

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  $\bar{X}$  van de uitkomsten  $X$ :

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

$$E(\bar{X}) = E(X) \quad \sigma(\bar{X}) = \frac{\sigma(X)}{\sqrt{n}}$$

**Binomiale verdeling**

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

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

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  $\mu$  en standaardafwijking  $\sigma$  geldt:

$$Z = \frac{X - \mu}{\sigma} \text{ is standaard-normaal verdeeld en } P(X < g) = P\left(Z < \frac{g - \mu}{\sigma}\right)$$

**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	$q(x) = \frac{f(x)}{g(x)}$	$q'(x) = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{(g(x))^2}$
kettingregel	$k(x) = f(g(x))$	$k'(x) = f'(g(x)) \cdot g'(x)$ of $\frac{dk}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx}$

**Logaritmen**

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

## Marathonloopsters

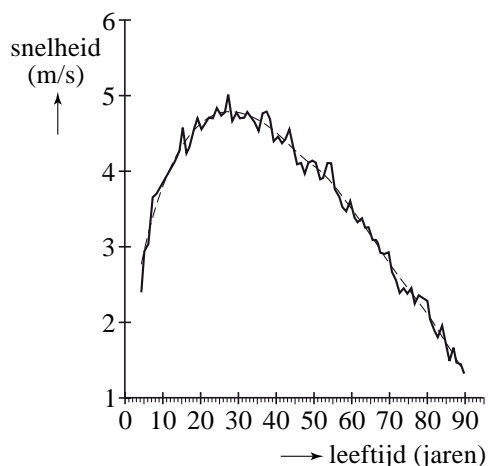
De Olympische hardloopwedstrijd met de grootste lengte is de marathon: ruim 42 kilometer, om precies te zijn 42 195 meter. De marathon wordt zowel door mannen als door vrouwen gelopen. In deze opgave concentreren we ons op de marathonloopsters.

De prestatie van een loopster geeft men in krantenberichten meestal weer door de tijd waarin de marathon is afgelegd, maar een even duidelijke maat is de gemiddelde snelheid over het gehele parcours. Dit noemen we kortweg de **snelheid**. Deze snelheid drukken we uit in m/s (meters per seconde).

- 3p 1 Een marathonloopster legt de marathon af in 2 uur, 43 minuten en 32 seconden. Bereken haar snelheid in m/s.

Elmer Sterken van de Rijksuniversiteit Groningen heeft onderzoek gedaan naar het verband tussen de snelheid van Amerikaanse marathonloopsters en hun leeftijden. Figuur 1 is afkomstig uit het rapport dat hij daarover geschreven heeft. In figuur 1 is voor iedere leeftijd weergegeven de hoogste snelheid ooit gelopen door een Amerikaanse (zie de 'zigzaglijn'). De geregistreerde leeftijden lopen van 6 tot en met 90 jaar<sup>1</sup>.

figuur 1



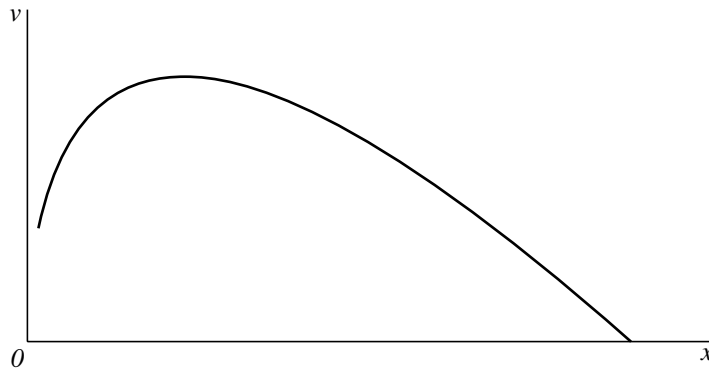
De 'zigzaglijn' is in figuur 1 benaderd door de grafiek met de formule:

$$v = 2,836 \cdot x^{0,665} - 1,390 \cdot x^{0,818}$$

Hierin is  $v$  de hoogste snelheid in m/s van marathonloopsters met een leeftijd van  $x$  jaar. In figuur 2 is de grafiek van  $v$  weergegeven.

noot 1 Figuur 1 is ontstaan door allerlei gegevens van verschillende loopafstanden (op verantwoorde manier) om te zetten naar de marathonlengte. Hierdoor zijn in deze figuur ook voor een marathon onwaarschijnlijk jonge leeftijden vermeld.

