

A full-page background image of a basketball player in mid-air, performing a jump shot or layup. The player is wearing a dark blue jersey with red and white accents. A large, semi-transparent white number '10' is overlaid on the left side of the image. In the background, a crowd of spectators is visible in a stadium setting, with a banner that reads 'London 2012' partially visible.

10

Statistiek gebruiken

Wat leer je?

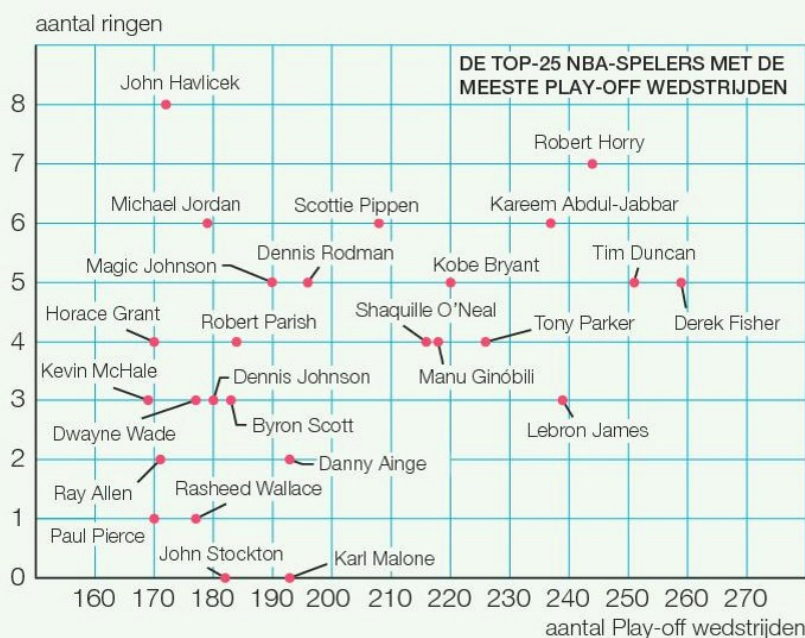
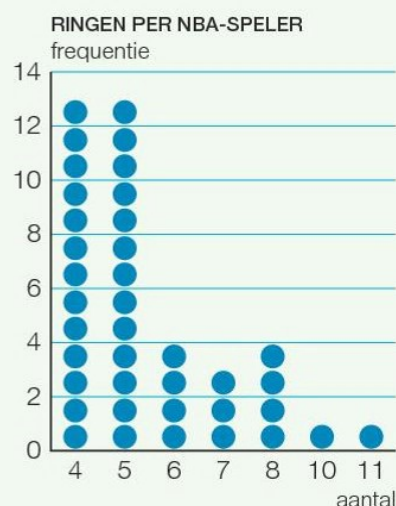
- Werken met data uit verschillende soorten diagrammen.
- De voorwaarden voor een oorzakelijk verband nagaan.
- Een geschikte associatiemaat voor het verschil tussen twee groepen kiezen.
- Vaststellen of een steekproef aselekt is.
- Werken met de normale verdeling en betrouwbaarheidsintervallen in contexten.



Beginopdracht NBA ringen

[▶ WERKBLAD] De in 1949 in de VS opgerichte National Basketball Association (NBA) is de meest prestigieuze basketbalcompetitie ter wereld. Aan het einde van elk seizoen worden zogenaamde Play-off wedstrijden gespeeld om te bepalen welke twee teams tegen elkaar spelen in de finale. De winnaars krijgen een ring en zijn voor een jaar kampioen van de NBA. In de figuur hiernaast staat de verdeling van het aantal ringen per speler, voor spelers met minstens vier ringen.

Jyri volgt de NBA nauwlettend. Hij heeft opgezocht wie in de top-25 van spelers met de meeste Play-off wedstrijden staan en hoeveel ringen zij wonnen. Het resultaat staat in de figuur hieronder. Deze figuur staat vergroot op het werkblad.



Bill Russell heeft het record aantal ringen voor spelers en won $\frac{1}{15}$ ring per Play-off wedstrijd die hij speelde.

- Onderzoek hoeveel Play-off wedstrijden Bill Russell speelde. Waarom komt hij niet voor in de figuur hierboven?

Bill Russell heeft het record aantal ringen voor spelers.

Volgens Jyri kun je nagaan wie het beste presteerde door te kijken wie per Play-off wedstrijd de meeste ringen won.

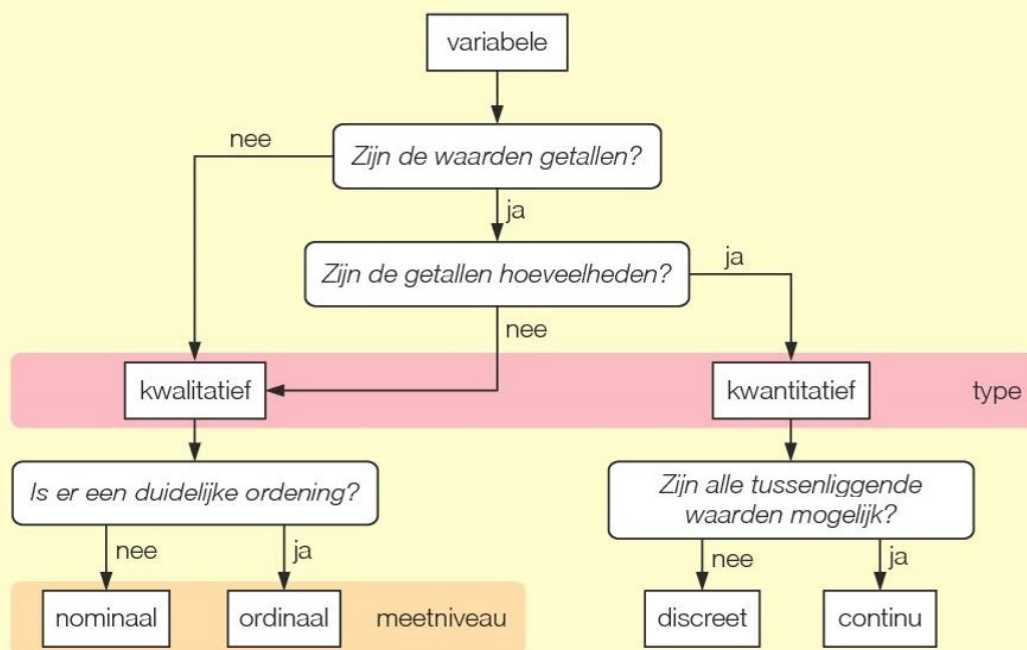
- Welke speler in de top-25 heeft volgens Jyri het beste gepresteerd?
- Kobe Bryant (overleden in 2020) is Jyri's favoriete speler. Hoeveel procent van de NBA spelers in de top-25 presteert volgens Jyri beter dan Kobe Bryant?
- De favoriete speler van Sanne staat in de top-25 en heeft het record aantal Play-off wedstrijden voor spelers met gemiddeld meer dan 70 Play-off wedstrijden per ring. Wie is de favoriete speler van Sanne?

Voorkennis Statistische variabelen

Theorie A De indeling van statistische variabelen

Bij statistiek heten de kenmerken die je onderzoekt variabelen. We maken onderscheid tussen kwalitatieve en kwantitatieve variabelen. De kwalitatieve variabelen hebben een meetniveau (nominaal of ordinaal). Bij kwantitatieve variabelen kun je onderscheid maken tussen discreet en continu.

Voor het indelen van een variabele kijk je naar de *waarden* die deze aan kan nemen en gebruik je het volgende schema.



In het voorbeeld zie je hoe je dit schema gebruikt. Noteer bij soortgelijke opgaven waarin een toelichting wordt gevraagd je antwoord zoals in het voorbeeld.

Voorbeeld

Bij een onderzoek is van 10 000 producten onder andere de variabele streepjescode gebruikt, met als waarden getallen van 13 cijfers.

Is deze variabele nominaal, ordinaal, discreet of continu? Licht toe.



Uitwerking

Wel getallen, geen hoeveelheden, geen duidelijke ordening. Dus nominaal.

- 1 In de vragenlijst hieronder komen een aantal variabelen voor. Geef bij elke variabele aan of deze nominaal, ordinaal, discreet of continu is. Licht toe.

Vragenlijst: Omcirkelen wat van toepassing is

geboortejaar: 2004

hoogst genoten opleiding: geen / middelbare school / mbo / hbo / wo

telefoonnummer: 0786127768

mogen we u mail sturen? ja / nee

- 2 Ibrahim, Senna en Wissal gaan voor een experiment herhaaldelijk met pijl en boog schieten op het doel in afbeelding 10.1. Ze hebben een lijstje gemaakt van de variabelen die ze willen verzamelen om later te analyseren.

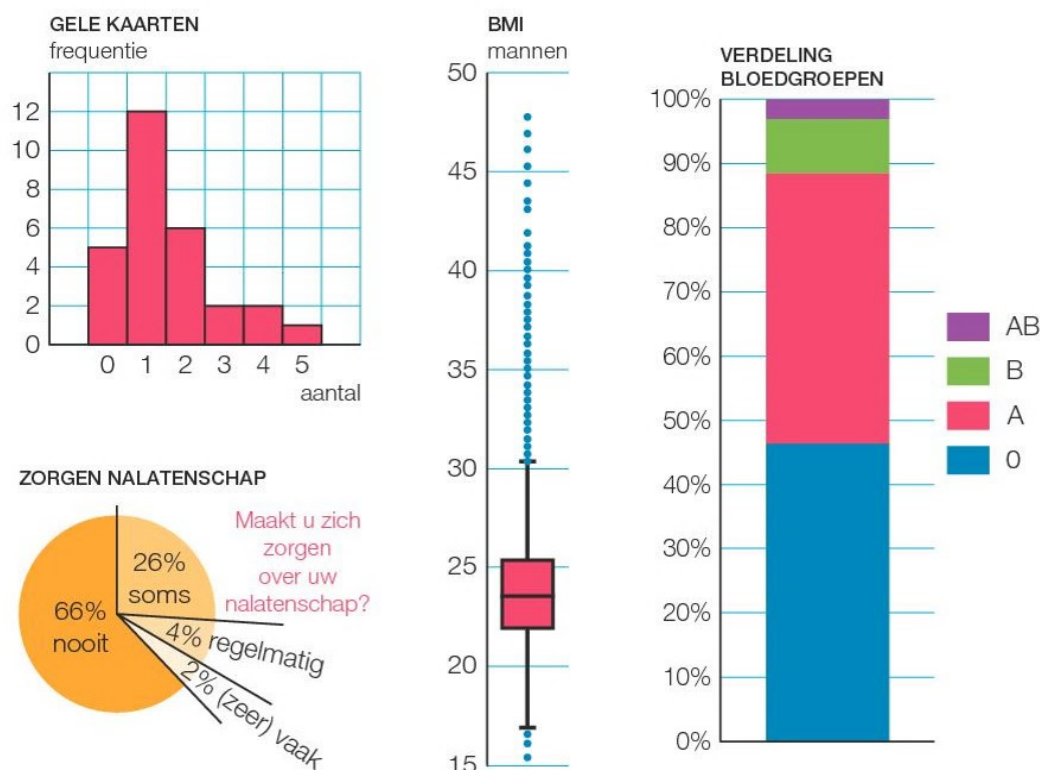


figuur 10.1

Geef van de volgende variabelen aan of ze nominaal, ordinaal, discreet of continu zijn. Licht toe.

- De naam van de schutter van de pijl.
- De afstand van de pijl tot het midden.
- Het aantal hele millimeters dat de pijl het gele gebied mist.
- De kleur van het gebied dat de pijl raakte.
- Of de pijl wel of niet dichterbij het midden zit dan de vorige pijl.
- De snelheid van de pijl in km/uur op gehelen afgerond.

- 3 Geef van elk van de diagrammen hieronder aan of de variabele kwalitatief of kwantitatief is. Geef ook het meetniveau of geef aan of de variabele discreet of continu is. Licht toe.



figuur 10.2

10.1 Statistische diagrammen



Een onderzoeker vraagt in een enquête onder andere naar het geslacht van de respondenten. Ze wil de resultaten van de enquête verwerken in enkele diagrammen.

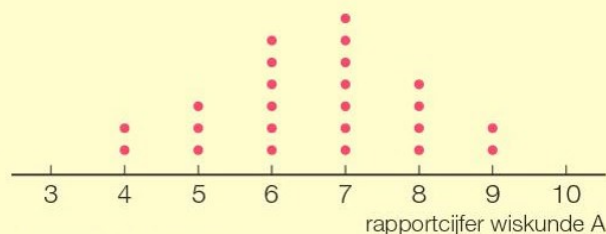
- a Kan ze bij de variabele geslacht een lijndiagram maken? Licht toe.
- b Kan ze bij de variabele geslacht een cirkeldiagram maken? Licht toe.

Theorie A Diagrammen bij data

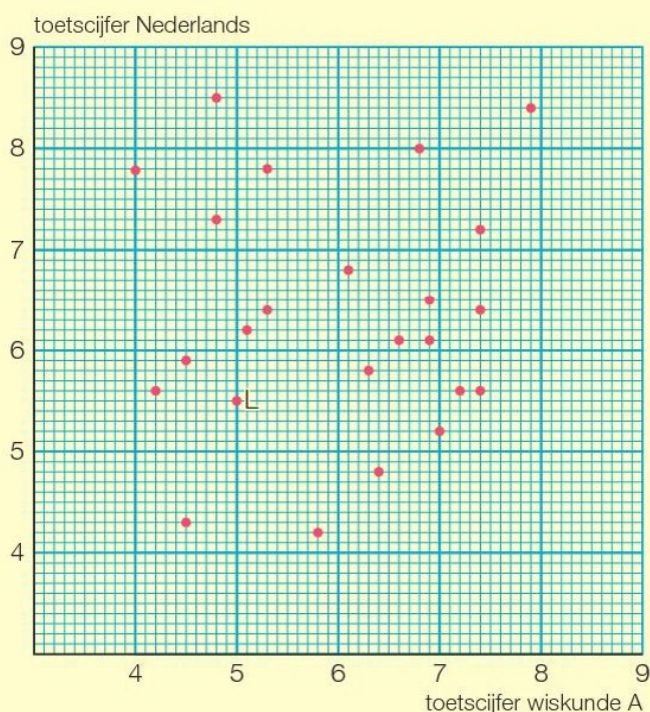
Voor de presentatie van data kun je kiezen uit verschillende typen diagrammen.

Bekijk je bijvoorbeeld de verdeling van rapportcijfers voor het vak wiskunde A, dan geeft de **dotplot** in figuur 10.2 een goed overzicht. Bekijk je echter de samenhang tussen de toetscijfers wiskunde A en de toetscijfers Nederlands, dan kun je beter gebruikmaken van een **spreidingsdiagram**. Een spreidingsdiagram wordt ook wel een **puntenwolk** genoemd. In het spreidingsdiagram in figuur 10.3 zie je dat Leonie (L) een 5 heeft gehaald voor de toets van wiskunde A en een 5,5 voor de toets van Nederlands.

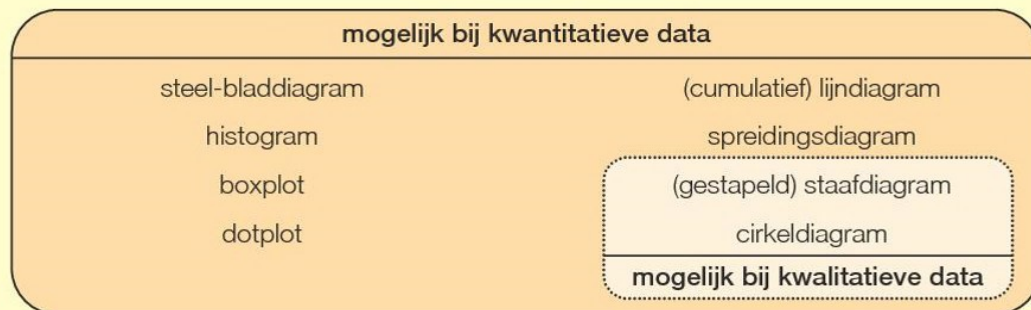
Als je te maken hebt met een kwalitatieve variabele, dan kun je alleen een staafdiagram of een cirkeldiagram maken. Heb je een kwantitatieve variabele dan kun je alle diagrammen maken, maar daarvoor moet je de data soms eerst bewerken. Zie de figuur hieronder.



figuur 10.2 dotplot



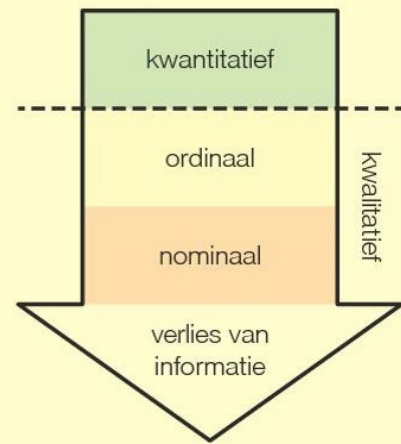
figuur 10.3 spreidingsdiagram



Bij het bewerken van data kan informatie verloren gaan. Als je bijvoorbeeld het gewicht van eieren in gram indeelt in de gewichtsklassen S, M, L en XL, dan is de data na omzetting kwalitatief in plaats van kwantitatief. Met deze indeling in gewichtsklassen is de variabele ordinaal geworden.

De omzettingen lopen, door verlies van informatie, in één richting. De richting is in het diagram hiernaast met een pijl aangegeven.

Het omzetten van een variabele kan gevolgen hebben voor het soort diagram dat je kunt maken. Als je bijvoorbeeld in plaats van het gewicht in gram de bijbehorende gewichtsklassen gebruikt, kun je hierbij geen boxplot meer tekenen. Er kan namelijk niet meer achterhaald worden wat de waarden zijn van het kleinste waarnemingsgetal, Q_1 , de mediaan, Q_3 en het grootste waarnemingsgetal.



figuur 10.4

Voorbeeld

Stefan heeft van alle leerlingen in zijn klas de lengte in cm nauwkeurig gemeten.

- Zijn de metingen van Stefan kwalitatief of kwantitatief?
- Stefan heeft in een cirkeldiagram met twee sectoren weergegeven welk deel van de leerlingen van zijn klas langer dan 1,70 m is. Licht toe dat de variabele nu kwalitatief is.
- Geef voor elk van de drie centrummaten aan of je deze uit het cirkeldiagram kunt aflezen.

Uitwerking

- kwantitatief
- De metingen waren hoeveelheden en zijn nu twee categorieën.
- Niet het gemiddelde en de mediaan, maar wel de modus.



Zie het voorbeeld.

