OVERZICHT FORMULES

Kansrekening

Voor toevalsvariabelen X en Y geldt: $E(X+Y)=E(X)+E(Y)$

Voor onafhankelijke toevalsvariabelen X en Y geldt:

$\sigma(X+Y)=\sqrt{\sigma^2(X)+\sigma^2(Y)}$

$\sqrt{n}$ -wet: bij een serie van n onafhankelijk van elkaar herhaalde experimenten geldt voor de som S en het gemiddelde $\overline{X}$ van de uitkomsten X:

$E(S)=n\cdot E(X) \qquad \sigma(S)=\sqrt{n}\cdot\sigma(X)$

$E(\overline{X})=E(X) \qquad \sigma(\overline{X})=\dfrac{\sigma(X)}{\sqrt{n}}$

Binomiale verdeling

Voor de binomiaal verdeelde toevalsvariabele X, waarbij n het aantal experimenten is en p de kans op succes per keer, geldt:

$\mathrm{P}(X=k)=\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}$ met k = 0, 1, 2, 3, …, n

Verwachting: $E(X)=n\cdot p$ Standaardafwijking: $\sigma(X)=\sqrt{n\cdot p\cdot(1-p)}$

Normale verdeling

Voor een toevalsvariabele X die normaal verdeeld is met gemiddelde μ en standaardafwijking σ geldt:

$Z=\dfrac{X-\mu}{\sigma}$ is standaard-normaal verdeeld en $\mathrm{P}(X<g)=\mathrm{P}(Z<\dfrac{g-\mu}{\sigma})$

Differentiëren

naam van de regel	**functie**	**afgeleide**
somregel	$s(x)=f(x)+g(x)$	$s'(x)=f'(x)+g'(x)$
productregel	$p(x)=f(x)\cdot g(x)$	$p'(x)=f'(x)\cdot g(x)+f(x)\cdot g'(x)$
quotiëntregel	$q(x)=\dfrac{f(x)}{g(x)}$	$q'(x)=\dfrac{f'(x)\cdot g(x)-f(x)\cdot g'(x)}{(g(x))^2}$
kettingregel	$k(x)=f(g(x))$	$k'(x)=f'(g(x))\cdot g'(x)$ of $\dfrac{\mathrm{d}k}{\mathrm{d}x}=\dfrac{\mathrm{d}f}{\mathrm{d}g}\cdot\dfrac{\mathrm{d}g}{\mathrm{d}x}$

Logaritmen

regel	voorwaarde
${}^{g}\log a + {}^{g}\log b = {}^{g}\log ab$	$g>0,\ g\neq 1,\ a>0,\ b>0$
${}^{g}\log a - {}^{g}\log b = {}^{g}\log \frac{a}{b}$	$g>0,\ g\neq 1,\ a>0,\ b>0$
${}^{g}\log a^{p} = p \cdot {}^{g}\log a$	$g>0,\ g\neq 1,\ a>0$
${}^{g}\log a = \frac{{}^{p}\log a}{{}^{p}\log g}$	$g>0,\ g\neq 1,\ a>0,\ p>0,\ p\neq 1$

Wikipedia

Wikipedia is een internationale internet-encyclopedie.
In maart 2012 bevatte de Nederlandstalige editie ruim één miljoen artikelen. In de tabel staan gegevens van 2012.

tabel

datum	22 maart	29 maart	5 april	12 april	19 april
aantal	1 033 414	1 034 660	1 035 882	1 037 184	1 038 340

Zoals in bovenstaande tabel te zien is, groeit het aantal artikelen flink. Sommigen beweren dat hier sprake is van lineaire groei, anderen houden het op exponentiële groei.

4p **1** Onderzoek elk van deze beweringen.

Over een langere periode bleek de groei sterker te worden: in de 23 weken van 19 april tot 27 september 2012 groeide de Nederlandstalige Wikipedia uit tot 1 120 987 artikelen.
Neem aan dat het aantal artikelen vanaf 19 april exponentieel groeide en in de toekomst met dezelfde factor blijft groeien.

