

Inhoud

Lijst van figuren	xiii
Lijst van tabellen	xix
1 Inleiding statistische verwerkingsmethoden	1
1.1 Inleiding	1
1.2 Empirisch onderzoek: mix van kennis en vaardigheden	2
1.3 Een case ter illustratie	6
1.4 Variabelen en observaties: bouwstenen van empirisch onderzoek	6
1.4.1 Nominale, ordinale, interval- en ratiovariabelen	8
1.4.2 Het degenereren en opwaarderen van meetniveaus	11
1.5 Samenhang, afhankelijkheid en causaliteit	12
1.5.1 Statistische samenhang	13
1.5.2 Statistische samenhang versus causaliteit	14
1.5.3 A priori en a posteriori causaliteit	16
1.6 Het causale relatieschema: netwerken met variabelen	17
1.6.1 Achtergrond van het relatieschema	17
1.6.2 Het opstellen van een relatieschema	20
1.6.3 Het relatieschema voor de voorbeeldcase	21
1.6.4 Model en werkelijkheid: <i>In & Uit</i>	23
1.7 Structuur van het boek	25
1.8 Literatuur	28
2 Introductie SPSS for Windows	31
2.1 De structuur van <i>SPSS for Windows</i>	31
2.2 Schermopbouw van <i>SPSS for Windows</i>	33
2.2.1 Onderdelen van het datavenster	33
2.2.2 Invoeren van gegevens in het datavenster	34
2.2.3 Aanpassen kenmerken variabelen in het datavenster	35
2.2.4 Invoeren van commando's in het syntaxvenster	36
2.3 Bestandsbeheer via het <i>File</i> -menu	39
2.3.1 Nieuwe vensters beginnen met <i>New</i>	40
2.3.2 Bestaande bestanden openen met <i>Open</i>	40
2.3.3 Ascii-gegevens inlezen met <i>Read Text Data</i>	43
2.3.4 Gegevens opslaan met <i>Save as</i>	45
2.3.5 Bestandsinformatie via <i>Display Data File Information</i>	46

2.3.6	Printen van informatie met <i>Print</i>	46
2.4	Aanpassingen maken via het <i>Edit</i> -menu	46
2.4.1	Overzicht van de <i>Edit</i> -instructies	47
2.4.2	Wijzigen voorkeursinstellingen	48
2.5	Algemene bewerkingen van gegevens via het <i>Data</i> -menu	51
2.5.1	Attribuutinformatie toevoegen aan variabelen	52
2.5.2	Bewerkingen aan observaties	53
2.5.3	Ontwerpen van een experimentele opzet	55
2.5.4	Algemene bewerkingen aan de observaties	58
2.6	Specifieke bewerkingen via het <i>Transform</i> -menu	60
2.6.1	Nieuwe variabelen maken met <i>Compute</i>	62
2.6.2	Hercoderen van uitkomsten met <i>Recode</i>	65
2.6.3	Rangnummers definiëren met <i>Rank Cases</i>	65
2.6.4	Overige instructies voor specifieke bewerkingen	68
2.7	Ondersteunende instructies via de <i>View</i> - en <i>Utilities</i> -menu's	68
2.8	Gegevensverwerking via de <i>Analyze</i> - en <i>Graphs</i> -menu's	70
3	Concepten statistische verwerkingsmethoden	73
3.1	Inleiding	73
3.2	Methoden van steekproeftrekken	74
3.2.1	De aselechte steekproef	74
3.2.2	De gestratificeerde steekproef	75
3.2.3	De getrapte steekproef	76
3.2.4	De systematische steekproef	76
3.2.5	Theorie en praktijk	77
3.3	Beschrijvende statistiek	78
3.3.1	Grafieken, tabellen en kengetallen	78
3.3.2	Kengetallen voor centrale en niet-centrale tendentie	81
3.3.3	Kengetallen voor de spreiding van de gegevens	88
3.4	Kansen, kansvariabelen en kansverdelingen	92
3.4.1	Kansen en relatieve frequenties	93
3.4.2	Kansvariabelen en kansverdelingen	95
3.4.3	Verwachtingen van kansvariabelen	100
3.4.4	Een uitgewerkt voorbeeld	105
3.5	Enkele veelgebruikte kansverdelingen	109
3.5.1	Discrete kansverdelingen	109
3.5.2	Continue kansdichtheidsverdelingen	118
3.6	Aselechte steekproeven en steekproeffuncties	128
3.6.1	De aselechte steekproef en steekproefvariatie	128
3.6.2	Steekproeffuncties	131
3.6.3	Kansverdelingen van steekproeffuncties	132
3.7	Inferentiële statistiek: schatten	138
3.7.1	Schatters en schattingen	139
3.7.2	Betrouwbaarheidsintervallen voor μ , p en σ^2	142
3.7.3	Schatten met een gestratificeerde steekproef	144
3.8	Inferentiële statistiek: toetsen	145
3.8.1	Handhaven en verwerpen	146