**figuur 2**



In het vervolg van deze opgave beschouwen we deze laatste grafiek en de formule voor  $v$  als een wiskundig model van de werkelijkheid.

Petra loopt vaker een marathon en hoopt binnenkort de marathon binnen 3 uur te volbrengen. Petra is 52 jaar.

- 3p **2** Kan een 52-jarige marathonloopster volgens dit model de marathon binnen 3 uur lopen? Licht je antwoord toe.

De grafiek van  $v$  heeft een maximum.

Volgens het model is er dus blijkbaar een leeftijd waarop marathonloopsters (gemiddeld) het beste presteren.

- 5p **3** Stel de afgeleide van  $v$  op en bereken hiermee deze leeftijd.

## Stoppen met roken

Veel mensen beginnen op jonge leeftijd met roken en proberen daar op latere leeftijd weer mee op te houden. Dat lukt niet altijd.

Het Centraal Bureau voor de Statistiek (CBS) publiceert regelmatig cijfers waarmee het rookgedrag van Nederlanders kan worden bestudeerd. In tabel 1 vind je enkele getallen.

**tabel 1**

rokers en aantallen sigaretten

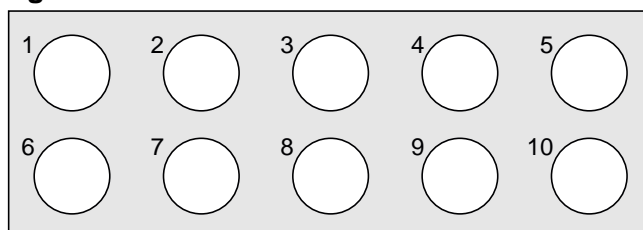
jaar	2001	2005
aantal Nederlanders, in miljoenen	16,0	16,3
percentage rokers	33,3%	29,5%
gemiddeld aantal sigaretten per roker per jaar	4526	4271

- 4p 4 Bereken met hoeveel procent het **totale** aantal gerookte sigaretten in 2005 is afgenomen ten opzichte van 2001.

Er zijn veel hulpmiddelen om minder te gaan roken of er zelfs helemaal mee te stoppen. Eén daarvan is het gebruik van tabletten van het merk Fumostop. Om na te gaan of Fumostop een middel is dat inderdaad helpt, wordt het volgende onderzoek uitgevoerd.

Uit alle zware rokers wordt aselekt een groep van 18 proefpersonen gekozen. Elke proefpersoon krijgt 10 tabletten die uiterlijk niet van elkaar verschillen. De tabletten zijn verpakt in doordrukstrips met bij elk tablet een nummer. Zie figuur 1.

**figuur 1**



Elke proefpersoon moet 10 dagen lang iedere dag bij het opstaan één willekeurig gekozen tablet innemen, het nummer van dat tablet noteren en bijhouden hoeveel sigaretten hij die dag rookt.

Wat de proefpersonen niet weten maar de onderzoekers wel, is dat 5 van de tabletten inderdaad van het merk Fumostop zijn. De andere 5 tabletten bevatten geen enkele werkzame stof. We geven de 'echte' tabletten aan met F en de andere tabletten met NF. Aan de genoteerde tabletnummers kunnen de onderzoekers zien wanneer de F- en de NF-tabletten ingenomen zijn.

Nico is één van de 18 proefpersonen. De mogelijkheid bestaat dat hij om de dag een F-tablet inneemt, waarmee bedoeld wordt dat hij de ene dag een F-tablet en de andere dag een NF-tablet inneemt.

- 4p 5 Bereken de kans dat hij om de dag een F-tablet inneemt.

