Wachttijden

Nieuwe wiskunde tweede fase
Profiel N&G en N&T
Freudenthal instituut

Bij de foto op de voorkaft:
Lange wachtrij van Zuidafrikaanse kiezers op de historische dag van de vrije verkiezingen (26 april 1995).

Wachttijden

Project:	Wiskunde voor de tweede fase
Profiel:	N&G en N&T
Domein:	Voortgezette Kansrekening
Klas:	VWO 6
Staat:	Eerste experimentele versie
Ontwerp:	Elisja Giepmans, Jan Smit, Martin Kindt

Hoofdstuk 1
Er zijn nog vijf wachtenden voor u

1: in de wachtrij

Je belt 06-8008 (Inlichtingen telefoonnummers binnenland) omdat je het telefoonnummer van je vriendin vergeten hebt. Als je de naam en het adres geeft, dan krijg je van inlichtingen het bijbehorende telefoonnummer. Soms word je meteen geholpen, maar dikwijls krijg je te horen: *er zijn nog ... wachtenden voor u*. Dat aantal wordt dan geleidelijk kleiner tot je eindelijk aan de beurt bent.
Zulke wachttoestanden kunnen zich voordoen bij alle instellingen die diensten verlenen, bestellingen opnemen, inlichtingen verstrekken, kaartjes verkopen, enzovoort. Het postkantoor, het station, de bakker, de kapper, de benzinepomp, maar ook in communicatienetwerken, de telefooncentrale, internet.

instroom en bedieningstijd

In zo'n bedieningssysteem heb je te maken met een 'stroom' van binnenkomende klanten. Deze stroom is niet regelmatig. Nu eens komen er een aantal vlak na elkaar, dan is het weer eens een tijdje rustig. Ook de tijd die nodig is om een klant te helpen is niet steeds even lang. De een komt alleen maar een kaartje kopen, de ander wil een buitenlandse reis bespreken. Zowel bij de instroom als bij de bedieningstijden kunnen allerlei toevallige afwijkingen optreden.

wachttijd-theorie

De *wachttijd-theorie* (Queueing theory) heeft als doel zulke bedieningssystemen te onderzoeken. Er wordt een wiskundig model opgezet voor zo'n systeem. In dat model wordt beschreven welke toevallige afwijkingen, met welke kansen er kunnen optreden in de instroom van de klanten en hun bedieningstijden. Kansrekening is de belangrijkste wiskunde die hierbij nodig is.
Als we het model goed begrijpen, kunnen we berekenen hoe lang een klant gemiddeld moet wachten en hoe lang de rij wachtenden gemiddeld zal zijn. Het doel van zulke berekeningen is om te komen tot voorspellingen en beslissingen (moet er een loket bij komen?, moet de verkeerssituatie worden verbeterd?, moet de service worden versneld? , enz.) Ook zonder moeilijke berekeningen kunnen we dikwijls een aardig beeld krijgen van de werking van het systeem door het na te spelen (*simulatie*). Het volgende sterk vereenvoudigde voorbeeld geeft hiervan een idee.

2: wachten met een dobbelsteen

In dit voorbeeld bekijken we een situatie van de 06-8008 inlichtingendienst. Stel voor het gemak dat voor iedere klant de bedieningstijd even lang is. *Eén minuut* laten we zeggen is de tijd nodig om een telefoonnummer bij een gegeven naam en adres te zoeken en iedere klant vraagt maar om één persoon die de inlichtingen verstrekt. In de tijd dat een klant geholpen wordt komen er 0, 1, 2, 3, ... nieuwe aanvragers die in de wachtrij gezet worden met de mededeling: ‘*Er zijn nog X wachtenden vóór u.*’
Stel dat de kansen op 0, 1, 2, 3, ... nieuwe opbellers in één minuut achtereenvolgens $a, b, c, d, \ldots$ zijn.
Om het eenvoudig te houden, nemen we aan dat alleen de eerste drie kansen positief zijn. Dat komt neer op: $a + b + c = 1$ ($d = \ldots = 0$).
Als we voor de kansen a, b en c getallen invullen, ligt het model vast. Een paar mogelijke keuzen voor het drietal a, b, c zijn bijvoorbeeld:

Geval 1 fi $a : b : c = 3 : 1 : 2$, dus: $a = \frac{1}{2}$; $b = \frac{1}{6}$; $c = \frac{1}{3}$.

Geval 2 fi $a : b : c = 2 : 1 : 3$, dus: $a = \frac{1}{3}$; $b = \frac{1}{6}$; $c = \frac{1}{2}$.

Geval 3 fi $a : b : c = 1 : 1 : 1$, dus: $a = b = c = \frac{1}{3}$.

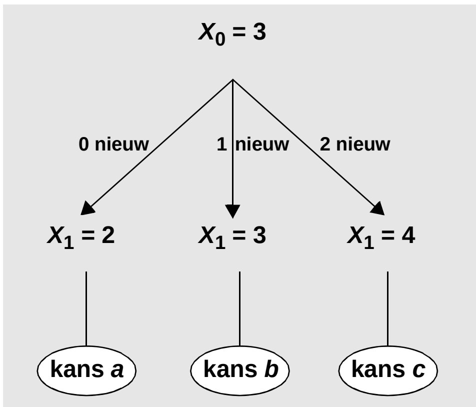
1 In welke van de drie gevallen denk je dat de wachttijd het langst zal zijn?

simuleren

Je kunt met dit model gaan experimenteren. Door een dobbelsteen te gooien, kunnen we steeds bepalen hoeveel nieuwe opbellers er in de volgende minuut bijkomen.
Bijvoorbeeld voor geval 1:
bij 1, 2 of 3 ogen, geen nieuwe aanvragen;
bij 4 ogen één aanvrager;
bij 5 of 6 ogen twee nieuwe aanvragers.

Voordat we zo’n experiment gaan uitvoeren, analyseren we eerst wat er kan gebeuren op de tijdstippen 0, 1, 2, ... (steeds één minuut na elkaar).
Noem die aantallen wachtenden op die tijdstippen X_0, X_1, X_2, ...
Of de rij wachtenden langer of korter wordt, hangt af van het aantal nieuwe opbellers.

2 Stel bijvoorbeeld dat $X_0 = 3$.
Dat wil zeggen dat de lengte van de rij op het moment dat we onze waarnemingen beginnen gelijk is aan 3. Dat wil zeggen: 1 aanvrager in behandeling, 2 aanvragers in de wachttoestand.
a. Bekijk het schema hiernaast en verklaar het.
b. Noem het aantal opbellers tijdens de eerste minuut Y_1.
Verklaar de formule: $X_1 = X_0 - 1 + Y_1$
c. Waarom klopt dit niet als $X_0 = 0$?



3 Verklaar onderstaande stap formules voor de verandering in de n^{de} minuut (dus van tijdstip $n - 1$ naar tijdstip n).

$$X_n = X_{n-1} - 1 + Y_n \qquad \text{als } X_{n-1} > 0$$
$$X_n = Y_n \qquad \text{als } X_{n-1} = 0$$

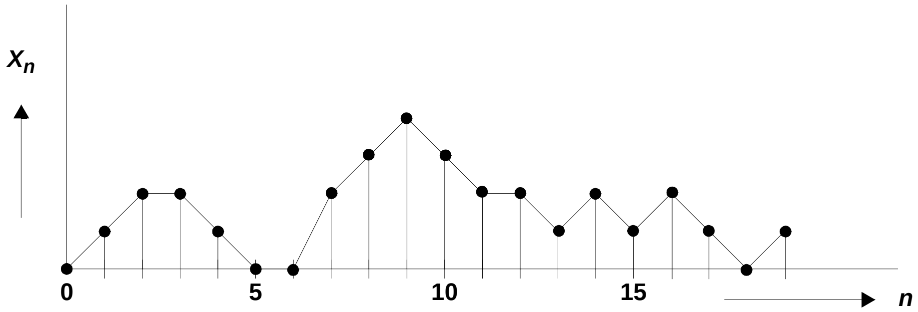
waarbij Y_n het aantal *aanvragen* in de n^{de} minuut voorstelt.

a **:** ***b*** **:** ***c*** **= 3 : 1 : 2**

Nu gaan we het proces simuleren met een dobbelsteen. In de tabel hieronder zie je het resultaat van zo’n simulatie. De bovenste regel geeft de tijd (in minuten). De tweede regel geeft de aantallen ogen geworpen met de dobbelsteen. De derde regel het bijbehorende aantal aanvragers, De vierde regel geeft X_n, het aantal personen in het systeem. De beginsituatie is: $X_0 = 0$

n	1				5					10					15					20
worp	4	5	4	1	3	2	6	6	5	1	1	4	3	6	2	6	3	2	4	
aanvragers	1	2	1	0	0	0	2	2	2	0	0	1	0	2	0	2	0	0	1	
X_n	1	2	2	1	0	0	2	3	4	3	2	2	1	2	1	2	1	0	1	

Dit kunnen we weergeven in een stippen-grafiek.



Merk op: de stippen geven de uitkomsten aan; de verbindingslijntjes horen niet bij de grafiek, maar zijn alleen getrokken om het verloop beter zichtbaar te maken.