3.8.2	Het 7-stappenschema	148
3.8.3	Fout van de tweede soort	152
3.8.4	Significantie en relevantie	154
3.9	Omvang van de steekproef	155
3.9.1	Bepalen steekproefomvang bij schatten	155
3.9.2	Bepalen steekproefomvang bij toetsen	158
4	Univariate methoden	161
4.1	Inleiding	161
4.2	Overzichten met <i>Reports</i> en <i>Tables</i>	162
4.2.1	Overzichten van uitkomsten met <i>Reports</i>	162
4.2.2	Nette tabellen met <i>Tables</i>	171
4.3	Het maken van grafieken via <i>Graphs</i>	176
4.3.1	Histogrammen met <i>Histogram</i>	176
4.3.2	Staafdiagrammen met <i>Bar</i>	179
4.3.3	Cirkeldiagrammen met <i>Pie</i>	182
4.4	Het analyseren van kwantitatieve gegevens	184
4.4.1	Het gebruik van <i>Descriptives</i> voor beschrijvende analyse	186
4.4.2	Het gebruik van <i>Descriptives</i> voor schatten	190
4.4.3	Het gebruik van <i>Descriptives</i> voor toetsen	193
4.4.4	Gedetailleerde analyses met <i>Explore</i>	200
4.4.5	Analyse van deelpopulaties met <i>Means</i>	203
4.5	Het analyseren van kwalitatieve gegevens	205
4.5.1	Het gebruik van <i>Frequencies</i> voor beschrijvende analyse	207
4.5.2	Het gebruik van <i>Frequencies</i> voor schatten	211
4.5.3	Het gebruik van <i>Frequencies</i> voor toetsen	212
4.6	Uitbijters	214
4.6.1	Het opsporen van uitbijters	215
4.6.2	Het behandelen van uitbijters	217
5	Samenhang tussen twee variabelen met gelijk meetniveau	219
5.1	Onderzoek naar paarsgewijze samenhangen	219
5.2	Overzicht bivariate onderzoekstechnieken	221
5.3	Spreidingsdiagrammen voor het weergeven van samenhangen	226
5.3.1	Het maken van spreidingsdiagrammen via <i>Scatter</i>	227
5.3.2	Spreidingsdiagrammen voor een indruk van de aard van samenhang	229
5.3.3	Spreidingsdiagrammen voor het opsporen van uitbijters	232
5.4	Samenhangen tussen twee kwantitatieve variabelen	235
5.4.1	Theoretische achtergrond van de correlatiecoëfficiënt	235
5.4.2	Het berekenen van correlatiecoëfficiënten via <i>Correlate/Bivariate</i>	245
5.4.3	Interpretatie van de uitvoer van <i>Correlate/Bivariate</i>	247
5.4.4	Het gebruik van de correlatiecoëfficiënt voor toetsen	250
5.4.5	De gekwadraterde correlatiecoëfficiënt als mate van verklaring	262
5.5	Samenhang tussen twee ordinale variabelen	264

5.6	Samenhang tussen twee nominale variabelen	265
5.6.1	Theoretische achtergrond van de χ^2 -toets	266
5.6.2	Het maken van kruistabellen met <i>Crosstabs</i>	280
5.6.3	Het gebruik van <i>Crosstabs</i> voor toetsen	285
5.6.4	Gebruik <i>Crosstabs</i> voor het beschrijven van aard en mate van samenhang	290
5.6.5	Kruistabellen voor geaggregeerde gegevens	294
5.7	Correlatie, rangcorrelatie en kruistabel	296
6	Vergelijken van gemiddelden van deelpopulaties	301
6.1	Inleiding	301
6.2	Vergelijken gemiddelden versus toetsen onafhankelijkheid	302
6.3	Het vergelijken van twee gemiddelden via de <i>t</i> -toets	303
6.3.1	Het uitvoeren van de <i>t</i> -toets met <i>Independent-Samples T Test</i>	304
6.3.2	Theoretische achtergrond van de <i>t</i> -toets voor onafhankelijke steekproeven	306
6.3.3	Het gebruik van <i>Independent Samples T Test</i> voor toetsen	316
6.3.4	Het gebruik van <i>Independent Samples T Test</i> voor schatten	330
6.4	Het vergelijken van twee of meer gemiddelden met variantieanalyse	332
6.4.1	Het uitvoeren van variantieanalyse met <i>One-Way Anova</i>	333
6.4.2	Theoretische achtergrond van variantieanalyse	336
6.4.3	Het gebruik van <i>One-Way Anova</i> voor toetsen	349
6.4.4	Het gebruik van <i>One-Way Anova</i> voor schatten	356
6.5	Vergelijken gemiddelden van afhankelijke steekproeven	369
6.5.1	De <i>t</i> -toets voor gepaarde waarnemingen	370
6.5.2	Variantieanalyse voor afhankelijke steekproeven	379
6.6	Verdelingsvrije toetsen voor het vergelijken van deelpopulaties	388
7	Doorkruisende variabelen	395
7.1	Inleiding	395
7.2	Een extreem voorbeeld van doorkruisendheid	396
7.3	Het belang van het opsporen van doorkruisende effecten	399
7.4	Methoden voor het opschonen van samenhangen	404
7.4.1	Procedure voor onderzoek naar verstorende effecten	405
7.4.2	Opschonen via vergelijken deelsteekproeven	407
7.4.3	Opschonen via de partiële of part correlatie	411
7.4.4	Opschonen via variantieanalyse met covariaat	415
7.5	Toepassingen voor de voorbeeldcase	418
7.5.1	Opschonen via vergelijken deelsteekproeven	418
7.5.2	Opschonen via de partiële of part correlatie	427
7.5.3	Opschonen via variantieanalyse met covariaat	428
7.6	Controleren voor verstorende invloeden via <i>matching</i>	432
7.7	Slotopmerkingen over doorkruisende invloeden	433
8	Meervoudige samenhangen: het 2-factor anova-model	435
8.1	Inleiding	435

