Les 4 Voorwaardelijke kansen, de Bayes regel en onafhankelijkheid

Sommige vragen uit de kanstheorie hebben een antwoord dat niet met de intuïtie van iedereen klopt. Een voorbeeld hiervoor is het *Monty-Hall probleem* ook bekend als *Geitenprobleem*:

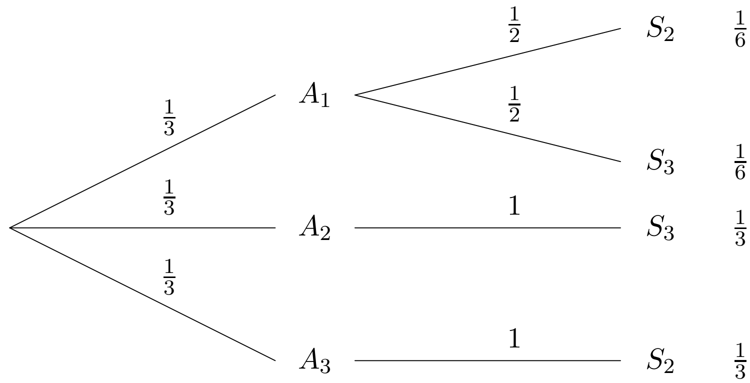
> Bij een TV-show valt er voor de kandidaat een auto te winnen. Het enige wat de kandidaat moet doen is uit drie deuren de goede deur te kiezen waar de auto achter staat. Achter de andere twee deuren zijn er geiten. Nadat de kandidaat een deur heeft gekozen, wordt deze niet meteen geopend, maar de showmaster (die weet waar de auto staat) opent een van de niet gekozen deuren en een geit blaat tegen het publiek (en de kandidaat). De vraag is nu: Is het voor de kandidaat verstandig is om bij zijn keuze te blijven, of is het gunstiger om te wisselen of maakt het niets uit.

Intuïtief zullen veel mensen denken, dat na het openen van een van de deuren met een geit daarachter de kans 50 : 50 is, dat de auto achter de door de kandidaat gekozen deur staat. Dus zou het niets uitmaken of de kandidaat wisselt of niet. In de VS heeft een journaliste, Marilyn vos Savant, de oplossing voor dit probleem in haar column in de tijdschrift *Parade* gepubliceerd. Deze vrouw heeft een van de hoogste IQ's ter wereld en haar antwoord was dat de kans op de auto groeit als de kandidaat wisselt. Haar column resulteerde in een lawine van boosaardige en verontwaardigde brieven, waaronder veel van wiskundigen, die het antwoord van vos Savant bespottelijk maakten. Als reactie op dit gebeuren werd in Duitsland door de journalist Gero von Randow in de weekkrant *Die Zeit* een artikel gepubliceerd, waarin hij het geitenprobleem en een oplossing met dezelfde conclusie als die van vos Savant voorstelde. Ook hier was de reactie opmerkelijk: Over weken kwamen er brieven binnen, waarin professoren, gepromoveerde en dergelijk 'geleerden' uitlegden waarom de oplossing van vos Savant en von Randow onzin is. Ook hier waren er behoorlijk veel wiskundigen bij.

Hoe zit het nu met de oplossing van het geitenprobleem? De reden waarom veel mensen voor de 50 : 50 oplossing kiezen is dat ze ervan uit gaan, dat de situatie na het openen van een van de deuren door de showmaster onafhankelijk is van wat er eerder is gebeurd. Dit is echter niet het geval! Als de kandidaat een deur met een geit daarachter heeft gekozen, heeft de showmaster geen keuze welke deur hij gaat openen, terwijl hij in het geval dat de kandidaat de deur met de auto heft gekozen twee mogelijkheden heeft.

We kunnen dit als volgt analyseren: Stel de kandidaat heeft deur 1 gekozen. De auto kan nu achter deur 1, 2 of 3 staan, deze gevallen noemen we A_1, A_2 en A_3 en we gaan ervan uit dat elk van deze gevallen een kans van $\frac{1}{3}$ heeft. In het geval A_1 kan de showmaster deur 2 of deur 3 openen. Deze gevallen noemen we S_2 en S_3 en omdat er geen verschil tussen de deuren (en de geiten) is, kunnen we aannemen dat S_2 en S_3 dezelfde kans $\frac{1}{2}$ hebben. De kans dat de auto achter deur 1 staat en de showmaster deur 2 opent is dus $\frac{1}{6}$, hetzelfde geldt voor het openen van deur 3. Maar in het geval A_2 heeft de showmaster geen keuze, hij moet deur 3 openen, dus is de kans voor dit geval $\frac{1}{3}$. Evenzo moet de showmaster in het geval A_3 deur 2 openen, dus is ook hier de kans $\frac{1}{3}$.

