Hoofdstuk 4 – Kansen

4.1 Randomheid

Herhalingen en kansen

Als je een munt opgooit (of zelfs als je een SRS trekt) kunnen de resultaten van tevoren voorspeld worden, omdat de uitkomsten zullen variëren wanneer je herhaaldelijk een munt opgooit of herhaaldelijk een steekproef trekt.
Kansen beschrijven alleen wat er op lange termijn gebeurt. Veel mensen verwachten dat kansuitkomsten op korte termijn al regelmatig zijn, terwijl dat niet zo is. Als je een munt opgooit, dan is het zo dat er na pas na heel vaak gooien een patroon van 50% kans op munt en 50% kans op kop ontstaat. Op korte termijn is dit vaak nog niet het geval.

Begrippen

- We noemen een fenomeen *random* als individuele uitkomsten onzeker zijn, maar er toch een duidelijk uitkomstenpatroon op lange termijn waar te nemen is. Er moet dan wel sprake zijn van vele herhalingen. Denk in dit verband maar aan het opgooien van een munt.
- De *kans (probability)* op een uitkomst van een random fenomeen is de proportie van het aantal keren dat de uitkomst voor zal komen na vele herhalingen. Bij een munt is de bijbehorende proportie voor munt dus 0.5 en voor kop geldt dezelfde proportie. Echte munten hebben echter kleine imperfecties waardoor de kans op kop niet precies 0.5 is. We noemen een munt *eerlijk (fair)* wanneer de kans op kop precies 0.5 is en de kans op munt ook 0.5.

Randomheid

Met het idee van randomheid kun je zelf experimenteren, door bijvoorbeeld meerdere malen een munt op te gooien. Je moet dan wel zeer vaak de munt opgooien om een patroon in kansen te ontdekken.

- Uitkomsten zijn random als herhalingen van dezelfde handeling *onafhankelijk (independent)* van elkaar zijn. Dit betekent dat de uitkomst van de eerste keer een munt opgooien geen invloed heeft op de uitkomst van de volgende keer. De kansen beïnvloeden elkaar dus niet.
- Het idee van kansberekeningen is empirisch. Simulaties beginnen met een gegeven kans en imiteren daarmee random gebeurtenissen. We kunnen een kans uit het dagelijks leven echter alleen schatten door vele herhalingen van dezelfde handeling te observeren.
- Toch zijn simulaties erg handig, omdat het niet handig is om een munt in de praktijk honderden keren op te gooien.

4.2 Kansmodellen

Een kansmodel

Een *kansmodel (probability model)* is de beschrijving van een random fenomeen in rekenkundige termen. Zo een model bevat altijd:

- Een lijst met alle mogelijke uitkomsten (bij een munt is dat dus de kop of munt).
- De kans op elke uitkomst (kop en munt hebben beide een kans van 0.5).

Steekproefruimten (sample spaces)

Een *steekproefruimte (S)* van een random fenomeen is de set van alle mogelijke uitkomsten. Bij een munt zijn dit dus kop en munt: S =(kop, munt). S is dus een opsomming van alle mogelijke uitkomsten van een random fenomeen. Een *gebeurtenis (event)* is een uitkomst (of een set van uitkomsten) van een random fenomeen. Een gebeurtenis is dus een kleiner onderdeel van de steekproefruimte. De kans op twee keer kop bij vier keer een munt werpen is een voorbeeld van een gebeurtenis. Dit wordt als volgt weergegeven:
A = (KKMM, KMKM, KMMK, MKKM, MKMK, MMKK)

Feiten over kansen

- *Elke kans bevindt zich tussen de 0 en de 1.* Als de kans op een gebeurtenis 0 is, betekent dit dat deze gebeurtenis nooit voorkomt. Is de kans op een gebeurtenis 1, dan komt deze juist in alle gevallen voor. Een gebeurtenis met kans 0,5 komt voor de helft van de keren voor.
- *Alle mogelijke uitkomsten bij elkaar hebben een kans van 1.* De optelling van alle mogelijkheden is dus altijd 1. Als er een kans van 0.5 op munt gooien is en een kans van 0.5 op kop gooien is, dan is dit samen 1.
- *Als twee gebeurtenissen geen gemeenschappelijke uitkomsten hebben, dan is de kans dat de ene of de andere voorkomt de optelling van de kansen op beide uitkomsten*. Als de ene gebeurtenis samengaat met een kans van 0.40 en de ander met een kans van 0.25, en de twee kunnen nooit samen voorkomen, dan komt één van de twee dus voor in 65% van de gevallen.
- *De kans dat een gebeurtenis niet voorkomt is 1 min de kans dat de gebeurtenis wel voorkomt*. Als er een kans van 0.4 is dat we munt gooien, dan is er 0.6 kans dat we geen munt gooien.