8.2	Uitvoeren van 2-factor variantieanalyse	436
8.3	Theoretische achtergrond 2-factor variantieanalyse	441
8.4	Interpretatie van de uitvoer van de 2-factor variantieanalyse	447
8.5	Nadere uitleg van het interactie-effect	452
8.6	Interactie-effecten en schijnverbanden	454
8.7	Variantieanalyse en regressieanalyse: het volgordeprobleem	457
8.7.1	Opschoningsprocedures in variantieanalyse en regressieanalyse	458
8.7.2	Illustratie van het volgordeprobleem	461
9	Regressieanalyse	471
9.1	Inleiding	471
9.2	Regressieanalyse in <i>SPSS for Windows</i>	473
9.3	Theoretische achtergrond regressieanalyse	478
9.3.1	Het populatieregressiemodel	479
9.3.2	Het schatten van de onbekende modelparameters	484
9.3.3	Het steekproefregressiemodel	488
9.3.4	Eigenschappen van de residuen	489
9.3.5	Voorspellingsintervallen	492
9.3.6	Verklarende bijdrage van het model als geheel	495
9.4	Interpretatie van de uitvoer	500
9.4.1	Plausibiliteit van de regressiecoëfficiënten	503
9.4.2	Standaardfouten en <i>t</i> -waarden regressiecoëfficiënten	504
9.4.3	Mate van verklaring R^2 en <i>F</i> -toets	510
9.4.4	Standaardrapportage van de resultaten	513
9.4.5	Problemen bij het beoordelen van de resultaten	514
9.5	Het toetsen van meervoudige restricties	515
9.6	Gestandaardiseerde regressiecoëfficiënten	517
9.7	Het probleem van niet-lineariteit	520
9.8	Het probleem van multicollineariteit	526
9.9	Dummyvariabelen en interactie-effecten	533
9.10	Analyse van de residuen	540
9.10.1	Een eerste indruk van de residuen	541
9.10.2	Nadere analyse residuen	546
9.10.3	Uitbijters en hun relatieve positie in de puntenwolk	554
9.10.4	Autocorrelatie	556
9.11	Gevaren van mechanistische modelbouw	562
9.12	Simultane relaties en het identificatieprobleem	565
9.13	Het beoordelen van de kwaliteit van regressiemodellen	569
	Bijlage A Gegevens van de voorbeeldcase	573
	Bijlage B Statistische tabellen	577
	Bibliografie	587
	Register	589

Hoofdstuk 1

Inleiding statistische verwerkingsmethoden

1.1 Inleiding

Statistiekboeken zijn vaak dik en ontoegankelijk, indrukwekkende aaneenrangingen van tabellen, plaatjes, formules, toepassingsmogelijkheden, uitzonderingen en voorwaarden voor toepassing. Ze zijn voor kwantitatief ingestelde studenten prettig leesbaar, maar ogen voor veel andere studenten als een onneembare vesting. Dan te bedenken dat de veelheid van aangeboden statistische technieken slechts een fractie is van wat allemaal beschikbaar is, zowel in termen van toepassingen als statistisch theoretische diepgang. Deze ongemakkelijke suggestie hangt mogelijk samen met de opzet van deze boeken, die aansluit bij de logische structuur van de begripsontwikkeling in de theoretische grondslag van al die praktische toepassingen.

Dit boek heeft een andere insteek, waarbij het accent ligt op de talloze problemen die zich voordoen bij de analyse van kwantitatieve gegevens. De opzet van dit boek volgt de structuur van de problemen die zich in opeenvolgende fasen van toegepast statistisch onderzoek voordoen en niet zozeer de structuur van de theoretische begripsontwikkeling. Dit heeft onder meer tot gevolg dat de rol van oefenmateriaal in de vorm van sommen in dit boek beperkt is. In plaats hiervan wordt de toepassing van statistische technieken uitgebreid geïllustreerd met een praktijkvoorbeeld. Oefenmateriaal in de vorm van opgaven wordt aangeboden in de vorm van een werkboek, dat los van dit leerboek onder dezelfde naam verkrijgbaar is.

Bij de uitwerking van het praktijkvoorbeeld wordt aandacht besteed aan uiteenlopende praktische problemen, zoals het vertalen van vage verhalen in een concrete statistische probleemstelling, het bouwen van modellen, het interpreteren en beschrijven van gegevens en het beoordelen van de relevantie en de plausibiliteit van statistische resultaten. De theoretische onderbouwing van de gebruikte statistische technieken blijft beperkt tot wat nodig is voor een

intuïtief begrip van de achtergrond van deze technieken en van de regels en uitzonderingen die bij toepassing van belang zijn. Waar nodig, wordt hiervoor verwezen naar andere tekstboeken.

Dit inleidende hoofdstuk is verder opgebouwd als volgt. In paragraaf 1.2 wordt de breedte van statistische verwerkingsmethoden geïllustreerd met een blik op de verschillende kennisrediënten van statistisch onderzoek. Daarna wordt in paragraaf 1.3 de voorbeeldcase toegelicht, die de leidraad vormt bij de statistische analyses in dit boek. In paragraaf 1.4 wordt ingegaan op de bouwstenen van empirisch onderzoek: de variabelen. Vervolgens wordt in paragrafen 1.5 en 1.6 aandacht besteed aan relaties tussen variabelen. Paragraaf 1.5 gaat hierbij in op het verschil tussen causaliteit en statistische samenhang, terwijl paragraaf 1.6 het schematisch uitbeelden van die relaties beschrijft. De structuur van het boek wordt uiteengezet in paragraaf 1.7. Paragraaf 1.8 besluit met een kort overzicht van relevante literatuur.

1.2 Empirisch onderzoek: mix van kennis en vaardigheden

Kwantitatief empirisch onderzoek is te typeren als het creatief combineren van uiteenlopende soorten kennis en vaardigheden. Kennis van statistische methoden en technieken is hierbij slechts een van de onderdelen; zij het een zeer wezenlijk onderdeel. In deze paragraaf worden verschillende kennisvereisten van empirisch onderzoek aangestipt om de boeiende rol van statistische methoden toe te lichten. Het gaat hierbij om een schets van de plaats van statistische methoden en technieken in de praktijk van het empirisch onderzoek. Een formele inbedding van statistische verwerkingsmethoden in een conceptueel kader vindt plaats in een volgend hoofdstuk.