4p **2** Bereken het aantal artikelen op 19 april 2014.

De relatief grote omvang van de Nederlandstalige Wikipedia is voor een deel te verklaren door het grote aantal door computers gegenereerde artikelen. Het zijn wel echte artikelen maar ze zijn erg kort en geven informatie die niet bijzonder interessant is. Een voorbeeld van zo'n artikel:

Miedzianów

Miedzianów is een dorp in de Poolse woiwodschap Groot-Polen. De plaats maakt deel uit van de gemeente Nowe Skalmierzyce en telt 200 inwoners.

Het valt niet op dat er zo veel van deze artikelen zijn. Alleen door in het beginscherm van Wikipedia een willekeurige pagina te vragen, komen deze 'computerartikelen' tevoorschijn.

In januari 2013 werd vastgesteld dat een derde deel van alle artikelen door computers gegenereerd was. Het aantal gewone artikelen groeide op dat moment exponentieel met een jaarlijkse toename van 5%. Het aantal computerartikelen groeide echter jaarlijks met 17%. Veronderstel dat de groei van beide soorten artikelen zich de jaren erna op dezelfde wijze voortzet.

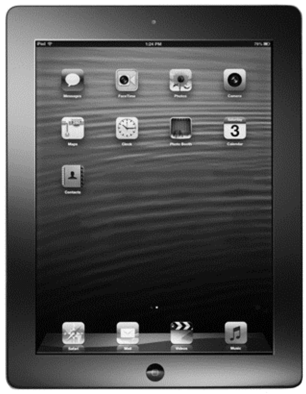
5p **3** Bereken na hoeveel jaar er meer computergegeneerde artikelen zullen zijn dan gewone artikelen. Geef je antwoord in maanden nauwkeurig.

Inmiddels wordt beweerd dat meer dan 40% van alle artikelen van de Nederlandstalige Wikipedia door een computer gegenereerd is. Bij een test in 2014 werden 50 willekeurige artikelen opgevraagd. Daarvan waren er 28 door een computer gegenereerd.

6p **4** Onderzoek met het toetsen van hypothesen met een significantieniveau van 1% of dit voldoende reden geeft om te veronderstellen dat meer dan 40% van de artikelen computerartikelen zijn.

Touchscreens

foto



Bij het ontwerpen van touchscreens (aanraakschermen) voor moderne media als tablets en mobiele telefoons besteedt men veel aandacht aan het gebruiksgemak. Gebruikers willen immers snel kunnen navigeren.
Op de foto zie je een touchscreen met een menu bestaande uit 13 knoppen.

De tijd die je nodig hebt om in een menu de juiste knop te vinden, hangt mede af van het aantal knoppen in het menu.
Volgens de psycholoog Hick kun je deze benodigde tijd T berekenen met de formule:

$$T(n) = b \cdot {}^{2}\log(n+1)$$

Hierbij is T de tijd in seconden, n het aantal knoppen in het menu en b een positieve constante die afhangt van de behendigheid van de gebruiker.

In deze opgave kijken we naar dit model van Hick.
Om de juiste knop te vinden op het touchscreen van de foto heeft Irene 8 seconden nodig.

3p **5** Bereken haar waarde van b in één decimaal nauwkeurig.

Pim is veel handiger met een touchscreen dan zijn vader. Hij kan in een menu met 16 knoppen even snel de juiste knop vinden als zijn vader in een menu met 4 knoppen. Dit betekent dat zijn b-waarde (b_p) kleiner is dan de b-waarde van zijn vader (b_v).

4p **6** Onderzoek of dit betekent dat de b-waarde van Pim precies half zo groot is als die van zijn vader.

