

2 Verdelingen



Inhoudsopgave

2.0	Geboortes	3
2.1	Frequentieverdelingen	4
2.2	Kans	15
2.3	Op den duur ...	20
2.4	Simulaties	28
2.5	Rekenen met kansen	32
2.6	De som van de kansen is 1	39
2.7	Voorwaardelijke kansen	42
2.8	Extra opgaven	48
2.9	Samenvatting Verdelingen	50

Bij dit hoofdstuk hoort een *digimap*. Daarin staan opgaven waarbij er iets met ict valt te beleven. Ze hebben hetzelfde nummer als de bijbehorende opgave in deze tekst. Je herkent ze aan het icoontje . Via de digimap kom je bij VU-Statistiek of op een website terecht.

De opgaven in de digimap kunnen op drie manieren worden gebruikt:

- niet,
- aanvullend op de corresponderende opgave in deze tekst,
- vervangend voor de corresponderende opgave in deze tekst.

Bij de opgaven waarvoor WB staat hoort een werkblad.

Colofon

© 2009 cTWO

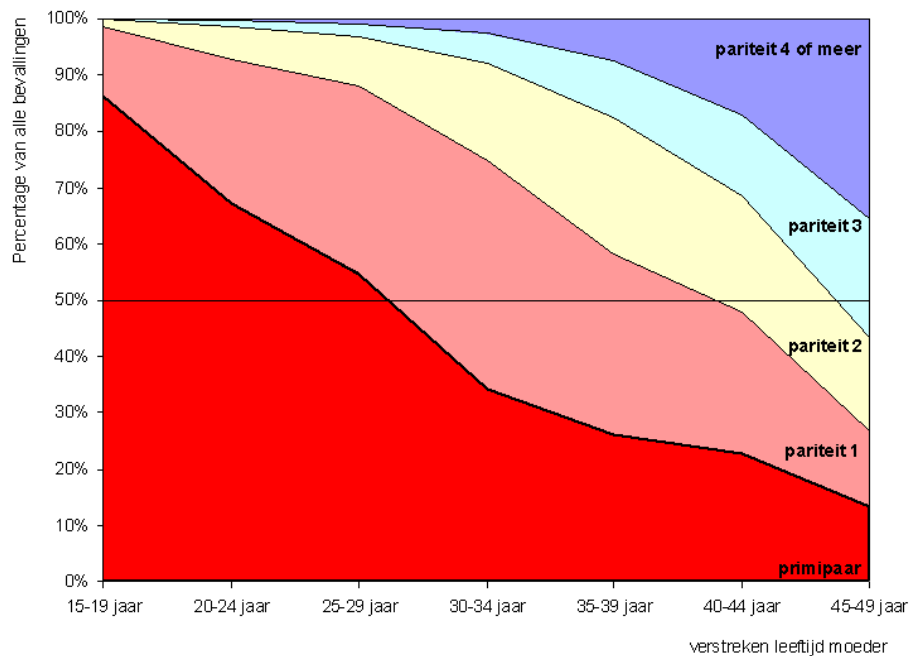
Experimentele uitgave Kansrekening en Statistiek, vwo, wiskunde A en C
versie 2 (januari. 2012)

auteurs Leon van den Broek, Maris van Haandel

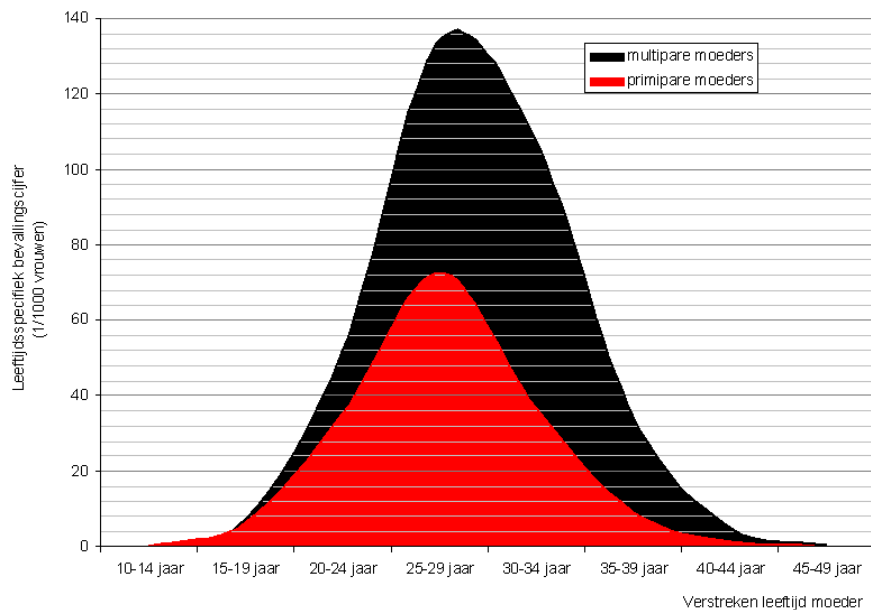
met medewerking van Simon Biesheuvel, Carel van de Giessen, Piet Versnel, Peter van Wijk

2.0 Geboortes

Procentuele verdeling van het hoeveelste kind bij een bevalling wordt geboren naar leeftijd van de moeder, Vlaams Gewest, 2006. Primipaar = eerste kind, pariteit 1 = tweede kind, enz.



Aantal bevallingen per 1.000 vrouwen, opgesplitst naar nieuwe moeders (eerste kind) en vrouwen die al een of meer kinderen hadden, Vlaams Gewest, 2006.



Bron: <http://www.zorg-en-gezondheid.be>

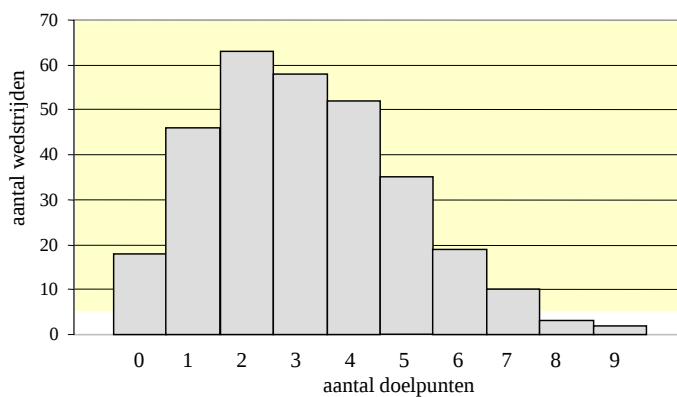
Wat kun je uit deze verdelingen aflezen?

2.1 Frequentieverdelingen

- WB** 1 In de eredivisie voetbal worden per seizoen 306 wedstrijden gespeeld. In 2006/2007 waren die als volgt *verdeeld* over het aantal doelpunten.

aantal doelpunten	0	1	2	3	4	5	6	7	8	9
aantal wedstrijden	18	46	63	58	52	35	19	10	3	2

Hieronder staat het bijbehorende histogram.

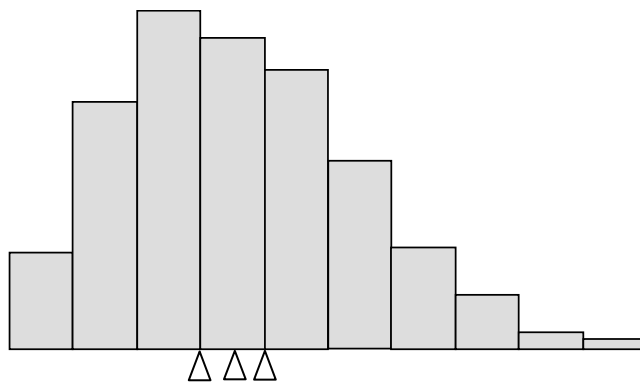


- a. In hoeveel procent van de wedstrijden werd niet gescoord?

Deze verdeling is niet symmetrisch, maar “scheef”.

- b. Wat betekent dat, denk je?

- c. Bereken het gemiddelde van het aantal doelpunten per wedstrijd in het seizoen 2006/2007.



Het gemiddelde is de "evenwichtswaarde". Daarbij hoort het volgende idee. Denk je het histogram in alsof het is uitgezaagd uit een houten plaat; plaats onder het histogram een wig op de plaats van het gemiddelde; dan balanceert het histogram netjes op deze wig.

d. Welk van de drie posities van de wig hierboven is de juiste? Denk aan het evenwicht. De verdeling kan van jaar tot jaar verschillen. Maar denk je dat de verdeling in andere jaren heel anders is? In het seizoen 1996/97 was de verdeling als volgt:

aantal doelpunten	0	1	2	3	4	5	6	7	8	9
aantal wedstrijden	17	48	79	66	53	23	11	6	2	1

We gaan de verdelingen in de twee jaren vergelijken.

e. Teken op het werkblad in dezelfde figuur als de verdeling van 2006/2007 het histogram van de verdeling in het seizoen 1996/1997.

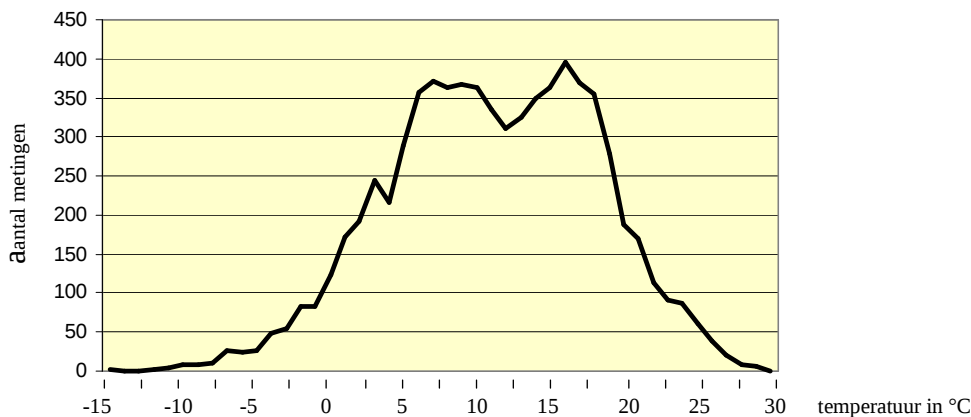
Voor het grootste deel vallen de histogrammen over elkaar heen.

f. Hoeveel procent zit niet in de overlap?

g. Vind je dat de verdelingen in de twee seizoenen veel verschillen?

2 Elk uur wordt in De Bilt de temperatuur gemeten. De resultaten van 8 uur 's ochtends in de jaren 1981 t/m 2000 geven de volgende verdeling.

a. Reken na dat er in die periode 7305 metingen zijn gedaan.



b. Hoe groot is de gemiddelde temperatuur ongeveer?

c. Er is iets verrassends (iets onlogisch) aan de verdeling. Wat?

d. Hoe groot schat jij het percentage van de dagen dat de temperatuur om 8 uur 's ochtends 20°C of hoger is.

 3 De meisjes en de jongens van het Amalia College zijn als volgt verdeeld over de vier profielen: 11 leerlingen in CM, 42 in EM, 60 in NG en 41 in NT:

a. Maak een staafdiagram van de verdeling van alle leerlingen van het Amalia College over de vier profielen.

Landelijk waren in 2007 de keuzes op het vwo als volgt over de profielen verdeeld.

14% CM, 31% EM, 36% NG en 19% NT.

b. Teken ook een staafdiagram bij deze landelijke verdeling.

- 4 De meest extreme verdelingen zijn die waarbij 100% van de leerlingen voor één profiel kiest (en de andere profielen dus geen klanten hebben). De meest vlakke verdeling is die waarbij alle vier de profielen evenveel leerlingen hebben.

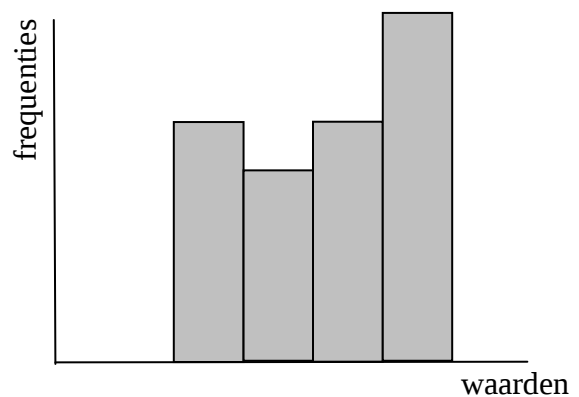
Stel dat op een school het profiel EM twee keer zo veel leerlingen heeft als CM en dat het profiel NG ook twee keer zo groot is als NT. Stel dat de M-profielen (CM en EM) anderhalf keer zoveel leerlingen hebben als de N-profielen.

Hoe is dan de frequentieverdeling?

Algemeen

In een groep letten we op een zekere eigenschap (variabele). Die eigenschap kan bijvoorbeeld vier waarden hebben. De groep is *verdeeld* over de vier waarden: elke waarde komt een zeker aantal keren voor. We spreken dan van een **frequentieverdeling**.

Als de verdeling in procenten van de totale groep is, spreken we van een **relatieve** frequentieverdeling. De som van de relatieve frequenties is 100 %.



- 5 Als je in een Engelse tekst van elk woord telt hoe vaak het voorkomt, dan blijkt het volgende. Het woordje "the" komt het vaakste voor, het woordje "of" komt op de tweede plaats, enzovoort. Hieronder staat de toptien:

1	2	3	4	5	6	7	8	9	10
the	of	and	to	a	in	that	it	is	was

Dit is zo voor elke (niet te korte) Engelse tekst. Georg Kingley Zipf, docent Duits aan de Harvard University heeft dit verschijnsel ontdekt en verder onderzocht. Dat was het begin van de kwantitatieve linguïstiek.

Maar er is meer aan de hand. Zipf constateerde ook dat "the" ongeveer 2 keer zo vaak voorkwam als "of" en 3 keer zo vaak als "and", 4 keer zo vaak als "to", enzovoort. Dat is als volgt kort samen te vatten: in een natuurlijke taal is de frequentie waarmee een woord voorkomt ongeveer omgekeerd evenredig met de rang van het woord in de frequentietabel.

Deze wet geldt zowel voor mensen met een grote woordenschat als voor mensen met een kleinere woordenschat.

Het woordje "the" kwam in een zekere tekst 1200 keer voor.

Teken een staafdiagram bij deze frequentieverdeling voor de tien meest voorkomende woordjes.

- 6 Alle kinderen die op de echopoli van kindercardiologie komen, krijgen een compleet echo-onderzoek waarbij diverse dimensies en bloedstroomsnelheden van het hart worden gemeten. Als blijkt dat het kind geen grote afwijkingen heeft aan het hart, worden deze kinderen "normaal" genoemd. Van deze normaal gescoorde kinderen wordt ook de lengte en het gewicht gemeten.

De gegevens zijn afkomstig van het Radboud Ziekenhuis te Nijmegen.

De gewichten zijn ingedeeld in klassen. Hieronder staat een histogram van de frequentieverdeling.

- a. Kun je uit onderstaand histogram aflezen om hoeveel kinderen het ongeveer gaat? Kies uit 220, 1120, 2440, 4440.

Het frequentiehistogram is verre van symmetrisch.

- b. Hoe komt dat?

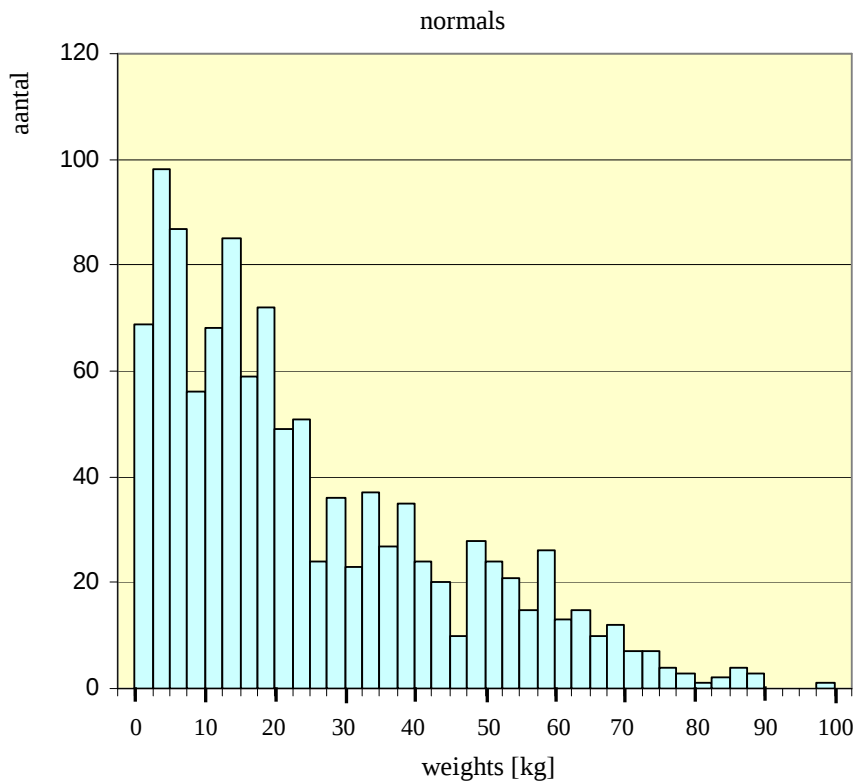
Het frequentiehistogram is niet zo regelmatig; het is nogal "springerig".

- c. Wat gebeurt er met die springerigheid als je de klassenbreedte twee keer zo groot neemt?

Bekijk het databestand via de digimap.



- d. Experimenteer met verschillende klassenbreedtes. Welke klassenbreedte heeft je voorkeur?



- 7 Old Faithful is de beroemdste en actiefste geiser in het Amerikaanse Yellowstone National Park (Wyoming). In een uitbarsting kan de geiser tot 32.000 liter kokend water zo'n 56 m hoog spuiten. Een uitbarsting duurt tussen de 1½ en 6 minuten. De toeschouwer moet soms wel geduld hebben, want de tussenpozen tussen twee uitbarstingen variëren tussen drie kwartier en twee uur.

Van 272 erupties van de Old Faithful is gegeven:

- de duur en
- de wachttijd vanaf de vorige eruptie, beide in minuten.