4 Gooi ook een paar keer met de dobbelsteen en kijk hoe de rij zich verder ontwikkelt.

5 Neem nu het geval 2 **(**dat wil zeggen: 1 of 2 ogen fi 0 klanten; 3 ogen fi 1 klant; 4, 5, of 6 ogen fi 2 klanten).
Welk verloop zal de rij worpen van hierboven geven? Vul de tabel in en maak een grafiek van X_n afhankelijk van n.

n	1				5					10					15					20
worp	4	5	4	1	3	2	6	6	5	1	1	4	3	6	2	6	3	2	4	
aanvragers																				
X_n																				

6 Doe hetzelfde voor geval 3.

n	1 5	10	15	20
worp	4 5 4 1 3	2 6 6 5 1	1 4 3 6 2	6 3 2 4
aanvragers				
X_n				

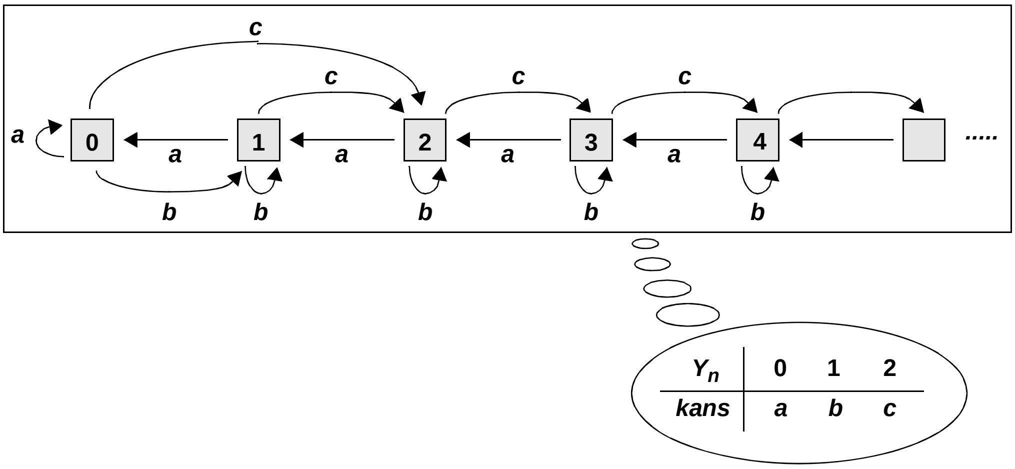
7 Wat is het verschil tussen de gevallen 1, 2 en 3?

8 Neem eens een langere periode, bijvoorbeeld 600 minuten.
a. Hoeveel aanvragen kun je ongeveer verwachten in 600 minuten in geval 1 ?
En hoeveel in de gevallen 2 en 3? (Ga er van uit dat de verschillende kanten van de dobbelsteen ongeveer even vaak zullen boven komen.)
Hoeveel klanten kunnen er in elk van de gevallen in 600 minuten maximaal worden geholpen ?
b. Bij geval 1 zien we dat de rij af en toe lengte nul heeft: $X_n = 0$ bij $n = 5, 6, 18$.
Is dit iets toevalligs of kun je voorspellen dat dit steeds weer zal gebeuren?
En hoe is dat bij de andere gevallen?

We bekeken de toestand van het systeem in ons voorbeeld op de tijdstippen 0, 1, 2, ... (steeds één minuut later). Dit noemt men een model met *discrete tijd*. Daarmee wordt aangegeven dat we net doen of de tijd met *stappen* voortschrijdt en niet continu. De toestand wordt bepaald door het aantal klanten in het systeem.
Bijvoorbeeld, toestand 3 wil zeggen: 3 klanten in het systeem : 2 in de wachttoestand en 1 die geholpen wordt.

toestanden schema

Dit wordt weergegeven in het volgende schema:

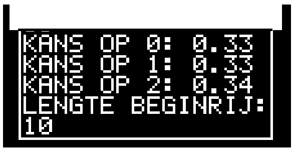


De pijlen geven de mogelijke overgangen met de bijbehorende ‘overgangskansen’.
Als het systeem bijvoorbeeld op tijd n in toestand 3 is ($X_n = 3$), dan kunnen we de volgende minuut overspringen naar toestand 2 of naar toestand 4 of in toestand 3 blijven, al naar gelang er 0, 2 of 1 klanten aankomen; dat gebeurt respectievelijk met kansen a, c en b.

computer-simulatie

Dat gedoe met dobbelstenen kost veel tijd. Met de computer kan het veel sneller. Met zijn ‘*random generator*’ (random = toeval) kan de computer getallen produceren die volkomen systeemloos (zo lijkt het tenminste) in het interval van 0 tot 1 vallen. Met behulp van die getallen kunnen we het werpen van een dobbelsteen nabootsen en zodoende in enkele seconden het verloop van een wachtrij simuleren. De plaatjes hieronder brengen dit in beeld voor de gevallen 1, 2 en 3.
Het is te begrijpen dat in geval 2 de wachtrij steeds langer wordt omdat de instroom groter is dan de verwerkingscapaciteit. In het toestandenschema zien we dat ook: de kans (c) om naar rechts te gaan is groter dan de kans (a) om naar links te gaan. Om een goede dienstverlening mogelijk te maken moet het systeem aangepast worden. Bijvoorbeeld door meer helpers in te schakelen.

Op de grafische rekenmachine is het mogelijk om het verloop van de wachtrijen te simuleren met het programmaatje LOKET (zie ommezijde).
Als je LOKET start kun je kansen opgeven voor 0, 1, 2 of 3 klanten die binnen komen (het programma vraagt alleen om de laatste kansen als de som van de kansen die reeds zijn opgegeven nog geen 1 is); bovendien moet je de lengte van de beginrij invoeren..
Hieronder zie je van drie verschillende situaties de uitvoer van het programma. Ieder vertikaal streepje is een wachtrij, horizontaal loopt de tijd.



9 Met het programma kun je bekijken welke gevolgen de verschillende kansen op de lengte van de wachtrij hebben.
Onderzoek onder welke voorwaarden de lengte van de wachtrij niet uit de hand loopt.

realistischer modellen

Het model dat we als voorbeeld genomen hebben is te simpel voor de meeste situaties in de realiteit.
Om betere modellen te ontwikkelen hebben we wat meer kansrekening nodig dus daar gaan we ons eerst mee bezighouden. We moeten wat meer weten over kansverdelingen en gemiddelden (verwachtingswaarden). We willen ook van het model met discrete tijd naar modellen met continue tijd waarbij de tijdstippen waarop klanten binnen komen op willekeurige plaatsen op de tijd-as kunnen vallen. Dit leidt tot de beroemde Poisson verdeling en het Poisson proces. Ook de bedieningstijd hoeft niet vast te zijn. Een veel gebruikte verdeling bij dit alles is de exponentiële verdeling. Na dat we in hoofdstuk 1, 2 en 3 vertrouwd zijn geraakt met de nodige begrippen uit de kansrekening en gezien hebben hoe we toevalsverschijnselen kunnen simuleren komen we in het laatste hoofdstuk weer terug naar wachttijdproblemen.

LOKET:

```
PROGRAM:LOKET
:0→A
:1→B
:1→C
:1→D
:Input "KANS OP
0: ",A
:Input "KANS OP
1: ",B
:If A+B<1
:Input "KANS OP
2: ",C
:If A+B+C<1
:Input "KANS OP
3: ",D
:Input "LENGTE B
EGINRIJ:",N
:0→Xmin
:94→Xmax
:-7→Ymin
:55→Ymax
:AxesOff
:ClrDraw
:FnOff
:PlotsOff
:Line(0,0,94,0)
:Line(0,0,0,N)
:For(X,0,94)
:rand→P
:0→Y
:If P>A
:Y+1→Y
:If P>(A+B)
:Y+1→Y
:If P>(A+B+C)
:Y+1→Y
:If P>(A+B+C+D)
:Y+1→Y
:N-1→L1(X+1)
:If N=0
:0→L1(X+1)
:Y→L2(X+1)
:N+Y-1→N
:If N<0
:0→N
:Line(X,0,X,N)
:End
```

Hoofdstuk 2
Kans en verwachting

1: Inleiding

toeval

Toeval speelt de hoofdrol in de processen die we willen bestuderen. Dikwijls is toeval een storende factor. Soms is het echter een zeer gewaardeerde gast. Dit laatste is het geval bij de kansspelen waar het toeval voor de nodige verrassingen zorgt. Geen wonder dus dat de wiskundige bestudering van het toeval, de kansrekening begonnen is met kansspelen. Al in de zestiende eeuw schreef de Italiaan Cardano (1501-1576) 'Liber de Ludo Aleae' ('Boek over Kansspelen'). Andere pioniers op het gebied van de kansrekening waren Galileï, Pascal en ook onze landgenoot Christiaan Huygens. Van zijn hand verscheen in 1657 de verhandeling “Van rekeningh in Spelen van Geluck”.

wie heeft de meeste kans ?

In de volgende spelsituaties zijn we geïnteresseerd in de vraag wie de meeste kans heeft om te winnen. Lees eerst de vier problemen goed door, want er is een verband.

1 Aad en Ben werpen om de beurt een dobbelsteen. Winnaar is degene die als eerste een zes gooit. Aad mag beginnen. Als A geen zes gooit dan is B aan de beurt. Als die ook geen zes gooit dan mag A weer. Net zolang tot er een zes komt. Je voelt wel aan dat A meer kans heeft om te winnen dan B.
Hoe groot zijn hun kansen precies?

2 Een hoed bevat zes kaarjes genummerd 1, 2, ..., 6.
Om de beurt nemen A en B kaartjes uit de hoed (blindelings natuurijk). Degene die (als eerste) kaartje zes pakt is winnaar.
Wie heeft de meeste kans om te winnen
a. als de trekkingen worden gedaan zonder teruglegging;
b. als de trekkingen met teruglegging plaatsvinden.

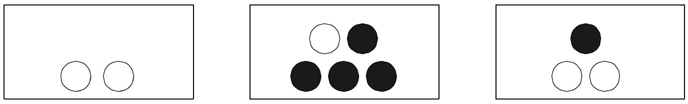
3 Een hoed bevat 36 kaartjes. Op ieder kaartje staat een rood nummer r en een zwart nummer z. Alle paren (r, z) met $r = 1, 2, \dots, 6$ en $z = 1, 2, \dots, 6$ zijn aanwezig.
De spelleider trekt blindelings een kaartje. Is het één van de 25 kaartjes zonder 6, dan moet opnieuw worden getrokken. Is het één van andere kaartjes en staat er een rode 6 op, dan wint A. Staat er geen rode 6 op en (dus) wel een zwarte, dan wint B.
a. Welke kaartjes zijn gunstig voor A? Welke voor B?
b. Hoe groot zijn hun kansen om te winnen bij de eerste trekking?
c. Hoe groot zijn hun uiteindelijke kansen om te winnen?
d. Maakt het wat uit of hier getrokken wordt mèt of zònder teruglegging?