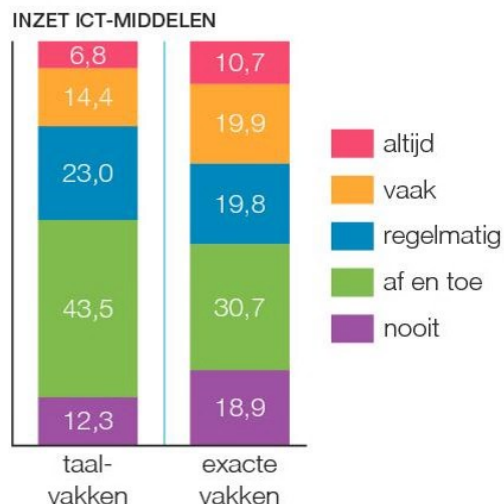
Stefan besluit nog een diagram te maken. Hij maakt daarom bij de variabele lengte een klassenindeling met als klassen:

$\langle \leftarrow; 1,70 \rangle$, $\langle 1,70; 1,90 \rangle$ en $\langle 1,90; \rightarrow \rangle$.

- Wat is het meetniveau van de variabele na de klassenindeling?
- Noem twee diagrammen die Stefan kan gebruiken om de data na de klassenindeling te presenteren.
- Waarom kan hij uit de klassenindeling niet meer de gemiddelde lengte berekenen?

3 In 2018 werd een onderzoek gedaan naar de mate van inzet van digitale middelen (ICT-middelen) bij de verschillende vakken op de middelbare school. In de figuur hiernaast zie je een diagram bij de resultaten van dit onderzoek.

- Wat is het meetniveau van de variabele 'mate van inzet'?
- Wat is de relatieve cumulatieve frequentie van *vaak* bij de taalvakken? En bij de exacte vakken?
- Welke centrum- en spreidingsmaten zijn bij dit diagram te gebruiken?
- Linda wil in een ander type diagram presenteren hoe de mate van inzet van ICT-middelen bij de taalvakken verdeeld is. Welk diagram zou je aanraden? Licht toe.



figuur 10.5

4 Samuel doet een klein onderzoek in de bovenbouw van zijn middelbare school. Hij vraagt daarbij 32 leerlingen naar het aantal liedjes dat ze dagelijks op hun mobiel streamen vanuit de databundel. Zijn resultaten staan verwerkt in het dubbel steel-bladdiagram hiernaast. Je ziet dat er één jongen is die dagelijks 25 liedjes streamt.

- Welke variabele is gebruikt voor het splitsen van de groepen links en rechts in het dubbel steel-bladdiagram? Wat is hiervan het meetniveau?

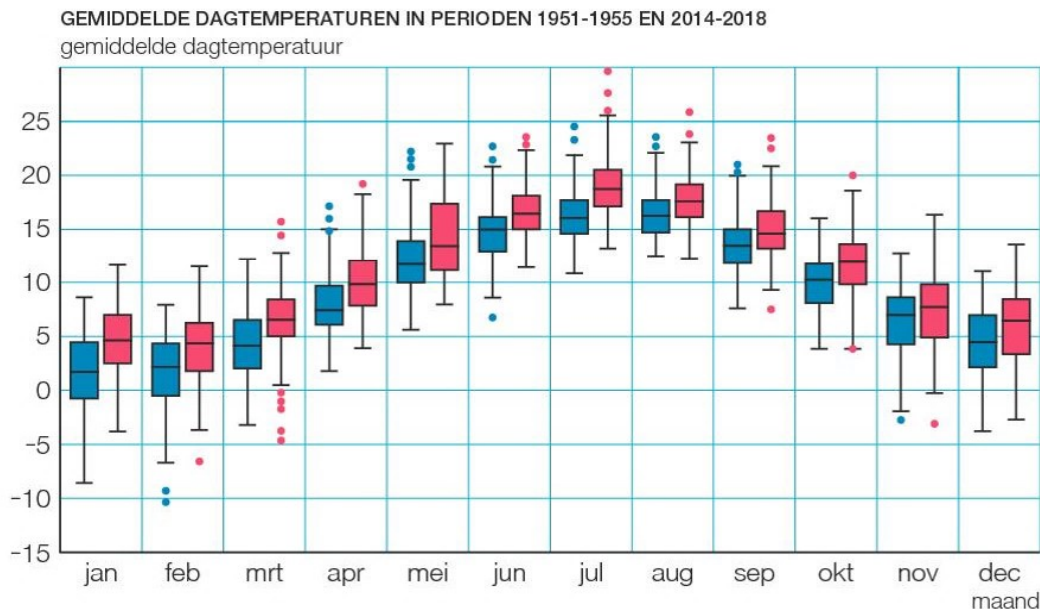
AANTAL DAGELIJKS BELUISTERDE LIEDJES VIA STREAMING															
jongens						meisjes									
3 1 0					0	0 0 0 0 0 0 0 0 0 3 5									
8 7 4 1 1 0					1	0 0 5									
5					2										
0					3										
0					4										
0 0					5	0									
0					6	0									
					7	5									
eenheden					tientallen	eenheden									

figuur 10.6

Samuel berekende $\bar{X}_1 = 22,7$ met $S_1 = 19,2$ en $\bar{X}_2 = 13,4$ met $S_2 = 23,9$.

- Koos hij voor groep 1 de jongens of de meisjes? Licht toe.
- Welke centrum- en spreidingsmaten had hij nog meer kunnen gebruiken om de groepen te vergelijken?
- Thom kreeg van Samuel het dubbel steel-bladdiagram toegestuurd. Hij wil onderzoeken of er een samenhang is tussen de leeftijd en het aantal muziknummers dat dagelijks gestreamd wordt. Hij heeft daarvoor ook nog een dotplot van de leeftijden bij dit onderzoek tot zijn beschikking. Heeft hij zo voldoende informatie voor een puntenwolk? Licht toe.

In figuur 10.7 zijn voor twee perioden van 5 jaar de gemiddelde Nederlandse dagtemperaturen in elke maand vergeleken met behulp van boxplots. In deze opgave noemen we het tijdvak 1951-1955 periode I en het tijdvak 2014-2018 periode II.



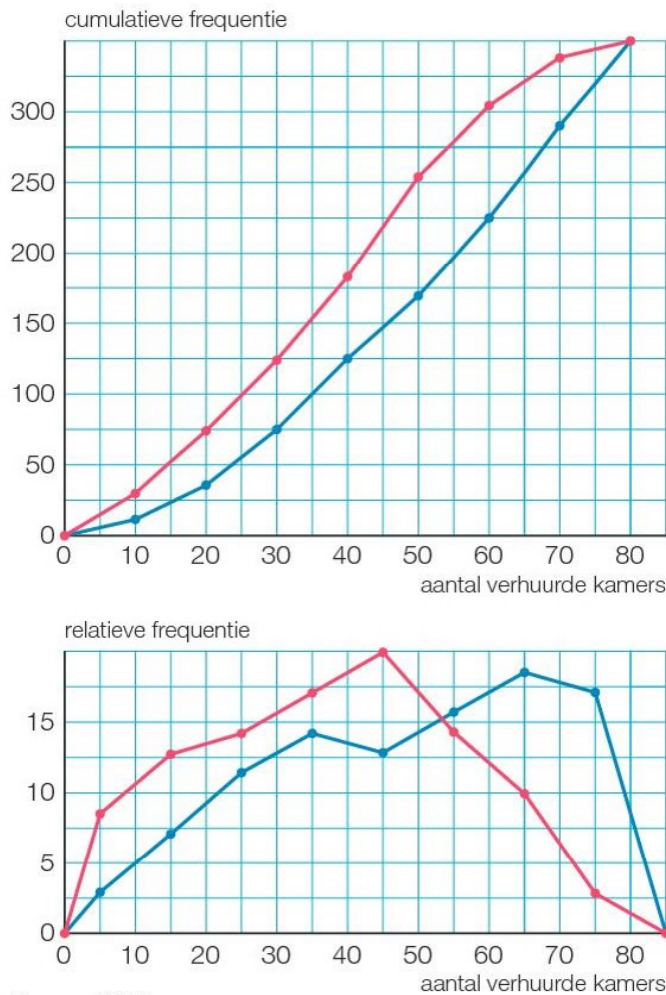
figuur 10.7

- Is de variabele gemiddelde dagtemperatuur kwalitatief of kwantitatief? Wat is het meetniveau van de variabele maand?
- In de figuur is de opwarming van de aarde goed te zien. De gemiddelde dagtemperaturen van periode I zijn doorgaans lager dan de gemiddelde dagtemperaturen van periode II. Welke boxplots horen bij periode II? Licht toe.

In de figuur zijn bij sommige boxplots stippen gezet. Dit zijn uitschieters die meer dan 1,5 keer de interkwartielafstand van de box af liggen.

- Yoeri is verbaasd dat de rode boxplot in januari geen uitschieter heeft. Bereken bij welke temperaturen er sprake zou zijn van een uitschieter.
- Noem een maand en periode waarin je verwacht dat het gemiddelde lager zal zijn dan de mediaan. Licht je antwoord toe.
- Marlijn maakt bij de gegevens in de figuur een boxplot bij elke periode. De twee boxplots die ze zo krijgt hebben verschillende spreidingsbreedten. Hoe groot is de grootste spreidingsbreedte? Neem de uitschieters hierbij niet mee.

Het gaat niet goed met hotel Bergzicht. Het hotel verdient vooral aan de verhuur van kamers en juist deze inkomstenbron is de afgelopen vijf jaar flink teruggelopen. In de diagrammen hieronder staan voor de jaren 2015 en 2019 de aantallen verhuurde kamers per dag verwerkt met als eerste klasse $[0, 10)$.



figuur 10.8

In de diagrammen is niet de data van elke dag van het jaar verwerkt.

a Hoeveel dagen werden in de diagrammen verwerkt? Licht toe.

Bij het maken van de diagrammen is men vergeten een legenda toe te voegen. Omdat bekend is dat de verhuur terugloopt, is toch te achterhalen welke lijndiagrammen bij welk jaar horen.

b Welke lijndiagrammen horen bij de kamerverhuur van 2019? Licht toe.

c Gebruik twee centrummaten om te demonstreren dat het slecht gaat met de kamerverhuur in hotel Bergzicht.

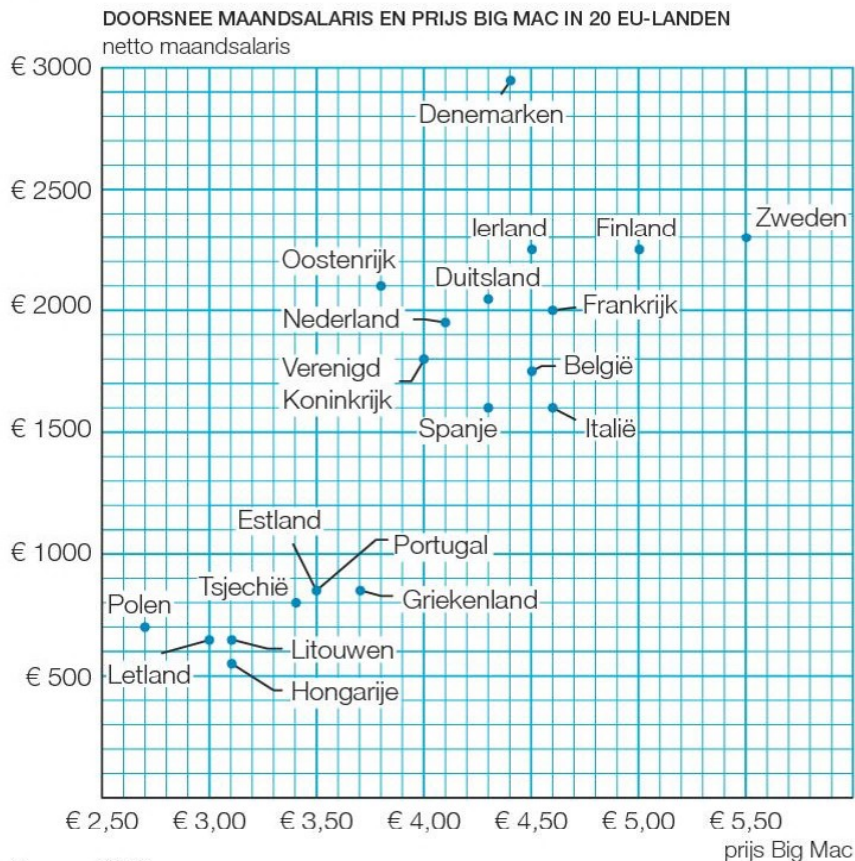
Doordat bij het verwerken van de data een klassenindeling is gebruikt, is niet meer precies af te lezen wat de mediaan van het aantal verhuurde kamers elk jaar is. De eigenaar van het hotel denkt zich te herinneren dat de medianen 40 en 53 waren.

d Onderzoek met de figuren of dit zou kunnen kloppen.

e Ans zegt dat je aan de cumulatieve lijndiagrammen ziet dat voor de klasse $[50, 60)$ de frequentie voor beide perioden gelijk is. Bert zegt dat je in de relatieve polygoenen ziet dat dit niet het geval is.

Wie heeft gelijk? Licht toe.

[▶ **WERKBLAD**] Niet alle EU-landen zijn even welvarend. Omdat elk EU-land McDonalds restaurants heeft en al deze restaurants de zogenaamde Big Mac serveren, kun je deze hamburger gebruiken om het welvaartsniveau tussen verschillende EU-landen te vergelijken. Zie de figuur met gegevens uit 2019. Deze figuur staat ook op het werkblad.



figuur 10.9

In het diagram wordt een verband gelegd tussen het doorsnee netto maandsalaris en de prijs van een Big Mac in verschillende EU-landen.

- Hoe heet dit type diagram?
- In de titel boven het diagram staat dat er 20 EU-landen zijn vergeleken, terwijl er maar 19 punten in het diagram staan. Leg uit waardoor dit komt.

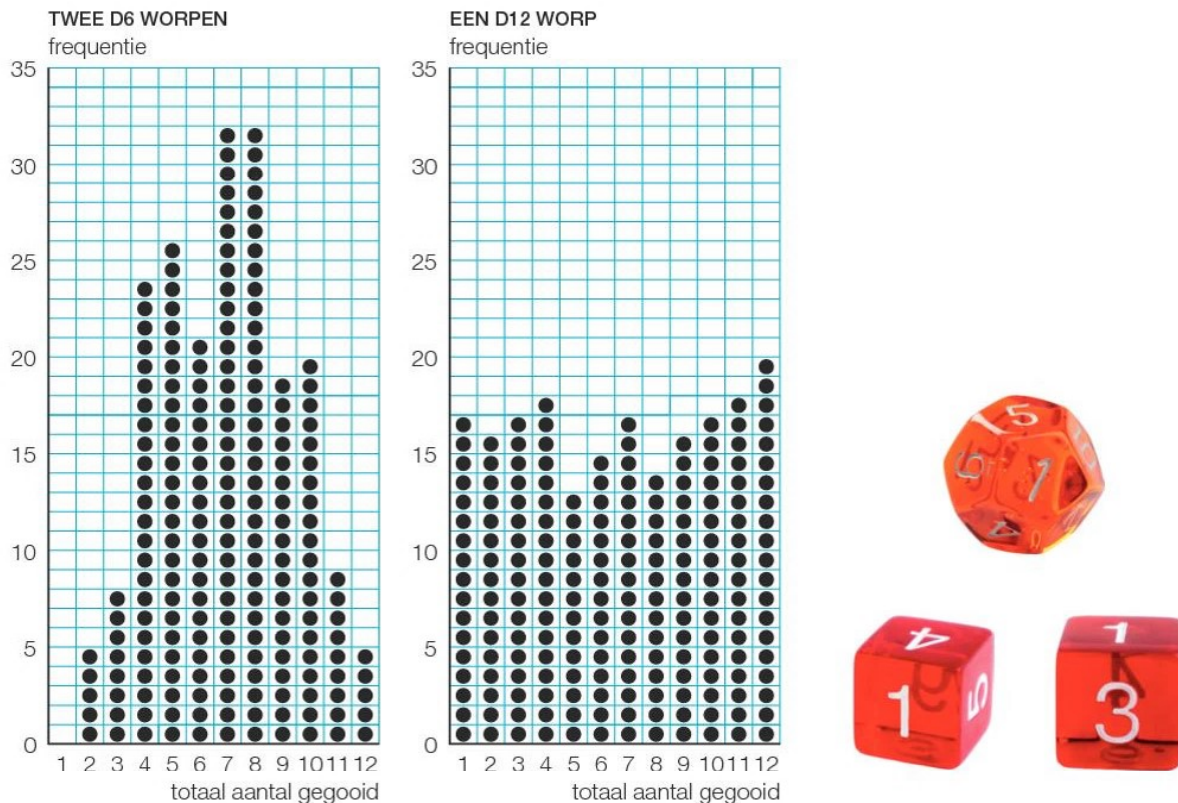
Op de verticale as staat het doorsnee netto maandsalaris. Uit de formulering is niet duidelijk welke centrummaat hierbij is gebruikt.

- Welke centrummaat is hiervoor niet geschikt? Licht toe. Gebruik dat de verdeling van maandsalarissen rechts-scheef is.
- Is de variabele netto maandsalaris kwalitatief of kwantitatief?
- Megan en Nicole vragen zich af in welk land ze de meeste Big Macs kunnen kopen voor een netto maandsalaris in dat land. Onderzoek welk land dat is.

In sommige landen is de Big Mac in verhouding tot het netto maandsalaris zo duur dat je er minder dan 250 keer een Big Mac van kunt kopen.

- Onderzoek in welke landen dit het geval is.

Roel en Michiel spelen een spel waarbij je twee opties hebt voor je dobbelsteenwerpen. Optie 1 is een worp met twee dobbelstenen met 6 vlakken (deze dobbelstenen worden ook wel D6 genoemd). Optie 2 is een worp met één dobbelsteen met 12 vlakken (zo'n dobbelsteen heet een D12). Ze voeren voor beide opties een experiment uit met 200 worpen. De resultaten zie je in de figuur hieronder.

**figuur 10.10**

a In welk type diagram staan de resultaten in de figuur gepresenteerd?

Bij het spel dat ze spelen zijn er situaties waarin een speler in totaal minstens 10 moet gooien. Er zijn ook situaties dat een speler niet minder dan in totaal 4 mag gooien.