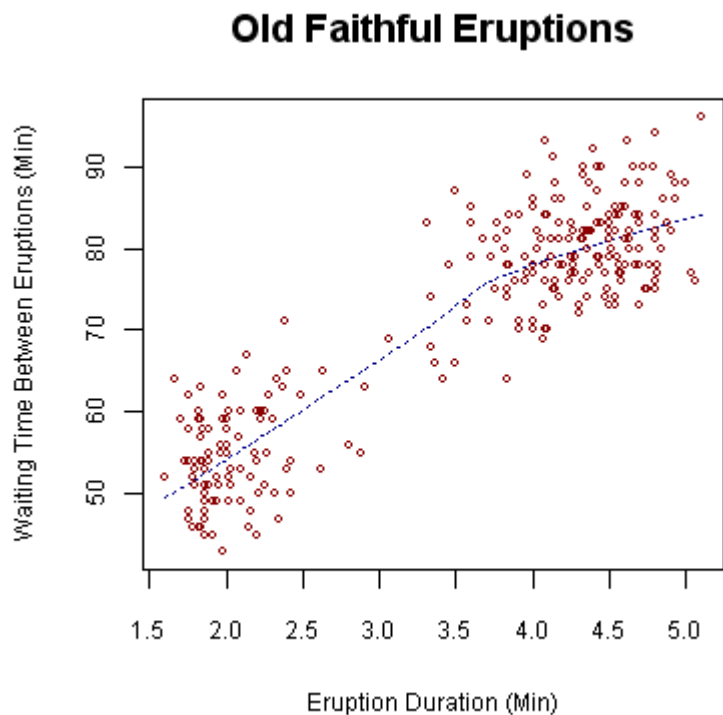
Bekijk de data via de digimap.

- Maak een histogram van de uren van de erupties. Kies zelf een geschikte klassen-indeling.
- Ook van de wachttijden.



De verdelingen van de duur en de wachttijd zijn allebei tweetoppig.

- Hoe zie je dat in nevenstaand diagram?



Van histogram naar polygoon

-  8 Hoe oud is een vrouw als ze haar eerste kind krijgt? De gegevens voor Vlaanderen (2006) staan in de tabel hieronder:

leeftijd	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54
frequentie	8	1113	5965	13302	6980	1874	237	5	2

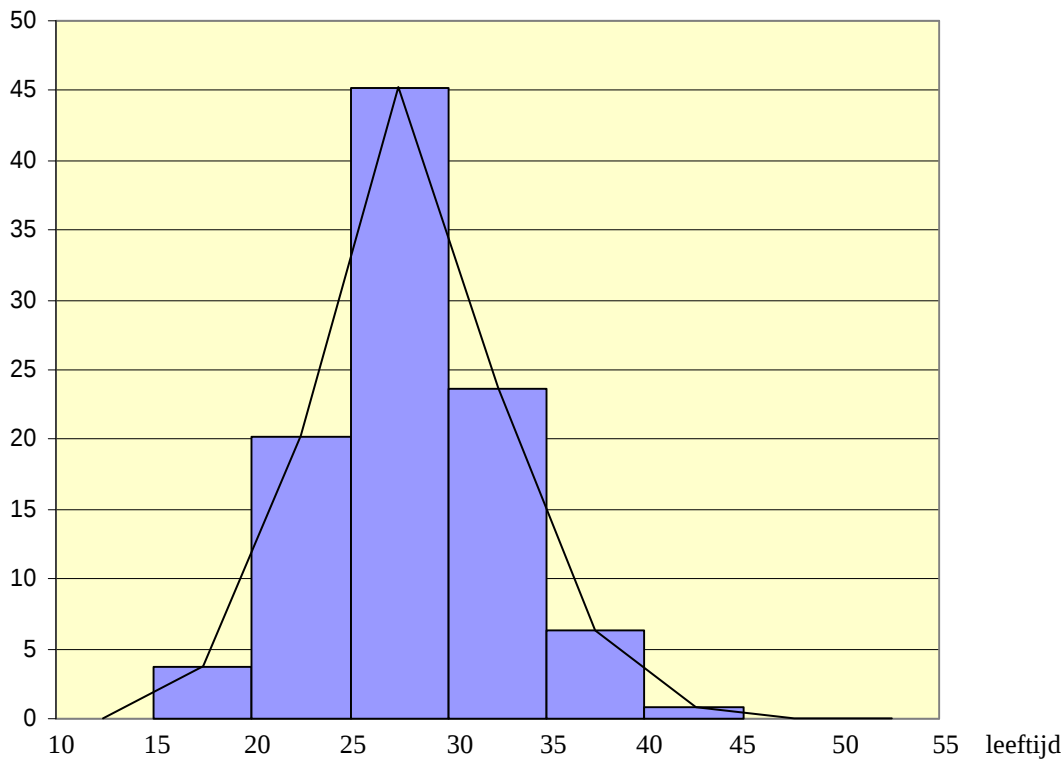
Het betreft in totaal 29486 vrouwen.

- a. Hoeveel procent van de vrouwen is jonger dan 25 jaar als ze hun eerste kind krijgen.

Als je wilt weten hoeveel procent van de vrouwen 20, 21 of 22 jaar oud is als ze hun eerste kind krijgen, heb je te maken met 3 van de 5 levensjaren in de klasse “20-24”. Een redelijke schatting van dat aantal vrouwen is $3/5$ van $5965 = 3579$. Dat klopt waarschijnlijk niet precies, maar het is het beste wat je op grond van de gegeven tabel kunt doen (zonder nadere informatie).

- b. Bepaal op grond van de tabel hoeveel procent ouder dan 26 maar jonger dan 31 jaar is?

We hebben er een procentueel histogram bij gemaakt.



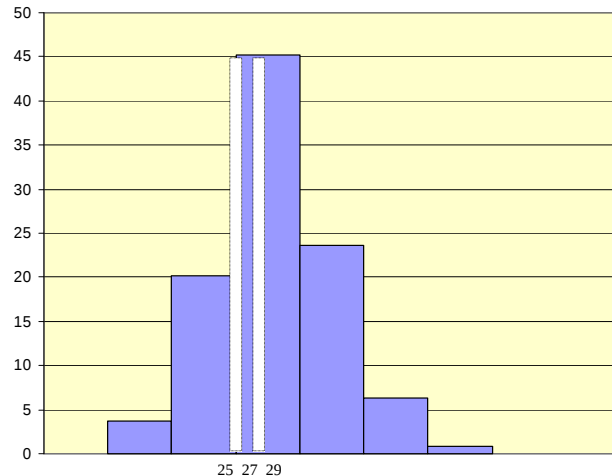
We hebben de middens van de bovenkanten van de opvolgende balken verbonden. Zodoende ontstaat er een zogenaamde **frequentiepolygoon**. Om de polygoon links en rechts op 0% te laten beginnen, is er aan weerszijden een lege klasse toegevoegd.

De hoogte van een staaf geeft het aantal of het percentage in de bijbehorende klasse. In plaats van de hoogte kun je ook op de *oppervlakte* letten.

Bijvoorbeeld in opgave 8:

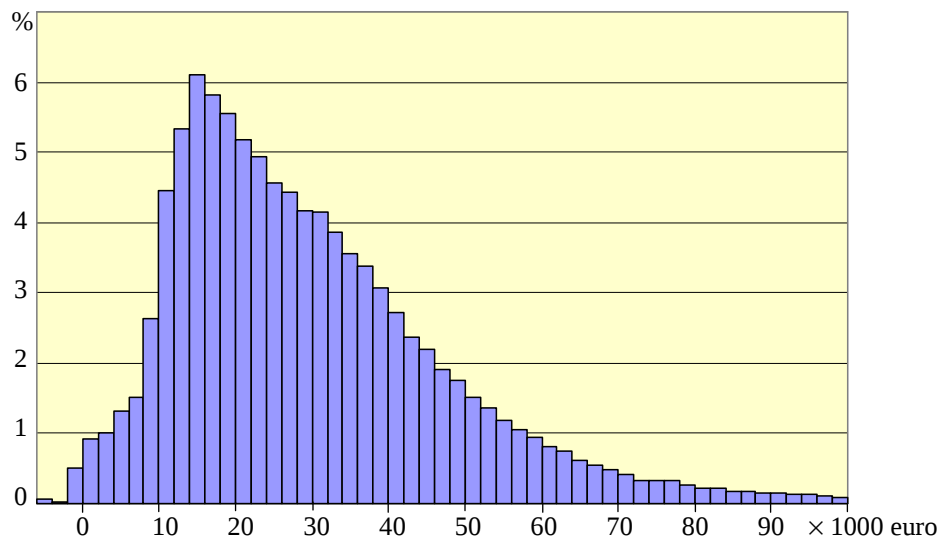
De oppervlakte van het histogram links van 27 jaar, inclusief 27 jaar zelf, geeft het percentage van de vrouwen van 27 jaar of jonger.

De totale oppervlakte van de balken is 100%.



9 Inkomensverdeling in Nederland

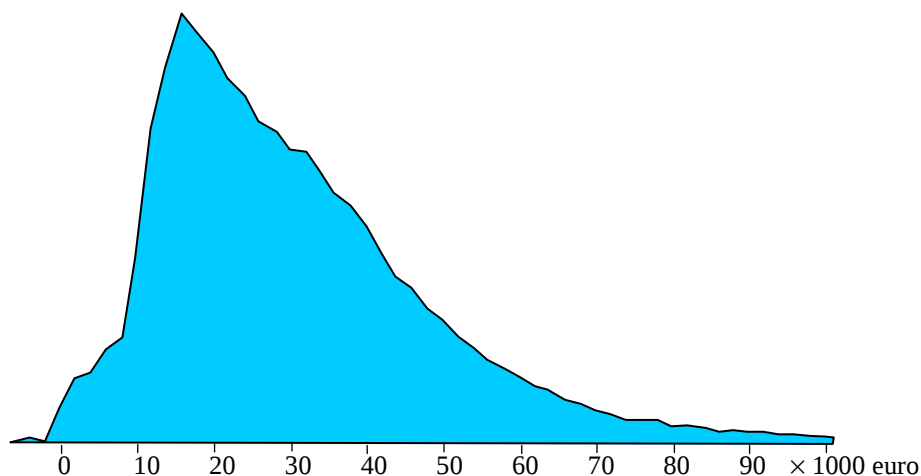
Het besteedbaar inkomen in Nederland is verdeeld in klassen van breedte 2000 euro. Er zijn 5000 gezinnen met een inkomen tussen -6000 en -4000 euro. Er zijn ook gezinnen met een besteedbaar inkomen boven de 100.000 euro, maar die zijn niet meer in de verdeling opgenomen. De gegevens zijn afkomstig van het CBS (2008)



a. Hoeveel procent van de gezinnen had een besteedbaar inkomen van minder dan 20.000 euro?

Merk op dat het totale percentage (van alle balken tezamen) 100% is.

Door de middens van de bovenkanten van de balken te verbinden, ontstaat de frequentiepolygoon.



- b. Hoe kun je – in principe – hieruit aflezen hoeveel procent van de gezinnen een besteedbaar inkomen onder de 20.000 euro heeft?

 Zie voor een aardige animatie:
<http://www.cbs.nl/nl-NL/menu/themas/inkomen-bestedingen/cijfers/extra/inkomensverdeling.htm>

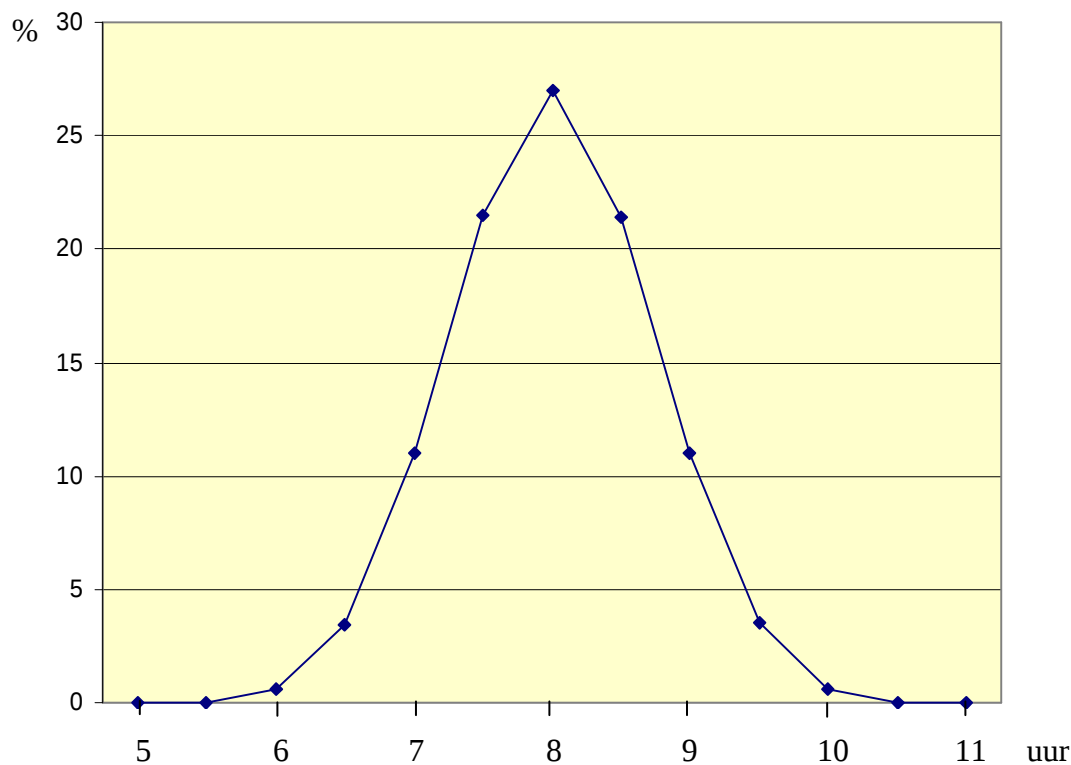
Opmerking

Het plaatje in opgave 2 is ook een frequentiepolygoon. Gevraagd werd het percentage te schatten van de dagen dat de temperatuur 's ochtends om 8 uur 20°C of hoger is. Daarvoor moest je schatten welk deel van de oppervlakte rechts van 20 °C ligt.

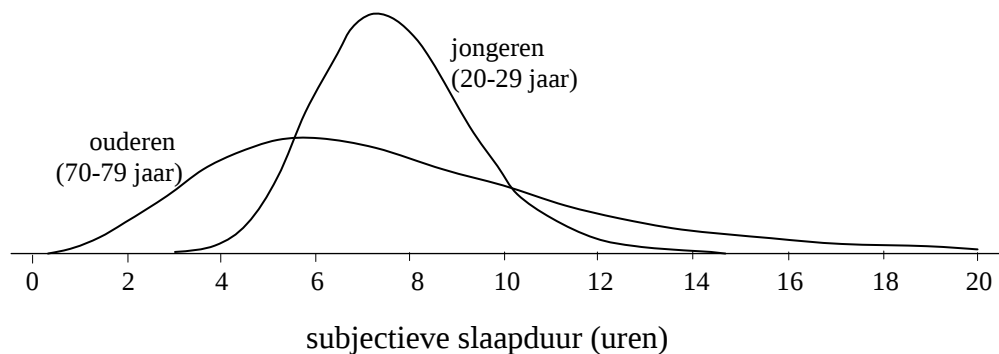
- 10 Hoeveel uur slaapt een mens? Tijdens een onderzoek is met een EEG (elektro-encefalogram) de slaapduur vastgesteld bij 147 volwassenen (16-71 jaar) gedurende 399 nachten. De gegevens zijn van werkdagen (op vrije dagen slaapt men gemiddeld 1,5 uur langer). Bron: *Waken en Slapen*, Paul A. M. van Dongen, isbn 90-74078-02-8.

Bekijk de frequentiepolygoon op de volgende bladzijde.

- a. Wat betekent de hoogste stip precies?
b. Bepaal op grond van de polygoon zo goed mogelijk hoeveel procent van de mensen meer dan 7 uur en minder dan 9 uur slaapt als hij/zij de volgende dag moet werken.



- 11 Hieronder zie je de relatieve (procentuele) verdelingen van de slaapduur van ouderen (70-79 jaar) en jongeren (20-29 jaar) (gebaseerd op onderzoek uit 1969). Het betreft de subjectieve slaapduur (zoals de mensen het zelf ervaren).



- a. Noem twee verschillen tussen deze verdelingen.

De **mediaan** (middelste) van de slaapduren is die waarde waarboven en waaronder 50% zit van alle slaapduren.

- b. Bepaal zo goed mogelijk de mediaan van de slaapduur bij de jongeren en bij de ouderen.

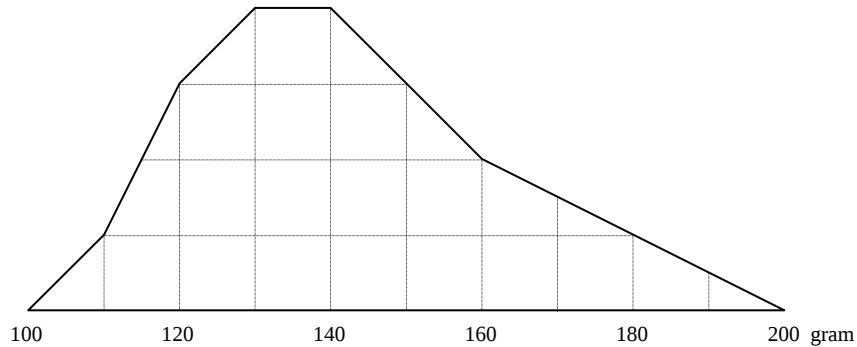
- c. De modale slaapduur is de slaapduur die het meest voorkomt.

Hoeveel procent ongeveer van de jongeren slaapt bovenmodaal? En hoeveel procent van de ouderen?

De twee krommen snijden elkaar bij 5,5. Anne zegt dat er evenveel ouderen als jongeren zijn die gemiddeld 5,5 uur slapen.

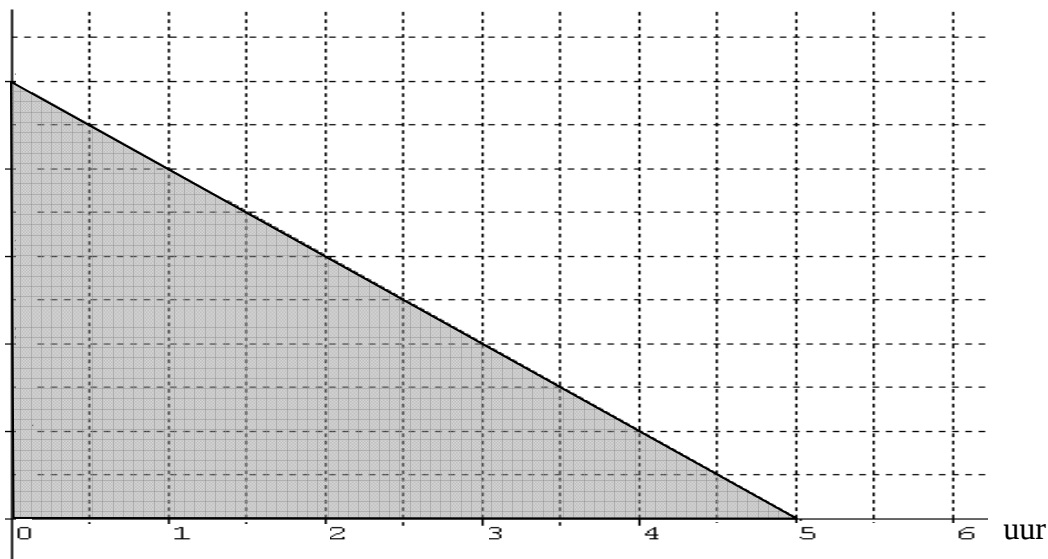
d. Is dat juist?

