

1

Waarvoor hebben we statistiek nodig

Soorten gegevens

Overzicht

- We gebruiken statistiek om onze gegevens te beschrijven, maar ook om vast te stellen hoe betrouwbaar de informatie is die op steekproeven is gebaseerd.
- Een variabele is elk meetbaar concept dat tussen individuen en cases varieert.
- Variabelen worden onderscheiden in nominale (ook wel categorische of kwalitatieve) variabelen en scorevariabelen (ook wel numerieke of kwantitatieve variabelen genoemd).
- Volgens de formele meettheorie zijn er nog meer soorten variabelen – nominale, ordinale, interval- en ratiovariabelen. Deze zijn over het algemeen niet zo belangrijk in de dagelijkse praktijk van het uitvoeren van statistische analyses.
- Nominale variabelen bestaan uit categorieën met alleen een bepaalde naam, terwijl scorevariabelen met een getal op een numerieke schaal gemeten worden en de grootte van de variabele aangeven.

1.1. Inleiding

Stel je eens een wereld voor waarin alles hetzelfde is; de mensen zijn in alle opzichten identiek. Ze dragen dezelfde kleren, ze eten dezelfde maaltijden, ze zijn vanaf de geboorte allemaal even lang, ze gaan allemaal naar dezelfde school met dezelfde leraren, dezelfde lessen en dezelfde faciliteiten, ze gaan allemaal dezelfde maand met vakantie, ze doen allemaal hetzelfde werk, ze wonen in dezelfde huizen, en de zon schijnt elke dag. Ze hebben geen seks zoals wij die kennen, omdat er geen geslachten zijn en iedereen met 30 jaar zichzelf kloont. In de tuinen staan dezelfde planten en de bodem is precies hetzelfde in elke tuin. Ze sterven allemaal op hun 75^e verjaardag en ze worden allemaal begraven in dezelfde soort houten kisten in identieke stukjes grond. Ze zijn allemaal even slim en ze hebben allemaal dezelfde persoonlijkheid. Hun genetische achtergrond varieert nooit. Wiskundig gezien zijn al deze kenmerken constanten. Als dit geen realistisch wereldbeeld lijkt te zijn, dan hebben we nieuws voor je – je hebt statistiek nodig! Alleen in een gestandaardiseerde wereld zou je geen statistiek nodig hebben – in een rijkelijk gevarieerde wereld is statistiek echter van essentieel belang.

Als er geen variatie zou bestaan, dan zou alles wat we over mensen te weten kunnen komen afgeleid kunnen worden uit de informatie die we over één enkele persoon krijgen. Er zouden geen problemen ontstaan met generaliseren, omdat alles wat waar is over Sandra Groen ook waar is voor ieder ander – iedereen heet tenslotte Sandra Groen!

Gelukkig ziet de wereld er in werkelijkheid niet zo uit. **Variabiliteit** is een essentieel kenmerk van het leven en de sociale wereld waarin we leven. We moeten de enorme variabiliteit zien te beheersen als we uitspraken over de werkelijke wereld willen doen. Statistiek is voor een groot deel bedoeld om de zin achter deze variabiliteit te ontdekken.

Statistische technieken vervullen drie hoofdfuncties:

1. Ze laten toe om de informatie die we uit een veelheid van bronnen vergaren samen te vatten. Een doelstelling van de statistiek is immers uit het zo duidelijk en effectief mogelijk presenteren van je onderzoeksgegevens. In die hoedanigheid beschrijft statistiek alleen maar de verzamelde informatie. Dit wordt bereikt met behulp van tabellen en grafieken voor het samenvatten van gegevens, en met eenvoudige formules die ingewikkelde gegevens omzetten in eenvoudige indices die de belangrijkste kenmerken van de gegevens numeriek beschrijven. Deze tak van de statistiek noemen we om voor de hand liggende redenen **beschrijvende statistiek** – hiermee wordt de informatie die je verzamelt zo nauwkeurig en zo bondig mogelijk beschreven. De eerste hoofdstukken van dit boek zijn voornamelijk gewijd aan beschrijvende statistiek.
2. Een andere tak van de statistiek is bij de meesten van ons veel minder bekend: de **verklarende statistiek**. Deze tak van statistiek heeft in feite te maken met efficiënt onderzoek doen. Als men vroeger iets over mensen te weten wilde komen, werd iedereen in het land benaderd om de nodige informatie te verzamelen. Dit gebeurt vandaag nog als een overheid een *volkstelling* houdt om iets te weten te komen over de bevolking van een land op een bepaald moment. Dit is een enorme en tijdrovende operatie die niet al te vaak kan worden herhaald. De meesten van ons zijn wel vertrouwd met het gebruik van relatief kleine **steekproeven** om een beeld te krijgen van de informatie die we zouden kunnen krijgen door iedereen te bestuderen. Dit gebeurt algemeen in enquêtes bij publiek, door bijvoorbeeld de antwoorden van een steekproef van 1000 mensen te gebruiken om de uitkomst van een nationale verkiezing te voorspellen. Het principe van steekproeven trekken is belangrijk, ook al kunnen steekproeven wel eens misleidend zijn. *Verklarende statistiek* gaat over de mate van betrouwbaarheid waarmee we van een steekproef naar de hele populatie kunnen generaliseren.
3. De hoeveelheid gegevens die een onderzoeker kan verzamelen kan in principe enorm groot zijn. Sommige statistische technieken maken het de onderzoeker mogelijk om door middel van een aantal krachtige methoden bepaalde trends te vinden in grote hoeveelheden gegevens. Het vereenvoudigen, analyseren en reduceren van gegevens zijn een paar van de namen die aan dit proces worden gegeven. Hoe je het ook noemt, het doel is hetzelfde – de betekenis zien te vinden in grote hoeveelheden gegevens die anders veel te verwarrend zouden zijn. Deze *methoden voor gegevensonderzoek* worden voornamelijk in de latere hoofdstukken behandeld.

1.2. Variabelen en metingen

Een **variabele** is een belangrijk basisbegrip in de statistiek. Het is ook een heel eenvoudig concept: *een variabele kan alles zijn wat varieert en wat gemeten kan worden*. Deze metingen hoeven *niet* echt overeen te komen met wat we in het dagelijks leven

onder metingen verstaan, zoals metingen van gewicht, afstand en temperatuur. Ook het geslacht van mensen is een variabele, omdat dit gemeten kan worden als hetzij mannelijk, hetzij vrouwelijk – en het geslacht varieert tussen mensen. Op dezelfde manier is oogkleur een variabele, omdat er in een groep mensen personen met bruine ogen zullen zijn, een aantal zal blauwe ogen hebben en een aantal zal groene ogen hebben. Meten kan dus ook bestaan uit alleen maar categoriseren. Klinisch psychologen kunnen in hun onderzoek verschillende diagnostische categorieën gebruiken, zoals schizofrenie, manische depressiviteit en angst. Deze diagnostische categorieën vormen een variabele omdat het verschillende mentale en emotionele problemen zijn die aan bepaalde personen kunnen worden toegekend. Zulke categoriseringsmethoden vormen een belangrijk soort metingen in de statistiek.

Een ander soort meting in de statistiek lijkt meer op het concept meting dat we in het dagelijks leven tegenkomen, waarin numerieke waarden (getallen) worden gegeven. Deze *numerieke* waarden of getalwaarden worden toegekend aan variabelen zoals gewicht, lengte, afstand, temperatuur en dergelijke – bijvoorbeeld 10 kilometer of 30 graden Celsius. Deze getalwaarden worden *scores* genoemd. In psychologisch onderzoek worden op min of meer dezelfde manier veel variabelen gemeten en *gekwantificeerd*. Voorbeelden zijn veel tests en schalen die gebruikt worden om intelligentie, persoonlijkheidstype, houdingen en mentale vaardigheden te bepalen. In de meeste van deze voorbeelden wordt aan personen een getal (of score) toegekend, waarmee bijvoorbeeld aangeduid wordt hoe neurotisch of hoe extravert een bepaald individu is. Psychologen zeggen bijvoorbeeld dat iemand een IQ van 112 of 93 heeft, of ze zeggen dat een bepaalde persoon een lage score 6 heeft op een schaal voor psychotisch gedrag. Gewoonlijk worden deze getallen gebruikt alsof ze exact hetzelfde zeggen als andere soorten metingen, zoals van gewicht of lengte. Hierover kunnen we uitspraken doen als: een persoon heeft een gewicht van 60 kilogram of een lengte van 1,8 meter.