Sommige gebruikers vinden een menu met veel knoppen onoverzichtelijk. Daarom deelt men een menu soms op in submenu's met minder knoppen. Als er bijvoorbeeld in totaal 18 knoppen zijn, kan de ontwerper ervoor kiezen om:

methode I één menu van 18 knoppen te maken

of

methode II een menu met 3 knoppen te maken, waarbij na elk van de 3 mogelijke keuzes weer een submenu met 6 knoppen verschijnt.

De gebruiker wint hiermee overzichtelijkheid want hij weet nu precies in welk submenu hij moet zoeken, maar hij verliest tijd omdat hij twee keer (in een menu) de juiste knop moet zien te vinden.

Als $b = 0{,}9$ duurt het keuzeproces bij methode II minstens 0,5 seconde langer dan bij methode I.

3p **7** Toon met behulp van de formule voor $T(n)$ aan dat dit juist is.

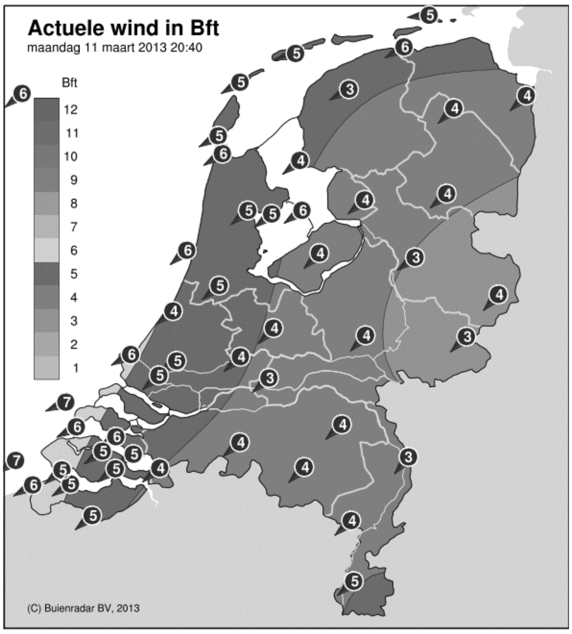
Uit de formule volgt dat één menu met alle knoppen altijd sneller werkt dan een opdeling in submenu's. Dus: één menu met $p \cdot q$ knoppen is altijd sneller dan een hoofdmenu met p knoppen gevolgd door p submenu's met elk q knoppen.

4p **8** Neem $b = 1$ en toon aan dat $T(p) + T(q)$ altijd groter is dan $T(p \cdot q)$.

Wind mee, wind tegen

Op de site buienradar.nl kun je verschillende weerkaarten bekijken. De kaarten bevatten actuele weergegevens zoals temperatuur, windkracht en windrichting. In de figuur hiernaast zie je de windkaart van Nederland op maandag 11 maart 2013 om 20:40 uur. Deze kaart is gebaseerd op gegevens van KNMI-meetstations die over Nederland zijn verspreid. Deze meetstations geven elke 10 minuten een nieuwe waarneming af.

figuur



In Nederland zijn er 53 officiële KNMI-meetstations.

2p **9** Bereken hoeveel waarnemingen er elke dag in totaal door de officiële meetstations aan het KNMI worden doorgegeven.

Als je in de ochtend van huis naar school fietst en in de middag terugfietst, kan de wind invloed hebben op je totale reistijd. Hoe dat zit, onderzoeken we in de rest van deze opgave.

Sylvia woont 10 km van school. Zij fietst elke schooldag. We gaan ervan uit dat als er geen wind is, haar snelheid constant 20 km/u is. Haar totale reistijd is op zo'n schooldag dus 1 uur.
Meestal waait het echter. We veronderstellen dat Sylvia altijd wind mee heeft op de heenweg en wind tegen op de terugweg en dat de wind de hele dag constant is. Dan is Sylvia's snelheid op de heenweg $20+w$ km/u en op de terugweg $20-w$ km/u. Hierbij geldt $0 \leq w < 20$.