- 12** Spaanse exportsinaasappelen zijn gewogen. Hieronder staat de frequentiepolygoon van de gewichten.



- Hoeveel procent weegt minder dan 160 gram? Schrijf ook op hoe je te werk bent gegaan.
- Hoeveel wegen de zwaarste 10 procent (ongeveer)?
- Hoe groot is de mediaan?

- 13** We bekijken het aantal uren dat een baan per dag telt in Luilekkerland. Dat varieert tussen 0 en 5 uur per dag. Hieronder zie je hoe de banen over de verschillende groottes verdeeld zijn.

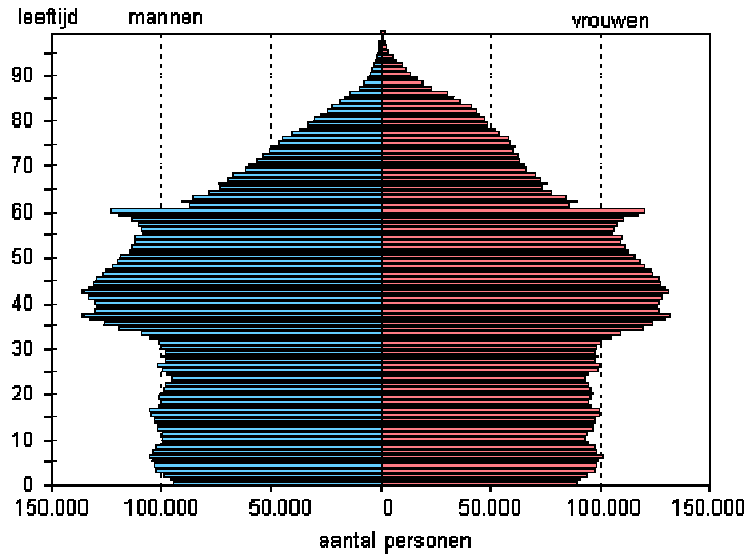


- Hoeveel procent van de banen wordt gerepresenteerd door één hokje?
- Hoeveel procent van de banen heeft een omvang tussen 1 en 2 uur per dag?
- Ga door een berekening na of een baan van 1,5 uur groter of kleiner dan de mediaan is.

- 14 Bekijk de bevolkingspiramide van Nederland op 1 januari 2007.

Een bevolkingspiramide is een grafische weergave van een frequentieverdeling, apart voor mannen en vrouwen.

- Op welk kenmerk zijn de mannen en de vrouwen verdeeld?
- Verklaar de "uitschieters" van 60-jarigen.
- Schat de mediaan van de leeftijden van de Nederlanders.
- Schat hoeveel procent ouder dan 80 jaar was (op 1 januari 2007).



Van een zekere leeftijd is 2 van de 3 Nederlanders vrouw.

- Welke leeftijd is dat (ongeveer)?



Op de site van het CBS kun je de ontwikkeling van de bevolkingspiramide in het verleden en toekomst volgen:

<http://www.cbs.nl/nl-NL/menu/themas/bevolking/cijfers/extra/piramide-fx.htm>

Opmerking

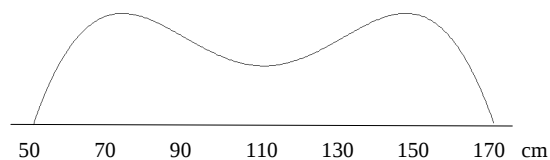
Langs de verticale as van de verdelingen in opgaven 9, 10, 12 en 13 staat niets. Dat kan ook niet. Je gebruikt de *oppervlakte* onder de grafiek, waarbij de totale oppervlakte op 100% is gesteld. Bij opgave 12 vertegenwoordigt één hokje 5%, bij opgave 13 is één hokje 2%.

Begrepen?

Hieronder staat de grafiek van de verdeling van de bevolking van Luilekkerland over de verschillende lengtes.

Op de horizontale as staat de lengte in cm.

- Wat weet je van de lengte van de mensen in Luilekkerland te vertellen?
- Wat kun je uit deze grafiek aflezen?
- Hoe doe je dat?



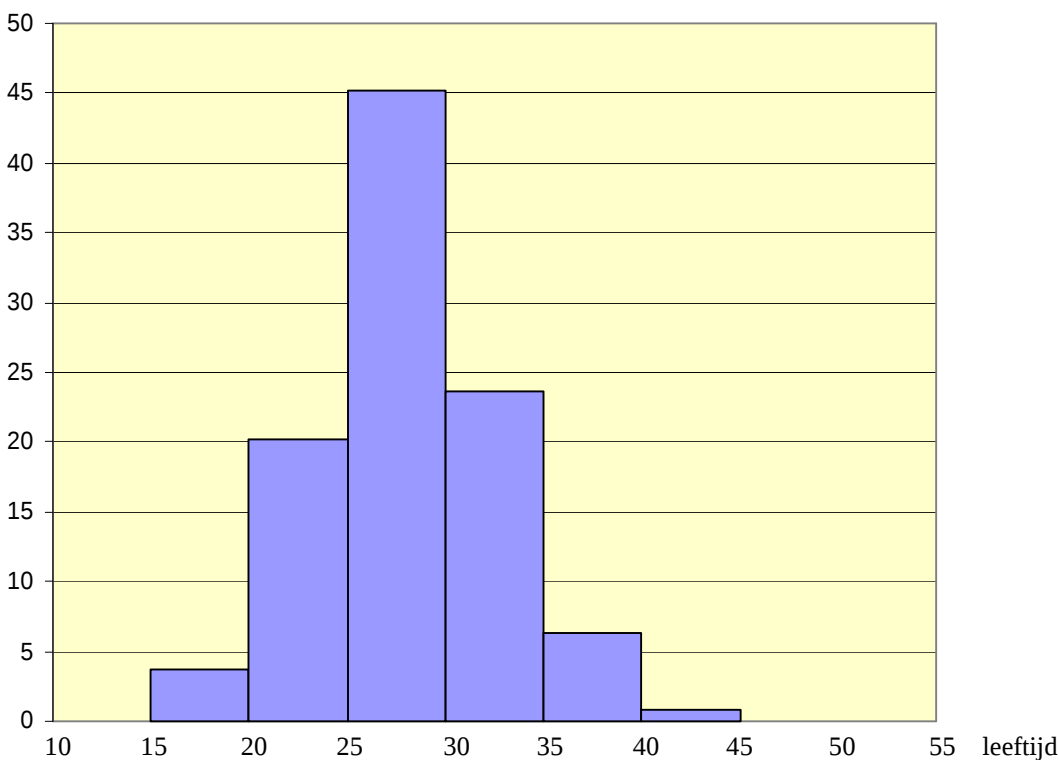
2.2 Kans

-  15 We gebruiken nog eens de context en de gegevens van opgave 8. Hieronder staat weer de tabel voor de 29486 bevallingen van het eerste kind in Vlaanderen, 2006.

leeftijd	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54
frequentie	8	1113	5965	13302	6980	1874	237	5	2

- a. In hoeveel procent van de bevallingen is de vrouw jonger dan 20 jaar?

Ook herhalen we het histogram.



- b. Wat betekent je antwoord op vraag a voor het histogram?

Een Vlaamse vrouw gaat zo dadelijk bevallen van haar eerste kind. We kennen de vrouw verder niet. Hoe oud kan ze zijn? Zeker tussen de 10 en 54 jaar.
In 3,8% van de bevallingen is de vrouw jonger dan 20 jaar.
We zeggen ook wel: **de kans dat de vrouw jonger dan 20 jaar is, is 3,8%.**

- c. Wat is de kans dat de vrouw jonger is dan 30 jaar en ouder dan 19 jaar?

-  **16** Hoeveel doelpunten vallen er per voetbalwedstrijd in de eredivisie? Dat staat voor het seizoen 2010/2011 in onderstaande tabel. Er zijn per seizoen 306 wedstrijden.

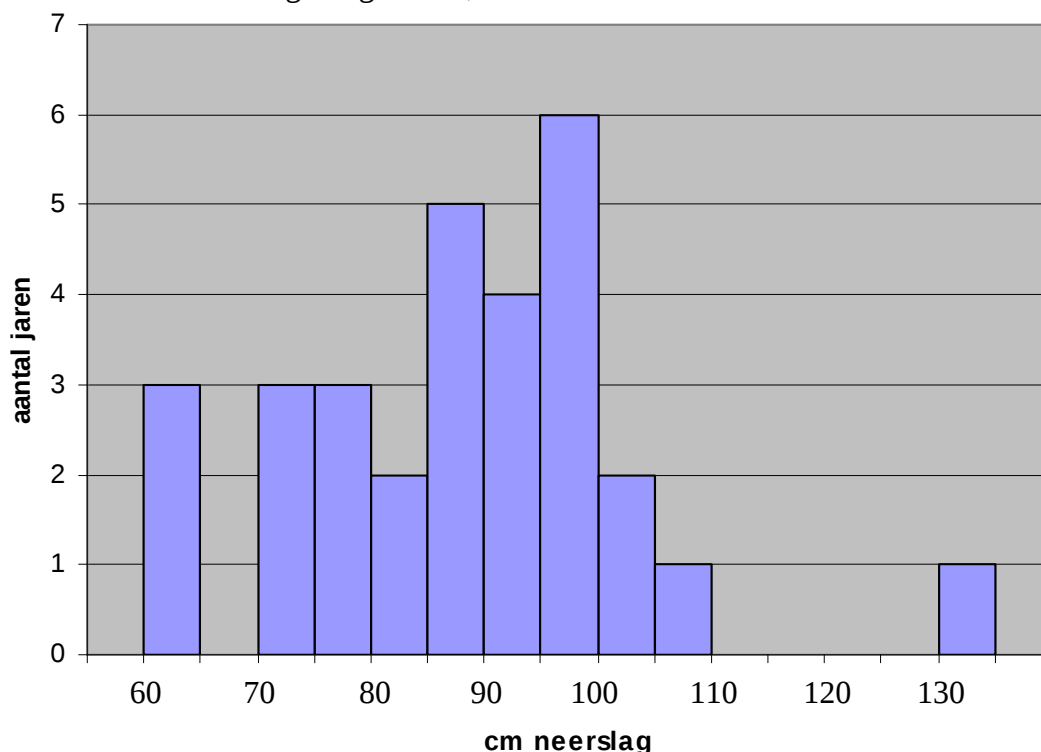
aantal doelpunten	0	1	2	3	4	5	6	7	8	9	10
aantal wedstrijden	18	34	66	65	53	35	21	7	5	1	1

We nemen aan dat deze tabel voor elk seizoen geldt.

Anne heeft geen verstand van voetballen en gaat met haar vriend naar een voetbalwedstrijd.

- Wat is de kans dat die in 0-0 eindigt?
- Wat is de kans dat Anne ten minste vijf doelpunten gaat zien?

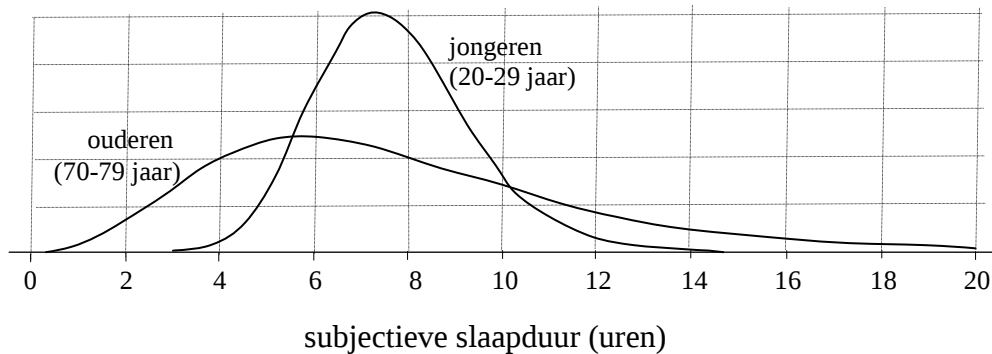
- WB**  **17** Jaarlijks wordt de hoeveelheid neerslag in De Bilt geregistreerd. Van de resultaten in de jaren 1980 t/m 2009 is een histogram gemaakt, met klassenbreedte 5 cm.



Neem aan dat het neerslagpatroon sindsdien niet is gewijzigd.

- Wat is dan de kans dat in het komend jaar meer dan 1 meter neerslag valt in De Bilt?
 - Wat is de kans dat er tussen de 9 dm en 1 meter neerslag valt?
 - Teken op het werkblad de frequentiepolygoon.
 - Leg uit dat de oppervlakte onder de polygoon even groot is als de totale oppervlakte van het histogram (dus 30).
 - Bepaal de kans in de vragen **a** en **b** nog eens met de polygoon.
- Merk op dat het histogram en de polygoon niet precies dezelfde kans geven.

- 18 Hieronder zie je nog eens de relatieve (procentuele) verdelingen van de slaapduur van ouderen (70-79 jaar) en jongeren (20-29 jaar).



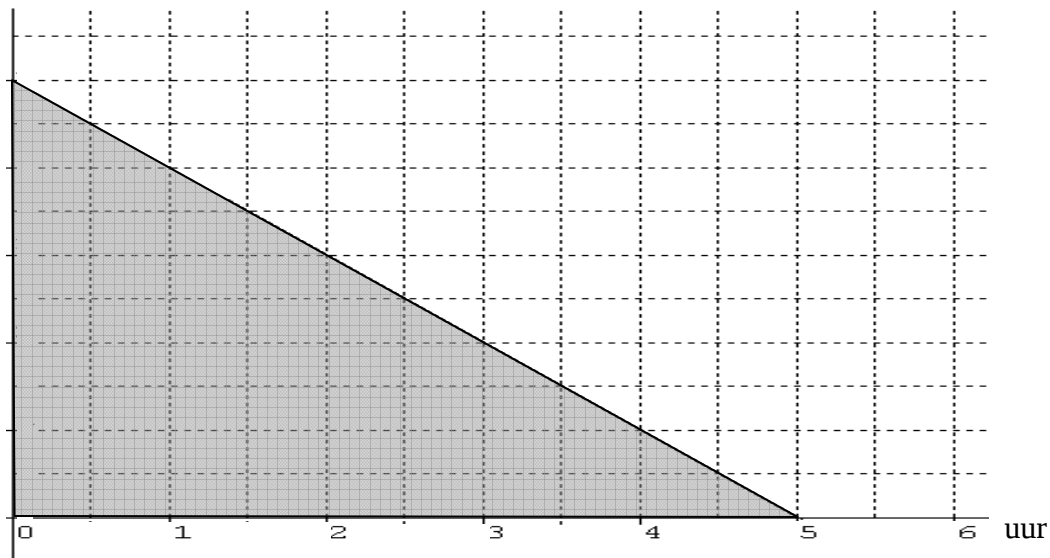
Jan maakt kennis met de opa van zijn nieuwe vriendin. Die behoort tot de ouderen.

- a. Wat is de kans ongeveer dat opa denkt dat hij minder dan 4 uur per nacht slaapt?

Jan is een jongere van 24 jaar. We kennen hem niet en weten niets van zijn slaapgewoontes.

- b. Wat is de kans ongeveer dat Jan tussen 6 en 10 uur per nacht denkt te slapen.

- 19 Luilekkerland bevalt ons wel; we brengen het opnieuw een bezoek. De omvang van een baan ligt in Luilekkerland tussen 0 en 5 uur. Hieronder zie je hoe de banen over de verschillende groottes verdeeld zijn. In opgave 13 vroegen we naar percentages; nu vragen we naar kansen. Verder zijn het precies dezelfde vragen.



- a. Wat is de kans dat een willekeurige Luilekkerlander een baan heeft met een omvang tussen 1 en 2 uur?
- b. Wat is de kans dat een willekeurige Luilekkerlander een baan heeft van ten hoogste 1,5 uur?

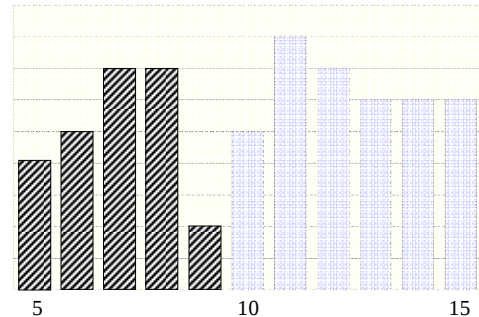
Een groep is verdeeld op een kenmerk.

Stel dat het kenmerk een discrete variabele is, die als waarden de gehele getallen van 5 t/m 15 kan aannemen.

Van de verdeling is het relatieve frequentie-staafdiagram gegeven: de percentages waarmee elk van de waarden 5 t/m 15 voorkomt zijn dus bekend.

De kans op een waarde kleiner dan 10 is dan de som van de percentages van de waarden 5 t/m 9.

Dat is $\frac{\text{oppervlakte van de staven bij 5, 6, 7, 8 en 9}}{\text{oppervlakte van alle staven}}$.

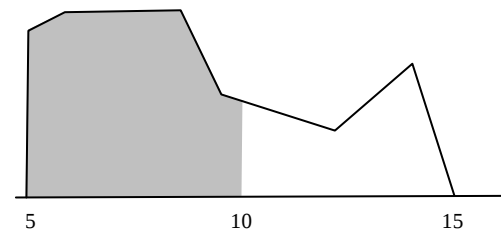


Stel dat het kenmerk een continue variabele is, die als waarden alle getallen van 5 t/m 15 kan aannemen.

Van de verdeling is de relatieve frequentie-polygoon gegeven: voor elk interval is het percentage bekend dat de waarde daarin ligt.

De kans op een waarde kleiner dan 10 is dan hoeveel procent het deel tussen 5 en 10 is van het geheel.