Kennis van praktijk en theorie

Verreweg de belangrijkste kennisvereisten in empirisch onderzoek zijn kennis van het praktische reilen en zeilen op het onderzoeksterrein en kennis van de relevante literatuur en theorievorming op het toepassingsgebied. Wanneer men bijvoorbeeld onderzoek doet naar het prijsgedrag van bloemengroothandelaren, heeft men kennis nodig van de wijze waarop de handel in bloemen is georganiseerd en van de dagelijkse handel en wandel van bloemenhandelaren zelf, maar ook van de literatuur op het gebied van industriële organisatie en van verschillende prijstheorieën. Wanneer men marktonderzoek verricht voor een bedrijf, heeft men tenminste een vaag inzicht nodig in de kenmerken van de doelgroep, de distributiekkanalen en de op de markt te brengen producten, maar ook van specifieke marketingtheorieën. Wanneer men onderzoek doet naar de implementatie van overheidsbeleid, heeft men inzicht nodig in de aard en de inhoud van voorgestelde maatregelen, de uitvoerende instanties en de doelgroepen van beleid, maar ook allerlei bestuurskundige theorieën

die op deze problematiek van toepassing zijn. Kennis van de praktijk, dat wil zeggen op wie of wat heeft het onderzoek betrekking, en kennis van de theorie, dat wil zeggen wat hebben anderen over het probleem bedacht en geschreven, zijn noodzakelijke onderdelen voor het doen van empirisch onderzoek. Ze bepalen de formulering van de probleemstelling, de wijze waarop problemen worden geanalyseerd en de praktische en theoretische bruikbaarheid van de onderzoeksresultaten.

Formuleren probleemstelling

Een tweede vereiste bij empirisch onderzoek is vaardigheid in het afbakenen van een onderzoeksterrein en het formuleren van heldere probleemstellingen. Voorbeelden zijn: verschilt het prijsgedrag tussen verzendexporteurs en lijnrijders in de bloemengroothandel? Of, wat is de invloed van demografische kenmerken, zoals leeftijd en sociale klasse, op de bereidheid om een nieuw ontwikkeld product te kopen? Hoewel deze vaardigheid vanzelfsprekend lijkt, blijkt uit ervaring dat velen onvoldoende in staat zijn om het waarom, hoe en waartoe van het onderzoek duidelijk te maken aan beoordelende instanties. Dit hangt wellicht samen met het feit dat een adequate probleemformulering inzicht vereist in zowel de situatie op het onderzoeksterrein, als de onderzoekstechnieken die geëigend zijn om de gestelde problemen te analyseren. Zelfs wanneer een probleemstelling is gedefinieerd door de opdrachtgever, is het nodig om tot een zelfstandige afbakening van het onderzoeksterrein te komen, aangezien er geen garantie is dat de opdrachtgever beschikt over voldoende inzicht in praktijk, theorie en mogelijke oplosmethoden om het probleem afdoende te kunnen structureren. Een zorgvuldig geformuleerde probleemstelling is in het geval van statistisch onderzoek bovendien van belang voor de aard en de omvang van de te verzamelen gegevens en voor de toepasbaarheid van specifieke statistische technieken.

Meetproblemen

Een derde ingrediënt van kwantitatief onderzoek is het omgaan met allerlei meetproblemen. Om te beginnen zal men maatstaven moeten verzinnen om empirische verschijnselen te meten. Zo zal men bij onderzoek naar de prijsvorming in de bloemengroothandel precies moeten vastleggen wat men bedoelt met de prijs en het product van bloemenhandelaren. Bij onderzoek naar de invloed van sociale status op de geneigdheid tot aankoop van een nieuw product, zal men goed moeten nadenken over de vraag hoe sociale status en koopgeneigdheid zinvol kunnen worden samengevat met numerieke grootheden.

Na het definiëren van maatstaven, moet een instrument worden ontwikkeld om de benodigde gegevens te verzamelen. In gedragswetenschappen zoals economie, sociologie en psychologie wordt hiertoe vaak gebruikgemaakt van enquêtes. In andere takken van wetenschap, zoals natuurkunde, biologie en agrarische wetenschappen, wordt veelal gebruikgemaakt van proefopstellingen al dan niet in laboratoria. Het grote voordeel van deze proefopstellingen zijn de betere controlemogelijkheden. In sommige gevallen wordt gebruik-

gemaakt van gegevens die door anderen, bijvoorbeeld het Centraal Bureau voor de Statistiek, zijn verzameld.

Bij dit alles moet een beslissing genomen worden over de aard van de te verzamelen gegevens: moeten de gegevens betrekking hebben op afzonderlijke personen of op groepen personen (we spreken dan van geaggregeerde gegevens); moeten de gegevens bestaan uit momentopnamen van processen in een bedrijf gevolgd door de tijd (tijdreeksgegevens) of uit waarnemingen van processen in verschillende bedrijven op een tijdstip (dwarsdoorsnede gegevens). Nadat de gegevens zijn verzameld, is inzicht nodig in de validiteit van de metingen: heeft men gemeten wat men wilde weten? Gemeten gegevens hebben lang niet altijd de beoogde betekenis. Wanneer een respondent in een enquête *ja* antwoordt, hoeft dit nog niet te betekenen dat deze ook *ja* bedoelt. Wanneer cijfers over het ziekteverzuim worden verzameld, dan hoeven tellingen gedaan door werkgevers geenszins overeen te stemmen met de opgaven die door de werknemers verstrekt worden. Wanneer historische verkoopcijfers worden verzameld, kan het uitmaken of de metingen hebben plaatsgevonden bij de productie-afdeling of bij de voorraadmagazijnen. Methoden om de validiteit van metingen te beoordelen en te verbeteren, vindt men in literatuur over onderzoeksmethoden en -technieken, bijvoorbeeld marktonderzoek, sociaal-wetenschappelijk onderzoek, kwantitatief-economisch en econometrisch onderzoek.