1.3. De belangrijkste soorten metingen

In statistiekboeken voor psychologen en sociologen heeft men het normaliter over verschillende soorten metingen – meestal **meetschalen** genoemd. *Voor vrijwel alle praktische toepassingen zijn er echter maar twee verschillende soorten metingen van belang in de statistiek.* Deze zijn al besproken, maar we willen er nog eens de nadruk op leggen:

1. *Score/numerieke meting* Dit is het toekennen van een *numerieke* waarde of getalwaarde aan een meting. Dit betreft de meeste fysieke en psychologische metingen. In het menswetenschappelijk jargon worden deze numerieke metingen *scores* genoemd. We zouden de IQ scores van vijf personen als in tabel 1.1 kunnen vastleggen. Elk van de numerieke waarden in de tabel geeft de *score* van het genoemde individu op de variabele IQ. Het lijkt misschien een triviaal punt, maar we willen erop wijzen dat de getallen informatie bevatten dat iemand met een IQ van 150 een hogere intelligentie heeft dan iemand met een IQ van 80. Met andere woorden, de getallen *kwantificeren* de variabele.
2. *Nominale meting/categoriemeting* Hierbij wordt vastgesteld tot welke categorie van variabele een bepaalde case behoort. Dit noemen we ook wel een *kwalitatieve meting*. Als we bijvoorbeeld het beroep of baan van een persoon willen meten,

moeten we vaststellen of deze een vrachtwagenchauffeur is, een professor sociologie, een incasseerder enzovoort. Dit noemen we een *nominale* meting, omdat de categorieën meestal met woorden worden omschreven, en ook namen krijgen. De categorie ‘vrachtwagenchauffeur’ is een naam of een verbale omschrijving van de soort case die in die categorie geplaatst moet worden. Er zit geen rangorde in de antwoordcategorieën, een vrachtwagenchauffeur is niet beter dan een incasseerder.

Tabel 1.1 IQ-scores van vijf personen

Individu	IQ-score
Stan	80
Mavis	130
Sanjit	150
Sharon	145
Peter	105

Tabel 1.2 Frequenties van voorkomen van verschillende beroepen

Beroepscategorie	Aantal of frequentie in de verzameling
Vrachtwagenchauffeurs	27
Professoren in de sociologie	10
Incasseerders	15
Andere beroepen	48

Merk op dat er geen getallen worden gebruikt bij het proces van categoriseren zelf. Iemand is vrachtwagenchauffeur of hij is het niet. *Je moet echter oppassen voor een mogelijke verwarring die kan ontstaan.* Stel dat je honderd mensen hebt van wie de beroepen bekend zijn, en je wilt bijvoorbeeld weten hoeveel daarvan vrachtwagenchauffeur zijn, hoeveel professoren in de sociologie enzovoort. Deze tellingen kunnen in een gegevenstabel als tabel 1.2 worden opgenomen. We wijzen erop dat de getallen in dit geval overeenkomen met een telling van de *frequentie* van voorkomen of het aantal cases dat in elk van de vier beroepscategorieën valt. *Het zijn geen scores*, maar frequenties. De getallen corresponderen niet met één enkele meting, maar ontstaan door het combineren van vele losse (nominale) metingen.

Het onderscheid tussen numerieke scores en frequenties is belangrijk, dus controleer altijd goed of dat wat numerieke scores in een tabel lijken te zijn in werkelijkheid geen frequenties zijn.

Maak er een gewoonte van om variabelen te benoemen als numerieke scores of als nominale categorieën. Hiermee zet je een belangrijke stap in de richting van statistisch denken. Dit is in feite alles wat je moet weten over verschillende soorten metingen. Bedenk echter wel dat anderen er ingewikkelder systemen op na houden. In de volgende paragraaf vertellen we iets meer over meetschalen.

1.3.1. Formele meettheorie

Het is waarschijnlijk niet nodig om de inhoud van deze paragraaf in detail te leren. Er staan echter termen in waarmee je vertrouwd moet zijn, omdat anderen die kunnen gebruiken.

Veel psychologen hebben het over vier verschillende meetschalen. Als concept zijn ze verschillend. Niettemin is in de meeste praktijksituaties waarin psychologen gebruiken van statistiek het onderscheid tussen de nominale categorie en de numerieke scores, zoals hierboven beschreven, voldoende.