Deze situatie kunnen we door het volgende boomdiagram beschrijven:



In het geval dat de showmaster deur 2 heeft geopend is de kans dus twee keer zo groot dat de auto achter deur 3 staat dan dat hij achter deur 1 staat. Hetzelfde geldt voor het geval dat de showmaster deur 3 heeft geopend. In elk geval is het dus verstandig dat de kandidaat van keuze verandert, want hierdoor wordt zijn kans op de auto twee keer zo groot.

We zullen later nog eens op het geitenprobleem terug komen en het antwoord uit de regel van Bayes afleiden. Maar eerst gaan we algemeen naar het probleem kijken dat de kans voor een uitkomst kan veranderen als aanvullende informatie over gerelateerde gebeurtenissen bekend wordt.

4.1 Voorwaardelijke kansen

Het idee dat de kans voor een uitkomst kan veranderen als we aanvullende informatie hebben, is zo natuurlijk dat we er meestal niet over nadenken. Bijvoorbeeld kan de kans op vorst op 30 april over de afgelopen 150 jaar eenvoudig afgelezen worden uit de tabellen van de weerkundige dienst. Als er bijvoorbeeld 10 keer in de afgelopen 150 jaren vorst op 30 april was, kunnen we aannemen dat de kans op vorst op 30 april 2005 ongeveer 6.67% is. Als aanvullende informatie kunnen we gebruiken dat er ook 10 keer vorst op 29 april is geweest en dat er in 5 jaren vorst op 29 en 30 april gevallen is. Zo ver maakt dit nog geen verschil voor de kans op vorst op 30 april 2005. Maar als er inderdaad vorst op 29 april 2005 valt, kunnen we zeggen dat de kans op vorst op 30 april 2005 opeens 50% is, want in 5 van de 10 jaren met vorst op 29 april was er ook vorst op 30 april.

De kans dat er vorst op 30 april valt, gegeven het feit dat er vorst op 29 april is, noemen we een *voorwaardelijke kans*.

Abstract gaan we dit zo beschrijven: Stel we willen de kans van $A \subseteq \Omega$ bepalen onder de voorwaarde dat $B \subseteq \Omega$ plaats vindt. Deze kans definiëren we als de kans dat A *en* B gebeuren, gegeven het feit dat B gebeurt. Als de kansen door relatieve frequenties gegeven zijn, dus $P(A) = \frac{|A|}{|\Omega|}$, hebben we $\frac{|A\cap B|}{|B|} = \frac{\frac{|A\cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{P(A\cap B)}{P(B)}$ en het laatste nemen we als definitie voor de voorwaardelijke kans:

Voor een kansverdeling P of Ω en $B \subseteq \Omega$ met $P(B) > 0$ noemen we

$$P(A \mid B) := \frac{P(A,B)}{P(B)} := \frac{P(A\cap B)}{P(B)}$$

de *voorwaardelijke kans op A gegeven B*.

Notatie: De kans voor het gemeenschappelijke optreden van de gebeurtenissen A en B wordt meestal met $P(A,B)$ in plaats van $P(A\cap B)$ genoteerd.

Om te rechtvaardigen, dat we $P(A \mid B)$ een *kans* noemen, moeten we even aantonen dat $P(\cdot \mid B)$ voor $P(B) > 0$ een kansverdeling is, waarbij we natuurlijk gebruiken dat $P(\cdot)$ al een kansverdeling is. (Voor $P(B) = 0$ is het onzin een kans onder de voorwaarde B te bekijken, want B gebeurt nooit.)

(i) $P(A \mid B) = \frac{P(A\cap B)}{P(B)} \geq 0$.

(ii) $P(\Omega \mid B) = \frac{P(\Omega\cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$.