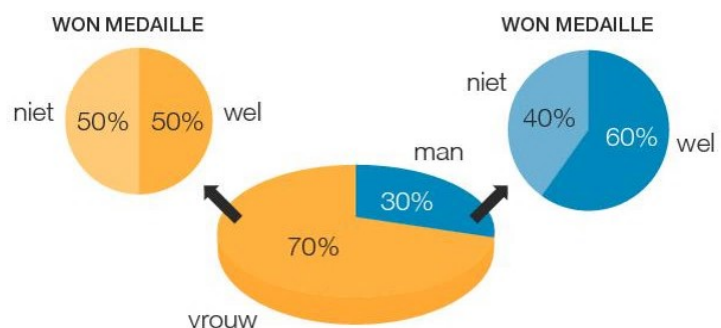
- b** Welke optie (1 of 2) zou je aanraden in elk van deze situaties?
- c** Welke standaardverdeling hoort bij de twee D6 worpen? En welke bij de D12 worp?
- d** Bij welke verdeling hoort de kleinste standaardafwijking? Licht toe.

E9
*

Bij een sportgala is een groot aantal sporters gevraagd of ze een olympische medaille hebben gewonnen. De resultaten staan in het diagram hiernaast.

Je ziet bijvoorbeeld dat van de ondervraagden op het gala 70% vrouw is.

Onderzoek welk percentage van de medaillewinnaars vrouw is.

**figuur 10.11**

Ook in de wetenschap komt fraude voor. Een voorbeeld is het onderzoek van een Amerikaanse voedingsdeskundige naar eetgedrag. Hij bestudeerde proefpersonen die onbeperkt pizza mochten eten. Eén groep moest daarvoor de volle prijs betalen, terwijl een andere groep korting kreeg. Het vermoeden van de onderzoeker was dat mensen die korting krijgen meer gaan eten dan goed voor ze is.

Na afloop legde hij de deelnemers een aantal stellingen voor, zoals ‘ik heb teveel gegeten’. Bij elke stelling moesten deelnemers een getal omcirkelen van 1 tot en met 9. Hier betekende 1 ‘helemaal niet mee eens’ en 9 ‘helemaal mee eens’.

- a Met welk meetniveau heb je hier te maken?
- b Noem twee soorten diagrammen die geschikt zijn om de antwoorden op de stellingen weer te geven.

In de tabel hieronder zijn de resultaten weergegeven per groep. Je ziet bijvoorbeeld dat er in de groep met korting 7 mensen waren die drie stukken pizza aten. Ook is per stelling berekend wat de gemiddelde score is.

- c Leg uit welk bezwaar is in te brengen tegen het berekenen van deze gemiddelden.

	korting			volle prijs		
aantal stukken pizza	één	twee	drie	één	twee	drie
aantal deelnemers	18	18	7	17	19	10
Ik heb meer pizza gegeten dan zou moeten.	2,63	4,82	6,00	1,76	3,53	4,40
Ik voel me schuldig over hoeveel pizza ik at.	2,39	3,44	3,71	2,26	1,68	2,90
Ik voel me ziek.	2,17	2,94	2,43	1,97	1,45	2,25
Ik heb teveel gegeten.	2,11	3,89	3,71	1,67	1,67	3,50

We gaan in detail kijken naar de gemiddelden in de tabel. Daarvoor geven we eerst een eenvoudig voorbeeld. Stel dat we een vraag hebben met antwoordmogelijkheden 1 - 2 - 3 en dat twee deelnemers een getal omcirkelen.

- d Onderzoek waarom 2,40 onmogelijk het gemiddelde kan zijn van de antwoorden van deze twee deelnemers.

Een kritische lezer ontdekte dat een gemiddelde in de derde kolom van de groep die de volle prijs betaalde niet kan kloppen.

- e Onderzoek welke waarde dat is.
- f Onderzoek welke waarden in de eerste kolom van de groep met korting en de eerste kolom van de groep die de volle prijs betaalde, onmogelijk zijn. Houd hierbij rekening met afronding.

Terugblik

Diagrammen bij data

Voor de presentatie van data kun je kiezen uit verschillende typen diagrammen. Welk diagram je kunt kiezen hangt af van het soort data dat je tot je beschikking hebt. Zie het overzicht hieronder.

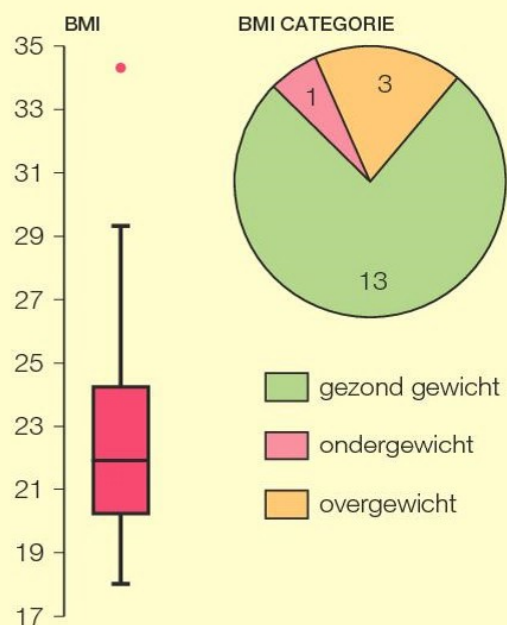
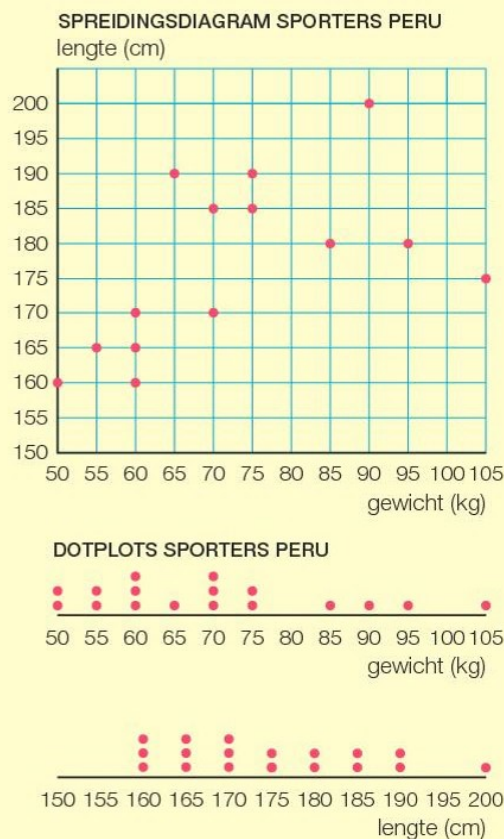


In het spreidingsdiagram zijn de kwantitatieve variabelen lengte en gewicht van enkele sporters uit Peru verwerkt. Hierin is te zien welk gewicht bij welke lengte hoort. Zo zie je bijvoorbeeld dat de langste sporter niet de zwaarste sporter was.

De dotplots geven op het eerste gezicht dezelfde informatie als het spreidingsdiagram, maar in de dotplots is niet meer te zien welk gewicht bij welke lengte hoort. In de dotplots is wel te zien dat het om 17 sporters gaat. Daardoor weet je dat in het spreidingsdiagram enkele metingen elkaar overlappen.

Met de kwantitatieve gegevens uit het spreidingsdiagram en de dotplots kun je van elk van de 17 sporters de BMI berekenen en deze verwerken in een boxplot.

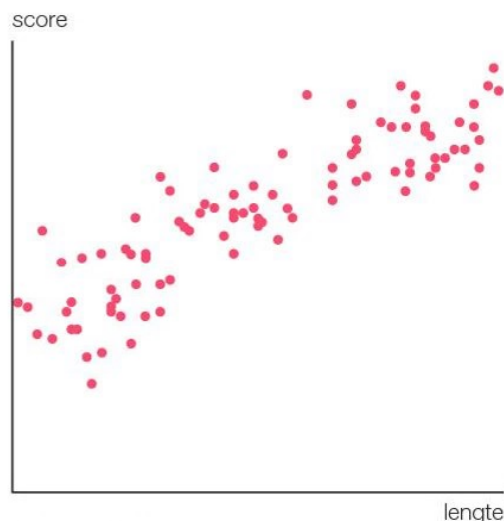
Na verwerking in een boxplot is hieruit geen gemiddelde of standaardafwijking meer te berekenen, er is dus informatie verloren gegaan. Soms is het voor een overzichtelijke presentatie nodig om informatie achterwege te laten. Kies je voor verwerking in een cirkeldiagram, dan kun je bijvoorbeeld een indeling in categorieën maken. De data in het cirkeldiagram is nu kwalitatief, waardoor nog maar weinig berekeningen gemaakt kunnen worden.



10.2 Causaliteit en groepen vergelijken

011
☐◎*

Op een middelbare school is een groot onderzoek gedaan onder de leerlingen van de klassen 1, 3 en 5 van de havo-afdeling. Van elke leerling uit deze klassen zijn gegevens over onder andere de lengte, het gewicht, de eetgewoonte en de mate van discipline in een database gezet. Nadat deze gegevens zijn verzameld, heeft elke leerling dezelfde algemene kennistoets gemaakt. Het bestuur van de school wil met deze gegevens onder andere onderzoeken waardoor leerlingen goed presteren op de kennistoets. Daartoe zijn in de database allerlei verbanden onderzocht. In figuur 10.12 is de score op de toets uitgezet tegen de lengte. Elk punt in het spreidingsdiagram stelt een leerling voor.



figuur 10.12

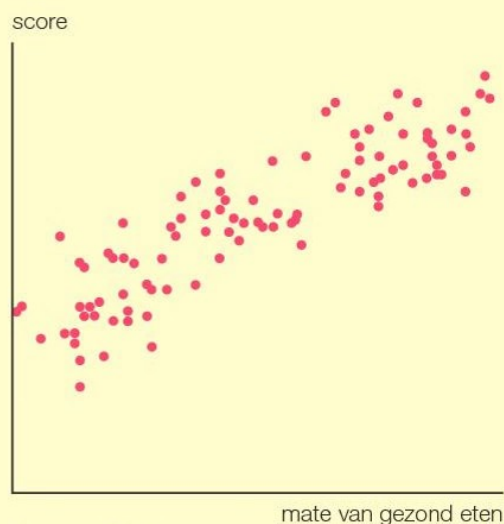
- a Licht toe hoe je in de figuur kunt zien dat er een samenhang is tussen de lengte en de score.
- b Wat denk je, is een grotere lengte een oorzaak van het beter scoren op de toets? Licht toe.

Theorie A Samenhang en causaliteit

Vaak kun je met gezond verstand bedenken dat een samenhang niet op een oorzaak wijst. De lengte van een scholier is bijvoorbeeld geen oorzaak van hoge scores op een algemene kennistoets.

Als blijkt dat het aantal jaren onderwijs wel de oorzaak is van goede scores, dan is er sprake van een *oorzakelijk verband* tussen het aantal jaren onderwijs en goede scores. In plaats van oorzakelijk verband spreken we ook wel van een *causaal verband* of *causaliteit*.

Bij het onderzoek uit opgave 11 is ook onderzocht of er een verband bestaat tussen gezond eten en scores op de toets. Ook hieruit komt een samenhang naar voren. Zie figuur 10.13. Om te weten te komen of deze samenhang causaal is, moet worden onderzocht of aan de volgende twee voorwaarden is voldaan.



figuur 10.13

Voorwaarden voor een causaal verband bij een samenhang

- 1 De oorzaak moet voorafgaan aan het gevolg.
- 2 Er mogen geen andere variabelen zijn die de samenhang verklaren.

In het geval van de samenhang tussen gezond eten en scores is aan de eerste voorwaarde voldaan: de toets is gemaakt nadat er is geconstateerd hoe gezond de leerling eet.

Om te onderzoeken of aan de tweede voorwaarde is voldaan, moet je allereerst variabelen bedenken die het verband tussen gezond eten en scores zouden kunnen verklaren. De discipline van leerlingen is zo'n mogelijke variabele. Leerlingen met meer discipline eten mogelijk gezonder en scoren wellicht beter op de kennistoets (doordat ze beter hun best doen op school).

Vervolgens onderzoek je of de variabelen die je hebt bedacht de samenhang verklaren. Om na te gaan of de variabele 'discipline' de samenhang verklaart, kun je als volgt te werk gaan. Verdeel de leerlingen in de categorieën lage discipline, matige discipline en hoge discipline en zet van elke categorie de score uit tegen de mate van gezond eten. Zo ontstaan de volgende drie spreidingsdiagrammen.



figuur 10.14

Je ziet aan de figuur dat er per categorie geen samenhang is tussen gezond eten en scores. Ook zie je dat leerlingen met meer discipline zowel gezonder eten als hoger scoren op de toets. Discipline is dus een variabele die de samenhang tussen gezond eten en scores verklaart. De variabele discipline heet in dat geval de **gemeenschappelijke factor**. Er is geen sprake van een causaal verband tussen gezond eten en scores, omdat niet aan de tweede voorwaarde is voldaan.

In het algemeen is het niet eenvoudig om aan te tonen dat wel aan de tweede voorwaarde is voldaan. In de praktijk komt onderzoek er vaak op neer dat mogelijke gemeenschappelijke factoren één voor één worden uitgesloten. Hoe meer variabelen worden uitgesloten, hoe waarschijnlijker het wordt dat het inderdaad om een oorzakelijk verband gaat.

Je kunt soms alle mogelijke gemeenschappelijke factoren in één keer uitsluiten met een handige onderzoeksopzet. Zie opgave 15. In de praktijk is dit vaak kostbaar.

Er hoeft daarentegen maar één gemeenschappelijke factor te zijn die de samenhang verklaart, zoals discipline, om te zeggen dat er geen sprake is van een causaal verband.

Tenzij anders vermeld gaan we er in de opgaven van uit dat als een oorzakelijk verband in de onderzochte groep bestaat, het ook in de populatie bestaat.

R12 Zie de theorie.



- a** Leg uit hoe je in figuur 10.14 kunt zien dat er binnen elk van de drie categorieën van discipline geen samenhang is tussen gezonder eten en scores.
- b** Hoe zie je aan de drie spreidingsdiagrammen in figuur 10.14 dat leerlingen met meer discipline zowel gezonder eten als hoger scoren op de toets?
- c** Op grond van figuur 10.13 overweegt het bestuur van de school om in de kantine gezonder eten aan te bieden om de prestaties op toetsen te verbeteren.
Wat vind jij hiervan?
- d** Joanne heeft een samenhang gevonden waarbij ze drie gemeenschappelijke factoren kan bedenken. Uit onderzoek blijkt dat geen van de drie de samenhang verklaart. Er is verder aan voorwaarde 1 voor een causaal verband voldaan.
Waarom weet Joanne toch niet helemaal zeker dat er sprake is van een causaal verband?

13

In de volgende gevallen is er wel sprake van een samenhang maar zie je met gezond verstand meteen dat er geen sprake is van een causaal verband. Geef steeds een gemeenschappelijke factor die de samenhang verklaart.

- a** Hoe meer ijsjes er in een week worden verkocht, hoe meer (bijna) verdrinkingsgevallen in die week.
- b** Mensen die een muts dragen zijn vaker verkouden.
- c** Hoe hoger het energieverbruik van gezinnen in een straat, hoe groter het aantal BMW's in de straat.
- d** Hoe meer brandweerlieden een brand bestrijden, hoe meer schade de brand aanricht.



14
□ ◎ *

Een gymleraar wil aantonen dat gewichtheffen een oorzaak is van grotere bovenarmen. Hij laat daarom de jongens tijdens de gymles gewichtheffen. De meisjes volgen het gewone gymlesprogramma. Na drie maanden stelt hij vast dat de jongens een grotere omvang van de bovenarmen hebben dan de meisjes.

- a Kun je op basis van de gekozen onderzoeksopzet concluderen dat gewichtheffen een oorzaak is van grotere bovenarmen? Licht toe.

Naar aanleiding van kritiek op zijn eerste onderzoeksopzet vraagt de gymleraar welke leerlingen uit zijn klassen willen gewichtheffen. Deze leerlingen gaan drie maanden gewichtheffen, de rest van de leerlingen volgt de gewone gymles. De leerlingen die meededen met gewichtheffen blijken na afloop grotere bovenarmen te hebben dan de rest van de leerlingen.

- b Kun je nu wel concluderen dat gewichtheffen een oorzaak is van grotere bovenarmen? Licht toe.

A15
□ ◎ *

Wiskundedocent mevrouw Vos volgt nauwgezet het boek in haar havo 5-klas bij wiskunde A. Naast het gewone huiswerk geeft ze elke les een aantal opgaven op uit de serie opgaven die vanaf bladzijde 258 achter in dit boek staan.

Ze constateert dat de leerlingen die vaker deze opgaven maken, hogere cijfers halen op toetsen. Ze vraagt zich af of er een oorzakelijk verband is tussen het maken van deze opgaven en de cijfers op toetsen.

- a Licht toe waarom inzet een gemeenschappelijke factor kan zijn bij de samenhang tussen het maken van de genoemde opgaven en de cijfers op toetsen.

Wetenschappers in de VS hebben in 2019 onderzocht of het voortdurend afwisselen van onderwerp bij het maken van wiskundeopgaven betere resultaten oplevert bij toetsen.

Hiervoor werden 800 scholieren aselekt in twee groepen ingedeeld. Bij groep 1 gingen de opgaven per les over hetzelfde onderwerp. Bij groep 2 werden naast de opgaven over het behandelde onderwerp ook opgaven aangeboden over andere onderwerpen. Na vier maanden werd in beide groepen dezelfde onverwachte toets afgenomen. Het gemiddelde in groep 1 was een 4,4 en in groep 2 een 6,5.

Bij aselekt indelen worden leerlingen door toeval ingedeeld in een groep.

- b Leg uit waarom je verwacht dat leerlingen met een grote inzet ongeveer gelijk verdeeld zijn over de twee groepen. Verwacht je dat inzet het verschil tussen de groepen verklaart? Licht toe.
- c Licht toe dat het door de opzet van het onderzoek onwaarschijnlijk is dat een andere gemeenschappelijke factor het verschil tussen de cijfers van de groepen verklaart.
- d Is aan de eerste voorwaarde voor een causaal verband bij een samenhang voldaan? Licht toe.
- e Leg uit of je op basis van de beschreven onderzoeksopzet kunt zeggen dat het maken van opgaven over afwisselende onderwerpen een oorzaak is van het halen van hoge cijfers voor wiskunde.

Je kent vier associatiematen voor het bepalen van de grootte van het verschil tussen twee groepen. Deze staan ook op het formuleblad op bladzijde 271.

- Wat heb je nodig van twee groepen om de effectgrootte te kunnen berekenen?
- In België hebben verhoudingsgewijs meer huishoudens een hond dan in Nederland. Jarno wil de grootte van dit verschil bepalen. Hij heeft beschikking over de gegevens in de tabel. Je ziet dat in de steekproef van 595 Belgische huishoudens er 144 een hond hebben. Kan Jarno de effectgrootte gebruiken? Licht toe.