Kansregels in statistische termen

De bovenstaande feiten zijn gebruikt om een aantal kansregels te formuleren. Deze regels worden hieronder beschreven.

- *Regel 1:* De kans P(A) die bij een gebeurtenis hoort is $0 \leq P(A) \leq 1$. Dit betekent dus dat de kans zich tussen de 0 en 1 bevindt.
- *Regel 2:* Als S de steekproefruimte in een kansmodel is, dan geldt: $P(S)=1$.
- *Regel 3:* Twee gebeurtenissen A en B zijn *disjunct* als ze geen gemeenschappelijke

uitkomsten zijn en dus nooit samen voor kunnen komen. Als A en B disjunct zijn, geldt: P(A of B)=P(A)+P(B). Dit wordt ook wel de *optelregel voor disjuncte gebeurtenissen* genoemd.
- *Regel 4:* Het complement van gebeurtenis A is de gebeurtenis waar A niet voorkomt. Het complement vinden we door de kans op gebeurtenis A af te trekken van 1. Dit noemen we ook wel de *complementregel.*

Om complementen en disjuncte gebeurtenissen beter te begrijpen kan het tekenen van een *Venndiagram* handig zijn. A en B zijn disjunct wanneer de gebieden elkaar niet overlappen. A en B zijn complement wanneer gebieden A en B elkaar niet overlappen maar waar er ook geen overige gebieden aanwezig zijn.

Kansen toewijzen aan een eindig aantal uitkomsten
De individuele uitkomsten van een random fenomeen zijn altijd disjunct. De optelregel voor disjuncte gebeurtenissen geeft aan hoe we kansen moeten toewijzen aan individuele uitkomsten. Deze uitkomsten kunnen vervolgens opgeteld worden om de kans op gebeurtenissen te beschrijven. Dit idee werkt goed als er een eindig (dus beperkt) aantal uitkomsten is.
- Wijs eerst een kans toe aan elke individuele uitkomst. Deze kansen moeten tussen de 0 en 1 zijn.
- De kans op een gebeurtenis is de optelling van de kansen voor de uitkomsten die deel uitmaken van de gebeurtenis.

Uitkomsten met dezelfde kans
Soms nemen we aan dat er een gelijke kans is op uitkomsten, omdat er een soort balans aanwezig is in een fenomeen. We nemen bijvoorbeeld aan dat er een kans van 0.5 is op het gooien van munt en een kans van 0.5 is op het gooien van kop. Je kunt bijvoorbeeld de kans uitzoeken dat een cijferreeks met een 1,2,3,4,5,6,7,8 of 9 begint. Er is een gelijke kans om al deze cijfers als eerste waar te nemen in een cijferreeks. Met al deze cijfers gaat dus een kans van $^1/_9$ samen. Deze kansen tellen allemaal op tot 1, zoals de kansregel beweert. De kans op het waarnemen van een 6 of hoger als eerste cijfer is $^4/_9$. Omdat de uitkomsten disjunct zijn, mogen de kansen op een 6,7,8 of 9 dus opgeteld worden.

Bij een random fenomeen met k aantal mogelijke uitkomsten (met dezelfde kans), is de kans op een specifieke uitkomst 1/k. De kans op gebeurtenis A is dan: P(A)= telling van de uitkomsten in A/ telling van de uitkomsten in S. Dit is hetzelfde als: telling van uitkomsten in A/k.
In de praktijk gaan veel uitkomsten niet samen met dezelfde kansen. De regel die gebruikt wordt voor een eindig aantal uitkomsten is daarom belangrijker.