(iii) Voor $A_1, A_2 \subseteq \Omega$ met $A_1\cap A_2 = \emptyset$ geldt $(A_1\cup A_2)\cap B = (A_1\cap B)\cup(A_2\cap B)$. Verder is $(A_1 \cap B)\cap(A_2\cap B) = \emptyset$ omdat $A_1\cap B$ een deelverzameling van A_1 en $A_2\cap B$ een deelverzameling van A_2 is. Daarom geldt: $P(A_1\cup A_2 \mid B) = \frac{P((A_1\cup A_2)\cap B)}{P(B)} = \frac{P((A_1\cap B)\cup(A_2\cap B))}{P(B)} = \frac{P(A_1\cap B)+P(A_2\cap B)}{P(B)} = \frac{P(A_1\cap B)}{P(B)} + \frac{P(A_2\cap B)}{P(B)} = P(A_1 \mid B) + P(A_2 \mid B)$.

Voorbeeld: Hier is een typisch voorbeeld van een vraag die met voorwaardelijke kansen te maken heeft:

Aan 1000 werknemers wordt gevraagd of ze een hoog of een laag salaris hebben. Van de werknemers geven 210 vrouwen aan een hoog salaris te hebben en 360 geven aan een laag salaris te hebben. Van de mannen blijken 210 een hoog en 220 een laag salaris te hebben. Deze gegevens vinden we in het volgende schema terug:

	hoog salaris	laag salaris	som
vrouw	0.21	0.36	0.57
man	0.21	0.22	0.43
totaal	0.42	0.58	1.00

De vraag is nu of vrouwen en mannen dezelfde kans op een hoog salaris hebben. De kans voor een vrouw om een hoog salaris te hebben is de voorwaardelijke kans $P(hoog \mid vrouw) = \frac{P(hoog \text{ en } vrouw)}{P(vrouw)} = \frac{0.21}{0.57} \approx 0.37$. Voor mannen is de kans $P(hoog \mid man) = \frac{P(hoog \text{ en } man)}{P(man)} = \frac{0.21}{0.43} \approx 0.49$ dus hebben mannen in dit voorbeeld een behoorlijk grotere kans op een hoog salaris dan vrouwen.

We kunnen voorwaardelijke kansen niet alleen maar voor twee deelverzamelingen van Ω maar ook algemeen voor n deelverzamelingen definiëren. Het idee hierbij is hetzelfde, we kijken naar de kans van het gemeenschappelijke optreden van de voorwaarden met een gebeurtenis, gedeeld door de kans voor de voorwaarden en krijgen dus:

$$P(A_n \mid A_1 \cap \ldots \cap A_{n-1}) = P(A_n \mid A_1, \ldots, A_{n-1}) = \frac{P(A_1, \ldots, A_n)}{P(A_1, \ldots, A_{n-1})}.$$

We hebben dus bijvoorbeeld $P(A_3 \mid A_1, A_2) = \frac{P(A_1,A_2,A_3)}{P(A_1,A_2)}$.

Omgekeerd kunnen we de kans voor het gemeenschappelijke optreden van gebeurtenissen (iteratief) door voorwaardelijke kansen uitdrukken en krijgen zo de zogeheten *kettingregel*:
$P(A_1, A_2) = P(A_2 \mid A_1) \cdot P(A_1)$,
$P(A_1, A_2, A_3) = P(A_3|A_1, A_2) \cdot P(A_1, A_2) = P(A_3|A_1, A_2) \cdot P(A_2|A_1) \cdot P(A_1)$,
en in het algemeen
$P(A_1, ..., A_n) = P(A_n|A_1, ..., A_{n-1}) \cdot P(A_{n-1}|A_1, ..., A_{n-2}) \cdot \ldots \cdot P(A_2|A_1) \cdot P(A_1)$.

4.2 Regel van Bayes

Omdat de doorsnede $A \cap B$ symmetrisch in A en B is, vinden we uit de definitie voor de voorwaardelijke kans dat

$$P(A \mid B) \cdot P(B) = P(A \cap B) = P(B \mid A) \cdot P(A)$$

en dit geeft de eenvoudigste vorm van de *regel van Bayes*, namelijk

$$P(B \mid A) = \frac{P(A \mid B) \cdot P(B)}{P(A)}.$$

De nut van deze regel ligt in het omdraaien van de rollen van voorwaarde en uitkomst. Denk hierbij bijvoorbeeld aan een test op een ziekte. Als de uitslag van de test gegeven is, zijn we geïnteresseerd in de kans dat we de ziekte hebben of niet. Maar bekend is alleen maar de nauwkeurigheid van de test die zegt met welke kans de test bij een gezonde mens het verkeerde resultaat geeft en andersom.