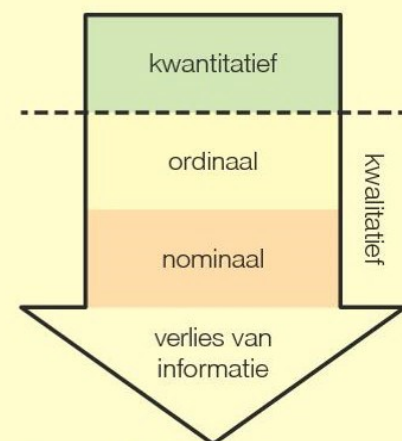
	hond		
	wel	niet	
België	144	451	595
Nederland	133	568	701
	277	1019	

Theorie B Groepen vergelijken

Op het formuleblad (zie bladzijde 271) bij het centraal examen staan vier associatiematen met de bijbehorende vuistregels. Je gebruikt deze om de grootte van het verschil tussen twee groepen te bepalen. Welke associatiemaat je gebruikt hangt af van de soort variabele en de beschikbare gegevens. Zie het volgende schema.



Het is soms mogelijk om een associatiemaat te gebruiken die bij een andere soort variabele hoort. In dat geval moet je de gegevens waarmee je begon meestal wel eerst bewerken, waardoor je een andere soort variabele krijgt. Je hebt gezien dat de omzetting naar een andere soort variabele slechts in één richting mogelijk is, omdat hierbij informatie verloren gaat. Zie figuur 10.15. Je kunt bijvoorbeeld kwantitatieve data van twee groepen zo bewerken dat je de phi-coëfficiënt kunt gebruiken. Maar je kunt bijvoorbeeld niet met ordinale variabelen beginnen en de effectgrootte gebruiken.



figuur 10.15

Voorbeeld

Aan havo 5-leerlingen van de scholen A en B is gevraagd of ze vinden dat Spaans in plaats van Duits gegeven zou moeten worden. Zie de tabel.

- a Wat is het meetniveau van de mening over Spaans als vak in plaats van Duits?
- b Tom wil het verschil in mening over Spaans als vak in plaats van Duits tussen school A en school B berekenen met behulp van de effectgrootte.

Leg uit waarom dat niet kan.

- c Onderzoek de grootte van het verschil in mening.

	school	
	A	B
helemaal niet mee eens	57	23
niet mee eens	28	41
mee eens	27	37
helemaal mee eens	16	21

Uitwerking

- a Het meetniveau is ordinaal.
- b De effectgrootte kan alleen bij een kwantitatieve variabele berekend worden, dus niet bij een ordinale variabele.
- c Ordinale variabele, meer dan 2 categorieën, dus max.Vcp.

	school A		school B		Vcp
	cum. freq.	cum. perc.	cum. freq.	cum. perc.	
helemaal niet mee eens	57	44,5%	23	18,9%	25,6%
niet mee eens	85	66,4%	64	52,5%	13,9%
mee eens	112	87,5%	101	82,8%	4,7%
helemaal mee eens	128	100%	122	100%	0%

max.Vcp = 25,6%

Omdat $20 < \text{max.Vcp} \leq 40$ is het verschil middelmatig.



Zie het voorbeeld.

Bij de variabele 'mening over Spaans in plaats van Duits als vak' is het ook mogelijk om de phi-coëfficiënt te berekenen. Daarvoor moeten de gegevens wel eerst bewerkt worden. Kies hiervoor de categorieën negatief en positief. Voor de categorie negatief tel je 'helemaal niet mee eens' en 'niet mee eens' bij elkaar op. Voor de categorie positief doe je hetzelfde met 'helemaal mee eens' en 'mee eens'.

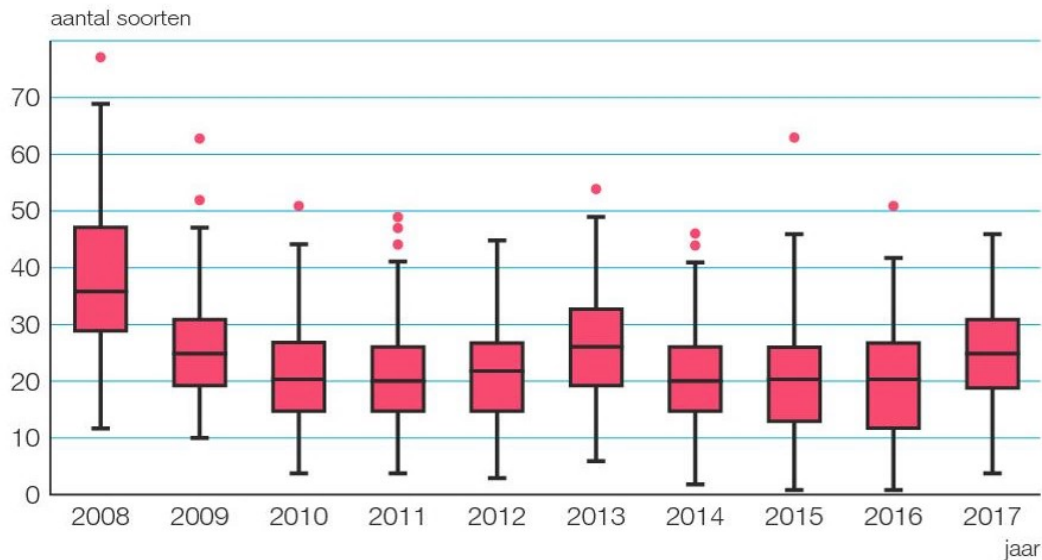
- a Neem de tabel hiernaast over en vul deze verder in.
- b Onderzoek de grootte van het verschil in mening.
- c Je hebt bij vraag b een andere conclusie getrokken dan in het voorbeeld. Welke conclusie is het beste te verdedigen? Licht toe.

	school	
	A	B
negatief	85	
positief		

18

Biologen hebben in Duitsland op een aantal locaties het aantal soorten insecten onderzocht in de periode 2008 tot en met 2017. Zie de figuur hieronder.

- Onderzoek met behulp van het formuleblad of het verschil in het aantal soorten insecten tussen 2008 en 2017 groot, middelmatig of gering is.
- Bepaal tussen welke tweetallen jaren het verschil groot is.
- Lisa beweert dat als je 2008 buiten beschouwing laat, de verschillen tussen alle jaren gering zijn. Ga na of ze gelijk heeft.



figuur 10.16

19

Werkt geweld om je doelen te bereiken als protestgroep? Om die vraag te beantwoorden werden 324 protestgroepen onderzocht die actief waren tussen 1900 en 2006. Zie de tabel hiernaast. In totaal waren 108 protestgroepen gewelddoos.

	gewelddoos	gewelddadig
mislukking	21%	62%
gedeeltelijk succes	25%	13%
succes	54%	25%

- Licht toe dat er 58 protestgroepen zowel gewelddoos als succesvol waren.
- Onderzoek met behulp van het formuleblad of het verschil tussen gewelddoze en gewelddadige protestgroepen groot, middelmatig of gering is.
- Bewerk de gegevens in de tabel zodat er twee categorieën zijn. Voeg daartoe 'gedeeltelijk succes' en 'succes' samen tot één categorie. Bereken vervolgens met behulp van het formuleblad of het verschil tussen gewelddoze en gewelddadige protestgroepen groot, middelmatig of gering is.



20
□ ⊙ *

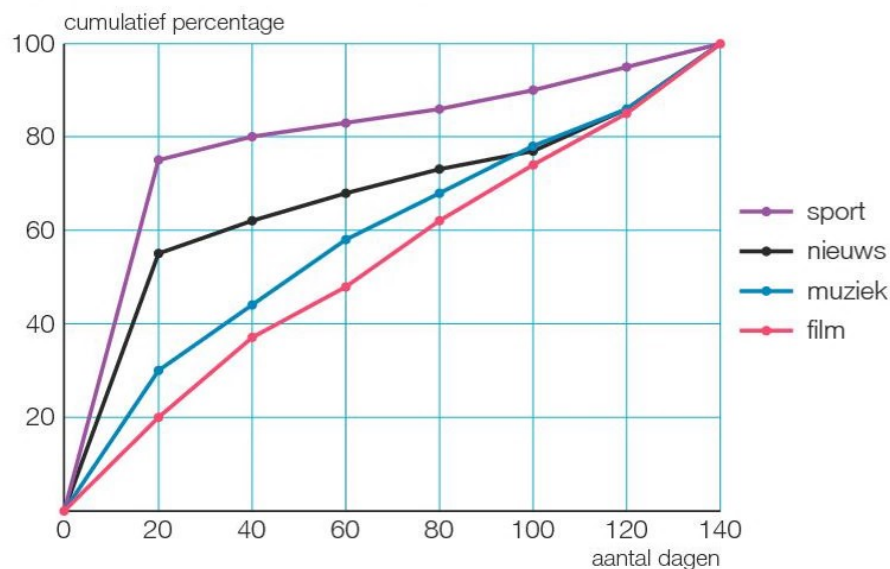
Hardlopers kunnen gebruikmaken van een exoskelet. Dat is een omhulling met kleine motortjes, die moet zorgen dat het lichaam minder energie nodig heeft om voort te bewegen. Onderzoekers lieten deelnemers met en zonder exoskelet op een lopende band 9 km/uur hardlopen. Hierbij werd gemeten hoeveel energie verbruikt werd. Omdat het gewicht van deelnemers geen rol mocht spelen is de verbruikte energie uitgedrukt in watt per kilogram. Zie de tabel hieronder. Bereken met behulp van het formuleblad of het verschil in energieverbruik met en zonder exoskelet groot, middelmatig of gering is.



	met exoskelet	zonder exoskelet
gemiddeld energieverbruik in watt per kg	8,70	9,57
standaardafwijking in watt per kg	0,83	0,79

21
□ ⊙ *

Hoe snel een filmpje op YouTube aan populariteit verliest, hangt af van het onderwerp. Onderzoekers hebben dit verloop van het aantal kijkers onderzocht voor 50 000 YouTube-video's over een periode van 140 dagen. De video's zijn ingedeeld in vier categorieën. In de figuur hieronder staan relatieve cumulatieve frequentiepolygonen van deze gegevens. De eerste klasse is $[0, 20)$.



figuur 10.17

- Verklaar het opvallende verschil in het verloop van het aantal kijkers tussen video's over sport en over film.
- Bepaal aan de hand van de figuur en het formuleblad of het verschil tussen muziekvideo's en elk van de andere categorieën groot, middelmatig of gering is.

A22

Er wordt vaak gedacht dat oorlog voeren een bezigheid is van mannen. Onderzoekers hebben deze aanname onderzocht door na te gaan of koningen relatief vaker oorlog voerden dan koninginnen. Daartoe bekeken ze in de periode 1480-1913 in verschillende koninkrijken per jaar wie er aan de macht was en of er oorlog werd gevoerd. In totaal werden er 3586 van zulke jaren bekeken. Hiervan was 596 jaar een koningin aan de macht. Verder werd in 889 jaren oorlog gevoerd, waarvan in 154 jaren een koningin aan de macht was.



- Vul de tabel in.
- Onderzoek of koningen of koninginnen relatief vaker oorlog voerden.
- Bepaal met behulp van het formuleblad of het verschil groot, middelmatig of gering is.

	oorlogsjaren	vredesjaren	
koningen			
koninginnen			

R23

In de tabel hiernaast zie je welke soort wiskunde is gekozen door jongens en meisjes op het Spinoza College. Amar berekent de phi-coëfficiënt en komt tot de conclusie dat er geen verschil is tussen jongens en meisjes wat betreft de keuze voor wiskunde A of B.

	jongens	meisjes
wiskunde A	20	24
wiskunde B	30	36

- Hoe komt hij tot deze conclusie?
- Bereken het percentage leerlingen dat wiskunde A kiest voor zowel de jongens als de meisjes. Komt dit overeen met de conclusie van Amar?

Het verschil tussen de groepen zou maximaal zijn als alle jongens wiskunde A hadden gekozen en alle meisjes wiskunde B. Of andersom.

- Bereken in beide gevallen de phi-coëfficiënt. Welke conclusie trek je over de minimale en maximale waarde van de phi-coëfficiënt?



Om de mate van geluk tussen landen te vergelijken wordt gebruikgemaakt van vragenlijsten. Hierop geven inwoners van een land aan hoe gelukkig ze zijn met een geheel rapportcijfer. In de tabel zijn de cumulatieve percentages gegeven van de scores van duizend inwoners van China en duizend inwoners van de VS in 2019.

geluk	1	2	3	4	5	6	7	8	9	10
China	0%	1%	2%	5%	17%	40%	65%	89%	98%	100%
VS	2%	2%	4%	8%	22%	34%	55%	80%	91%	100%

- Pascal denkt dat in China 11% van de mensen hun geluk een 8 of hoger geeft. Volgens Daniek moet dat 35% zijn. Wie heeft er gelijk? Licht toe.
- Bepaal met behulp van het formuleblad of het verschil in geluk tussen China en de VS groot, middelmatig of gering is.
- Als mensen hun geluk een 5 of lager geven, noemen we dat 'ongelukkig'. Bij een 6 of hoger, noemen we dat 'gelukkig'. Vul de kruistabel in en onderzoek met behulp van het formuleblad of het verschil in geluk groot, middelmatig of gering is.



	ongelukkig	gelukkig	
China			
VS			
			2000

In deze opgave kijken we naar de formule van de effectgrootte

$$E = \frac{\bar{X}_1 - \bar{X}_2}{\frac{1}{2}(S_1 + S_2)} \text{ met } \bar{X}_1 \geq \bar{X}_2.$$

Beredeneer hoe uit de formule van E volgt dat

- de effectgrootte altijd groter dan of gelijk is aan 0
- als $\bar{X}_1$ toeneemt en $\bar{X}_2$, S_1 en S_2 gelijk blijven, de effectgrootte toeneemt
- als S_1 toeneemt en $\bar{X}_1$, $\bar{X}_2$ en S_2 gelijk blijven, de effectgrootte afneemt
- als $\bar{X}_1$ toeneemt en S_1 afneemt en $\bar{X}_2$ en S_2 en gelijk blijven, de effectgrootte toeneemt
- als S_1 en S_2 beide verdubbelen en $\bar{X}_1$ en $\bar{X}_2$ gelijk blijven, de effectgrootte halveert.

Door de verzuring van oceanen sterven koraalriffen in hoog tempo af. De afbraak van koraal kan geremd worden door er vissen naartoe te lokken. Wetenschappers hebben daarom onderzocht of er meer vissen naar een koraalrif zwemmen als er geluiden van gezonde koraalriffen worden afgespeeld. Ze selecteerden hiervoor 33 aangetaste koraalriffen voor de kust van Australië, die ze aselekt in drie groepen verdeelden. In groep 1 werd het geluid van gezonde koraalriffen afgespeeld met onderwaterluidsprekers, in groep 2 stonden alleen de luidsprekers zonder dat er geluid uit kwam en in groep 3 werd niets gedaan.



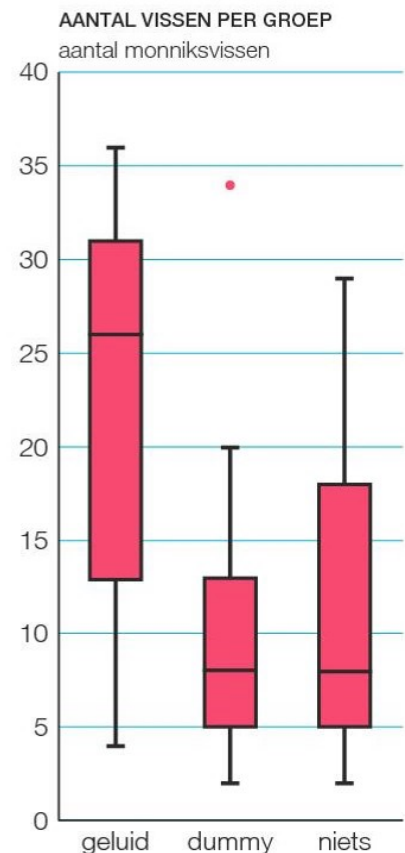
- a** Leg uit waarom ervoor gekozen is om in groep 2 luidsprekers te plaatsen waar geen geluid uit kwam.

Gedurende 40 dagen is er meerdere keren per dag gemeten hoeveel monniksvissen zich in de koraalriffen bevinden. De resultaten van één dag van het onderzoek staan in de figuur en de tabel. De term ‘dummy’ verwijst naar groep 2, waar een luidspreker stond die geen geluid maakte. Je ziet dat voor de groep ‘dummy’ een punt buiten de boxplot ligt. Dit punt ligt meer dan anderhalf keer de interkwartielafstand boven het derde kwartiel.

- b** Toon dit aan met een berekening.
c Saffa beweert dat er in groep 3 meer metingen zijn met meer dan 18 vissen dan metingen met minder dan 5 vissen.
 Leg uit of ze gelijk heeft.
d Onderzoek met het formuleblad of het verschil tussen groep 1 en elk van de beide andere groepen groot, middelmatig of gering is. Doe dat op twee verschillende manieren.

Ook voor de overige dagen van het onderzoek zijn er in de groep ‘geluid’ meer vissen geteld dan in de andere twee groepen.

- e** Leg uit of je op basis van de onderzoeksopzet en de resultaten mag concluderen dat geluid een oorzaak is van de aanwezigheid van meer monniksvissen.

