5.4 Normaal of niet?

Inleiding

Als je de gewichten van zo'n 1000 aselect gekozen vrouwen uit de leeftijdsgroep van $20- < 30$ jaar meet, krijg je ongeveer een klokvormige frequentieverdeling. Maar mag je nu zeggen dat er sprake is van een normale verdeling?
Er bestaat normaal waarschijnlijkheidspapier. Op dat papier wordt een cumulatief relatief frequentiepolygoon een rechte lijn als er van een normale verdeling sprake is. Met behulp van de vuistregels kun je vanaf dat papier dan het gemiddelde en de standaardafwijking aflezen.

Figuur 1

Je leert in dit onderwerp

- een cumulatief relatief frequentiepolygoon tekenen op normaal waarschijnlijkheidspapier om te onderzoeken of de gegevens normaal zijn verdeeld;
- verwachtingswaarde of standaarddeviatie van een normaal verdeelde variabele berekenen;
- kansen berekenen bij som/verschil van twee normaal verdeelde variabelen.

Voorkennis

- werken met normale verdelingen;
- kansen berekenen bij normale verdelingen;
- grenswaarden terugzoeken bij normale kansen.

Verkennen

Opgave V1

Bij een landelijk onderzoek zijn de gewichten bepaald van 1000 aselect gekozen volwassen vrouwen van $20- < 30$ jaar, zie tabel. De frequentieverdeling lijkt op die van een normale verdeling. Kun je nu zonder meer een normale verdeling als rekenmodel gebruiken voor deze groep vrouwen?

a Ga na of deze frequentieverdeling voldoet aan de vuistregels voor een normale verdeling.

b Is het voldoen aan de vuistregels voldoende reden om te concluderen dat een normale verdeling een bruikbaar rekenmodel is?

gewicht	frequentie
35-<40	10
40-<45	15
45-<50	25
50-<55	75
55-<60	75
60-<65	125
65-<70	150
70-<75	175
75-<80	150
80-<85	100
85-<90	50
90-<95	25
95-<100	15
100-<105	10

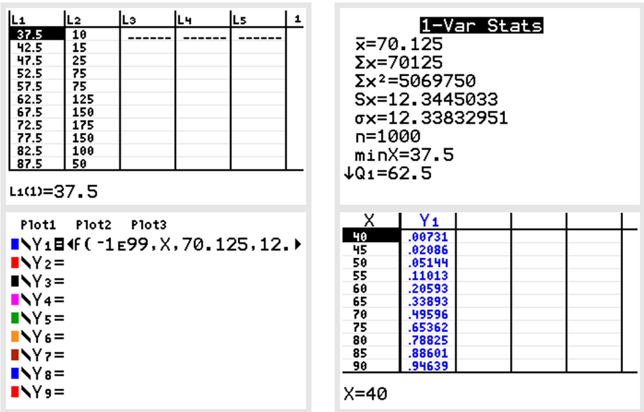
Figuur 2

Uitleg

Bij een landelijk onderzoek zijn de gewichten bepaald van 1000 aselect gekozen volwassen vrouwen van 20− < 30 jaar, zie tabel. De frequentieverdeling lijkt op die van een normale verdeling met $\mu(L) = 70{,}125$ en $\sigma(L) = 12{,}338$. Kun je nu zonder meer een normale verdeling als rekenmodel gebruiken voor deze groep vrouwen?

gewicht	frequentie
35-<40	10
40-<45	15
45-<50	25
50-<55	75
55-<60	75
60-<65	125
65-<70	150
70-<75	175
75-<80	150
80-<85	100
85-<90	50
90-<95	25
95-<100	15
100-<105	10

Figuur 3



Figuur 4

Zet je op **normaal waarschijnlijkheidspapier** kansen van de vorm $P(L \leq 40 | \mu = 70{,}125 \text{ en } \sigma = 12{,}338)$ uit, dan krijg je een rechte lijn. Elke zuivere cumulatieve normale verdeling wordt op normaal waarschijnlijkheidspapier een rechte lijn.