4 Het volgende probleem komt uit het genoemde werk van Christiaan Huygens.
A en B werpen om beurten met twee dobbelstenen. A wint als hij 6 ogen werpt.
B wint als hij 7 ogen werpt. A heeft het voordeel dat hij mag beginnen. B heeft het voordeel dat het iets gemakkelijker is 7 ogen te werpen (kans 6/36) dan 6 ogen te werpen (kans 5/36).
a. Verklaar de kansen 6/36 en 5/36.
b. Het probleem wie hier de meeste kans heeft is nogal lastig. Huygens had de oplossing er niet bijgeschreven. Verschillende grote wiskundigen uit zijn tijd (zoals Jakob Bernoulli) hebben later een oplossing ervan gepubliceerd. Als jij er nu niet uitkomt, hoef je je zeker niet te schamen. Aan het eind van hoofdstuk 2 moet je het dan nog maar eens proberen.

Andere vragen die kunnen opkomen bij de voorgaande situaties, zijn: hoe lang kun je verwachten dat het spel gaat duren? Of: hoe vaak verwacht je dat de dobbelstenen moeten worden geworpen. Hoe groot is de kans dat de eerste 6 pas na de tiende worp komt? Dit soort vragen zullen verderop in dit hoofdstuk worden behandeld.

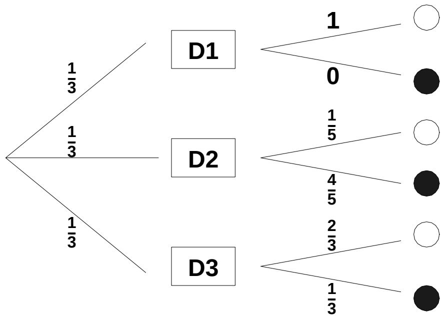
boomdiagrammen

Bij veel kansproblemen helpt het als je de situatie verduidelijkt met een boomdiagram. De vertakkingen geven de mogelijkheden die er bij iedere stap kunnen optreden en de bijgeschreven breuken geven de kansen aan. Een typisch voorbeeld is het volgende.
Gegeven zijn drie dozen met witte en zwarte ballen.



In de eerste doos zitten twee witte ballen. In de tweede doos één witte bal en vier zwarte. In de derde doos twee witte en één zwarte bal.

Kies blindelings en doos (kans $^1/_3$ voor iedere doos) en kies uit de gekozen doos een bal (gelijke kansen voor alle ballen in die doos.



De kans dat je een zwarte bal hebt is:

$$\frac{1}{3}\ 0+\frac{1}{3}\ \frac{4}{5}+\frac{1}{3}\ \frac{1}{3} = \frac{17}{45} = 0,377\ldots$$

5 **a.** Bereken ook op dezelfde manier de kans op een witte bal. Controleer of de som van beide kansen 1 is.

b. Er zijn evenveel witte ballen als zwarte ballen in de drie dozen samen. Waarom is de kans op een zwarte bal niet even groot als de kans op een witte?

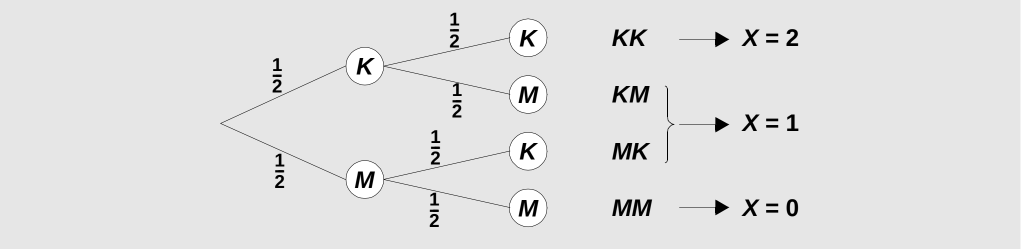
6 **a.** Hoe kun je de ballen (5 witte en 5 zwarte) verdelen over de drie dozen zó dat de kans op een witte bal even groot wordt als de kans op een zwarte? (Er zijn veel mogelijkheden.)

7 Je hebt 20 ballen met nummers 1, 2, … , 20. Hoe verdeel je ze over de drie dozen (in iedere doos minstens één bal) als je wilt dat de kans op bal 13 zo klein mogelijk wordt?

8 Drie kaartjes. Het eerste kaartje is aan beide kanten wit. Het tweede kaartje is aan de ene kant wit, aan de andere kant rood. Het derde kaartje is aan beide kanten rood. Deze drie kaartjes gaan in een zak. Op goed geluk neemt men een kaartje uit de zak. We zien alleen de rode bovenkant van het kaartje. Hoe groot is de kans dat de onderkant ook rood is?

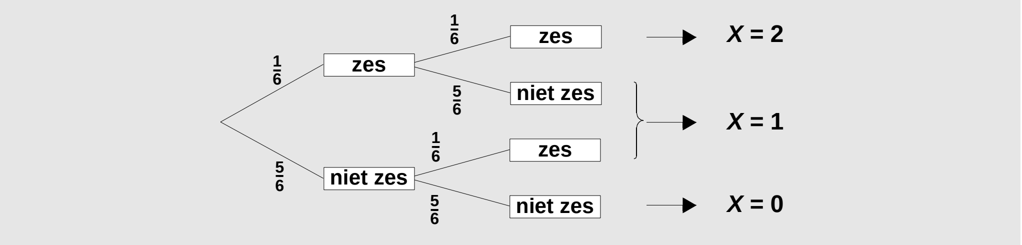
2: De binomiale verdeling

kop of munt Een munt die wordt opgegooid valt op de ene kant of op de andere kant met gelijke kans. Dat is wat men een 'eerlijke munt' noemt.
De kant die boven ligt heet *Kop* of *Munt*, kortweg *K* of *M*.
Er worden twee munten geworpen. Dat kan opleveren: *KK*, *KM*, *MK* of *MM*.
Stel: X = het aantal keren kop.



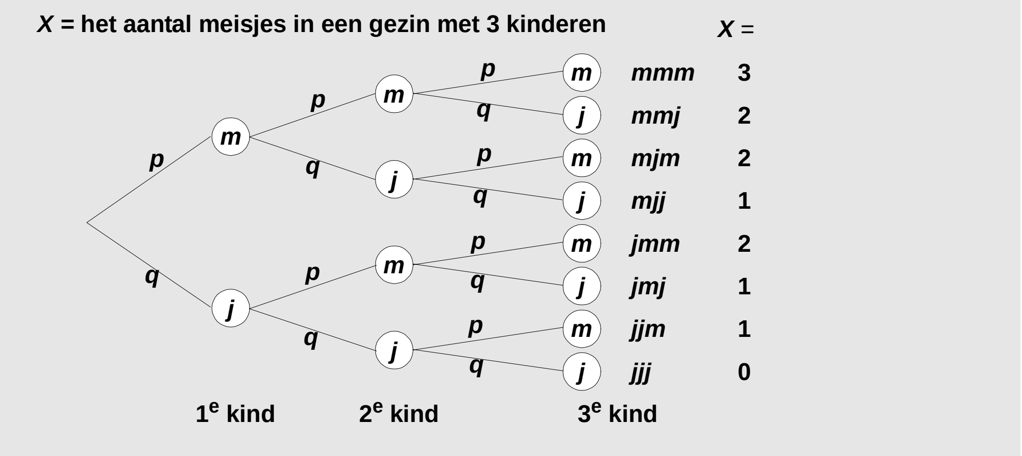
X neemt de waarden 0, 1 of 2 aan met respectievelijke kansen $\frac{1}{4}$, $\frac{2}{4}$, $\frac{1}{4}$
We schrijven: $P(X = 0) = \frac{1}{4}$, $P(X = 1) = \frac{1}{2}$, $P(X = 2) = \frac{1}{4}$.

2 dobbelstenen Een worp met 2 dobbelstenen (2 worpen met 1 dobbelsteen kan ook)
Het gaat nu om het *aantal zessen* dat we daarbij krijgen.
Noem dat aantal X.



9 Verklaar uit het boomdiagram dat de kansen $P(X = 0)$, $P(X = 1)$ en $P(X = 2)$ zich verhouden als 25 : 10 : 1.

gezin met 3 kinderen Bij iedere geboorte is de kans op een meisje p en de kans op een jongen q.
(Volgens statistische gegevens is q ongeveer 0,49 en p ongeveer 0,51).



10 $P(X = 3) = p \cdot p \cdot p = p^3$.

a. Druk ook $P(X = 2)$, $P(X = 1)$ en $P(X = 0)$ uit in p en q.

b. Hoe verhouden zich de kansen op 0, 1, 2 en 3 meisjes in het geval voor $p = q = \frac{1}{2}$?

multiple choice

De “Australian Mathematics Competition”, in Nederland ingevoerd onder de naam Kangoeroe wedstrijd bestaat uit 30 vragen. De eerste tien vragen zijn gemakkelijk. Daarna komen tien vragen die ook nog best te doen zijn maar wat meer inspanning vragen. De laatste tien vragen zijn behoorlijk “tough”. (In het kader zie je drie van zulke vragen).

1984 S. 25

On a tiny, remote island where the death sentence still exists, a man can be granted mercy after receiving a death sentence in the following way:
He is given 18 white balls and 6 black balls. He must divide them among three boxes with at least one ball in each box. Then blindfolded, he must choose a box atrandom, and then a ball from within this box. He receives mercy only if the chosen ball is white. The probability of the man receiving mercy (provided he has distributed the balls in the most favourable manner) is

(A) $\frac{11}{12}$ (B) $\frac{3}{4}$ (C) $\frac{10}{11}$ (D) $\frac{8}{11}$ (E) $\frac{1}{4}$

1989 S.30

Two forgetful friends agree to meet in a coffee shop one afternoon but each has forgotten the agreed time. Each remembers that the time was somewhere between 2 pm and 5 pm. Each decides to go to the coffee shop at a random time between 2 pm and 5 pm, wait half an hour, and leave if the other doesn’t arrive. What is the probability that they meet?