De vier ‘theoretische’ meetschalen zijn de volgende. De nummers 2, 3 en 4 zijn verschillende soorten *numerieke* scores.

1. **Nominale** *categorisering* Hierbij worden cases in categorieën met een bepaalde *naam* geplaatst – nominaal verwijst duidelijk naar namen. Dit is precies hetzelfde als ons nominale meet- of categoriseringsproces. Er zit geen rangorde tussen de verschillende categorieën.
2. **Ordinale** *meting (of rangordemeting)* In dit geval wordt aangenomen dat de waarden van de numerieke scores ons weinig anders vertellen dan welke de kleinste is, de op een na kleinste enzovoort, tot de grootste. Met andere woorden, we kunnen de scores in een *rangorde* plaatsen (vandaar de term *ordinaal*), van de kleinste tot de grootste. Soms wordt dit ook rangordemeting genoemd, omdat we rangordenummers kunnen toekennen aan de eerste, tweede, derde, vierde, vijfde enzovoort in volgorde van de kleinste naar de grootste numerieke waarde. Er is een trapsgewijs verschil in waarde tussen de variabelen, alleen is het verschil tussen iedere trede niet gelijk. Deze rangordenummers hebben de numerieke waarden 1, 2, 3, 4, 5 enzovoort. Voorbeelden hiervan zul je later in dit boek tegenkomen, vooral in de hoofdstukken 7 en 18. Er zijn echter weinig menswetenschappers die gegevens direct als rangordenummers verzamelen.
3. **Intervalmeting** De basisgedachte hierachter is dat in sommige gevallen de intervallen tussen getallen op een numerieke schaal gelijk van grootte zijn. Dus als we de afstand meten op een centimeterschaal, dan is de afstand tussen 0 en 1 centimeter precies even groot als het verschil tussen 4 en 5 centimeter of tussen 11 en 12 centimeter op die schaal. Dit is de voor de hand liggende methode bij verschillende fysische metingen, zoals die van een temperatuur.
4. **Ratiometing** Dit is precies hetzelfde als een intervalmeting met één belangrijke extra voorwaarde. Een ratioschaal heeft een absoluut nulpunt dat als 0 wordt gemeten. De meeste fysische metingen, zoals die van afstand en gewicht, hebben absolute nulpunten. Zo is de nul op een meetlint de kleinste afstand die je kunt hebben – er is geen afstand tussen twee samenvallende punten. Met dergelijke meetschalen kun je verhoudingen tussen metingen berekenen. Zo is bijvoorbeeld een plaats die 20 kilometer ver ligt twee keer zo ver weg als een plaats die maar 10 kilometer ver ligt. Een gebouw dat 15 meter hoog is, is half zo hoog als een gebouw dat 30 meter hoog is. (Niet alle fysische metingen hebben een nulpunt dat een absoluut nulpunt is – dat is bijvoorbeeld niet het geval voor verschillende temperatuurmaten. Temperaturen die in graden Celsius of Fahrenheit worden gemeten hebben waarden die als nulpunt worden aangemerkt. Deze nulpunten corresponderen echter niet met de laagst mogelijke temperatuur die je kunt hebben. Het is daarom zinloos om bijvoorbeeld te zeggen dat het bij 20 graden Celsius twee keer zo warm is als bij 10 graden Celsius.)

Het is niet eenvoudig om de drie laatstgenoemde meetmethoden eenduidig toe te passen bij psychologische metingen. Omdat de meeste psychologische scores geen direct waarneembare fysieke basis hebben, is het onmogelijk om vast te stellen of ze uit gelijke intervallen bestaan of dat ze een absoluut nulpunt hebben. Dit probleem heeft jarenlang voor grote controverse en verwarring onder psychologen gezorgd. Tegenwoordig wordt bij gebruik van door de meeste psychologen het verschil tussen de verschillende soorten numerieke scores genegeerd. Maar sommige psychologen hechten eraan om te benadrukken dat bepaalde gegevens het beste als rangordegegevens kunnen worden gezien, omdat ze de eigenschappen missen die kenmerkend zijn voor interval/ratiogegevens. Deze psychologen zullen eerder dan anderen geneigd zijn de statistische methoden die in hoofdstuk 18 worden beschreven en de rangcorrelatiecoëfficiënt (hoofdstuk 7) te gebruiken. Met andere woorden, voor precies dezelfde gegevens zullen verschillende psychologen verschillende statistische methoden gebruiken. Meestal zal dit weinig uitmaken voor de uitkomst van de statistische analyses – de resultaten. In het algemeen zul je er weinig problemen van ondervinden als je de drie onderverdelingen van het meten van numerieke scores negeert bij het gebruik van statistiek in de praktijk. De uitzonderingen hierop worden in de hoofdstukken 7 en 18 besproken. Omdat psychologen zelden of nooit gegevens verzamelen in de vorm van rangordennummers, hebben dergelijke overwegingen geen effect op de hoofdstukken 2 tot en met 6.