De regel van Bayes wordt vaak op een iets slimmere manier toegepast. Hiervoor wordt de deelverzameling $B \subseteq \Omega$ in verschillende gevallen onderverdeeld die elkaar uitsluiten, dus we schrijven $B = \cup_{i=1}^{n} B_i$ met $B_i \cap B_j = \emptyset$ als $i \neq j$. Een belangrijk speciaal geval hiervoor is $B = B_1 \cup B_2$ met $B_2 = B \setminus B_1 = B_1^c$. We noemen B_2 het *complement* van B_1 in B.

Er geldt:

$$P(A \cap B) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)$$

en dus

$$P(A \mid B) = \frac{1}{P(B)} \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i).$$

In het bijzonder kunnen we in het geval dat $A \subseteq B$ de *totale kans* $P(A)$ berekenen als $P(A) = P(A \cap B) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)$ en het belangrijkste geval hiervoor is $B = \Omega$, d.w.z. we delen alle mogelijke uitkomsten in een aantal klassen van uitkomsten op.

We kunnen nu de regel van Bayes algemeen formuleren:

Regel van Bayes: Zij $B \subseteq \Omega$ met $B = \cup_{i=1}^{n} B_i$ en $B_i \cap B_j = \emptyset$ als $i \neq j$. Verder zij $A \subseteq B$. Dan geldt

$$P(B_j \mid A) = \frac{P(A \mid B_j) \cdot P(B_j)}{P(A)} = \frac{P(A \mid B_j) \cdot P(B_j)}{\sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)}$$

Om de abstracte concepten duidelijk te maken, passen we de regel van Bayes op een aantal voorbeelden toe.

Voorbeeld 1: De uitkomst van een HIV-test noemen we A als de test positief was en A^c als de test negatief was. Het geïnfecteerd zijn noemen we I en het niet geïnfecteerd zijn I^c. Over de kwaliteit van de test is bekend, dat hij voor geïnfecteerden in 99% van de gevallen een positief resultaat oplevert en voor niet geïnfecteerden in 99.9% van de gevallen een negatief resultaat. We hebben dus $P(A \mid I) = 0.99$, $P(A^c \mid I) = 0.01$ en $P(A^c \mid I^c) = 0.999$, $P(A \mid I^c) = 0.001$. Verder nemen we aan dat 1 uit 10000 mensen HIV-geïnfecteerd is, dus $P(I) = 0.0001$ en $P(I^c) = 0.9999$. De vraag is nu, hoe groet bij een positieve HIV-test de kans is, inderdaad geïnfecteerd te zijn, dus hoe groot de voorwaardelijke kans $P(I \mid A)$ is. Met de regel van Bayes hebben we

$$P(I \mid A) = \frac{P(A \mid I) \cdot P(I)}{P(A)} = \frac{P(A \mid I) \cdot P(I)}{P(A \mid I) \cdot P(I) + P(A \mid I^c) \cdot P(I^c)}$$
$$= \frac{0.99 \cdot 0.0001}{0.99 \cdot 0.0001 + 0.001 \cdot 0.9999} \approx 9.0\%.$$

Deze verrassend lage kans is opmerkelijk maar toch goed te begrijpen. Als we 10000 mensen testen, dan is er gemiddeld 1 HIV-geïnfecteerde mens bij en die krijgt waarschijnlijk ook een positieve test-uitslag. Maar bij de 9999 niet-geïnfecteerden zal de test in 0.1% van de gevallen een (verkeerd) positief resultaat opleveren, dus komen er nog 10 positieve resultaten bij. Als we dus naar de 11 positieve resultaten kijken, is dit alleen maar in één geval veroorzaakt door een geïnfecteerde, maar in 10 gevallen door een test-fout.

Merk op dat er in dit soort vragen vaak verkeerd geargumenteerd wordt. Dit vind je zelfs in wetenschappelijke publicaties, bijvoorbeeld in de medicijn of in de rechtsgeleerdheid terug. Denk hier bijvoorbeeld aan een misdadiger waarbij de schuld door een DNA-analyse wordt bewezen. Het probleem is, dat zelfs bij een test met een hoge nauwkeurigheid het aantal verkeerde uitslagen vaak hoger is dan het aantal van de gezochte zeldzame uitkomsten.