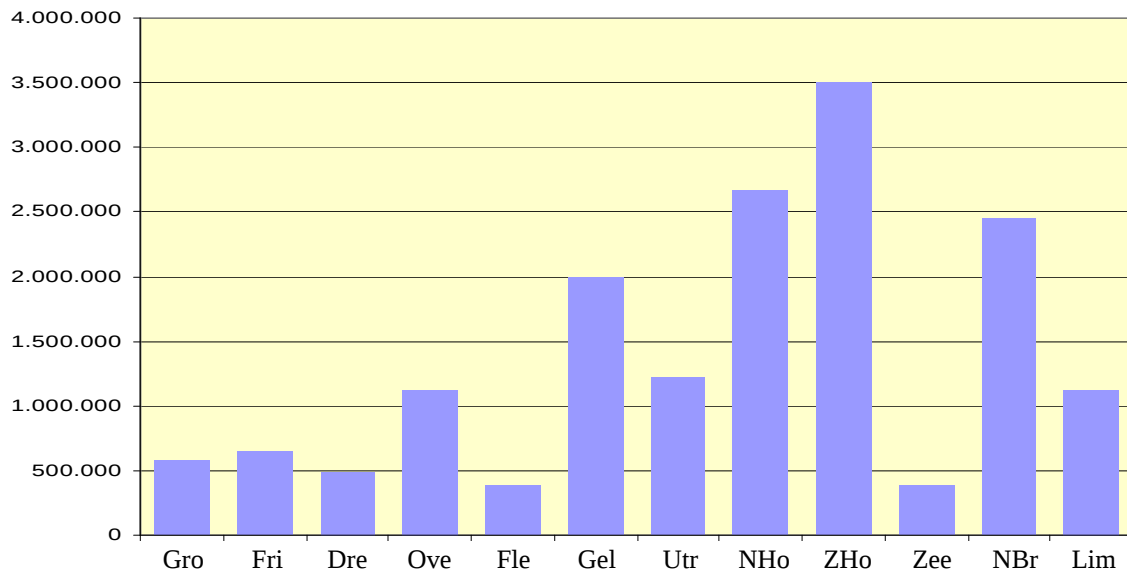
Dat is $\frac{\text{oppervlakte links van 10}}{\text{totale oppervlakte}}$.



Een kans laat zich uitdrukken als een percentage (tussen 0% en 100%) en ook als een (verhoudings-) getal (tussen 0 en 1).

Begrepen?

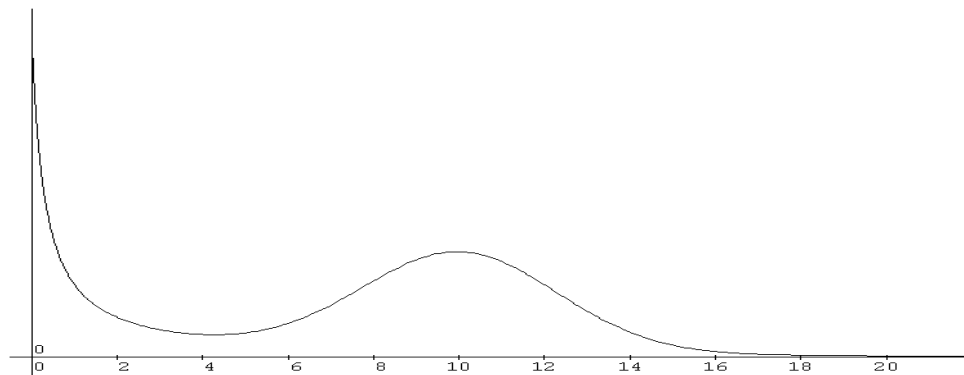
De bevolking van Nederland is verdeeld over twaalf provincies, zoals te zien is in onderstaand staafdiagram.



Wat is de kans dat de ministerpresident in 2050 afkomstig is uit een zuidelijke provincie (Noord Brabant, Limburg of Zeeland)?

Neem aan dat alle provincies gelijkwaardig zijn. Dat wil zeggen dat er niet bij voorbaat een voorkeur dat de ministerpresident uit een bepaald deel van Nederland komt.

Hoe oud wordt een kat? Dat kun je hieronder aflezen.

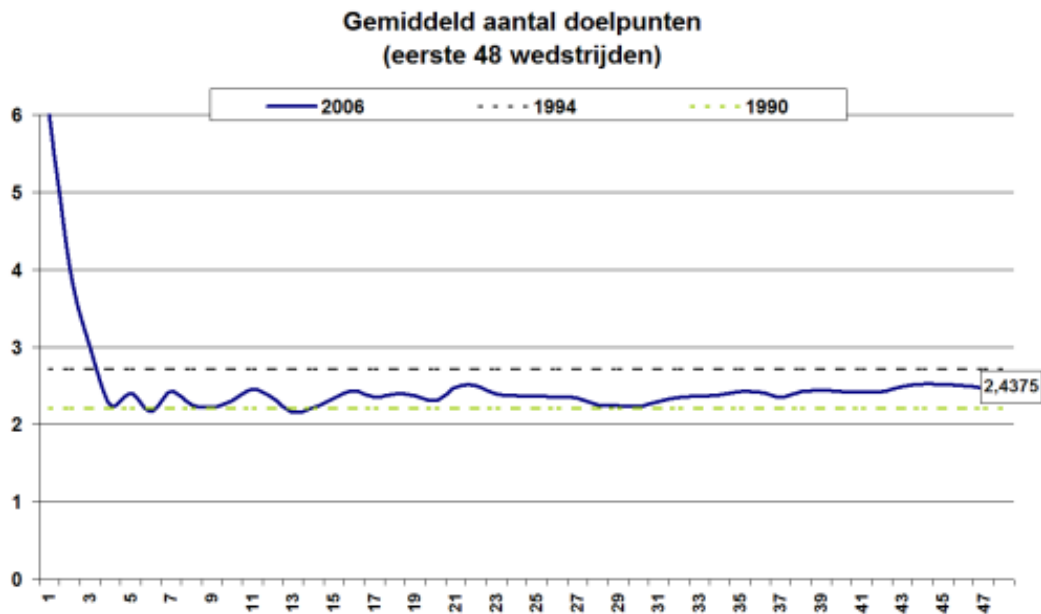


Er is zojuist weer een katje geboren.

Wat is de kans ongeveer dat het katje ouder dan 12 jaar wordt?

2.3 Op den duur ...

-  20 In 2010 stond de onderstaande grafiek op het internet. Hij betreft het gemiddelde van het aantal doelpunten per wedstrijd dat gemaakt is tijdens het WK Voetbal 2006 in Duitsland. Horizontaal is het aantal gespeelde wedstrijden uitgezet, verticaal het gemiddelde aantal doelpunten *tot dan toe*.



- a. Hoeveel doelpunten werden er in de tweede wedstrijd gemaakt?
b. Hoeveel doelpunten werden er in de zevende wedstrijd gemaakt?

Wat opvalt is dat de schommelingen in de grafiek in het begin nogal heftig zijn, maar dat ze verderop veel minder voorkomen. En dat is niet (alleen maar) toevallig.

- c. Leg dat uit.

Series herhalingen



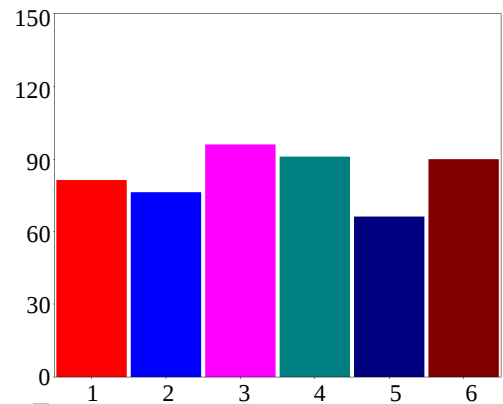
- 21 Iemand heeft 500 keer met een dobbelsteen geworpen. Het resultaat staat hiernaast.

Dit is goed te simuleren met VU-statistiek, Simulatie, Dobbelstenen (via de digimap).

66 van de 500 keer viel de dobbelsteen op 5 ogen. Anne zegt dat de kans op 5 ogen dus

$$\frac{66}{500} = 0,132 \text{ is.}$$

- Wat is daarop je commentaar?
- Hoe vaak verwacht je 5 ogen in een volgende serie van 500 worpen?

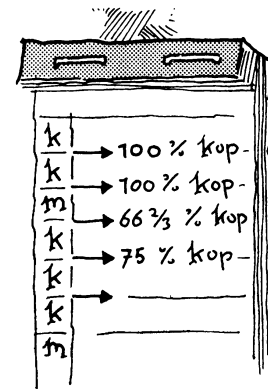


- 22 Er is een serie van twintig worpen met een munt gedaan:

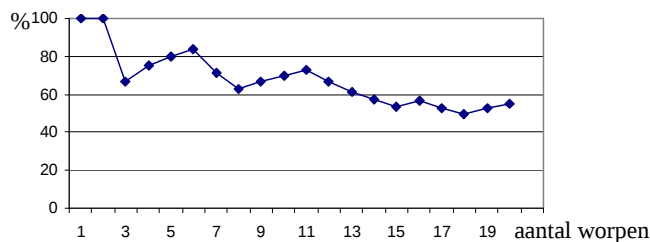
k k m k k k m m k k k m m m m k m m k k.

Na elke worp werd het percentage kop genoteerd dat tot dan toe was gegooid. Na de eerste worp en na de tweede worp was dat dus 100%. Na drie worpen hadden we: k k m. Toen was het percentage kop $66\frac{2}{3}\%$.

- Wat werd het percentage kop na zes worpen? En na de twintigste worp?

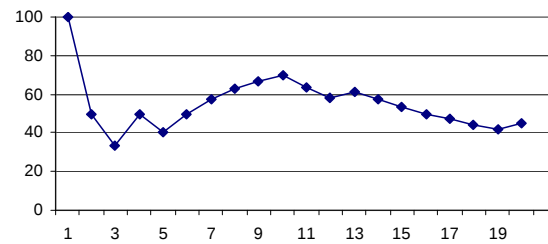
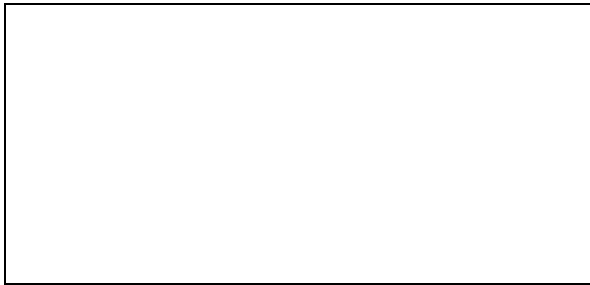


Hieronder staat een grafiek van het verloop van het percentage:



- Controleer je antwoorden op vraag a in de grafiek.
- Stel dat er nog een eenentwintigste worp zou volgen. Wat wordt het percentage kop als deze eenentwintigste worp kop zou zijn? En wat als hij munt zou zijn?

We laten de computer nog twee series van twintig worpen produceren.



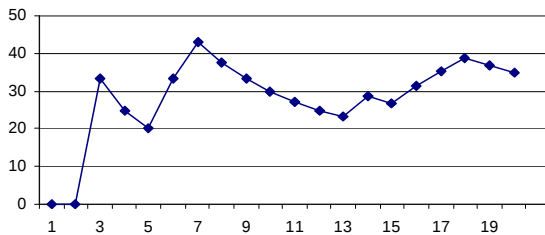
De twee plaatjes lijken op elkaar, maar zijn niet *precies* hetzelfde: het begin is heel verschillend, de staarten lijken op elkaar.

d. Waarom verbaast je dat niet?

Een munt valt gemiddeld even vaak op kop als op munt. Op de computer is dat gemakkelijk te veranderen. Stel dat we kop gemiddeld 9 keer zo vaak laten optreden als munt. Weer doen we een serie van 20 worpen en tekenen daar een plaatje bij.

e. Wat zal het grote verschil zijn tussen dit plaatje en de vorige plaatjes?

We veranderen de gemiddelde frequentie van kop opnieuw. Dat leverde het volgende plaatje op.



f. Wat is de verhouding van de frequenties *kop* : *munt* nu ongeveer, denk je?

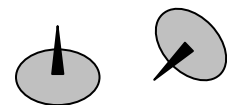
Wat is de kans op kop ongeveer?

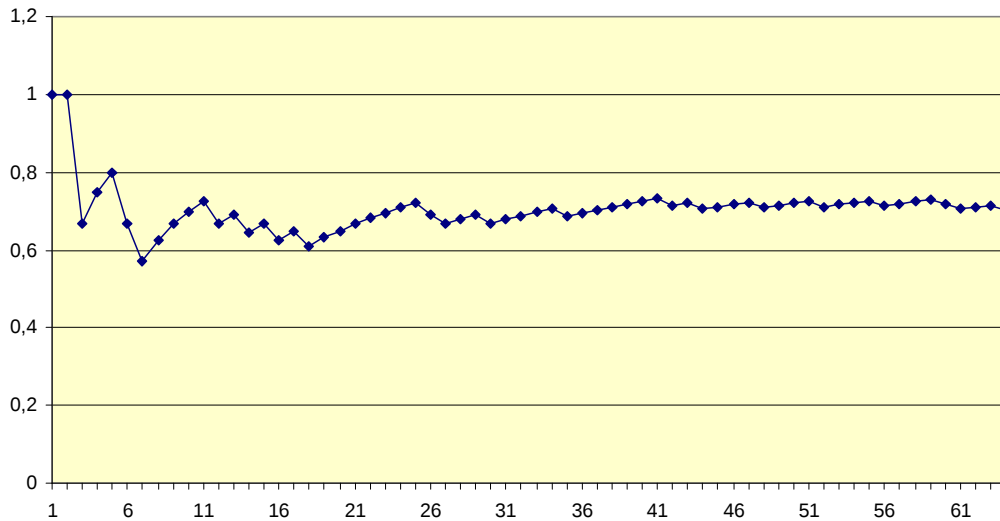


23 We kunnen de computer ook met een dobbelsteen laten gooien. Steeds kijken we of het aantal ogen 1 is of niet. We doen een serie van 1000 worpen. Na elke worp bepalen we het percentage 1'en tot dan toe. Je kunt daar weer een plaatje bij tekenen.

a. Wat kun je van dat plaatje zeggen?

Bij een dobbelsteen en een munt is bekend met welke gemiddelde frequentie de mogelijke uitkomsten optreden. Tenminste, als er niet met de dobbelsteen of met de munt geknoeid is (dan komt elk van de mogelijke uitkomsten gemiddeld even vaak voor). Als je met een punaise werpt, kan die in twee standen eindigen: met de punt omhoog of met de punt op de grond. In hoeveel procent van de worpen die twee standen voorkomen, is niet bekend. Hieronder staat het plaatje waarbij we de relatieve frequentie-tot-dan-toe van de stand "punt omhoog" hebben gevolgd.





- b. Wat zijn de eerste tien worpen? (Kort af: "punt omhoog" = H , "punt omlaag" = L)
- c. Hoe zie je gemakkelijk in het plaatje bij een worp in welke stand de punaise viel?
- d. Hoe groot schat jij dat de relatieve frequentie op den duur van de stand "punt omhoog" is?
Hoe groot is de kans op "punt omhoog" ongeveer?

24 Wordt het een jongen of een meisje?

Er worden ongeveer evenveel jongens als meisjes geboren, maar niet precies evenveel. De verhouding *jongen: meisje* is ongeveer 513 : 487.

- a. Hoe kan men dat weten?
- b. In Nederland worden jaarlijks zo'n 196000 kinderen geboren. Hoeveel jongens en hoeveel meisjes ongeveer?
- c. Hoeveel jongens ongeveer worden er per 1000 meisjes geboren?
- d. Wat is de kans op een jongen?
- e. Je zou dus verwachten dat er meer mannen dan vrouwen zijn. Maar dat is juist niet zo: op 1 januari 2011 waren er 8.243.482 mannen en 8.412.317 vrouwen in Nederland. Heb je hier een verklaring voor?

- 25 Als je twee raszuivere groene erwten kruist, krijg je 100% raszuivere groene nakomelingen. (Raszuiver betekent dat de erwt alleen drager is van de groene eigenschap.) Als je twee raszuivere gele erwten kruist, krijg je 100% raszuivere gele nakomelingen. Wat gebeurt er nu als je een raszuivere groene erwt kruist met een raszuivere gele erwt? De nakomelingen zijn dan allemaal geel, maar niet raszuiver! Dat blijkt uit de tweede nakomelingen (de nakomelingen van de nakomelingen); daar zijn zowel groene als gele exemplaren bij. Wel zijn er meer gele dan groene.

De Tsjechische monnik Gregor Mendel deed uitgebreide experimenten met erwten. Hij bestudeerde de overerving van zeven verschillende eigenschappen. Dat waren: vorm en kleur van de zaden, vorm en kleur van de peulen, lengte van de stengels. De tweede nakomelingen telden 6022 exemplaren met gele zaden en 2001 met groene zaden. Dit en de andere aantallen staan in het overzicht hieronder.

Eigenschappen van de ouders	Eerste nakomelingen	Tweede nakomelingen
Ronde zaden  X Gerimpelde zaden 	Ronde zaden 	5474  : 1850 
Gele zaden  X Groene zaden 	Gele zaden 	6022  : 2001 
Rode bloemen  X Witte bloemen 	Rode bloemen 	705  : 224  (3,15 : 1)
Gladde peulen  X Ingeseerde peulen 	Gladde peulen 	882  : 299 
Groene peulen  X Gele peulen 	Groene peulen 	428  : 152 
Losse bloemen  X Trosvormige bloemen 	Losse bloemen 	615  : 207 
Lange stengels  X Korte stengels 	Lange stengels 	787  : 277 



Gregor Mendel
1822 - 1884

- Wat is de verhouding tussen de aantallen met gele zaden en met groene zaden afgerond?
- Hoe is die verhouding bij elk van de andere zes eigenschappen ongeveer?

Mendel trok op grond van deze resultaten de volgende conclusie. Bij kruising tussen twee raszuivere variëteiten - de een met een zekere eigenschap, de ander zonder die eigenschap - zullen de tweede generatie nakomelingen die eigenschap wel of niet hebben. En wel in een vaste verhouding.

- Tot welke verhouding besloot Mendel, denk je?

Als je er vanuit gaat dat er inderdaad een vaste verhouding is, kun je van alle kenmerken de grote aantallen optellen en ook de kleine aantallen, en de verhouding tussen de twee sommen bekijken.

- Bevestigt die verhouding je antwoord op vraag c?
- Is het niet verontrustend dat bij geen van de zeven eigenschappen die Mendel onderzocht deze mooie verhouding precies uitkwam?

De wet van de grote aantallen

Voorbeeld

Je gooit met een (oneerlijke) munt die met kans 0,4 op kop valt.

