X en f_Y aan. Als $Z = X + Y$ is, heeft Z een absoluut continue verdeling met dichtheidsfunctie*

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) dy, z \in \mathbb{R}$$

Bewijs We bekijken de verdelingsfunctie van Z. Voor elke $t \in \mathbb{R}$ geldt:

$$F_Z(t) = \mathbb{P}\{Z \le t\} = \mathbb{P}\{(X, Y) \in M_t\},$$

waar $M_t = \{(x, y) \in \mathbb{R}^2 : x + y \le t\}$.Vermits de gezamenlijke dichtheidsfunctie van X en Y gelijk aan $f_X(x) f_Y(y)$ is, zien we dat

$$F_Z(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{t-x} f_Y(y) dy f_X(x) dx = \int_{-\infty}^{\infty} \int_{-\infty}^{t} f_Y(z - x) dz f_X(x) dx,$$

waar we in de y-integraal $y = z - x$ hebben gesubstitueerd. Verwisselen we nu de volgorde van de twee integralen, volgt dat

$$F_Z(t) = \int_{-\infty}^{t} \int_{-\infty}^{\infty} f_Y(z - x) f_X(x) dx dz =: \int_{-\infty}^{t} g(z) dz.$$

Dus is de verdelingsfunctie van Z gelijk aan deze van de kansmaat met dichtheidsfunctie g wat weer impliceert dat de verdeling $\mathbb{P}_Z$ van Z gelijk aan deze kansmaat is.
De tweede formule voor f_Z volgt analoog. □

Voorbeelden

(1) Als X en Y onafhankelijke toevalsvariabelen zijn zodanig dat $\mathbb{P}_X$ = gamma(α_1, β) en $\mathbb{P}_Y$ = gamma(α_2, β), heeft $Z = X + Y$ een gamma$(\alpha_1 + \alpha_2, \beta)$-verdeling.

Dit impliceert in verband met stelling 2.9 dat als $Y_1, \dots, Y_n$ onafhankelijke toevalsvariabelen met standaard-normaalverdelingen zijn, de toevalsvariabele $\sum_{j=1}^n Y_j^2$ een gamma($n/2, 2$)-verdeling heeft. Men noemt deze verdeling ook de **chi-kwadraat-verdeling** *met n vrijheidsgraden.*

Bewijs Uit stelling 3.11 volgt onmiddellijk dat

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx \\ &= c_1 \int_0^z x^{\alpha_1 - 1} e^{-x/\beta} (z-x)^{\alpha_2 - 1} e^{-(z-x)/\beta} dx \\ &= c_1 z^{\alpha_1 + \alpha_2 - 1} e^{-z/\beta} \int_0^z z^{-1} (x/z)^{\alpha_1 - 1} (1 - x/z)^{\alpha_2 - 1} dx \\ &= c_2 z^{\alpha_1 + \alpha_2 - 1} e^{-z/\beta}, z > 0, \end{aligned}$$

waar de constanten c_1 en $c_2 = c_1 \int_0^1 u^{\alpha_1 - 1}(1-u)^{\alpha_2 - 1} du$ afhankelijk van α_1, α_2 en β zijn. (We substitueren in de laatste integraal $x = zu$.)

We zien dat $f_Z = cg$ waar $c > 0$ een constante en g de dichtheidsfunctie van de gamma$(\alpha_1 + \alpha_2, \beta)$-verdeling is. Dit impliceert dat $c = 1$ en het bewijs is compleet. □

(2) Als X, Y onafhankelijke toevalsvariabelen met $\mathbb{P}_X$ = normaal(μ_1, σ_1^2) en $\mathbb{P}_Y$ = normaal(μ_2, σ_2^2) zijn, hebben we: $\mathbb{P}_{X+Y}$ = normaal$(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Dit is een speciaal geval van opmerking 2 op het einde van hoofdstuk 3.4. We kunnen dit ook alternatief via stelling 3.11 bewijzen.

Definitie 3.6 *Stel Z, W onafhankelijke toevalsvariabelen zodanig dat $Z \sim$ normaal $(0,1)$ en $W \sim$ chi-kwadraat(m). Beschouw $T = \frac{Z}{\sqrt{W/m}}$. Dan noemen we de verdeling van de toevalsvariabele T de* **student-t-verdeling** *met m vrijheidsgraden.*

We kunnen nu de volgende stelling bewijzen:

Stelling 3.12 *Zij $n \geq 2$. Stel $Y_1, \dots, Y_n$ onafhankelijke normaal(μ, σ^2)-verdeelde toevalsvariabelen, $\overline{Y} = \sum_{j=1}^n Y_j/n, S^2 = \sum_{j=1}^n (Y_j - \overline{Y})^2/(n-1)$.*
Dan heeft de toevalsvariabele $T := \sqrt{n}(\overline{Y} - \mu)/S$ een t-verdeling met $(n-1)$ vrijheidsgraden.

Bewijs We veronderstellen weer dat $Y_i = \sigma Z_i + \mu, 1 \leq i \leq n$, waar $Z_1, \dots, Z_n$ onafhankelijk en standaard-normaalverdeeld zijn. Dan geldt er natuurlijk dat $T = \sqrt{n}\overline{Z}/S_z$, waar $S_z^2 = \sum_{i=1}^n (Z_i - \overline{Z})^2/(n-1)$.

Stelling 3.9 impliceert dat de twee toevalsvariabelen $\overline{Z}$ en S_z^2 onafhankelijk zijn. Verder geldt er dat $\sqrt{n}\overline{Z} = \xi_1$ en $S_z^2 = \sum_{i=2}^n \xi_i^2/(n-1)$, waar de toevalsvariabelen $\xi_i, 1 \leq i \leq n$ onafhankelijk en standaard-normaalverdeeld zijn.

Bovendien kunnen we uit stelling 2.9 afleiden dat de toevalsvariabelen ξ_i^2 gamma$(1/2, 2)$-verdeeld zijn, $2 \leq i \leq n$. Vermits deze variabelen ook onafhankelijk zijn volgt uit de convolutieformule voor gammaverdeelde variabelen dat $(n-1)S_z^2 = \sum_{i=2}^n \xi_i^2 \sim$ gamma$((n-1)/2, 2)$= chi-kwadraat$(n-1)$ en de stelling is bewezen. □

Uit het volgende lemma blijkt dat de t-verdeling absoluut continu is en er bestaat een expliciete formule voor de dichtheidsfunctie.

Lemma 3.3 *Stel Y, W onafhankelijke toevalsvariabelen zodanig dat $Y \sim$ normaal $(0,1)$ en $W \sim$ chi-kwadraat(m). Dan heeft $T = \frac{Y}{\sqrt{W/m}}$ een absoluut continue verdeling met dichtheidsfunctie*

$$t \to \frac{\Gamma((m+1)/2)}{\sqrt{m\pi}\Gamma(m/2)}(1+t^2/m)^{-(m+1)/2}$$

Bewijs Vermits W de dichtheidsfunctie $f_W(w) = (w/2)^{(m-2)/2}e^{-w/2}/(2\Gamma(m/2)), w \geq 0$ heeft, kunnen we via stelling 2.10 concluderen dat $Z = \sqrt{W/2}$ de dichtheidsfunctie

$$f_Z(z) = 2z^{m-1}e^{-z^2}/\Gamma(m/2), z > 0$$

Verder heeft $X = Y/\sqrt{2}$ de dichtheidsfunctie $f_X(x) = e^{-x^2}/\sqrt{\pi}, x \in \mathbb{R}$. Uit de onafhankelijkheid van de twee variabelen X en Z volgt dat ook X/Z een absoluut continue verdeling heeft waar de dichtheidsfunctie $g = f_{X/Z}$ door de volgende (algemene) formule bepaald is:

$$g(u) = \int_0^\infty z f_X(uz) f_Z(z) dz, u \in \mathbb{R}.$$

(Dit kan men via een modificatie van het argument bewijzen wat we in het bewijs van de convolutieformule uit stelling 3.11 hebben gebruikt.) Dus geldt in ons geval:

$$\begin{aligned} g(u) &= 2\int_0^\infty z^m e^{-z^2(1+u^2)} dz/(\Gamma(m/2)\sqrt{\pi}) \\ &= (1+u^2)^{-(m+1)/2}\int_0^\infty y^{(m-1)/2}e^{-y}dy/(\Gamma(m/2)\sqrt{\pi}) \text{ (subst. } y = z^2(1+u^2)) \\ &= \frac{\Gamma((m+1)/2)}{\Gamma(m/2)\sqrt{\pi}}(1+u^2)^{-(m+1)/2}, u \in \mathbb{R} \end{aligned}$$

Gezien $T = m^{-1/2}X/Y$, is het evident dat T de dichtheidsfunctie $g(t/\sqrt{m})/\sqrt{m}, t \in \mathbb{R}$ heeft, waarmee het lemma bewezen is. □

3.7 Conditionele verdelingen

Het doel van dit laatste gedeelte van hoofdstuk 3 is het de conditionele (of voorwaardelijke) verdeling van een toevalsvector $X : \Omega \to \mathbb{R}^{d_1}$ gegeven dat een andere d_2-dimensionale toevalsvector Y gelijk aan een bepaalde waarde $y \in \mathbb{R}^{d_2}$ is te definiëren. Als X en Y discreet zijn, is dit zonder meer mogelijk via voorwaardelijke kansen zoals in hoofdstuk 1.

Dus onderstel dat X, Y discrete toevalsvectoren met gezamenlijke kansfunctie $p(x,y)$ zijn en $y \in \mathbb{R}^{d_2}$ zo gekozen is dat geldt $p_Y(y) = \mathbb{P}\{Y = y\} > 0$. Dan hebben we voor $A \subset \mathbb{R}^{d_1}$:

$$\mathbb{P}(\{X \in A\}|\{Y = y\}) = \mathbb{P}\{X \in A, Y = y\}/\mathbb{P}\{Y = y\} = \sum_{x \in A} p(x,y)/p_Y(y) =: \sum_{x \in A} p_{X|Y}(x|y)$$

Het is evident dat $x \to p_{X|Y}(x|y)$ ook een kansfunctie is. (Het linkerlid is gelijk aan 1 als we $A = S$ stellen waar $S = \{x : \mathbb{P}\{X = x\} > 0$ hoogstens aftelbaar is.)

Definitie 3.7 *Gegeven twee discrete toevalsvectoren $X : \Omega \to \mathbb{R}^{d_1}$ en $Y : \Omega \to \mathbb{R}^{d_2}$ met respectieve kansfuncties p_X en p_Y, definiëren we de* **conditionele kansfunctie** *van X gegeven $Y = y$, waar $p_Y(y) > 0$, als*

$$p_{X|Y}(x|y) = \mathbb{P}(\{X = x\}|\{Y = y\}), x \in \mathbb{R}^{d_1}$$

We noemen de door $p_{X|Y}(\cdot|y)$ bepaalde discrete kansmaat de conditionele verdeling van X gegeven $Y = y$.

Voorbeeld Beschouw twee onafhankelijke binomiaalverdeelde toevalsvariabelen X_1, X_2 met respectieve parameters n_1, p en n_2, p. Stel $X = X_1$ en $Y = X_1 + X_2$. Vind de conditionele verdeling van X gegeven $Y = m$.

Oplossing We noteren eerst dat Y binomiaal$(n_1 + n_2, p)$-verdeeld is (oef.)
Verder geldt voor $0 \le m_1 \le m \le n_1 + n_2$:

$$\begin{aligned}
\mathbb{P}\{X = m_1, Y = m\} &= \mathbb{P}\{X_1 = m_1, X_2 = m - m_1\} \\
&= \binom{n_1}{m_1} p^{m_1}(1-p)^{n_1-m_1} \binom{n_2}{m-m_1} p^{m-m_1}(1-p)^{n_2-m+m_1} \\
&= \binom{n_1}{m_1}\binom{n_2}{m-m_1} p^m (1-p)^{n_1+n_2-m}
\end{aligned}$$

Vermits geldt $p_Y(m) = \binom{n_1+n_2}{m} p^m(1-p)^{n_1+n_2-m}$ volgt dat

$$p_{X|Y}(m_1|m) = \frac{\binom{n_1}{m_1}\binom{n_2}{m-m_1}}{\binom{n_1+n_2}{m}}, \ m_1 = 0, \dots, m.$$

Dus is deze conditionele verdeling gelijk aan de hypergeometrische verdeling.

Als de verdeling van $(X, Y) : \Omega \to \mathbb{R}^{d_1+d_2}$ een dichtheidsfunctie $f : \mathbb{R}^{d_1+d_2} \to [0, \infty[$ heeft, volgt zoals in stelling 3.4 dat ook X en Y dichtheidsfuncties hebben. Als f_Y de dichtheidsfunctie van Y is, hebben we $f_Y(y) = \int_{\mathbb{R}^{d_1}} f(x, y)dx$ (dus een d_1-dimensionale integraal).
We definiëren dan conditionele verdelingen zoals in het discrete geval, waar we gewoon de kansfuncties door de dichtheidsfuncties vervangen.

Definitie 3.8 *Als $X : \Omega \to \mathbb{R}^{d_1}$ en $Y : \Omega \to \mathbb{R}^{d_2}$ toevalsvectoren zijn met een (gezamenlijke) dichtheidsfunctie $(x, y) \to f(x, y)$, en $y \in \mathbb{R}^{d_2}$ zodanig is dat $f_Y(y) > 0$, is de* **conditionele dichtheidsfunctie** *van X gegeven $Y = y$ gedefinieerd door*

$$f_{X|Y}(x|y) = f(x, y)/f_Y(y), x \in \mathbb{R}$$

De door $f_{X|Y}(\cdot|y)$ bepaalde kansmaat heet de **conditionele verdeling** *van X gegeven $Y = y$.*

Voorbeelden (1) Zij $(X, Y) : \Omega \to \mathbb{R}^2$ een toevalsvector met gezamenlijke dichtheidsfunctie

$$f(x, y) = y^{-1}e^{-x/y}e^{-y}, x, y > 0 \text{ en } f(x, y) = 0 \text{ elders}$$

Bepaal $\mathbb{P}(\{X > 1\}|\{Y = y\}), y > 0$.

Oplossing We bepalen eerst de conditionele verdeling van X gegeven $Y = y$.
De marginale dichtheidsfunctie van Y is gelijk aan $f_Y(y) = \int_0^\infty y^{-1}e^{-x/y}e^{-y}dx = e^{-y}, y > 0$. (Dit is de exponentiële verdeling met parameter 1.)
Als we de formule voor de conditionele dichtheidsfunctie gebruiken, volgt onmiddellijk dat

$$f_{X|Y}(x|y) = y^{-1}e^{-x/y}, x > 0.$$

Dit is de exponentiële verdeling met parameter y en het volgt dat

$$\mathbb{P}(\{X > 1\}|\{Y = y\}) = \int_1^\infty y^{-1}e^{-x/y}dx = e^{-1/y}.$$

(2) Beschouw nu een 2-dimensionale toevalsvector (X, Y) met een normaal$(0, \Sigma)$-verdeling. Wat is dan de conditionele verdeling van Y gegeven $X = x$ voor $x \in \mathbb{R}$?

Oplossing Gezien de matrix Σ symmetrisch en positief definiet is, geldt er $\Sigma = \begin{bmatrix} a^2 & ab\rho \\ ab\rho & b^2 \end{bmatrix}$, waar $a, b > 0$ en $|\rho| < 1$. Voor de inverse matrix hebben we dan

$$\Sigma^{-1} = \frac{1}{a^2b^2(1-\rho^2)}\begin{bmatrix} b^2 & -ab\rho \\ -ab\rho & a^2 \end{bmatrix}$$

Daarmee kunnen we de dichtheidsfunctie f van (X, Y) opschrijven zoals volgt,

$$f(x, y) = \frac{1}{2\pi ab\sqrt{1-\rho^2}} \exp(-(b^2x^2 + a^2y^2 - 2ab\rho xy)/(2a^2b^2(1-\rho^2)))$$

Uit lemma 3.2 volgt er verder dat $f_X(x) = \exp(-x^2/(2a^2))/(\sqrt{2\pi}a), x \in \mathbb{R}$.
Bijgevolg, hebben we

$$\begin{aligned} f_{Y|X}(y|x) &= \frac{1}{\sqrt{2\pi}b\sqrt{1-\rho^2}} \exp(x^2/(2a^2) - (b^2x^2 + a^2y^2 - 2ab\rho xy)/(2a^2b^2(1-\rho^2))) \\ &= \frac{1}{\sqrt{2\pi}b\sqrt{1-\rho^2}} \exp(-(\rho^2b^2x^2 + a^2y^2 - 2ab\rho xy)/(2a^2b^2(1-\rho^2))) \\ &= \frac{1}{\sqrt{2\pi}b\sqrt{1-\rho^2}} \exp(-(y - \rho(b/a)x)^2/(2b^2(1-\rho^2))) \end{aligned}$$

Dit is de 1-dimensionale normaal$(\rho bx/a, b^2(1-\rho^2))$-verdeling die verschillend is van de marginale normaal$(0, b^2)$-verdeling van Y (zie lemma 3.2) als $\rho \neq 0$.

Hoofdstuk 4

Verwachtingswaarden

4.1 Definitie en eigenschappen

We definiëren eerst verwachtingswaarden voor elementaire toevalsvariabelen, dan voor algemene positieve toevalsvariabelen en ten slotte voor semi-integreerbare toevalsvariabelen. (Verwachtingswaarden bestaan niet voor alle toevalsvariabelen.)