1.4. Kernpunten

- Stel jezelf steeds de vraag wat voor soort meting je van plan bent te gaan uitvoeren – is het een numerieke score van een variabele of gaat het om het plaatsen van individuen in categorieën?
- Veronderstel nooit zonder meer dat een getal een numerieke score is. Zonder controle zou het een frequentie van waarnemingen van een bepaalde categorie kunnen zijn.
- Helder denken is van belang voor het gebruiken van statistiek – er zal zelden van je verwacht worden dat je veel creativiteit in je statistisch werk laat zien. Een goed begrip van de exacte betekenis van termen is een duidelijk voordeel bij het werken met statistiek.

1.5. Kernbegrippen

Beschrijvende statistiek De verzameling van technieken en procedures die data uit steekproeven of populaties beschrijven: grafische voorstellingen, spreidingsmaten, maten voor centrale tendens, scheefheid en platheid van verdelingen.

Intervalmeting De intervallen tussen de getallen op een numerieke schaal zijn gelijk van grootte, zonder absoluut nulpunt.

Meetschaal De schaal waarop je een meting afneemt.

Nominaal meetniveau Een meetniveau dat enkel een kwalitatief onderscheid maakt tussen de verschillende waarden van een variabele.

Ordinaal meetniveau Meetniveau waarbij enkel een ordening van de waarden kan worden opgesteld zonder dat de afstand tussen de waarden enige betekenis heeft.

Ratiometing De intervallen tussen de getallen op een numerieke schaal zijn gelijk van grootte, met een absoluut nulpunt.

Steekproef Een deel van een populatie dat wordt onderzocht in onderzoek en waaruit conclusies worden getrokken over de populatievariabelen.

Variabiliteit Een niet wiskundige term om te verwijzen naar de grootte van de verschillen in waarnemingen.

Verklarende statistiek De verzameling van statistische toetsen die, via het toetsen van hypothesen en het onderzoek van steekproeven, generaliserende uitspraken doen over populaties. Wordt ook wel toetsende statistiek, inductieve of analytische statistiek genoemd en gaat dus een stap verder dan de beschrijvende statistiek die geen generalisering doet.

2

Variabelen beschrijven

Tabellen en grafieken

Overzicht

- Tabellen en grafieken zijn belangrijke aspecten van de beschrijvende statistiek (de beschrijving van de belangrijkste kenmerken van de gegevens). Gegevens met deze mate van detaillering onderzoeken is een essentieel stadium van elke statistische analyse, dat nooit mag worden overgeslagen.
- In dit hoofdstuk laten we zien hoe je tabellen en grafieken voor individuele variabelen kunt maken en presenteren.
- Statistische tabellen en grafieken moeten op effectieve wijze informatie over je gegevens naar de lezer overbrengen. Pas op voor een te complexe weergave.
- Het soort gegevens (nominaal of score) bepaalt in belangrijke mate welk soort tabel of grafiek het meest geschikt is.
- Als de gegevens nominaal zijn, dan zijn eenvoudige frequentietabellen, staafdiagrammen of cirkeldiagrammen het meest geschikt. De frequenties zijn dan gewoon het aantal gevallen in elk van de afzonderlijke categorieën.
- Als de gegevens uit scores bestaan, zijn frequentietabellen of histogrammen geschikt. Om de presentatie overzichtelijk te houden en trends duidelijk naar voren te laten komen, kun je het beste de gegevens in banden (of gebieden) van naast elkaar liggende scores weergeven.