De proefpersonen kiezen hun tabletten iedere dag dus volledig aselekt. Het kan dus gebeuren dat een proefpersoon de eerste dag een van de tabletten met nummer 1 of nummer 2 kiest.

- 4p **6** Bereken hoe groot de kans is dat 6 of meer proefpersonen op de eerste dag van het onderzoek een van de tabletten met nummer 1 of 2 kiezen.

De onderzoekers vermoeden dat het gebruik van F-tabletten leidt tot het roken van minder sigaretten. Om dat na te gaan, wordt van elke proefpersoon bijgehouden hoeveel sigaretten hij in totaal heeft gerookt op de vijf dagen met een F-tablet en op de vijf dagen met een NF-tablet. Het resultaat vind je in tabel 2.

**tabel 2**

aantal sigaretten

<b>proefpersoon</b>	1	2	3	4	5	6	7	8	9
<b>bij gebruik van F-tabletten</b>	106	90	109	72	103	118	124	103	89
<b>bij gebruik van NF-tabletten</b>	112	108	132	92	96	120	145	129	101

<b>proefpersoon</b>	10	11	12	13	14	15	16	17	18
<b>bij gebruik van F-tabletten</b>	87	92	145	101	100	97	112	104	101
<b>bij gebruik van NF-tabletten</b>	104	127	138	124	121	139	100	93	118

- 6p **7** Onderzoek met behulp van een tekentoets of er voldoende aanleiding is om het vermoeden van de onderzoekers te bevestigen. Neem hierbij als significantieniveau 5%.

Van de mensen die in 2006 rookten, rookte 24,5% per dag 20 sigaretten of meer. Rokers rookten toen gemiddeld 11,4 sigaretten per dag. Tine wil onderzoeken of het aantal sigaretten per dag normaal verdeeld zou kunnen zijn. Ze bedenkt de volgende aanpak: "Als er sprake is van een normale verdeling, dan kan ik de bijbehorende standaardafwijking berekenen. Daarna kan ik nagaan of die waarde – in combinatie met dat gemiddelde 11,4 – tot een conclusie leidt."

- 4p **8** Bereken die standaardafwijking en toon daarmee aan dat het aantal sigaretten dat een roker per dag in 2006 rookte, niet normaal verdeeld kan zijn.

## Boomgroei

Naar de groei van bomen is veel onderzoek gedaan. Dat heeft geleid tot een goed inzicht in het verband tussen de hoogte van een boom en de leeftijd van die boom. In de bosbouw wordt voor veel bomen de te verwachten hoogte berekend met de formule van Chapman-Richards:

$$h = a(1 - b^t)^c$$

Hierin is  $h$  de hoogte van een boom in meters en  $t$  de leeftijd in jaren. De waarden van de getallen  $a$ ,  $b$  en  $c$  hangen af van de soort boom. De getallen  $a$ ,  $b$  en  $c$  zijn positief. In tabel 1 zijn deze waarden voor enkele boomsoorten weergegeven.

tabel 1

boom	$a$	$b$	$c$
Japane lariks	23,743	0,9603	1,22770
zomereik	39,143	0,9867	0,96667
Amerikaanse eik	29,026	0,9790	0,80820
berk	43,281	0,9876	0,95040
grove den	24,426	0,9656	1,59980

Het verband tussen de hoogte en de leeftijd van de zomereik wordt dus gegeven door de formule:

$$h = 39,143(1 - 0,9867^t)^{0,96667}$$

De zomereik wordt op den duur veel groter dan de Amerikaanse eik, maar in de eerste levensjaren groeit de Amerikaanse eik veel sneller.

- 5p **9** Toon door berekening aan dat volgens de formule van Chapman-Richards de Amerikaanse eik in het vierde levensjaar ruim 20 cm méér groeit dan de zomereik.