Definitie 4.1 *Zij $X : \Omega \to \mathbb{R}$ een elementaire toevalsvariabele met de verschillende waarden $x_1, \ldots, x_m$. Dan wordt de* **verwachtingswaarde** *van X (Notatie: $\mathbb{E}[X]$) gedefinieerd door*

$$\mathbb{E}[X] = \sum_{i=1}^{m} x_i \mathbb{P}\{X = x_i\}.$$

Voorbeelden

1. Zij $A \in \mathcal{F}$ een gebeurtenis met $\mathbb{P}(A) = p$ en $X = I_A$ een Bernoulli(p)-variabele. Dan geldt: $\mathbb{E}[X] = 0\mathbb{P}\{X = 0\} + 1\mathbb{P}\{X = 1\} = \mathbb{P}(A) = p$.

2. Zij nu X een binomiaal(n, p)-variabele. Dan geldt:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{j=0}^{n} j\mathbb{P}\{X = j\} = \sum_{j=1}^{n} j\binom{n}{j} p^j (1-p)^{n-j} \\
&= np \sum_{j=1}^{n} \frac{(n-1)!}{(n-j)!(j-1)!} p^{j-1}(1-p)^{n-j} \\
&= np \sum_{i=0}^{n-1} \binom{n-1}{i} p^i (1-p)^{n-1-i} = np(p + (1-p))^{n-1} = np.
\end{aligned}$$

   We kunnen deze verwachtingswaarde echter op een meer efficiënte manier berekenen. De toevalsvariabele X heeft een representatie $X = \sum_{i=1}^{n} I_{A_i}$, waar A_i onafhankelijke gebeurtenissen zijn met $\mathbb{P}(A_i) = p, 1 \le i \le n$. Uit de lineariteit van de verwachtingswaarde (deel (ii) van de volgende stelling) volgt er dan onmiddellijk dat $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[I_{A_i}] = np$.

Stelling 4.1

(i) *Zij* $A_i \in \mathcal{F}, 1 \leq i \leq n$ *een partitie van* Ω. *Dan geldt voor de elementaire toevalsvariabele* $X = \sum_{i=1}^n y_i I_{A_i}$: $\mathbb{E}[X] = \sum_{i=1}^n y_i \mathbb{P}(A_i)$.

(ii) *Als* X, Y *elementaire toevalsvariabelen zijn, dan hebben we voor* $\alpha, \beta \in \mathbb{R}$

$$\mathbb{E}[\alpha X + \beta Y] = \alpha \mathbb{E}[X] + \beta \mathbb{E}[Y].$$

(iii) *Als* X *elementair is en* $f : \mathbb{R} \to \mathbb{R}$ *een willekeurige afbeelding is, dan hebben we*

$$\mathbb{E}[f(X)] = \sum_x f(x)\mathbb{P}\{X = x\}.$$

(iv) *Als* $X \leq Y$ *elementaire toevalsvariabelen zijn, hebben we:* $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

(v) $X_1, \ldots, X_n$ *elementaire, onafhankelijke toevalsvariabelen* $\Rightarrow \mathbb{E}[\prod_{i=1}^n X_i] = \prod_{i=1}^n \mathbb{E}[X_i]$.

Bewijs (i) Duid de de verschillende waarden van X door $x_1, \ldots, x_m$ aan, waar $m \leq n$ moet zijn. Dan is het evident dat $y_i \in \{x_1, \ldots, x_m\}, 1 \leq i \leq n$ en we kunnen concluderen dat

$$\sum_{i=1}^n y_i \mathbb{P}(A_i) = \sum_{j=1}^m \sum_{i:y_i=x_j} y_i \mathbb{P}(A_i) = \sum_{j=1}^m x_j \mathbb{P}(\biguplus_{i:y_i=x_j} A_i) = \sum_{j=1}^m x_j \mathbb{P}\{X = x_j\} = \mathbb{E}[X].$$

(ii) Noteer de verschillende waarden van X (Y) als $x_i, 1 \leq i \leq m$ $(y_j, 1 \leq j \leq n)$. Stel

$$A_i = \{X = x_i\}, 1 \leq i \leq m \text{ en } B_j = \{Y = y_j\}, 1 \leq j \leq n.$$

Dan is $A_i \cap B_j, 1 \leq i \leq m, 1 \leq j \leq n$ een partitie van Ω en we kunnen uit deel (i) concluderen dat

$$\mathbb{E}[\alpha X + \beta Y] = \mathbb{E}[\sum_{i=1}^m \sum_{j=1}^n (\alpha x_i + \beta y_j) I_{A_i \cap B_j}] = \sum_{i=1}^m \sum_{j=1}^n (\alpha x_i + \beta y_j) \mathbb{P}(A_i \cap B_j).$$

Als we de laatste dubbelsom opsplitsen, zien we dat de bovenstaande verwachtingswaarde gelijk is aan

$$\alpha \sum_{i=1}^m x_i \sum_{j=1}^n \mathbb{P}(A_i \cap B_j) + \beta \sum_{j=1}^n y_j \sum_{i=1}^m \mathbb{P}(A_i \cap B_j),$$

hetgeen we wegens de additiviteit van de kansmaat $\mathbb{P}$ kunnen schrijven als

$$\alpha \sum_{i=1}^m x_i \mathbb{P}(A_i) + \beta \sum_{j=1}^n y_j \mathbb{P}(B_j) = \alpha \mathbb{E}[X] + \beta \mathbb{E}[Y].$$

(iii) Als $x_1, \ldots, x_m$ de verschillende waarden van X zijn, dan volgt: $f(X) = \sum_{i=1}^m f(x_i) I_{\{X=x_i\}}$ en deel (i) impliceert dat $\mathbb{E}[f(X)] = \sum_{i=1}^m f(x_i)\mathbb{P}\{X = x_i\}$.

(iv) Uit de definitie van de verwachtingswaarde volgt onmiddellijk dat voor elke positieve elementaire toevalsvariabele $Z \geq 0$ geldt: $\mathbb{E}[Z] \geq 0$.
Stel nu $Z = Y - X$. Dan volgt wegens (ii): $0 \leq \mathbb{E}[Z] = \mathbb{E}[Y] - \mathbb{E}[X]$ en dus, $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

(v) Schrijf $x_{i,j_i}, 1 \le j_i \le m_i$ voor de verschillende waarden van $X_i, 1 \le i \le n$ en beschouw de partitie $\bigcap_{i=1}^n \{X_i = x_{i,j_i}\}, 1 \le j_i \le m_i, 1 \le i \le n$ van Ω. Uit (ii) volgt dan weer:

$$\begin{aligned}
\mathbb{E}[\prod_{i=1}^n X_i] &= \sum_{j_1=1}^{m_1} \dots \sum_{j_n=1}^{m_n} (\prod_{i=1}^n x_{i,j_i}) \mathbb{P}(\bigcap_{i=1}^n \{X_i = x_{i,j_i}\}) \\
&= \sum_{j_1=1}^{m_1} \dots \sum_{j_n=1}^{m_n} \prod_{i=1}^n (x_{i,j_i} \mathbb{P}\{X_i = x_{i,j_i}\}) \text{ (wegens onafhankelijkheid)} \\
&= \prod_{i=1}^n (\sum_{j_i=1}^{m_i} x_{i,j_i} \mathbb{P}\{X_i = x_{i,j_i}\}) = \prod_{i=1}^n \mathbb{E}[X_i].
\end{aligned}$$

Daarmee is stelling 4.1 bewezen. □

Definitie 4.2 *Zij $X \ge 0$ een (niet noodzakelijk elementaire) toevalsvariabele. Dan definiëren we de verwachtingswaarde van X door*

$$\mathbb{E}[X] = \sup\{\mathbb{E}[Y] : 0 \le Y \le X,\ Y \text{ elementair}\} \in [0, \infty].$$

Opmerkingen

1. Als X elementair is, levert deze definitie dezelfde waarde voor $\mathbb{E}[X]$ als onze eerste definitie. (Dit volgt onmiddellijk uit deel (iv) van stelling 4.1.)

2. Verder volgt direct uit de definitie dat verwachtingswaarden monotoon zijn, dwz als $0 \le X_1 \le X_2$ algemene toevalsvariabelen zijn, geldt weer $\mathbb{E}[X_1] \le \mathbb{E}[X_2]$.

Stelling 4.2 (Stelling van de monotone convergentie)
Zij $0 \le X_1 \le X_2 \le \dots$ een monotone rij toevalsvariabelen. Stel $X = \lim_{n\to\infty} X_n$. Dan geldt:

$$\mathbb{E}[X] = \lim_{n\to\infty} \mathbb{E}[X_n].$$

Bewijs Het is triviaal (zie opmerking (2)) dat de rij $\mathbb{E}[X_n], n \ge 1$ monotoon niet-dalend is en beneden $\mathbb{E}[X]$ blijft. Dus bestaat $\lim_n \mathbb{E}[X_n]$ in het interval $[0, \mathbb{E}[X]]$ en het is voldoende aan te tonen dat er geldt: $\boxed{\lim_{n\to\infty} \mathbb{E}[X_n] \ge \mathbb{E}[X].}$
Om dit te bewijzen, kiezen we een willekeurige elementaire toevalsvariabele $0 \le Y \le X$.
Als $Y \ne 0$, geldt er $Y = \sum_{i=1}^m y_i I_{A_i}$, waar $y_1, \dots, y_m$ de verschillende waarden van Y zijn met $y_i \ne 0$, $A_i = \{Y = y_i\}, 1 \le i \le m$ en $m \ge 1$.
Zij $0 < \epsilon < 1$. Stel verder $A_{i,n} = \{X_n > y_i(1-\epsilon)\} \cap A_i, 1 \le i \le m$. Dan geldt natuurlijk: $X_n \ge \sum_{i=1}^m y_i(1-\epsilon) I_{A_{i,n}}$ en bijgevolg

$$\mathbb{E}[X_n] \ge (1-\epsilon)\mathbb{E}[\sum_{i=1}^m y_i I_{A_{i,n}}] \ge (1-\epsilon)\sum_{i=1}^m y_i \mathbb{P}(A_{i,n}), n \ge 1.$$

Wegens de monotonie van de rij $X_n, n \ge 1$ is $A_{i,n}, n \ge 1$ een stijgende rij van gebeurtenissen hetgeen via stelling 1.1 impliceert dat als $n \to \infty$ geldt: $\mathbb{P}(A_{i,n}) \nearrow \mathbb{P}(\bigcup_{n=1}^\infty A_{i,n}) = \mathbb{P}(\{\sup_{n\ge1} X_n > y_i(1-\epsilon)\} \cap A_i)$. Wegens $\sup_{n\ge1} X_n = X$ en $Y \le X$ geldt er dat

$$\{\sup_{n\ge1} X_n > y_i(1-\epsilon)\} \cap A_i = \{X > y_i(1-\epsilon)\} \cap A_i = A_i, 1 \le i \le m.$$

Dus: $\lim_{n\to\infty} I\!E[X_n] \geq (1-\epsilon)\sum_{i=1}^m y_i \mathbb{P}(A_i) = (1-\epsilon) I\!E[Y]$ en $\lim_n I\!E[X_n] \geq I\!E[Y]$ (we kunnen $\epsilon > 0$ willekeurig klein kiezen). Gezien deze ongelijkheid voor elke elementaire toevalsvariabele $0 \leq Y \leq X$ geldt (triviaal als $Y = 0$), volgt $\lim_n I\!E[X_n] \geq I\!E[X]$ en de stelling is bewezen. □

Als eerste toepassing van de stelling van de monotone convergentie tonen we nu hoe men de verwachtingswaarde van discrete positieve toevalsvariabelen algemeen via de kansfunctie kan berekenen.

Lemma 4.1 *Zij $X : \Omega \to [0,\infty[$ een discrete toevalsvariabele met kansfunctie p_X. Dan geldt er: $I\!E[X] = \sum_{x\geq 0} x p_X(x)$.*

Bewijs. Indien X elementair is, is dit de definitie van $I\!E[X]$. Dus veronderstel dat de drager $S = \{x : p_X(x) > 0\}$ oneindig en dus gelijk is aan $\{x_n : n \geq 1\}$ waar x_n een rij in $[0,\infty[$ is zodanig dat $x_m \neq x_n$ als $m \neq n$.
Stel

$$X_n = \sum_{i=1}^n x_i I_{\{X=x_i\}}, n \geq 1.$$

Dan geldt $X_n \nearrow X$ en bijgevolg

$$I\!E[X] = \lim_{n\to\infty} I\!E[X_n] = \lim_{n\to\infty}\left(\sum_{i=1}^n x_i \mathbb{P}\{X = x_i\}\right) = \sum_{i=1}^\infty x_i p_X(x_i),$$

waarmee het lemma bewezen is. □

Voorbeelden

1. Zij X Poisson(λ)-verdeeld. Dan geldt:

$$I\!E[X] = \sum_{k=0}^\infty kP\{X = k\} = \sum_{k=1}^\infty \lambda^k e^{-\lambda}/(k-1)! = \lambda.$$

2. $X \sim$ geometrisch(p) $\Rightarrow I\!E[X] = \sum_{k=1}^\infty k(1-p)^{k-1}p = pg'(1-p)$, waar $g(x) = \sum_{k=0}^\infty x^k = (1-x)^{-1}, |x| < 1$. Dus: $I\!E[X] = 1/p$.

3. $X \sim$ neg.-binomiaal(r,p) $\Rightarrow I\!E[X] = r/p$. (oef.)
   [*Hint:* $X = \sum_{i=1}^r X_i$, waar $X_i \sim$ geometrisch(p) ,$1 \leq i \leq r$.]

Het volgende lemma is cruciaal voor wat volgt. Het toont dat elke (niet noodzakelijk discrete) positieve toevalsvariabele een representatie heeft als een limiet van een monotone rij elementaire toevalsvariabelen.

Lemma 4.2 *Zij $X \geq 0$ een toevalsvariabele. Dan bestaat er een stijgende rij elementaire toevalsvariabelen $0 \leq X_1 \leq X_2 \leq \dots$ zodanig dat $X_n \to X$ als $n \to \infty$.*

Bewijs Stel $X_n := 2^{-n}[2^n X] \wedge n$, $n \geq 1$, waar $[x]$ het gehele deel van het positieve getal x is (i.e., $[x] = m, m \leq x < m+1, m = 0,1,\dots$).
Dan geldt $X_n = f_n \circ X$, waar $f_n = nI_{[n,\infty[} + \sum_{i=1}^{n2^n}(i-1)2^{-n}I_{[(i-1)2^{-n},i2^{-n}[}$ een Borel-meetbare afbeelding van $\mathbb{R}$ naar $\mathbb{R}$ is.
(Dit volgt direct uit stelling 2.3.iii in verband met het triviale feit dat indicatorfuncties van intervallen Borel-meetbaar zijn.)

Een toepassing van stelling 2.2.ii levert dan dat X_n $\mathcal{F}$-meetbaar is (omdat X als een toevalsvariabele $\mathcal{F}$-meetbaar is.) Gezien X_n maar de waarden $i2^{-n}, 0 \leq i \leq n2^n$ kan aannemen, hebben we een rij elementaire toevalsvariabelen.
Het is evident dat deze rij niet-dalend is en puntsgewijs naar X convergeert. □

Lemma 4.2 in combinatie met de stelling van de monotone convergentie maakt het nu mogelijk de resultaten die we al voor verwachtingswaarden van elementaire toevalsvariabelen hebben verkregen, voor algemene positieve toevalsvariabelen te bewijzen.

Lemma 4.3
(i) *Als* $X, Y \geq 0$ *toevalsvariabelen zijn en* $\alpha, \beta \geq 0$, *geldt er:*
$$\mathbb{E}[\alpha X + \beta Y] = \alpha \mathbb{E}[X] + \beta \mathbb{E}[Y].$$
(ii) *Als* $X_1, \ldots, X_n \geq 0$ *onafhankelijke toevalsvariabelen zijn, hebben we:*
$$\mathbb{E}[\textstyle\prod_{i=1}^n X_i] = \prod_{i=1}^n \mathbb{E}[X_i],$$
waarbij $\infty \cdot a = a \cdot \infty = \infty, a > 0, \infty \cdot 0 = 0.$

Bewijs (i) Kies elementaire toevalsvariabelen $X_n \nearrow X$ en $Y_n \nearrow Y$.
Dan geldt natuurlijk: $\alpha X_n + \beta Y_n \nearrow \alpha X + \beta Y$ hetgeen via de stelling van de monotone convergentie en stelling 4.1.ii impliceert

$$\mathbb{E}[\alpha X + \beta Y] = \lim_{n\to\infty} \mathbb{E}[\alpha X_n + \beta Y_n] = \alpha \lim_{n\to\infty} \mathbb{E}[X_n] + \beta \lim_{n\to\infty} \mathbb{E}[Y_n]$$

Een tweede toepassing van de stelling van de monotone convergentie toont dan dat het rechterlid gelijk aan $\alpha\mathbb{E}[X] + \beta\mathbb{E}[Y]$ is.