In een serie van 1000 worpen zal het aantal keer kop in de buurt van 400 liggen.

De relatieve frequentie van het aantal kop zal dus dicht bij $\frac{400}{1000}$ liggen. (Maar precies $\frac{400}{1000}$ zou wel erg toevallig zijn.)

Bij een experiment kan iets gebeuren of niet. Als het gebeurt, spreken we van een "treffer".

Veronderstel dat de kans op een treffer 0,4 is.

Als het experiment vaak herhaald wordt, nadert het gemiddelde aantal treffers altijd tot 0,4.

Dit staat bekend als de *wet van de grote aantallen*.

De wet van de grote aantallen is voor het eerst geformuleerd in 1689 door de Zwitser Jacob Bernoulli.

Deze wet zegt dat je de ware kans op een treffer kunt achterhalen door (zeer) vaak het experiment te herhalen. Hoe vaak, dat weet je niet van tevoren.



Theoretische kansen

De kans op vijf ogen bij een worp met een dobbelsteen is $\frac{1}{6}$.

We gaan nader op deze uitspraak in.

Als je vaak met een dobbelsteen werpt, zal je ongeveer 1 op de 6 keer 5 ogen krijgen. Bij de ene serie zul je er wat minder vijven hebben, bij een andere wat meer.

In de praktijk blijken dobbelstenen niet perfect te zijn. De meeste hebben een systematisch verschil met die "1 op 6". De afwijking is weliswaar niet zo groot, maar duidelijk aanwezig.

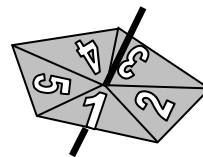
Toch gaat men er bij berekeningen vanuit dat de zes grensvlakken van een dobbelsteen gemiddeld even vaak boven zullen komen te liggen (even waarschijnlijk zijn). Er is immers geen duidelijke reden waarom 5 ogen vaker zou voorkomen (een grotere kans zou hebben) dan 2 ogen. En dan moet de totale kans 1 (=100%) dus eerlijk verdeeld worden over de zes mogelijke aantallen: allemaal $\frac{1}{6}$. In de wiskunde gaan we, als we met een dobbelsteen werpen, ervan uit dat die ideaal is: de kans op elk van de aantallen ogen is exact $\frac{1}{6}$.

Zo ook bij andere kansmechanismen.

26 a. Wat zijn de kansen bij het werpen van een munt?

Een regelmatige vijfhoek is verdeeld in vijf driehoeken, genummerd 1, 2, 3, 4 en 5. Steek een pin door het middelpunt en je hebt een kanstol.

b. Wat zijn de kansen bij het draaien van deze kanstol?



Iemand pakt (willekeurig) een kaart uit een kaartspel en let op de "kleur" (klaveren, ruiten, harten, schoppen).

c. Wat zijn de kansen?

Iemand pakt (willekeurig) een kaart uit een kaartspel en let op de "rang" (2 t/m 10, Boer, Vrouw, Heer, Aas)

d. Wat zijn de kansen?

In de trommel van de lotto zitten de ballen 1 t/m 45. Er wordt een bal uit de trommel getrokken; we letten op het nummer.

e. Wat zijn de kansen?

Iemand beantwoordt een driekeuzevraag puur op de gok. Eén van de drie alternatieven is goed. We letten op goed/fout.

f. Wat zijn de kansen?

De randomgenerator van je rekenmachine produceert "toevalsgetallen" van tien cijfers achter de komma, bijvoorbeeld 0,1058347538.

g. Wat is de kans op elk van die toevalsgetallen?

Bij een loterij op een braderie worden 100 loten verkocht. Miss Braderie zal tien winnende loten trekken. Iemand koopt een lot en wint een prijs of niet.

h. Wat zijn de kansen?

Iemand kiest willekeurig een getal van twee cijfers, dus van 10 t/m 99. We letten op het aantal even cijfers. Dat kan 0, 1 of 2 zijn.

i. Wat zijn de kansen?

Bij een toevalsexperiment zijn er verschillende uitkomsten mogelijk. De totale kans van 100% is *verdeeld* over de verschillende uitkomsten. Die verdeling noemen we een **kansverdeling**.

Veronderstel dat er zeven uitkomsten mogelijk zijn. Als die gelijkwaardig zijn (gemiddeld even vaak voorkomen), zeggen we dat elk van die uitkomsten kans $\frac{1}{7}$ heeft.

Als drie van de zeven uitkomsten speciaal zijn (en de andere vier niet), is de kans op een speciale uitkomst $\frac{3}{7}$.

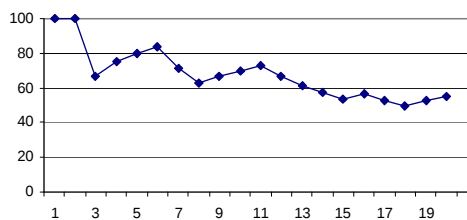
Soms is het verstandig kansen als decimale breuken te schrijven (en die af te ronden):

$\frac{1}{7} = 0,142857... \approx 14\%$ en $\frac{3}{7} = 0,428571... \approx 43\%$.

Begrepen?

Iemand werpt herhaaldelijk met een (eventueel niet geheel zuivere) munt. Hiernaast staat het verloop van de relatieve frequentie van het aantal keer kop t/m de twintigste worp.

- Hoe groot schat jij de kans op kop?
- Hoeveel zal je schatting na de eenentwintigste worp veranderen?
- Hoe ziet het verloop van de relatieve frequentie van het aantal keer munt eruit t/m de twintigste worp?



We werpen met drie dobbelstenen en willen weten wat de kans is op drie gelijke aantallen ogen. Hiernaast staan alle mogelijke resultaten.

- Wat is de kans op 543?
- Wat is de kans op drie gelijken?
- Zeg precies hoe je de kans op drie kop bij het werpen met drie munten kunt achterhalen?

111 112 113 114 115 116
131 132 133 134 135 136
151 152 153 154 155 156

211 212 213 214 215 216
231 232 233 234 235 236
251 252 253 254 255 256

311 312 313 314 315 316
331 332 333 334 335 336
351 352 353 354 355 356

411 412 413 414 415 416
431 432 433 434 435 436
451 452 453 454 455 456

511 512 513 514 515 516
531 532 533 534 535 536
551 552 553 554 555 556

611 612 613 614 615 616
631 632 633 634 635 636
651 652 653 654 655 656

121 122 123 124 125 126
141 142 143 144 145 146
161 162 163 164 165 166

221 222 223 224 225 226
241 242 243 244 245 246
261 262 263 264 265 266

321 322 323 324 325 326
341 342 343 344 345 346
361 362 363 364 365 366

421 422 423 424 425 426
441 442 443 444 445 446
461 462 463 464 465 466

521 522 523 524 525 526
541 542 543 544 545 546
561 562 563 564 565 566

621 622 623 624 625 626
641 642 643 644 645 646
661 662 663 664 665 666

2.4 Simulatie

27 Voetbaluitslagen

Vanaf het seizoen 1956-1957 speelden er achttien clubs in de eredivisie voetbal. Dat geeft 306 wedstrijden per jaar. Alle uitslagen daarvan kun je vinden op www.koningvoetbal.nl

- Ga naar die site en bepaal hoe groot de kans ongeveer is op de uitslag 0 – 0.
- Welk gelijkspel heeft de grootste kans?

28 Gratis toegang tot Chimpie Champ

Bij de entree van de overdekte kinderspeelplaats Chimpie Champ krijg je een kaart. Er zijn zes verschillende kaarten. Als je vier verschillende hebt, mag je een keer gratis naar binnen.

Hoeveel keer kan het duren voordat je vier verschillende kaarten hebt? Minimaal vier keer natuurlijk. Maar om in vier keer alle vier de kaarten ook echt te hebben, moet je veel geluk hebben. Trouwens om in vijf of zes keer vier verschillende te hebben, is ook nog niet zo zeker. Anderzijds heb je wel heel veel pech als je in bijvoorbeeld tien keer nog niet vier verschillende hebt.

Door heel vaak naar Chimpie Champ te gaan en te kijken na hoeveel keer je vier verschillende kaarten hebt, kun je de kansverdeling van het aantal keer achterhalen. Maar dat is niet praktisch. Handiger is om de bezoeken aan Chimpie Champ te **simuleren** (na te bootsen).

- Hoe zou je dat met een dobbelsteen kunnen doen?
Beschrijf precies hoe het gaat.

Een andere vraag is: hoeveel verschillende kaarten heb je na zes bezoeken?

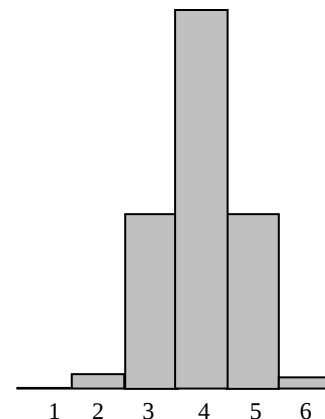
- Wat zijn de mogelijke aantallen?
- Hoe zou je het krijgen van zes kaarten kunnen simuleren om achter de kansverdeling van het aantal verschillende kaarten te komen?

Hiernaast staat een kanshistogram van het aantal verschillende kaarten na zes bezoeken.

- Wat is de kans ongeveer dat je na zes keer nog geen vier verschillende hebt?

 Je kunt het kaarten-krijgen bij Chimpie Champ simuleren met VU-statistiek, Simulaties, Toevalsregen (via de digimap).

- Controleer daarmee of het kanshistogram hierboven ongeveer klopt.



- 29 In de herfst van 2009 had Albert Heijn een reclameactie. Bij elke 15 euro aan boodschappen kreeg de klant een figuur uit het sprookje Sneeuwwitje en de zeven dwergen van de gebroeders Grimm. Er waren er vijftien: Sneeuwwitje, de Prins, de boze Koningin, de Heks, de zeven dwergen: Doc, Sleepy, Happy, Grumpy, Sneezzy, Bashful en Dopey, de Jager en nog drie dieren uit het bos: de Schildpad, de Wasbeer en het Hert.

Het werd een rage om alle vijftien de figuren te bemachtigen, tot aan ruilbeurzen toe. Maar van ruilen zien we nu even af.



We buigen ons over de vraag: Hoe vaak moet je 15 euro aan boodschappen uitgeven om met kans 50% alle vijftien de sprookjesfiguren te bemachtigen?

Die vraag is niet zo maar te beantwoorden. Anne denkt dat de helft van de mensen met 30 keer boodschappen van 15 euro wel alle vijftien de sprookjesfiguren zal hebben verzameld.

a. Denk je dat Anne te optimistisch of te pessimistisch is?



Je kunt dit simuleren met VU-Statistiek, Simulaties, Toevalsregen (via de digimap).

b. Doe dat.



30 Trekkingen lotto 2009

De lotto is een bekend kansspel. Er vallen zes genummerde balletjes en een gekleurd balletje uit een trommel: de zogenaamde trekking.

De speler heeft zes nummers op zijn lottoformulier aangekruist en ook een kleur.

Hoe meer overeenkomst er is met de trekking, des te hoger de prijs.

In Nederland zijn er 45 balletjes: 1 t/m 45, en zes kleuren: geel, groen, blauw, oranje, rood en paars. De nummers hebben allemaal evenveel kans om in een trekking voor te komen en de kleuren hebben ook allemaal evenveel kans.

Elk zaterdag is er een trekking, en de laatste zaterdag van de maand zijn er twee.



a. Ga na dat er 64 trekkingen per jaar zijn.

Hieronder zie je van elke trekking van 2009 hoe vaak elk nummer en elke kleur voorkwamen.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
11	9	9	1	3	15	6	9	10	11	15	5	14	5	17

16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
9	7	11	5	8	7	6	6	11	12	10	6	4	8	8

31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
10	11	5	9	9	6	11	7	9	3	8	12	8	10	8

geel	groen	blauw	oranje	rood	paars
11	12	10	12	9	10

- Hoe vaak zal elk nummer gemiddeld voorkomen in 64 trekkingen?
- Wat is dus de kans dat bijvoorbeeld nummer 23 voorkomt in een trekking?
- Welke nummers waren in 2009 onderbedeeld?

In de tabel hieronder vind je hoeveel van de nummers 0 keer voorkwam, hoeveel 1 keer, hoeveel 2 keer, enzovoort.

aantal keer voorkomen	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
aantal nummers	0	1	0	2	1	4	5	3	6	7	4	6	2	0	1	2	0	1

- Maak een frequentiestafdiagram waarbij het-aantal-keer-voorkomen van een nummer horizontaal wordt uitgezet.
- Bevestigt deze verdeling het antwoord op vraag **b**?

-  **31** Ga naar het computerprogramma VU-statistiek, Simulaties, Steekproeven (via de digimap).
- Hoeveel verschillende aantallen blauwe vierkantjes zijn er mogelijk bij een steekproef van 10.

De steekproef kun je willekeurig vaak herhalen. Daarna krijg je een frequentiestafdiagram te zien van het aantal blauwe vierkantjes in de steekproeven.

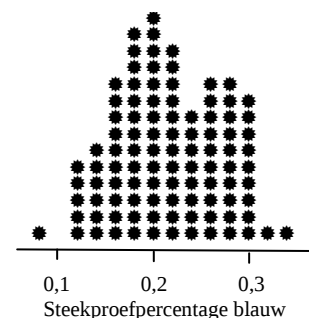
- Experimenteer met het programma.

We kiezen voor de volgende instellingen:

- populatieomvang: 5000
- steekproefomvang: 50
- aantal steekproeven: 100.

Stel dat we nevenstaand resultaat krijgen.

- Wat betekent de geïsoleerde stip links?
- Schat het percentage blauw.
- Hoe groot is de kans ongeveer op 15 of meer blauwe vierkantjes in de steekproef?



Begrepen?

Met VuStat kun je ook het werpen met een munt simuleren. We stellen de kans op kop in op 0,4.

We doen een serie van 50 worpen en tellen in die serie het aantal keer kop.

We maken 40 van deze series, met nevenstaand resultaat:

aantal kop	frequentie
10-11	1
12-13	0
14-15	3
16-17	9
18-19	4
20-21	13
22-23	6
24-25	3
26-27	0
28-29	1

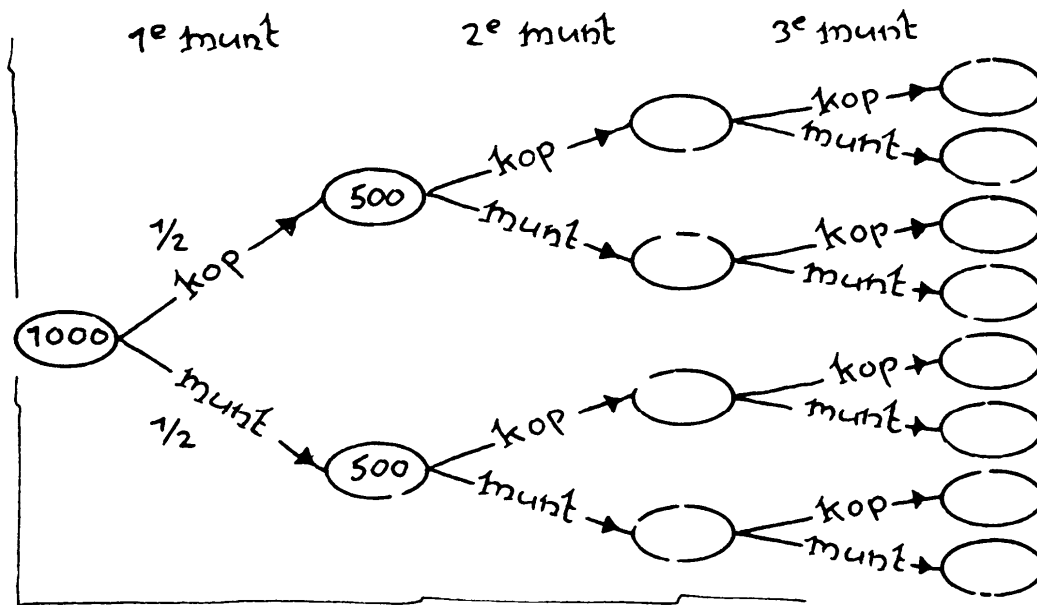
- Hoe groot schat jij op grond van deze simulatie de kans dat er in een serie van 50 meer kop dan munt geworpen wordt?
- Hoe kun je de kans op meer kop dan munt in een serie van 50 worpen met deze munt met grotere zekerheid bepalen?

Anne schat op grond van het resultaat dat de kans op 18 of 19 kop in een serie van 50 worpen met deze munt 10% is.

- Hoe komt Anne aan haar schatting, denk je?
- Geef commentaar op Annes schatting, gezien de hele lijst van frequenties.
- Hoe groot schat jij de kans op 18 of 19 kop in een serie van 50 worpen met deze munt op grond van het resultaat?

2.5 Rekenen met kansen

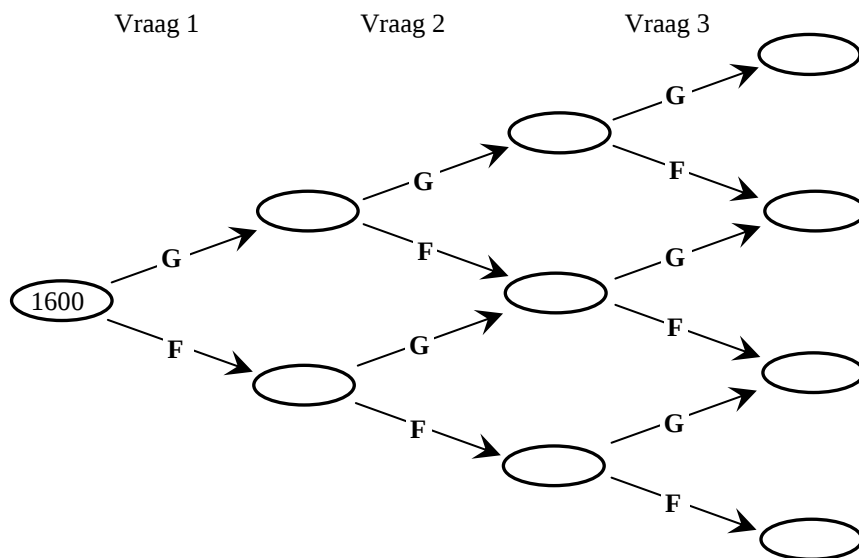
WB 32 Stroomdiagrammen helpen je om kansen te berekenen. We werpen 1000 keer met drie munten en maken daar als volgt een stroomdiagram bij. We letten eerst op de eerste munt: valt die op kop, dan gaan we bij de eerste wegensplitsing naar boven; valt die op munt, dan gaan we naar beneden. Daarna letten we op de tweede munt; valt die op kop dan gaan we bij de tweede wegensplitsingen naar boven; anders gaan we naar beneden. Evenzo voor de derde munt bij de derde wegensplitsingen. Hieronder zie je het stroomdiagram.