Onafhankelijkheid en de vermenigvuldigingsregel
De derde kansregel stelt dat als de *één of de ander* van twee gebeurtenissen, A en B, afzonderlijk van elkaar voorkomen, dat ze dan disjunct genoemd kunnen worden. De vierde kansregel beschrijft dat de kans dat *beide* gebeurtenissen, A en B, samen voor kunnen komen. Stel dat je een munt twee keer werpt. Je wilt graag weten hoe vaak je munt hebt gegooid. De bijbehorende

kansen zijn dus: P(A) de eerste werp geeft munt en P(B) de tweede werp geeft munt. De gebeurtenissen A en B zijn in dit geval niet disjunct, je kunt ze niet optellen. De kans op twee keer munt is niet 0.5 + 0.5 = 1. Ze komen samen voor als beide worpen munt opleveren. Wij willen de kans berekenen dat beide gebeurtenissen (A en B) *beide* munt opleveren. De twee gebeurtenissen zijn niet disjunct, maar wel *onafhankelijk.* De kans dat twee keer munt geworpen zal worden is 0.5x0.5=0.25. Dit is dan ook meteen de laatste kansregel:

Regel 5: Twee gebeurtenissen A en B zijn *onafhankelijk* als het voorkomen van de ene gebeurtenis geen invloed heeft op het voorkomen van de andere gebeurtenis. Als A en B onafhankelijk zijn, dan geldt: P(A en B)= P(A)P(B). Dit noemen we ook wel de *vermenigvuldigingsregel voor onafhankelijke gebeurtenissen.* Deze regel geldt alleen voor onafhankelijke gebeurtenissen en kan dus niet voor disjuncte gebeurtenissen gebruikt worden. Disjuncte gebeurtenissen kunnen nooit onafhankelijk zijn. Het is belangrijk om disjuncte gebeurtenissen niet te verwarren met onafhankelijke gebeurtenissen. Als A en B onafhankelijk zijn, dan zijn de complementen van A en B ook onafhankelijk.

4.3 Random variabelen

Randomheid

Steekproefruimten hoeven niet uit cijfers te bestaan. Als je vier keer een munt werpt dan kunnen we de uitkomsten ook in letters beschrijven, bijvoorbeeld: MMKK. Als we het aantal koppen tellen, dan is dat in dit voorbeeld dus X=2. Bij vier keer een munt werpen, is er een kans dat er 0,1,2,3 en 4 keer kop wordt gegooid. Als je dan het aantal koppen wilt tellen, dan neemt X dus een andere waarde aan. We noemen X een random variabele.

- Een *random variabele* heeft een numerieke waarde die bij een random fenomeen hoort.
- Random variabelen korten we vaak af met hoofdletters, zoals X en Y. Als een random variabele X een random fenomeen beschrijft, dan is de steekproefruimte van S een lijst van de mogelijke uitkomsten van de random variabele.

Discrete random variabelen

Er zijn twee manieren om kansen toe te wijzen aan gebeurtenissen: discrete random variabelen en continuerende random variabelen.

Een *discrete random variabele* X heeft een eindig aantal mogelijke waarden. De *kansdistributie* van X is een lijst van de waarden en de kansen. Bij de eerste waarde van X hoort een kans, bij de tweede waarde van X hoort een kans enz. De waarden van X noemen we x_1, x_2, x_3… x_k. De bijbehorende kansen zijn p_1, p_2, p_3… p_k. De kansen moeten aan twee voorwaarden doen:

(1) Elke kans moet tussen de 0 en 1 zijn.

(2) Alle kansen samen (p_1+ p_2 + p_3 +…+ p_k) moeten optellen tot 1. De kans op een gebeurtenis kan gevonden worden door de benodigde kansen die bij bepaalde X-waarden horen op te tellen.

Continuerende random variabelen
Wanneer we een tabel van random cijfers tussen de 0 en 9 gebruiken, dan is het resultaat een discrete random variabele. Er is dan evenveel kans om random 1 van de 10 cijfers te trekken. Het is echter ook denkbaar dat we geïnteresseerd zijn in het trekken van een getal tussen de 0 en 1. Je kunt bijvoorbeeld geïnteresseerd zijn in de vraag hoe groot de kans is dat we iets tussen de 0.3 en 0.7 trekken. Er is oneindig aantal mogelijkheden tussen 0 en 1 en daarom kunnen we niet zomaar een aantal kansen optellen, zoals we dat wel kunnen doen bij discrete random variabelen. We wijzen de kansen bij een continuerende random variabele aan gebeurtenissen toe middels *gebieden onder een dichtheidscurve.*

1. Een *continuerende random variabele* X kan alle waarden in een interval van getallen aannemen.
2. De *kansdistributie* van X wordt beschreven met een dichtheidscurve. De kans op een gebeurtenis is het gebied onder de dichtheidscurve en boven de waarden van X die samengaan met de gebeurtenis.
3. Alle continuerende kansdistributies wijzen een kans van 0 toe aan elke individuele uitkomst. De dichtheidscurve die het meest voor continuerende random variabelen worden gebruikt is de Normaalverdeling. Normaalverdelingen worden ook wel kansdistributies genoemd. Als X de $N(\mu,\sigma)$ -distributie heeft, dan is de gestandaardiseerde variabele: $z= (x-\mu)/\sigma$. Deze gestandaardiseerde variabele heeft een gemiddelde van 0 en een standaard deviatie van 1: $N(0,1)$.