Maak je van de gegeven frequentieverdeling een cumulatieve relatieve frequentieverdeling en zet je die uit op normaal waarschijnlijkheidspapier, dan zou je een rechte lijn moeten krijgen (je zet de cumulatieve relatieve frequenties uit tegen de rechter klassengrenzen). Je zult zien, dat de cumulatieve relatieve frequentieverdeling en de normale verdeling goed overeenkomen! Kennelijk zijn deze gewichten ongeveer normaal verdeeld!

Opgave 1

Bestudeer de **Uitleg**. Neem (print) een aantal bladen normaal waarschijnlijkheidspapier. Ga uit van een normale verdeling met $\mu = 70{,}125$ en $\sigma = 12{,}338$.

a Teken op normaal waarschijnlijkheidspapier de waarden van $P(L \leq g)$ bij deze normale verdeling voor $g = 40, 45, 50, \ldots, 105$. Denk er om dat alle kansen als percentages moeten worden gegeven.

b Wordt je grafiek een rechte lijn?

c Waar in je figuur vind je μ terug? Kun je ook σ terugvinden?

Opgave 2

Neem nu de tabel met de werkelijke gewichten van de 1000 vrouwen.

a Maak hierbij een tabel met cumulatieve relatieve frequenties.

b Zet deze cumulatieve relatieve frequenties uit tegen de bovengrenzen van elke klasse op het normaal waarschijnlijkheidspapier waar de normale verdeling van de vorige opgave op staat.

c Verschilt je grafiek veel van de grafiek van de normale verdeling? Waarom moet je de bovengrenzen van de klassen gebruiken?

d Kun je concluderen dat de gewichten van deze 1000 vrouwen normaal zijn verdeeld?

Theorie en voorbeelden

Om te onthouden

Zet je bij een normaal verdeelde stochast X met verwachting $\mu(X)$ en standaardafwijking $\sigma(X)$ op **normaal waarschijnlijkheidspapier** kansen van de vorm $\text{P}(X \leq g)$ uit tegen g, dan krijg je een rechte lijn.

Maak je van een gegeven frequentieverdeling een cumulatieve relatieve frequentieverdeling en zet je die uit op normaal-waarschijnlijkheidspapier, dan zou je een rechte lijn moeten krijgen als de frequenties normaal zijn verdeeld. De cumulatieve relatieve frequenties moeten tegen de **boven-grenzen** van de klassen worden uitgezet.

Vaak liggen op het **normaal waarschijnlijkheidspapier** de punten van de cumulatieve relatieve frequentieverdeling niet precies op een rechte lijn. Trek dan een rechte lijn die zo goed mogelijk bij de getekende punten past. Je benadert op die manier de frequentieverdeling door de normale kansverdeling die bij die lijn hoort.
Schat de verwachtingswaarde door af te lezen welk getal er bij 50% hoort.
Omdat één van de twee vuistregels zegt dat bij een normale verdeling 68% in het interval $[\mu - \sigma, \mu + \sigma]$ ligt, is bij 84% de waarde van $\mu + \sigma$ af te lezen. Bepaal zo σ.

Is van twee verschillende stochasten X en Y bekend dat ze normaal verdeeld zijn met $\mu(X) \neq \mu(Y)$ en/of $\sigma(X) \neq \sigma(Y)$ en moeten beide stochasten met elkaar vergeleken worden, dan kun je beide stochasten **standaardiseren**.
Elke normaal verdeelde stochast X is om te zetten naar de standaardnormaal verdeelde stochast Z met: $z = \frac{x - \mu}{\sigma}$. Dit heet de **z-waarde**.

Met behulp van z-waarden van twee verschillende stochasten zijn deze stochasten met elkaar te vergelijken.