(ii) Stel $X_i^{(m)} = f_m \circ X_i, 1 \leq i \leq n, m \geq 1$, waar f_m zoals in het bewijs van lemma 4.1 gedefinieerd is. Dan geldt $X_i^{(m)} \nearrow X_i$ als $m \to \infty$, $1 \leq i \leq n$ en bijgevolg
$$\textstyle\prod_{i=1}^n X_i^{(m)} \nearrow \prod_{i=1}^n X_i.$$
Verder zijn voor elke vaste $m \geq 1$ de elementaire toevalsvariabelen $X_i^{(m)}, 1 \leq i \leq n$ als functies van de onafhankelijke toevalsvariabelen $X_i, 1 \leq i \leq n$ zelf onafhankelijk zodat we via monotone convergentie en stelling 4.1.v kunnen concluderen dat

$$\mathbb{E}[\prod_{i=1}^n X_i] = \lim_{m\to\infty} \mathbb{E}[\prod_{i=1}^n X_i^{(m)}] = \prod_{i=1}^n (\lim_{m\to\infty} \mathbb{E}[X_i^{(m)}])$$

Een verdere toepassing van de stelling van de monotone convergentie toont dan dat het rechterlid gelijk aan $\prod_{i=1}^n \mathbb{E}[X_i]$ is en het lemma is bewezen. □

We gaan nu een zeer belangrijke representatie van de verwachtingswaarde van algemene positieve toevalsvariabelen als een bepaalde Riemann-integraal bewijzen. Via deze kunnen we dan ook een formule voor $\mathbb{E}[X]$ vinden als X absoluut continu verdeeld is.

Stelling 4.3 *Zij* $X : \Omega \to [0, \infty[$ *een toevalsvariabele. Dan geldt:*

$$\mathbb{E}[X] = \int_0^\infty \mathbb{P}\{X \geq x\} dx.$$

Opmerking. Wegens de monotonie van $x \to \mathbb{P}\{X \geq x\}$ bestaat voor elke $t > 0$ de Riemann-integraal $I(t) = \int_0^t \mathbb{P}\{X \geq x\}dx$ en we definiëren $\int_0^\infty \mathbb{P}\{X \geq x\}dx = \lim_{t\to\infty} I(t) \in [0, \infty]$.

Bewijs. Zij de monotone rij $X_n, n \geq 1$ zoals in lemma 4.2 gedefinieerd, i.e. $X_n = 2^{-n}[2^n X] \wedge n$. Dan geldt:

$$\mathbb{E}[X] = \lim_{n\to\infty} \mathbb{E}[X_n] = \lim_{n\to\infty} a_n$$

waar

$$\begin{aligned}
a_n &= n\mathbb{P}\{X \geq n\} + \sum_{k=2}^{n2^n} \frac{k-1}{2^n}\mathbb{P}\left\{\frac{k-1}{2^n} \leq X < \frac{k}{2^n}\right\} \\
&= n\mathbb{P}\{X \geq n\} + \sum_{k=2}^{n2^n} \frac{k-1}{2^n}\mathbb{P}\left\{X \geq \frac{k-1}{2^n}\right\} - \sum_{k=1}^{n2^n} \frac{k-1}{2^n}\mathbb{P}\left\{X \geq \frac{k}{2^n}\right\} \\
&= n\mathbb{P}\{X \geq n\} + \sum_{k=1}^{n2^n-1} 2^{-n}\mathbb{P}\left\{X \geq \frac{k}{2^n}\right\} - (n - 2^{-n})\mathbb{P}\{X \geq n\} \\
&= \sum_{k=1}^{n2^n} 2^{-n}\mathbb{P}\left\{X \geq \frac{k}{2^n}\right\} \leq \sum_{k=1}^{n2^n} \int_{(k-1)/2^n}^{k/2^n} \mathbb{P}\{X \geq t\}dt = \int_0^n \mathbb{P}\{X \geq t\}dt.
\end{aligned}$$

Analoog volgt,

$$a_n \geq \sum_{k=1}^{n2^n} \int_{k/2^n}^{(k+1)/2^n} \mathbb{P}\{X \geq t\}dt \geq \int_0^n \mathbb{P}\{X \geq t\}dt - 2^{-n}.$$

Dus,

$$\mathbb{E}[X] = \lim_{n\to\infty} a_n = \lim_{n\to\infty} \int_0^n \mathbb{P}\{X \geq t\}dt = \int_0^\infty \mathbb{P}\{X \geq t\}dt,$$

waarmee de stelling bewezen is. □

Opmerking. Bovenstaande stelling impliceert onmiddellijk dat als $X, Y : \Omega \to [0.\infty[$ toevalsvariabelen zijn zodanig dat $\mathbb{P}_X = \mathbb{P}_Y$, er geldt: $\mathbb{E}[X] = \mathbb{E}[Y]$.
Dus de verwachtingswaarde van een niet-negatieve toevalsvariabele X is bepaald door zijn verdeling $\mathbb{P}_X$.

Gegeven een toevalsvariabele $X : \Omega \to \mathbb{R}$ definiëren we

$$X^+ := X \vee 0 \text{ en } X^- := (-X) \vee 0.$$

Dan is X^+ $\mathcal{F}$-meetbaar omdat geldt $X^+ = g \circ X$, waar $g(x) = x \vee 0, x \in \mathbb{R}$ continu en bijgevolg Borel-meetbaar is. Dit impliceert dat ook $X^- = (-X)^+$ $\mathcal{F}$-meetbaar is. Bijgevolg zijn X^+ en X^- niet-negatieve toevalsvariabelen waarvoor verwachtingswaarden gedefinieerd zijn. We noteren verder dat geldt:

$$X = X^+ - X^- \text{ en } |X| = X^+ + X^-.$$

Definitie 4.3 *Zij $X : \Omega \to \mathbb{R}$ een toevalsvariabele.*

(i) *Als tenminste één van de twee verwachtingswaarden $\mathbb{E}[X^+], \mathbb{E}[X^-]$ eindig is, zeggen we dat X* **semi-integreerbaar** *is en we stellen $\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-] \in [-\infty, \infty]$, waarbij $\infty - a = \infty, a - \infty = -\infty, a < \infty$.*

(ii) *Indien zowel $\mathbb{E}[X^+]$ als $\mathbb{E}[X^-]$ eindig zijn, zeggen we dat X* **integreerbaar** *is. In dit geval geldt $\mathbb{E}[X] \in \mathbb{R}$.*

Opmerkingen.

1. Uit de definitie en lemma 4.3.i volgt onmiddellijk dat een toevalsvariabele $X : \Omega \to \mathbb{R}$ integreerbaar is als en slechts als $\mathbb{E}[|X|] < \infty$. Bovendien hebben we in dit geval $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$.

2. Verder is het evident dat we algemeen hebben als $X, Y : \Omega \to \mathbb{R}$ toevalsvariabelen met identieke verdelingen zijn, X semi-integreerbaar is als en slechts als Y semi-integreerbaar is. In dit geval hebben we: $\mathbb{E}[X] = \mathbb{E}[Y]$.

3. Een speciaal geval van de laatste eigenschap is de volgende: We zeggen dat $X = Y$ bijna overal (b.o.) indien $\mathbb{P}\{X \neq Y\} = 0$. Dan geldt natuurlijk $\mathbb{P}_X = \mathbb{P}_Y$ en bijgevolg is X semi-integreerbaar als en slechts als Y semi-integreerbaar is. We hebben dan weer dat $\mathbb{E}[X] = \mathbb{E}[Y]$.

In het discrete geval geldt verder:

Stelling 4.4 *Zij $X : \Omega \to \mathbb{R}$ een discrete toevalsvariabele. Deze is integreerbaar als en slechts als $\sum_x |x| p_X(x) < \infty$. In dit geval geldt er:*

$$\mathbb{E}[X] = \sum_x x p_X(x).$$

Bewijs. Dit volgt onmiddellijk uit lemma 4.1. □

Om het analoge resultaat voor toevalsvariabelen met absoluut continue verdelingen te bewijzen, gebruiken we stelling 4.3.

Stelling 4.5 *Zij $X : \Omega \to \mathbb{R}$ een toevalsvariabele met een absoluut continue verdeling. Deze is integreerbaar als en slechts als $\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty$. In dit geval geldt er:*

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx.$$

Bewijs. We bepalen eerst $\mathbb{E}[X^+]$. Gezien $\mathbb{P}\{X^+ \geq t\} = \mathbb{P}\{X \geq t\}, t > 0$ volgt via stelling 4.3,

$$\begin{aligned}
\mathbb{E}[X^+] &= \int_0^{\infty} \mathbb{P}\{X \geq t\} dt \\
&= \int_0^{\infty} \int_t^{\infty} f_X(x) dx dt \\
&= \int_0^{\infty} \int_0^{x} 1 dt f_X(x) dx = \int_0^{\infty} x f_X(x) dx.
\end{aligned}$$

Wegens $X^- = (-X)^+$ en $f_{-X}(x) = f_X(-x)$ zien we na een substitutie dat

$$\mathbb{E}[X^-] = \int_0^{\infty} x f_{-X}(x) dx = \int_0^{\infty} x f_X(-x) dx = -\int_{-\infty}^{0} x f_X(x) dx.$$

Dus hebben we:

$$\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-] = \int_0^{\infty} x f_X(x) dx + \int_{-\infty}^{0} x f_X(x) dx = \int_{-\infty}^{\infty} x f_X(x) dx,$$

waarmee de stelling bewezen is. □

Voorbeelden

1. $X \sim \text{gamma}(\alpha, \beta) \Rightarrow \mathbb{E}[X] = \alpha\beta$. (oef.)

2. $X \sim \text{normaal}(\mu, \sigma^2) \Rightarrow \mathbb{E}[X] = \mu$. (oef.)

3. Als X een Cauchy-verdeling heeft, i.e. X absoluut-continu met dichtheidsfunctie $f_X(x) = (\pi(1+x^2)^{-1}, x \in \mathbb{R}$, bestaat $\mathbb{E}[X]$ niet, omdat

$$\mathbb{E}[X^+] = \mathbb{E}[X^-] = \int_0^\infty x/(\pi(1+x^2))dx = \infty.$$

We bewijzen nu enkele algemene resultaten over verwachtingswaarden.

Stelling 4.6

(i) **(Leariteit van de verwachtingswaarde)** *Indien X, Y integreerbare toevalsvariabelen zijn, is voor elke keuze van $\alpha, \beta \in \mathbb{R}$ de toevalsvariabele $\alpha X + \beta Y$ ook integreerbaar en er geldt: $\mathbb{E}[\alpha X + \beta Y] = \alpha\mathbb{E}[X] + \beta\mathbb{E}[Y]$.*

(ii) **(Monotonie van de verwachtingswaarde)** *Gegeven twee integreerbare toevalsvariabelen $X \leq Y$, hebben we $\mathbb{E}[X] \leq \mathbb{E}[Y]$.*

(iii) **(Onafhankelijkheid)** *Als $X_1, \ldots, X_n$ onafhankelijke integreerbare toevalsvariabelen zijn, is de toevalsvariabele $\prod_{i=1}^n X_i$ integreerbaar en $\mathbb{E}[\prod_{i=1}^n X_i] = \prod_{i=1}^n \mathbb{E}[X_i]$.*

Bewijs (i) Gezien $|\alpha X + \beta Y| \leq |\alpha||X| + |\beta||Y|$ volgt uit lemma 4.3.i in verband met de monotonie van verwachtingswaarden voor positieve toevalsvariabelen (zie opmerking (2) na de definitie):

$$\mathbb{E}[|\alpha X + \beta Y|] \leq |\alpha|\mathbb{E}[|X|] + |\beta|\mathbb{E}[|Y|] < \infty.$$

Dus is $\alpha X + \beta Y$ integreerbaar.

Om de leariteit van de verwachtingswaarden te bewijzen, tonen we eerst:

(1) $\boxed{X, Y \text{ integreerbaar} \Rightarrow \mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]}$

Om (1) te bewijzen, noteren we dat

$$X + Y = (X+Y)^+ - (X+Y)^- = X^+ - X^- + Y^+ - Y^-$$

wat natuurlijk impliceert:

$$(X+Y)^+ + X^- + Y^- = (X+Y)^- + X^+ + Y^+.$$

Uit lemma 4.3.i volgt dan

$$\mathbb{E}[(X+Y)^+] + \mathbb{E}[X^-] + \mathbb{E}[Y^-] = \mathbb{E}[(X+Y)^-] + \mathbb{E}[X^+] + \mathbb{E}[Y^+].$$

Vermits de zes bovenstaande verwachtingswaarden eindig zijn (de drie toevalsvariabelen $X, Y, X+Y$ zijn integreerbaar) kunnen we concluderen dat

$$\begin{aligned} \mathbb{E}[X+Y] &= \mathbb{E}[(X+Y)^+] - \mathbb{E}[(X+Y)^-] \\ &= \mathbb{E}[X^+] - \mathbb{E}[X^-] + \mathbb{E}[Y^+] - \mathbb{E}[Y^-] = \mathbb{E}[X] + \mathbb{E}[Y]. \end{aligned}$$

We beweren nu dat verder geldt:

(2) $\boxed{X \text{ integreerbaar}, \alpha \in \mathbb{R} \Rightarrow \mathbb{E}[\alpha X] = \alpha \mathbb{E}[X]}$
hetgeen in verband met betrekking (1) de lineariteit van de verwachtingswaarde impliceert.

Om (2) te bewijzen, kiezen we eerst $\alpha \geq 0$ en noteren dat dan geldt:

$$(\alpha X)^+ = \alpha X^+, \ (\alpha X)^- = \alpha X^-.$$

Dus: $\mathbb{E}[\alpha X] = \mathbb{E}[(\alpha X)^+] - \mathbb{E}[(\alpha X)^-] = \alpha \mathbb{E}[X]$.
Als $\alpha < 0$, geldt:

$$0 = \mathbb{E}[\alpha X + (-\alpha)X] \overset{(1)}{=} \mathbb{E}[\alpha X] + \mathbb{E}[(-\alpha)X] = \mathbb{E}[\alpha X] - \alpha \mathbb{E}[X].$$

Dus is (2) ook juist als $\alpha < 0$ en deel (i) is bewezen.
(ii) Min of meer triviaal. ($\mathbb{E}[Y] = \mathbb{E}[X] + \underbrace{\mathbb{E}[Y - X]}_{\geq 0} \geq \mathbb{E}[X]$.)

(iii) Stel $Z_i^{(j)} = \begin{cases} X_i^+ & \text{als } j = 0 \\ X_i^- & \text{als } j = 1 \end{cases}, j = 0, 1; 1 \leq i \leq n.$
Dan hebben we

$$Z := \prod_{i=1}^n X_i = \prod_{i=1}^n (Z_i^{(0)} - Z_i^{(1)}) = \sum_{\underline{j} \in \{0,1\}^n} (-1)^{\sum_{i=1}^n j_i} \prod_{i=1}^n Z_i^{(j_i)},$$

waar voor elke $\underline{j} \in \{0,1\}^n$ de positieve toevalsvariabelen $Z_i^{(j_i)}, 1 \leq i \leq n$ onafhankelijk en integreerbaar (wegens $Z_i^{(j)} \leq |X_i|$) zijn.
Dus kunnen we uit lemma 4.3.ii en deel (i) van onze stelling afleiden dat Z integreerbaar is.
Bovendien volgt dat

$$\mathbb{E}[Z] = \sum_{\underline{j} \in \{0,1\}^n} (-1)^{\sum_{i=1}^n j_i} \prod_{i=1}^n \mathbb{E}[Z_i^{(j_i)}] = \prod_{i=1}^n (\mathbb{E}[Z_i^{(0)}] - \mathbb{E}[Z_i^{(1)}]) = \prod_{i=1}^n \mathbb{E}[X_i]$$

en stelling 4.6 is bewezen. □

Stelling 4.7 (Gedomineerde convergentie)
Zij $X_n, n \geq 1$ *een rij toevalsvariabelen die naar een toevalsvariabele* X *convergeert. Veronderstel dat er een integreerbare toevalsvariabele* $Y : \Omega \to [0, \infty[$ *bestaat zodanig dat voor één* $n_0 \geq 1$ *geldt:* $\boxed{|X_n| \leq Y, n \geq n_0.}$ *Dan is* X *integreerbaar en* $\mathbb{E}[X_n] \to \mathbb{E}[X]$ *als* $n \to \infty$.

Bewijs Het is evident dat onder de bovenstaande voorwaarde geldt: $|X| = \lim_n |X_n| \leq Y$ en dus $\mathbb{E}[|X|] \leq \mathbb{E}[Y] < \infty$, hetgeen impliceert dat X integreerbaar is.

Verder geldt: $X = \liminf_n X_n = \lim_n Z_n$, waar $Z_n = \inf_{k \geq n} X_k, n \geq 1$ een stijgende rij van toevalsvariabelen is. Stel $Z_n' = Z_n + Y, n \geq n_0$. Dan is $0 \leq Z_n'$ een stijgende rij van toevalsvariabelen met $Z_n' \nearrow X + Y$ als $n \to \infty$ en de stelling van de monotone convergentie impliceert dat

$$\lim_{n \to \infty} \mathbb{E}[Z_n'] = \mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y].$$

Wegens $|Z_n| \leq Y, n \geq n_0$ zijn deze toevalsvariabelen integreerbaar en bijgevolg geldt:

$$\mathbb{E}[Z'_n] = \mathbb{E}[Z_n] + \mathbb{E}[Y], n \geq n_0$$

Het is nu evident dat $\mathbb{E}[Z_n] \to \mathbb{E}[X]$ als $n \to \infty$.