Voorbeeld 2: Een student moet bij een tentamen een multiple-choice vraag met n mogelijkheden oplossen. Als hij voorbereid is, zal zijn antwoord juist zijn, als niet zal hij willekeurig een antwoord gokken en dus een kans van $\frac{1}{n}$ op een juiste antwoord hebben. De kans dat de student voorbereid is, zij p. Voor de docent is het nu interessant om de kans te bepalen, dat de student inderdaad voorbereid was, als hij een juiste antwoord heeft gegeven. Als we een juiste antwoord met J en een voorbereide student met V betekenen hebben we dus:

$$P(V \mid J) = \frac{P(J \mid V) \cdot P(V)}{P(J \mid V) \cdot P(V) + P(J \mid V^c) \cdot P(V^c)}$$

$$= \frac{1 \cdot p}{1 \cdot p + \frac{1}{n}(1-p)} = \frac{np}{np + (1-p)}.$$

Het is duidelijk dat dit voor grote waarden van n dicht bij 1 ligt, want dan is $(1-p)$ tegen np te verwaarlozen. Maar voor $n = 4$ en $p = 0.5$ hebben we bijvoorbeeld $P(V \mid J) = \frac{4}{5} = 80\%$ en voor $n = 4$ en $p = 0.2$ geldt al $P(V \mid J) = \frac{1}{2} = 50\%$. Als de docent dus weet dat gewoon maar een vijfde van de studenten voorbereid is, weet hij ook dat de helft van de goede antwoorden goede gokken zijn.

Voorbeeld 3: In de automatische spraakherkenning gaat het erom, gegeven een akoestisch signaal X het woord w te vinden dat hier het beste bij past, d.w.z. waarvoor de voorwaardelijke kans $P(w \mid X)$ maximaal is. Hiervoor gebruiken we ook de regel van Bayes en schrijven

$$P(w \mid X) = \frac{P(X \mid w) \cdot P(w)}{P(X)}.$$

Omdat we alleen maar aan het woord met de hoogste kans geïnteresseerd zijn, kunnen we de noemer gewoon vergeten, omdat die voor elk woord hetzelfde is. In de teller geeft $P(X \mid w)$ de kans, dat een zeker woord w tot het signaal X lijdt. Deze kans wordt tijdens het *training* van een systeem bepaald, waarbij een aantal mensen het woord spreekt en uit de zo verkregen signalen een kansverdeling geschat wordt. De kans $P(w)$ is de totale kans dat een woord gesproken wordt. Dit noemen we de a-priori kans voor het woord, en deze kansen worden als relatieve frequenties op heel grote tekst-corpora (bijvoorbeeld 10 jaar NRC Handelsblad) bepaald.

Hetzelfde principe geldt trouwens voor de meeste soorten van patroonherkenning (beeld-herkenning, handschrift-herkenning).

Voorbeeld 4: We komen nog eens terug op het Monty-Hall probleem. Stel de kandidaat heeft deur 1 gekozen, dan nemen we aan dat de showmaster deur 2 heeft geopend (S_2), het geval S_3 geeft een analoog resultaat. We zijn nu geïnteresseerd in de kansen $P(A_1 \mid S_2)$ en $P(A_3 \mid S_2)$, dus de voorwaardelijke kansen dat de auto achter deur 1 of deur 3 staat, gegeven het feit dat de showmaster deur 2 heeft geopend. Er geldt

$$P(A_1 \mid S_2) = \frac{P(S_2 \mid A_1) \cdot P(A_1)}{P(S_2 \mid A_1) \cdot P(A_1) + P(S_2 \mid A_2) \cdot P(A_2) + P(S_2 \mid A_3) \cdot P(A_3)}$$

$$= \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 0 + 1 \cdot \frac{1}{3}} = \frac{1}{3}.$$

Evenzo berekenen we de kans $P(A_3 \mid S_2)$ als

$$P(A_3 \mid S_2) = \frac{P(S_2 \mid A_3) \cdot P(A_3)}{P(S_2 \mid A_1) \cdot P(A_1) + P(S_2 \mid A_2) \cdot P(A_2) + P(S_2 \mid A_3) \cdot P(A_3)}$$

$$= \frac{1 \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 0 + 1 \cdot \frac{1}{3}} = \frac{2}{3}.$$

We zien dus weer dat het voor de kandidaat verstandig is om naar deur 3 te wisselen, omdat de kans dat de auto daar achter zit twee keer zo groot is.