(A) $^{11}/_{36}$ (B)$^{1}/_{3}$ (C)$^{1}/_{6}$ (D)$^{25}/_{36}$ (E)$^{1}/_{36}$

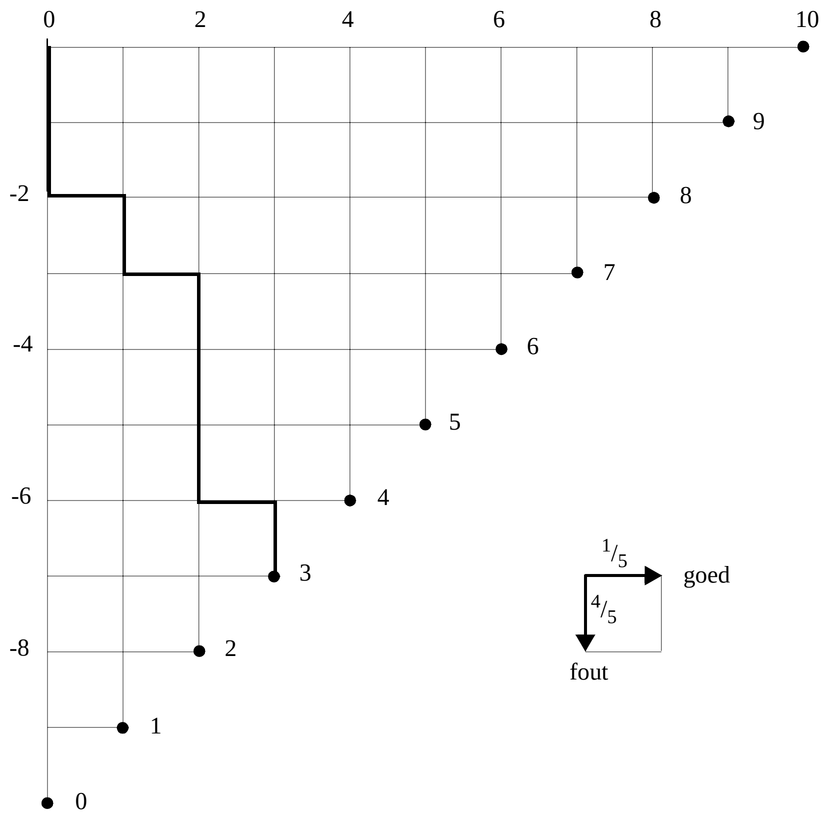
1985 S.29

An honest tennis promoter decides to stage an exhibition match between two players, whose names are to be drawn at random from a pool of N players, n of whom are Australians. He draws the two names and before announcing them is asked by a reporter “Will the match involve at least one Australian?” After looking at the drawn names he says “Yes”. The chance of both players being an Australian is

(A) $\frac{n-1}{2N-n-1}$ (B) $\frac{n-1}{N-1}$ (C) $\frac{n+1}{2N-n+1}$ (D) $\frac{3n-1}{2N+n-1}$ (E) $\frac{2n-1}{2N-1}$

Bruce is redelijk door de eerste 20 opgaven gekomen, maar hij heeft nu nog maar vijf minuten over voor de laatste 10 vragen, veel te kort dus. Dan maar gokken. Op goed geluk kiest Bruce bij iedere vraag één van de vijf antwoorden. We mogen erop rekenen dat hij bij iedere vraag een kans van 1 op 5 heeft goed te kiezen. Zodoende hoopt Bruce nog een paar punten ‘binnen te halen’ uit deze laatste opgaven. Als X het aantal goede antwoorden voorstelt in de laatste 10 opgaven, dan is X minimaal 0 en maximaal 10.

Het resultaat kan worden weergegeven in het diagram op de volgende bladzij.
Begin linksboven in het punt (0,0). Is het eerste antwoord goed dan ga je één stap naar rechts. Is het fout, dan één stap naar beneden. Net zo bij de volgende vragen. Is het ‘werk’ van Bruce nagekeken dan heb je een pad van (0,0) naar bijvoorbeeld (3,-7); dat betekent dan 3 goede en 7 foute antwoorden.



We kunnen dat pad ook noteren als een rijtje nullen en enen, 0 voor fout, 1 voor goed. Zo wordt het pad in de figuurvoorgesteld door: **0010100010**.
Dit zegt: de antwoorden op vragen 3, 5 en 9 waren goed, de andere fout.

11 a. Verklaar dat het betreffende rijtje de kans $(\frac{1}{5})^3(\frac{4}{5})^7$ heeft.

b. Die kans is niet al te groot. Hoeveel % ongeveer?

c. Is dit nu de kans op 3 goede antwoorden?

Er zijn een heleboel verschillende rijtjes met drie keer een **1** en zeven keer een **0.**
Evenveel als er paden (met lengte 10) zijn van (0,0) naar (3,-7).
Een manier om die paden te tellen is via *recursie*.
Om in (3, -7) aan te komen, moet je òf (2 , -7) òf (3 , -6) passeren!

het aantal paden van (0 , 0) naar (3, -7) **=** **het aantal paden van (0 , 0) naar (2, -7)** **+** **het aantal paden van (0 , 0) naar (3, -6)**

12 a. Wat kun je zeggen van het aantal paden van (0 , 0) naar (2, -7)?

b. Je komt zo steeds terugredenerend op de aantallen paden van (0 , 0) tot (0 , -1) en (1, 0) en die aantallen zijn naturlijk 1. Door vervolgens het optelprincipe vaak genoeg toe te passen, vind je dat het gezochte aantal paden van (0 , 0) naar (3 , -7) gelijk moet zijn aan 120. Controleer dit.

c. De kans op 3 goede en 7 foute antwoorden bij het op de gok beantwoorden van de vragen is ruim 20%. Verklaar dit via een berekening.

directe manier

Er is ook een directe manier om het aantal paden met lengte 10 van (0 , 0) naar (3, -7) te berekenen.

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**

Elk pad correspondeert met een *combinatie* van drie nummers uit een serie van tien. Door drie vakjes grijs te maken, kies je zo'n combinatie.
Omdat je voor het eerste grijze vakje de keus hebt uit 10, voor het tweede vakje de keus uit 9 en voor het derde de keus uit 8, lijkt het dat er $10 \cdot 9 \cdot 8$ mogelijkheden zijn.

13 $10 \cdot 9 \cdot 8 = 720$, maar hiervoor heb je ontdekt dat er 120 paden zijn!
Waar zit de fout? Hoe corrigeer je dat?

Het aantal 'combinaties van drie uit tien' noteren we zo: $\binom{10}{3}$ of zó: 10 C 3.
De directe manier om dit getal te berekenen is:

$$10C3 = \binom{10}{3} = \frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1} = 120$$

Je kunt de met de GR de uitkomst van 10 C 3 aldus vinden:

> Toets in 10.
> Kies dan in het MATH menu optie PRB (= probability) en daarin nCr .
> Toets dan 3 in gevolgd door ENTER en je krijgt de uitkomst.

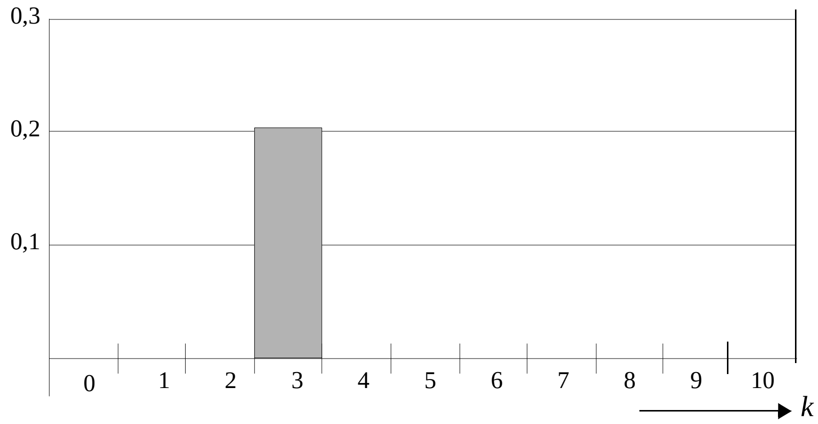
Terug naar de 10 vragen. De kans op "3 goed" bij 10 vragen, als bij iedere vraag de kans op een goed antwoord $^1/_5$ is, is gelijk aan

$$P(X = 3) = \binom{10}{3}\left(\tfrac{1}{5}\right)^3\left(\tfrac{4}{5}\right)^7 = 120 \cdot \frac{4^7}{5^{10}} = 0{,}2013...$$

14 Vul de tabel zelf verder in: $P(X = k) = \binom{10}{k}\left(\tfrac{1}{5}\right)^k\left(\tfrac{4}{5}\right)^{10-k}$ voor $k = 0, 1, ... , 10$

Controleer of de kansen bij elkaar opgeteld 1 geven. Completeer het kanshistogram.

k	$P(X = k)$
0	
1	
2	
3	0,2013
4	
5	
6	
7	
8	
9	
10	



driehoek van Pascal

Als je bij elk kruispunt in het rooster op bladzij 14 het aantal paden naar dat punt opschrijft, krijg je, zij het in gekantelde vorm, (een begin van) de driehoek van Pascal die je al bent tegengekomen in ‘Eindeloze Regelmaat’.

$$\begin{array}{ccccccccccccccc}
 & & & & & & & 1 & & & & & & & \\
 & & & & & & 1 & & 1 & & & & & & \\
 & & & & & 1 & & 2 & & 1 & & & & & \\
 & & & & 1 & & 3 & & 3 & & 1 & & & & \\
 & & & 1 & & 4 & & 6 & & 4 & & 1 & & & \\
 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & \\
 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & \\
1 & & 7 & & 21 & & 35 & & 35 & & 21 & & 7 & & 1
\end{array}$$

enz.