Gezien $X_n \geq Z_n, n \geq 1$ volgt verder dat

$$\mathbb{E}[X] = \liminf_{n\to\infty} \mathbb{E}[Z_n] \leq \liminf_{n\to\infty} \mathbb{E}[X_n] \tag{4.1}$$

Als we in betrekking (4.1) X door $-X$ en X_n door $-X_n$ vervangen, volgt:

$$-\mathbb{E}[X] = \mathbb{E}[-X] \leq \liminf_{n\to\infty} \mathbb{E}[-X_n] = -\limsup_{n\to\infty} \mathbb{E}[X_n].$$

Dus

$$\mathbb{E}[X] \geq \limsup_{n\to\infty} \mathbb{E}[X_n] \tag{4.2}$$

Betrekkingen (4.1) en (4.2) impliceren samen dat $\mathbb{E}[X_n] \to \mathbb{E}[X]$ als $n \to \infty$ en de stelling van de gedomineerde convergentie is bewezen. □

Opmerking De bovenstaande stelling impliceert in het bijzonder dat als X_n een uniform begrensde rij van toevalsvariabelen is (dwz, $\exists\, K > 0, |X_n| \leq K, n \geq 1$) die naar een (begrensde) toevalsvariabele X convergeert, geldt: $\mathbb{E}[X_n] \to \mathbb{E}[X]$ als $n \to \infty$. Dit speciale geval van de stelling van de gedomineerde convergentie noemt men ook de *stelling van de begrensde convergentie.*

We vermelden nog de volgende generalisatie van stellingen 4.4 en 4.5 die men analoog als deze resultaten kan bewijzen.

Stelling 4.8 *Zij $X : \Omega \to \mathbb{R}$ een toevalsvariabele en $g : \mathbb{R} \to \mathbb{R}$ een Borel-meetbare afbeelding zodanig dat $\mathbb{E}[|g(X)|] < \infty$. Dan geldt:*

$$\mathbb{E}[g(X)] = \begin{cases} \sum_x g(x)p_X(x) & \textit{(discreet geval)} \\ \int_{-\infty}^{\infty} g(x)f_X(x)dx & \textit{(continu geval)} \end{cases}$$

4.2 Variantie, covariantie en moment-genererende functies

Definitie 4.4 *Zij $X : \Omega \to \mathbb{R}$ een toevalsvariabele zodanig dat voor één $k \in \{1, 2, \ldots\}$ geldt: $\mathbb{E}[|X|^k] < \infty$. Dan bestaat $\mathbb{E}[X^k]$ en we noemen deze grootheid het k-de* **moment** *van X. Verder heet $\mathbb{E}[|X|^k]$ het k-de* **absolute moment** *van X.*

Opmerking Vermits geldt $|X|^j \leq |X|^k + 1, 1 \leq j \leq k$ zien we dat als het k-de moment van een toevalsvariabele X bestaat, ook de momenten van de orde $j \leq k$ bestaan.

Definitie 4.5 *Zij $X : \Omega \to \mathbb{R}$ een toevalsvariabele met $\mathbb{E}[X^2] < \infty$.*
Dan definiëren we de **variantie** *van X (Notatie: $Var(X)$) door $Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$.*

Eigenschappen

1. Natuurlijk geldt $\mathrm{Var}(X) \geq 0$.
2. $\mathbb{P}\{X = \mathbb{E}[X]\} = 1 \Rightarrow \mathrm{Var}(X) = 0$.
3. $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X),\ a, b \in \mathbb{R}$.
4. $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. Dus ook $\mathbb{E}[X]^2 \leq \mathbb{E}[X^2]$.
5. (*Ongelijkheid van Chebyshev*) $\mathbb{P}\{|X - \mathbb{E}[X]| \geq \epsilon\} \leq \mathrm{Var}(X)/\epsilon^2$.
   (Om dit te bewijzen, noteren we dat $I_{\{|X-\mathbb{E}[X]|\geq\epsilon\}} \leq (X - \mathbb{E}[X])^2/\epsilon^2$ en de ongelijkheid volgt uit de monotonie van de verwachtingswaarde.)
6. $\mathrm{Var}(X) = 0 \Rightarrow \mathbb{P}\{X = \mathbb{E}[X]\} = 1$. Dus (zie (2)) is de variantie van een toevalsvariabele gelijk aan 0 als en slechts als X deterministisch is.
   (Bewijs: $\mathbb{P}\{X \neq \mathbb{E}[X]\} = \lim_n \mathbb{P}\{|X - \mathbb{E}[X]| \geq 1/n\} \overset{(5)}{\leq} \limsup_n n^2\mathrm{Var}(X) = 0$.)

Voorbeelden

1. $X \sim \text{normaal}(0,1) \Rightarrow \mathbb{E}[X] = 0 \Rightarrow \mathrm{Var}(X) = \mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 e^{-x^2/2}dx/\sqrt{2\pi} = 1$. (partiële integratie)
2. Zij nu $Y \sim \text{normaal}(\mu, \sigma^2)$. Dan geldt: $X = (Y - \mu)/\sigma \sim \text{normaal}(0,1)$ en bijgevolg $\mathrm{Var}(X) = \mathrm{Var}(Y)/\sigma^2 = 1$. Dus: $\mathrm{Var}(Y) = \sigma^2$.
3. $X \sim \text{Poisson}(\lambda) \Rightarrow \mathbb{E}[X] = \lambda \Rightarrow \mathrm{Var}(X) = \mathbb{E}[X^2] - \lambda^2$. Verder geldt:
$$\mathbb{E}[X^2] = \mathbb{E}[X] + \mathbb{E}[X(X-1)] = \lambda + \sum_{k=2}^{\infty} k(k-1)\lambda^k e^{-\lambda}/k! = \lambda + \lambda^2$$
   en het volgt dat $\mathrm{Var}(X) = \lambda$.

De momenten en de variantie laten zich vaak efficiënter berekenen als de toevalsvariabele X een eindige moment-genererende functie heeft.

Definitie 4.6 *Als $X : \Omega \to \mathbb{R}$ een toevalsvariabele is die voldoet aan de voorwaarde*
$$\exists t_0 > 0 : R(t) := \mathbb{E}[e^{tX}] < \infty, |t| \leq t_0,$$
zeggen we dat X een **eindige moment-genererende functie** *(=mgf) heeft.*

Stelling 4.9 *Zij $X : \Omega \to \mathbb{R}$ een toevalsvariabele die een eindige mgf heeft. Dan geldt er:*
(i) $\mathbb{E}[|X|^n] < \infty, \forall n \geq 1$.
(ii) $R(t) = \sum_{j=0}^{\infty} \mathbb{E}[X^j]t^j/j!, |t| \leq t_0$.
(iii) $\mathbb{E}[X^j] = R^{(j)}(0), j = 0, 1, 2, \ldots$

Bewijs (i) Als $0 < s \le t_0$ geldt: $\mathbb{E}[e^{s|X|}] \le R(-s) + R(s) < \infty$. Dit impliceert in verband met de triviale ongelijkheid $e^{s|x|} \ge |x|^n s^n/n!$:

$$\mathbb{E}[|X|^n] \le n! s^{-n} \mathbb{E}[e^{s|X|}] < \infty, n = 1, 2, \ldots$$

(ii) Zij $t \in [-t_0, t_0]$ vast. Stel $Z_n = \sum_{j=0}^n X^j t^j/j!$. Dan geldt natuurlijk: $Z_n \to e^{tX}$ als $n \to \infty$. Vermits we ook hebben:

$$|Z_n| \le \sum_{j=0}^n |t|^j |X|^j/j! \le e^{|t||X|},$$

waar de toevalsvariabele $e^{|t||X|}$ integreerbaar is (zie bewijs van (i)), kunnen we via de stelling van de gedomineerde convergentie concluderen dat

$$\sum_{j=0}^n t^j \mathbb{E}[X^j]/j! = \mathbb{E}[Z_n] \to \mathbb{E}[e^{tX}] = R(t), |t| \le t_0.$$

(iii) Gezien we een representatie van $R(t), |t| \le t_0$ als een machtreeks hebben gevonden, volgt uit de analyse dat alle afgeleiden $R^{(k)}(t)$ van R voor $|t| < t_0$ en $k = 1, 2, \ldots$ bestaan. Verder geldt:

$$R^{(k)}(t) = \sum_{j=k}^{\infty} \mathbb{E}[X^j] t^{j-k}/(j-k)!$$

wat natuurlijk (iii) impliceert. □

Opmerking Stel $L(t) := \log R(t), |t| < t_0$, waar $\log s$ de natuurlijke logaritme van $s > 0$ betekent. Dan volgt onmiddellijk uit (iii) (en het triviale feit dat $R(0) = 1$):

$$\mathbb{E}[X] = L'(0), \mathrm{Var}(X) = L''(0).$$

Dus kunnen we als $L(t)$ bestaat de verwachtingswaarde en de variantie door differentiëren van de functie L verkrijgen.

Voorbeelden

1. Als $X \sim$ Poisson(λ), geldt $\mathbb{E}[e^{tX}] = \sum_{k=0}^{\infty} e^{tk}\lambda^k e^{-\lambda}/k! = e^{\lambda(e^t-1)}, t \in \mathbb{R}$. Dus $L(t) = \lambda(e^t - 1)$ en $L'(t) = L''(t) = \lambda e^t \Rightarrow L'(0) = \mathbb{E}[X] = \mathrm{Var}(X) = L''(0) = \lambda$, wat we boven al via de definities van verwachtingswaarde en variantie hebben bewezen.

2. Als $X \sim$ exponentieel(λ), geldt: $\mathbb{E}[e^{tX}] = \int_0^{\infty} e^{tx} e^{-x/\lambda}\lambda^{-1} dx = (1 - \lambda t)^{-1}, |t| < 1/\lambda$. Dus: $R(t) = \sum_{k=0}^{\infty} \mathbb{E}[X^k] t^k/k! = \sum_{k=0}^{\infty} \lambda^k t^k, t < 1/\lambda$, wat natuurlijk impliceert (de coëfficiënten van een machtreeks zijn uniek): $\mathbb{E}[X^k] = k!\lambda^k, k = 1, 2, \ldots$ en $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \lambda^2$.

3. Als $X \sim$ normaal$(0,1)$, geldt (oef.) $\mathbb{E}[e^{tX}] = e^{t^2/2}, t \in \mathbb{R}$ en als we deze functie als een machtreeks opschrijven volgt dat $\mathbb{E}[X^{2k}] = 2^{-k}(2k)!/k!, k = 1, 2, \ldots$

Onderstel dat X, Y toevalsvariabelen met eindige tweede momenten zijn, i.e. $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$. Dan is ook het product XY van deze toevalsvariabelen (wegens $|XY| \le (X^2 + Y^2)/2$) integreerbaar en we kunnen de **covariantie** van X, Y definiëren als

$$\mathrm{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

De volgende eigenschappen van de covariantie zijn evident.

Lemma 4.4 *Onderstel dat X, Y, Z toevalsvariabelen zijn zodanig dat $\mathbb{E}[X^2+Y^2+Z^2] < \infty$. Dan geldt:*

(i) X, Y *onafhankelijk* $\Rightarrow \mathrm{cov}(X,Y) = 0$.
(ii) $\mathrm{cov}(X,X) = \mathrm{Var}(X)$.
(iii) $\mathrm{cov}(X,Y) = \mathrm{cov}(Y,X)$.
(iv) $\mathrm{cov}(aX+b,Y) = a\,\mathrm{cov}(X,Y)$.
(v) $\mathrm{cov}(X+Y,Z) = \mathrm{cov}(X,Z) + \mathrm{cov}(Y,Z)$.

We kunnen nu de volgende belangrijke formule voor de variantie van de som van n toevalsvariabelen bewijzen.

Stelling 4.10 *Onderstel dat $X_1, \ldots, X_n$ toevalsvariabelen zijn met eindige tweede momenten. Dan geldt:*

$$\mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{Var}(X_i) + 2\sum_{1\le i<j\le n} \mathrm{cov}(X_i, X_j).$$

Opmerking Stelling 4.10 impliceert natuurlijk via lemma 4.4.i dat als de toevalsvariabelen $X_1, \ldots, X_n$ paarsgewijs onafhankelijk zijn, geldt: $\mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{Var}(X_i)$.

Bewijs van stelling 4.10 Met behulp van lemma 4.4 volgt:

$$\begin{aligned}
\mathrm{Var}(\sum_{i=1}^n X_i) &\overset{(ii)}{=} \mathrm{cov}(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i) \overset{(v)}{=} \sum_{i=1}^n \mathrm{cov}(X_i, \sum_{i=1}^n X_i) \\
&\overset{(iii)}{=} \sum_{i=1}^n \mathrm{cov}(\sum_{j=1}^n X_j, X_i) \overset{(v)}{=} \sum_{i=1}^n\sum_{j=1}^n \mathrm{cov}(X_j, X_i) \\
&\overset{(ii),(iii)}{=} \sum_{i=1}^n \mathrm{Var}(X_i) + 2\sum_{1\le i<j\le n} \mathrm{cov}(X_i, X_j). \quad \square
\end{aligned}$$

Voorbeelden

1. Als $X = \sum_{i=1}^n I_{A_i}$, waar A_i onafhankelijke gebeurtenissen zijn met $\mathbb{P}(A_i) = p, 1 \le i \le n$ en dus $X \sim$ binomiaal(n,p), volgt dat $\mathrm{Var}(X) = \sum_{i=1}^n \mathrm{Var}(I_{A_i}) = np(1-p)$.

2. Als we uit een vaas met M blauwe en N gele ballen n ballen ($n \le N+M$) zonder teruglegging trekken, dan heeft de toevalsvariabele $X = \#$ getrokken blauwe ballen een hypergeometrische verdeling (zie voorbeeld (5) in 2.2). In principe, kunnen we dan de variantie met behulp van de kansfunctie van X berekenen.
   Maar we weten ook dat $X = \sum_{i=1}^n I_{A_i}$, als A_i de gebeurtenis is dat de i-de getrokken bal blauw is, $1 \le i \le n$. Zoals in het voorbeeld na stelling 1.3 kan men bewijzen dat $\mathbb{P}(A_i) = p := M/(N+M), 1 \le i \le n$. Deze gebeurtenissen zijn *niet* onafhankelijk, zodat we nu in tegenstelling tot het eerste voorbeeld ook met het tweede gedeelte van de formule uit stelling 4.10 moeten rekening houden.
   We noteren eerst dat $\mathrm{cov}(I_{A_i}, I_{A_j}) = \mathrm{cov}(I_{A_1}, I_{A_2}), i \ne j$. Verder geldt:

$$\begin{aligned}
\mathrm{cov}(I_{A_1}, I_{A_2}) &= \mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1)\mathbb{P}(A_2) \\
&= \mathbb{P}(A_1)(\mathbb{P}(A_2|A_1) - \mathbb{P}(A_2)) \\
&= p(\frac{M-1}{N+M-1} - \frac{M}{N+M}) = -\frac{p(1-p)}{N+M-1},
\end{aligned}$$

hetgeen in verband met stelling 4.10 impliceert dat

$$\mathrm{Var}(X) = np(1-p) + n(n-1)\mathrm{cov}(I_{A_1}, I_{A_2}) = np(1-p)\frac{N+M-n}{N+M-1}.$$

Definitie 4.7 (correlatiecoëfficiënt)
Als X, Y toevalsvariabelen zijn met $\sigma_X^2 := Var(X) \in]0,\infty[$ en $\sigma_Y^2 := Var(Y) \in]0,\infty[$ definiëren we de correlatiecoëfficiënt van X en Y als

$$\rho_{X,Y} := \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$$

Stelling 4.11 *Zij X, Y zoals in definitie 4.7. Dan geldt voor de correlatiecoëfficiënt:*
(i) $\rho_{X,Y} \in [-1,1]$
(ii) $\rho_{X,Y} = 1 \Leftrightarrow \exists a > 0, b \in \mathbb{R} : \mathbb{P}\{Y = aX + b\} = 1$.
(iii) $\rho_{X,Y} = -1 \Leftrightarrow \exists a < 0, b \in \mathbb{R} : \mathbb{P}\{Y = aX + b\} = 1$.

Bewijs (i) Stelling 4.10 in verband met lemma 4.4 impliceert onmiddellijk dat
$0 \le \mathrm{Var}(X/\sigma_X - Y/\sigma_Y) = 2 - 2\mathrm{cov}(X/\sigma_X, Y/\sigma_Y) = 2(1-\rho_{X,Y})$. Dus : $\rho_{X,Y} \le 1$.
Wegens $1 \ge \rho_{-X,Y} = -\rho_{X,Y}$, volgt dan ook $\rho_{X,Y} \ge -1$.
(ii) Als $\rho_{X,Y} = 1$ impliceert de formule in het bewijs van (i) dat $\mathrm{Var}(X/\sigma_X - Y/\sigma_Y) = 0$. Dus (eigenschap (6) van de variantie) bestaat er een $c \in \mathbb{R}$ ($c = \mathbb{E}[X/\sigma_X - Y/\sigma_Y]$) zodanig dat

$$1 = \mathbb{P}\{X/\sigma_X - Y/\sigma_Y = c\} = \mathbb{P}\{Y = (\sigma_Y/\sigma_X)X - c\sigma_Y\}.$$

Omgekeerd volgt uit $\mathbb{P}\{Y = aX + b\} = 1$:

$$\mathrm{cov}(X,Y) = \mathrm{cov}(aX+b, X) = a\sigma_X^2 \text{ en } \sigma_Y^2 = \mathrm{Var}(aX+b) = a^2\sigma_X^2,$$

hetgeen wegens $a > 0$ impliceert dat $\rho_{X,Y} = 1$.
(iii) Gebruik (ii) en het feit dat $\rho_{-X,Y} = -\rho_{X,Y}$. □

Opmerking Als $\rho_{X,Y} > 0$ ($\rho_{X,Y} < 0$) zeggen we dat er een positieve (negatieve) correlatie tussen X en Y bestaat. Als X, Y onafhankelijk zijn, geldt $\rho_{X,Y} = 0$, maar het omgekeerde geldt niet.