Elke boom groeit, maar die groei gaat steeds langzamer. We kunnen dit nagaan met behulp van de afgeleide van  $h$ . Dat gaan we hier doen voor de zomereik. Voor de snelheid waarmee deze boomsoort groeit, kijken we naar de formule voor de afgeleide:

$$h' = \frac{0,507 \cdot 0,9867^t}{(1 - 0,9867^t)^{0,03333}}$$

Aan de vorm van deze afgeleide kun je door redeneren, dus **zonder** berekeningen uit te voeren, zien dat de zomereik steeds groeit en dat die groei steeds langzamer gaat.

- 6p **10** Laat zien (door een redenering) hoe uit de formule van de afgeleide volgt dat de zomereik steeds groeit en dat deze groei steeds langzamer gaat.

Voor de formule voor de zomereik hebben we gebruik gemaakt van de waarden van  $a$ ,  $b$  en  $c$  uit tabel 1. Maar niet alle zomereiken hebben de waarde 39,143 voor  $a$ .

Factoren zoals klimaat en bodemgesteldheid beïnvloeden de waarde van  $a$ . Chapman en Richards gaan er in hun model van uit dat de waarden van  $b$  en  $c$  uitsluitend afhangen van de boomsoort.

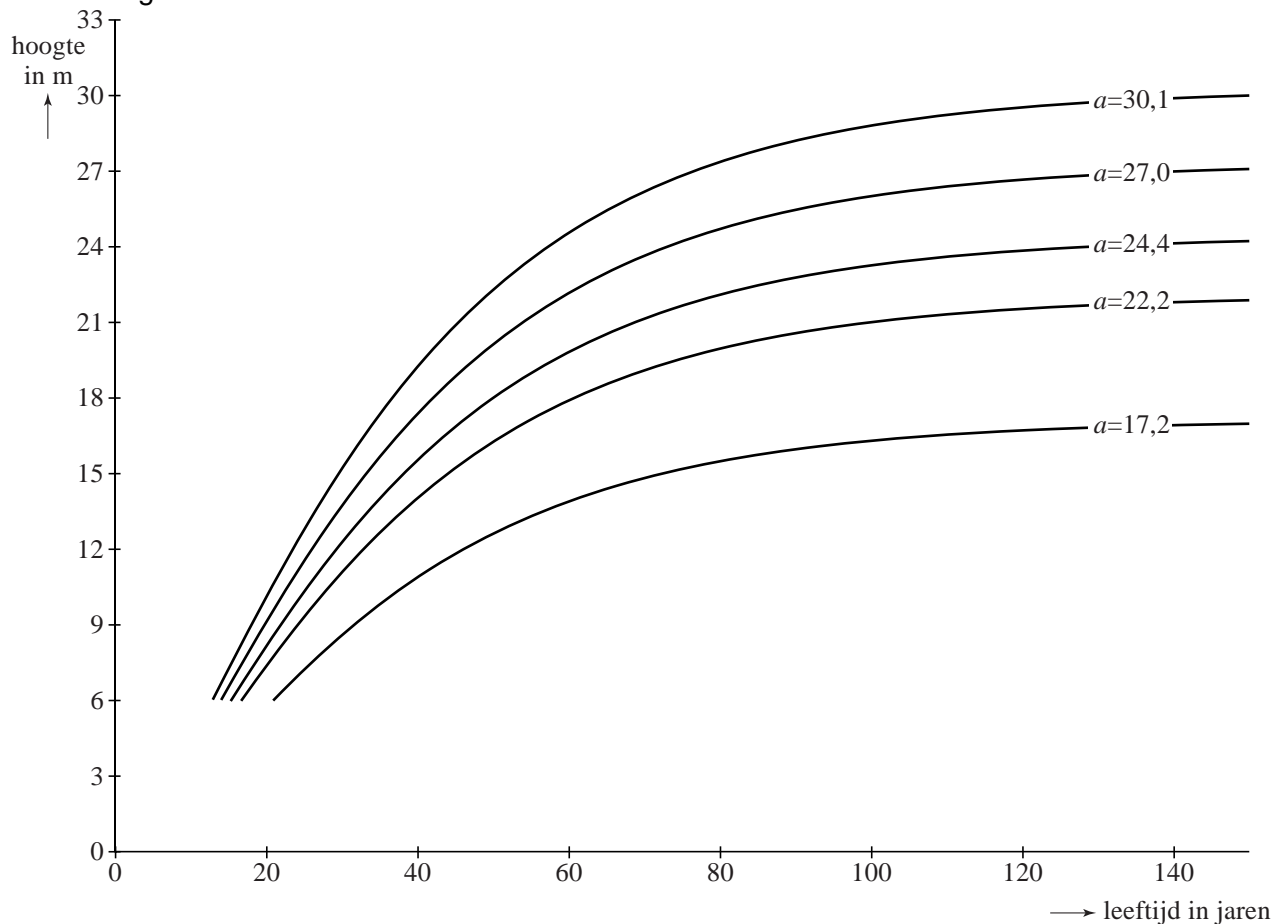
Vaak weet men niet van tevoren welke waarde van  $a$  een boom heeft. Om de waarde van  $a$  voor een boom te bepalen, laat men de boom eerst een aantal jaren groeien.