4.3 Onafhankelijkheid

Nu dat we goed naar voorwaardelijke kansen hebben gekeken kunnen we ook zeggen wat het betekent dat twee uitkomsten onafhankelijk zijn. Intuïtief zullen we zeggen, dat twee uitkomsten A en B onafhankelijk zijn, als de kans voor A niet ervan afhangt of B optreed of niet. Met de voorwaardelijke kans kunnen we dit zo formuleren:

Twee uitkomsten A en B heten onafhankelijk als $P(A) = P(A \mid B)$.
Equivalent hiermee is dat $P(A \cap B) = P(A) \cdot P(B)$.

De equivalentie van de twee formuleringen volgt uit de definitie van de voorwaardelijke kans, want $P(A \cap B) = P(A \mid B) \cdot P(B)$ geeft $P(A) = P(A \mid B) \Leftrightarrow P(A \cap B) = P(A \mid B) \cdot P(B) = P(A) \cdot P(B)$. Omdat ook $P(A \cap B) = P(B \mid A) \cdot P(A)$ geldt, volgt hieruit ook dat $P(A) = P(A \mid B) \Leftrightarrow P(B) = P(B \mid A)$, dus het maakt niets uit welke voorwaardelijke kans we bekijken.

Een eenvoudig voorbeeld zijn de soorten en kleuren in een kaartspel. De kans om uit een kaartspel met 52 kaarten een aas te trekken is $\frac{1}{13}$, de kans om een kaart van kleur klaver te trekken is $\frac{1}{4}$. De doorsnede van de uitkomsten *aas* en *klaver* is alleen maar de kaart *klaver aas* en de kans om deze kaart te trekken is $\frac{1}{52} = \frac{1}{13} \cdot \frac{1}{4}$. Omdat we ook elke andere soort of kleur hadden kunnen kiezen, toont dit aan, dat de soorten en de kleuren onafhankelijk zijn.

In een ander voorbeeld kijken we naar een familie met twee kinderen. We vragen ons af of de uitkomsten

$A:$ er is een meisje en een jongen $\qquad$ $B:$ er is hoogstens een meisje

onafhankelijk zijn. Als we m voor een meisje en j voor een jongen schrijven, zijn de mogelijkheden voor de twee kinderen (m, m), (m, j), (j, m) en (j, j). We zien makkelijk dat $P(A) = \frac{1}{2}$ en $P(B) = \frac{3}{4}$, maar $P(A \cap B) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{3}{4} = \frac{3}{8}$. Dus zijn de uitkomsten A en B niet onafhankelijk.

Als we de familie nu van twee naar drie kinderen uitbreiden maar dezelfde uitkomsten bekijken, is de situatie veranderd. De mogelijkheden voor de drie kinderen zijn nu (m, m, m), (m, j, m), (j, m, m), (j, j, m), (m, m, j), (m, j, j), (j, m, j) en (j, j, j). In dit geval is $P(A) = \frac{3}{4}$, $P(B) = \frac{1}{2}$ en $P(A \cap B) = \frac{3}{8} = P(A) \cdot P(B)$, dus zijn de uitkomsten nu inderdaad onafhankelijk.

Aan de hand van dit voorbeeld zien we, dat soms uitkomsten kanstheoretisch onafhankelijk zijn, die we in het echte leven niet onafhankelijk zouden noemen.

De onafhankelijkheid van uitkomsten A en B heeft ook nuttige consequenties voor de complementen A^c en B^c. Er geldt namelijk dat met (A, B) ook de paren (A, B^c), (A^c, B) en (A^c, B^c) onafhankelijk zijn. Dit kunnen we met behulp van een paar eenvoudige manipulaties van de betrokken verzamelingen uit $P(A \cap B) = P(A) \cdot P(B)$ concluderen:

$P(A \cap B^c) = P(A \cup B) - P(B) = P(A) + P(B) - P(A \cap B) - P(B) = P(A) - P(A \cap B) = P(A) - P(A) \cdot P(B) = P(A)(1 - P(B)) = P(A) \cdot P(B^c)$. Dit werkt evenzo voor $P(A^c \cap B)$.

$P(A^c \cap B^c) = P((A \cup B)^c) = 1 - P(A \cup B) = 1 - P(A) - P(B) + P(A \cap B) = 1 - P(A) - P(B) + P(A) \cdot P(B) = (1 - P(A))(1 - P(B)) = P(A^c) \cdot P(B^c)$.