We bekijken nog k-dimensionale toevalsvectoren $\vec{X} = (X_1, \ldots, X_k)^t : \Omega \to \mathbb{R}^k$. Deze heten integreerbaar indien de toevalsvariabelen $X_1, \ldots, X_k$ allemaal integreerbaar zijn en we definiëren in dit geval de verwachtingswaarde door $\mathbb{E}[\vec{X}] = (\mathbb{E}[X_1], \ldots, \mathbb{E}[X_k])^t$.
Als $\mathbb{E}[X_i^2] < \infty, 1 \le i \le k$ definiëren we verder een (k,k)-matrix $\Sigma = (\Sigma_{i,j})_{1 \le i,j \le k}$ door

$$\Sigma_{i,j} := \mathrm{cov}(X_i, X_j), 1 \le i, j \le k.$$

We noemen deze matrix de **covariantiematrix** van X. (Notatie: $\mathrm{Cov}(\vec{X})$.)

Stelling 4.12 *Zij $\vec{X} : \Omega \to \mathbb{R}^k$ een toevalsvector met covariantiematrix Σ. Dan is Σ symmetrisch en positief semi-definiet.*

Bewijs Σ is symmetrisch omdat geldt $\text{cov}(X_i, X_j) = \text{cov}(X_j, X_i), i, j \in \{1, \ldots, k\}$. (lemma 4.4.iii)
Om te tonen dat Σ positief semi-definiet is, kiezen we een vector $\underline{z} = (z_1, \ldots, z_k)^t \in \mathbb{R}^k$. Dan geldt wegens lemma 4.4.:

$$\underline{z}^t \, \Sigma \, \underline{z} = \sum_{i=1}^k \sum_{j=1}^k z_i \Sigma_{i,j} z_j = \sum_{i=1}^k \sum_{j=1}^k z_i \text{cov}(X_i, X_j) z_j = \text{Var}(\sum_{i=1}^k z_i X_i) \geq 0$$

en de stelling is bewezen. □

Opmerking Men kan bewijzen dat als Σ niet positief-definiet is en dus rang(Σ)= $m < k$, er een deelruimte V van $\mathbb{R}^k$ met dim(V)=m en een vector $\underline{a} \in \mathbb{R}^k$ bestaan zodanig dat

$$\mathbb{P}\{\vec{X} \in \underline{a} + V\} = 1.$$

Voorbeeld Zij $\vec{X} = (X_1, \ldots, X_k)^t \sim \text{normaal}(\underline{\mu}, \Sigma)$, waar Σ een symmetrische, positief definitie (k, k)-matrix is. Dan geldt:

$$I\!E[\vec{X}] = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_k \end{bmatrix} \text{ en Cov(X)} = \Sigma.$$

Om dit te bewijzen, gebruiken we het feit dat $\vec{X}$ een representatie als

$$\begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_k \end{bmatrix} + A \begin{bmatrix} Z_1 \\ \vdots \\ Z_k \end{bmatrix}$$

heeft, waar A een symmetrische (k,k)-matrix is met $A^2 = \Sigma$ en $Z_1, \ldots, Z_k$ onafhankelijke standaard-normaalverdeelde toevalsvariabelen zijn (zie lemma 3.1). Dit impliceert dat

$$I\!E[\vec{X}] = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_k \end{bmatrix} + I\!E[A \cdot \vec{Z}] = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_k \end{bmatrix} + A \cdot I\!E[\vec{Z}]$$

en ook dat Cov($\vec{X}$)=A·Cov($\vec{Z}$)·A^t (waar de formule voor $I\!E[\vec{X}]$ direct uit de lineariteit van de (1-dimensionale) verwachtingswaarde volgt en de formule voor Cov($\vec{X}$) uit lemma 4.4).
Het is evident dat $I\!E[\vec{Z}] = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ en verder geldt (wegens de onafhankelijkheid) Cov($\vec{Z}$)= I (= de k-dimensionale identiteitsmatrix). Dus geldt $I\!E[\vec{X}] = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_k \end{bmatrix}$ en Cov($\vec{X}$)= $A \cdot A^t = A^2 = \Sigma$.

Alternatief kunnen we dit ook via de volgende k-dimensionale versie van stelling 4.8 bewijzen.

Stelling 4.13 *Zij $X_1, \ldots, X_k$ toevalsvariabelen met een discrete of een gezamenlijk continue verdeling. Zij $h : \mathbb{R}^k \to \mathbb{R}$ en Borel-meetbare afbeelding zodanig dat $I\!E[|h(X_1, \ldots, X_k)|] < \infty$. Dan geldt:*

$$I\!E[h(X_1, \ldots, X_k)] = \begin{cases} \sum_{x_1} \cdots \sum_{x_k} h(x_1, \ldots, x_k) p_{X_1, \ldots, X_k}(x_1, \ldots, x_k) & \textit{(discreet geval)} \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(x_1, \ldots, x_k) f_{X_1, \ldots, X_k}(x_1, \ldots, x_k) dx_1 \ldots dx_k & \textit{(continu geval)} \end{cases}$$

4.3 Conditionele verwachtingswaarden

We definiëren in dit gedeelte van hoofdstuk 4 de conditionele verwachtingswaarde van een toevalsvariabele X gegeven dat een toevalsvector $Y : \Omega \to \mathbb{R}^d$ gelijk aan een bepaalde waarde $y \in \mathbb{R}^d$ is. We moeten ons hier beperken tot de twee gevallen waar X en Y discreet of gezamenlijk continu zijn.

Later kunnen we dan met behulp van de maattheorie ($\to$ bachelor 3) conditionele verwachtingswaarden algemeen definiëren.

Definitie 4.8 *Zij $X : \Omega \to \mathbb{R}$ een toevalsvariabele en zij $Y : \Omega \to \mathbb{R}^d$ een toevalsvector die discreet of gezamenlijk continu zijn. Als $\mathbb{E}[|X|] < \infty$ en $y \in \mathbb{R}^d$ zodanig is dat $p_Y(y) > 0$ (in het discrete geval) of $f_Y(y) > 0$ (in het continue geval), definiëren we de* **conditionele verwachtingswaarde** *van X gegeven $Y = y$ door*

$$\mathbb{E}[X \| Y = y] = \begin{cases} \sum_x x p_{X|Y}(x|y) & \textit{(discreet geval)} \\ \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx & \textit{(continu geval)} \end{cases}$$

Voorbeeld Neem twee gezamenlijk continue toevalsvariabelen X, Y met dichtheidsfunctie

$$f(x,y) = \begin{cases} \beta^{-2} e^{-y/\beta} & 0 < x < y < \infty \\ 0 & \text{elders} \end{cases}$$

Dan geldt:

$$f_Y(y) = \int_0^y \beta^{-2} e^{-y/\beta} dx = \beta^{-2} y e^{-y/\beta}, y > 0,$$

waaruit blijkt dat

$$f_{X|Y}(x|y) = 1/y, 0 < x < y$$

Dus is de conditionele verdeling van X gegeven $Y = y$ uniform op het interval $(0, y)$ en de conclusie is dat $\mathbb{E}[X \| Y = y] = y/2, y > 0$.

Als we $g(y) = \mathbb{E}[X \| Y = y]$ stellen indien $p_Y(y) > 0$ (discreet geval) of $f_Y(y) > 0$ (continu geval) en $g(y) = 0$ elders, verkrijgen we een Borel-meetbare afbeelding $g : \mathbb{R}^d \to \mathbb{R}$.

We definiëren dan de **conditionele verwachtingswaarde van X, gegeven** Y door

$$\mathbb{E}[X \| Y] = g \circ Y$$

Dus $\mathbb{E}[X \| Y] : \Omega \to \mathbb{R}$ is een *toevalsvariabele* en als $Y(\omega) = y$ geldt er

$$\mathbb{E}[X \| Y](\omega) = \mathbb{E}[X \| Y = y].$$

Stelling 4.14 *Zij $X : \Omega \to \mathbb{R}$ een toevalsvariabele en zij $Y : \Omega \to \mathbb{R}^d$ een toevalsvector die discreet of gezamenlijk continu zijn. Voor elke Borel-verzameling $C \in \mathcal{R}^d$ geldt:*

$$\mathbb{E}[X I_{\{Y \in C\}}] = \mathbb{E}[\mathbb{E}[X \| Y] I_{\{Y \in C\}}].$$

In het bijzonder: $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \| Y]]$.

Bewijs We bewijzen de stelling in het continue geval.
Stel $N = \{f_Y = 0\}$. Dan is $N \in \mathcal{R}^d$ (omdat f_Y Borel-meetbaar is) en $\mathbb{P}\{Y \in N\} = 0$. Dit impliceert dat $XI_{\{Y\in C\}} = XI_{\{Y\in C\cap N^c\}}$ b.o. en dus $\mathbb{E}[XI_{\{Y\in C\}}] = \mathbb{E}[XI_{\{Y\in C\cap N^c\}}]$. De tweede verwachtingswaarde kunnen we wegens stelling 4.13 ook opschrijven als

$$\int_{-\infty}^{\infty}\int_{C\cap N^c} xf(x,y)dydx = \int_{C\cap N^c}\int_{-\infty}^{\infty} xf_{X|Y}(x|y)dxf_Y(y)dy$$

Het rechterlid is gelijk aan

$$\int_{C\cap N^c} \mathbb{E}[X\|Y=y]f_Y(y)dy = \int_{\mathbb{R}^d} I_C(y)g(y)f_Y(y)dy,$$

waar $g(y) = \mathbb{E}[X\|Y = y]$ als $f_Y(y) > 0$ en $g(y) = 0$ elders. Dus $\mathbb{E}[X\|Y] = g\circ Y$. Een toepassing van stelling 4.13 toont ten slotte dat de laatste term gelijk is aan

$$\mathbb{E}[(I_C\circ Y)\mathbb{E}[X\|Y]] = \mathbb{E}[I_{\{Y\in C\}}\mathbb{E}[X\|Y]],$$

waarmee de stelling (in het continue geval) bewezen is. □

Voorbeelden

1. Als X, Y toevalsvariabelen zijn met een gezamenlijk continue verdeling bepaald door de dichtheidsfunctie $f_{X,Y}(x,y) = \beta^{-2}e^{-y/\beta}, 0 < x < y < \infty$ en $f_{X,Y}(x,y) = 0$ elders, dan hebben we al aangetoond dat $\mathbb{E}[X\|Y = y] = y/2, y > 0$.
   Bijgevolg is $\mathbb{E}[X\|Y] = Y/2$. De toevalsvariabele Y heeft een gamma$(2,\beta)$-verdeling zodat $\mathbb{E}[Y] = 2\beta$. Het volgt nu dat $\mathbb{E}[X] = \mathbb{E}[Y/2] = \beta$.

2. We bekijken nog een keer de geometrische verdeling met parameter p. Zij X een toevalsvariabele met deze verdeling en stel $Y = I_A$, waar A de gebeurtenis is dat het eerste experiment een succes is. (X is het aantal experimenten totdat we de eerste keer succes hebben.) Dan geldt natuurlijk:

   $$\mathbb{E}[X\|Y=1] = 1 \text{ en } \mathbb{E}[X\|Y=0] = \mathbb{E}[X]+1.$$

   Dus kunnen we via stelling 4.14 concluderen dat

   $$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}[\mathbb{E}[X\|Y]] = \mathbb{P}\{Y=1\}\mathbb{E}[X\|Y=1] + \mathbb{P}\{Y=0\}\mathbb{E}[X\|Y=0]\\ &= p + (1-p)(\mathbb{E}[X]+1),\end{aligned}$$

   hetgeen impliceert dat $\mathbb{E}[X] = 1/p$ of $\mathbb{E}[X] = \infty$. Het is niet te moeilijk te zien dat $\mathbb{E}[X]$ eindig moet zijn en we kunnen dus op deze manier de verwachtingswaarde van X bepalen zonder resultaten over machtreeksen te gebruiken.

3. Beschouw een normaalverdeelde toevalsvector $(X,Y):\Omega\to\mathbb{R}^2$ met verwachting 0 en covariantie matrix $\Sigma = \begin{bmatrix} a^2 & ab\rho \\ ab\rho & b^2\end{bmatrix}$, waar $a, b > 0$ en $|\rho| < 1$.
   Dan is de conditionele verdeling van Y gegeven $X = x$ de normaalverdeling met verwachtingswaarde $\rho bx/a$ en variantie $b^2(1-\rho^2)$ (→ voorbeeld 2 op het einde van hoofdstuk 3.7). Dus, $\mathbb{E}[Y\|X = x] = \rho bx/a$ wat impliceert dat $\mathbb{E}[XY\|X = x] = x\mathbb{E}[Y\|X = x] = \rho x^2 b/a$. Dus, $\mathbb{E}[XY\|X] = \rho X^2 b/a$ en $\mathbb{E}[XY] = \mathbb{E}[X^2]\rho b/a = \rho ab$. Gezien $\mathbb{E}[X] = \mathbb{E}[Y] = 0$ en $\mathbb{E}[X^2] = a^2, \mathbb{E}[Y^2] = b^2$ (→ lemma 3.2) hebben we $\text{cov}(X,Y) = \rho ab$ en $\rho_{X,Y} = \rho$.

4.4 Verwachtingswaarden en integralen

We hebben gezien hoe men de verwachtingswaarde van een toevalsvariabele $X : \Omega \to \mathbb{R}$ op een kansruimte $(\Omega, \mathcal{F}, \mathbb{P})$ kan definiëren. Deze is natuurlijk afhankelijk van de kansmaat $\mathbb{P}$. Soms bekijken we ook meetbare ruimten $(\Omega, \mathcal{F})$, waarop verschillende kansmaten gedefinieerd zijn. In dit geval duiden we de verwachtingswaarde van een toevalsvariabele X t.o.v. een kansmaat P aan door $E_P[X]$. Men gebruikt ook een notatie als een integraal:

$$E_P[X] = \int X dP = \int_\Omega X(\omega) P(d\omega).$$

(In de maattheorie zullen we zien dat er een algemeen begrip van integraal bestaat waarvan verwachtingswaarden en de Lebesgue-integraal speciale gevallen zijn.)

We bekijken nu verwachtingswaarden op de meetbare ruimte $(\mathbb{R}^d, \mathcal{R}^d)$.
Als $g : \mathbb{R}^d \to \mathbb{R}$ Borel-meetbaar is, is dit een toevalsvariabele op de kansruimte $(\mathbb{R}^d, \mathcal{R}^d, \mu)$, waar $\mu : \mathcal{R}^d \to [0,1]$ een willekeurige kansmaat kan zijn.
We kunnen dan altijd $E_\mu[g^+] = \int_{\mathbb{R}^d} g^+(x)\mu(dx)$ en $E_\mu[g^-] = \int_{\mathbb{R}^d} g^-(x)\mu(dx)$ berekenen. Als minstens één van deze twee verwachtingswaarden eindig is, dan is g μ-semi-integreerbaar en we hebben dat

$$E_\mu[g] = \int_{\mathbb{R}^d} g(x)\mu(dx) = \int_{\mathbb{R}^d} g^+(x)\mu(dx) - \int_{\mathbb{R}^d} g^-(x)\mu(dx).$$

Indien $d = 1$, schrijven we meestal $\int_{-\infty}^\infty g(x)\mu(dx)$ in plaats van $\int_\mathbb{R} g(x)\mu(dx)$.
De volgende stelling toont dat gegeven een toevalsvector $X : \Omega \to \mathbb{R}^d$, we verwachtingswaarden van toevalsvariabelen van de gedaante $g \circ X : \Omega \to \mathbb{R}$ als μ-integralen kunnen berekenen, waar $\mu = \mathbb{P}_X$ de verdeling van X is.