Elk getal in de driehoek kan via de stap-voor-stap-manier worden berekend, maar je hebt inmiddels gezien dat er ook een *directe* manier is. De getallen in de driehoek geven de aantallen combinaties van k uit n.
De driehoek van Pascal kan ook zó worden geschreven:

$$\begin{array}{ccccccccccccccc}
 & & & & & & & \binom{0}{0} & & & & & & & \\
 & & & & & & \binom{1}{0} & & \binom{1}{1} & & & & & & \\
 & & & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & & & \\
 & & & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & & & \\
 & & & \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} & & & \\
 & & \binom{5}{0} & & \binom{5}{1} & & \binom{5}{2} & & \binom{5}{3} & & \binom{5}{4} & & \binom{5}{5} & & \\
 & \binom{6}{0} & & \binom{6}{1} & & \binom{6}{2} & & \binom{6}{3} & & \binom{6}{4} & & \binom{6}{5} & & \binom{6}{6} & \\
\binom{7}{0} & & \binom{7}{1} & & \binom{7}{2} & & \binom{7}{3} & & \binom{7}{4} & & \binom{7}{5} & & \binom{7}{6} & & \binom{7}{7}
\end{array}$$

enz.

binomiaal-coëfficiënt

De getallen $\binom{n}{k}$ of nCk worden *binomiaalcoëfficiënten* genoemd.
$\binom{n}{k}$ kun je uitspreken als ‘n boven k‘.

Voor n en k geheel en niet negatief en k £ n, geldt:

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{k \cdot (k-1) \cdot \ldots \cdot 1}$$

(k factoren in de teller en in de noemer).
Voor het herhaalde produkt

$$k \cdot (k-1) \cdot (k-2) \cdot \ldots \cdot 1$$

schrijft men $k!$ (spreek uit k faculteit).
Bovenstaande formule wordt ook vaak in deze vorm gehanteerd:

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

Uit deze formule volgt de symmetrie-eigenschap:

$$\binom{n}{k} = \binom{n}{n-k}$$

15 Wat levert die laatste formule op voor $k = 2$ en $n = 5$? Hoe kun je de gelijkheid 'zien' in de driehoek van Pascal?

De randgetallen in de driehoek van Pascal zijn alle gelijk aan 1.
Dat betekent : voor $n = 0, 1, 2, 3, ...$

$$\binom{n}{0} = \binom{n}{n} = 1$$

Om de formule met de faculteiten ook in die gevallen te laten opgaan is de afspraak $0! = 1$ gemaakt.
Hieronder volgt een overzicht van de betekenis van de binomiaalcoëfficiënten:

$\binom{n}{k}$ **is het getal met rangnummer *k* in de rij met rangnummer *n* van de driehoek van Pascal;**

voorbeeld: rij 5 in de driehoek van Pascal = 1 5 10 10 5 1

$$\uparrow \binom{5}{2}$$

$\binom{n}{k}$ **is het aantal paden in een rooster van lengte *n*, met *k* 'horizontale stappen' en *n* - *k* 'verticale stappen ';**

voorbeeld: er zijn 10 paden met lengte 5 van punt (0 , 0) naar punt (2 , 3)

$\binom{n}{k}$ **is het aantal verschillende rijtjes met *k* 'nullen' en *n* - *k* 'enen';**

voorbeeld: er zijn 10 rijtjes met 2 nullen en 3 enen:

00111, 01011, 01101, 01110, 10011, 10101, 10110, 11001, 11010, 11100

$\binom{n}{k}$ **is het aantal combinaties van *k* elementen uit een verzameling van *n* elementen;**

voorbeeld: er zijn 10 combinaties van 2 uit $\{a_1, a_2, a_3, a_4, a_5\}$, namelijk: :

$\{a_1, a_2\}$,$\{a_1, a_3\}$, $\{a_1, a_4\}$, $\{a_1, a_5\}$, $\{a_2, a_3\}$, $\{a_2, a_4\}$, $\{a_2, a_5\}$, $\{a_3, a_4\}$, $\{a_3, a_5\}$, $\{a_4, a_5\}$

$\binom{n}{k}$ **is de coëfficiënt van $p^{n-k} q^k$ in de ontwikkeling van $(p + q)^n$';**

voorbeeld:

$$(p+q)^5 = \binom{5}{0}p^5 + \binom{5}{1}p^4 q + \binom{5}{2}p^3 q^2 + \binom{5}{3}p^2 q^3 + \binom{5}{4}p q^4 + \binom{5}{5}q^5$$
$$= p^5 + 5p^4 q + 10p^3 q^2 + 10p^2 q^3 + 5p q^4 + q^5$$

binomiale verdeling

In de voorbeelden van deze paragraaf ging het steeds om een proeven waarbij in elke proef slechts twee uitkomsten tellen.
- bij het opgooien van een geldstuk: kop of munt;
- bij het werpen met een dobbelsteen: zes of géén zes;
- bij een geboorte: jongen of meisje;
- bij een meerkeuzevraag: goed of fout.

Een dergelijke proef kan een aantal keren achtereen, zeg *n* keer, worden uitgevoerd,
Bij iedere proef kijken we of één bepaalde uitkomst optreedt.
Die uitkomst noemen we voor het gemak '*succes*'.
(Dat is alleen maar een manier van spreken. Het houdt geen waardering in. Zou je op een andere uitkomst letten dan noem je die uitkomst succes.)
Bij iedere proef is de kans op 'succes' hetzelfde, zeg *p*.
De kans op geen succes in één proef is dan 1 - *p* = *q*.
De *n* proeven zijn onderling *onafhankelijk*. Dit wordt in paragraaf 5 nog nader uitgelegd**.**
We tellen nu het torale aantal successen in de *n* proeven. Dat aantal noemen we *X*.
***X* kan de waarden 0, 1, 2, ..., *n* aannemen, met kansen:**

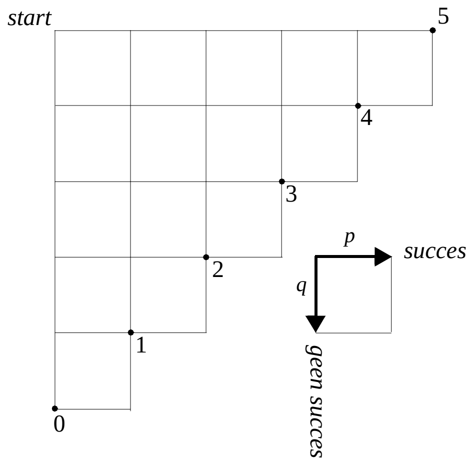
$$P(X = k) = \binom{n}{k} p^k q^{n-k}$$

voor *k* = 0, 1, 2, ... , *n*.

Voorbeeld:
n = 5

Bij 5 proeven (n = 5) zijn de kansen op 0, 1, 2, 3, 4 of 5 successen achtereenvolgens:

k	$P(X = k)$
0	q^5
1	$5pq^4$
2	$10p^2q^3$
3	$10p^3q^2$
4	$5p^4q$
5	p^5
	+
	1

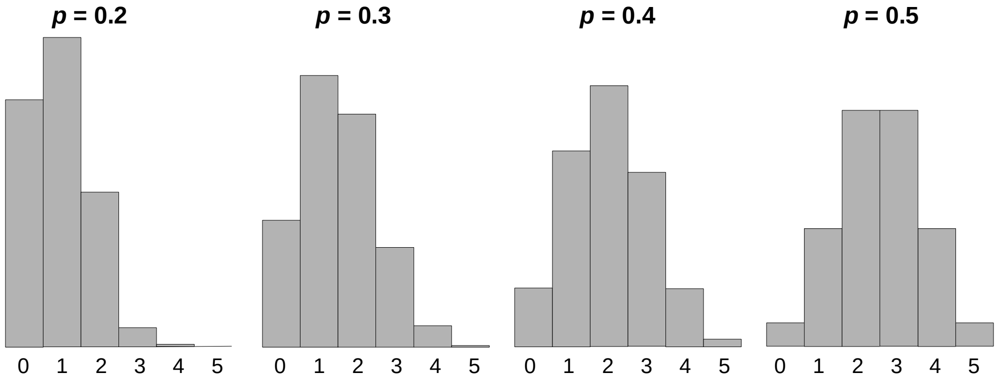


Dit is wat men noemt een *binomiale kansverdeling*.
De serie van n proeven noemt men een *binomiaal experiment*
Zo'n binomiaal experiment kun je voorstellen als een (toevallige) wandeling in een rooster: met kans p ga je bij een kruispunt naar rechts, met kans q naar beneden.
Hierboven zie je de mogelijke eindpunten bij n = 5.
De kansen in de tabel zijn juist de kansen om in de punten 0, 1, ... , 5 aan te komen.

16 De som van deze zes kansen moet gelijk zijn aan 1. Hoe kun je dat met algebra verifiëren? (Aanwijzing: raadpleeg het overzicht op de vorige bladzij).
Je zou kunnen zeggen dat de 'totale kans 1' opgedeeld is in een aantal porties.

17 Welke van de bovenstaande zes uitkomsten heeft de grootste kans in het geval p = 0.2?

18 Een kansverdeling kun je voorstellen doormiddel van een *kanshistogram*.
Hieronder zie je de kans histogrammen voor de binomiale verdeling met n = 5 en p achtereenvolgens 0.2, 0.3, 0.4 en 0.5



a. Controleer of je antwoord bij opgave **17** in overeenstemming is met het eerste kanshistogram.
b. Voor p = 0.5 is het kanshistogram symmetrisch. Hoe verklaar je dat? En hoe verhouden zich in dit geval de zes kansen?
c. Hoe kun je nu snel de kanshistogrammen vinden voor p = 0.6, 0.7, en 0.8 ?