In ongeveer de helft van de gevallen zal de eerste munt op kop vallen en ga je in de eerste wegensplitsing naar boven. Je verwacht dat dat 500 van de 1000 keer zal gebeuren. En je verwacht dus dat je 500 keer naar beneden gaat bij de eerste splitsing. Die getallen 500 zijn allebei al in het stroomdiagram geschreven. Ook staan de kansen (fracties) bij de wegen.

- Geef op het werkblad aan hoe de aantallen zich verder over de wegen verdelen en schrijf de kansen bij de overige wegen.
- Kleur in het stroomdiagram de route die hoort bij drie keer kop. Schrijf aan het eind van die route KKK.
- In hoeveel van de 1000 worpen verwacht je 3 kop? Wat is dus de kans op 3 kop?
- Schrijf ook achter de overige routes het resultaat. Met bijvoorbeeld KKM geef je aan dat de eerste twee munten op kop zijn gevallen en de derde op munt.
- In hoeveel van de 1000 worpen verwacht je 2 kop en 1 munt? Wat is dus de kans op 2 kop en 1 munt?
- Wat is de kans op 1 kop en 2 munt?
- Wat is de kans op 3 munt?
- Wat is de som van de hierboven berekende kansen?

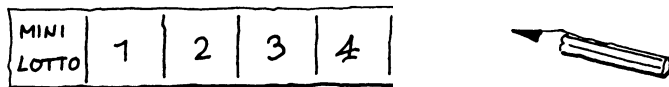
- WB** 33 Een kleine test bestaat uit drie vierkeuzevragen. Je kunt deze vragen op de gok beantwoorden. We zijn geïnteresseerd in het aantal goede antwoorden. We stellen ons voor dat maar liefst 1600 mensen de test op de gok maken. Daarbij hoort het volgende stroomdiagram.



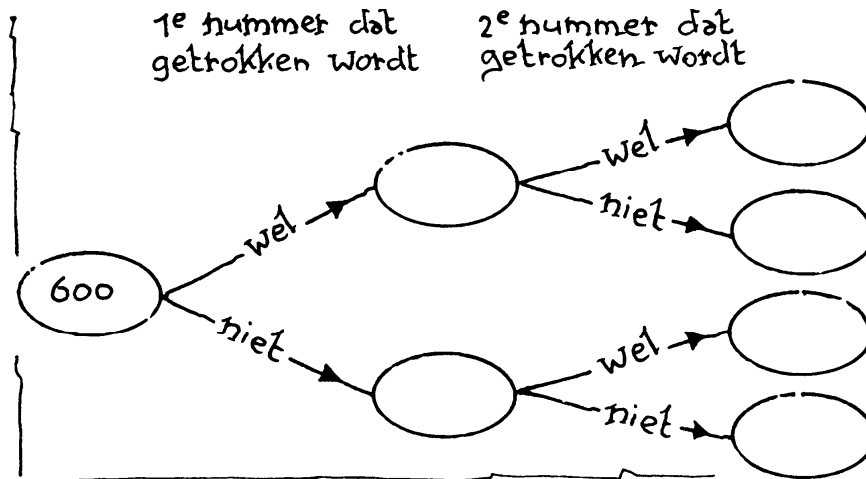
Merk op dat de routes GGF, GFG en FGG samenkomen in hetzelfde eindpunt.

- Vul het stroomdiagram verder in en zet de aantallen goede antwoorden achter de eindpunten.
- Wat is de kans dat een gokker alle drie de vragen goed gokt.
- Wat is de kans dat een gokker precies één vraag fout heeft?
- Wat is de kans dat een gokker precies twee vragen fout heeft?
- Wat is de kans dat je alle drie de vragen fout hebt?
- Hoe kun je de antwoorden op vraag c, d, e en f controleren? Doe dat.
- Waarom zijn we, denk je, met 1600 mensen zijn begonnen? Maakt het wat uit voor de hierboven berekende kansen als je met een ander aantal begint?
- Hoeveel mensen hebben de tweede vraag goed? Wat is dus de kans dat een gokker de tweede vraag goed heeft?

WB 34 De minilotto is een spel waarbij je twee nummers moet omcirkelen op een formulier:

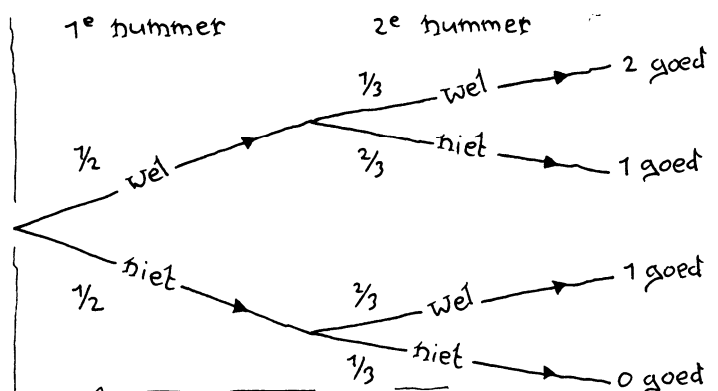


Slechts één combinatie van twee cijfers is winnend. We laten maar liefst 600 mensen dit spel spelen. Hieronder staat een stroomdiagram bij deze situatie.



- Vul het stroomdiagram verder in. Zet de kansen bij de pijlen en schrijf de resultaten achteraan: 0 goed, 1 goed of 2 goed.
- Wat is de kans dat je de winnende combinatie van twee cijfers hebt omcirkeld.
- Bereken ook de kans op 0 cijfers goed en de kans op 1 cijfer goed.
- Je speelt dertig dagen mee in de minilotto, elke dag met één formulier. Hoe vaak verwacht je dan te winnen?

35 Het is niet nodig om bij de minilotto een heel stroomdiagram te tekenen; met behulp van een zogenaamde **kansboom** kun je ook de kansen uitrekenen. Een kansboom krijg je door bij een stroomdiagram de aantallen weg te laten; er staan alleen nog kansen bij de takken. Hiernaast staat een kansboom voor de minilotto.



Kansen op bepaalde uitkomsten bereken je door de kansen langs de takken te vermenigvuldigen: bijvoorbeeld de kans op 2 cijfers goed is $\frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$.

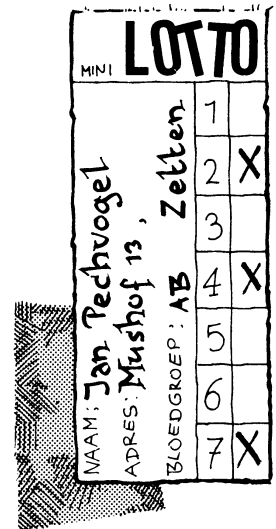
- a. Leg uit dat je de kans dat precies één van jouw aangekruiste nummers wordt getrokken als volgt kunt berekenen: $\frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{2}{3}$.
- b. Bereken ook de kans dat geen van jouw aangekruiste nummers wordt getrokken.

Het doel van deze paragraaf is dat je kansen leert berekenen met behulp van kansbomen. Misschien lukt jou dat nu nog niet. Werk dan voorlopig nog met een stroomdiagram. In dat geval moet je wel zelf nadenken over een geschikt aantal om mee te beginnen.

- 36 Bij een iets grotere minilotto worden drie nummers uit de getallen 1 tot en met 7 getrokken. Op het spelformulier kruist de speler drie nummers uit 1 tot en met 7 aan.
 - a. Maak bij deze minilotto een kansboom. Schrijf de kansen bij de takken. Geef achter alle takken het behaalde resultaat aan: 0 goed, 1 goed, 2 goed of 3 goed.
 - b. Wat is de kans op 3 goed bij deze minilotto?
 - c. Geef ook de kans op 0 goed, de kans op 1 goed en de kans op 2 goed.

Bij de echte lotto worden 6 getallen uit de 45 getrokken. De speler kruist ook 6 getallen op zijn formulier aan.

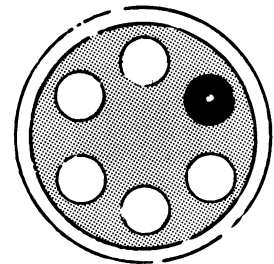
- d. Wat is de kans op 6 goed bij de echte lotto? Je kunt (een deel) van een kansboom tekenen.



- 37 Een pistool heeft een draaibaar magazijn waarin plaats is voor zes kogels. Bij Russisch roulette wordt één van de zes plaatsen met een kogel geladen. De persoon die dit "spel" speelt geeft het magazijn een flinke draai, zet de loop tegen zijn slaap en haalt de trekker over.

Iemand neemt zich voor dit "spel" zes keer te spelen.

- a. Je kunt dit nabootsen met een dobbelsteen. Leg uit hoe.



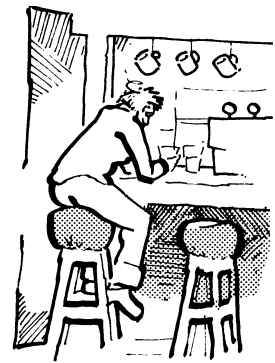
We willen de kans te weten komen om een serie van zes keer Russisch roulette te overleven.

- b. Speel daarvoor twintig series van zes keer met een dobbelsteen. Noteer bij elke serie of je dood bent of niet. Hoe groot schat jij de kans om een serie van zes keer Russisch roulette te overleven?
- c. Vergelijk jouw resultaat met dat van klasgenoten. Hoe groot schat jij nu de kans om een serie van zes keer Russisch roulette te overleven?
- d. Teken een (deel van een) kansboom bij dit "spel" en bereken daarmee de gevraagde kans.

- 38 Stef en Leon gaan een avondje stappen. Wat later op de avond komen ze een café binnen waar aan de bar negen lege barkrukken op een rij staan. Ze zijn wat aangeschoten. Daardoor nemen ze, zonder verder op elkaar te letten, op een kruk plaats. Er is een kans dat Stef en Leon naast elkaar gaan zitten.

We gaan een kansboom tekenen waarmee we deze kans kunnen berekenen. We letten eerst op de kruk waar Stef op gaat zitten.

- Welke twee soorten krukken moet je daarbij onderscheiden?
- Maak de kansboom en bereken daarmee de kans dat Stef en Leon naast elkaar komen te zitten.



- 39 Als Anne van school naar huis fietst, komt zij twee verkeerslichten tegen. Het eerste verkeerslicht staat met kans 0,4 op groen en met kans 0,6 op rood (de tijd dat het licht op oranje staat verwaarlozen we). Als dit verkeerslicht op rood staat, kan Anne in de helft van de gevallen toch doorrijden, omdat er geen verkeer aankomt (en dat doet ze dan ook). Het tweede verkeerslicht staat met kans 0,3 op groen en met kans 0,7 op rood. Als dit licht op rood staat, kan Anne in 40% van de gevallen toch doorrijden.
- Maak een bijbehorende kansboom.
 - Bereken de kans dat Anne bij beide verkeerslichten gaat stoppen.
 - Bereken de kans dat Anne bij beide verkeerslichten gaat doorrijden.

Bij de volgende opgaven staat er niet meer bij of je een kansboom of een stroomdiagram moet maken. Er wordt van jou verwacht dat je hier zelf over nadenkt en vervolgens doet wat jij het beste vindt.

- 40 Het CBR (Centraal Bureau Rijvaardigheidsbewijzen) publiceert jaarlijks de gegevens van het percentage mensen dat hun rijbewijs haalt. Omdat de kans dat je de eerste keer slaagt kleiner is dan de kans dat je de tweede keer slaagt, zijn de gegevens uitgesplitst naar het aantal pogingen.

	Slagingspercentage
1 ^e keer op	48%
2 ^{de} keer op	65%
3 ^{de} keer op	72%
daarna	80%

- Wat is de kans dat een willekeurige Nederlander in twee keer zijn rijbewijs haalt?
- Wat is de kans dat iemand zijn rijbewijs na drie keer rijexamen doen nog steeds niet heeft?

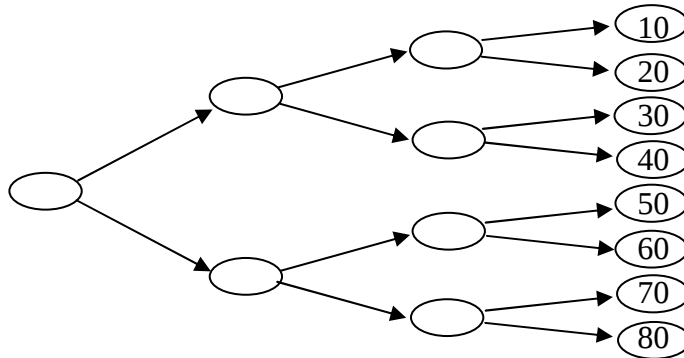
In 2009 nam het CBR 407.773 rijexamens voor rijbewijs B af. Er werden 209.182 eerste examens en 198.591 herexamens afgenomen.

- Hoeveel personen slaagden in 2009 naar verwachting de eerste keer?
- Kun je zeggen hoeveel personen de tweede keer slaagden?

- 41** Jeanine speelt even goed tennis als haar moeder. Een partijtje tegen haar moeder wint ze met kans 0,5. Ze speelt veel beter dan haar vader. Een partijtje tegen haar vader wint ze met kans 0,8. Jeanine krijgt van haar ouders een bromfiets als ze in een serie van drie partijen, afwisselend tegen haar vader en haar moeder, twee maal achter elkaar wint.
- a.** Tegen wie zou jij beginnen als je Jeanine was, tegen haar vader of tegen haar moeder? En waarom?
 - b.** Bepaal de kans dat Jeanine de bromfiets wint als ze tegen haar vader begint.
 - c.** Bepaal ook de kans dat ze de brommer wint als ze tegen haar moeder begint.
- 42** Anne heeft in principe elke woensdagmiddag bijles van de heer Nijdam. Maar Anne is nogal ziekelijk: gemiddeld moet ze 30% van de bijlessen afzeggen. De heer Nijdam is een drukbezet man; hij is gemiddeld 20% van de woensdagen verhinderd.
- a.** Wat is de kans dat de heer Nijdam van drie opeenvolgende woensdagen er twee verhinderd is?
 - b.** Bereken de kans dat de bijles op een willekeurige woensdag niet doorgaat.
- 43** In een volière in een dierenzaak zitten drie exemplaren van een zeldzame vogelsoort: 1 mannetje en 2 vrouwtjes. Het geslacht van zo'n vogeltje kun je op afstand niet zien. Daarvoor moet je het eerst uit de volière halen en van dichtbij bekijken. Willem wil de twee vrouwtjes kopen. De verkoper haalt steeds één vogeltje uit de volière. Blijkt dat het mannetje te zijn, dan stopt hij het weer terug in de volière; blijkt het een vrouwtje te zijn, dan geeft hij het aan Willem. Vervolgens pakt de verkoper opnieuw een vogeltje; net zo lang tot Willem zijn twee vrouwtjes heeft. Elke keer dat de verkoper een vogeltje uit de volière pakt, noemen we een trekking. Bij elke trekking hebben de vogeltjes die nog in de kooi zitten evenveel kans om gepakt te worden.
- a.** Wat is de kans dat bij de tweede trekking een vrouwtje wordt gepakt?
 - b.** Bepaal de kans dat er twee trekkingen nodig zijn om de twee vrouwtjes te pakken.
 - c.** Bepaal de kans dat er drie trekkingen nodig zijn om de twee vrouwtjes te pakken.
- 44** Een bommelding blijkt in 50 van de 100 gevallen loos alarm te zijn. Na de melding gaat de opsporingsdienst aan het werk en wanneer er werkelijk een bom aanwezig is, vindt men deze in 95 van de 100 gevallen.
- a.** Hoe groot is de kans dat er na een willekeurige bommelding geen bom wordt gevonden?
 - b.** Als er geen bom gevonden wordt (dat is in 2100 van de 4000 gevallen), wat is dan de kans dat er toch een bom aanwezig is?

Begrepen?

Hieronder staat het stroomdiagram van een toevalsexperiment dat bestond uit drie stappen. Als in een stap de pijl naar boven gaat spreken van “plus” en anders van “min”. Alleen aan de toppen zijn de (verwachte) aantallen ingevuld.



- Vul de andere aantallen in.
- Maak er een kansboom bij.
- Wat is de kans op 2 minnen en 1 plus?

We bekijken twee experimenten. Het ene experiment heeft mogelijke uitkomsten 1, 3 en 5 met kansen achtereenvolgens $\frac{1}{3}$, $\frac{1}{2}$ en $\frac{1}{6}$, het andere heeft uitkomsten 2 en 4 met kansen achtereenvolgens $\frac{2}{5}$ en $\frac{3}{5}$.

- Maak een kansboom.
- Wat is de kans dat het tweede experiment een lagere uitkomst geeft dan het eerste?

2.6 De som van de kansen is 1

-  45 In een hoed zitten zes kaarten. Op elke kaart staat één cijfer. De cijfers zijn 6, 7, 7, 8, 8 en 8. Uit de hoed worden zonder te kijken twee kaartjes getrokken (dus zonder terugleggen).
- Teken een kansboom hierbij (met acht “toppen”, dat zijn de uiteinden). Vermeld bij elke top het resultaat en de kans op dat resultaat (het resultaat *eerste kaart 7, tweede kaart 8* geef je aan met 7,8).

Als je de kans wilt weten dat er ten minste één kaartje wordt getrokken zonder het cijfer 7, kun je zeven kansen optellen. Maar het kan handiger.

- Hoe?
- Wat is de kans op minstens één kaartje met het cijfer 8? Wat is de kans op geen enkel kaartje met het cijfer 8. Hoe kloppen die kansen met elkaar?

We letten nu op de som S van de cijfers die op de twee getrokken kaartjes staan.

- Welke uitkomsten kan S allemaal hebben?
- Wat is de kans dat $S = 13$? En wat is de kans dat $S = 14$, dat $S = 15$ en dat $S = 16$? Hoe kloppen deze vier kansen met elkaar?

- 46 Anne, Barbara en Cleo zijn vriendinnen. Ze willen graag weten op welke dag van de week ze geboren zijn. Dat kunnen ze uitzoeken met een eeuwigdurende kalender die ze op internet hebben gevonden. Ze vragen zich af hoeveel van hen zondagskinderen zijn.
- Maak een bijbehorende kansboom. (Hoeveel toppen?)
 - Bereken handig de kans dat één of meer van hen een zondagskind is.