Stelling 4.15 *Zij $X : \Omega \to \mathbb{R}^d$ een toevalsvector met verdeling μ en $g : \mathbb{R}^d \to \mathbb{R}$ Borel-meetbaar. Dan bestaat $\mathbb{E}[g \circ X]$ als en slechts als g μ-semi-integreerbaar is. In dit geval hebben we,*

$$\mathbb{E}[g \circ X] = \int_\Omega g(X(\omega))\mathbb{P}(d\omega) = \int_{\mathbb{R}^d} g(x)\mu(dx).$$

Bewijs. Het is voldoende de stelling voor niet-negatieve Borel-meetbare functies te bewijzen. Dus zij $g \geq 0$. Dan is $g \circ X$ een niet-negatieve toevalsvariabele en dus geldt er wegens stelling 4.3,

$$\mathbb{E}[g \circ X] = \int_0^\infty \mathbb{P}\{g \circ X \geq t\} dt = \int_0^\infty \mathbb{P}\{X \in g^{-1}[t, \infty[\} dt.$$

Aangezien $\mu(B) = \mathbb{P}_X(B) = \mathbb{P}\{X \in B\}, B \in \mathcal{R}$, kunnen we de laatste integraal ook schrijven als

$$\int_0^\infty \mu(g^{-1}([t, \infty[)) dt = \int_0^\infty \mu\{g \geq t\} dt = E_\mu[g] = \int_{\mathbb{R}^d} g(x)\mu(dx),$$

waarmee de stelling bewezen is. □

Als $X : \Omega \to \mathbb{R}$ een toevalsvariabele is waarvoor $\mathbb{E}[X]$ bestaat, krijgen we als speciaal geval van de bovenstaande stelling dat $\mathbb{E}[X] = \int_{-\infty}^\infty x\mu(dx)$, hetgeen weer toont dat de verwachtingswaarde van een toevalsvariabele X bepaald is door de verdeling $\mu = \mathbb{P}_X$ van deze toevalsvariabele.

Hoofdstuk 5

Enkele belangrijke limietstellingen

5.1 De zwakke wet van de grote getallen

Zij $X_n, n \geq 1$ een rij onafhankelijke, identiek verdeelde toevalsvariabelen (waar de laatste voorwaarde betekent dat geldt $\mathbb{P}_{X_1} = \mathbb{P}_{X_k}, k \geq 1$). Stel $S_n = \sum_{k=1}^n X_k, n \geq 1$ en bekijk het gemiddelde S_n/n van de eerste n toevalsvariabelen. Uit onze ervaring in het dagelijkse leven (bv, als we lotto spelen of naar casino's gaan), weten we dat als n groot wordt, deze gemiddelden zich beginnen te stabiliseren, dwz deze toevalsvariabelen convergeren.

Zoals in de meeste andere richtingen van de wiskunde, bestaan er in de kanstheorie verschillende soorten van convergentie. We bekijken eerst convergentie in kans.

Definitie 5.1 (Convergentie in kans) *Zij $Z_n, n \geq 1$ een rij toevalsvariabelen. Dan zeggen we dat Z_n in kans naar een toevalsvariabele Z convergeert (Notatie: $Z_n \xrightarrow{\mathbb{P}} Z$) indien geldt:*

$$\mathbb{P}\{|Z_n - Z| \geq \epsilon\} \to 0 \text{ als } n \to \infty, \ \forall\, \epsilon > 0.$$

Voorbeeld

Zij $X_n, n \geq 1$ zoals boven, i.e. onafhankelijk en identiek verdeeld. Onderstel dat $\mathbb{E}[X_1^2] < \infty$. Dan geldt: $S_n/n \xrightarrow{\mathbb{P}} \mathbb{E}[X_1]$.

Om dit te tonen, noteren we eerst dat $\mathbb{E}[S_n/n] = \sum_{i=1}^n \mathbb{E}[X_i]/n = \mathbb{E}[X_1]$. Dus volgt uit Chebyshevs ongelijkheid en stelling 4.10 dat

$$\mathbb{P}\{|S_n/n - \mathbb{E}[X_1]| \geq \epsilon\} \leq \mathrm{Var}(S_n/n)/\epsilon^2 = \mathrm{Var}(X_1)/(n\epsilon^2) \to 0.$$

Opmerkingen

1. In het bovenstaande voorbeeld hebben we formeel niet nodig dat het tweede moment van X_1 eindig is. Dit was wel nodig om Chebyshevs ongelijkheid te kunnen gebruiken. Dus is het voor de hand liggend te vragen of er een ander bewijs bestaat, zodat we de bovenstaande convergentie in kans onder de voorwaarde $\mathbb{E}[|X_1|] < \infty$ kunnen verkrijgen.

2. Verder is het bovenstaande argument nog goed als we in plaats van "onafhankelijk" veronderstellen dat de toevalsvariabelen **paarsgewijs** onafhankelijk zijn.

3. Algemener kan men zich ook afvragen wanneer $S_n/n \xrightarrow{\mathbb{P}} a$ geldt, waar a een constante is (die niet noodzakelijk een verwachtingswaarde is).

Om deze problemen op te lossen, bewijzen we eerst een resultaat dat aantoont wanneer voor paarsgewijs onafhankelijke (niet noodzakelijk identiek verdeelde) toevalsvariabelen X_n geldt $S_n/a_n \xrightarrow{\mathbb{P}} 0$, waar $a_n \nearrow \infty$. Met meer gevorderde technieken uit de kanstheorie kan men ook tonen dat de voorwaarden van onze stelling nodig zijn als de toevalsvariabelen X_n onafhankelijk zijn.

Stelling 5.1 *Zij $X_n, n \geq 1$ een rij paarsgewijs onafhankelijke toevalsvariabelen en zij $S_n = \sum_{j=1}^n X_j$. Veronderstel dat voor een rij $a_n \nearrow \infty$ geldt:*
(i) $\lim_{n\to\infty} \sum_{j=1}^n \mathbb{P}\{|X_j| \geq a_n\} = 0$
(ii) $\lim_{n\to\infty} a_n^{-1} \sum_{j=1}^n \mathbb{E}[X_j I_{\{|X_j|<a_n\}}] = 0$ *en*
(iii) $\lim_{n\to\infty} a_n^{-2} \sum_{j=1}^n \mathrm{Var}(X_j I_{\{|X_j|<a_n\}}) = 0$
Dan volgt: $S_n/a_n \xrightarrow{\mathbb{P}} 0$.

Bewijs Stel $X_{n,j} := X_j I_{\{|X_j|<a_n\}}, 1 \leq j \leq n, S_{n,n} := \sum_{j=1}^n X_{n,j}, n \geq 1$. Dan geldt:

$$\begin{aligned} \mathbb{P}\{|S_n| \geq \epsilon a_n\} &\leq \mathbb{P}\{|S_{n,n}| \geq \epsilon a_n\} + \mathbb{P}(\bigcup_{j=1}^n \{X_{n,j} \neq X_j\}) \\ &\leq \mathbb{P}\{|S_{n,n}| \geq \epsilon a_n\} + \sum_{j=1}^n \mathbb{P}\{|X_j| \geq a_n\}, \end{aligned}$$

waar de laatste term wegens (i) naar nul convergeert. Verder impliceert (ii) voor grote n,

$$\mathbb{P}\{|S_{n,n}| \geq \epsilon a_n\} \leq \mathbb{P}\{|S_{n,n} - \mathbb{E}[S_{n,n}]| \geq \epsilon a_n/2\}.$$

Als we nu zoals boven de ongelijkheid van Chebyshev en stelling 4.10 gebruiken waar we rekening houden met het feit dat de toevalsvariabelen $X_{n,j}, 1 \leq j \leq n$ paarsgewijs onafhankelijk zijn als functies van de variabelen $X_j, 1 \leq j \leq n$, volgt uit (iii),

$$\begin{aligned} \mathbb{P}\{|S_{n,n} - \mathbb{E}[S_{n,n}]| \geq \epsilon a_n/2\} &\leq 4\epsilon^{-2}\mathrm{Var}(S_{n,n})/a_n^2 \\ &= 4\epsilon^{-2} a_n^{-2} \sum_{j=1}^n \mathrm{Var}(X_j I_{\{|X_j|<a_n\}}) \to 0 \end{aligned}$$

en de stelling is bewezen. □

Als we de bovenstaande stelling met $a_n = n$ toepassen, verkrijgen we de zwakke wet van de grote getallen (= weak law of large numbers, WLLN).

Stelling 5.2 (WLLN) *Zij $X_n, n \geq 1$ een rij paarsgewijs onafhankelijke toevalsvariabelen met identieke verdelingen. Veronderstel dat*
(i) $n\mathbb{P}\{|X_1| \geq n\} \to 0$ *als* $n \to \infty$.
(ii) $\mathbb{E}[X_1 I_{\{|X_1|<n\}}] \to 0$ *als* $n \to \infty$.
Dan geldt: $S_n/n \xrightarrow{\mathbb{P}} 0$.

Bewijs (i) en (ii) stemmen met de voorwaarden (i) en (ii) van stelling 5.1 (voor het speciale geval waar de toevalsvariabelen identieke verdelingen hebben en $a_n = n$) overeen. Voorwaarde (iii) betekent in dit geval dat

$$\mathrm{Var}(X_1 I_{\{|X_1|\le n\}})/n \to 0 \text{ als } n \to \infty.$$

Om dit te bewijzen noteren we dat

$$\begin{aligned}
\mathrm{Var}(X_1 I_{\{|X_1|<n\}}) &\le \mathbb{E}[X_1^2 I_{\{|X_1|<n\}}] = \sum_{m=1}^{n} \mathbb{E}[X_1^2 I_{\{m-1\le|X_1|<m\}}] \\
&\le \sum_{m=1}^{n} m^2 \mathbb{P}\{m-1 \le |X_1| < m\} \le 2\sum_{m=1}^{n}(\sum_{k=1}^{m} k)\mathbb{P}\{m-1\le|X_1|<m\} \\
&= 2\sum_{k=1}^{n} k \sum_{m=k}^{n} \mathbb{P}\{m-1\le|X_1|<m\} = 2\sum_{k=1}^{n} k\mathbb{P}\{k-1\le|X_1|<n\} \\
&\le 2+4\sum_{k=2}^{n}(k-1)\mathbb{P}\{|X_1|\ge k-1\} = 2+4\sum_{m=1}^{n-1} m\mathbb{P}\{|X_1|\ge m\}.
\end{aligned}$$

Wegens (i) geldt: $\frac{1}{n}\sum_{m=1}^{n} m\mathbb{P}\{|X_1|\ge m\} \to 0$ als $n\to\infty$ en de WLLN is bewezen.□

Gevolg 5.1 *Zij $X_n, n \ge 1$ een rij paarsgewijs onafhankelijke, identiek verdeelde toevalsvariabelen. Veronderstel dat $\mathbb{E}[|X_1|] < \infty$. Dan geldt: $S_n/n \xrightarrow{\mathbb{P}} \mathbb{E}[X_1]$.*

Bewijs Het is voldoende het gevolg voor het geval $\mathbb{E}[X_1] = 0$ te bewijzen. (Vervang indien nodig de toevalsvariabelen X_n door $X_n - \mathbb{E}[X_1], n \ge 1$.)
Wegens de monotonie van de verwachtingswaarde geldt:

$$n\mathbb{P}\{|X_1|\ge n\} = \mathbb{E}[nI_{\{|X_1|\ge n\}}] \le \mathbb{E}[|X_1|I_{\{|X_1|\ge n\}}]$$

Vermits geldt $|X_1|I_{\{|X_1|\ge n\}} \to 0$ en $|X_1|I_{\{|X_1|\ge n\}} \le |X_1|$ kunnen we via de stelling van de gedomineerde convergentie concluderen dat $\mathbb{E}[|X_1|I_{\{|X_1|\ge n\}}] \to 0$ en bijgevolg is aan (i) voldaan. Gezien $\mathbb{E}[X_1] = 0$ hebben we $|\mathbb{E}[X_1 I_{\{|X_1|<n\}}]| = |\mathbb{E}[X_1 I_{\{|X_1|\ge n\}}]| \le \mathbb{E}[|X_1|I_{\{|X_1|\ge n\}}] \to 0$, waaruit blijkt dat ook aan voorwaarde (ii) voldaan is. Dus $S_n/n \xrightarrow{\mathbb{P}} 0$ en het gevolg is bewezen. □

Opmerking Er bestaan onafhankelijke en identiek verdeelde toevalsvariabelen $X_n, n \ge 1$ zodanig dat $\mathbb{E}[|X_1|] = \infty$ en $S_n/n \xrightarrow{\mathbb{P}} 0$.

Voorbeeld: Stel $c = \sum_{n=2}^{\infty} \frac{1}{n^2 \log n}$ en zij μ de discrete kansmaat gegeven door

$$\mu\{n\} = \mu\{-n\} = \frac{1}{2cn^2 \log n}, n = 2, 3, \ldots$$

Als $X_n, n \ge 1$ een rij onafhankelijke toevalsvariabelen met deze verdeling is, geldt voor $n \ge 2$,

$$\begin{aligned}
n\mathbb{P}\{|X_1|\ge n\} &= n\sum_{m=n}^{\infty} \frac{1}{cm^2 \log m} \le \frac{n}{c\log n}\sum_{m=n}^{\infty} m^{-2} \\
&\le \frac{n}{c\log n}\int_{n-1}^{\infty} x^{-2}dx \le \frac{2}{c\log n} \to 0
\end{aligned}$$

Gezien μ symmetrisch is (i.e. $\mu(A) = \mu(-A), A \in \mathcal{R}$) en bijgevolg $\mathbb{P}_{X_1} = \mathbb{P}_{-X_1}$ volgt onmiddellijk dat $\mathbb{E}[X_1 I_{\{|X_1|<n\}}] = 0\ \forall n \ge 1$. Dus is aan de voorwaarden van de WLLN voldaan en het geldt: $S_n/n \xrightarrow{\mathbb{P}} 0$. Maar $\mathbb{E}[X_1^+] = \mathbb{E}[X_1^-] = \frac{1}{2c}\sum_{n=2}^{\infty} \frac{1}{n\log n} = \infty$.

5.2 De sterke wet van de grote getallen

Definitie 5.2 (Convergentie bijna overal) *Zij $Z_n, n \geq 1$ een rij toevalsvariabelen. Dan zeggen we dat Z_n bijna overal (of met kans 1) naar een toevalsvariabele Z convergeert indien geldt:*

$$\mathbb{P}\{\omega : Z_n(\omega) \to Z(\omega)\} = 1.$$

(Notatie: $Z_n \to Z$ b.o.)

Het volgende lemma geeft een karakterisering van deze convergentie.

Lemma 5.1 *Als $Z, Z_n, n \geq 1$ toevalsvariabelen zijn, dan zijn equivalent:*

(A) $Z_n \to Z$ *b.o.*

(B) $\lim_{n\to\infty} \mathbb{P}(\bigcup_{m=n}^{\infty}\{|Z_m - Z| \geq \epsilon\}) = 0 \ \ \forall \epsilon > 0.$

Bewijs Stel $\Delta_\infty := \limsup_{n\to\infty} |Z_n - Z|$. Dan hebben we natuurlijk:

$$\text{(A)} \iff \mathbb{P}\{\Delta_\infty = 0\} = 1 \iff \mathbb{P}\{\Delta_\infty \geq \epsilon\} = 0 \, \forall \, \epsilon > 0.$$

De niet-triviale implicatie "$\Leftarrow$" in de laatste equivalentie volgt wegens

$$\mathbb{P}\{\Delta_\infty \neq 0\} = \mathbb{P}\left(\bigcup_{n=1}^{\infty}\{\Delta_\infty \geq 1/n\}\right) \leq \sum_{n=1}^{\infty} \mathbb{P}\{\Delta_\infty \geq 1/n\}.$$

Verder geldt er $\Delta_n \searrow \Delta_\infty$, waar $\Delta_n := \sup_{m\geq n} |Z_m - Z|, n \geq 1$.
Bijgevolg, hebben we: $\{\Delta_n \geq \epsilon\} \searrow \{\Delta_\infty \geq \epsilon\}$ en dus $\mathbb{P}\{\Delta_\infty \geq \epsilon\} = \lim_{n\to\infty} \mathbb{P}\{\Delta_n \geq \epsilon\}$. Dit betekent dat

$$\boxed{\text{(A)} \iff \mathbb{P}\{\Delta_n \geq \epsilon\} \to 0 \text{ als } n \to \infty \ \forall \, \epsilon > 0.}$$

Aangezien

$$\mathbb{P}(\bigcup_{m=n}^{\infty}\{|Z_m - Z| \geq \epsilon\}) \leq \mathbb{P}\{\Delta_n \geq \epsilon\} \leq \mathbb{P}(\bigcup_{m=n}^{\infty}\{|Z_m - Z| \geq \epsilon/2\}) \ \forall \, \epsilon > 0,$$

volgt onmiddellijk dat (A) en (B) equivalent zijn. □

Gevolg 5.2 *Als $Z, Z_n, n \geq 1$ toevalsvariabelen zijn zodanig dat $Z_n \to Z$ b.o., hebben we: $Z_n \xrightarrow{\mathbb{P}} Z$.*

Bewijs Als $Z_n \to Z$ b.o., volgt via lemma 5.1 en de monotonie van de kansmaat $\mathbb{P}$:

$$\mathbb{P}\{|Z_n - Z| \geq \epsilon\} \leq \mathbb{P}(\bigcup_{m=n}^{\infty}\{|Z_m - Z| \geq \epsilon\}) \to 0, \forall \epsilon > 0.$$

Dus bestaat ook convergentie in kans en het lemma is bewezen. □

We geven nu een voorbeeld waar convergentie in kans bestaat, maar niet bijna overal.