19 Een meerkeuze-toets bestaat uit 6 vragen, elk met 4 mogelijke antwoorden.
a. Bereken de kans op precies 4 goede antwoorden bij zuiver gokken.
b. Bereken de kans op ten minste 4 goede antwoorden bij zuiver gokken.

20 Afke en Bert gooien ieder vier munten op.
a. Hoe groot is de kans dat ze beide hetzelfde aantal keren kop gooien?
(Stel X het aantal kop bij de vier munten van A, Y het aantal kop van B. Dan wordt gevraagd: $P(X = Y)$)
b. Bepaal $P(X > Y)$ de kans dat Afke meer kop werpt dan Bert.
Gebruik hierbij de symmetrie van de kansverdeling.
c. Bepaal $P(X + Y = 4)$.
d. Probeer te verklaren wat de antwoorden van **a** en **c** gelijk zijn?

21 **a.** Bereken de kans om in 4 worpen met een dobbelsteen minstens één zes te gooien.
b. De kans om in 8 worpen minstens 2 zessen te gooien.
c. De kans om in 24 worpen met twee dobbelstenen minstens één keer dubbel zes te gooien.

extra opgave

22 De uitwerking van $(q + p)^n = \binom{n}{0}q^n + \binom{n}{1}q^{n-1}p + \ldots$ geeft als termen de kansen bij de binomiale kansverdeling met *parameters n* en p.
Omdat $q + p = 1$ volgt hieruit dat de som van alle kansen bij die verdeling onderdaad gelijk is aan 1.
a. De termen van $(q - p)^n$ zijn afwisselend positief en negatief. Leg uit dat de positieve termen staan voor kansen op even aantallen successen.
b. Bewijs nu dat bij een binomiaal experiment de kans op een even aantal successen gelijk is aan $\frac{1}{2}(1 + (q-p)^n)$.

Aanwijzing: gebruik ook de ontwikkeling van $(q + p)^n$.

3: Wachten op succes

wachten op de eerste 6

Hoe vaak moet een dobbelsteen geworpen worden om een zes te krijgen?
Dat is niet te voorspellen. Soms komt de zes al bij de eerste worp. Dikwijls duurt het veel langer. Een enkele keer zijn er meer dan 20 worpen nodig.
Stel X = *het aantal worpen tot (en met) de eerste zes*. Door wat experimenteren krijgen we een indruk van de grote spreiding in de uitkomsten van X .
Tabel 1 geeft de resultaten van 100 herhalingen van dit experiment (10 series van 10). In de laatste serie moest er een keer 27 keer geworpen worden om een zes te krijgen.

Tabel 1: aantal worpen om een zes te krijgen

											som serie
serie 1	9	5	3	4	10	10	6	25	1	4	77
serie 2	2	7	1	3	4	4	1	8	2	6	38
serie 3	17	6	6	2	4	1	14	21	7	1	79
serie 4	3	1	3	7	15	1	6	3	1	1	41
serie 5	2	1	11	4	8	6	8	11	4	2	57
serie 6	2	5	1	6	2	5	9	1	1	17	49
serie 7	15	2	7	3	1	1	4	12	2	4	51
serie 8	7	2	9	11	18	5	5	15	14	2	78
serie 9	7	5	1	8	4	12	13	9	12	22	93
serie 10	6	1	27	1	18	10	5	9	4	18	99
											662

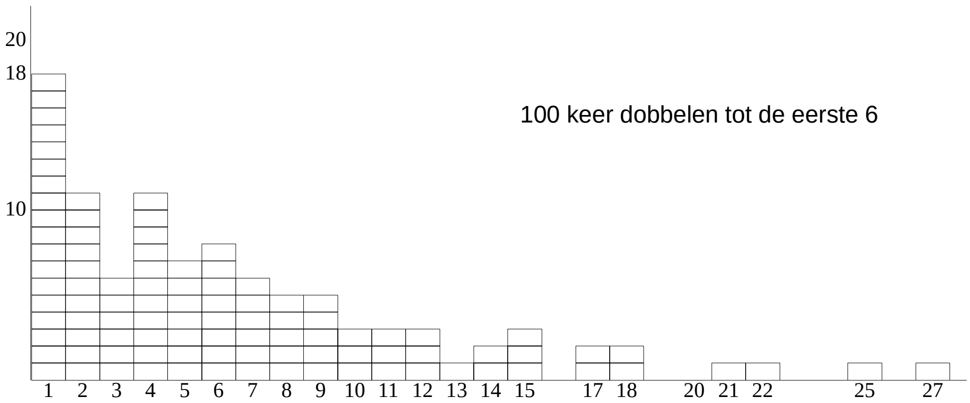
Tabel 2 geeft de frequenties waarmee de verschillende waarden van X optraden in deze 100 proeven.
(De proeven zijn niet echt met een dobbelsteen gedaan maar gesimuleerd met de TI-82: seq (1 + int(ln(rand)/ln(516)), A, 1, 10, 1) fi L_1 geeft een serie van 10 gesimuleerde waarden van X.)

Tabel 2: frequenties

1	18	**11**	3	**21**	1
2	11	**12**	3	**22**	1
3	6	**13**	1	**23**	-
4	11	**14**	2	**24**	-
5	7	**15**	3	**25**	1
6	8	**16**	-	**26**	-
7	6	**17**	2	**27**	1
8	5	**18**	2	**28**	-
9	5	**19**	-	**29**	-
10	3	**20**	-	**30**	-

Deze frequenties worden overzichtelijk weergegeven in een frequentie histogram

frequentie histogram.



Zijn deze resultaten in overeenstemming met de kansverdeling van X?
In 18 van de 100 proeven kwam de eerste zes al bij de eerste worp; de kans op die gebeurtenis : $P(X=1) = {}^1/_6$.
In 100 proeven verwacht je dus (${}^{100}/_6 = 16{,}66$), ongeveer 16 à 17 keer de waarde 1.
Het aantal 18 is daar goed mee in overeenstemming.

23 Nu ook de andere kansen $P(X = k)$ voor $k = 2, 3, \ldots$

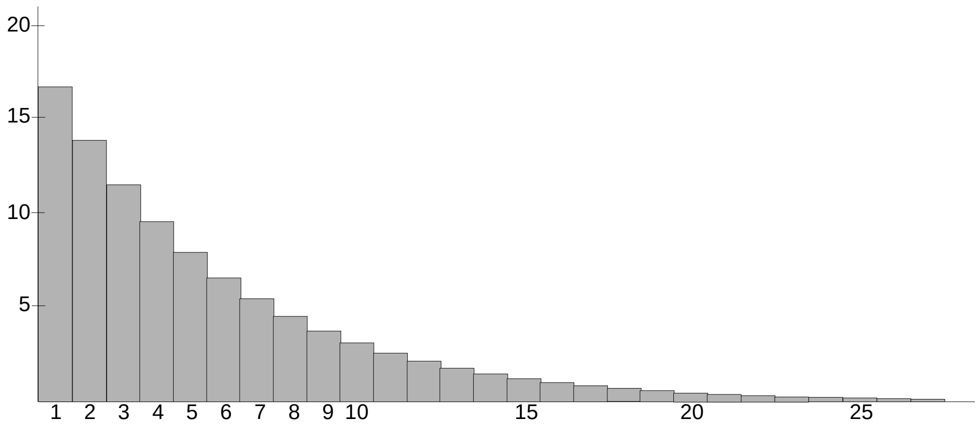
a. Verklaar de formule: $P(X = k) = \frac{1}{6} \left(\frac{5}{6}\right)^{k-1}$

b. Vul de kansentabel (met kansen in %, afgerond op honderste procenten) een stukje verder in

1	16,67	11	2,69	21	0,43
2	13,89	12		22	
3					
4					

kans-histogram

Kanshistogram voor aantal worpen t/m eerste 6 (kansen in %)



Het histogram van de kansen ziet er veel regelmatiger uit dan het histogram van de frequenties. Hadden we 1000 proeven gedaan in plaats van 100 dan was het frequentiehistogram waarschijnlijk mooier geworden, dat wil zeggen: beter lijkend op het kanshistogram.

24 Bij de 100 proeven waren er 4 waarbij meer dan 20 worpen nodig waren.

a. Hoe groot is de kans $P(X > 20)$ dat er bij zo'n proef meer dan 20 worpen nodig zijn? (N.B. Bedoeld is één enkele proef, dus niet te verwarren met de volgende vraag.)

b. Hoe groot is de kans dat er bij minstens één van de 100 proeven meer dan 20 worpen nodig zijn?

c. Hoe groot is de kans dat er bij minstens één van de honderd proeven meer dan 30 worpen nodig zijn?

25 X stelt weer het aantal worpen voor tot en met de eerste zes.

a. Hoe groot is de kans $P(X = n)$ voor $n = 1, 2, 3, \ldots$?

b. Verklaar: $P(X > n) = \left(\frac{5}{6}\right)^n$ voor $n = 1, 2, 3, \ldots$

c. De kans dat je meer dan 100 beurten moet wachten tor er een zes komt is heel erg klein, maar die gebeurtenis is niet geheel onmogelijk. Het is zelfs mogelijk dat er nooit een zes verschijnt (Mogelijk in een denkbeeldige zin, in een situatie waar onbeperkt lang gedobbeld kan worden.) De kans hierop is gelijk aan:

$$\lim_{n \to \infty} P(X > n)$$

Bereken die limiet.

Opmerking:
Bovenstaande limiet noteren we kortweg zó: $P(X = \infty)$.
Uit **25c** volgt: $P(X = \infty) = 0$.
De gebeurtenis '$X = \infty$', *er komt nooit een zes*, is een voorbeeld van een gebeurtenis die weliswaar mogelijk is, maar toch kans nul heeft!
Bij kansverdelingen met oneindig veel uitkomsten is 'onmogelijk' dus niet synoniem met 'kans is nul'.

Bij het examen voor het rijbewijs kan er van alles misgaan. Stel dat je bij ieder examen kans p hebt om te slagen. Bij het T-de examen gaat het goed. "U bent geslaagd", zegt de examinator. T kan de waarden 1, 2, 3, ... aannemen. De kansverdeling van T wordt gegeven door:

k	$P(X = k)$
1	p
2	qp
3	q^2p
4	q^3p
5	q^4p
...	...

