



**H. Quaadgras**

# **STATISTIEK IN BEDRIJF**

**1e druk**



Noordhoff Uitgevers



## **Statistiek in bedrijf**





# Statistiek in bedrijf

H. Quaadgras

Noordhoff Uitgevers Groningen | Houten

Omslagontwerp: Dorèl en anderen  
Omslagbeeld: Di>ja

Eventuele op- en aanmerkingen over deze of andere uitgaven kunt u richten aan: Noordhoff Uitgevers bv, Afdeling Hoger Onderwijs, Antwoordnummer 13, 9700 VB Groningen, e-mail: [info@noordhoff.nl](mailto:info@noordhoff.nl)

2 3 4 5 / 10 09

© 2006 Noordhoff Uitgevers bv Groningen/Houten, The Netherlands.

Behoudens de in of krachtens de Auteurswet van 1912 gestelde uitzonderingen mag niets uit deze uitgave worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever. Voor zover het maken van reprografische verveelvoudigingen uit deze uitgave is toegestaan op grond van artikel 16h Auteurswet 1912 dient men de daarvoor verschuldigde vergoedingen te voldoen aan Stichting Reprorecht (postbus 3060, 2130 KB Hoofddorp, [www.cedar.nl/reprorecht](http://www.cedar.nl/reprorecht)). Voor het overnemen van gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (artikel 16 Auteurswet 1912) kan men zich wenden tot Stichting PRO (Stichting Publicatie- en Reproductierechten Organisatie, postbus 3060, 2130 KB Hoofddorp, [www.cedar.nl/pro](http://www.cedar.nl/pro)).

*All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.*

ISBN (ebook) 978-90-01-84433-2  
ISBN 978-90-01-70528-2  
NUR 916



## Woord vooraf

*Statistiek in bedrijf* is in eerste instantie bedoeld voor SPD-studenten bedrijfsstatistiek. De behandelde stof omvat dan ook de onderwerpen van het examenprogramma daarvan.

De opzet van het boek is statistiek voor studenten toegankelijk te maken. Daarmee is het boek ook geschikt voor een breder publiek.

Ik wil graag mijn collega's van Markus Verbeek Praehep bedanken voor hun inbreng in wat eerst nog een syllabus zou worden. Hans Slabbe-koorn heeft alle versies van de hoofdstukken vanuit het onderwijskundig perspectief grondig doorgenomen en suggesties gegeven voor het verbeteren van tekst en structuur met de student als uitgangspunt. En heel belangrijk: zonder steun van thuis was dit boek er niet gekomen.

Op- of aanmerkingen worden op prijs gesteld. Deze kunnen doorgegeven worden aan de uitgever:

Noordhoff Uitgevers  
Postbus 58  
9700 MB Groningen

Zutphen, mei 2006  
Herman Quaadgras





# Inhoud

Studiewijzer 11

- 1 Statistiek, een verkenning 13**
  - 1.1 Basisbegrippen 14
  - 1.2 Variabelen 18
    - Begrippenlijst 21
    - Oefenopgaven 22
  
- 2 Primaire en secundaire nauwkeurigheid, afrondingsregels 23**
  - 2.1 Primaire nauwkeurigheid 24
  - 2.2 Secundaire nauwkeurigheid 28
    - Begrippenlijst 33
    - Oefenopgaven 34
  
- 3 Frequentieverdelingen 35**
  - 3.1 Begrippen bij frequentietabellen 36
  - 3.2 Opstellen van een frequentietabel 38
  - 3.3 Histogram 48
  - 3.4 Frequentiepolygoon 56
  - 3.5 Cumulatieve frequentie 61
  - 3.6 Cumulatieve frequentiepolygoon 62
  - 3.7 Concentratiecurve 64
  - 3.8 Het gebruiken van soorten frequenties 68
  - 3.9 Interpoleren 68
  - 3.10 Stamdiagram 70
  - 3.11 Staafdiagram, samengesteld staafdiagram en cirkeldiagram 72
    - Begrippenlijst 78
    - Oefenopgaven 79
  
- 4 Centrummaten 81**
  - 4.1 Het begrip centrummaat 82
  - 4.2 Centrummaten bij losse waarnemingen 82
  - 4.3 Centrummaten bij een klassenindeling 86
  - 4.4 Centrummaten en transformeren van massa's 96
  - 4.5 Centrummaten en samenvoegen van massa's 97
  - 4.6 Keuze van een centrummaat 97
  - 4.7 Ligging van mediaan, modus en rekenkundig gemiddelde in de verdeling 99
  - 4.8 Bijzondere gemiddelden 99
    - Begrippenlijst 103
    - Oefenopgaven 104