Op een dag geldt $w=5$. Sylvia's totale reistijd is die dag langer dan 1 uur.

4p **10** Bereken hoeveel minuten haar totale reistijd die dag langer is dan 1 uur.

Sylvia's totale reistijd T in uren wordt gegeven door de formule:

$$T = \frac{400}{400 - w^2}$$

De formule voor T kan worden gevonden door een formule voor de reistijd voor de heenweg en een formule voor de reistijd voor de terugweg op te stellen en deze formules bij elkaar op te tellen.

5p **11** Stel deze formules op en toon daarmee aan dat de bovenstaande formule voor T juist is.

Op een dag is Sylvia's totale reistijd 1 uur en 20 minuten.

3p **12** Bereken de waarde van w op die dag.

Met de formule voor Sylvia's totale reistijd kun je zonder te rekenen beredeneren dat haar totale reistijd op een dag met wind groter is dan op een dag zonder wind.

3p **13** Geef zo'n redenering.

Dat de totale reistijd toeneemt als w toeneemt, kun je ook aantonen met behulp van de afgeleide van T.

5p **14** Stel een formule op voor de afgeleide van T en toon daarmee aan dat de totale reistijd toeneemt als w toeneemt.

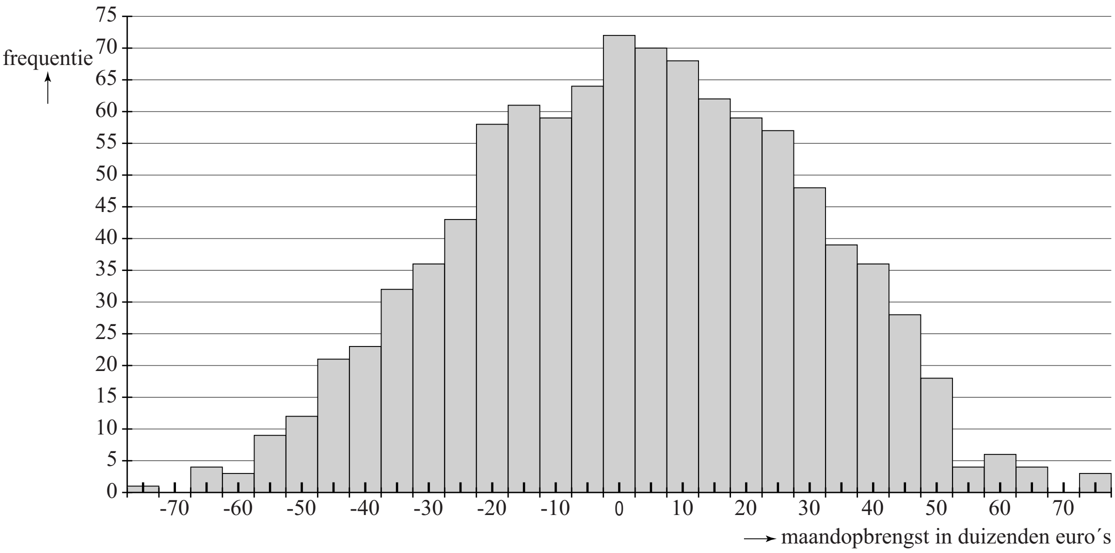
Financieel risico

Value-at-Risk-model

Sinds de financiële crisis van 2009 zijn banken verplicht hun financiële risico's extra in de gaten te houden. Een veelgebruikte manier om financieel risico in te schatten is het Value-at-Risk-model (VaR). Dit is een statistisch model dat het mogelijke verlies op een aandelenportefeuille van een bank berekent.