Kortom:

$$\boldsymbol{P(T = k) = p \cdot q^{k-1}} \textbf{ voor } \boldsymbol{k = 1, 2, 3, \ldots\ldots}$$

geometrische verdeling

Dit geldt voor de algemene situatie van een rij onafhankelijke proeven, die ieder met kans p een succes opleveren. Het nummer T van de proef waarbij voor het eerst een succes optreedt, heeft de bovenstaande kansverdeling.
Merk op dat de kansen

$$P(T = 1),\ P(T = 2),\ P(T = 3),\ \ldots$$

een meetkundige rij (met reden q) vormen.
Daarom noemt men dit de *geometrische* (= meetkundige) kansverdeling.
Men zegt: T heeft een geometrische verdeling (met parameter p).

26 Verifieer dat geldt: de 'som' van alle kansen van de geometrische verdeling = 1.
Kortweg:

$$\sum_{k=1}^{\infty} P(T = k) = 1$$

27 Arie en Bart spelen een aantal partijen tennis.
Bij iedere partij heeft Arie een kans a om te winnen; voor Bart is de kans om te winnen b (= 1 - a). Ze spelen net zoveel partijen tot ieder ten minste één maal gewonnen heeft. Het aantal partijen dat daarvoor nodig is, noemen we X.
a. Bepaal $P(X = 2)$ en $P(X = 3)$.
b. Wat betekent precies $X > 3$? Bepaal $P(X > 3)$.

28 Bij sommige sporten (voetbal, schaken) kan een partij onbeslist eindigen (gelijk spel, remise). Stel de kans op winst voor A = a, de kans op winst voor B = b en de kans op onbeslist = c. Er geldt nu dus: $a + b + c = 1$. Het aantal partijen totdat ieder tenminste één keer gewonnen heeft, stellen we weer X.
Bepaal weer $P(X = 2)$ en $P(X = 3)$**.**

29 Terug naar de situatie van **27.** Arie en Bart spelen nu net zo lang tot één van hen twee keer achter elkaar gewonnen heeft. Y is het aantal hiervoor benodigde partijen.
Bepaal $P(Y = k)$ voor $k = 2, 3, \ldots$.

4: Verwachtingswaarde

Bij een kansspel (bijvoorbeeld lotto of toto) moet je een zeker bedrag betalen om mee te doen: de *inzet*. Je doet dat in de hoop op een uitbetaling of prijs die de inzet vele malen overtreft. Meestal krijg je niets. Als je heel vaak mee zou spelen en de gemiddelde uitkering blijkt meer te zijn dan je inzet, is het spel gunstig. De ondernemer (onderneming) die het spel organiseert, bekijkt de zaak van de andere kant. Voor hem is het spel gunstig als de ontvangen inzetten meer zijn dan het uit te keren prijzengeld.

Een voorbeeld.

chuck a luck

Het spel Chuck a Luck komt uit Amerika. De deelnemer zet één dollar in op één van de nummers 1, 2, 3, 4, 5 of 6. Daarna mag hij drie dobbelstenen gooien.
Komt zijn nummer 1, 2 of 3 keer boven, dan krijgt hij 1, 2 of 3 dollar uitbetaald en bovendien krijgt hij zijn inzet terug.
Vertoont geen van de drie dobbelstenen het nummer waarop is ingezet dan is die inzet voor het ‘huis’.
De winst van de speler kan dus bedragen -1, 1, 2 of 3 dollar, al naar gelang 0, 1, 2 of 3 dobbelstenen het gewenste nummer vertonen.
Is dit spel gunstig voor de speler of voor het huis? Dat is natuurlijk een lange termijn kwestie!

30 Stel het winstbedrag voor door W en bereken achtereenvolgens $P(W = -1)$, $P(W = 1)$, $P(W = 2)$ en $P(W = 3)$, dus de kansverdeling van W.

gemiddelde winst

Een speler doet een groot aantal keren mee, zeg n keer en zet steeds in op 5.
Daarbij zijn n_k ($k = 0, 1, 2, 3$) spelen waarbij het nummer 5 op k van de drie dobbelstenen verschijnt.
Er geldt dus; $n_0 + n_1 + n_2 + n_3 = n$. De getallen n_k worden de *frequenties* van de verschillende uitkomsten genoemd. De totale winst van de speler is dan

$$-1 \quad \frac{n_0}{n} + 1 \quad \frac{n_1}{n} + 2 \quad \frac{n_2}{n} + 3 \quad \frac{n_3}{n}$$

Als n heel groot is, zal de *relatieve frequentie* $\frac{n_k}{n}$ niet veel verschillen van de kans p_k (dat wil zeggen van de kans dat k van de drie dobbelstenen het gewenste nummer vertonen).
Als het spel dus heel vaak gespeeld wordt, kunnen we *verwachten* dat de gemiddelde winst dicht bij de waarde:

$$-1 \quad p_0 + 1 \quad p_1 + 2 \quad p_2 + 3 \quad p_3$$

zal liggen.

verwachtingswaarde

Dit getal noemt men dan ook de *verwachtingswaarde* van de winst (of verwachte winst).
Notatie: $E(W)$ (E van *expectation*). Dus: $E(W) = -1 \quad p_0 + 1 \quad p_1 + 2 \quad p_2 + 3 \quad p_3$
Algemeen:
Als W de waarden $w_1, w_2, w_3, \ldots$ kan aannemen met kansen $p_1, p_2, p_3, \ldots$ dan

$$E(W) = w_1 p_1 + w_2 p_2 + w_3 p_3 + \ldots$$

31 a. Bereken in voor het Chuck a Luck spel de verwachtingswaarde $E(W)$.
b. Verifieer dat ‘het huis’ op den duur 8% van alle inzetten opstrijkt.

32 a. Terug naar opgave **8** van hoofdstuk 1. Verklaar dat geldt:
de verwachtingswaarde voor het aantal aanvragen per minuut is gelijk aan $b + 2c$.
b. Hoe groot is de verwachtingswaarde voor de drie gevallen 1, 2 en 3 dus?

gemiddelde versus verwachting

We moeten onderscheid maken tussen de *gemiddelde winst* en de *verwachte winst*.
Een gemiddelde wordt berekend uit een *eindig* aantal waarnemingen (proeven, spelen) en hangt af van het *toeval*. Het kan hoger of lager uitvallen.
De verwachtingswaarde is een exacte waarde berekend uit de kansverdeling en is onafhankelijk van toevallige uitkomsten.
De gemiddelde winst uit n spelen kan opgevat worden als een *schatting* voor het ideale gemiddelde: de verwachte winst, net zoals de relatieve frequenties n_k/n van de uitkomsten opgevat kunnen worden als schattingen voor de kansen p_k op die uitkomsten.
Dit wordt gedemonstreerd door het volgende simulatie experiment. Er zijn 12 series van 216 spelen geproduceerd.

	1	2	3	4	5	6	7	8	9	10	11	12	Totaal
n_0	125	123	120	137	126	126	135	140	114	136	118	118	1518
n_1	73	76	83	63	71	74	68	64	84	66	76	81	879
n_2	17	14	13	16	18	14	13	11	18	14	20	14	182
n_3	1	3	0	0	1	2	0	1	0	0	2	3	13
W	-15	-10	-11	-42	-16	-18	-41	-51	+6	-42	+4	0	-236

Bij iedere serie is ook de winst: $-1\ \ n_0 + 1\ \ n_1 + 2\ \ n_2 + 3\ \ n_3$ bepaald.
Omdat de verwachte frequenties 125, 75, 15 en 1 zijn (in overeenstemming met de kansen 125/216, 75/216, 15/216 en 1/216) zou de verwachte winst in een serie van 216 spelen dus $(-1 \cdot 125) + (2 \cdot 15) + (3 \cdot 1) = -17$ zijn.
De totale winst in de 12 series ($12 \cdot 216 = 2592$ spelen) is gelijk aan –236 waar uit een gemiddelde winst: $-236/2592 = -0,091$ volgt, wat niet veel verschilt van de verwachte winst: $-17/216 = -0,078$.

definitie

De algemene regel voor de berekening van de verwachtingswaarde $E(X)$ van X is :

Als X de waarden $x_1, x_2, x_3, \ldots$ kan aannemen met kansen $p_1, p_2, p_3, \ldots$,

dan is $E(X)$ de verwachting van X bepaald door:

$$E(X) = \sum_{k=1}^{\infty} x_k\ p_k$$

Merk op: als er eindig veel waarden kunnen worden aangenomen dan geldt vanaf een zeker rangnummer N: als $n > N$, dan $x_n = 0$ en $p_n = 0$.
Over terminologie:
Een grootheid X die afhankelijk van het toeval allerlei waarden aan kan nemen, noemt men een *toevalsvariabele* (of stochastische variabele). In de Engelse taal spreekt men van *random variable.*

33 X is het aantal ogen bij een worp met een dobbelsteen.
Verklaar: $E(X) = 3\frac{1}{2}$.

34 In een zak zitten zes bankbiljetten. Drie briefjes van tien, twee van vijfentwintig en één van honderd. Blindelings neem je één bankbiljet uit de zak (ieder briefje heeft dezelfde kans, nemen we aan). De waarde van het getrokken biljet noemen we X.
Toon aan: $E(X) = 30$

waarschuwing De verwachte waarde is niet zoiets als de meest waarschijnlijke waarde!
Het hoeft zelfs geen mogelijke waarde te zijn!
De dobbelsteen geeft nooit $3\frac{1}{2}$ en uit de zak in **32** komt nooit een briefje van 30.

35 Werp drie dobbelstenen. X is het aantal zessen dat verschijnt. Bereken $E(X)$.

36 Y is het *totale aantal* ogen bij een worp met 2 dobbelstenen.
Y kan dus de waarden 2, 3, ..., 12 aannemen.
a. Teken het kanshistogram van Y.
b. Bepaal $E(Y)$.