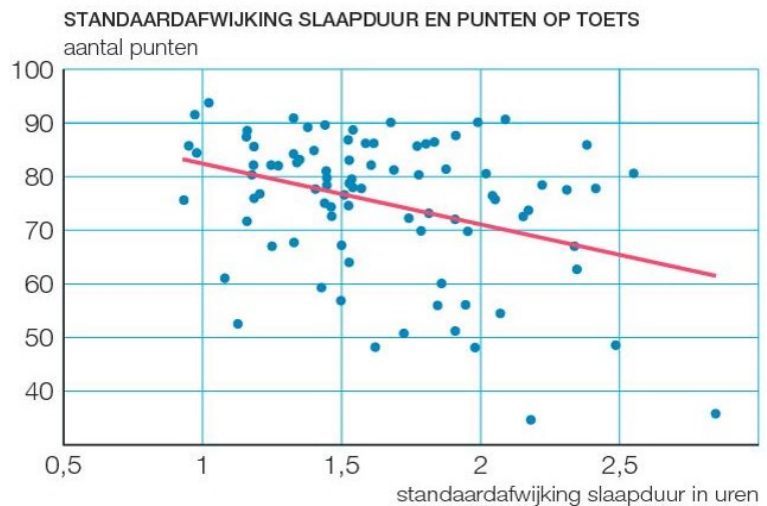
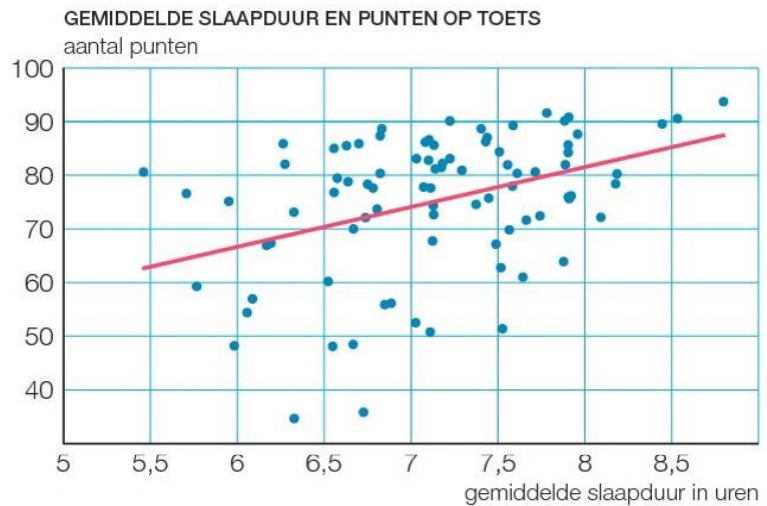


figuur 10.18

	$\bar{X}$	S
geluid	21,6	9,8
dummy	11,3	9,0
niets	12,1	8,8

Haal je betere cijfers als je langer slaapt? Om dat te onderzoeken kregen 100 studenten aan de Amerikaanse universiteit MIT drie maanden voor een scheikundetoets een fitbit mee. Dat is een horloge dat onder andere bijhoudt hoelang je slaapt. De resultaten staan in de figuren hiernaast. De rode lijnen geven de samenhang weer.

- a Heeft de student met de hoogste score ook de langste gemiddelde slaapduur? Sliep de student met de laagste score gemiddeld ook het kortst?
- b Leg bij de volgende uitspraken uit of ze volgen uit de figuren.
 - I Studenten met een gemiddelde slaapduur van minstens acht uur haalden tenminste 70 punten.
 - II Studenten die vaker ongeveer even lang slapen, scoren beter.



figuur 10.19

Er was geen samenhang tussen het aantal uren slaap op de nacht voor de toets en de score op de toets.

Om ook het effect van slapen op lange termijn vast te stellen, werden de studenten in twee groepen ingedeeld. Studenten links van de mediaan werden 'slechte slapers' genoemd en studenten rechts van de mediaan 'goede slapers'. De goede slapers scoorden gemiddeld 6,57 punten beter op de toets dan de slechte slapers. De standaardafwijking van de slechte slapers was 11,01 punten en die van de goede slapers 13,71 punten.

- c Bereken met behulp van het formuleblad hoe groot het verschil in score tussen goede en slechte slapers is.
- d Kun je concluderen dat voor deze studenten een lange slaapduur de oorzaak is van goede scores? Licht toe.

Koeien worden soms duizenden keren per dag gestoken door vliegen, waardoor de koeien gewicht verliezen en minder melk geven. Dit levert de veehouderij wereldwijd een kostenpost van miljarden euro's op.

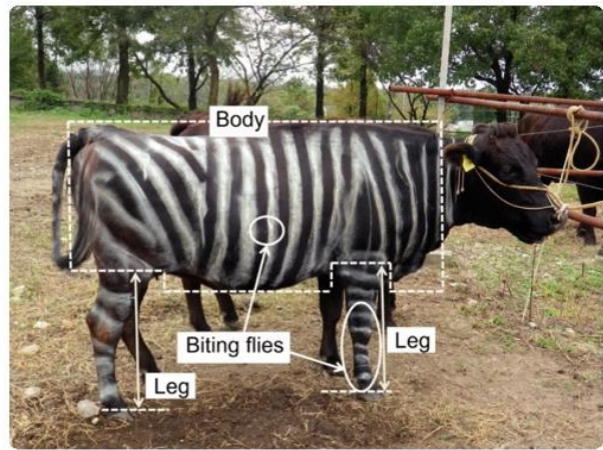
Bekend is dat zebra's minder last hebben van vliegen. Bij een onderzoek in Japan werden daarom witte strepen op zes Zwarte Wagyu koeien geverfd. Deze koeien werden vergeleken met zes Zwarte Wagyu koeien zonder strepen. De

wetenschappers telden op basis van foto's en videobeelden hoe vaak koeien reageerden op vliegen (schudden van de kop, stampen, flapperen met de oren, enzovoort) en hoeveel vliegen op de foto's te zien waren.

- a** Noem twee kwalitatieve variabelen die een rol spelen in dit onderzoek.

Van alle koeien werd elke 30 minuten een foto gemaakt om het aantal vliegen te tellen. Deze gegevens staan in de tabel hiernaast.

- b** Bereken met behulp van het formuleblad de grootte van het verschil tussen het totale aantal vliegen bij de twee groepen koeien. Gebruik dat de standaardafwijking van het totale aantal vliegen voor gestreepte koeien 35,2 is en voor zwarte koeien 68,6.



	gemiddeld aantal vliegen per foto	
	lichaam	poten
zwarte koeien	42,1	86,7
gestreepte koeien	15,3	40,2

De foto's van de koeien werden van slechts één zijde genomen. Het aantal vliegen is dus twee keer zo groot als op een foto. Verder is bekend dat de vliegen gemiddeld elke vier minuten een keer steken.

De Japanse boer Joji vindt het verven van zijn Zwarte Wagyu koeien veel gedoe, maar weet dat er een flink verminderde melkproductie is als zijn koeien gemiddeld meer dan 2000 keer per uur op lichaam of poten worden gestoken. Dat wil hij voorkomen.

- c** Zou je hem op basis van de onderzoeksresultaten aanraden zijn koeien te verven? Licht toe met een berekening.

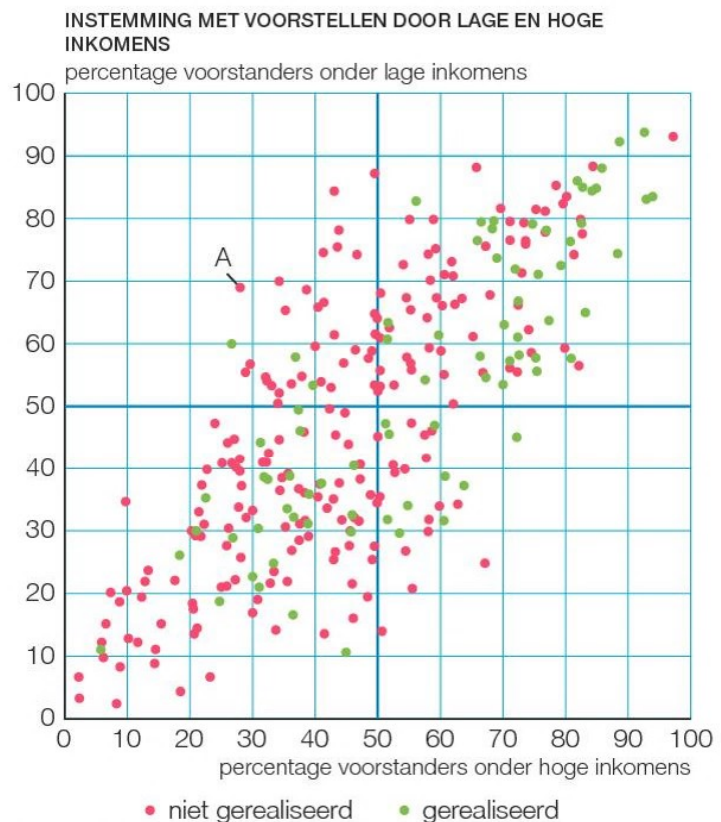
Bij het onderzoek was ook een derde groep Zwarte Wagyu koeien opgenomen die zwart waren geverfd.

- d** Waarom hebben de wetenschappers dit gedaan?

Een onderzoeker trekt de conclusie dat gestreepte koeien minder vaak gestoken worden. Door de gekozen onderzoeksopzet is deze conclusie echter te algemeen.

- e** Noem hiervoor twee aspecten uit de gekozen onderzoeksopzet.

Veel Nederlanders met een laag inkomen hebben het gevoel dat hun mening niet telt. Onderzoekers wilden weten of ze gelijk hebben. Ze verzamelden opiniepeilingen uit de periode 1979 tot en met 2012, waarin Nederlanders werd gevraagd wat ze van een voorstel vinden. Er werden voorstellen voorgelegd zoals: 'Er moet een referendum over de afschaffing van de euro komen.' Vervolgens maakten ze onderscheid tussen de meningen van mensen met een laag inkomen en een hoog inkomen. Elk punt in het spreidingsdiagram is een voorstel. Op de assen is het percentage dat een voorstel steunt voor beide groepen af te lezen.



figuur 10.20

- a** Voorstel A in het spreidingsdiagram ging over de verhoging van het minimumloon.
 Hoeveel procent van de mensen met een hoog inkomen was voorstander van dit voorstel? En hoeveel procent bij de lage inkomens?

Het spreidingsdiagram is in vier vlakken verdeeld. In het vlak rechtsboven staan voorstellen waar zowel de lage als de hoge inkomens in meerderheid voorstander van zijn.

- b** Geef de betekenis van het vlak linksboven.

De kleur van de punten geeft aan of een voorstel gerealiseerd is. Rood betekent dat het voorstel niet is gerealiseerd en groen betekent dat het wel is gerealiseerd. In het vlak linksboven staan 38 voorstellen.

- c** Bepaal hoeveel procent van deze voorstellen gerealiseerd is.

Van de voorstellen in het vlak rechtsonder zijn 10 van de 30 voorstellen gerealiseerd. Er is een verschil tussen de vlakken rechtsonder en linksboven wat betreft het realiseren van voorstellen.

- d** Toon met behulp van percentages aan dat er verschil is en geef de praktische betekenis van dit verschil.
e Bepaal met behulp van het formuleblad of dit verschil groot, middelmatig of gering is.

Tarik stelt vast dat de voorstellen in het vlak rechtsboven vaak zijn gerealiseerd en dat deze ook door mensen met lage inkomens worden gesteund. Volgens hem toont dit aan dat er wel naar hen geluisterd wordt.

- f** Waarom kan Tarik deze conclusie niet trekken?

Terugblik

Samenhang en causaliteit

Om bij twee variabelen vast te stellen of er bij een samenhang ook sprake is van causaliteit onderzoek je twee voorwaarden.

1. De oorzaak moet voorafgaan aan het gevolg.
2. Er mogen geen andere variabelen zijn die de samenhang verklaren.

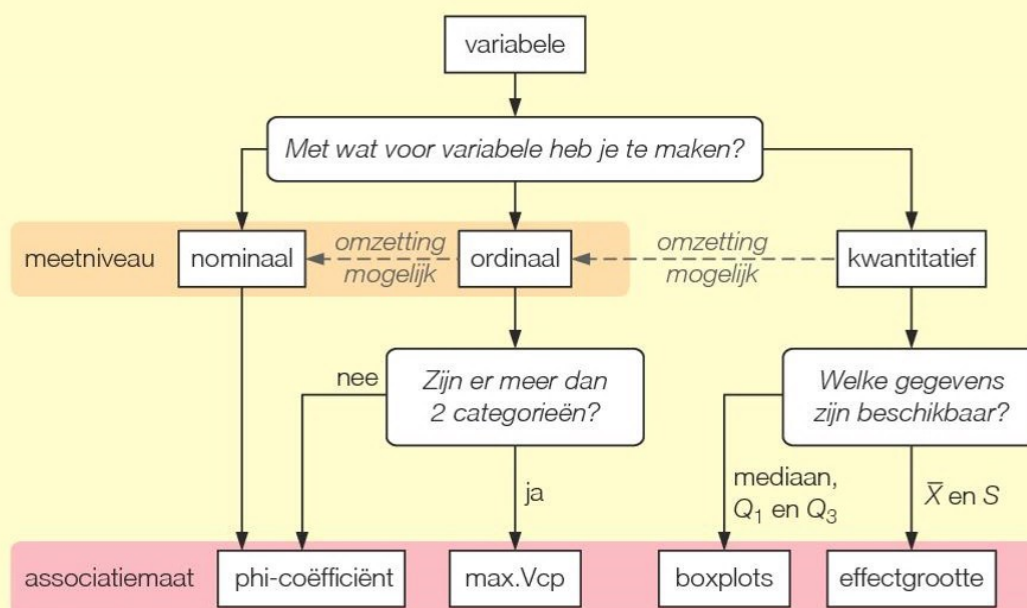
Wanneer uit een onderzoek blijkt dat mensen die meer zitten vaker diabetes hebben, dan wil dat nog niet zeggen dat veel zitten een oorzaak is van diabetes.

Het zou immers kunnen dat mensen met diabetes minder mobiel zijn en dus door hun ziekte meer zitten (voorwaarde 1). Het zou ook kunnen dat de variabele overgewicht de samenhang verklaart (voorwaarde 2).

Mensen met overgewicht zitten wellicht meer omdat ze moeilijker bewegen en krijgen mogelijk ook vaker diabetes. Een variabele die de samenhang verklaart, is een gemeenschappelijke factor.

Groepen vergelijken

Om de grootte van een verschil tussen twee groepen te bepalen, gebruik je een geschikte associatiemaat. Gebruik voor de keuze van een associatiemaat het onderstaande schema.



Welke associatiemaat je kunt gebruiken, hangt af van het soort variabele waar je mee te maken hebt en welke gegevens beschikbaar zijn. Het is soms mogelijk om een associatiemaat te gebruiken die bij een andere soort variabele hoort. In dat geval moet je de gegevens waarmee je begon meestal wel eerst bewerken.

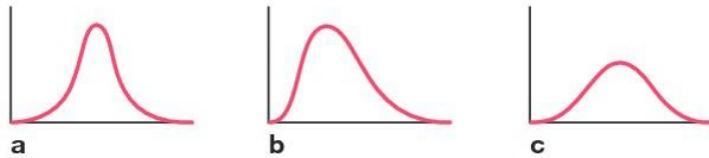
Om het verschil tussen groepen in een associatiemaat uit te drukken gebruik je het formuleblad op bladzijde 271. Hier vind je ook de vuistregels waarmee je aangeeft of het verschil gering, middelmatig of groot is. Dit formuleblad staat ook op het voorblad bij het centraal examen.

10.3 De normale verdeling en betrouwbaarheidsintervallen



Bij welke van de onderstaande figuren hoort geen normale verdeling?

Licht toe.



figuur 10.21

Theorie A De normale verdeling bij waarnemingsgetallen

Wanneer afwijkingen van het gemiddelde willekeurig optreden, ontstaat een normale verdeling. Dit proces is in figuur 10.22 nog eens geïllustreerd door het bord van Galton. In dit bord komen de blauwe balletjes op zwarte pinnen terecht, waarna ze door toeval steeds naar links of rechts vallen. De balletjes onderaan het bord vormen zo bij benadering een normale verdeling.

Je bent de normale verdeling tegengekomen bij waarnemingsgetallen en bij steekproeven. In de volgende opgaven komen situaties uit het dagelijks leven voor waarbij door toeval afwijkingen kunnen optreden.

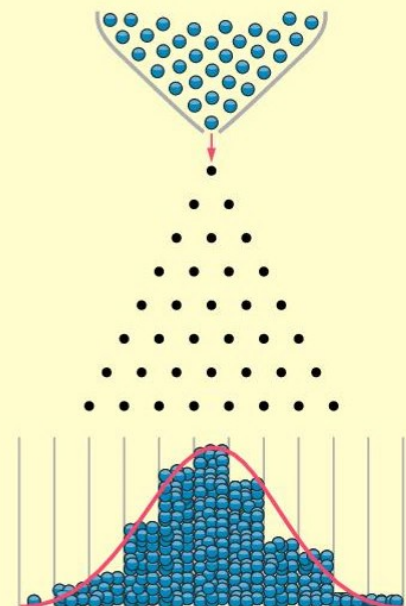
Een blauw balletje in het bord van Galton stelt hier telkens een waarnemingsgetal voor. Denk bijvoorbeeld aan de lengte van volwassen Nederlandse vrouwen.

Een blauw balletje stelt dan de lengte van een volwassen Nederlandse vrouw voor.

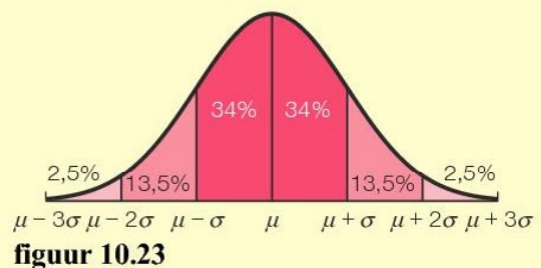
De normale verdeling bij steekproeven komt aan bod in theorie B.

Je kent de vuistregels bij de normale verdeling:

- ongeveer 68% van de waarnemingsgetallen ligt tussen $\mu - \sigma$ en $\mu + \sigma$
- ongeveer 95% van de waarnemingsgetallen ligt tussen $\mu - 2\sigma$ en $\mu + 2\sigma$
- bijna 100% van de waarnemingsgetallen ligt tussen $\mu - 3\sigma$ en $\mu + 3\sigma$.



figuur 10.22



figuur 10.23

Bij berekeningen met de vuistregels mag je de woorden ongeveer en bijna weglaten. Zo mag je bij een normale verdeling met $\mu = 49$ en $\sigma = 2$ zeggen dat 16% van de waarnemingsgetallen kleiner dan 47 is.

- 31** Het gewicht van een groep van 8000 producten is normaal verdeeld met gemiddelde 500 gram en standaardafwijking 15 gram.
- Welk percentage van de producten weegt minder dan 485 gram?
 - Hoeveel producten wegen meer dan 530 gram?
 - Wat weet je van de gewichten van de middelste 95% van deze groep producten?

Van een andere partij van 3940 producten is niet bekend of het gewicht van de producten normaal verdeeld is. Wel is bepaald dat het gemiddelde 400 gram is en de standaardafwijking 15 gram. Het blijkt dat 97 van deze producten lichter dan 370 gram zijn. Dit is ongeveer gelijk aan wat je op basis van een normale verdeling zou verwachten.

- Toon dat aan.

- 32** Hiernaast zie je een tabel met de lengte van mannelijke deelnemers aan de Olympische Spelen van Rio in 2016. Met de percentages uit de tabel en het gegeven dat er 190 mannelijke deelnemers langer dan 2,02 m waren, kun je berekenen hoeveel mannen deelnamen aan de spelen in Rio.