De maandopbrengsten van 1000 maanden van een aandelenportefeuille zijn verzameld. Deze opbrengsten zijn weergegeven in een frequentieverdeling. Zie de figuur. Langs de horizontale as zijn de maandopbrengsten gezet in duizenden euro's. Uit de figuur kun je bijvoorbeeld afleiden dat er 68 keer een maandwinst is behaald tussen € 7 500 en € 12 500. Ook kun je afleiden dat er 43 keer een verlies was tussen € 22 500 en € 27 500.

figuur



Deze figuur staat vergroot op de uitwerkbijlage. Op basis van de figuur wordt een model gemaakt. In dat model wordt gesteld dat de kans op een verlies de komende maand van € 42 500 of meer gelijk is aan 5%. Op basis van de figuur kan ook een schatting worden gemaakt van de kans dat de volgende maand het verlies € 17 500 of meer zal zijn.

4p **15** Bereken deze schatting in procenten met behulp van de figuur op de uitwerkbijlage.

De € 42 500 uit het voorbeeld wordt de 5%-VaR van de maandopbrengst van deze aandelenportefeuille genoemd: het bedrag dat je in een maand met een kans van 5% ten minste riskeert te verliezen.
De VaR kan ook berekend worden voor andere tijdsperioden dan een maand, bijvoorbeeld een week of 10 dagen en ook voor andere kanspercentages, bijvoorbeeld 1%.

De weekopbrengst van een bepaalde aandelenportefeuille is normaal verdeeld met een gemiddelde van € 752 en een standaardafwijking van € 2500. De 1%-VaR van de weekopbrengst is het verlies dat men met een kans van 1% de komende week ten minste zal lijden.

3p **16** Bereken de 1%-VaR van de weekopbrengst van deze aandelenportefeuille in euro's nauwkeurig.

In het zogenoemde akkoord van Basel staan eisen voor de minimale hoeveelheid kapitaal die banken moeten aanhouden om ervoor te zorgen dat ze niet te snel in financiële problemen komen. Vereenvoudigd kan gesteld worden dat het minimaal vereiste kapitaal gelijk is aan driemaal de 1%-VaR van de tiendaagse opbrengst.

De dagopbrengst van een aandelenportefeuille van een bepaalde bank is normaal verdeeld met een gemiddelde van 380 000 euro en een standaardafwijking van 1,4 miljoen euro.
Aangenomen wordt dat de dagopbrengsten onafhankelijk van elkaar zijn. De som van de dagopbrengsten is in dat geval ook normaal verdeeld.

5p **17** Bereken de minimale hoeveelheid kapitaal die de bank volgens het akkoord van Basel moet aanhouden.

Risico bij leningen
Behalve met risico's bij de aandelenportefeuille heeft een bank ook te maken met het risico dat leningen niet terugbetaald worden. Om dit risico in te schatten worden personen die bij een bank geld geleend hebben verdeeld op grond van bepaalde kenmerken in risicocategorieën.
In de tabel zie je een voorbeeld van zo'n indeling voor een sterk vereenvoudigde situatie.