4.4 Gemiddelden en varianties van random variabelen

Het gemiddelde van een random variabele
Het gemiddelde van variabele X is een gemiddelde van alle mogelijk waarden van X. Het is echter niet zo dat er een even grote kans moet zijn dat elke uitkomst voorkomt. Het gemiddelde van een kansdistributie beschrijft welke waarde voor een variabele X gevonden zou worden op lange termijn.

Het *gemiddelde van een kansdistributie* wordt aangeduid met μ. Om onszelf eraan te herinneren dat we praten over het gemiddelde van X (en bijvoorbeeld niet het gemiddelde van een populatie) gebruiken we de notatie μ_x. Soms wordt het gemiddelde in dit verband ook wel de *verwachte waarde* van X genoemd. Deze term kan misleidend zijn, aangezien een waarde van X niet per se dichtbij het gemiddelde hoeft te liggen.

Discrete variabelen en continuerende variabelen
Het *gemiddelde* van een discrete variabele wordt gevonden door alle waarden van X te vermenigvuldigen met alle bijbehorende kansen en deze allemaal op te tellen.

Het gemiddelde van een continuerende random variabele wordt gevonden door te kijken naar de bijbehorende dichtheidscurve. Het gemiddelde is het punt waarop de curve balanceert als de curve van vast materiaal gemaakt zou zijn. Het gemiddelde ligt precies in het midden van symmetrische dichtheidscurven, zoals bij Normaalverdelingen. De precieze berekening van het gemiddelde van een curve met een afwijking naar links of rechts wordt met ingewikkelde rekenkundige formules uitgevoerd. Ook bij een curve met een afwijking naar links of rechts is het gemiddelde het balanspunt van de curve. Deze is alleen lastiger te ontdekken dan bij een symmetrische curve het geval is.

De wet van grote getallen

De *wet van grote getallen* stelt dat als het aantal observaties stijgt, de waarde van μ benaderd zal worden. Het moet dan wel gaan om onafhankelijke observaties die random uit de populatie getrokken worden. De benaderde waarde zal in de buurt van μ blijven liggen. Deze wet geldt voor *elke* populatie. De wet van grote getallen stelt dus dat grote steekproeven waarden met zich meebrengen die erg lijken op populatiewaarden.

Stel: we willen weten hoe lang Nederlandse vrouwen tussen de 15 en 25 gemiddeld (μ) zijn. Deze μ is de μ_x van de random variabele X, die verkregen wordt door een jonge vrouw random te kiezen en haar lengte te meten. Om μ te schatten kiezen we een SRS van jonge vrouwen en gebruiken we het steekproefgemiddelde $\bar{x}$ als schattingsmethode: μ is een *parameter* en $\bar{x}$ is een *statistiek.* Statistieken die door middel van steekproeven verkregen worden zijn random variabelen, omdat hun waarden variëren als er opnieuw een steekproef wordt getrokken. De steekproevendistributies van statistieken zijn eigenlijk de kansdistributies van deze random variabelen. Natuurlijk is $\bar{x}$ nooit helemaal gelijk aan μ en verschillende steekproeven geven vaak verschillende statistieken. Waarom is dan toch een goede schatter van het populatiegemiddelde? Het antwoord is dat $\bar{x}$ een foutloze schatter is en dat we de spreiding van $\bar{x}$ kunnen beïnvloeden aan de hand van de grootte van de steekproef. Als we steeds meer mensen onderzoeken, dan is het te *garanderen* dat $\bar{x}$ zich dichtbij het populatiegemiddelde zal gaan bevinden.

Hoe groot moet een steekproef dan zijn?
Er is geen eenduidig antwoord voor deze vraag. Hoeveel observaties gedaan moeten worden hangt namelijk af van de spreiding van de random uitkomsten. Hoe meer spreiding er in de uitkomsten waarneembaar is, hoe meer observaties nodig zijn om te garanderen dat $\bar{x}$ dichtbij μ zal liggen.