- Bereken dit aantal en rond af op honderdtallen.

De lengtes zijn bij benadering normaal verdeeld. De aantallen wijken iets af ten opzichte van wat je zou verwachten op basis van een normale verdeling met een gemiddelde van 1,82 meter en een standaardafwijking van 10 cm.

- Laat zien dat de percentages in de tabel weinig verschillen van de vuistregels bij de normale verdeling.

LENGTE MANNEN RIO

lengte in m	percentage
<1,52	0,1%
1,52-1,62	1,9%
1,62-1,72	12,7%
1,72-1,82	33,2%
1,82-1,92	35,1%
1,92-2,02	13,8%
2,02-2,12	2,9%
>2,12	0,1%

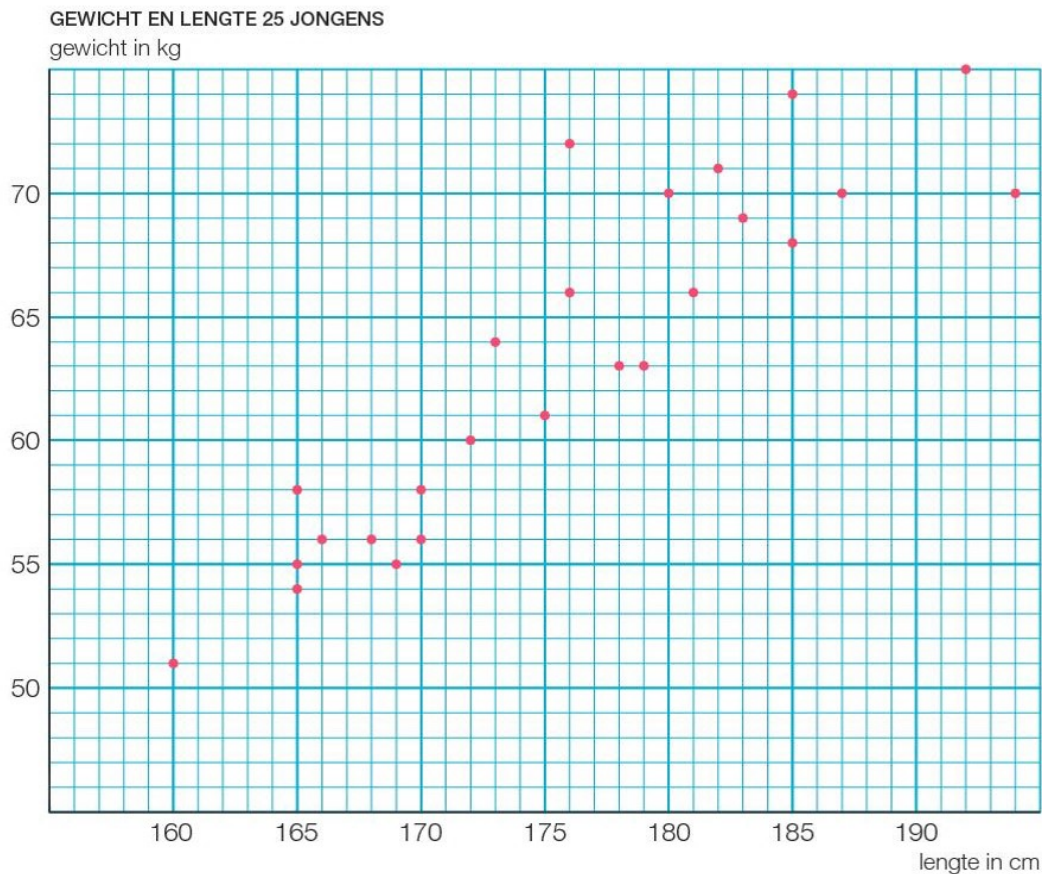
- 33** Van een partij olijven is de diameter normaal verdeeld met een gemiddelde van 18 mm. Internationale kwaliteitseisen stellen dat 95% van de olijven in de partij een diameter tussen 17 en 19 mm moet hebben.

- Wat zou op basis van deze informatie de standaardafwijking van de diameter bij deze olijven moeten zijn?
- Een onderzoeker wil nagaan of grotere olijven over het algemeen sneller bederven.

Welk type diagram is hiervoor geschikt en welke informatie is nodig om dit diagram te maken?



Bij een onderzoek is van 25 jongens het gewicht en de lengte gemeten. De resultaten staan in het spreidingsdiagram.



figuur 10.24

- Maarten wil van zowel de lengte als het gewicht de modus bepalen. Lukt dit? Licht toe.
- Heeft de jongen met de mediane lengte ook het mediane gewicht? Licht toe.
- De onderzoeker maakt op basis van de lengte vijf even grote groepen, oplopend van de 20% kortste jongens tot en met de 20% langste jongens. Bij welke van deze groepen is de spreidingsbreedte het grootst?
- Hoeveel procent van de jongens hoort zowel bij de langste 12% als bij de zwaarste 12% van de onderzoeksgroep?

De onderzoeker vraag zich af of de lengte van de deelnemende jongens bij benadering normaal verdeeld is. Zij berekent dat hun gemiddelde lengte 176 cm is met een standaardafwijking van 9 cm.

Om een van de vuistregels te controleren bekijkt ze het aantal jongens met een lengte in de klasse [167, 185].

- Welke vuistregel controleert ze hiermee en wat is haar conclusie?

Hoe harder de muziek, hoe mooier het klinkt, denken veel producenten van muziek. Ze willen daarom dat de liedjes van hun artiesten zo hard mogelijk worden afgespeeld. Ze kunnen natuurlijk niet bepalen hoe hard jij het volume zet. Maar muziek heeft ook een digitale hardheid, die producenten wel kunnen aanpassen. Deze digitale hardheid wordt uitgedrukt in dBFS. Het hardste geluid dat mogelijk is, heeft een waarde van 0 dBFS, terwijl minder harde geluiden negatieve waarden hebben. Om zoveel mogelijk tonen in de buurt van 0 dBFS te krijgen, veranderen producenten de spreiding van de digitale hardheid van hun liedjes en stellen ze vervolgens de hardste toon in op 0 dBFS.

a Maken ze daartoe de spreiding groter of kleiner? Licht toe.

In de figuur hiernaast zie je voor een aantal jaren de verdeling van de gemiddelde digitale hardheid in dBFS van duizenden liedjes. Een onderzoeker concludeert dat de hardheid in geen enkel jaar normaal verdeeld is.

b Geef een mogelijke statistische redenering die de onderzoeker hiervoor gebruikt kan hebben.

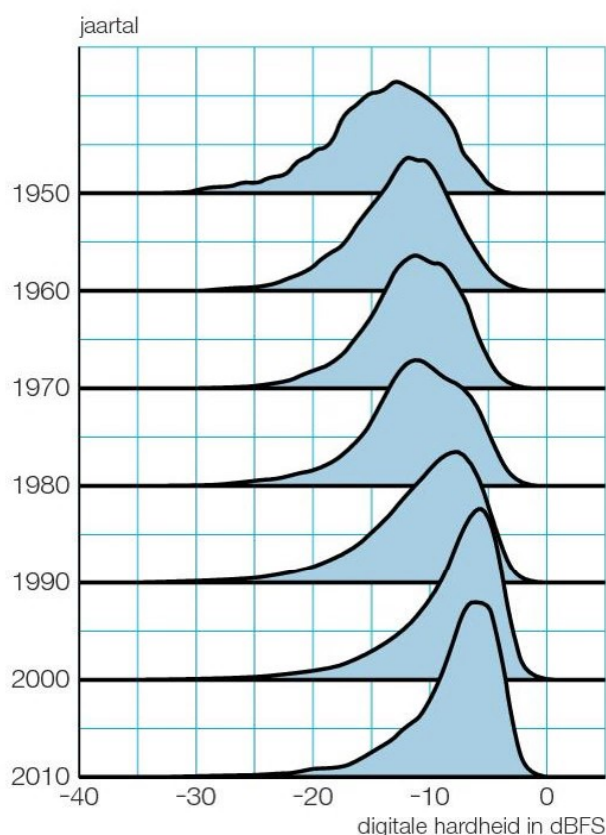
De ontwikkeling van de digitale hardheid van muziek kan beschreven worden met de formule $H = -14 + 0,067t$. Hierin is H de gemiddelde digitale hardheid in dBFS en t de tijd in jaren met $t = 0$ in 1950.

c Leg met behulp van figuur 10.25 uit of de keuze voor het gemiddelde als centrummaat een overschatting of een onderschatting van de digitale hardheid geeft.

d Bereken in welk jaar volgens de formule de gemiddelde digitale hardheid 0 dBFS zou zijn en beschrijf hoe liedjes in dat jaar zouden klinken.

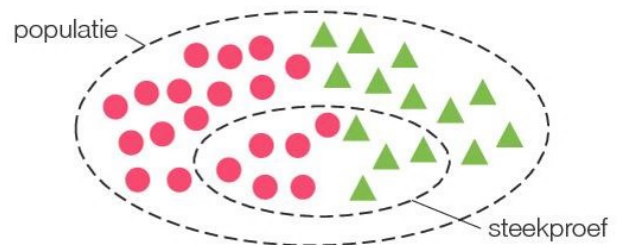
Zover zal het niet komen, want sinds 2010 neemt de gemiddelde digitale hardheid weer iets af en geldt de formule voor H dus niet meer. Een aanleiding hiervoor was het album *Death Magnetic* van Metallica. Het verschil tussen de hardste en zachtste tonen was hier door de producent zo klein gemaakt dat volgens sommige fans de muziek vervormd was. Streamingdiensten zoals Spotify hebben naar aanleiding van deze kritiek hun beleid aangepast. Ze berekenen sindsdien de gemiddelde digitale hardheid van een liedje en spelen dit af op -14 dBFS.

e Leg uit waarom bij liedjes van het album *Death Magnetic* van Metallica door dit beleid de hardste toon minder hard klinkt dan bij liedjes met een grote spreiding.



figuur 10.25

In de figuur hiernaast zie je welke steekproef uit een populatie is genomen. Toon met een berekening aan dat de proportie rode bolletjes in de steekproef verschilt van die in de populatie.



figuur 10.26

Theorie B De normale verdeling bij steekproeven

De populatie is de totale groep waarop een onderzoek zich richt. Het is soms onwenselijk en bijna altijd te veel werk om de hele populatie te onderzoeken. Daarom wordt een steekproef genomen. Als een steekproef aselekt is, heeft elk element in de populatie een even grote kans om in de steekproef te komen.

Als je een kwantitatieve variabele hebt onderzocht, kun je een gemiddelde berekenen. Het gemiddelde van een populatie geven we aan met μ en het gemiddelde van een steekproef met $\bar{X}$.

Doordat een aselechte steekproef door toeval tot stand komt, wijkt het steekproefgemiddelde door toeval af van het populatiegemiddelde. Als er, in gedachten, heel veel aselechte steekproeven worden genomen, ontstaat een normale verdeling van heel veel steekproefgemiddeldes. Deze verdeling heet de steekproevenverdeling. In het bord van Galton stelt een blauw balletje in deze situatie een steekproefgemiddelde voor.

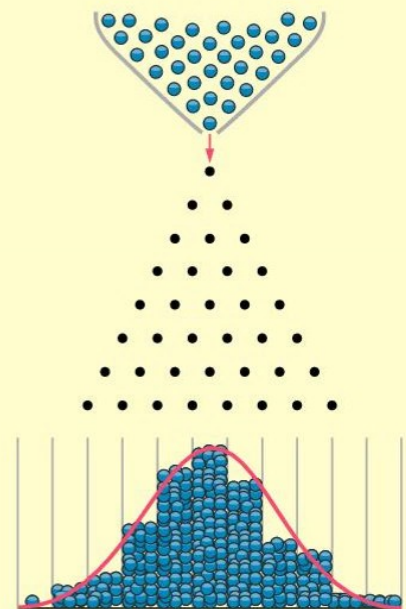
Als je een kwalitatieve variabele hebt onderzocht, kun je geen gemiddelde berekenen. Je kunt wel de proportie van het onderzochte kenmerk berekenen. De populatieproportie p bereken je met

$$p = \frac{\text{aantal elementen met het kenmerk in de populatie}}{\text{totaal aantal elementen in de populatie}}.$$

De steekproefproportie $\hat{p}$ bereken je met

$$\hat{p} = \frac{\text{aantal elementen met het kenmerk in de steekproef}}{\text{totaal aantal elementen in de steekproef}}.$$

Ook bij heel veel steekproefproporties uit aselechte steekproeven, ontstaat een normale verdeling. Ook deze verdeling heet een steekproevenverdeling. In het bord van Galton stelt een blauw balletje in deze situatie een steekproefproportie voor.



figuur 10.27

R37
☐◎*

In een ziekenhuis liggen 1000 mensen, waarvan 50 een levensbedreigende ziekte hebben. In een steekproef van 80 kinderen op de kinderafdeling blijken er 4 een levensbedreigende ziekte te hebben.

- a Bereken de populatieproportie en de steekproefproportie.
- b Is de steekproef aselekt? Licht toe.

38
☐◎*

In de onderstaande gevallen is er geen sprake van een aselekte steekproef.

- a Met PISA-scores worden de schoolprestaties van Nederlandse 15-jarige scholieren gemeten. PISA selecteert hiertoe aselekt scholen voor een toets. In 2018 deed slechts 61% van de geselecteerde scholen mee. Verwacht je dat de gemiddelde score in de steekproef lager of hoger is dan in de populatie? Licht toe.
- b In 2018 onderzocht de Nederlandse Voedsel en Warenautoriteit (NVWA) 93 boerenbedrijven die werden verdacht van fraude bij de registratie van dieren. Hiervan bleken 45 inderdaad fraude te hebben gepleegd. Verwacht je dat de steekproefproportie lager of hoger is dan de populatieproportie van alle boerenbedrijven? Licht toe.



39
☐◎*

Er wordt vaak beweerd dat 97% van de klimaatwetenschappers denkt dat opwarming van de aarde door de mens wordt veroorzaakt. Een van de bronnen voor dit percentage is een steekproef van 79 klimaatwetenschappers, waarvan 77 de mens als oorzaak voor de opwarming zien.

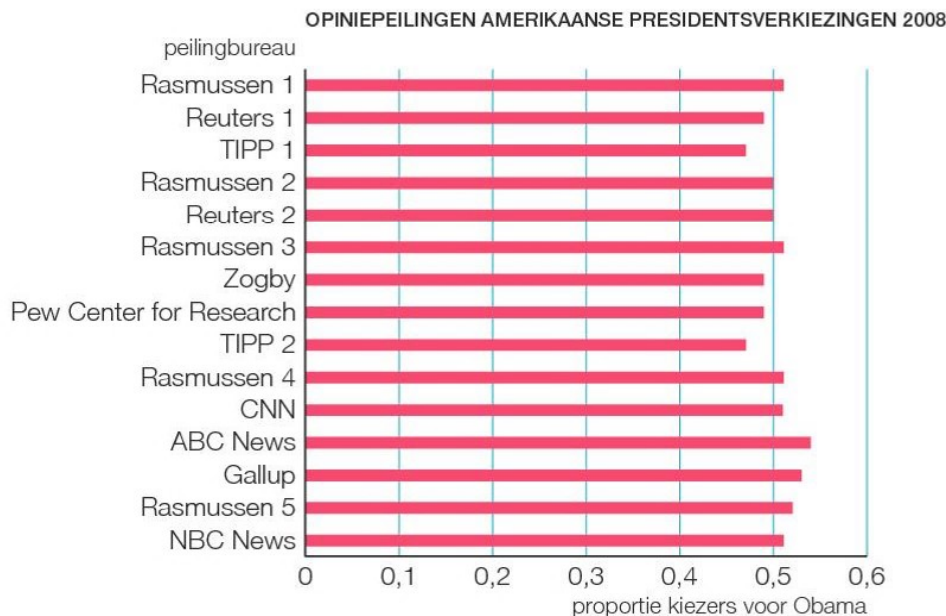
- a Benoem wat in deze peiling de populatie en wat de steekproef is.

Voor een andere studie zijn geen wetenschappers ondervraagd, maar zijn de samenvattingen gelezen van alle 12000 wetenschappelijke artikelen over klimaatverandering uit de periode 1991-2011. Van deze artikelen gaf twee derde geen mening over de bewering dat de mens bijdraagt aan klimaatverandering. Uit de overige artikelen bleek wel een mening en hiervan stelde 97,1% dat de mens bijdraagt aan klimaatverandering.

- b Bram ziet 0,971 als een populatieproportie. Wat is volgens Bram de populatie?
- c Sabine ziet 0,971 als een steekproefproportie bij een steekproef uit de populatie van alle 12000 artikelen. Leg uit of deze steekproef aselekt is.



In de week voorafgaand aan de Amerikaanse verkiezingen van 2008 hebben 15 peilingbureaus een steekproef genomen om het stemgedrag van Amerikanen die gaan stemmen te peilen. In figuur 10.28 zie je voor elke steekproef de proportie kiezers die op Barack Obama zegt te gaan stemmen.



figuur 10.28

- a** Leg uit waarom de populatie voor de onderzoekers niet bestond uit Amerikanen die mogen stemmen.
 Wat was voor hen wel de populatie?
- b** Aan de steekproef van CNN deden 1000 mensen mee.
 Bereken hoeveel daarvan hebben aangegeven op Obama te gaan stemmen.
- c** De verkiezingen lieten zien dat de populatieproportie 0,51 was.
 Leg uit waarom de verkiezingsuitslag gelijk is aan de populatieproportie.

We gaan ervan uit dat de 15 steekproeven dezelfde omvang hebben, dat ze aselekt zijn genomen en dat de stemvoorkeur tijdens deze week niet veranderd is.

- d** Benthe denkt dat hieruit volgt dat alle steekproeven dezelfde steekproefproportie $\hat{p}$ zouden moeten hebben.
 Licht toe waarom dat niet zo is.

Als er heel veel aselekte steekproeven met $n = 1000$ worden genomen bij $p = 0,51$ ontstaat een steekproevenverdeling die normaal verdeeld is met standaardafwijking 0,016.

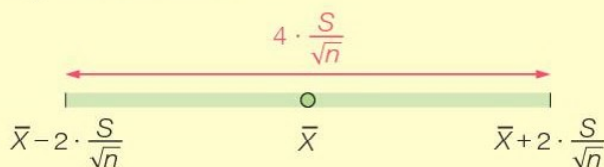
- e** Wat weet je van de middelste 95% van de steekproefproporties van deze steekproeven?
- f** Wat weet je van de 5% steekproefproporties die het meeste afwijken van de populatieproportie?
- g** Twee van de 15 steekproeven uit figuur 10.28 liggen op een afstand groter dan 2σ van de populatieproportie.
 Welke steekproeven zijn dat?