- 5**     **Spreidingsmaten** 105
  - 5.1     Het begrip spreidingsmaat 106
  - 5.2     Spreidingsmaten bij losse waarnemingen 106
  - 5.3     Spreidingsmaten bij klassenindelingen 114
  - 5.4     Boxplot 119
  - 5.5     Transformeren van spreidingsmaten 120
  - 5.6     Keuze spreidingsmaat 122
    - Begrippenlijst 123
    - Oefenopgaven 124
  
- 6**     **Indexcijfers** 125
  - 6.1     Het begrip indexcijfer 126
  - 6.2     Enkelvoudige indexcijfers 127
  - 6.3     Samengestelde indexcijfers 133
  - 6.4     Productie-indexcijfers 142
  - 6.5     Index reëel loon, nominaal loon, consumentenprijsindex 145
  - 6.6     Arbeidsproductiviteit 145
    - Begrippenlijst 148
    - Oefenopgaven 149
  
- 7**     **Elementaire kansberekening** 151
  - 7.1     Kansbegrippen 152
  - 7.2     Elementaire kansregels 154
  - 7.3     Trekkingen met en zonder terugleggen 169
  - 7.4     Combinatoriek 175
  - 7.5     Binomiale kansen 180
  - 7.6     Verwachtingswaarde en standaarddeviatie 183
  - 7.7     Aanvullende voorbeelden 185
    - Begrippenlijst 189
    - Oefenopgaven 191
  
- 8**     **Normale verdeling** 193
  - 8.1     Veelgebruikte symbolen en vuistregels normale verdeling 194
  - 8.2     Standaardnormale verdeling 198
  - 8.3     Verwachtingswaarde en standaarddeviatie van de som en van het gemiddelde (bij trekkingen uit één populatie) 215
  - 8.4     Binomiale kansen benaderen met de normale verdeling 217
    - Begrippenlijst 223
    - Oefenopgaven 224
  
- 9**     **Toetsen en schatten met behulp van steekproeven** 227
  - 9.1     Toetsen van een populatiegemiddelde 228
  - 9.2     Schatten van een populatiegemiddelde 237
  - 9.3     Bepalen van de steekproefgrootte (bij het schatten van een populatiegemiddelde) 239
  - 9.4     Toetsen van een populatieproportie 241
  - 9.5     Schatten van een populatieproportie 243
  - 9.6     Bepalen van de steekproefgrootte (bij het schatten van een populatieproportie) 244
  - 9.7     Chi-kwadraat toetsen ( $\chi^2$ -toetsen) 246
    - Begrippenlijst 250
    - Oefenopgaven 252

**10**      **Correlatie en regressie** 253

- 10.1    Soorten correlatie 254
- 10.2    Spreidingsdiagram 255
- 10.3    Regressielijn 257
- 10.4    Correlatiecoëfficiënt 263
- 10.5    Twee regressielijnen 264
- 10.6    Time-lag 270
- Begrippenlijst 274
- Oefenopgaven 275

**11**      **Tijdreeksanalyse** 277

- 11.1    Componenten van tijdreeksen 278
- 11.2    Trend 278
- 11.3    Seizoenspatroon 286
- 11.4    Restcomponent 295
- 11.5    Voor seizoensinvloeden gecorrigeerde reeks 297
- 11.6    Prognoses 300
- Begrippenlijst 303
- Oefenopgaven 304

Register 307



## Studiewijzer

*Statistiek in bedrijf* is een boek voor hbo-studenten, SPD in het bijzonder.

In het eerste hoofdstuk verkennen we het vakgebied Statistiek. Daarna komen aan de orde: primaire en secundaire nauwkeurigheid, afrondingsregels; frequentieverdelingen; centrummaten; spreidingsmaten; indexcijfers; elementaire kansberekening; normale verdeling; toetsen en schatten met behulp van steekproeven; correlatie en regressie en tijdreeksanalyse.