37 Je werpt drie keer met een geldstuk. X is het aantal keren kop.
a. Teken het kanshistogram van X.
b. Bepaal $E(X)$.

38 *Best of five*. Bij een tenniswedstrijd wordt zolang gespeeld totdat één van beide spelers 3 sets heeft gewonnen. Het aantal sets X kan dus 3, 4 of 5 zijn.
Beschouw de wedstrijd als een kansspel, met vaste kansen voor elk van de beide spelers om een set te winnen.
a. Bepaal $E(X)$ in het geval de spelers even sterk zijn.
b. Stel de spelers A en B hebben respectievelijk winstkansen p en $q = 1 - p$.
Druk $E(X)$ uit in p.
c. In welke situatie is $E(X)$ maximaal?

39 Gegeven een loterij met 4000 loten van 1 gulden.
De hoofdprijs is *f* 1000,-, de tweede en de derde prijs zijn beide *f* 500, - en dan zijn er nog 5 prijzen van *f* 100,-
Iemand koopt twee loten. Wat is de verwachtingswaarde van zijn winst (verlies)?

somregel De verwachtingswaarde van het aantal ogen X bij het werpen met 1 dobbelsteen is $3\frac{1}{2}$.

De verwachtingswaarde van het aantal ogen Y bij het werpen met 2 dobbelsteen is 7.
Je hebt dit in **35** berekend via de kansverdeling van Y.

Maar het kan ook zo: $3\frac{1}{2} + 3\frac{1}{2} = 7$.

Dit laatste berust op de zogenaamde somregel voor verwachtingswaarden.
Die regel zegt:

> **Als X_1 en X_2 toevalsvariabelen zijn , met $E(X_1) = \mathrm{m}_1$ en $E(X_2) = \mathrm{m}_2$,**
>
> **dan is $Y = X_1 + X_2$ een toevalsvariabele met $E(Y) = \mathrm{m}_1 + \mathrm{m}_2$**

In het voorbeeld van de dobbelstenen stelt X_1 het aantal ogen van de eerste steen en X_2 dat van de tweede steen voor.

Er geldt blijkbaar: $E(Y) = E(X_1) + E(X_2)$.

Bovensrtaande regel is waar voor elk tweetal toevalsvariabelen. Zij geldt ook als er sprake is van meer dan twee toevalsvariabelen. (Zelfs voor een som met oneindig veel (positieve) termen kan worden bewezen dat de somregel voor de verwachtingswaarden geldt.)
Dat is erg handig, want nu kun je bijvoorbeeld de verwachting van X, het aantal ogen bij vier dobbelstenen, heel eenvoudig berekenen. Immers $X = X_1 + X_2 + X_3 + X_4$ en dus:

$$E(X) = E(X_1) + E(X_2) + E(X_3) + E(X_4) = 4 \cdot 3\tfrac{1}{2} = 14.$$

We hebben de (tamelijk ingewikkelde) kansverdeling van X hierbij niet nodig.

binomiale verdeling

Een ander voorbeeld is de verwachtingswaarde van het aantal successen bij de *binomiale verdeling*.
Voor het geval $n = 1$ geldt:

k	$P(X = k)$
0	q
1	p

Dus: $E(X) = q \quad 0 + p \quad 1 = p$

Voor $n = 2$ geldt:

k	$P(X = k)$
0	q^2
1	$2pq$
2	p^2

Dus: $E(X) = q^2 \quad 0 + 2pq \quad 1 + p^2 \quad 2 = ...$

40 a. Controleer dat in het laatste geval de verwachtingswaarde gelijk is aan $2p$.
b. Een slimme manier om dit resultaat te krijgen is met behulp van de somregel. Stel X_1 is het aantal successen in de eerste en X_2 is het aantal successen in de tweede proef. Er geldt dan $X = X_1 + X_2$.
Hoe volgt nu uit het bovenstaande dat $E(X) = 2p$?
c. Laat zien dat de verwachtingswaarde van het aantal successen bij een binomiale verdeling met $n = 3$ gelijk is aan $3p$.

Het resultaat van **41** is nauwelijks verrassend. Als een operatie met slaagkans p honderd keer wordt uitgevoerd, dan is de verwachtingswaarde van het aantal geslaagde operaties gelijk $100p$.
Algemeen: als X het aantal successen in n proeven met succeskans p, dan:

$$E(X) = np$$

Zonder somregel is het bewijs ingewikkeld, maar met de somregel gaat het gemakkelijk. Het totaal aantal successen X is de som van de bijdragen van de afzonderlijke proeven: $X = X_1 + X_2 + ... + X_n$, waarbij $X_k = 1$ als de k-de proef een succes geeft en $X_k = 0$ als dit niet zo is.
Omdat $E(X_k) = p$ voor $i = 1, 2, ... , n$, krijgen we als som van de verwachtingen np.

extra opgave

41 *Bloedonderzoek.*
Het bloed van een groot aantal (N) mensen moet worden onderzocht op de aanwezigheid van een bepaald virus. Van ieder is een bloedmonster afgenomen. Als ieder monster apart onderzocht wordt, zijn er N analyses nodig. Die vallen meestal negatief uit, want het virus wordt maar bij een kleine fractie (1 à 2%) van deze mensen aangetroffen. Omdat deze analyses vrij tijdrovend en kostbaar zijn, heeft men het volgende bedacht: Men verdeelt de N monsters in groepen van k. Van zo'n groep k neemt men van ieder monster een beetje en doet dat bij elkaar.
Blijkt dat mengsel *negatief* dan zijn die k personen gezond en is men met één analyse klaar.
Is het mengsel *positief* dan moeten alle k monsters apart onderzocht worden, dus dan zijn er bij deze groep $k + 1$ analyses nodig.
Neem aan dat bij iedere persoon er een kans p is dat het virus aanwezig is (en dat er onafhankelijkheid is tussen de personen).

a. De kans dat in een mengsel van het bloed van k personen het virus wordt aangetroffen is gelijk aan $1 - (1 - p)^k$. Toon dit aan.
b. Het aantal analyses dat gedaan moet worden bij een groep van k personen is een toevalsvariabele X die de waarden 1 en $k + 1$ kan hebben. Druk $E(X)$ uit in p en k.
c. De optimale groepsgrootte k is die waarbij de verwachtingswaarde van het totale aantal proeven per groep minimaal is. Stel dat $p = 0.02$. Bepaal met behulp van de GR wat de optimale aantal bloedmonsters per groep is bij deze keuring.

wachten op het eerste succes

Neem weer een serie onafhankelijke proeven met elk succeskans p.
Bij proef nummer T krijgen we voor het eerst een succes.
T heeft een geometrische verdeling:

$$P(T = k) = p \quad q^{k-1}$$

De vraag is nu wat de verwachtingswaarde van T is.

42 We bekijken eerst een eenvoudig voorbeeld.
Je werpt met een geldstuk net zo lang tot Kop boven komt.
De verwachtingswaarde is de limietsom:

$$1 \quad \frac{1}{2} + 2 \quad \frac{1}{4} + 3 \quad \frac{1}{8} + 4 \quad \frac{1}{16} + 5 \quad \frac{1}{32} + \ldots$$

a. Verklaar dit.
b. Bereken die som (aanwijzing: zie het boekje 'Eindeloze regelmaat').

Opgave **42b** kan ook worden opgelost met de somregel voor de verwachtingswaarde.
We tellen de proeven: de eerste proef, de tweede, de derde, ... We stoppen zodra er een succes is. De eerste proef moet altijd gedaan worden. De tweede hoeft alleen als de eerste mislukt. De derde als de eerste en de tweede beide mislukken, enzovoorts.
De stochast X_k geeft aan of de k^e proef gedaan wordt of niet.
$X_k = 1$ als de k^e proef gedaan wordt en 0 als dat niet hoeft.
Dan is duidelijk dat geldt: $T = X_1 + X_2 + X_3 + \ldots$.
Immers, als bijvoorbeeld bij de 5e proef het eerste succes komt dan hebben X_1 tot en met X_5 de waarde 1 en vanaf $k = 6$ zijn alle X_k's nul, dus $T = 1 + 1 + 1 + 1 + 1 + 0 + 0 + \ldots = 5$.
Omdat X_k alleen waarde 0 of 1 kan hebben, is $E(X_k) = \left(\frac{1}{2}\right)^{k-1}$ (de k^e proef wordt gedaan als de $k - 1$ voorgaande proeven mislukken, anders hoeft het niet).

Met de somregel krijgen we dus

$$E(T) = E(X_1) + E(X_2) + E(X_3) + \ldots = 1 + \frac{1}{2} + \frac{1}{4} + \ldots = 2$$

43 Toon aan dat bij een succeskans *p* de verwachtingswaarde van *T* gelijk is aan:

$$\frac{1}{1-q} = \frac{1}{p}$$

Je hebt de keus uit twee mogelijkheden:
de directe manier; je stuit dan op de limietsom van een reken-meetkundige rij;
de manier met de hulpvariabelen en de somregel: je krijgt dan de limietsom van een meetkundige rij.

opmerkingen Het resultaat van **43** is heel aannemelijk: hoe kleiner de succeskans hoe meer proeven er gemiddeld nodig zijn om een succes te bereiken.

44 Een eerlijke munt wordt zolang geworpen tot beide kanten minstens één keer boven zijn gekomen.

a. Bepaal de verwachting van het aantal worpen.

b. Als de munt niet eerlijk is – hij valt met kans *p* kop en met kans $q = 1 - p$ munt. Wat is nu het verwachte aantal worpen tot beide kanten zijn opgetreden? Voor welke *p* is die verwachting minimaal?

45 Je gooit net zo lang met een dobbelsteen totdat alle uitkomsten 1, 2, ... , 6 ten minste één keer aan de beurt zijn geweest. Wat is de verwachtingswaarde van het aantal worpen dat daarvoor nodig is?