Gegeven is het 95%-betrouwbaarheidsinterval $0,3 \pm 2 \cdot 0,02$.
 Wat is de breedte van dit 95%-betrouwbaarheidsinterval?

Theorie C Betrouwbaarheidsintervallen

Bij een onderzoek wordt doorgaans maar één steekproef genomen. Op basis daarvan wil je iets zeggen over de populatie. Als je bijvoorbeeld het steekproefgemiddelde $\bar{X}$ berekent, dan is dat meestal niet gelijk aan het populatiegemiddelde μ . Het is zelfs onbekend hoe ver $\bar{X}$ bij het populatiegemiddelde vandaan ligt. Maar als de steekproef aselekt is, kun je wel een interval geven rondom $\bar{X}$, waar het populatiegemiddelde met 95% betrouwbaarheid binnen valt. Dit heet het 95%-betrouwbaarheidsinterval omdat het in 95% van de steekproeven het populatiegemiddelde bevat.

De formule van het 95%-betrouwbaarheidsinterval voor het populatiegemiddelde bij steekproefgemiddelde $\bar{X}$, steekproefstandaardafwijking S en steekproefomvang n is $\bar{X} \pm 2 \cdot \frac{S}{\sqrt{n}}$ oftewel $\left[\bar{X} - 2 \cdot \frac{S}{\sqrt{n}}, \bar{X} + 2 \cdot \frac{S}{\sqrt{n}} \right]$. De breedte van dit interval is $4 \cdot \frac{S}{\sqrt{n}}$. Zie de figuur hieronder.



figuur 10.29

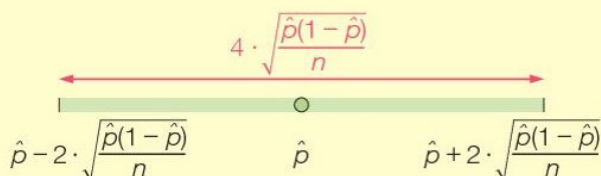
Bij een kwalitatieve variabele kun je voor de populatie en de steekproef alleen een proportie berekenen. De formule van het 95%-betrouwbaarheidsinterval voor de populatieproportie p bij

steekproefproportie $\hat{p}$ en steekproefomvang n is $\hat{p} \pm 2 \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

oftewel $\left[\hat{p} - 2 \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + 2 \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right]$.

De breedte van dit interval is $4 \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. Zie de figuur hieronder.

Je mag ook gebruiken $[\hat{p} - 2\sigma, \hat{p} + 2\sigma]$, want $\sigma = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.



figuur 10.30

Voorbeeld

Uit onderzoek van het RIVM in 2018 blijkt dat slechts 47% van de Nederlanders ouder dan 4 jaar aan de bewegingsrichtlijnen voldoet. Neem aan dat voor dit onderzoek een aselechte steekproef van 5000 respondenten werd genomen.

- a Bereken hoeveel het RIVM er met 95% betrouwbaarheid maximaal naast zit. Geef je antwoord in procenten.
- b Het RIVM wil dat de breedte van het 95%-betrouwbaarheidsinterval maximaal 2% is.
Bereken welke steekproefomvang hierbij hoort. Ga uit van $\hat{p} = 0,5$.

Aanpak

- a Omdat $\hat{p} = 0,47$ het midden van het interval is, zit het RIVM er met 95% betrouwbaarheid maximaal $2 \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ naast.
- b Gebruik de formule voor de breedte met $\hat{p} = 0,5$ en breedte = 0,02 en los de vergelijking op die je zo krijgt.

Uitwerking

- a $\hat{p} = 0,47$ en $n = 5000$ geeft $2 \cdot \sqrt{\frac{0,47 \cdot 0,53}{5000}} \approx 0,014$.
Het RIVM zit er dus met 95% betrouwbaarheid maximaal 1,4% naast.
- b Los op $4 \cdot \sqrt{\frac{0,5 \cdot 0,5}{n}} = 0,02$.
Voer in $y_1 = 4 \cdot \sqrt{\frac{0,5 \cdot 0,5}{x}}$ en $y_2 = 0,02$.
De optie snijpunt geeft $x = 10000$.
De steekproefomvang moet dus minstens 10000 zijn.



Zie het voorbeeld.

- a Stel dat naar aanleiding van vraag b besloten wordt een steekproef van 10000 personen te nemen en dat blijkt dat $\hat{p} = 0,48$ in plaats van het uitgangspunt $\hat{p} = 0,5$.
Is het interval dan nauwkeurig genoeg voor het RIVM? Licht toe met een berekening.
- b Bereken bij de steekproef onder 10000 respondenten met $\hat{p} = 0,48$ het 95%-betrouwbaarheidsinterval bij het percentage Nederlanders ouder dan 4 jaar dat voldoet aan de richtlijn. Rond af op één decimaal.

Een betrouwbaarheidsinterval noemen we nauwkeuriger als het minder breed is.

Let op: er wordt een 95%-betrouwbaarheidsinterval bij een percentage gevraagd.

43 Bereken het 95%-betrouwbaarheidsinterval voor het populatiegemiddelde bij

- a $\bar{X} = 120$, $S = 36$ en $n = 144$
- b $\bar{X} = 64,4$; $S = 15,5$ en $n = 100$
- c $\bar{X} = 18$ en $\frac{S}{\sqrt{n}} = 0,32$.

44 Bereken het 95%-betrouwbaarheidsinterval voor de populatieproportie bij

- a $\hat{p} = 0,4$ en $n = 240$
- b $\hat{p} = 0,48$ en $n = 380$
- c $\hat{p} = \frac{5}{9}$ en $n = 10\,000$.

Rond in een betrouwbaarheidsinterval proporties af op drie decimalen, tenzij anders wordt gevraagd.

A45 Bereken het 95%-betrouwbaarheidsinterval voor een steekproef

- a met 260 deelnemers met gemiddelde 23,4 en steekproefstandaardafwijking 1,9
- b van 700 personen waarbij 64% het onderzochte kenmerk had
- c een steekproef van 1000 voorstellingen waarbij in totaal 5350 personen voor het einde van de film vertrokken en waarbij de standaardafwijking van het aantal vertrokken personen per voorstelling 3,2 was.

Rond in een betrouwbaarheidsinterval gemiddelden af op één decimaal meer dan de gegeven standaardafwijking, tenzij anders wordt gevraagd.

46 Bereken welke steekproefomvang minimaal nodig is voor een 95%-betrouwbaarheidsinterval met

- a $\hat{p} = 0,3$ en een maximale breedte van het betrouwbaarheidsinterval van 0,04
- b $S = 4$ en een maximale breedte van het betrouwbaarheidsinterval van 0,8.

47 Bereken welke steekproefomvang nodig is voor een 95%-betrouwbaarheidsinterval met

- a $\hat{p} = 0,6$ en een maximale afwijking van 3%
- b een steekproefgemiddelde van 80 kg, een steekproefstandaardafwijking van 2 kg en een maximale afwijking van 0,5 kg.

A48
☐ ⊙ *

Gegeven zijn de volgende 95%-betrouwbaarheidsintervallen bij een populatiegemiddelde.

P: $[112,1; 124,1]$ en Q: $117 \pm 2 \cdot 3,3$

- a** Welke van deze betrouwbaarheidsintervallen is het meest nauwkeurig? Licht toe.

Bij betrouwbaarheidsinterval Q was S gelijk aan 66.

- b** Bereken de steekproefomvang bij betrouwbaarheidsinterval Q.
c Bereken hoe groot de steekproefomvang bij betrouwbaarheidsinterval Q moet zijn, zodat het betrouwbaarheidsinterval net zo nauwkeurig is als bij P.

R49
⊙ *

Een onderzoeker wil weten wat de proportie linkshandige Nederlanders is. Neem aan dat de Nederlandse bevolking uit 17 miljoen mensen bestaat en dat de onderzoeker uitgaat van een steekproefproportie van 0,1.

- a** Bereken hoeveel mensen hij moet opnemen in zijn steekproef voor een 95%-betrouwbaarheidsinterval met een maximale breedte van 0,02.

De onderzoeker vraagt zich af of de proportie linkshandigen in de wereld hetzelfde is als in Nederland. Hij gaat er vanuit dat de wereldbevolking uit 8 miljard mensen bestaat en dat de steekproefproportie weer 0,1 is. Hij berekent dat zijn steekproef net zo groot moet zijn als bij vraag a. Jori kan dit niet geloven, omdat de wereldbevolking veel groter is dan de bevolking van Nederland. Hij denkt dat de onderzoeker een rekenfout heeft gemaakt. Laura beweert dat de steekproefomvang in beide gevallen hetzelfde is. Volgens haar kun je dit zien aan de formule voor het berekenen van het 95%-betrouwbaarheidsinterval.

- b** Leg uit wie volgens jou gelijk heeft.

Er worden nooit aselechte steekproeven genomen van de wereldbevolking.

- c** Leg uit waarom het praktisch gezien moeilijk is om een aselechte steekproef te nemen van de wereldbevolking.



E50*****

In 2014 publiceerde NRC handelsblad een onderzoek naar zelfcitatie in het werk van een Nederlandse econoom. Zelfcitatie is het kopiëren van zinnen uit eigen werk zonder bronvermelding. In een steekproef vond een NRC-journalist 8 artikelen waarin zelfcitatie plaatsvond. Neem aan dat deze steekproef aselect was. Een commissie onderzocht vervolgens alle 261 artikelen van de econoom en vond dat 60 hiervan zelfcitatie bevatten.

a Bereken de populatieproportie p .

De NRC-journalist heeft niet vermeld wat de steekproefomvang n was. Salma wil met 95% betrouwbaarheid berekenen wat het minimale en het maximale aantal artikelen in de steekproef kan zijn geweest. Hiertoe lost

ze de vergelijkingen $p - 2 \cdot \sqrt{\frac{p(1-p)}{n}} = \frac{8}{n}$ en $p + 2 \cdot \sqrt{\frac{p(1-p)}{n}} = \frac{8}{n}$ op.

b Met welke vergelijking berekent ze het maximale aantal artikelen in de steekproef?

c Bereken met 95% betrouwbaarheid wat het minimale en maximale aantal artikelen in de steekproef kan zijn geweest.

R51**□ ⊗ ***

De onderstaande stellingen gaan over de betekenis van het 95%-betrouwbaarheidsinterval $[0,715; 0,885]$.

Welke stellingen zijn juist?

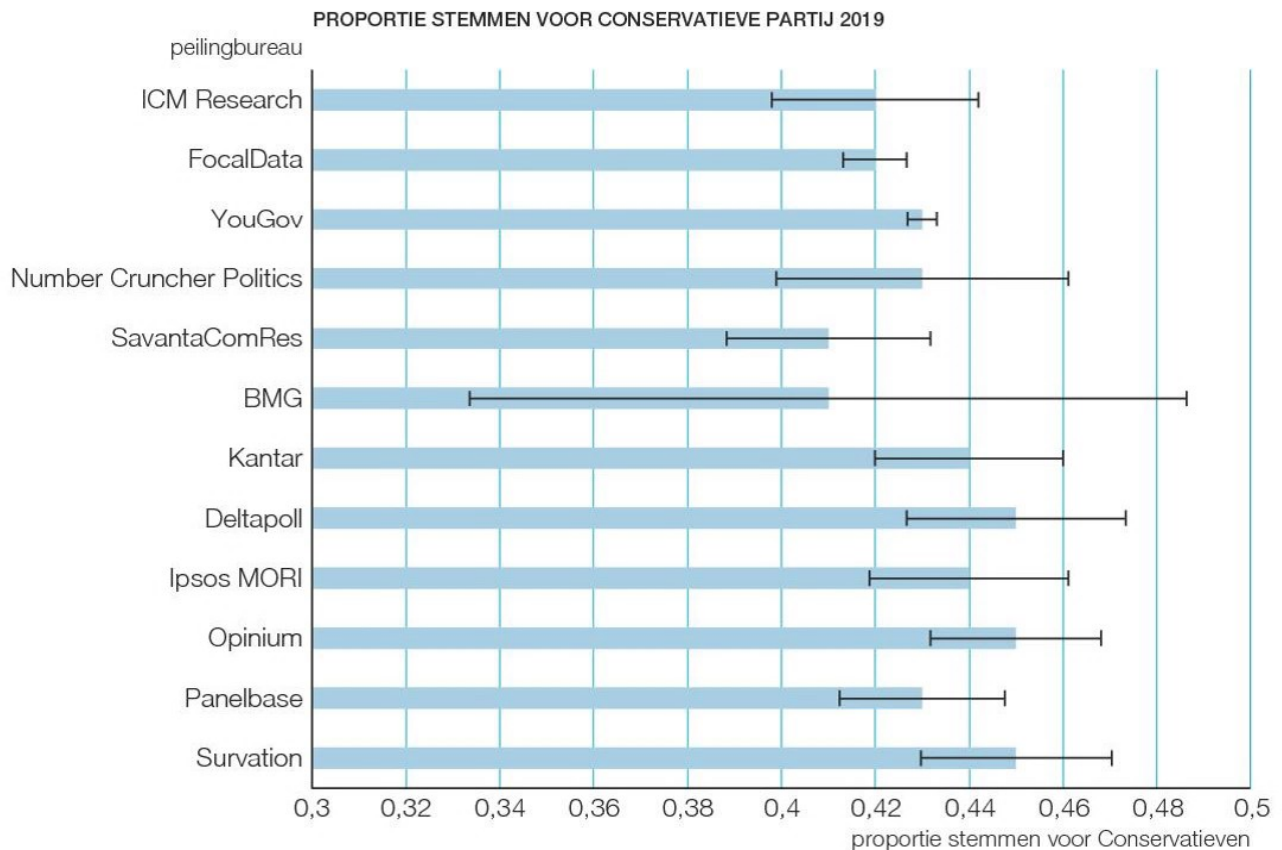
- I Je weet van de populatieproportie met 95% betrouwbaarheid dat deze 0,8 is.
- II Je weet van de populatieproportie met 95% betrouwbaarheid dat deze in het interval $[0,715; 0,885]$ ligt.
- III Je weet van de steekproefproportie met 100% betrouwbaarheid dat deze 0,8 is.
- IV Je weet van de steekproefproportie met 95% betrouwbaarheid dat deze 0,8 is.

52**□ ⊗ ***

Als er bij de onderstaande situaties sprake is van een aselecte steekproef uit de populatie, bereken dan het 95%-betrouwbaarheidsinterval. Als er geen sprake is van een aselecte steekproef uit de populatie, licht dan toe waarom dat niet zo is.

- a** In 2012 waren er onaangekondigde controles bij een aselecte steekproef van 119 piercing- en tatoeageshops uit een lijst van alle dergelijke shops in Nederland. Hiervan bleken er 36 niet aan alle eisen te voldoen.
- b** Voor een onderzoek naar hacken onder Nederlandse jongeren wordt aselect een steekproef van 1044 jongeren genomen uit een panel van 33 000 mensen. Deze panelleden hebben zich aangemeld om aan onderzoeken deel te nemen. Van de jongeren in de steekproef zegt 14,2% wel eens een mobiele telefoon gehackt te hebben.
- c** Voor een onderzoek naar de beeldvorming van een zanger op Twitter heeft een onderzoeker toegang tot alle data van Twitter. Hierin staan 5000 berichten over de zanger, waarvan 1800 negatief blijken te zijn.

In de week voorafgaand aan de verkiezingen in het Verenigd Koninkrijk in 2019 zijn peilingen uitgevoerd. In de figuur zie je de proportie stemmen voor de conservatieve partij volgens 12 peilingbureaus. Aan het einde van de staven is steeds het 95%-betrouwbaarheidsinterval weergegeven.



figuur 10.31

- a** Voor de peiling van Deltapoll zijn 1800 mensen ondervraagd en was de steekproefproportie 0,45.
Bereken de breedte van het 95%-betrouwbaarheidsinterval dat bij deze peiling hoort en bepaal of deze ongeveer overeenkomt met de figuur.
- b** Aan de grootste peiling deden 105 612 mensen mee en aan de kleinste 166.
Leg op basis van de figuur uit welke organisatie bij welke steekproefomvang hoort.
- c** Bepaal met behulp van de figuur de steekproefomvang van de steekproef van Kantar.

Bij de verkiezingen wonnen de conservatieven met 44% procent van de stemmen. Dus $p = 0,44$ kan gezien worden als de populatieproportie.

- d** Bij BMG ligt de populatieproportie binnen het 95%-betrouwbaarheidsinterval.
Waarom is dat niet zo indrukwekkend?
- e** Bij hoeveel procent van de organisaties ligt de populatieproportie binnen het 95%-betrouwbaarheidsinterval? Is dat meer of minder dan je zou verwachten? Geef een mogelijke verklaring.

A54

In een land zijn er voor elke autobezitter 2 mensen zonder auto. In een steekproef blijken 18 van de 80 ondervraagden een auto te hebben. De populatieproportie valt buiten het 95%-betrouwbaarheidsinterval.

- Toon dit met een berekening aan.
- Lieke beweert dat hieruit volgt dat de steekproef niet aselekt was. Leg uit waarom je het wel of niet eens bent met Lieke.

A55

Hoewel katten voor de gezelligheid in huis worden gehaald zijn het eigenlijk roofdieren. Er bestaan daarom zorgen over de invloed van jagende huiskatten op beschermde diersoorten.

Onderzoekers willen graag weten hoe groot het jachtgebied van katten is. Om deze oppervlakte te bepalen hebben onderzoekers GPS-trackers bevestigd aan een steekproef van 925 huiskatten die naar buiten mogen. Neem aan dat de steekproef aselekt is.

De gemiddelde oppervlakte van het jachtgebied van de katten in de steekproef was $0,019 \text{ km}^2$ met een standaardafwijking van $0,012 \text{ km}^2$.

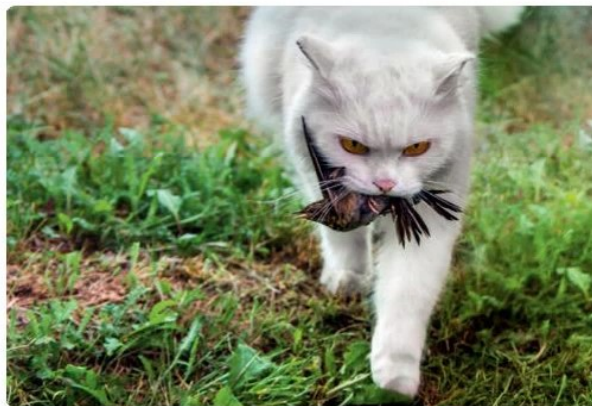
- Beredeneer op basis van deze gegevens dat de oppervlakte van het jachtgebied in de steekproef niet normaal verdeeld kan zijn.
- Bereken het 95%-betrouwbaarheidsinterval voor de oppervlakte van het jachtgebied in km^2 .