Voorbeeld 1

De tabel geeft de diameters (in mm) van machinaal geproduceerde moeren. Ga na, dat deze diameters normaal zijn verdeeld en bereken het gemiddelde en de standaardafwijking.

diameter	percentage
12,8-<12,9	0,1
12,9-<13,0	2,1
13,0-<13,1	13,6
13,1-<13,2	34,1
13,2-<13,3	34,0
13,3-<13,4	13,6
13,4-<13,5	2,0
13,5-<13,6	0,1

Figuur 5

Antwoord

Met de GR vind je $\mu(M) = 13{,}20$ en $\sigma(M) = 0{,}10$ waarin M de diameter van een moer voorstelt.

Op normaal waarschijnlijkheidspapier verschillen de cumulatieve relatieve frequentieverdeling vanuit de tabel en die gemaakt vanuit een normale verdeling met $\mu(M) = 13{,}20$ en $\sigma(M) = 0{,}10$ vrijwel niet van elkaar. (Maak beide op een blad normaal waarschijnlijkheidspapier.)

Conclusie: M is normaal verdeeld met $\mu(M) = 13{,}20$ en $\sigma(M) = 0{,}10$.

Opgave 3

In **Voorbeeld 1** zie je een tabel met diameters van machinaal geproduceerde moeren.

a Reken zelf het gemiddelde en de standaarddeviatie na.

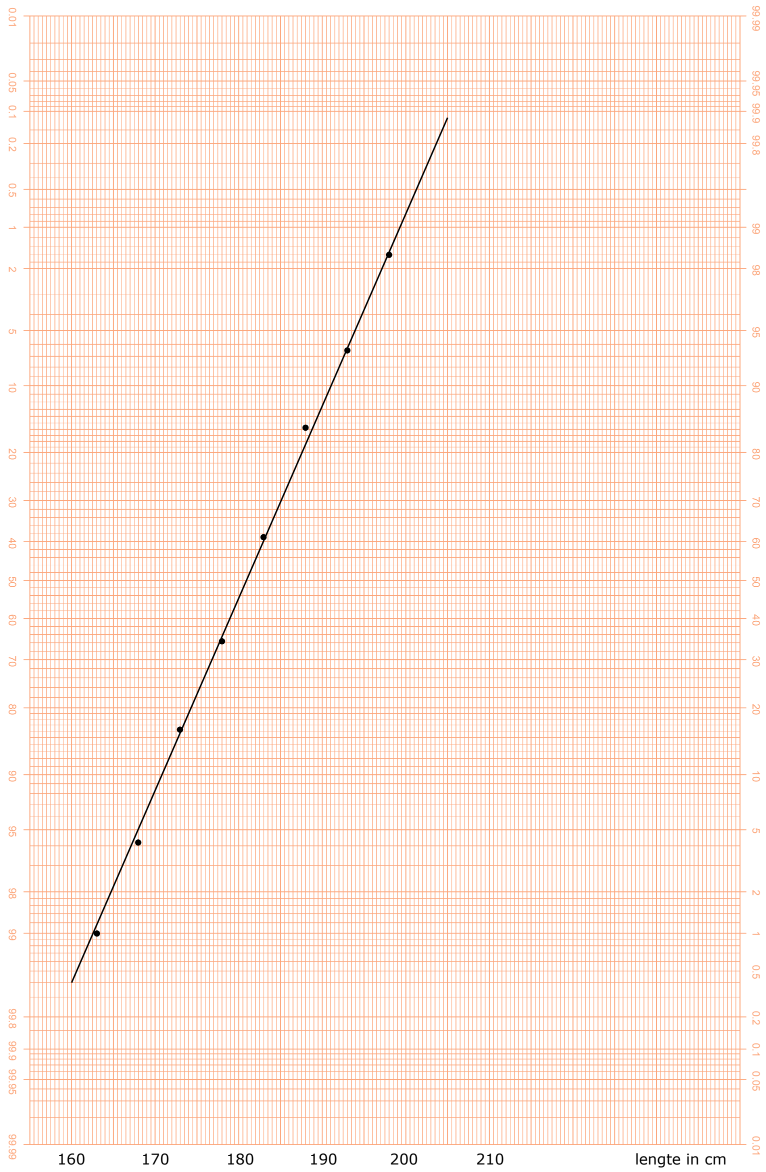
b Teken op normaal waarschijnlijkheidspapier de cumulatieve normale verdeling bij dit gemiddelde en deze standaardafwijking.