Daarna meet men de boom op en berekent men welke waarde van  $a$  past bij de groei van die boom. Men gaat ervan uit dat die waarde van  $a$  daarna niet meer verandert.

- 3p **11** Een zomereik bereikt op de leeftijd van 10 jaar een hoogte van 6,18 meter. Bereken de waarde van  $a$  die hierbij hoort.

Afhankelijk van de waarde van  $a$  krijgen we verschillende groeiformules. In figuur 1 zie je de grafieken van enkele groeiformules van de grove den. De waarde van  $a$  staat er steeds bij vermeld.

**figuur 1**  
grove den



Het ziet er naar uit dat je aan de waarde van  $a$  kunt zien hoe groot de verschillende grove dennen uiteindelijk worden.

- 4p **12** Onderzoek aan de hand van de formule of dit inderdaad het geval is voor de grafiek die hoort bij  $a = 30,1$ .

Als je naar figuur 1 kijkt, kun je je afvragen of deze grafieken door de oorsprong  $(0, 0)$  gaan als we ze verder naar links zouden doortekenen. Dit is inderdaad het geval.

Sterker nog: dit is het geval voor **alle** grafieken die horen bij de algemene formule  $h = a(1 - b^t)^c$  van Chapman-Richards.

- 4p **13** Beredeneer, dus zonder getallenvoorbeelden te gebruiken, dat **alle** grafieken die horen bij de formule van Chapman-Richards door de oorsprong gaan.



## Inkomen

Het CBS (Centraal Bureau voor de Statistiek) besteedt elk jaar aandacht aan de verdeling van de inkomens van huishoudens in Nederland. In tabel 1 is voor het jaar 2004 weergegeven hoeveel huishoudens in een bepaalde inkomensklasse zaten.

**tabel 1**

besteedbaar inkomen in euro's	aantal huishoudens in duizendtallen
0 – 10 000	490
10 000 – 20 000	2057
20 000 – 30 000	1777
30 000 – 40 000	1309
40 000 – 50 000	687
50 000 – 70 000	460
meer dan 70 000	197

Met behulp van lineair interpoleren kun je een schatting maken van het percentage huishoudens met een besteedbaar inkomen van ten hoogste 27 000 euro.

- 5p **14** Schat dat percentage huishoudens met behulp van lineair interpoleren.

De verdeling van de inkomens is geen normale verdeling, zelfs niet bij benadering. Dat kun je bijvoorbeeld zien door het histogram<sup>1)</sup> te tekenen voor de inkomensklassen 0 – 10 000 tot en met 40 000 – 50 000.

- 4p **15** Teken dit histogram en leg met behulp daarvan uit dat deze frequentieverdeling niet kan worden benaderd met een normale verdeling.

Toch is er wel een manier om tabel 1 in verband te brengen met een normale verdeling. Dat gaat als volgt:

We geven de inkomens aan met  $X$  en berekenen van alle inkomens de logaritme. Die noemen we  $L$ , dus  $L = \log(X)$ . We krijgen dan voor  $L$  een klassenindeling met andere grenzen. Omdat  $\log(0)$  niet bestaat, nemen we bij de eerste klasse als linkergrens  $\log(1)$ .