Er waren drie katten in de steekproef die per dag jaagden op een gebied van meer dan 1 km^2 . Een van deze katten was Max, waarvan een route te zien is op de foto hiernaast. Gedurende 50 aselekt gekozen dagen is de oppervlakte van het jachtgebied van Max bijgehouden. Het 95%-betrouwbaarheidsinterval dat hoort bij de oppervlakte van zijn jachtgebied in km^2 is $[1,273; 3,478]$. Hieruit volgt dat S ongeveer gelijk is aan 3,9.

- Bereken S in drie decimalen.

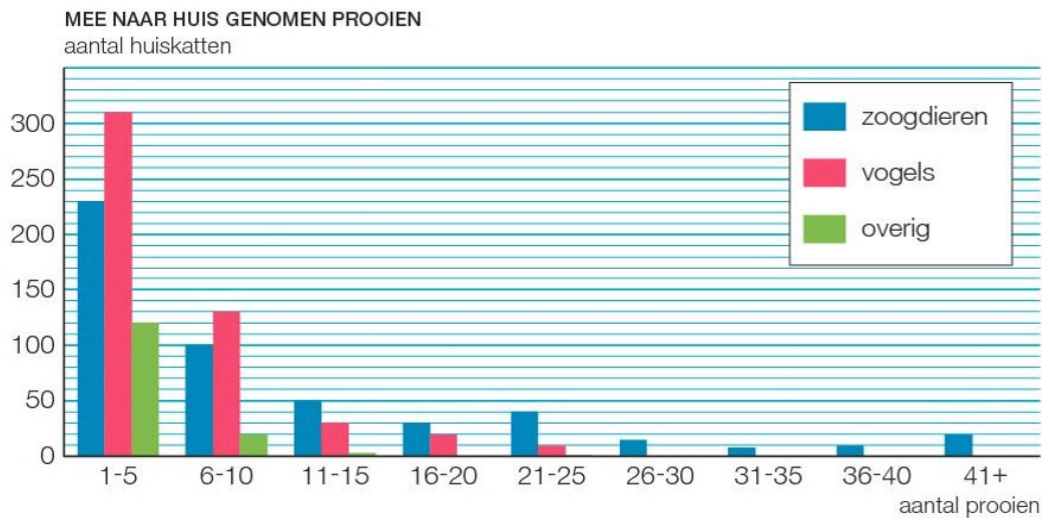
Door het jachtgebied van Max meer dan 50 dagen te onderzoeken, kun je een nauwkeuriger betrouwbaarheidsinterval opstellen.

- Bereken hoeveel dagen het onderzoek moet duren om een 95%-betrouwbaarheidsinterval met een breedte van 1 km^2 te krijgen. Neem aan dat $S = 3,9 \text{ km}^2$.



A56
  *

In Groot-Brittannië werd een representatieve steekproef van 986 kattenbezitters genomen. Zij hielden gedurende vijf maanden bij hoeveel prooien hun kat mee naar huis nam. Neem aan dat elke kattenbezitter één kat heeft. De resultaten staan in het staafdiagram.



figuur 10.32

Amel trekt op grond van het staafdiagram twee conclusies bij deze gegevens.

- I Alle huiskatten hebben minstens één vogel mee naar huis genomen.
- II Er werden meer vogels dan zoogdieren mee naar huis genomen.
- a Toon met berekeningen aan dat je deze conclusies niet kunt trekken.
- b Geef het 95%-betrouwbaarheidsinterval voor de proportie huiskatten in Groot-Brittannië die vogels mee naar huis nemen.

Het totale aantal mee naar huis genomen prooien in de periode van de steekproef is 14 370. Het totale aantal huiskatten in Groot-Brittannië is 9 miljoen. Neem verder aan dat 25% van de prooien mee naar huis wordt genomen en dat de katten gedurende het jaar evenveel prooien vangen.

- c Bereken het 95%-betrouwbaarheidsinterval voor het totale aantal prooien in miljoenen per jaar dat door huiskatten wordt gevangen in Groot-Brittannië. Gebruik dat $S = 14,1$ mee naar huis genomen prooien per kat.

A57
 *

Zie opgave 56.

Een onderzoeker wil een 95%-betrouwbaarheidsinterval opstellen voor het aantal prooien per jaar van de 2,6 miljoen Nederlandse huiskatten. Op basis van onder andere de bovenstaande studie verwacht ze een gemiddelde van 140 prooien per kat per jaar. Hoeveel kattenbezitters moeten in haar steekproef zitten, zodat haar 95%-betrouwbaarheidsinterval een maximale breedte van 50 miljoen heeft? Ga ervan uit dat de kattenbezitters allemaal één kat hebben en gebruik dat $S = 134$.

Terugblik

De normale verdeling

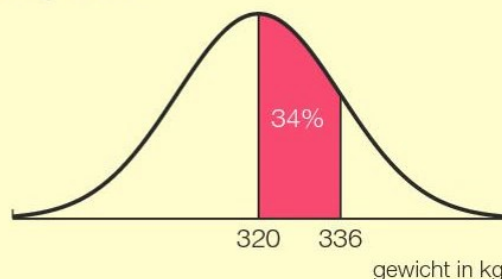
Bij een normale verdeling horen de volgende vuistregels.

- ongeveer 68% van de waarnemingsgetallen ligt tussen $\mu - \sigma$ en $\mu + \sigma$
- ongeveer 95% van de waarnemingsgetallen ligt tussen $\mu - 2\sigma$ en $\mu + 2\sigma$
- bijna 100% van de waarnemingsgetallen ligt tussen $\mu - 3\sigma$ en $\mu + 3\sigma$

Je komt de normale verdeling tegen in twee situaties:

- bij sommige hoeveelheden uit het dagelijks leven
- bij steekproefgemiddelden en steekproefproporties, die een steekproevenverdeling vormen als er heel veel steekproeven worden genomen.

Je mag bij een normale verdeling van gewichten met $\mu = 320$ kg en $\sigma = 16$ kg zeggen dat 34% van de gewichten tussen 320 kg en 336 kg ligt.



Werken met betrouwbaarheidsintervallen

Het 95%-betrouwbaarheidsinterval voor het populatiegemiddelde is

$\left[\bar{X} - 2 \cdot \frac{S}{\sqrt{n}}, \bar{X} + 2 \cdot \frac{S}{\sqrt{n}} \right]$. Hierin is $\bar{X}$ het steekproefgemiddelde, S de

steekproefstandaardafwijking en n de steekproefomvang.

De breedte van dit interval is $4 \cdot \frac{S}{\sqrt{n}}$.

Het 95%-betrouwbaarheidsinterval voor de populatieproportie is

$\left[\hat{p} - 2 \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + 2 \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right]$. Hierin is $\hat{p}$ de

steekproefproportie en n de steekproefomvang.

De breedte van dit interval is $4 \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.

Wil je dat de breedte van een 95%-betrouwbaarheidsinterval maximaal 1,2 is bij steekproefgemiddelde 12,6 en steekproefstandaardafwijking 5,2

dan bedenk je dat de breedte $4 \cdot \frac{S}{\sqrt{n}}$ is met $S = 5,2$.

Je lost op $4 \cdot \frac{5,2}{\sqrt{n}} = 1,2$.

Voer in $y_1 = 4 \cdot \frac{5,2}{\sqrt{x}}$ en $y_2 = 1,2$.

De optie snijpunt geeft $x = 300,44...$

De steekproefomvang moet dus minstens 301 zijn.

Eindopdracht Dunks in de WNBA

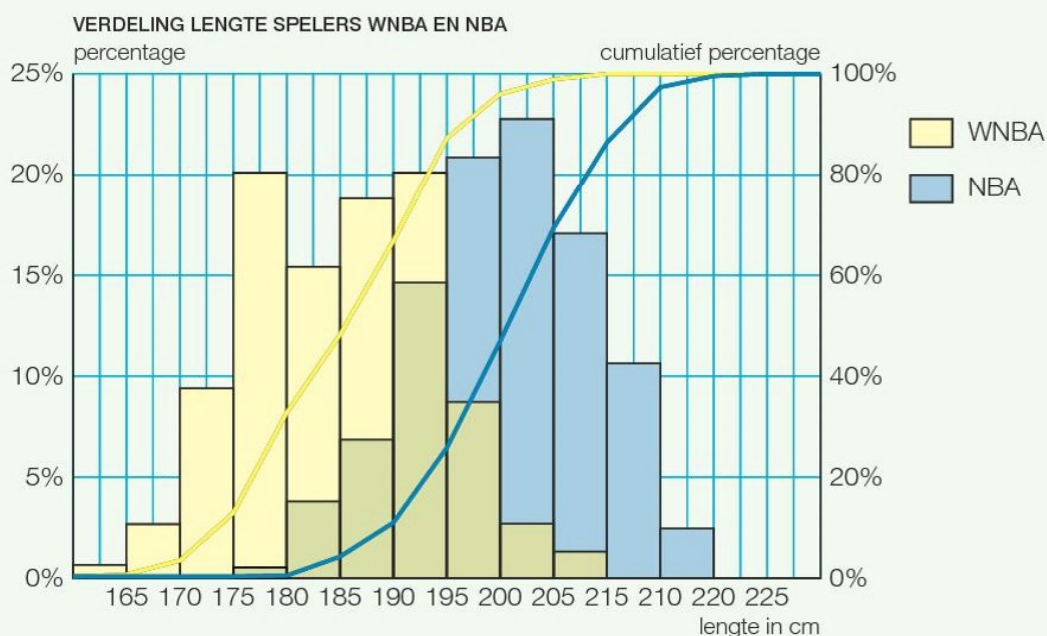
De Women's National Basketball Association (WNBA) is de bekendste Amerikaanse basketbalcompetitie voor dames. De WNBA werd in 1996 opgericht en heeft een opzet die sterk lijkt op die van de NBA.

Wanneer een basketbalspeler door te springen de basketbal boven de ring weet te brengen en vervolgens in een neerwaartse beweging scoort, spreekt men van een (slam)dunk. In de NBA komt deze wijze van scoren vaak voor, terwijl er in de WNBA minder dan tien spelers zijn die op deze wijze kunnen scoren. De recordhouder aantal dunks binnen de WNBA is Brittney Griner.

Het is niet verrassend dat weinig WNBA-spelers dunken, omdat ze doorgaans een stuk kleiner zijn en de ring op dezelfde hoogte (305 cm) hangt als in de NBA. In de figuur hieronder staan gegevens over de lengte van spelers uit de NBA en WNBA in 2019.



Brittney Griner dunkt tijdens haar debuut-wedstrijd in de WNBA



Ga in deze opgave uit van de volgende aannames:

- Spelers in de WNBA springen 48,8 cm en in de NBA 67,1 cm.
- Het bereik van een speler is 34% meer dan zijn of haar lengte.
- Om de bal te kunnen dunken moet je 10 cm boven de ring komen.
- Onderzoek hoe hoog de ring in de WNBA zou moeten hangen om te zorgen dat (op basis van de aannames hierboven) hetzelfde deel van de spelers kan dunken als in de NBA.

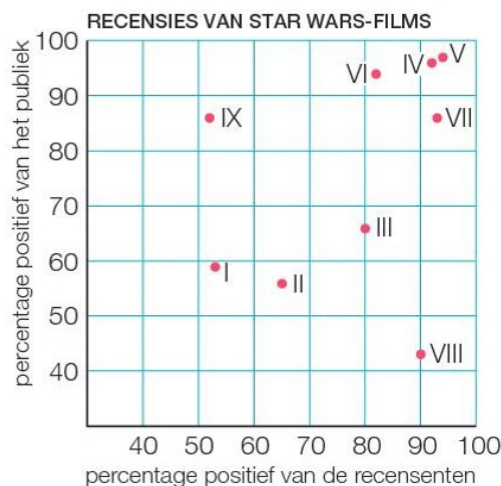


Diagnostische toets

10.1 Statistische diagrammen

1 Op de website Rotten Tomatoes worden door zowel recensenten als het gewone publiek films beoordeeld met 'fresh' (positief) of 'rotten' (negatief). Zie het spreidingsdiagram met beoordelingen van negen Star Wars-films.

- Is een beoordeling op Rotten Tomatoes een kwalitatieve of een kwantitatieve variabele? En het aantal mensen dat positief is over een Star Wars-film per groep?
- Bij welke film is het verschil tussen het oordeel van het publiek en de recensenten het grootst? Bepaal de grootte van dit verschil.
- Voor welke films is het percentage positieve beoordelingen door recensenten groter dan door het publiek?



figuur 10.33

10.2 Causaliteit en groepen vergelijken

2 Wetenschappers verzamelden gegevens van 40 monarchvlinders die in gevangenschap zijn opgegroeid. Ze vergeleken deze met een groep van 41 monarchvlinders in het wild. Onder de vlinders in gevangenschap waren 24 mannetjes en bij die in het wild 27.

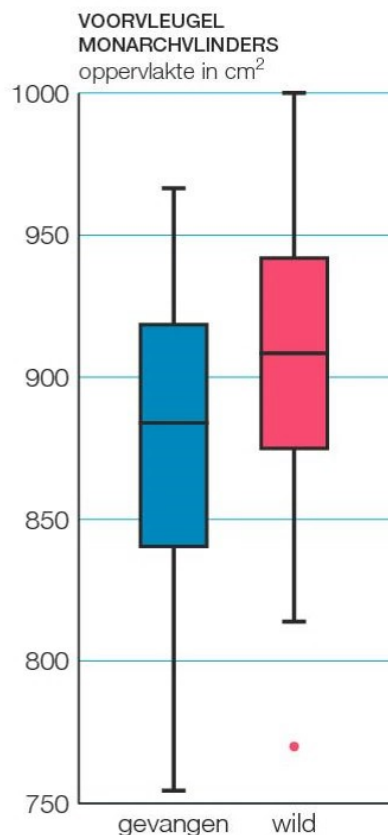
- Is de variabele geslacht kwalitatief of kwantitatief?
- Onderzoek met behulp van het formuleblad of het verschil in geslacht tussen monarchvlinders in gevangenschap en in het wild groot, middelmatig of gering is.

Van beide groepen vlinders is de oppervlakte van een voorvleugel gemeten. Zie de tabel en boxplots.

VOORVLEUGELOPPERVLAKE

	gevangen	wild
$\bar{X}$	877,1 cm ²	903,1 cm ²
S	55,7 cm ²	73,3 cm ²

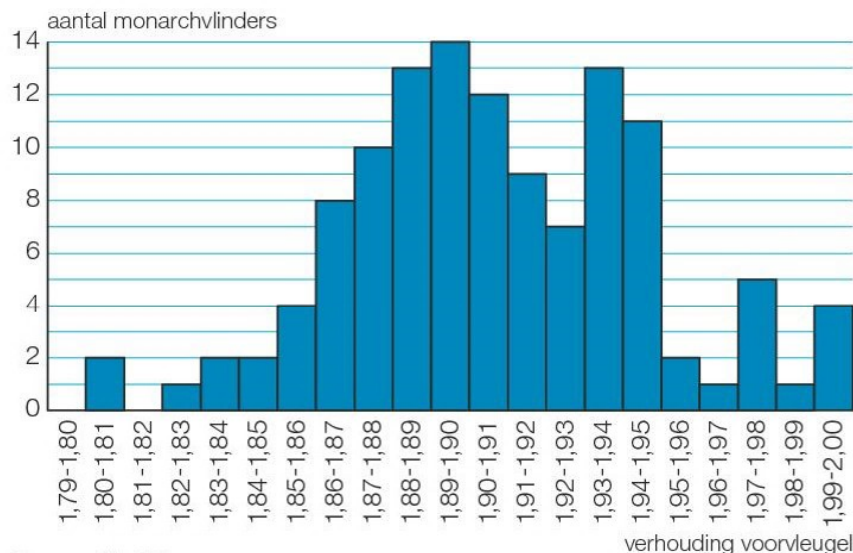
- Onderzoek met twee associatiematen de grootte van het verschil in voorvleugeloppervlakte tussen de groepen monarchvlinders.
- Mag je op grond van dit onderzoek concluderen dat gevangenschap voor deze vlinders een oorzaak is van een kleinere oppervlakte van de voorvleugel? Licht toe.



figuur 10.34

10.3 De normale verdeling en betrouwbaarheidsintervallen

- 3** Van 123 monarchvlinders is van een voorvleugel de verhouding gemeten van de lengte tot de breedte. Zie de figuur hieronder.



figuur 10.35

- a** Ligt de mediaan in de modale klasse? Licht toe.
b Geef twee redenen waarom dit geen normale verdeling is.

Bij de steekproef was de gemiddelde verhouding 1,91 met $\sigma = 0,04$.

- c** Bereken hoeveel procent van de vlinders met een verhouding van minder dan 1,83 je zou verwachten als er sprake zou zijn van een normale verdeling. Komt dit percentage ongeveer overeen met figuur 10.35?

- 4** Het longfonds heeft in 2019 een steekproef gehouden onder Nederlanders met een longziekte. Van de 220 ondervraagden gaven 55 aan verhuisd te zijn naar een plek met schonere lucht. Neem aan dat de steekproef aselekt is.

- a** Bereken het 95%-betrouwbaarheidsinterval voor het percentage personen dat is verhuisd naar een plek met schonere lucht.
b Waarom is de steekproef van het longfonds geen aselekte steekproef voor een onderzoeker die wil weten hoeveel Nederlanders zijn verhuisd naar een plek met schonere lucht?

- 5** Voor een onderzoek uit 2017 naar *fast fashion* is aan 50 volwassen Nederlanders gevraagd hoeveel schoenen en kledingstukken zij bezitten en hoeveel daarvan ongebruikt zijn. Zie de tabel hiernaast. Neem aan dat de steekproef aselekt is.

- a** Bereken het 95%-betrouwbaarheidsinterval voor het gemiddeld aantal jassen van een volwassen Nederlander. Gebruik dat $S = 2,2$.

	totaal	ongebruikt
schoenen	482	117
broeken	475	124
t-shirts	1282	353
jumpsuits	49	21
jassen	284	55

Een kledingfabrikant wil een 95%-betrouwbaarheidsinterval met een breedte van maximaal 4% opstellen voor het percentage ongebruikte jumpsuits in Nederland. De fabrikant gaat uit van $\hat{p} = 0,5$.

- b** Bereken het aantal personen dat in de steekproef moet zitten.