c Teken op hetzelfde papier de cumulatieve relatieve frequentieverdeling van de moeren.

d Ga na, dat beide redelijk goed overeen komen. Welke conclusie kun je trekken?

Voorbeeld 2

De lengteverdeling van Nederlandse mannen boven 20 jaar is bij benadering klokvormig. In deze figuur op normaal waarschijnlijkheidspapier zie je hoe deze verdeling wordt benaderd door een rechte lijn.



Figuur 6

Bepaal vanuit deze figuur het gemiddelde en de standaardafwijking van de lengte X van de Nederlandse man in cm.

Antwoord

Op deze lengteverdeling van Nederlandse mannen boven 20 jaar zijn de lengtes bij 50% en 84% af te lezen:

- bij 50% zit de gemiddelde lengte van $\mu \approx 181$ cm;
- bij 84% zit volgens de vuistregels $\mu + \sigma \approx 189$ cm.

Je vindt een gemiddelde van ongeveer 181 cm met een standaardafwijking van $189 - 181 = 8$ cm.

Opgave 4

Wanneer een cumulatieve relatieve frequentieverdeling op normaal waarschijnlijkheidspapier vrijwel een rechte lijn oplevert is er sprake van een normale verdeling. In **Voorbeeld 2** kun je zien hoe je dan het gemiddelde en de standaardafwijking van de verdeling kunt aflezen uit de figuur.
Ga zelf na dat de in het voorbeeld vermelde waarden inderdaad correct zijn.

Opgave 5

In een fabriek worden kilopakken suiker machinaal gevuld. Volgens de Europese norm mag niet meer dan 2,5% van de pakken suiker minder dan 1000 gram bevatten. Open het bestand **de vulgewichten van 100 pakken suiker**.

a Maak een tabel met cumulatieve relatieve frequenties van deze vulgewichten. Gebruik klassen met een klassenbreedte van 1 gram.

b Bereken het gemiddelde en de standaarddeviatie van deze vulgewichten in één decimaal nauwkeurig.

c Teken op normaal waarschijnlijkheidspapier de cumulatieve relatieve frequentieverdeling.

d Zijn de vulgewichten (bij goede benadering) normaal verdeeld? Zo ja, trek dan de rechte lijn die hoort bij deze normale verdeling.

e Laat zien dat het gemiddelde vulgewicht en de bijbehorende standaarddeviatie die je uit de figuur afleest overeen komen met de berekende waarden.

f Welke vulgewichten hebben de 10% zwaarste pakken suiker? Lees je antwoord uit de figuur af.

Voorbeeld 3

In **Voorbeeld 1** zag je dat de diameters M van een bepaalde soort moeren normaal zijn verdeeld met $\mu(M) = 13{,}20$ en $\sigma(M) = 0{,}10$ mm.
Bij deze moeren horen bouten waarvan de diameters B ook normaal zijn verdeeld: $\mu(B) = 13{,}05$ en $\sigma(B) = 0{,}10$ mm.
Deze bouten passen nog in de moeren als hun diameters maximaal 0,25 mm kleiner zijn dan die van de moeren.
Hoeveel procent van de bouten past niet?