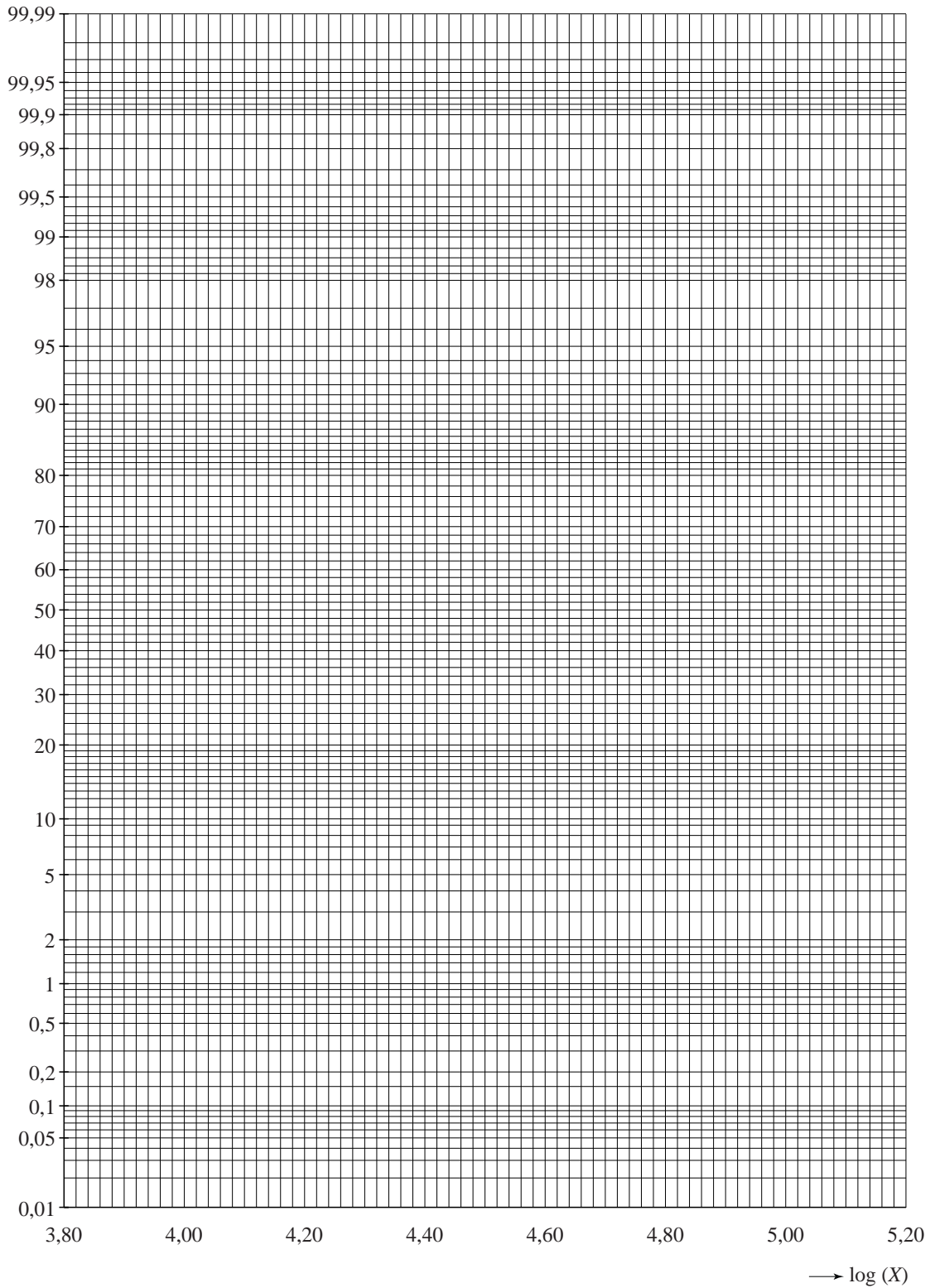
Deze klassenindeling levert wel (bij benadering) een normale verdeling op.