We kunnen het begrip van onafhankelijkheid ook naar stochasten uitbreiden: Voor twee stochasten X, Y zij $A_x := \{\omega \in \Omega \mid X(\omega) = x\}$ en $B_y := \{\omega \in \Omega \mid Y(\omega) = y\}$. We noemen de uitkomsten A_x en B_y onafhankelijk als $P(A_x \cap B_y) = P(A_x) \cdot P(B_y)$. Maar in de taal van stochasten heet dit dat

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$$

en we noemen twee stochasten X, Y onafhankelijk als dit voor alle paren (x, y) geldt.

Tot nu toe hebben we het alleen maar over de onafhankelijkheid van *twee* uitkomsten gehad. Als we meerdere uitkomsten bekijken, zijn er verschillende mogelijkheden om hun onafhankelijkheid te definiëren:

(1) We noemen de n uitkomsten $A_1, \ldots, A_n$ *paarsgewijs onafhankelijk* als $P(A_i \cap A_j) = P(A_i) \cdot P(A_j)$ voor alle $i \neq j$.

(2) We noemen n uitkomsten $A_1, \ldots, A_n$ *onafhankelijk* als $P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot \ldots \cdot P(A_{i_k})$ voor elke deelverzameling $\{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$.

Als we de begrippen op deze manier definiëren is het duidelijk dat onafhankelijke uitkomsten ook paarsgewijs onafhankelijk zijn. Het omgekeerde geldt niet, wat aan het volgende tegenvoorbeeld duidelijk wordt:

We dobbelen met twee dobbelstenen en bekijken de kansen van de volgende uitkomsten:

A_1 : de eerste dobbelsteen toont een oneven getal,

A_2 : de tweede dobbelsteen toont een oneven getal,

A_3 : de som van de getallen is even.

We hebben $P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}$ en $P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{4}$, dus zijn de uitkomsten paarsgewijs onafhankelijk. Maar $P(A_1 \cap A_2 \cap A_3) = P(A_1 \cap A_2)$ omdat de som van twee oneven getallen even is, dus is $P(A_1 \cap A_2 \cap A_3) \neq P(A_1) \cdot P(A_2) \cdot P(A_3) = \frac{1}{8}$ en dus zijn de drie uitkomsten niet onafhankelijk.

We zouden bij de definitie van onafhankelijkheid voor meerdere uitkomsten ook kunnen hopen dat het voldoende is om $P(A_1 \cap \ldots \cap A_n) = P(A_1) \cdot \ldots \cdot P(A_n)$ te eisen, maar het volgende tegenvoorbeeld laat zien dat hieruit niet eens volgt dat de A_i paarsgewijs onafhankelijk zijn: We werpen een munt drie keer en kijken naar de volgende uitkomsten:

A_1 : de eerste worp toont kop,

A_2 : er valt vaker kop dan munt,

A_3 : de laatste twee worpen leveren hetzelfde resultaat.

Door naar de mogelijke uitkomsten te kijken zien we dat $P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}$ en dat $P(A_1 \cap A_2 \cap A_3) = \frac{1}{8}$. Aan de andere kant hebben we $P(A_1 \cap A_2) = \frac{3}{8}$, dus zijn A_1 en A_2 niet (paarsgewijs) onafhankelijk. De andere paren zijn wel onafhankelijk, want $P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{4}$.

4.4 Bernoulli-model

Een belangrijke toepassing van de onafhankelijkheid van uitkomsten is de herhaalde uitvoering van een experiment. We nemen aan dat we in de uitkomstenruimte Ω een deelverzameling $A \subseteq \Omega$ van gunstige uitkomsten hebben. Bij de eenmalige uitvoering van het experiment is de kans op een gunstige uitkomst gegeven door $p = \frac{|A|}{|\Omega|}$. De kans voor een ongunstige uitkomst is dan $1-p$. Als we het experiment twee keer uitvoeren is de kans dat we twee gunstige uitkomsten hebben de kans van de doorsnede van een gunstige uitkomst bij de eerste keer en een gunstige uitkomst bij de tweede keer. Omdat we ervan uitgaan dat het eerste en het tweede experiment onafhankelijk zijn, kunnen we de kans voor de doorsnede als product van de enkele kansen berekenen, dus als $p \cdot p = p^2$.