Figuur 7

Antwoord

Kijk naar het verschil $V = M - B$ van de diameters van een bout en een moer. Dit verschil is ook normaal verdeeld met

- $\mu(V) = \mu(M - B) = \mu(M + \text{-}B) =$
  $= \mu(M) + \mu(\text{-}B) = \mu(M) - \mu(B) = 13{,}20 - 13{,}05 = 0{,}15$ mm.
- $\sigma(V) = \sigma(M - B) = \sigma(M + \text{-}B) = \sqrt{(\sigma(M))^2 + (\sigma(\text{-}B))^2}$
  $= \sqrt{(\sigma(M))^2 + (\sigma(B))^2} = \sqrt{(0{,}10)^2 + (0{,}10)^2} \approx 0{,}14$ mm.

De bouten passen als $0 \leq V \leq 0{,}25$.
De kans hierop is $\text{P}(0 \leq V \leq 0{,}25 | \mu = 0{,}15 \text{ en } \sigma = 0{,}14) \approx 0{,}62$.
Conclusie: 38% van de bouten past niet in de moeren.

Opgave 6

In **Voorbeeld 3** vind je de gegevens van machinaal geproduceerde moeren en bijbehorende bouten.

a Het verschil van de diameters van de bouten en de moeren is ook normaal verdeeld. Welk gemiddelde en welke standaarddeviatie horen daar bij?

b Reken zelf het percentage bouten na dat niet in de moeren past.

Het bedrijf dat deze bevestigingsmiddelen maakt, produceert nog veel meer bouten en moeren in diverse afmetingen. Voor een ander type bout met bijpassende moer geldt: de diameter van de moer is normaal verdeeld met een gemiddelde van 8,10 mm en een standaarddeviatie van 0,05 mm en de diameter van de bout is normaal verdeeld met een gemiddelde van 8,05 mm en een standaardafwijking van 0,03 mm. De bouten passen in de moeren als het verschil van de diameter van de moer en de bout minder dan 0,02 mm is.

c Hoeveel procent van deze bouten is te dik voor de moeren?

d Hoeveel procent van deze bouten past niet in de bijbehorende moer?

Ook de gewichten van de bouten en moeren van deze tweede soort zijn normaal verdeeld: het gemiddelde gewicht van de moer is 5,0 gram met een standaardafwijking van 0,2 gram en het gemiddelde gewicht van de bout is 7,3 gram met een standaardafwijking van 0,3 gram. Ze worden verpakt in dozen. In elke doos zitten 100 bouten en moeren.

e Hoe zwaar zijn de bouten en moeren in zo'n doos gemiddeld? En welke standaardafwijking hoort daar bij?

f Hoeveel procent van deze dozen heeft een totale inhoud van meer dan 1235 gram?

Verwerken

Opgave 7

Deze tabel geeft de bloeddruk in mmHg (millimeter kwikdruk) van een groep mannen en een groep vrouwen.

bloeddruk	mannen	vrouwen
105	2	1
110	4	3
115	6	5
120	16	15
125	15	12
130	6	6
135	7	7
140	7	7
145	7	8
150	2	5
155	1	3
160	1	2
165	15	1

Tabel 1

a Bereken van beide groepen de gemiddelde bloeddruk en de standaardafwijking van de bloeddruk.

b Welke klassenindeling is hier gehanteerd?