Bij het bestuderen van de stof wordt geadviseerd de volgorde van de hoofdstukken van het boek aan te houden. Dit kan door het hoofdstuk eerst globaal door te nemen om een eerste indruk te krijgen. Daarna kan hetzelfde hoofdstuk nauwgezet doorgenomen worden, en zouden de voorbeelden nagerekend kunnen worden. Ten slotte kunnen de oefenopgaven gemaakt worden die aan het einde van ieder hoofdstuk staan.

De antwoorden van de oefenopgaven zijn te vinden op de site [www.statistiekinbedrijf.wolters.nl](http://www.statistiekinbedrijf.wolters.nl). Op deze site is ook aanvullende informatie te vinden, met name over de kengetallen bij hoofdstuk 6 en de Chi-kwadraat-toetsen bij hoofdstuk 9. Deze onderwerpen worden niet expliciet voor het SPD-examen gevraagd, maar plaatsen de leerstof in een wat breder perspectief. Op de site zijn ook extra opgaven voor docenten te vinden.

Als voorbereiding op het SPD-examen kunnen de margewoorden en de begrippenlijst achter in ieder hoofdstuk een hulpmiddel zijn. Verder is het maken van veel examenopgaven belangrijk. De oefenopgaven in de hoofdstukken zijn namelijk wel geschikt om te bepalen of de stof begrepen is, maar niet representatief voor een examen. Examens zijn te vinden op de site van de Associatie of in examenbundels.



# Statistiek, een verkenning

## 1

- 1.1 Basisbegrippen
- 1.2 Variabelen

Statistiek wordt tegenwoordig veel gebruikt. Je kunt geen krant of tijdschrift openslaan of er worden wel cijfers over bijvoorbeeld werkloosheid, verkeersongevallen of beursnoteringen gepresenteerd met tabellen en/of grafieken. Vaak worden deze cijfers nog verder toegelicht met bepaalde analysemethoden, allemaal onderdelen van statistiek. Ook bedrijven en instellingen maken veel gebruik van statistiek. Het uiteindelijke doel hiervan is het inzichtelijk aanleveren van informatie om beter onderbouwde beslissingen te kunnen nemen.

## 1.1 Basisbegrippen

In deze paragraaf wordt verteld wat statistiek is en wordt een aantal basisbegrippen uit de statistiek beschreven.

### 1.1.1 Omschrijving statistiek

Een omschrijving van statistiek is: een manier om informatie te krijgen uit gegevens.

#### Voorbeeld 1.1

Een bedrijf wil weten hoe het eigen nettowinstpercentage zich verhoudt tot de nettowinstpercentages van branchegenoten. Van 55 branchegenoten zijn de volgende gegevens bekend (zie tabel 1.1).

Tabel 1.1 Nettowinstpercentages

2,4%	4,4%	3,6%	1,2%	8,7%
3,1%	4,5%	11,4%	9,5%	3,9%
8,4%	7,0%	4,2%	5,5%	3,1%
0,3%	0,1%	3,1%	3,3%	0,4%
10,1%	1,6%	8,5%	5,0%	5,3%
9,3%	3,4%	5,7%	6,3%	7,3%
3,8%	7,5%	7,5%	2,6%	2,5%
3,8%	7,4%	3,9%	10,8%	5,8%
5,9%	11,0%	9,1%	4,9%	4,2%
5,2%	4,7%	2,7%	7,8%	3,7%
6,6%	6,3%	0,8%	7,6%	6,2%

Wat nu te doen met deze (ruwe) gegevens? Zijn er meer branchegenoten met een hoger nettowinstpercentage? Is het gemiddelde nettowinstpercentage van de branche hoger, enzovoort.

Kortom: Statistiek houdt zich ermee bezig welke informatie we uit deze gegevens kunnen halen.

De gegevens van de 55 branchegenoten, de verzameling met als elementen de 55 nettowinstpercentages, wordt een *statistische massa* genoemd, of ook wel populatie. De termen statistische massa en populatie worden door elkaar gebruikt. Definitie:

**Statistische massa** || Een statistische massa is een verzameling gelijksoortige elementen.

Het begrip statistische massa wordt veel gebruikt in de statistiek. Een hiervan afgeleide omschrijving van het begrip statistiek is:

**Statistiek** || Statistiek is een systematische methode die zich bezighoudt met het bestuderen van massale verschijnselen (statistische massa's) met het doel cijfermatige informatie daarover te verkrijgen.



Denk bij dit laatste bijvoorbeeld aan het berekenen van een gemiddelde.