Modelbouw

Een vierde element van kwantitatief onderzoek, dat nauw samenhangt met de interpretatie van gevonden resultaten, is modelbouw. Stel dat uit een onderzoek naar arbeidsomstandigheden in een bepaald bedrijf blijkt dat vrouwen een korter gemiddeld ziekteverzuim vertonen dan mannen. Mag men dan concluderen dat dit verschil een direct gevolg is van het geslacht van respondenten? Een alternatieve verklaring voor het geconstateerde verschil is dat in de gehanteerde steekproef vrouwen jonger zijn dan mannen. Zodra de samenhang tussen ziekteverzuim en geslacht wordt gecorrigeerd voor dit leeftijdseffect, zouden mannen en vrouwen dezelfde gemiddelde ziekteduur kunnen vertonen. Een ander voorbeeld betreft het voorspellen van verkoopvolumes van een detailhandelsvestiging. Kunnen dergelijke voorspellingen worden gedaan op basis van uitsluitend historische verkoopcijfers of moet tevens rekening worden gehouden met mogelijke veranderingen in prijs, promotie-inspanningen en andere instrumenten van de marketingmix? Om deze vragen te kunnen beantwoorden is het nodig om modellen te maken van de processen die zich in de ogen van de onderzoeker afspelen in de werkelijkheid.

Bij kwantitatief onderzoek moet dus ook aandacht besteed worden aan methoden van modelontwikkeling en aan criteria voor het kiezen tussen verschillende modellen. In veel statistiekboeken blijven deze vragen buiten beschouwing. De rekenprocedures worden gedetailleerd aan de orde gesteld, maar de betekenisgeving wordt beschouwd als niet behorend tot het vakterrein van de statistiek. Wanneer men echter beweert dat statistiek een belangrijk hulpmiddel is in praktische onderzoekssituaties, is deze houding natuurlijk niet verdedigbaar. In dit boek wordt dan ook vrij uitgebreid op dit soort vragen

ingegaan, met name wanneer er een relatie ligt met de meer technische aspecten van statistische verwerkingsmethoden.

Gebruik computersoftware

Een vijfde vereiste voor kwantitatief onderzoek is kennis van statistische computerprogrammatuur. Het is tegenwoordig niet meer nodig om statistische berekeningen met de hand te maken, vanwege de beschikbaarheid van rekenmachines en razendsnelle, eenvoudig te bedienen computers. Er bestaan tal van computerprogramma's voor het uitvoeren van statistische bewerkingen. De voorbeelden in dit boek zijn gebaseerd op *SPSS for Windows*. Andere veelgebruikte programma's, zowel in statistiekboeken als in de praktijk, zijn: *SAS*, *Minitab* en *Stata*.

De beschikbaarheid van computerprogrammatuur is niet zonder meer een zegen. Maar al te doet de 'wet van het apparaat' opgeld: geef iemand een computer en men rekt net zolang tot de meest onzinnige correlaties tevoorschijn komen, die achteraf worden voorzien van redeneringen om de gevonden samenhangen plausibel te maken. Bovendien genereren softwareprocedures doorgaans veel meer resultaten dan voor de beantwoording van de onderzoeksvragen strikt noodzakelijk is. Het identificeren van de benodigde informatie is dan een belangrijke vaardigheid.

In dit boek wordt gepleit voor het opstellen van dergelijke redeneringen vooraf. Dit maakt dat niet-verwachte correlaties die toch blijken te bestaan, worden onderzocht op de mogelijkheid dat het onzincorrelaties zijn. Op eenzelfde manier zal aandacht worden gegeven aan verwachte correlaties die niet blijken te bestaan. Op ieder ogenblik van de analyse wordt men gedwongen de logische structuur van het probleem voor ogen te houden.

Kennis statistische methoden en technieken

Ten slotte heeft men kennis van statistische technieken zelf nodig. Men zal moeten weten wanneer welke statistische technieken kunnen worden toegepast en wat de betekenis is van de uitkomsten. Hiertoe is een meer dan oppervlakkige kennis nodig van de theoretische achtergrond van statistische methoden en technieken.

Kennis van statistische verwerkingsmethoden is slechts een van de benodigde ingrediënten van kwantitatief onderzoek. Voor het verkrijgen van betrouwbare conclusies zijn alle vijf genoemde kenniselementen noodzakelijk: kennis van het onderzoeksveld, de meetmethoden, de statistische theorie, de statistische computerprogrammatuur en de valkuilen bij de interpretatie. Hierbij geldt dat de keten zo sterk is als de zwakste schakel, hetgeen de statistiek maakt tot een moeilijk, maar uitdagend vak.

1.3 Een case ter illustratie

Voor het toelichten van uiteenlopende problemen van statistisch onderzoek wordt in dit boek gebruikgemaakt van een fictieve casus. Het nadeel van een fictieve casus is dat er een zekere afstand ontstaat tot echt empirisch onderzoek. Een voordeel is echter dat de samenhang tussen verschillende onderzoekstechnieken duidelijker kan worden voorgesteld. Het onderwerp van de case heeft betrekking op het meten van de verkoopprestaties van vertegenwoordigers van een groot bedrijf.