De wet van kleine getallen

De wet van grote getallen beschrijft wat er op *lange termijn* gebeurt. Als je vier keer een munt werpt, dan kan het zo zijn dat er vier keer munt uitkomt, terwijl we weten dat er een kans van 0.5 bestaat om munt te gooien. Deze kans van 0.5 uit zich echter nog niet op korte termijn. Dat patroon wordt pas zichtbaar na honderden keren werpen. Vaak verwachten we op korte termijn ook een regelmatig patroon te ontdekken, terwijl dat in werkelijkheid niet gebeurt. Deze verwachting wordt ook wel de wet van kleine getallen genoemd.

Regels voor gemiddelden

Er gelden twee regels voor gemiddelden van random variabelen:

- *Regel 1:* als X een random variabele is en a en b vastgestelde getallen zijn, dan geldt: $\mu_{a+bX} = a + b\mu_X$.
- *Regel 2:* als X en Y random variabelen zijn, dan geldt: $\mu_{X+Y} = \mu_X + \mu_Y$.

Variantie van een random variabele (σ_X^2)

De variantie is het gemiddelde van de gekwadrateerde afwijkingen $(X - \mu_X)^2$ van de variabele tot het gemiddelde (μ_X). De variantie van een discrete random variabele is handmatig uit te rekenen, terwijl dat voor een continuerende random variabele alleen met lastige rekenkundige formules lukt. De variantie van een discrete random variabele wordt hieronder beschreven.

Voorbeeld:

Je hebt verschillende waarden voor X ($x_1, x_2, x_3 \ldots x_k$) met verschillende bijbehorende kansen ($p_1 + p_2 + p_3 + \ldots + p_k$). Het gemiddelde wordt gevonden door elke waarde van X te vermenigvuldigen met elke bijbehorende proportie. Vervolgens moeten alle uitkomsten bij elkaar opgeteld worden. De *variantie* van X is: $\sigma_X^2 = (x_1 - \mu_X)^2 p_1 + (x_2 - \mu_X)^2 p_2 + (x_3 - \mu_X)^2 p_3 + \ldots + (x_k - \mu_X)^2 p_k$. De standaard deviatie σ_X wordt gevonden door de wortel uit de variantie te trekken.

Regels voor varianties en standaarddeviaties

Er gelden vier regels voor varianties en standaarddeviaties van random variabelen.

- *Regel 1:* als X een random variabele is en a en b vastgestelde cijfers zijn, dan geldt: $\sigma^2_{a+bX} = b^2\sigma^2_x$
- *Regel 2:* als X en Y onafhankelijke random variabelen zijn, dan geldt: $\sigma^2_{X+Y} = \sigma_x^2 + \sigma_y^2$. Ook geldt: $\sigma^2_{X-Y} = \sigma^2_x + \sigma^2_y$. Dit wordt ook wel de *optelregel voor varianties van onafhankelijke random variabelen* genoemd.
- *Regel 3:* als X en Y een correlatie ρ hebben, dan geldt: $\sigma^2_{X+Y} = \sigma_x^2 + \sigma_y^2 + 2\rho\sigma_X\sigma_Y$. Ook geldt: $\sigma^2_{X-Y} = \sigma_X^2 + \sigma_Y^2 - 2\rho\sigma_X\sigma_Y$. Dit is de *algemene optelregel voor varianties van random variabelen*. Om de standaard deviatie te vinden moet de wortel uit de variantie getrokken worden.

4.5 Algemene kansregels

Algemene optelregels

We weten dat als twee gebeurtenissen disjunct zijn, dat dan de kans op de ene of de andere P (A of B) een optelling van de afzonderlijke kansen is: P (A) + P (B). Wat gebeurt er als er meer dan twee gebeurtenissen zijn of als de gebeurtenissen niet disjunct zijn? In dit soort gevallen zijn algemene optelregels van toepassing.