Uit de omschrijving blijkt al dat bestudering van een opzichzelfstaand geval, bijvoorbeeld het saldo op je laatste bankafschrift, eigenlijk niet tot de statistiek hoort. We noemen dit *casuïstiek*: casus is geval.

## Casuïstiek

### 1.1.2 Onderzoeksaspecten statistiek

In statistisch onderzoek proberen we dus cijfermatige informatie te verkrijgen. Deze informatie kan op één statistische massa betrekking hebben, maar we kunnen ook meerdere massale verschijnselen met elkaar vergelijken.

#### Onderzoeksaspecten van één statistische massa

De cijfermatige (kwantitatieve) informatie zal vaak betrekking hebben op één of meer van de volgende aspecten van een massaal verschijnsel:

- omvang
- structuur
- ontwikkeling in de tijd.

#### *Omvang*

## Omvang

De omvang betreft het aantal elementen van de statistische massa. (In voorbeeld 1.1 is de omvang 55 elementen.)

#### *Structuur*

## Centrum

Bij de structuur kan naar een aantal aspecten worden gekeken:

- Is er een centrale tendens (centrum), een gebied waar veel waarnemingen liggen en zo ja, waar ligt dat? (Veel winstpercentages tussen de 2% en 6%)
- Hoe zijn de waarnemingen gespreid (spreiding)? Liggen de waarden dicht bij het centrum? (Veel winstpercentages rond de 4%)
- Liggen de waarnemingen symmetrisch rond dit centrum of zijn er naar één kant uitschieters (scheefheid)? (Veel waarnemingen beneden of boven een bepaald winstpercentage?)
- Is de verdeling gelijkmatig (homogeen), of juist heel onregelmatig? (Gelijkmatige verdeling van de winstpercentages?)

## Spreiding

## Scheefheid

## Homogeen

#### *Ontwikkeling in de tijd*

## Ontwikkeling in de tijd

Bij ontwikkeling in de tijd gaat het erom of we een regelmaat kunnen ontdekken in het verloop van een bepaald verschijnsel, bijvoorbeeld de procentuele groei van de kwartaalomzetten van een bedrijf over een aantal jaren, of een jaarlijks terugkerend patroon in die kwartaalomzetten.

#### Massale verschijnselen vergelijken

We kunnen ook massale verschijnselen met elkaar vergelijken. We kunnen kijken naar:

- overeenkomsten of verschillen van soortgelijke verschijnselen bij verschillende massa's;
- samenhang (correlatie) tussen verschijnselen.

## Overeenkomsten of verschillen

### *Overeenkomsten of verschillen tussen soortgelijke verschijnselen*

Bij het vergelijken van overeenkomsten of verschillen van soortgelijke verschijnselen zouden we weer kunnen kijken naar de eerdergenoemde aspecten van een massaal verschijnsel, bijvoorbeeld het vergelijken van de ontwikkeling van de kwartaalomzetten van twee bedrijven over een aantal jaren. Als de grootte van de omzetten van beide bedrijven erg veel verschilt, zou met percentages (of met indexcijfers) kunnen worden gewerkt.

## Samenhang

### *Samenhang (correlatie)*

Bij samenhang onderzoeken we of er een verband is tussen verschijnselen, bijvoorbeeld tussen het reclamebudget en de omzet van een bedrijf (een hoger reclamebudget zal, hopelijk, leiden tot een hogere omzet), twee verschillende verschijnselen dus. Er kan ook samenhang zijn tussen soortgelijke verschijnselen, bijvoorbeeld tussen de jaaromzetten van een bepaald bedrijf en de jaaromzetten van de totale branche.

### 1.1.3 **Verzamelen van gegevens**

## Primaire statistiek

Als duidelijk is welk aspect van het massaal verschijnsel we gaan bekijken, dan is de volgende stap het verzamelen van de benodigde gegevens van de populatie. Verzamelen we de benodigde gegevens door zelf waarnemingen in de populatie te doen, dan heet dit *primaire statistiek*. Het is ook mogelijk dat we gegevens gebruiken die al eerder door de onderzoeker of door een andere instantie (bijvoorbeeld het Centraal Bureau voor de Statistiek, CBS) voor een ander doel zijn verzameld. Dan spreekt men van *secundaire statistiek*.