Vorbereiding

Ga aan de hand van hoofdstuk 1 nog eens na wat een variabele is. En als je nog niet helemaal thuis bent in de nominale vorm van meten (categoriseren) en van het gebruik van numerieke scores bij een meting, lees dan de desbetreffende gedeelten ook nog eens door.

2.1. Inleiding

Je weet waarschijnlijk veel meer over statistiek dan je zou denken. De massamedia gebruiken geregeld statistische tabellen en grafieken; kinderen raken op school bekend met statistische tabellen en grafieken. Het is heel belangrijk om vaardig te kunnen omgaan met deze technieken, omdat onderzoekers grote hoeveelheden gegevens van talloze personen opvragen. Als we 100 personen vragen naar hun leeftijd, geslacht, bur-

gerlijke staat (gescheiden, getrouwd, alleenstaand), aantal kinderen en hun beroep, dan levert dit al 500 stukjes afzonderlijke informatie op. Hoewel dit nog niet veel voorstelt in vergelijking met veel onderzoek dat wordt verricht, is het niet erg zinvol om deze 500 metingen in je onderzoeksverslag weer te geven. Zulke onverwerkte informatie noemen we de **ruwe gegevens**. Een statistische analyse moet echter meer doen dan alleen maar de ruwe ingrediënten tonen. Het is nodig om de gegevens op een manier structuur te geven die de belangrijkste trends *op effectieve wijze presenteert*. Als je geen structuur in je gegevens aanbrengt, kun je net zo goed de lezer kopieën van je vragenlijsten of waarnemingsschema's geven, zodat ze die zelf kunnen interpreteren.

Er zijn maar heel weinig regels voor het maken van tabellen en grafieken in de statistiek, zolang ze maar duidelijk en beknopt zijn; ze moeten snel de belangrijkste trends in de gegevens kunnen overbrengen. Tabellen en grafieken die de taak van communicatie niet vereenvoudigen zijn zinloos. De beste manier om na te gaan of de tabellen en grafieken goed zijn, is om anderen te vragen uit te zoeken wat ze betekenen. Tabellen die voor anderen onduidelijk zijn, zijn in het algemeen overbodig.

Beschrijvende statistiek bestaat voornamelijk uit betrekkelijk eenvoudige visuele en numerieke technieken om de belangrijkste kenmerken van je gegevens te beschrijven. Onderzoekers kunnen beschrijvende statistieken maken om de belangrijkste kenmerken van de gegevens naar anderen door te geven, maar in eerste instantie worden ze door de onderzoekers zelf gebruikt om de verdeling van de reacties van de deelnemers in het onderzoek te kunnen begrijpen. Je moet een beschrijvende statistische analyse nooit beschouwen als een overbodig of triviaal stadium in het onderzoek. Zo'n analyse geeft waarschijnlijk meer informatie dan enig ander aspect van gegevensanalyse.

Het onderscheid tussen nominale gegevens (categoriegegevens) en numerieke scores dat in het vorige hoofdstuk werd beschreven, is van belang voor het kiezen van de juiste tabellen en grafieken.

ONDERZOEKSKWESTIE

Een van de fouten die je heel gemakkelijk kunt maken bij een onderzoek is om deelnemers in je onderzoek in staat te stellen meer dan een antwoord op een vraag te geven. Als je bijvoorbeeld mensen vraagt om hun favoriete tv-programma te noemen en elke proefpersoon meer dan een antwoord mag geven, zul je erachter komen dat het heel lastig is om de gegevens grondig te analyseren. Neem voorlopig maar van ons aan dat statistiek in het algemeen niet erg goed kan omgaan met meervoudige antwoorden. Zeker, het is mogelijk om tabellen en grafieken op te stellen, maar het wordt lastig om sommige van de meer geavanceerde statistische procedures hierop toe te passen. Soms lees je opmerkingen dat de totalen in een tabel meer zijn dan het aantal deelnemers aan het onderzoek. Dat komt dan doordat de onderzoeker meervoudige antwoorden voor een variabele toegelaten heeft. Zorg er dus voor dat de deelnemers aan je onderzoek voor elke variabele die je meet maar een antwoord kunnen geven, om te voorkomen dat je een kuil voor jezelf graaft. Als je de analyse van je gegevens in detail plant voordat je ze verzamelt, kun je op eventuele moeilijkheden anticiperen.