- Een *unie* van een verzameling van gebeurtenissen is de gebeurtenis dat minstens één van de alle gebeurtenissen voorkomt.
- Als gebeurtenissen A, B en C disjunct zijn en dus geen uitkomsten met elkaar gemeen hebben, dan geldt: P(één of meer van A, B,C) = P (A) + P (B) + P (C). Deze regel is ook toepasbaar in situaties waarbij sprake is van nog veel meer gebeurtenissen.
- De kans op gebeurtenis A of B kan ook anders gevonden worden: P(A of B) = P (A) + P (B) -P(A en B). Als A en B disjunct zijn, dan is de kans op P(A en B) nul. Dit deel van de formule valt in dat geval weg. We houden dan de oorspronkelijke regel voor disjuncte gebeurtenissen over. Deze regel is eerder besproken.

Conditionele kansen

Van een *conditionele kans* wordt gesproken als we kijken naar de kans op een bepaalde gebeurtenis, gegeven het feit dat een andere gebeurtenis is voorgekomen. Je kunt bijvoorbeeld de vraag stellen hoeveel studenten Psychologie studeren gegeven dat ze man zijn. Een conditionele kans noteren we als P(A|B). Dit is te vertalen naar de kans op gebeurtenis A, gegeven dat gebeurtenis B voorkomt.

De kans dat gebeurtenissen A en B samen voorkomen wordt uitgerekend met de formule: P(A en B)= P (A) P (B | A). Hier is P(B | A) de conditionele kans dat B voorkomt, als gebeurtenis A waar is.

Bijvoorbeeld:

29% van de internetgebruikers downloadt muziek, en 67% van de downloaders maakt het niet uit of de muziek een copyright heeft. Dus het percentage van internetgebruikers die muziek downloadt (gebeurtenis A) en het niet uitmaakt of er copyright op zit (gebeurtenis B) is 67% van 29% oftewel: (0.67)(0.29) = 0.1943 = 19,43%

Of in statestiektaal:

P (A en B) = P (A) x P (B | A)

Wanneer de kans op gebeurtenis A groter dan 0 is, dan wordt de *conditionele kans* van B, gegeven A, gevonden met de formule: P (B | A)= P(A en B)/ P(A)

Bijvoorbeeld:

We hebben de tabel:

Leeftijd	**Studie**
	Voltijd
15 – 19	0.21
20 - 24	0.32
25 – 35	0.10
30 of meer	0.05

We willen de kans weten dat een student tussen de 15 en 19 jaar is, gegeven het een voltijd student is. De kans dat een student tussen de 15 en 19 jaar is en voltijd student is 0.21. Dus P (A en B) = 0.21

P (B) = 0.21 + 0.32 + 0.10 + 0.05 = 0.68

P (A|B) = P (A en B) / P (B)

= 0.21 / 0.68
= 0.31
Dus de kans dat een student tussen de 15 en 19 jaar is gegeven het een voltijd student te zijn is 31%
Als gebeurtenissen A en B *onafhankelijk* zijn, dan geldt: P(B | A)= P (B).

Intersectie
De *intersectie* van een verzameling gebeurtenissen is de gebeurtenis waarbij *alle* gebeurtenissen uit de verzameling voorkomen. De intersectie voor gebeurtenis A, B en C is daarom:
P (A en B en C) = P (A) P (B | A) P (C | A en B). Van belang zijn dus de kans op A, de kans op B gegeven A en de kans op C gegeven A en B.

Bijvoorbeeld:
5% van de sporters op de middelbare school blijft op hetzelfde niveau sporten op de universiteit. Van deze is 1,7% van de topsport. 40% van de sporters die doorgaan in universiteit en daarna topsport gaat doen heeft een carrière in sport van meer dan 3 jaar
A = doorgaan in universiteit
B = topsport
C = carrière van meer dan 3 jaar.

P (A) = 0.05
P (B|A) = 0.017
P (C| A en B) = 0.4

De kans dat A, B, en C voorkomen is
P (A en B en C) = P(A) P(B|A) (C|A en B)
= 0.05 x 0.017 x 0.4 = 0.00034
Dus maar 3 uit ieder 10,000 middelbare school sporten gaan verder op hetzelfde niveau op de universiteit en krijgen een sportcarrière van minimaal 3 jaar.

De regel van Bayes
Stel dat er A_1, A_2,.., A_k disjuncte gebeurtenissen zijn die allemaal een kans van boven de 0 hebben en samen optellen tot 1. Stel dat C een andere gebeurtenis is waarvan de kans niet 0 of 1 is. In dat geval kan de regel van Bayes toegepast worden:

P (| C) =

Onafhankelijke gebeurtenissen
Twee gebeurtenissen A en B met positieve kansen zijn onafhankelijk als:
P (B | A) = P (B)