- 6p **16** Maak de frequentieverdeling die bij  $L$  hoort en toon met behulp van het normaal waarschijnlijkheidspapier op de uitwerkbijlage aan dat deze frequentieverdeling bij benadering normaal verdeeld is.

noot 1 Een histogram wordt ook wel staafdiagram genoemd.

**uitwerkbijlage**

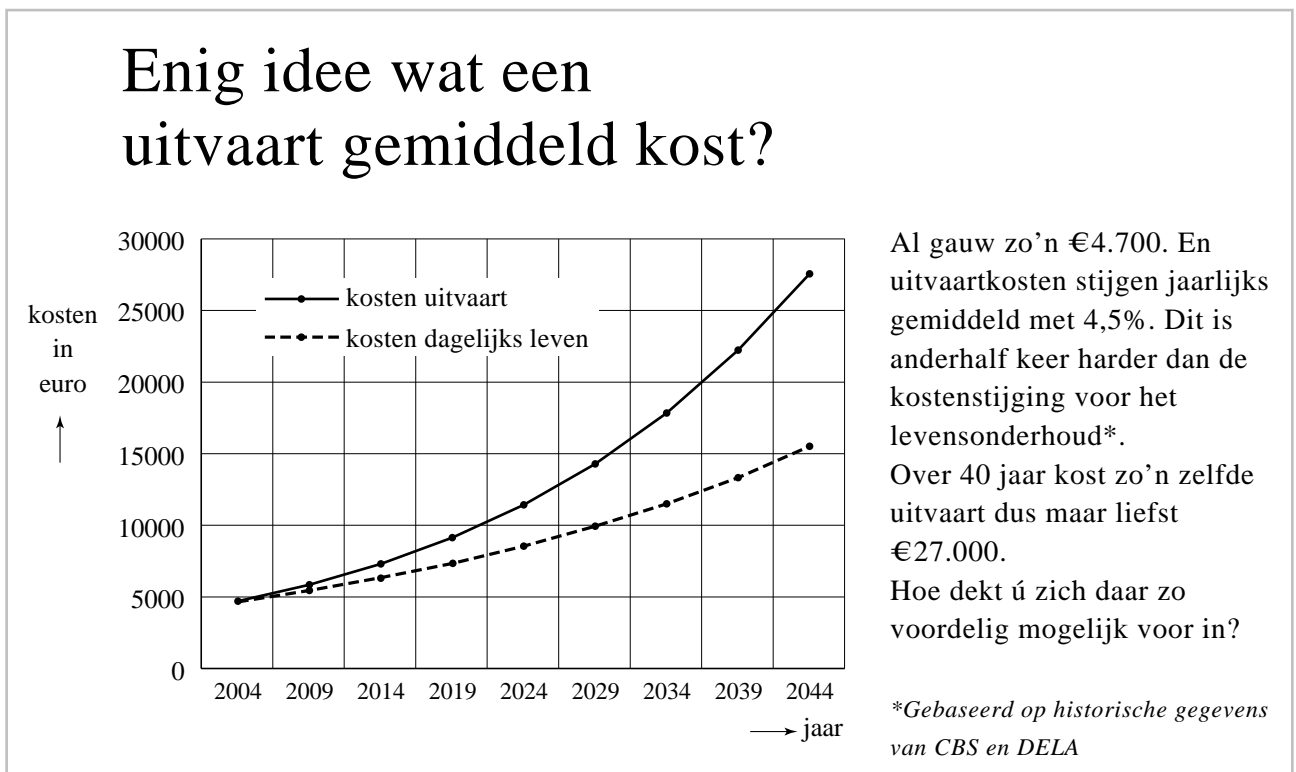
**16** Normaal waarschijnlijkheidspapier