Casus: verkoopprestaties vertegenwoordigers

Een al oud familiebedrijf heeft een goede naam opgebouwd met het fabriceren van technisch eenvoudige instrumenten. Door fusie met een ander bedrijf heeft men zich verzekerd van een concurrerende positie in het computertijdperk. Het resultaat van de fusie is een nieuw bedrijf met twee divisies die in verschillende markten opereren. De ene divisie produceert goederen met een relatief lage technische complexiteit, terwijl de andere divisie actief is in markten met producten van een hoog technisch niveau. Korte tijd na voltooiing van de fusie is bij de directie een sterke behoefte ontstaan aan inzicht in de verkoopprestaties van het personeel. Een grootscheepse prestatiemeting stuit echter op onoverkomelijke weerstanden bij de ondernemingsraad. De directie besluit daarop genoeg te nemen met een beperkte steekproef van 90 vertegenwoordigers van de verkoopafdelingen van de twee divisies. De anonimiteit van het onderzoek maakt dat de verkoopprestaties van deze vertegenwoordigers zonder al te veel bezwaren kunnen worden geanalyseerd.

Toelichting casus

Per vertegenwoordiger is een zestal kenmerken verzameld: verkoopprestatie, intelligentie, sociale vaardigheid, diensttijd, opleiding en afdeling. In tabel 1.1 staat een overzicht van deze variabelen. De verkoopprestatie van vertegenwoordigers wordt weergegeven door de variabele *Verkoop*. De waarden van *Verkoop*, die zijn afgeleid van de gerealiseerde omzetten, liggen rondom een score van 100. De factoren die deze verkoopprestatie mogelijk verklaren, zijn: de intelligentie van vertegenwoordigers (*IQ*), de sociale vaardigheid (*Sociaal*), het aantal maanden dat een verkoper in dienst is (*Dienst*), type opleiding (*Opleid*) en ten slotte de afdeling of divisie waar de vertegenwoordiger werkt (*Branche*). In bijlage A zijn de uitkomsten van de gehele steekproef afgedrukt.

1.4 Variabelen en observaties: bouwstenen van empirisch onderzoek

Empirisch onderzoek vindt plaats op basis van waargenomen kenmerken van personen, bedrijven, objecten en verschijnselen. De eenheden waarop het onderzoek betrekking heeft, personen, bedrijven et cetera, worden ook wel aan-

Tabel 1.1 Overzicht variabelen in de voorbeeldcase ($n = 90$)

Variabele	Omschrijving en eventuele codering
<i>Verkoop</i>	Verkoopprestatie van vertegenwoordigers gemeten op een schaal rondom 100, 999 = <i>missing value</i>
<i>IQ</i>	Indicator van het IQ van vertegenwoordigers op een schaal rondom 100
<i>Sociaal</i>	Indicator van het sociaal gedrag van vertegenwoordigers op een schaal van 1 tot 100
<i>Dienst</i>	Aanstellingsduur in maanden
<i>Opleid</i>	Opleidingsgraad van vertegenwoordigers: 1 = lage vooropleiding, 2 = hoge vooropleiding, 3 = alleen interne cursussen
<i>Branche</i>	Technische ontwikkeling van de afdeling waar de vertegenwoordiger werkt: 1 = hoog (elektronica), 2 = laag (apparaten)

geduid met onderzoekselementen. De waargenomen kenmerken van de onderzoekselementen worden samengevat met variabelen: grootheden die verschillende waarden kunnen hebben voor verschillende elementen. De verschillende kenmerken van een enkel onderzoekselement worden samengevat in een observatie. Deze definities van variabelen en observaties worden hierna toegelicht aan de hand van het praktijkvoorbeeld.

Variabelen en observaties in de voorbeeldcase

Het gebruik van variabelen en observaties leidt in de praktijk tot een herkenbare structuur van gegevensopslag. Een voorbeeld staat in tabel 1.2, waarin de waargenomen kenmerken van zeven vertegenwoordigers uit de voorbeeldcase worden getoond.

In de kolommen staan de variabelen, de uitkomsten van een waargenomen kenmerk voor alle vertegenwoordigers. Zo geeft de kolom met opschrift 'Verkoop' de uitkomsten van de variabele *Verkoop* voor de verschillende vertegenwoordigers: 121, 104, 104, 95, 98, 106, 999. De waarde '999' betekent dat de verkoopprestatie van de betreffende vertegenwoordiger niet bekend (*missing*) is, dergelijke ontbrekende waarden worden vaak ook met een '.' aangegeven. In de rijen staan de observaties, de uitkomsten van alle kenmerken van een enkele vertegenwoordiger. Op de rij met observatienummer 3 staan de waargenomen kenmerken van een vertegenwoordiger met een verkoopprestatie van 104, een IQ gelijk aan 99, een score sociale vaardigheid gelijk aan 86, een dienstduur van 40 maanden, een opleidingsniveau bestaande uit alleen interne cursussen, die werkzaam is in de apparatenbranche.

Variabelen en kansvariabelen

In de statistische theorie waarop de statistische methoden en technieken zijn gebaseerd, hebben we te maken met *kansvariabelen*: variabelen waarvan de waarde pas bekend is op het moment dat de waarneming wordt verricht. Kenmerkend voor kansvariabelen is dat men wel inzicht heeft in het bereik van

Tabel 1.2 Voorbeeld structuur gegevensopslag in de voorbeeldcase

Observatie	Variabelen					
	Verkoop	IQ	Sociaal	Dienst	Opleid	Branche
1	121	119	86	37	2	1
2	104	102	63	8	3	1
3	104	99	86	40	3	2
4	95	92	72	23	1	1
5	98	103	78	58	1	2
6	106	109	67	23	2	1
7	999	103	76	4	2	2

mogelijke uitkomsten, maar dat men niet weet welke uitkomst de kansvariabele in een specifiek geval zal hebben. Wanneer men bijvoorbeeld onderzoek doet naar de doelmatigheid van busvervoer, dan is het verschil tussen werkelijke aankomsttijd en gestelde aankomsttijd een kansvariabele: voorafgaand aan de feitelijke meting kan dit tijdsverschil verschillende waarden aannemen en is de uiteindelijke waarde onbekend. Nadat de metingen zijn verricht, is het waargenomen tijdsverschil een variabele.