tabel

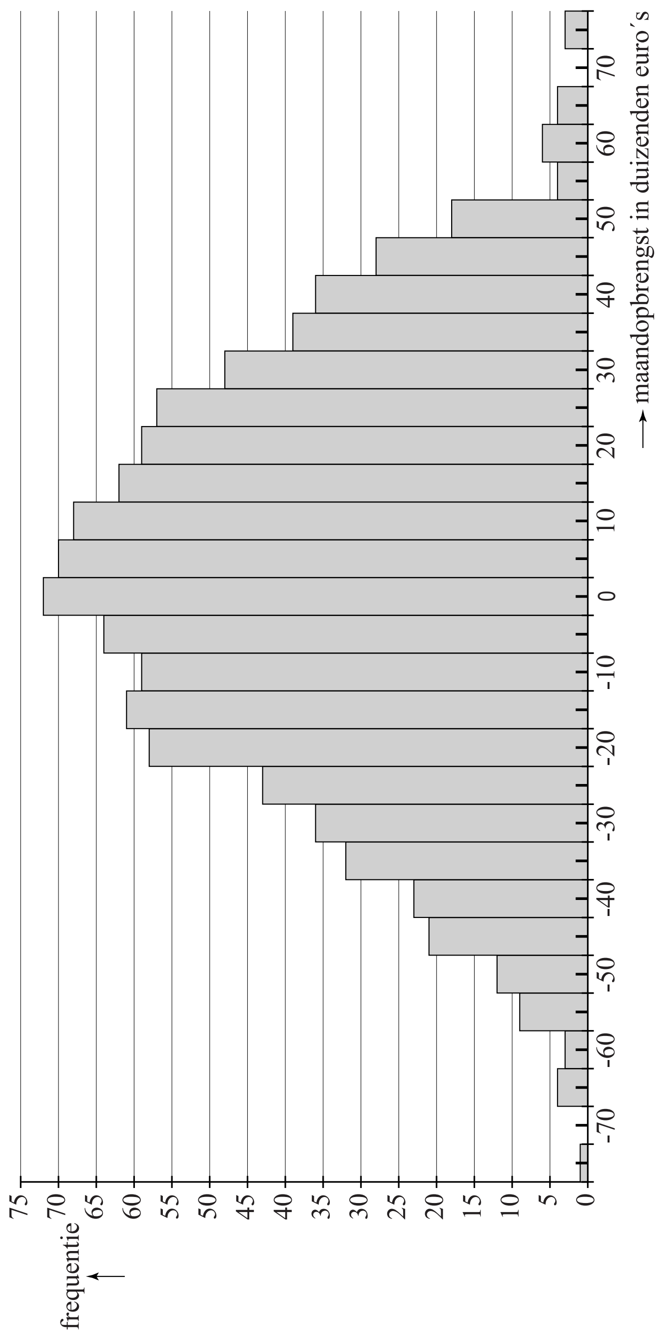
categorie	aantal personen	kans dat de lening terugbetaald wordt
A	850	95%
B	530	80%
C	260	60%

In dit vereenvoudigde model zijn er drie categorieën en gaan we ervan uit dat een lening ofwel helemaal terugbetaald wordt ofwel helemaal niet.

4p **18** Bereken hoe groot de kans volgens het model is dat meer dan de helft van de personen in categorie C zijn lening **niet** terugbetaalt. Geef je antwoord in vier decimalen.

uitwerkbijlage

15



Vreemde dobbelstenen

De investeerder Warren Buffett houdt van dobbelspelletjes met ongebruikelijke dobbelstenen. Hij daagt Bill Gates, de oprichter van Microsoft, uit voor een spelletje waarbij ze allebei een dobbelsteen mogen werpen. Degene met het hoogste ogenaantal wint.

Ze gebruiken drie dobbelstenen: een blauwe, een groene en een rode. De ogenaantallen staan in tabel 1.

tabel 1

blauw	3 3 3 3 3 6
groen	2 2 2 5 5 5
rood	1 4 4 4 4 4

Warren laat Bill als eerste een dobbelsteen kiezen, en nadat Bill de blauwe pakt, kiest Warren de rode dobbelsteen.

3p **19** Bereken de kans dat Warren wint.

Even later spelen Warren en Bill weer tegen elkaar, maar de spelregels zijn veranderd. Er zijn nu twee blauwe, twee groene en twee rode dobbelstenen. Warren kiest twee dobbelstenen van gelijke kleur, waarna Bill twee andere dobbelstenen van gelijke kleur moet kiezen. De winnaar is degene met de hoogste som van zijn ogenaantallen.

Warren begint. Hij kiest de twee rode dobbelstenen. De kansverdeling voor de som van zijn ogenaantallen staat in tabel 2.

tabel 2

som	2	5	8
kans	$\frac{1}{36}$	$\frac{10}{36}$	$\frac{25}{36}$

Bill kiest de twee groene dobbelstenen.

6p **20** Bereken de kans dat Bill wint.