Voorbeeld. Zij U een uniform(0,1)-verdeelde toevalsvariabele en stel

$$X_1 = I_{]0,1]} \circ U, X_{2^k+j-1} = I_{](j-1)/2^k, j/2^k]} \circ U, 1 \leq j \leq 2^k, k = 1, 2 \ldots$$

Dan geldt: $\mathbb{P}\{X_n \neq 0\} \to 0$ en dus $X_n \xrightarrow{\mathbb{P}} 0$. Maar $\limsup_{n\to\infty} X_n = I_{]0,1]}$, waaruit blijkt dat de convergentie niet bijna overal kan bestaan.

Stelling 5.3 (De sterke wet van de grote getallen (= SLLN))
Zij $X_n, n \geq 1$ een rij paarsgewijs onafhankelijke, identiek verdeelde toevalsvariabelen zodanig dat $\mathbb{E}[|X_1|] < \infty$. Stel $S_n = \sum_{j=1}^n X_j, n \geq 1$. Dan geldt: $S_n/n \to \mathbb{E}[X_1]$ b.o.

Opmerking Men kan tonen dat S_n/n **niet** bijna overal kan convergeren als $\mathbb{E}[|X_1|] = \infty$. Dus weten we dat als S_n/n bijna overal convergent is, het eerste moment van X_1 moet bestaan en de limiet is dan gelijk aan $\mathbb{E}[X_1]$. In bepaalde gevallen is het wel nog mogelijk dat S_n/n in kans convergeert hoewel $\mathbb{E}[|X_1|] = \infty$ (zie voorbeeld in 5.1).

Bewijs (Als $\mathbb{E}[X_1^2] < \infty$.)
We bekijken eerst niet-negatieve toevalsvariabelen. Dus, veronderstel dat $X_n \geq 0, n \geq 1$. Merk op dat dan S_n monotoon niet-dalend is.
(STAP 1) Stel $n_k = k^2$. Gezien de toevalsvariabelen X_j paarsgewijs onafhankelijk zijn, hebben we

$$\mathrm{Var}(S_{n_k}) = n_k \mathrm{Var}(X_1) \leq n_k \mathbb{E}X_1^2.$$

Als we de ongelijkheid van Chebyshev toepassen volgt er voor $\epsilon > 0$,

$$\mathbb{P}\{|S_{n_k}/n_k - \mathbb{E}X_1| > \epsilon\} \leq \epsilon^{-2}\mathbb{E}X_1^2/n_k,$$

hetgeen natuurlijk impliceert dat $\sum_{k=1}^\infty \mathbb{P}\{|S_{n_k}/n_k - \mathbb{E}X_1| > \epsilon\} < \infty, \epsilon > 0$.
Dan volgt er voor $\epsilon > 0$,

$$\mathbb{P}(\bigcup_{k=N}^\infty \{|S_{n_k}/n_k - \mathbb{E}X_1| > \epsilon\} \leq \sum_{k=N}^\infty \mathbb{P}\{|S_{n_k}/n_k - \mathbb{E}X_1| > \epsilon\} \to 0 \text{ als } N \to \infty,$$

en dus (Lemma 5.1)

$$S_{n_k}/n_k \to \mathbb{E}X_1 \text{ b.o.}$$

(STAP 2) Wegens de monotonie van S_n kunnen we voor $n_{k-1} \leq n \leq n_k$ concluderen:

$$\frac{n_{k-1}}{n_k}\frac{S_{n_{k-1}}}{n_{k-1}} \leq \frac{S_n}{n} \leq \frac{n_k}{n_{k-1}}\frac{S_{n_k}}{n_k}$$

Merk op dat $n_k/n_{k-1} \to 1$ als $k \to \infty$. Dus kunnen we onmiddellijk uit de convergentie van de deelrij S_{n_k}/n_k concluderen dat $S_n/n \to \mathbb{E}X_1$ b.o. als $n \to \infty$. Daarmee is de sterke wet van de grote getallen bewezen voor niet-negatieve toevalsvariabelen.
(STAP 3) Stel $X_n^+ = X_n \vee 0$ en $X_n^- = (-X_n) \vee 0, n \geq 1$. De twee rijen $X_n^+, n \geq 1$ en $X_n^-, n \geq 1$ zijn paarsgewijs onafhankelijk en we hebben natuurlijk:

$$S_n/n = \sum_{j=1}^n X_j^+/n - \sum_{j=1}^n X_j^-/n.$$

Als we de sterke wet van de grote getallen voor niet-negatieve toevalsvariabelen toepassen, zien we dat met kans 1,

$$S_n/n \to \mathbb{E}X_1^+ - \mathbb{E}X_1^- = \mathbb{E}X_1 \text{ als } n \to \infty$$

en de stelling is bewezen. □

Opmerking Het speciale geval van de sterke wet van de grote getallen wat we pas hebben bewezen, impliceert dat relatieve frequenties (zie 1.1) bijna overal convergeren. Om dit in te zien, bekijken we een rij onafhankelijke gebeurtenissen A_n met $\mathbb{P}(A_n) = p$ en stellen $X_n = I_{A_n}$. Dan zijn de X_n onafhankelijke Bernoulli(p)-variabelen met $\mathbb{E}[X_n^2] = \mathbb{E}[X_n] = p < \infty$. Het volgt dat $n_A/n = \sum_{i=1}^n X_i/n$ bijna overal naar $p = \mathbb{P}(A_1)$ convergeert.

5.3 De centrale limietstelling

We gaan in dit hoofdstuk het misschien belangrijkste resultaat uit de kanstheorie te bewijzen, de centrale limietstelling. Dit resultaat toont dat bepaalde kansen betreffende sommen van onafhankelijke toevalsvariabelen X_n naar kansen ten opzichte van de normaalverdeling convergeren. Dus als $S_n = \sum_{i=1}^n X_i$ en n groot genoeg is, kunnen we zowel in het discrete geval als in het absoluut continue geval kansen van het type $\mathbb{P}\{S_n \in A\}$ benaderen door "normale" kansen $\mathbb{P}\{T_n \in A\}$, waar T_n een normaalverdeelde variabele is.

Formeel gaat het hier over "convergentie in distributie" die we nu gaan definiëren.

Definitie 5.3 (Convergentie in distributie) *Zij $Z_n, n \geq 1$ een rij toevalsvariabelen. Dan zeggen we dat Z_n naar een toevalsvariabele Z in distributie (of in verdeling) convergeert indien er voor de verdelingsfuncties F_n van Z_n en F van Z geldt:*

$$F_n(x) \to F(x), \forall\, x \notin D(F),$$

waar $D(F) = \{x \in \mathbb{R} : F \text{ is niet continu in } x\}$.

We hebben al in hoofdstuk 2 getoond (zie stelling 2.5) dat $D(F) = \{x \in \mathbb{R} : \mathbb{P}\{Z = x\} > 0\}$ en het is evident dat deze verzameling hoogstens aftelbaar is (zie bewijs van stelling 1.4).
De volgende stelling geeft een karakterisatie van convergentie in distributie die we in het vervolg altijd door $\stackrel{d}{\to}$ gaan aanduiden.

Stelling 5.4 *De volgende drie uitspraken zijn equivalent:*

(a) $Z_n \stackrel{d}{\to} Z$ *als* $n \to \infty$.

(b) Als $g : \mathbb{R} \to \mathbb{R}$ continu en begrensd is, geldt er $\mathbb{E}[g \circ Z_n] \to \mathbb{E}[g \circ Z]$ als $n \to \infty$.

(c) Als $g : \mathbb{R} \to \mathbb{R}$ uniform continu en begrensd is, geldt er $\mathbb{E}[g \circ Z_n] \to \mathbb{E}[g \circ Z]$ als $n \to \infty$.

Bewijs $\boxed{\text{(a)} \Rightarrow \text{(b)}}$ Zij $0 < \epsilon < 1/3$. Gezien $F(x) \to 1$ als $x \to \infty$ en $F(x) \to 0$ als $x \to -\infty$ (zie stelling 2.5) en $D(F)$ hoogstens aftelbaar is, kunnen we een $K > 0$ vinden zodanig dat $-K, K \notin D(F)$ en $F(K) = \mathbb{P}\{Z \leq K\} \geq 1 - \epsilon, F(-K) = \mathbb{P}\{Z \leq -K\} \leq \epsilon$, hetgeen natuurlijk impliceert dat

$$\mathbb{P}\{-K < Z \leq K\} = F(K) - F(-K) \geq 1 - 2\epsilon. \tag{5.1}$$

Aangezien $F_n(K) \to F(K)$ en $F_n(-K) \to F(-K)$ als $n \to \infty$, volgt er ook dat vanaf een bepaalde $n_0 = n_0(K) \geq 1$ geldt,

$$\mathbb{P}\{-K < Z_n \leq K\} = F_n(K) - F_n(-K) \geq 1 - 3\epsilon. \tag{5.2}$$

De restrictie van de continue functie g tot het compacte interval $[-K, K]$ is *uniform* continu. Dus bestaat er een $\delta = \delta_\epsilon > 0$ zodanig dat

$$|g(x_1) - g(x_2)| \leq \epsilon \text{ als } |x_1 - x_2| \leq \delta \text{ en } x_1, x_2 \in [-K, K]. \tag{5.3}$$

Stel $m = \max\{j \geq 0 : j\delta < 2K\}$.
Beschouw de volgende partitie $I_j, -m \leq j \leq m + 1$ van $[-K, K]$,

$$I_{-m} = [-K, -m\delta/2[, I_j = [(j-1)\delta/2, j\delta/2[, -m+1 \leq j \leq m, I_{m+1} = [m\delta/2, K].$$

Kies in elk interval I_j ($j \neq -m, m+1$) een getal $t_j \notin D(F)$ (wat mogelijk is omdat $D(F)$ hoogstens aftelbaar is). Stel verder $t_{m+1} = K, t_{-m} = -K$. Merk op dat dan geldt

$$0 < t_{j+1} - t_j \leq \delta, -m \leq j \leq m \text{ en } t_j \notin D(F), -m \leq j \leq m+1.$$

Stel $h(x) = \sum_{j=-m}^{m} g(t_j) I_{]t_j, t_{j+1}]}(x), x \in \mathbb{R}$. Wegens (5.3) geldt dan

$$\sup_{x \in]-K,K]} |g(x) - h(x)| \leq \epsilon \tag{5.4}$$

Aangezien $h(x) = 0$ als $x \notin]-K, K]$ kunnen we nu concluderen dat

$$\begin{aligned} |\mathbb{E}[g \circ Z_n] - \mathbb{E}[g \circ Z]| &\leq |\mathbb{E}[h \circ Z_n] - \mathbb{E}[h \circ Z]| + |\mathbb{E}[g \circ Z_n] - \mathbb{E}[h \circ Z_n]| \\ &\quad + |\mathbb{E}[g \circ Z] - \mathbb{E}[h \circ Z]| \\ &\leq |\mathbb{E}[h \circ Z_n] - \mathbb{E}[h \circ Z]| + 2\epsilon + \mathbb{E}[|g \circ Z_n| I_{\{Z_n \notin]-K,K]\}}] \\ &\quad + \mathbb{E}[|g \circ Z| I_{\{Z \notin]-K,K]\}}] \end{aligned}$$

Kies een $M > 0$ zodanig dat $|g(x)| \leq M, x \in \mathbb{R}$ (de functie g is begrensd). Dan volgt wegens (5.1), (5.2) en (5.4) dat voor $n \geq n_0$ geldt,

$$|\mathbb{E}[g \circ Z_n] - \mathbb{E}[g \circ Z]| \leq |\mathbb{E}[h \circ Z_n] - \mathbb{E}[h \circ Z]| + (2 + 5M)\epsilon. \tag{5.5}$$

We kunnen ϵ willekeurig klein kiezen en dus is het voldoende te tonen dat

$$\mathbb{E}[h \circ Z_n] \to \mathbb{E}[h \circ Z] \text{ als } n \to \infty.$$

h is een elementaire functie en dus hebben we

$$\mathbb{E}[h \circ Z_n] = \sum_{j=-m}^{m} g(t_j)(F_n(t_{j+1}) - F_n(t_j)),$$

waar $F_n(t_j) \to F(t_j), -m \leq j \leq m+1$ (omdat $t_j \notin D(F)$).
We zien nu dat

$$\mathbb{E}[h \circ Z_n] \to \sum_{j=-m}^{m} g(t_j)(F(t_{j+1}) - F(t_j)) = \mathbb{E}[h \circ Z],$$

en daarmee is (b) bewezen, hetgeen natuurlijk ook (c) impliceert.

$\boxed{\text{(c)} \Rightarrow \text{(a)}}$ Neem $x \notin D(F)$. We moeten tonen dat

$$\mathbb{P}\{Z_n \leq x\} \to \mathbb{P}\{Z \leq x\} \text{ als } n \to \infty.$$

Zij $\epsilon > 0$. Kies een uniform continue functie $g_{x,\epsilon} : \mathbb{R} \to \mathbb{R}$ zodanig dat

$$I_{]-\infty,x]} \leq g_{x,\epsilon} \leq I_{]-\infty,x+\epsilon]}.$$

Dan geldt wegens monotonie van de verwachtingswaarde en (b) dat

$$\limsup_{n\to\infty} \mathbb{P}\{Z_n \leq x\} \leq \limsup_{n\to\infty} \mathbb{E}[g_{x,\epsilon} \circ Z_n] = \mathbb{E}[g_{x,\epsilon} \circ Z] \leq \mathbb{P}\{Z \leq x + \epsilon\}.$$

Dit geldt voor elke $\epsilon > 0$. Aangezien $\mathbb{P}\{Z \le x + \epsilon\} \to \mathbb{P}\{Z \le x\}$ als $\epsilon \searrow 0$, kunnen we concluderen dat

$$\limsup_{n\to\infty} \mathbb{P}\{Z_n \le x\} \le \mathbb{P}\{Z \le x\}.$$

Analoog volgt:

$$\liminf_{n\to\infty} \mathbb{P}\{Z_n \le x\} \ge \liminf_{n\to\infty} I\!E[g_{x-\epsilon,\epsilon} \circ Z_n] = I\!E[g_{x-\epsilon,\epsilon} \circ Z] \ge \mathbb{P}\{Z \le x - \epsilon\}, \epsilon > 0.$$

Aangezien $\lim_{\epsilon \searrow 0} \mathbb{P}\{Z \le x - \epsilon\} = \mathbb{P}\{Z < x\} = \mathbb{P}\{Z \le x\}$ als $x \notin D(F)$, geldt er ook

$$\liminf_{n\to\infty} \mathbb{P}\{Z_n \le x\} \ge \mathbb{P}\{Z \le x\},$$

waarmee (a) bewezen is. $\square$

Opmerking Het is gekend dat er een C^∞-functie $0 \le f \le 1$ bestaat zodat

$$f(x) = 1, x \le 0 \text{ en } f(x) = 0, x \ge 1.$$

We kunnen dan in het bewijs "(c) $\Rightarrow$ (a)" ook de functie $g_{x,\epsilon}(z) = f((z - x)/\epsilon), z \in \mathbb{R}$ nemen. Deze is een C^∞-functie waarvoor ook alle afgeleiden begrensd zijn (omdat $g_{x,\epsilon}^{(m)} \ne 0$ voor $m \ge 1$ alleen op het compact interval $[x, x + \epsilon]$ mogelijk is).

Dit betekent dat we convergentie in distributie hebben indien $I\!E[g \circ Z_n] \to I\!E[g \circ Z]$ voor alle begrensde functies $g : \mathbb{R} \to \mathbb{R}$ waarvoor alle afgeleiden bestaan en begrensd zijn.

Als Z een continue verdelingsfunctie heeft, volgt uit de definitie van convergentie in distributie dat de verdelingsfuncties F_n van Z_n puntsgewijs naar deze van Z convergeren. In dit geval kan men nog meer bewijzen, namelijk dat de convergentie uniform is.

Stelling 5.5 *Zij $Z_n, n \ge 1$ een rij toevalsvariabelen die in distributie naar een toevalsvariabele Z convergeert. Als $F(x) = \mathbb{P}\{Z \le x\}, x \in \mathbb{R}$ continu is, hebben we voor de verdelingsfuncties $F_n(x) = \mathbb{P}\{Z_n \le x\}, x \in \mathbb{R}$ van Z_n,*

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \to 0 \text{ als } n \to \infty.$$

Bewijs Zij $0 < \epsilon < 1/2$ en zij $m = m_\epsilon \ge 1$ het unieke natuurlijke getal zodanig dat

$$m\epsilon < 1 \le (m + 1)\epsilon.$$

Kies $x_1 < \ldots < x_m$ zodanig dat $F(x_i) = i\epsilon, 1 \le i \le m$. (Deze bestaan omdat F continu is.) Stel $x_0 = -\infty, x_{m+1} = \infty$. Dan hebben we natuurlijk,

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| = \max_{1\le i\le m+1} \sup_{x_{i-1}\le x\le x_i} |F_n(x) - F(x)| =: \max_{1\le i\le m+1} \Delta_{n,i},$$

waar we $F_n(-\infty) = F(-\infty) = 0$ en $F_n(\infty) = F(\infty) = 1$ stellen.