Tijd: variabele en kansvariabele

Een bijzonder voorbeeld van een grootte die zowel een variabele als een kansvariabele kan zijn, is *tijd*. Wanneer tijd wordt opgevat als tijdsduur, bijvoorbeeld de rondetijd tijdens een schaatswedstrijd, dan is ze een kansvariabele: uiteenlopende waarden zijn denkbaar en de uiteindelijke waarde is vooraf onbekend. Wanneer tijd wordt opgevat als kalendertijd, dan is *tijd* een gewone variabele: de grootte neemt verschillende waarden aan, die op elk moment vooraf volledig bekend zijn.

Variabelen: consequenties verschillende meetniveaus

Een deel van de complexiteit van statistiek komt voort uit het feit dat variabelen verschillende meetniveaus kunnen hebben. Statistische technieken die geschikt zijn voor toepassing op variabelen met het ene meetniveau, zijn vaak niet geschikt voor het analyseren van variabelen met andere meetniveaus. Het gevolg is dat in de loop van de tijd een grote verscheidenheid aan technieken is ontstaan, waardoor het zicht op de gemeenschappelijke gedachtegang wordt belemmerd. Om door de bomen het bos te blijven zien, is het van belang de verschillende meetniveaus van variabelen uit elkaar te houden.

1.4.1 Nominale, ordinale, interval- en ratiovariabelen

Het is gebruikelijk om vier niveaus te onderscheiden waarop variabelen kunnen worden gemeten: het nominale, het ordinale, het interval- en het ratio-meetniveau. Figuur 1.1 toont een illustratie van de vier meetniveaus en de

hiermee samenhangende toegestane wiskundige bewerkingen. Nominale en ordinale variabelen worden ook wel aangeduid als *kwalitatieve* variabelen. De uitkomsten van deze variabelen verwijzen naar wederzijds uitsluitende groepen van elementen. De interval- en ratiovariabelen worden ook wel aangeduid als *kwantitatieve* variabelen. De uitkomsten van deze variabelen zijn verkregen uit tel- en meetprocessen. Hierna worden de vier meetniveaus meer gedetailleerd besproken.

Nominaal meetniveau

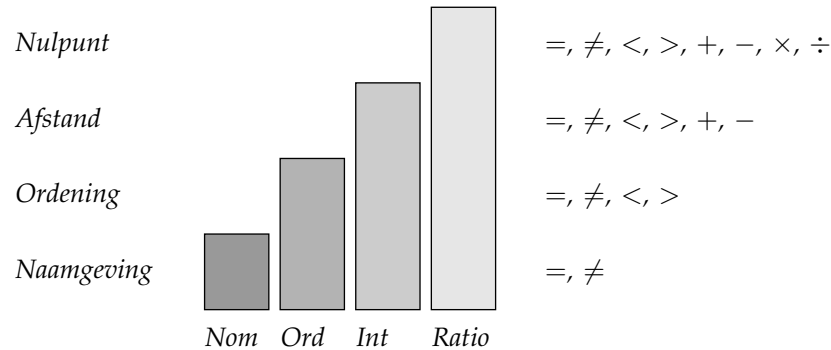
De *nominale* schaal is het laagste meetniveau waarop variabelen beschikbaar kunnen zijn. De uitkomsten van nominale variabelen hebben alleen betekenis als naamgeving. De enige voorwaarde die aan het nominale meetniveau wordt gesteld, is dat de verschillende uitkomsten verwijzen naar wederzijds uitsluitende en gezamenlijk uitputtende categorieën. Die enige bewerking die op de uitkomsten van nominale variabelen kan worden gedaan is vergelijken, dat wil zeggen beoordelen of twee uitkomsten al dan niet tot dezelfde categorie behoren ($=$, \neq).

Voorbeelden van nominale variabelen zijn provincie, geslacht, kleur ogen. Een voorbeeld in onderzoek naar tevredenheid over online boekingen van vliegtickets is de variabele 'vliegmaatschappij' met uitkomsten: 1 = British Airways, 2 = Ryanair, 3 = Lufthansa, 4 = Air France, et cetera. Voorbeelden van nominale variabelen in de voorbeeldcase zijn de opleidingsgraad van vertegenwoordigers (*Opleid*) en de afdeling waar men werkt (*Branche*); zie tabellen 1.1 en 1.2.

Merk op dat de gebruikte codering geen enkele natuurlijke betekenis heeft, anders dan de verwijzing naar de specifieke categorie waartoe een onderzoekselement behoort. De opmerking is van belang omdat in de praktijk nominale variabelen vaak met oplopende nummers zijn gecodeerd. Hiermee wordt een ordening in de uitkomsten gesuggereerd, die er feitelijk niet is. Een voorbeeld is de codering van de uitkomsten van de variabele *Opleid* in de voorbeeldcase. De uitkomsten 1, 2 en 3 verwijzen naar drie soorten opleiding (lage en hoge vooropleiding en interne cursussen), waarmee drie groepen vertegenwoordigers worden onderscheiden. De gedachte dat interne cursussen meer zouden voorstellen dan een hoge of lage vooropleiding is slechts suggestief, veroorzaakt door de gebruikte codering, ze wordt niet ondersteund door de informatie in tabel 1.1.

Ordinaal meetniveau

Daarnaast kunnen variabelen beschikbaar zijn op een *ordinale* schaal. De uitkomsten van ordinale variabelen verwijzen naar wederzijds uitsluitende categorieën onderzoekselementen, evenals de uitkomsten van nominale variabelen. In tegenstelling tot de uitkomsten van nominale variabelen impliceren de uitkomsten van ordinale variabelen een rangordering. Met de uitkomsten van ordinale variabelen kan dus niet alleen worden vastgesteld of onderzoekselementen tot dezelfde categorie behoren ($=$, \neq), maar kunnen onderzoekselementen ook geordend worden ($<$, $>$).