Over kansbomen

Bij opgave 46 kun je twee kansbomen maken.

- Door voor elke dag een aparte tak te nemen; en dat voor elk van de drie kinderen. Dan krijg je een kansboom van $7 \cdot 7 \cdot 7 = 343$ toppen. Bij elke tak is de kans $\frac{1}{7}$.
Alle 343 toppen hebben kans $\frac{1}{7} \cdot \frac{1}{7} \cdot \frac{1}{7} = \frac{1}{343}$.
- Door twee mogelijkheden te onderscheiden: zondag en niet-zondag. Dan krijg je een kansboom met $2 \cdot 2 \cdot 2 = 8$ toppen. De helft van de takken heeft kans $\frac{1}{7}$, de andere helft kans $\frac{6}{7}$.
De kansen bij de toppen zijn nu $\frac{1}{343}, \frac{6}{343}, \frac{6}{343}, \frac{36}{343}, \frac{6}{343}, \frac{36}{343}, \frac{36}{343}$ en $\frac{216}{343}$.



- 47 Harrie en Ferdie spelen een tennispartij. Zo'n partij bestaat uit een aantal sets. Winnaar van de partij is degene die het eerst twee sets gewonnen heeft. Harrie en Ferdie spelen even sterk; bij elke set hebben ze beiden kans $\frac{1}{2}$ om te winnen.
- Hoeveel sets worden er minstens gespeeld? En hoogstens?
 - Teken een kansboom bij deze tennispartij. Schrijf bij elk uiteinde de bijbehorende uitkomst (met HFH kun je aangeven dat Harrie de eerste set wint, Ferdie de tweede en Harrie weer de derde). Schrijf er ook de kans bij.
 - Wat is de kans dat de partij 2 sets duurt? En dat de partij 3 sets duurt?
- 48 We gaan verder met het tennissen van Harrie en Ferdie.
Na hun eerste partij besluiten Harrie en Ferdie om nog een partij te spelen, maar nu gaan ze door tot een van hen drie sets gewonnen heeft.
- Zo'n partij wordt ook wel *best of five* genoemd. Kun jij dat uitleggen?
 - Wat is de kans dat Harrie de eerste drie sets wint?
 - Wat is de kans dat de partij maar drie sets duurt?
 - Wat is de kans dat de partij minstens vier sets duurt?

Bij een kansexperiment is de som van de kansen op alle mogelijke uitkomsten gelijk aan 1.

Als bij een kansexperiment bijvoorbeeld de mogelijke uitkomsten 7, 8, 9, 10 en 11 zijn, dan is het volgende vaak handig:

- de kans op minstens waarde 8 = $1 - \text{de kans op waarde 7}$
- de kans op hoogstens waarde 10 = $1 - \text{de kans op waarde 11}$

- 49 Het water in de rivieren staat erg hoog. De dijk om een polder heeft drie zwakke plekken. Deskundigen schatten dat voor elk van de plekken de kans 0,2 is dat de dijk daar doorbreekt. Als de dijk op een van de plekken doorbreekt, komt de hele polder onder water te staan.
- Bereken de kans dat de dijk op geen van de drie zwakke plekken doorbreekt.
 - Bereken de kans dat de polder onder water komt te staan op twee manieren: één keer als de som van zeven kansen en één keer door gebruik te maken van *de som van de kansen is 1*.
- 50 Je gooit net zo lang met een dobbelsteen tot je alle aantallen ogen minstens één keer boven hebt gehad.
- Hoe groot is de kans dat je daar precies 6 worpen voor nodig hebt?
 - Wat is de kans dat je daar minstens 7 worpen voor nodig hebt?
- 51 Een steenfabriek levert trottoirtegels in pakketten van 15 tegels. Een aannemer bestelt 50 pakketten. Omdat er wel eens tegels tussen zitten die veel breuken vertonen, voert de aannemer vooraf een controle uit. Hij kiest 5 van de 50 pakketten en controleert daarin de tegels. Als alle tegels in deze 5 pakketten goed zijn gaat de bestelling door. Neem aan dat bij de bestelling 6 pakketten zitten met één of meer gebroken tegels.
- Bereken de kans dat de bestelling niet doorgaat.



52 Bij een finalewedstrijd voetbal die ook na verlenging gelijk staat, worden strafschoppen genomen. Op grond van de statistieken weten we dat het Nederlands elftal bij het nemen van een strafschop 70% kans heeft om deze te benutten. We bekijken een serie van vijf strafschoppen van het Nederlands elftal.

a. Bereken de kans dat het Nederlands team tijdens deze serie een of meer strafschoppen mist.

Een tegenstander van het Nederlands elftal weet met 90% kans een strafschop te benutten. Beide teams nemen elk vijf strafschoppen.

b. Bereken de kans dat minstens één van beide teams alle strafschoppen weet te benutten.

c. Bereken de kans dat minstens één van de beide teams een of meer strafschoppen mist.

Begrepen?

Iemand doet een serie van tien worpen met een muntstuk. Van elke worp noteert hij de uitkomst: K(op) of M(unt).

Hiernaast staan twee mogelijke series.

KMMKMKKKMK

MKMKMKKKKM

- Bereken de kans dat er twee of meer keer na elkaar dezelfde uitkomst voorkomt.
- Dezelfde vraag als het muntstuk niet zuiver is, maar met twee keer zoveel kans op K valt als op M.

2.7 Voorwaardelijke kansen

53 Vraag 16 van de Nationale Wetensquiz 2003 luidde als volgt:

Met een steekproef testen we de deelnemers aan de tiende Nationale Wetensquiz op een verboden pepmiddel. Stel dat tien procent van de deelnemers het pepmiddel gebruikt. De test is slechts voor negentig procent zuiver.

Een deelnemer blijkt pep-positief.

Hoe groot is de kans dat hij het pepmiddel daadwerkelijk heeft gebruikt?

- a) *Minder dan vijftig procent.*
- b) *Vijftig procent.*
- c) *Meer dan vijftig procent.*

Voor alle duidelijkheid: met de zinsnede "*de test is voor slechts negentig procent zuiver*" wordt bedoeld dat de test in 10% van de gevallen een verkeerde uitslag geeft. Dus 10% van de gebruikers wordt pep-negatief getest en 10% van de niet-gebruikers wordt pep-positief getest.

- a. Teken een stroomdiagram bij deze test uitgaande van 200 deelnemers waarvan 10% gebruikers.
- b. Hoeveel procent van de deelnemers wordt pep-positief getest?
- c. Hoeveel procent van de deelnemers wordt pep-positief getest en is ook gebruiker?
- d. Wat is het goede antwoord op vraag 16 van de Nationale Wetensquiz 2003? Leg uit.

In opgave 53 was er sprake van twee percentages:

- het percentage met een positieve uitslag überhaupt,
- het percentage met een positieve uitslag onder de pep-gebruikers.

Bij het tweede percentage gaat het niet over de hele populatie, maar over een select deel daarvan. Er wordt dan een *voorwaarde* aan de deelnemer gesteld. Zo'n voorwaarde heeft meestal invloed op het percentage.

Met tabellen

Met behulp van gegevens in tabellen kunnen we bepaalde kansen eenvoudig aflezen. Soms staan er aantallen in de tabel, soms ook kansen. In het eerste geval spreken we van een frequentietabel, in het tweede geval van een kanstabel.

Voorbeeld 1. Kleurenblindheid



Uit een onderzoek onder 1000 personen (527 mannen en 473 vrouwen) komen de volgende gegevens over kleurenblindheid en geslacht naar voren. Op de onderste rij en in de rechterkolom staan de randtotalen.

	Man	Vrouw	
Kleurenblind	42	7	49
Niet kleurenblind	485	466	951
	527	473	1000

Met behulp van de tabel kunnen we snel een aantal kansen schatten:

- De kans dat iemand kleurenblindheid is, is $49/1000 = 0,049 = 4,9\%$.
- De kans dat iemand kleurenblind en man is, is $42/1000 = 0,042 = 4,2\%$.

Je kunt ook kijken naar kansen in een deel van de data, bijvoorbeeld alleen naar de mannen of alleen naar de kleurenblinde mensen.

- De kans dat een man kleurenblind is, is $42/527 = 0,080 = 8,0\%$.

- 54 a. Wat is de kans dat een vrouw kleurenblind is?
b. Wat is de kans dat een kleurenblind persoon een vrouw is?

Het is duidelijk dat kleurenblindheid *afhankelijk* is van het geslacht. Bij mannen komt kleurenblindheid immers veel vaker voor dan bij vrouwen.

Mucoviscidose (taaislijmziekte) is de meest voorkomende erfelijke, niet-besmettelijke longziekte. Ongeveer 1400 Nederlanders hebben deze ziekte. Mucoviscidose is onafhankelijk van het geslacht.

- c. Hoeveel Nederlandse vrouwen ongeveer hebben mucoviscidose?

- 55 Het al dan niet sport beoefenen is in Nederland nagenoeg onafhankelijk van het geslacht.

- a. Zeg in eigen woorden wat dat betekent.

Of men een dagelijkse roker is, is wel afhankelijk van het geslacht.

- b. Zeg in eigen woorden wat dat betekent.

-  56 Een autoreparatiebedrijf heeft twee verfspuiters in dienst, Willem en John. Willem spuit 60% van alle auto's. Bij regelmatige controles is gebleken dat 4% van de auto's die door Willem gespoten werden, niet aan de kwaliteitseisen voor spuitwerk voldeed. Bij John was dit 8%.
- Maak een frequentietabel bij bovenstaande gegevens. Ga bijvoorbeeld uit van 1000 auto's.
 - Hoeveel procent van de auto's voldoet niet aan de kwaliteitseisen?
 - Hoeveel procent van de auto's die aan de kwaliteitseisen voldoen, is door Willem gespoten?
 - Hoeveel procent van de auto's die niet aan de kwaliteitseisen voldoen, is door Willem gespoten?

- 57 De meeste slagmannen bij honkbal hebben een voorkeur voor linkshandige of rechtshandige werpers. Zo ook John, die bij een linkshandige werper slechts 15% honkslagen slaat en bij een rechtshandige werper 35%. In de competitie waarin John speelt is 20% van de werpers linkshandig.
- Maak een frequentietabel bij deze gegevens.
 - Wat is de kans dat John een honkslag slaat tegenover een willekeurig gekozen werper?
 - Hoeveel procent van John's honkslagen slaat hij tegenover een rechtshandige werper?



58 De Mantoux-test

Zo nu en dan vindt in ons land een tuberculose-uitbraak plaats. Om mensen in de omgeving die ook besmet zijn op te sporen wordt de zogenaamde Mantoux-test gebruikt. Deze test is rond 1900 ontwikkeld door de Franse arts Charles Mantoux. De test werkt niet feilloos: van degenen met een tuberculosebesmetting reageert 98% positief op de test. Bovendien reageert 1% van degenen die niet besmet zijn ook positief op de test. Daarom worden mensen die positief reageren nogmaals getest met dezelfde test, die een nieuwe uitslag geeft, onafhankelijk van de eerste test. Stel dat er in een stad met 250.000 inwoners 100 gevallen van tuberculosebesmetting zijn (dat is 0,04%).

We maken een tabel voor het geval dat alle inwoners van die stad getest worden.

- Vul de frequentietabel hiernaast in.

	Positief	Negatief	
Wel TBC			
Geen TBC			

- Wat is de kans op tuberculose na een positieve uitslag van de test?
- Wat is de kans op tuberculose na een negatieve uitslag van de test?
- Wat is de kans op tuberculose na een positieve test en een positieve hertest?

Met kansen

Als je de kansen weet, kun je die omzetten in aantallen, door uit te gaan van een geschikt totaal aantal. Welk totaal aantal je kiest doet er niet toe. Je kunt daarom ook met kansen rekenen. Schrijf die kansen in de tabel als (decimale) breuken, niet als percentages. Percentages werken vaak verwarrend. Is de gevraagde kans eenmaal berekend, dan mag deze natuurlijk weer als percentage worden genoteerd.

Voorbeeld 2. Defecte producten

Machine I, machine II en machine III maken alledrie hetzelfde product. Machine I neemt 50% van de productie voor zijn rekening, machine II 30% en machine III 20%. Machine I produceert 1% defecten, machine II 2% en machine III 3%.

Om de gegevens overzichtelijk weer te geven maken we een kanstabel. Voor de drie machines staan hierin de fracties goede en defecte producten. Eerst zetten we rechtsonder de totale kans 1 neer. Vervolgens leiden we uit de gegevens de randtotalen 0,5 , 0,3 en 0,2 af. Vervolgens kunnen we de kansen van de binnenste cellen van de tabel berekenen. Ten slotte kunnen we hiermee de randtotalen van de fracties goede en defecte producten berekenen. Dit alles leidt tot de volgende kanstabel (met kansen in drie decimalen):

	Machine I	Machine II	Machine III	
Goed	0,495	0,294	0,194	0,983
Defect	0,005	0,006	0,006	0,017
	0,500	0,300	0,200	1,000

Zodoende hebben we de volgende kansen bepaald.

Het percentage goede producten = de kans op een goed product $0,983 = 98,3\%$.

De kans dat een defect product door machine II is gemaakt is $0,006/0,017 = 0,353 = 35,3\%$.



Als je het lastig vindt met kansen te werken, kun je ook een frequentietabel maken door van een aantal van bijvoorbeeld 1000 producten uit te gaan.

Notatie

Kansen worden vaak kort en bondig als volgt genoteerd.

- Gewone kansen: “De kans op een goed product” wordt $P(\text{goed product})$.
- Voorwaardelijke kansen: “De kans dat een defect product door machine II is gemaakt” wordt $P(\text{machine II} \mid \text{defect product})$.

Voor de verticale streep staat de eigenschap waarnaar we op zoek zijn: hier de producten die door machine II zijn geproduceerd. Na de verticale streep staat de voorwaarde: hier de defecte producten. We zijn dus geïnteresseerd in het deel van de defecte producten dat door machine II is geproduceerd.

De letter P staat voor ‘Probabilitas’; dat is Latijn voor ‘kans’.

30% van de eredivisiewedstrijden voetbal eindigt in een gelijkspel. Dus $P(\text{gelijkspel}) = 0,3$.

Hoe wordt deze kans beïnvloed in de volgende voorwaardelijke kansen?

$P(\text{gelijkspel} \mid \text{een van de ploegen is Ajax})$,

$P(\text{gelijkspel} \mid \text{ruststand is } 4 - 0)$,

$P(\text{gelijkspel} \mid \text{de spelers ontvangen een winstpremie van 10000 euro})$

Maak bij de volgende opgaven telkens eerst zelf een geschikte (frequentie- of kans-)tabel.

- 59 Van de Nederlandse automobilisten is 58% man en 42% vrouw. Van de mannelijke automobilisten ontvangt jaarlijks 55% één of meerdere bekeuringen wegens te hard rijden. Bij de vrouwelijke automobilisten is dat percentage slechts 10%.
- Hoeveel procent van de Nederlandse automobilisten ontvangt jaarlijks één of meerdere bekeuringen wegens te hard rijden?
 - Hoeveel procent van de bekeurde automobilisten is een man?

- 60** Het vak statistiek wordt op een bepaalde faculteit afgesloten met een tentamen en eventueel een hertentamen. Op basis van resultaten uit de afgelopen jaren is bekend dat 55% van de studenten uiteindelijk (na een eventueel hertentamen) een voldoende haalt voor dit vak. Van de studenten die gedurende de collegeperiode regelmatig opgaven geoefend hebben, haalt 80% uiteindelijk een voldoende voor dit vak. Het percentage studenten dat regelmatig oefent wordt geschat op 35%.

Iemand beweert dat oefenen voor statistiek weinig zin heeft.

- a. Wordt de bewering door deze gegevens ondersteund?
- b. Hoeveel procent van de studenten die uiteindelijk een voldoende hebben voor statistiek, heeft regelmatig geoefend?

- 61** Voor de productie van gevoelige druktoetsen heeft een fabriek twee machines in bedrijf. De machines zijn even goed: de kans op een foute druktoets is bij beide machines 0,65%. Machine I neemt 60% van de productie voor zijn rekening, machine II 40%.

- a. Maak een kanstabel:

	Goed	Fout	
Machine I			
Machine II			

- b. Is het goed-zijn van een druktoets onafhankelijk van de machine? Dat wil zeggen: is de kans dat een druktoets goed is op beide machines hetzelfde?

We bekijken dit algemener. Veronderstel dat het goed-zijn van een druktoets onafhankelijk is van de machine. Veronderstel dat de kansen door welke machine een druktoets gemaakt is en of een druktoets goed of fout is bekend zijn.

- c. Hoe kun je de kanstabel dan verder maken?

- 62** Iemand werpt twee keer met een dobbelsteen.

- a. Is het aantal ogen bij de tweede worp onafhankelijk van het aantal ogen bij de eerste worp?

Iemand trekt twee kaarten uit een stapeltje van vier kaarten: twee boeren en twee azen.

- b. Is het resultaat (boer of aas) bij de tweede kaart onafhankelijk van het resultaat bij de eerste kaart? Licht je antwoord toe.

Begrepen?

In een land leven twee bevolkingsgroepen, de Langen en de Korten. In beide groepen komt een ziekte voor. Hieronder staan de kansen op het hebben van de ziekte, apart voor de Langen en de Korten

Zo is $P(\text{Lang en Niet}) = x$.

	Wel	Niet
Lang	0,1	x
Kort	0,2	y

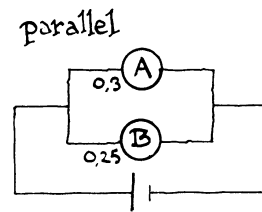
- Wat is het verband tussen x en y ?
- Druk $P(\text{Wel} \mid \text{Lang})$ uit in x ?
- Wat weet je van x en y als het hebben van de ziekte onafhankelijk is van de bevolkingsgroep?
- Wat weet je van x en y als een Lange meer risico heeft de ziekte te krijgen dan een Korte?

2.8 Extra opgaven

- 63 Een zekere elektrische stroomkring is opgebouwd uit twee elementen: A en B. De kans dat element A (in een bepaalde periode) uitvalt is 0,3. De kans dat element B uitvalt is 0,25.

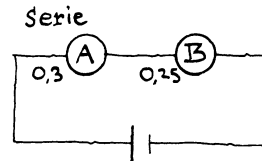
We bekijken eerst het geval dat A en B parallel geschakeld zijn.

- Teken een kansboom voor het al of niet uitvallen van de stroom in de kring.
- Bereken de kans dat de stroom uitvalt.
- Bereken de kans dat de stroom niet uitvalt.