## Verzekering

Verzekeringsmaatschappijen maken op verschillende manieren reclame voor allerlei verzekeringen. Een voorbeeld daarvan vind je in figuur 1 hieronder. Daar zie je een deel van een reclamefolder die in 2004 huis aan huis werd verspreid. In de folder legt de verzekeraar uit dat de kosten voor een uitvaart sneller stijgen dan de kosten voor het levensonderhoud. Ook wordt de ontwikkeling van beide kostensoorten met een grafiek in beeld gebracht.

figuur 1



Uitgaande van een jaarlijkse kostenstijging met 4,5% berekende men de kosten in 2044. De uitvaartkosten stijgen van € 4700 in 2004 tot ongeveer € 27 000 in 2044.

Het bedrag in 2044 is afgerond op duizendtallen.

3p **17** Bereken dit bedrag in euro's nauwkeurig.

Met "anderhalf keer harder" bedoelt de schrijver van de folder dat de jaarlijkse procentuele stijging van de kosten voor een uitvaart 1,5 keer zo groot is als die van de kosten voor het levensonderhoud. Daardoor zullen de kosten voor het levensonderhoud in de periode 2004–2044 stijgen met een percentage dat aanzienlijk kleiner is dan 474% (het stijgingspercentage van de uitvaartkosten). Dit is in de folder ook grafisch weergegeven.

3p **18** Bereken met hoeveel procent de kosten voor het levensonderhoud volgens de folder zullen toenemen in de periode 2004–2044.

Op de achterzijde van de folder doet de verzekeraar een aanbod om de kosten van een uitvaart te verzekeren. Wie aan deze verzekering deelneemt, betaalt maandelijks een premie. De hoogte daarvan hangt af van de leeftijd van de verzekeringnemer op het moment dat de verzekering wordt afgesloten. De verzekeraar bepaalt die hoogte onder andere op grond van de gemiddelde leeftijd van overlijden in Nederland. Voor iemand die volgens deze inschatting van de verzekeraar gemiddeld nog 40 jaar lang premie zal betalen, bedraagt de premie € 4,79 per maand.

De verzekeraar gaat deze premiegelden beleggen. De jaarlijkse procentuele stijging op zijn beleggingen wordt het **jaarlijkse rendement** van de beleggingen genoemd.

De premies en de winst van de beleggingen kunnen samen uitgroeien tot een bedrag waarmee de uitvaart kan worden betaald.

De verzekeraar maakt bij zijn berekeningen gebruik van de volgende formule:

$$\text{Opbrengst} = b \cdot \frac{r^n - 1}{r - 1}$$

In deze formule is  $b$  de maandelijkse premie,  $r$  de groeifactor van de beleggingen per maand en  $n$  het aantal maanden dat er (naar verwachting) premie moet worden betaald.

Uit de folder blijkt dat de verzekeraar ervan uitgaat dat bij een maandelijkse premie van € 4,79 de opbrengst na 40 jaar minstens € 27 000 zal zijn.

6p **19** Onderzoek hoeveel het **jaarlijkse rendement** op zijn beleggingen dan moet zijn.

Als de maandelijkse premie verhoogd wordt of als er gedurende meer maanden betaald wordt, zal de opbrengst natuurlijk toenemen.

Voor de formule  $\text{Opbrengst} = b \cdot \frac{r^n - 1}{r - 1}$  geldt dan ook:

I Als  $r$  en  $n$  gelijk blijven en  $b$  toeneemt, dan neemt  $b \cdot \frac{r^n - 1}{r - 1}$  toe.

en

II Als  $b$  en  $r$  gelijk blijven en  $n$  toeneemt, dan neemt  $b \cdot \frac{r^n - 1}{r - 1}$  toe.

We gaan ervan uit dat altijd geldt:  $r > 1$ .

De juistheid van de beweringen I en II kun je beredeneren door uitsluitend

gebruik te maken van de formule  $\text{Opbrengst} = b \cdot \frac{r^n - 1}{r - 1}$ , dus **zonder** te rekenen

of te differentiëren.

4p **20** Toon door redeneren met behulp van de formule de juistheid van de beweringen I en II aan.