## Secundaire statistiek

Ook moet er een keus worden gemaakt of de hele populatie onderzocht gaat worden of dat we maar een deel van de populatie gaan bekijken. In dat laatste geval nemen we dan een zogenoemde *steekproef* uit de populatie.

### 1.1.4 **Steekproeven**

Het doel van een steekproef is door het onderzoeken van slechts een deel van de populatie, eigenschappen van de hele populatie te leren kennen. Belangrijk hierbij is de vraag wat de betrouwbaarheid van uitspraken over die eigenschappen is. Kunnen we aannemen dat de gevonden eigenschappen van de steekproef ook gelden voor de hele populatie?

## Representatief

Als we de populatie onderzoeken door middel van een steekproef, dan moet de steekproef wel representatief zijn. Een steekproef moet aan drie voorwaarden voldoen om representatief te zijn:

- 1 afgestemd op het onderzoek;
- 2 voldoende groot;
- 3 aselechte trekking uit de relevante doelmassa.

We bespreken deze voorwaarden afzonderlijk.

### Afgestemd op het onderzoek

Allereerst moeten we natuurlijk de goede populatie onderzoeken. Als we wat over het computergebruik van de jeugd onder de achttien jaar willen weten, dan moeten we geen bejaarden ondervragen.

Verder kan het verstandig zijn om de populatie op te delen in lagen (deelpopulaties), en uit deze lagen in de juiste verhouding steekproeven te nemen.

#### Gestratificeerde steekproef

Als we bijvoorbeeld wat over bepaalde aspecten van de gezondheid willen weten en we vermoeden dat het actief deelnemen aan sport hiermee te maken heeft, en we weten dat de bevolking onderverdeeld is in actieven en niet-actieven in een verhouding van 3 : 2, dan nemen we ook in de steekproef 60% actieven en 40% niet-actieven op, dezelfde verhouding. Dit heet een gelaagde of *gestratificeerde steekproef* (stratum = laag).

#### Getrapte steekproef

Een andere manier van steekproeven nemen is de *getrapte steekproef*. Een voorbeeld hiervan is de Nederlandse bevolking onderzoeken door eerst met een steekproef een aantal gemeenten te selecteren en daaruit weer met een steekproef een aantal inwoners. Meestal spelen kostenoverwegingen een rol bij de keuze.

### Voldoende groot

Om op basis van een steekproef van drie personen iets te zeggen over de eetgewoonten van de totale Nederlandse bevolking, is natuurlijk absurd.

Voor de omvang van een steekproef bestaan geen absolute criteria, maar de steekproef mag in ieder geval niet té klein zijn. In hoofdstuk 9 wordt op de steekproefgrootte teruggekomen.

#### Aselect

### Aselect

Bij een aselechte steekproef moet elk element van de populatie dezelfde kans hebben om in de steekproef voor te komen. Als we bijvoorbeeld het telefoonboek gebruiken om een steekproef te nemen, dan komen we hierin onder anderen geen mensen met een geheim telefoonnummer tegen, terwijl die groep voor een bepaald onderzoek best van belang kan zijn.

### 1.1.5 Onderverdeling statistiek

#### Verklarende (of inductieve) statistiek

In de vorige subparagraaf werd aan het begin de vraag gesteld in hoeverre we eigenschappen van de hele populatie kunnen leren kennen met behulp van een steekproef. Het deel van de statistiek dat zich hiermee bezighoudt, wordt *verklarende (of inductieve) statistiek* genoemd. De verklarende statistiek doet naar aanleiding van een deel van de populatie uitspraken over de hele populatie en de betrouwbaarheid daarvan. Dit onderdeel wordt in hoofdstuk 9 over kansberekening behandeld.

#### Beschrijvende (of deductieve) statistiek

Het andere deel van de statistiek (statistiek wordt vaak in twee delen opgesplitst) wordt *beschrijvende (of deductieve) statistiek* genoemd. Onder beschrijvende statistiek valt onder meer:

- het verzamelen van gegevens;

- de overzichtelijke presentatie van het waarnemingsmateriaal door middel van tabellen en grafieken;
- het bepalen van de karakteristieke grootheden, zoals een gemiddelde en een spreidingsmaat;
- het leggen van verbanden.

De beschrijvende statistiek houdt zich dus bezig met het overzichtelijk weergeven van informatie over de populatie. Dit gebeurt met behulp van tabellen, grafieken en samenvattende maten (bijvoorbeeld een gemiddelde).