Merk op dat de eis dat herhalingen van een experiment onafhankelijk zijn een voorwaarde voor de opzet van het experiment is. Als je bijvoorbeeld de kans wilt bepalen waarmee een vaccinatie tot de uitbraak van een ziekte lijdt mag je bij het herhalen van het experiment geen mensen nemen die al bij de vorige keer gevaccineerd zijn, omdat deze een hoger aantal antilichamen hebben en dus een kleinere kans lopen dat de ziekte uitbreekt.

Als we ervan uitgaan dat het herhalen van een experiment onafhankelijke uitkomsten heeft, dan is de kans dat we bij m herhalingen k keer een gunstige uitkomst hebben gegeven door de binomiale verdeling:

$$b(m, p; k) = \binom{m}{k} p^k (1-p)^{m-k}.$$

De kans dat de eerste k uitkomsten gunstig zijn is namelijk p^k en de kans dat de laatste $m-k$ uitkomsten ongunstig zijn is $(1-p)^{m-k}$. Nu kunnen we de gunstige uitkomsten nog op $\binom{m}{k}$ manieren over de m experimenten verdelen.

De beschrijving van uitkomsten door onafhankelijke herhaling van een experiment noemen we het *Bernoulli-model*.

Belangrijke begrippen in deze les

- voorwaardelijke kans
- regel van Bayes
- onafhankelijkheid, paarsgewijs onafhankelijk
- Bernoulli-model

Opgaven

16. Een socioloog wil de kans bepalen dat mensen een keer een winkeldiefstal hebben gepleegd. Omdat mensen op een rechtstreekse vraag waarschijnlijk niet eerlijk zouden antwoorden heeft hij de volgende opzet verzonnen: Elke persoon krijgt 10 kaarten waarvan op 4 de vraag staat:
Heb je ooit een winkeldiefstal gepleegd?
en op de andere 6 de vraag
Heb je nog nooit een winkeldiefstal gepleegd?
De mensen worden nu gevraagd om toevallig één van de tien kaarten te trekken, het (waarheidsgetrouwe) antwoord op een briefje te schrijven en alleen maar dit briefje aan de onderzoeker te geven. Zo hoeft niemand om zijn anonimiteit te vrezen.

    Bij 1000 testpersonen krijgt de onderzoeker 516 keer het antwoord *ja* en 484 keer het antwoord *nee*. Hoe kan hij nu de gezochte kans berekenen en wat is deze kans?

17. Er wordt met twee dobbelstenen gedobbeld. Gegeven de informatie dat de twee dobbelstenen verschillende getallen tonen (bijvoorbeeld in een spel waar je bij gelijke getallen nog een keer dobbelt), wat is de kans dat de som oneven is?

18. In een zak zitten drie munten, waarvan twee eerlijk zijn maar de derde heeft twee kop-zijden. Er wordt blindelings een munt getrokken, vervolgens wordt deze munt twee keer geworpen, waarbij twee keer kop valt. Bepaal de kans, dat de getrokken munt een eerlijke munt is.

    Hoe zit het met het geval dat in de zaak een miljoen in plaats van drie munten zitten, waarvan weer één oneerlijk is. Nu werp je twintig keer in plaats van twee keer en krijgt twintig keer het resultaat kop. Hoe groot is nu de kans dat de getrokken munt een eerlijke munt is.

19. In sommige studies is er na het eerste semester een advies aan de studenten die weliswaar niet bindend is. Neem aan dat in een (zware) studie gemiddeld 40% van de studenten vroegtijdig afhaken. Het blijkt dat van de afhakende studenten 90% een negatief studieadvies kregen, terwijl slechts 1% van de studenten die afstuderen een negatief advies hadden. Wat is de kans dat een student met negatief studieadvies wel in dit vak zou afstuderen?

20. Bij een rechtbank zal een leugendetector geraadpleegd worden. Het is bekend dat voor een schuldige verdachte de detector in 90% van de gevallen het juiste resultaat (schuldig) geeft en voor een onschuldige verdachte in 99% van de gevallen het resultaat onschuldig. Uit de statistieken van de belastingdienst is bekend dat 5% van de burgers in hun belastingaangifte ernstig bedriegen. Bij een verdachte geeft de leugendetector aan dat de man/vrouw schuldig is. Wat is de kans, dat de verdachte toch onschuldig is?