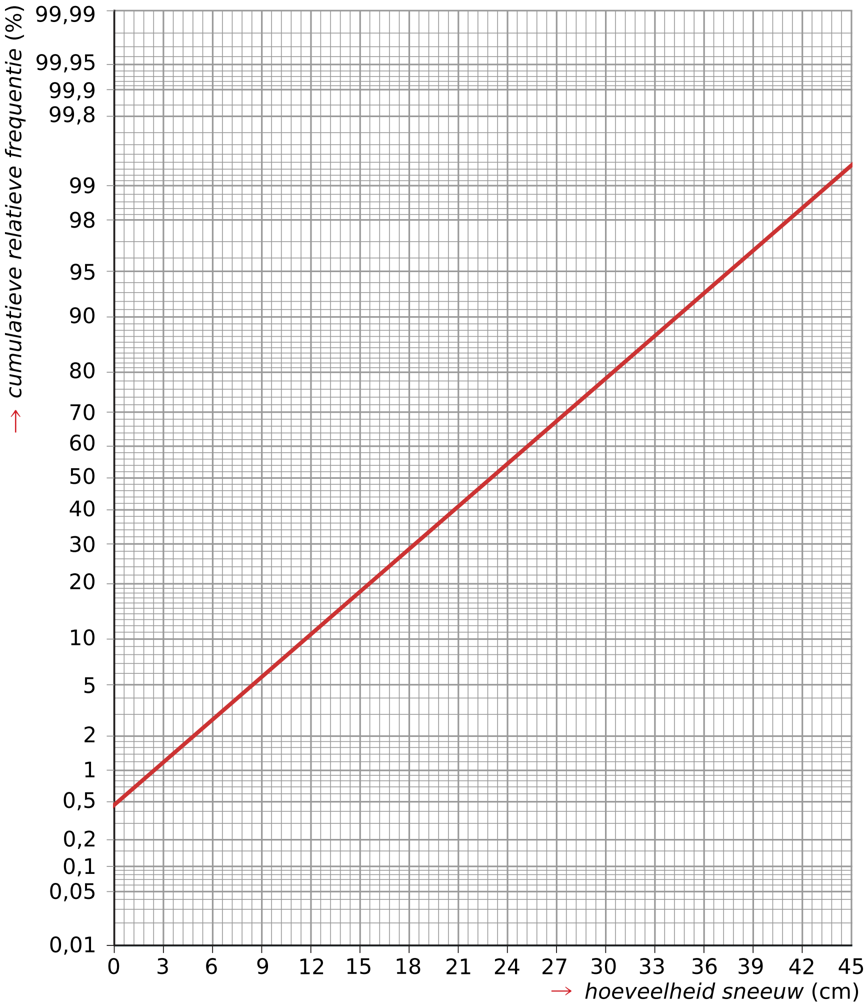
c Laat met behulp van normaal waarschijnlijkheidspapier zien dat de bloeddruk van de mannen niet normaal is verdeeld.

d Je kunt wel een rechte lijn trekken die de verdeling zo goed mogelijk benaderd. Doe dat en lees het gemiddelde en de standaardafwijking af die bij die lijn passen. Wijken de waarden veel af van de berekende waarden?

e Is de bloeddruk van de vrouwen wel normaal verdeeld?

Opgave 8

In een noordelijk gelegen stad sneeuwt het regelmatig in het winterseizoen.
Van de hoeveelheid sneeuw die afgelopen winter per dag dat het sneeuwde viel, is een grafiek op normaal-waarschijnlijkheidspapier gemaakt.



Figuur 8

a Leg uit hoe je uit deze grafiek kunt aflezen hoeveel procent van de dagen dat het sneeuwde, er meer dan 18 cm sneeuw per dag viel.

b Leg uit waarom de hoeveelheid sneeuw die per sneeuwdag viel, normaal verdeeld is en bepaal de gemiddelde hoeveelheid sneeuw die per sneeuwdag viel. Bepaal ook de bijbehorende standaardafwijking.

Opgave 9

Open het bestand **Enkele lichaamsmetingen van 5001 vrouwen uit 1947**. Hierin zie je een tabel met lichaamslengtes in cm van de 5001 vrouwen uit het onderzoek in 1947 van Freudenthal en Sittig in opdracht van De Bijenkorf.

a Bereken de gemiddelde lichaamslengte en de standaarddeviatie.

b Teken op normaal waarschijnlijkheidspapier de bijbehorende cumulatieve relatieve frequentieverdeling.

c Zijn de lichaamslengtes bij benadering normaal verdeeld?

d 95% van de lichaamslengtes zit tussen $\mu - a$ en $\mu + a$. Hoe groot is a? Lees je antwoord uit de figuur af.

e Welke minimale lengte hebben de 16% grootste lichaamslengtes? Lees je antwoord uit de figuur af.