Het leggen van verbanden (zie het laatste punt van beschrijvende statistiek) wordt ook wel het derde deel van statistiek genoemd en heet dan *relationele statistiek*. Hieronder vallen de onderwerpen correlatie en analyse van tijdreeksen die in de hoofdstukken 10 en 11 behandeld worden.

## Relationele statistiek

### 1.2 Variabelen

Een belangrijk basisbegrip in statistiek is het begrip variabele. Vaak hebben we in de statistiek (zowel in de theorie als in de praktijk) te maken met onbekende grootheden, kenmerken van de populatie die verschillende waarden kunnen aannemen. Een onbekende grootheid noemen we een *variabele*. Een voorbeeld van een variabele is de lengte van een groep volwassen mensen, deze zal waarschijnlijk niet voor allemaal gelijk zijn. Meten we de lengte van een persoon, dan heet de uitkomst de *waarde van de variabele*.

## Variabele

#### 1.2.1 Soorten variabelen

We onderscheiden kwalitatieve en kwantitatieve variabelen. Over het algemeen kan worden gesteld dat het bij een *kwalitatieve variabele* om een eigenschap gaat die niet in een getal of grootheid is uit te drukken. Bijvoorbeeld: iemands beroep, het merk van een auto, kleuren, geuren, enzovoorts. Anders gezegd: een kwalitatieve variabele heeft geen meetbaar sorteerkennmerk.

## Kwalitatieve variabele

Een *kwantitatieve variabele* is een eigenschap die wel in een getal of grootheid (met een eenheid) is uit te drukken. Bijvoorbeeld iemands lengte, 172 cm, of een aantal mensen. Een kwantitatieve variabele heeft wel een meetbaar sorteerkennmerk.

## Kwantitatieve variabele

Kwantitatieve variabelen kunnen we onderscheiden in continu of discreet.

## Continue variabele

Bij een *continue* variabele zijn tussen twee uitkomsten alle tussenliggende waarden mogelijk. Het gaat dan om lengten, gewichten, temperaturen, geldwaarden en dergelijke. Bijvoorbeeld tussen 73,2 kilogram en 73,3 kilogram ligt 73,25 kg, maar ook 73,2749724921 kg.

## Discrete variabele

Bij een *discrete* variabele kan tussen twee uitkomsten slechts een beperkt aantal waarden worden aangenomen. Vaak betreft het een variabele die alleen in gehele aantallen kan worden gemeten. Bijvoorbeeld:

het aantal cursisten in een klas is discreet. Tussen 23 en 25 cursisten bevindt zich alleen de meetwaarde 24 cursisten.

### 1.2.2 Meetniveau variabelen

Voor de behandeling van latere stof is ook nog het zogenoemde meetniveau van de variabele met bijbehorende meetschaal belangrijk.

We onderscheiden vier typen meetschalen (bij vier meetniveaus). De eerste twee hebben betrekking op kwalitatieve variabelen, de laatste twee op kwantitatieve. De vier meetschalen zijn:

- 1 nominale schaal;
- 2 ordinale schaal;
- 3 intervalschaal;
- 4 ratio- of verhoudingschaal.

We behandelen deze meetschalen (en meetniveaus) hierna achtereenvolgens.

#### Nominale schaal

Op de nominale schaal worden de waarnemingsuitkomsten van een *kwalitatieve* variabele alleen onderscheiden naar naam (Latijn: nomen). De waarnemingsuitkomsten zijn dus niet te rangschikken in een logische volgorde. Denk aan codenummers. Het geslacht bijvoorbeeld is een kwalitatieve variabele die wordt ondergebracht op een nominale schaal. De schaal uit dit voorbeeld kent maar twee schaalpunten: man en vrouw. Daarom spreken we hier van een tweepuntsschaal. Maar ook de variabele land van Europa had als voorbeeld genomen kunnen worden.

#### Ordinale schaal

Op de ordinale schaal krijgen de waarnemingsuitkomsten van een *kwalitatieve* variabele een naam én kunnen ze worden gerangschikt in een logische volgorde, een rangorde. Voor een indeling naar rang (een kwalitatieve variabele) van een groep militairen is bijvoorbeeld de volgende vierpuntsschaal bruikbaar: 'soldaat, korporaal, onderofficier, officier'. Er is een ordening tussen deze categorieën (soldaat lager dan korporaal, enzovoort) en daarom is deze vierpuntsschaal een ordinale schaal. Een ander voorbeeld is het aantal Michelin-sterren (hoewel dit een kwantitatieve variabele lijkt).