Neem nu het geval dat A en B in serie geschakeld zijn.

- Teken een kansboom voor het al of niet uitvallen van de stroom in de kring.
- Bereken de kans dat de stroom niet uitvalt.
- Bereken de kans dat de stroom wel uitvalt.



- 64 Op de lijn Arnhem-Utrecht is de kans 20% dat je in de trein gecontroleerd wordt op een geldig vervoersbewijs (per enkele reis). Door de week reist Jaap 's ochtends van Arnhem naar Utrecht en 's middags weer terug naar Arnhem. Jaap besluit op een maandag om geen kaartje te kopen voor de heen- en terugreis.
- Hoe groot is de kans dat Jaap die dag tegen de lamp loopt?

Jaap werd die maandag niet gecontroleerd. Hij besluit daarom ook de rest van de week zonder kaartje te reizen.

- Hoe groot is de kans dat hij gedurende die week tegen de lamp loopt?

- 65 Een klas telt 25 leerlingen. Ze vormen een hechte groep en vieren elkaars verjaardag uitbundig. Bereken de kans dat er twee (of meer) leerlingen op dezelfde dag jarig zijn. (Neem maar aan dat het bekend is dat geen van de leerlingen op 29 februari geboren is. Er zijn dan voor elke leerling nog 365 dagen over om te zijn geboren.) Schrik niet terug voor flink wat rekenwerk op de GR.

- 66 Een bedrijf neemt regelmatig nieuw personeel aan. Na verloop van tijd blijkt slechts 60% hiervan geschikt voor het werk, dus 40% is ongeschikt. Om beter te kunnen selecteren, worden kandidaten voortaan psychologisch getest. Van de geschikte kandidaten komt 90% door de test, terwijl van de ongeschikte kandidaten 20% door de test komt.
- Maak een kanstabel bij bovenstaande gegevens. Kies zelf een geschikt aantal kandidaten.
 - Bereken de kans dat een willekeurige kandidaat door de test komt.
 - Bereken de kans dat een kandidaat die afgewezen wordt, toch geschikt is.

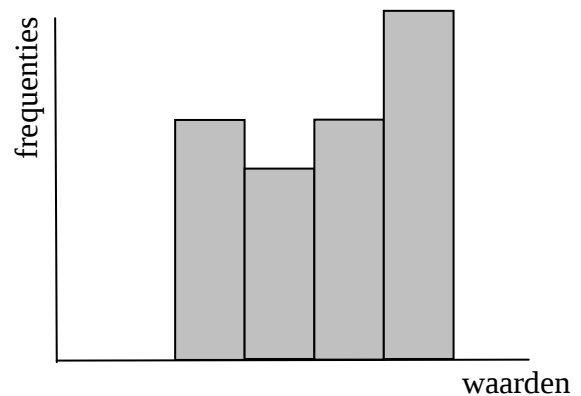
- 67** De faculteit VPL kent drie studierichtingen: Verkeerskunde (V), Planologie (P) en Logistiek (L). Van de mannelijke studenten is 29,1% bij Verkeerskunde ingeschreven, 27,2% bij Planologie en 43,7% bij Logistiek. Van de vrouwelijke studenten is 34,4% bij Verkeerskunde ingeschreven, 38,5% bij Planologie en 27,1% bij Logistiek. Van de totale studentenpopulatie is 77,6% man.
- a.** Bepaal voor elke studierichting het percentage van de totale studentenpopulatie dat deze richting volgt.
 - b.** Bepaal bij elk van de drie studierichtingen de man/vrouw-verdeling in procenten.
- 68** 20% van de veroorzakers van verkeersongevallen heeft alcohol gebruikt. Een kwart van de weggebruikers die met drank op een ongeval veroorzaken is jonger dan 25 jaar oud. Van de weggebruikers die een ongeval veroorzaken zonder dat van drankgebruik sprake is, is 35% jonger dan 25 jaar oud. De 18-24 jarige weggebruikers noemen we jongeren en de overige weggebruikers noemen we ouderen.
- a.** Hoeveel procent van de verkeersongevallen wordt veroorzaakt door ouderen?
- Iemand beweert dat met name bij jongeren het alcoholgebruik leidt tot veel ongelukken.
- b.** Wordt de bewering door deze gegevens ondersteund?

2.9 Samenvatting Verdelingen

Paragraaf 2.1 Frequentieverdelingen

In een groep letten we op een zekere eigenschap (variabele). Die eigenschap kan bijvoorbeeld vier waarden hebben. De groep is *verdeeld* over de vier waarden: elke waarde komt een zeker aantal keren voor. We spreken van een **frequentieverdeling**.

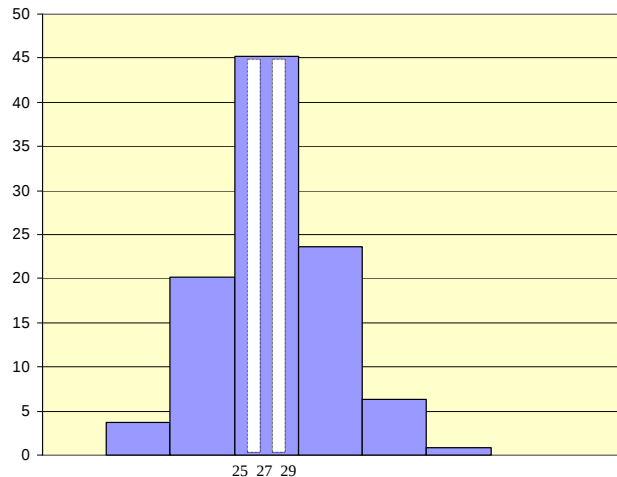
Als de verdeling in procenten van de totale groep is, spreken we van een **relatieve** frequentieverdeling. De som van de relatieve frequenties is 100%.



Bij een frequentieverdeling kunnen we een **staafdiagram** of een **frequentiehistogram** maken.

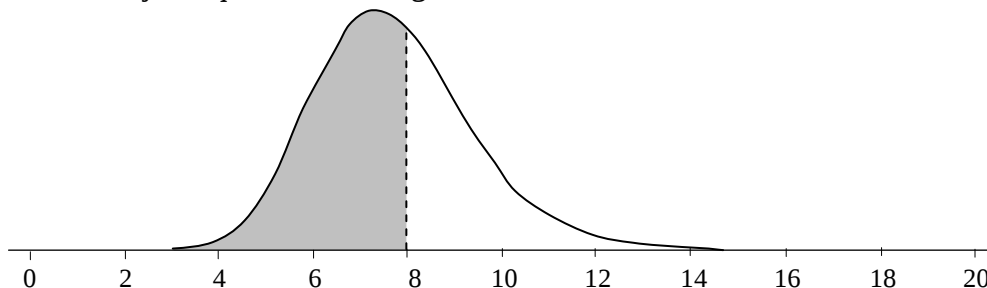
De hoogte van een staaf geeft het aantal of het percentage in de bijbehorende klasse. In plaats van de hoogte kun je ook op de *oppervlakte* letten.

Voorbeeld: de totale oppervlakte van het histogram links van 27, inclusief 27 zelf, geeft het percentage van de gevallen dat de waarde 27 of lager is.



Als we in een frequentiehistogram de middens van de opvolgende balken verbinden, ontstaat een **frequentiepolygoon**. Als in een frequentiehistogram de balken heel smal zijn, is de frequentiepolygoon goed te benaderen met een continue gladde kromme.

Een zekere variabele (bijvoorbeeld gewicht, tijdsduur, bedrag) neemt waarden aan tussen 1 en 8. Hieronder staat zijn frequentieverdeling:



Hoe vaak de grootheid een waarde kleiner dan 8 heeft, wordt gegeven door de oppervlakte van het grijze gebied.

Preciezer: het percentage van het gebied onder de kromme dat grijs is, is gelijk aan het percentage van de gevallen dat de variabele kleiner dan 8 is.

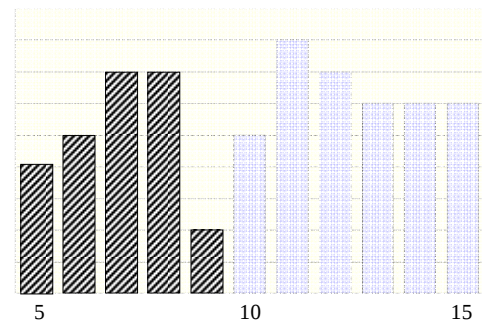
Paragraaf 2.2 Kans

Als in 3,8% van de gevallen een variabele een waarde kleiner dan 20 heeft, zeggen we: de **kans** dat de variabele kleiner dan 20 is, is 3,8%.

Van een discrete variabele staat hiernaast een relatief frequentie-staafdiagram.

De kans op een waarde kleiner dan 10 is de som van de percentages van de waarden 5 t/m 9.

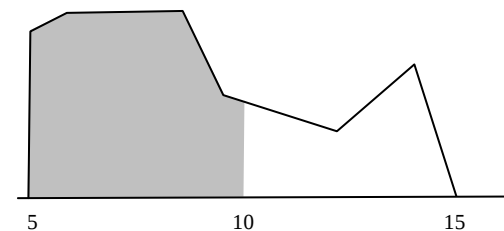
Dat is $\frac{\text{oppervlakte van de staven bij 5, 6, 7, 8 en 9}}{\text{oppervlakte van alle staven}}$.



Van een continue variabele staat hiernaast een relatieve frequentiepolygoon.

De kans op een waarde kleiner dan 10 is dan hoeveel procent het deel tussen 5 en 10 is van het geheel.

Dat is $\frac{\text{oppervlakte links van 10}}{\text{totale oppervlakte}}$.



Een kans laat zich uitdrukken als een percentage (tussen 0% en 100%) en ook als een breuk (tussen 0 en 1).

Paragraaf 2.3 Op den duur ...

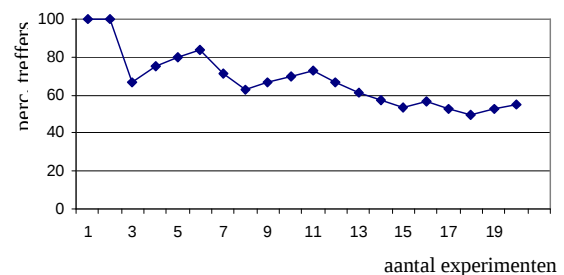
De wet van de grote aantallen

Veronderstel dat bij een zeker experiment de kans op een treffer 0,4 is.

Als het experiment vaak herhaald wordt, nadert het gemiddelde aantal treffers altijd tot 0,4.

We voeren het experiment bij herhaling uit en volgen het verloop van het aantal treffers.

Dan zal op den duur het percentage treffers zeer dicht in de buurt komen van de kans op een treffer.



Theoretische kansen

De kans op vijf ogen bij een worp met een dobbelsteen is $\frac{1}{6}$.

De kans op kop bij een worp met een muntstuk is $\frac{1}{2}$.

Een randomgenerator produceert de cijfers 0 t/m 9. De kans op het cijfer 5 is $\frac{1}{10}$.

Bij een toevalsexperiment zijn er verschillende uitkomsten mogelijk. De totale kans van 100% is *verdeeld* over de verschillende uitkomsten. We spreken van een **kansverdeling**.

Veronderstel dat er zeven uitkomsten mogelijk zijn. Als die gelijkwaardig zijn (gemiddeld even vaak voorkomen) zeggen we dat elk van die uitkomsten kans $\frac{1}{7}$ heeft.

Paragraaf 2.4 Simulaties

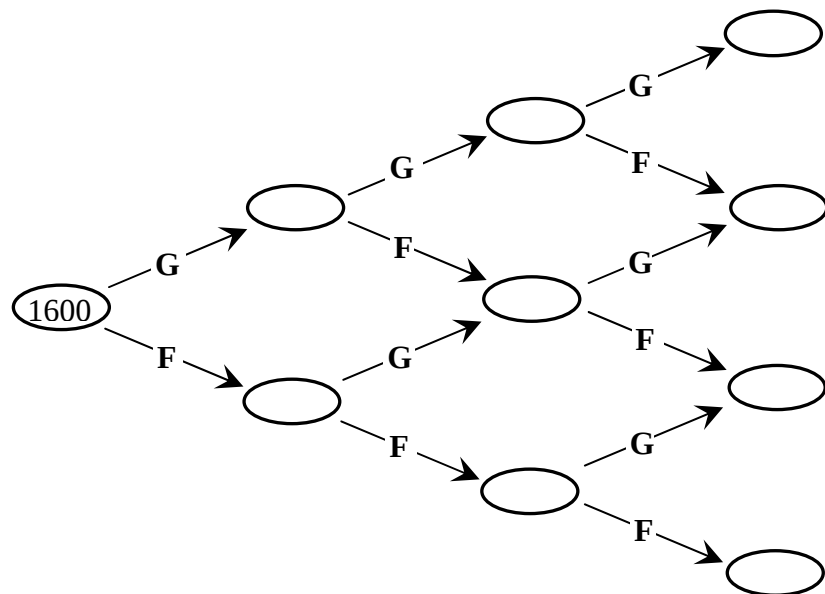
Soms is het ondoenlijk een experiment in werkelijkheid vaak uit te voeren. Om dan toch zicht te krijgen op de kansen maken we gebruik van *simulatie*. Dat wil zeggen dat we het experiment nabootsen. Hierbij kan de computer worden ingeschakeld of de randomgenerator op de GR.

Paragraaf 2.5 Rekenen met kansen

Een experiment wordt opgesplitst in deelexperimenten. We voeren het experiment in gedachten een groot aantal keren uit, zeg 1600 keer.

In een stroomdiagram splitst dat aantal zich na elk deelexperiment. Bij elke splitsing vraag je je af: "hoe vaak mag ik verwachten dat de ene uitslag voorkomt, en hoe vaak de andere".

Aan de toppen (de uiteinden) van het diagram kun je tellen hoe vaak je mag verwachten dat een speciale uitkomst optreedt.

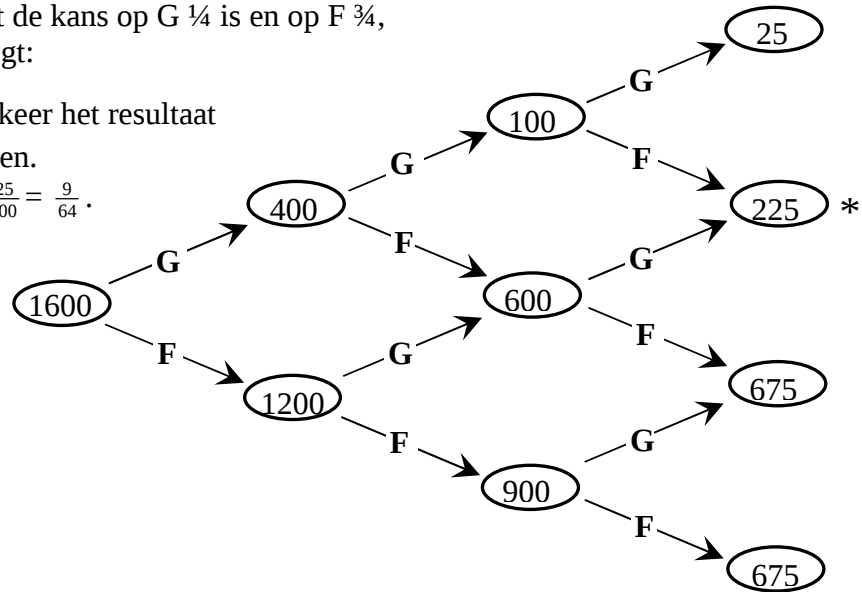


Als bij elk deexperiment de kans op G $\frac{1}{4}$ is en op F $\frac{3}{4}$,
dan gaat de stroom als volgt:

Bij * lees je af dat je 225 keer het resultaat

"2G en 1F" mag verwachten.

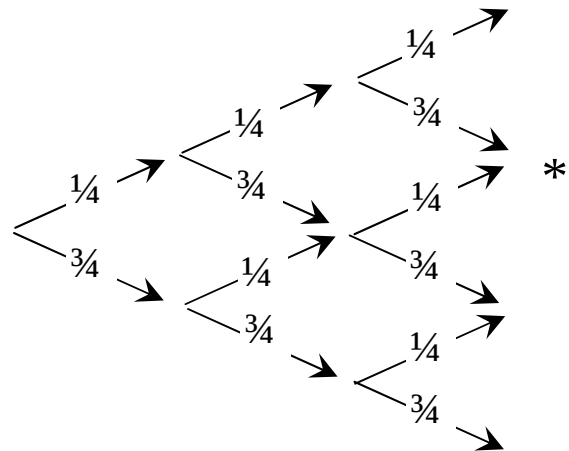
De kans op 2G en 1F is $\frac{225}{1600} = \frac{9}{64}$.



Je hoeft niet een (geschikt) aantal experimenten
te kiezen om in het stroomdiagram te beginnen.
Je kunt ook rekenen met de fracties (kansen)
langs de wegen van het diagram. In het
voorbeeld wordt dat zoals hiernaast.

De kans op het resultaat * laat zich als volgt
berekenen:

$$\frac{1}{4} \cdot \frac{1}{4} \cdot \frac{3}{4} + \frac{1}{4} \cdot \frac{3}{4} \cdot \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} = \frac{9}{64}.$$



Paragraaf 2.6 De som van de kansen is 1

Bij een kansexperiment is de som van de kansen op alle mogelijke uitkomsten gelijk aan 1.

Als bij een kansexperiment bijvoorbeeld de mogelijke uitkomsten 7, 8, 9, 10 en 11 zijn, dan is het volgende vaak handig:

- de kans op minstens waarde 8 = 1 – de kans op waarde 7
- de kans op hoogstens waarde 10 = 1 – de kans op waarde 11

Paragraaf 2.7 Voorwaardelijke kansen

In een groep van 20 jongens en 10 meisjes komt een eigenschap voor bij 12 jongens en 4 meisjes.

	eig.	niet	totaal
jongen	12	8	20
meisje	4	6	10
totaal	16	14	30

Dan

- heeft $12/20 \times 100\% = 60\%$ van de jongens die eigenschap,
- is $12/16 \times 100\% = 75\%$ van de mensen met die eigenschap jongen.

Met kansen:

- de kans dat een jongen die eigenschap heeft is $P(\text{eigenschap} \mid \text{jongen}) = 0,6$,
- de kans dat de drager van de eigenschap een jongen is is $P(\text{jongen} \mid \text{eigenschap}) = 0,75$.

Men spreekt hier van **voorwaardelijke kansen**; de voorwaarden zijn “jongen” respectievelijk “eigenschap”.