Opgave 10

Limburgse kaas wordt verkocht in pakjes van 200 g. De snijmachine is zo afgesteld dat het gewicht van de pakjes normaal verdeeld is met een gemiddelde van 202,5 g en een standaardafwijking van 4,0 g. Er gaan 50 pakjes in een doos. Die dozen hebben een gemiddeld gewicht van 145,0 gram met een standaardafwijking van 5,5 gram.

a Hoeveel bedraagt het gemiddelde gewicht van zo'n doos met 50 pakjes kaas? En welke standaardafwijking hoort daar bij?

b Hoeveel procent van deze dozen met kaas weegt minder dan 10250 gram?

Toepassen

Opgave 11: Voedselzoekende vogels

Vogels die hun voedsel in bomen en struiken zoeken, doen dat vaak bij voorkeur op een specifieke hoogte. Gedurende een winter zijn in een bos voedselzoekende vogels geobserveerd. In de tabel staat de verdeling over verschillende hoogtes van 400 waarnemingen bij pimpelmezen.

hoogte in meter	$< 1{,}5$	$1{,}5- < 3$	$3- < 5$	$5- < 7$	$7- < 10$	$10- < 15$	> 15
aantal waarnemingen	24	26	51	72	122	92	13

Tabel 2

Toon aan dat de waargenomen hoogtes bij benadering normaal verdeeld zijn. Gebruik hiervoor normaal waarschijnlijkheidspapier. Lees uit je tekening af hoe groot het gemiddelde en de standaardafwijking van deze verdeling zijn. Geef beide antwoorden tot op dm nauwkeurig. Licht je werkwijze toe.

(bron: examen vwo wiskunde A1,2 2002, eerste tijdvak)

Testen

Opgave 12

Open het bestand **Enkele lichaamsmetingen van 5001 vrouwen uit 1947**. Hierin zie je een tabel met kniehoogtes in cm van de 5001 vrouwen uit het onderzoek in 1947 van Freudenthal en Sittig in opdracht van De Bijenkorf.

a Bereken de gemiddelde kniehoogte en de standaarddeviatie.

b Teken op normaal waarschijnlijkheidspapier de bijbehorende cumulatieve relatieve frequentieverdeling.

c Ga na, dat de gemiddelde kniehoogte en de standaardafwijking die je uit de figuur kunt aflezen overeen komen met de berekende waarden.

d Zijn de kniehoogtes bij benadering normaal verdeeld?

e 60% van de kniehoogtes zit tussen $\mu - a$ en $\mu + a$. Hoe groot is a? Lees je antwoord uit de figuur af.

f Welke minimale lengte hebben de 15% grootste kniehoogtes? Lees je antwoord uit de figuur af.

Opgave 13

Literpakken melk worden machinaal gevuld. Zo'n pak heeft een inhoud die normaal is verdeeld met een gemiddelde van 1,010 liter en een standaardafwijking van 0,006 liter. De vulmachine staat ingesteld op het vullen van gemiddeld 1,005 liter melk in zo'n pak met een standaardafwijking van 0,004 liter. Ook dit vulvolume is normaal verdeeld.

a In hoeveel procent van de gevallen gaat bij het vullen melk verloren?

b Het gemiddelde vulvolume kan worden ingesteld. Hoeveel moet dit bedragen opdat in niet meer dan 1% van de gevallen melk verloren gaat bij het vullen van een pak?

Practicum

Hier vind je een **leeg blad normaal-waarschijnlijkheidspapier** in drie kleuren.

- **normaal waarschijnlijkheidspapier (paars)**
- **normaal waarschijnlijkheidspapier (oranje)**
- **normaal waarschijnlijkheidspapier (grijs)**