#### Intervalschaal

Op de intervalschaal worden de waarnemingsuitkomsten van een *kwantitatieve* variabele afgezet. De keuze van het nulpunt is willekeurig. Alleen het verschil tussen twee getallen van een interval heeft een eenduidige betekenis, niet de verhouding. Bijvoorbeeld de temperatuurschaal van Celsius; je kunt niet zeggen 20 °C is twee maal zo warm als 10 °C, wel kun je zeggen dat het verschil tussen 20 °C en 15 °C gelijk is aan het verschil tussen 15 °C en 10 °C. De temperatuurschaal van Fahrenheit heeft trouwens een ander nulpunt.

### **Ratioschaal**

Ook op de ratioschaal worden de waarnemingsuitkomsten van een *kwantitatieve* variabele afgezet, maar er is sprake van een natuurlijk nulpunt. Daardoor kunnen niet alleen verschillen tussen waarnemingsuitkomsten worden vergeleken, maar is het ook zinvol hun onderlinge verhouding te bepalen. De lengte van mensen is een kwantitatieve variabele die op een ratioschaal wordt afgezet. Het nulpunt van een ratioschaal is niet willekeurig gekozen. Daarom is een lengte van 1,80 meter wel 1,5 maal zo lang als een lengte van 1,20 meter.

## Begrippenlijst

<b>Aselect</b>	Een steekproef is aselect als elk element van de populatie dezelfde kans heeft om in de steekproef voor te komen.
<b>Beschrijvende statistiek</b>	Het deel van statistiek dat zich bezighoudt met het overzichtelijk weergeven van informatie over de populatie.
<b>Casuïstiek</b>	Het bestuderen van een opzichzelfstaand geval.
<b>Continue variabele</b>	Variabele waarbij tussen twee uitkomsten ook alle tussenliggende waarden mogelijk zijn.
<b>Discrete variabele</b>	Variabele waarbij tussen twee uitkomsten slechts een beperkt aantal waarden mogelijk is.
<b>Gestratificeerde steekproef</b>	Steekproef waarbij de populatie opgedeeld is in lagen, waaruit de steekproef in dezelfde verhouding wordt gekozen.
<b>Getrapte steekproef</b>	Steekproef waarbij uit een steekproef weer een steekproef genomen wordt.
<b>Kwalitatieve variabele</b>	Een eigenschap die niet in een getal of grootheid is uit te drukken.
<b>Kwantitatieve variabele</b>	Een eigenschap die in een getal of grootheid is uit te drukken.
<b>Statistiek</b>	Methode die zich bezighoudt met het informatie verkrijgen over statistische massa's.
<b>Statistische massa</b>	Een verzameling gelijksoortige elementen, ook wel populatie genoemd.
<b>Primaire statistiek</b>	Manier om informatie te krijgen uit gegevens die verzameld zijn door zelf waarnemingen in de populatie te doen.
<b>Secundaire statistiek</b>	Manier om informatie te krijgen uit gegevens die al eerder door een ander voor een ander doel verzameld zijn.
<b>Verklarende statistiek</b>	Het deel van statistiek dat zich bezighoudt met de vraag in hoeverre we eigenschappen van de populatie met behulp van een steekproef kunnen leren kennen.
<b>Variabelen</b>	Vooralsnog onbekende grootheden, kenmerken van een populatie, die verschillende waarden kunnen aannemen, maar door meting vastgesteld kunnen worden.

## Oefenopgaven

- 1.1 We willen door middel van een steekproef van 1 000 personen de mening van de Nederlandse vakantieganger weten over het comfort van busreizen naar buitenlandse vakantiebestemmingen. We ondervragen hiervoor 1 000 busreizigers naar buitenlandse vakantiebestemmingen.

Welke problemen kunnen we tegenkomen als we kijken naar de voorwaarde 'aselect' voor een representatieve steekproef?

- 1.2 a Geef eigen voorbeelden van een kwalitatieve variabele.  
b Geef eigen voorbeelden van een kwantitatieve variabele.
- 1.3 Geef bij de variabelen uit opgave 1.2 aan op welk meetniveau of op welke meetschaal ze betrekking hebben.