Wegens de monotonie van F en F_n hebben we voor $1 \le i \le m + 1$,

$$\begin{aligned} \Delta_{n,i} &\le (F_n(x_i) - F(x_{i-1})) \vee (F(x_i) - F_n(x_{i-1})) \\ &\le [(F_n(x_i) - F(x_i)) \vee (F(x_{i-1}) - F_n(x_{i-1}))] + \epsilon, \end{aligned}$$

en bijgevolg geldt er,

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \le \max_{0\le i\le m+1} |F_n(x_i) - F(x_i)| + \epsilon = \max_{1\le i\le m} |F_n(x_i) - F(x_i)| + \epsilon.$$

Aangezien $F_n(x_i) \to F(x_i)$ als $n \to \infty$ $(1 \le i \le m)$, zien we dat

$$\limsup_{n\to\infty} \|F_n - F\|_\infty \le \epsilon.$$

Dit impliceert dat de convergentie uniform is omdat we ϵ willekeurig klein kunnen kiezen. □

We noteren nog dat convergentie in kans convergentie in distributie impliceert. Dus veronderstel dat $Z_n \xrightarrow{\mathbb{P}} Z$ en zij $g : \mathbb{R} \to \mathbb{R}$ een uniform continue en begrensde functie. Kies voor $\epsilon > 0$, $\delta = \delta_\epsilon > 0$ zodanig dat

$$|x_1 - x_2| < \delta \Rightarrow |g(x_1) - g(x_2)| < \epsilon.$$

Dan volgt er onmiddellijk,

$$|\mathbb{E}[g(Z_n)] - \mathbb{E}[g(Z)]| \le 2\|g\|_\infty \mathbb{P}\{|Z_n - Z| \ge \delta\} + \epsilon,$$

en dus wegens convergentie in kans dat

$$\limsup_{n\to\infty} |\mathbb{E}[g(Z_n)] - \mathbb{E}[g(Z)]| \le \epsilon.$$

Dit impliceert convergentie in distributie omdat we weer ϵ willekeurig klein kunnen kiezen. Aangezien convergentie bijna overal ook convergentie in kans impliceert, hebben we het volgende schema: $\boxed{Z_n \to Z \text{ b.o.} \Rightarrow Z_n \xrightarrow{\mathbb{P}} Z \Rightarrow Z_n \xrightarrow{d} Z.}$

We zijn nu in staat de centrale limietstelling te bewijzen. We doen dit eerst onder een sterkere voorwaarde ($\mathbb{E}|X_1|^3 < \infty$) omdat het idee van het bewijs dan iets duidelijker te zien is. Op het einde van dit hoofdstuk tonen we dan de centrale limietstelling onder de zwakkere voorwaarde dat het tweede moment van X_1 eindig is. (In meer gevorderde cursussen kanstheorie zie je dat deze voorwaarde ook nodig is voor de centrale limietstelling.)

Stelling 5.6 (De centrale limietstelling (= CLT))
Zij $X_n, n \ge 1$ een rij onafhankelijke identiek verdeelde toevalsvariabelen met eindige tweede momenten. Stel $S_n := \sum_{i=1}^n X_i$ en $\mu = \mathbb{E}[X_1], \sigma^2 = \mathrm{Var}(X_1)$. Zij verder Y een standaardnormaalverdeelde toevalsvariabele. Als $\sigma^2 > 0$, hebben we,

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} Y.$$

Opmerking De CLT in verband met stelling 5.5 impliceert onmiddellijk dat als $I \subset \mathbb{R}$ een interval is, T_n normaalverdeeld is met $\mathbb{E}T_n = n\mu$ en $\mathrm{Var}(T_n) = n\sigma^2, n \ge 1$ er geldt:

$$|\mathbb{P}\{S_n \in I\} - \mathbb{P}\{T_n \in I\}| \to 0.$$

Bovendien is deze convergentie ook uniform over de klasse van de intervallen in $\mathbb{R}$.

In het algemeen kunnen we echter **niet** concluderen dat $\mathbb{P}\{S_n \in A\} - \mathbb{P}\{T_n \in A\} \to 0$ voor elke Borelverzameling $A \in \mathcal{R}$. Een eenvoudig tegenvoorbeeld is de binomiaalverdeling, die we ook door de normaalverdeling kunnen benaderen, omdat binomiaal(n,p)-variabelen een representatie hebben als een som $\sum_{i=1}^n X_i, n \geq 1$, waar $X_1,\ldots,X_n$ onafhankelijke Bernoulli(p)-variabelen zijn. Als $T_n \sim$ normaal$(np, np(1-p))$ dan geldt voor $A = \{0,1,2,\ldots\}$: $\mathbb{P}\{S_n \in A\} \equiv 1$ en $\mathbb{P}\{T_n \in A\} \equiv 0$, omdat de normaalverdeling absoluut continu is.

Om de CLT te bewijzen hebben we nog een lemma nodig.

Lemma 5.2 *Laat X, Y, Z onafhankelijke toevalsvariabelen zijn en $g : \mathbb{R} \to \mathbb{R}$ een begrensde functie waarvoor de afgeleiden g', g'', g''' bestaan en begrensd zijn.*
Als $\mathbb{E}[|X|^3], \mathbb{E}[|Y|^3] < \infty$, $\mathbb{E}[X] = \mathbb{E}[Y]$ en $\mathbb{E}[X^2] = \mathbb{E}[Y^2]$, geldt:

$$\mathbb{E}[g(Z+X)] - \mathbb{E}[g(Z+Y)] \leq \frac{1}{6}\|g'''\|_\infty(\mathbb{E}[|X|^3] + \mathbb{E}[|Y|^3]).$$

Bewijs We noteren eerst dat de afbeelding g continu en bijgevolg Borel-meetbaar is. Dus zijn $g(Z+X)$ en $g(Z+Y)$ begrensde toevalsvariabelen waarvoor verwachtingswaarden bestaan. Uit Taylors stelling volgt dat

$$|g(z+t) - (g(z) + tg'(z) + t^2 g''(z)/2)| \leq \|g'''\|_\infty |t|^3/6, t, z \in \mathbb{R}.$$

Uit de monotonie van de verwachtingswaarde in verband met stelling 4.6.iii volgt verder dat

$$\mathbb{E}[g(X+Z)] \leq \mathbb{E}[g(Z)] + \mathbb{E}[X]\mathbb{E}[g'(Z)] + \mathbb{E}[X^2]\mathbb{E}[g''(Z)]/2 + \|g'''\|_\infty \mathbb{E}[|X|^3]/6$$

en

$$\mathbb{E}[g(Y+Z)] \geq \mathbb{E}[g(Z)] + \mathbb{E}[Y]\mathbb{E}[g'(Z)] + \mathbb{E}[Y^2]\mathbb{E}[g''(Z)]/2 - \|g'''\|_\infty \mathbb{E}[|Y|^3]/6.$$

Gezien $\mathbb{E}[X] = \mathbb{E}[Y]$ en $\mathbb{E}[X^2] = \mathbb{E}[Y^2]$, is de bovenstaande ongelijkheid nu evident. □

Bewijs van de CLT (onder de extra-voorwaarde $\mathbb{E}[|X_1|^3] < \infty$)
Zonder beperking van de algemeenheid kunnen we veronderstellen dat $\mu = 0$ en $\sigma = 1$. (Vervang de toevalsvariabelen X_j door $(X_j - \mu)/\sigma$ indien nodig.)
In dit geval moeten we tonen dat $S_n/\sqrt{n} \xrightarrow{d} Y$, waar Y standaard normaal is. Daarvoor is het voldoende te bewijzen (zie opmerking na stelling 5.4) dat

$$\mathbb{E}[g(S_n/\sqrt{n})] \to \mathbb{E}[g(Y)]$$

voor elke begrensde functie waarvoor alle afgeleiden bestaan en begrensd zijn.
Neem een rij onafhankelijke standaard-normaalverdeelde toevalsvariabelen $Y_n, n \geq 1$ die onafhankelijk is van de rij $X_n, n \geq 1$. Stel $T_n = \sum_{j=1}^n Y_j$. Dan is ook $T_n/\sqrt{n}$ standaard-normaalverdeeld (zie voorbeeld (2) in 3.6) en we hebben

$$\mathbb{E}[g(Y)] = \mathbb{E}[g(T_n/\sqrt{n})]$$

zodat het voldoende is te tonen

$$\mathbb{E}[g(S_n/\sqrt{n})] - \mathbb{E}[g(T_n/\sqrt{n})] \to 0$$

Stel $X_{n,j} = X_j/\sqrt{n}, Y_{n,j} = Y_j/\sqrt{n}, 1 \le j \le n$. Dan geldt er met de notatie $\sum_{j=n}^m a_j = 0$ als $n > m$,

$$\begin{aligned}
|\mathbb{E}[g(S_n/\sqrt{n})] - \mathbb{E}[g(T_n/\sqrt{n})]| &= |\mathbb{E}[g(\sum_{j=1}^n X_{n,j})] - \mathbb{E}[g(\sum_{j=1}^n Y_{n,j})]| \\
&= |\sum_{j=1}^n \{\mathbb{E}[g(\sum_{i=j}^n X_{n,i} + \sum_{i=1}^{j-1} Y_{n,i})] - \mathbb{E}[g(\sum_{i=j+1}^n X_{n,i} + \sum_{i=1}^{j} Y_{n,i})]\}| \\
&\le \sum_{j=1}^n |\mathbb{E}[g(Z_{n,j} + X_{n,j})] - \mathbb{E}[g(Z_{n,j} + Y_{n,j})]|,
\end{aligned}$$

waar $Z_{n,j} := \sum_{i=j+1}^n X_{n,i} + \sum_{i=1}^{j-1} Y_{n,i}, 1 \le j \le n$.
Lemma 5.2 impliceert dat de laatste som begrensd is door

$$\frac{\|g'''\|_\infty}{6} \sum_{j=1}^n (\mathbb{E}[|X_{n,j}|^3] + \mathbb{E}[|Y_{n,j}|^3]) = \frac{\|g'''\|_\infty}{6} \frac{\mathbb{E}[|X_1|^3] + \mathbb{E}[|Y_1|^3]}{\sqrt{n}},$$

hetgeen natuurlijk de CLT onder de extra-voorwaarde $\mathbb{E}[|X_1|^3] < \infty$ impliceert. □

Toepassing (Roulette) Een gokker in Las Vegas zet altijd 1$ op "rood". Wat is (in benadering) de kans dat hij/zij na (a) 100 (b) 1000 spelletjes niets verloren heeft?

Oplossing Zij $S_n = \#$ successen $\sim$ binomiaal$(n, 9/19)$, waar n het aantal spelletjes is. Dan geldt voor de toevalsvariabele S_n^* = winst/verlies: $S_n^* = S_n - (n - S_n) = 2S_n - n$.Gezien S_n een representatie heeft als een som van n onafhankelijke Bernoulli(9/19)-variabelen, kunnen we via de CLT concluderen dat

$$\mathbb{P}\{S_n^* \ge 0\} = \mathbb{P}\{S_n \ge n/2\} \approx 1 - \Phi\left(\frac{n/2 - 9n/19}{\sqrt{n(9/19)(10/19)}}\right) = 1 - \Phi(\sqrt{n/360})$$

Dus als $n = 100$ is de kans nog rond 0,2981 terwijl als $n = 1000$ de kans ongeveer gelijk aan 0,0918 is.

Algemeen bewijs van de CLT Het idee is zoals in het bewijs van de WLLN de toevalsvariabelen X_j door begrensde (afgekapte) toevalsvariabelen te vervangen.
We veronderstellen weer dat $\mathbb{E}[X_1] = 0, \mathbb{E}[X_1^2] = 1$.
Stel $X_j^{(n)} := X_j I_{\{|X_j| \le \sqrt{n}\}}, 1 \le j \le n$ en $S_n' := \sum_{j=1}^n X_j^{(n)}, n \ge 1$.
Zij $g : \mathbb{R} \to \mathbb{R}$ weer een begrensde continue functie waarvoor ook alle afgeleiden bestaan en begrensd zijn. Dan volgt er onmiddellijk dat

$$|\mathbb{E}[g(S_n/\sqrt{n})] - \mathbb{E}[g(S_n'/\sqrt{n})]| \le 2\|g\|_\infty \mathbb{P}\{S_n \ne S_n'\} \tag{5.6}$$

Verder geldt:

$$\mathbb{P}\{S_n \ne S_n'\} \le \mathbb{P}(\bigcup_{j=1}^n \{X_j \ne X_j^{(n)}\}) \le n\mathbb{P}\{|X_1| > \sqrt{n}\} \le \mathbb{E}[X_1^2 I_{\{|X_1| > \sqrt{n}\}}], \tag{5.7}$$

waar de laatste term naar nul convergeert (stelling van de gedomineerde convergentie).
Dus is het voldoende te bewijzen dat als Z standaard normaal verdeeld is, geldt:

$$|\mathbb{E}[g(S_n'/\sqrt{n})] \to \mathbb{E}[g(Z)]| \text{ als } n \to \infty. \tag{5.8}$$

Zij $\{Z_n, n \geq 1\}$ een rij onafhankelijke standaard-normaal verdeelde toevalsvariabelen. Stel

$$Y_j^{(n)} = \sigma_n Z_j + \mu_n, 1 \leq j \leq n, n \geq 1,$$

waar $\mu_n = \mathbb{E}[X_1^{(n)}]$, $\sigma_n^2 = \mathrm{Var}(X_1^{(n)})$.
Zij verder $T_n' := \sum_{j=1}^n Y_j^{(n)}, n \geq 1$. Dan volgt zoals in het vorige bewijs dat

$$|\mathbb{E}[g(S_n'/\sqrt{n})] - \mathbb{E}[g(T_n'/\sqrt{n})]| \leq \frac{\|g'''\|_\infty}{6} \frac{\mathbb{E}|X_1^{(n)}|^3 + \mathbb{E}|Y_1^{(n)}|^3}{\sqrt{n}}. \tag{5.9}$$

Verder hebben we voor elke $0 < \delta < 1$,

$$\mathbb{E}|X_1^{(n)}|^3 \leq \mathbb{E}[X_1^2]\delta\sqrt{n} + \mathbb{E}|X_1|^3 I_{\{\delta\sqrt{n} < |X_1| \leq \sqrt{n}\}} \leq \sqrt{n}(\delta + \mathbb{E}X_1^2 I_{\{|X_1| > \delta\sqrt{n}\}}).$$

De stelling van de gedomineerde convergentie impliceert dat $\mathbb{E}X_1^2 I_{\{|X_1|>\delta\sqrt{n}\}}$ naar nul convergeert en bijgevolg geldt er voor elke $\delta > 0$, $\limsup_{n\to\infty} \mathbb{E}|X_1^{(n)}|^3/\sqrt{n} \leq \delta$, hetgeen natuurlijk impliceert dat

$$\mathbb{E}|X_1^{(n)}|^3/\sqrt{n} \to 0 \text{ als } n \to \infty.$$

Gebruiken we de convexiteit van de functie $t \to |t|^3$, zien we dat

$$\mathbb{E}|Y_1^{(n)}|^3/\sqrt{n} \leq 4(\sigma_n^3 \mathbb{E}|Z_1|^3 + |\mu_n|^3)/\sqrt{n}.$$

Merk op dat

$$\sqrt{n}|\mu_n| = \sqrt{n}|\mathbb{E}[X_1 I_{\{|X_1|>\sqrt{n}\}}]| \leq \sqrt{n}\mathbb{E}[|X_1| I_{\{|X_1|>\sqrt{n}\}}] \leq \mathbb{E}[X_1^2 I_{\{|X_1|>\sqrt{n}\}}] \to 0$$

en dat

$$\sigma_n^2 = \mathbb{E}[X_1^2 I_{\{|X_1|\leq\sqrt{n}\}}] - \mu_n^2 \to \mathbb{E}[X_1^2] = 1.$$

Aangezien $\mathbb{E}|Z_1|^3 < \infty$, kunnen we concluderen dat ook

$$\mathbb{E}|Y_1^{(n)}|^3/\sqrt{n} \to 0 \text{ als } n \to \infty.$$

Uit betrekking (5.9) volgt dan dat

$$|\mathbb{E}[g(S_n'/\sqrt{n})] - \mathbb{E}[g(T_n'/\sqrt{n})]| \to 0 \text{ als } n \to \infty$$

en het is voldoende te tonen dat

$$\mathbb{E}[g(T_n'/\sqrt{n})] \to \mathbb{E}[g(Z)] \text{ als } n \to \infty.$$

Uit de definitie van T_n' (en de convolutieformule voor de normaalverdeling) volgt er dat $T_n' \sim$ normaal$(n\mu_n, n\sigma_n^2)$ en verder dat $T_n'/\sqrt{n} \sim$ normaal$(\sqrt{n}\mu_n, \sigma_n^2)$. Dit impliceert dan dat

$$\mathbb{E}[g(T_n'/\sqrt{n})] = \mathbb{E}[g(\sqrt{n}\mu_n + \sigma_n Z)].$$

We weten reeds dat $\sqrt{n}\mu_n \to 0$ en dat $\sigma_n \to 1$. Dus convergeert $\sqrt{n}\mu_n + \sigma_n Z$ naar Z (bijna) overal en ook in distributie. Het laatste impliceert dat $\mathbb{E}[g(\sqrt{n}\mu_n + \sigma_n Z)]$ naar $\mathbb{E}[g(Z)]$ convergeert, waarmee de centrale limietstelling bewezen is. □

Referenties

Georgii, Hans-Otto: *Stochastics*, DeGruyter (2008).
Jacod, Jean and Protter, Philip: *Probability Essentials*, Springer (2000).
Ross, Sheldon: *A first course in probability*, 9th edition, Pearson International (2013).