Figuur 1.1 Schema verschillende meetniveaus met toegestane bewerkingen

Voorbeelden van ordinale variabelen zijn: de antwoordcategorieën ‘sterk mee oneens/oneens/neutral/mee eens/sterk mee eens’ bij enquêtevragen; de uitkomsten ‘laag/laag-midden/hoog-midden/hoog’ bij een beschrijving van sociale klassen; het rangnummer van aankomst bij een beschrijving van het resultaat van een hardlooptrace; de klassen ‘klein/midden/groot’ bij een beschrijving van bedrijfsomvang.

Bij ordinale variabelen is weliswaar sprake van een natuurlijke volgorde van de uitkomsten, maar niet van een natuurlijke afstand tussen de uitkomsten. Het is dus niet juist te stellen dat het verschil tussen antwoordcategorieën ‘sterk oneens’ en ‘oneens’ even groot is als het verschil tussen ‘neutral’ en ‘eens’.

Interval meetniveau

Voorts kunnen variabelen zijn gemeten op een *intervalschaal*. Bij intervalvariabelen is er een natuurlijke interpretatie van de afstanden tussen de uitkomsten, naast de eerder beschreven natuurlijke ordening. Met uitkomsten van intervalvariabelen kunnen niet alleen onderzoekselementen worden vergeleken ($=$, \neq) en geordend ($<$, $>$), maar kunnen ook optellingen en verschillen op een zinvolle manier worden bepaald en gebruikt ($+$, $-$).

Voorbeelden van intervalvariabelen zijn: de temperatuur gemeten op een willekeurige niet-Kelvinschaal; en in het geval van het doelmatigheidsonderzoek bij het busvervoer de schoenmaat van de buschauffeur; het cijfer voor een multiple-choicetentamen, waarvan het aantal fouten dat nog net een voldoende oplevert, achteraf wordt vastgesteld zodanig dat 75% van de klas slaagt.

Bij intervalvariabelen is weliswaar sprake van een afstandsbeleg, waardoor het vergelijken van afstanden tussen uitkomsten betekenis heeft, maar niet van een natuurlijk nulpunt. Hierdoor hebben verhoudingen van uitkomsten geen natuurlijke interpretatie. Het is dus correct vast te stellen dat het verschil tussen schoenmaten 36 en 37 even groot is als dat tussen 42 en 43, maar het is niet juist te stellen dat schoenmaat 44 tweemaal zo groot is als schoenmaat 22.

Ratio meetniveau

Het beoordelen van verhoudingen tussen uitkomsten is wel mogelijk bij variabelen die zijn gemeten op een *ratioschaal*. De uitkomsten van dergelijke variabelen hebben een natuurlijke ordening, een natuurlijke afstand en een natuurlijk nulpunt.

Variabelen gemeten op een *ratioschaal* verschaffen de meest rijke informatie. Men kan de uitkomsten niet alleen vergelijken ($=$, \neq), ordenen ($<$, $>$), optellen en aftrekken ($+$, $-$), maar men kan er ook op een zinvolle manier vermenigvuldigingen en verhoudingsgetallen mee bepalen (\times , \div).

In het geval van het eerder gedefinieerde verschil tussen de werkelijke en aangekondigde vertrektijd van bussen, is het correct om op te merken dat iemand die reist met een 15 minuten te late bus driemaal zo lang heeft moeten wachten als iemand die reist met een 5 minuten te late bus. Andere voorbeelden van *ratiovariabelen* zijn: leeftijd, inkomen, het aantal werkzame personen in een bedrijf, het aantal ziekteverzuimdagen van een werknemer, et cetera. In het geval van de voorbeeldcase is de verkoopprestatie van medewerkers (*Verkoop*) gemeten op een *ratioschaal*.

Voor de meeste statistische procedures is het onderscheid tussen een *interval-schaal* en een *ratioschaal* van weinig betekenis, omdat het absolute nulpunt wordt vervangen door een kunstmatig nulpunt zoals het gemiddelde van de steekproef. In dit boek worden beide schalen dan ook door elkaar gebruikt.

1.4.2 Het degenereren en opwaarderen van meetniveaus

Variabelen op een *ratioschaal* kunnen altijd worden beschouwd als gemeten op een *interval-schaal*. Er wordt dan eenvoudigweg niet gebruikgemaakt van de beschikbare informatie over een interpreteerbaar nulpunt. Variabelen op een *interval-schaal* kunnen op hun beurt altijd worden beschouwd als gemeten op een *ordinale schaal*. De natuurlijke interpretatie van de afstand tussen twee uitkomsten wordt dan niet benut. Variabelen op een *ordinale schaal* kunnen altijd worden beschouwd als gemeten op een *nominale schaal*. Er wordt dan niet gebruikgemaakt van de natuurlijke interpretatie van de ordening in de uitkomsten. Statistische technieken voor toepassingen van *nominale variabelen* kunnen op deze manier worden gebruikt voor *alle* soorten van variabelen. Echter, de prijs die voor het degenereren van meetniveaus wordt betaald, is het ongebruikt laten van aanwezige informatie.

Het is niet correct om het meetniveau van variabelen op te waarderen. Het is dus niet juist om een *nominale* variabele te behandelen als een *ordinale* variabele. Immers, er wordt dan impliciet een eigenschap ('volgorde') toegevoegd aan de variabele, waarover geen informatie bekend is. Evenmin is het juist om een *ordinale* variabele te behandelen als gemeten op een *interval-schaal*. Immers, er wordt dan impliciet informatie over afstanden toegevoegd zonder dat deze is waargenomen. Het is dus niet zonder meer correct om een variabele met antwoordcategorieën 'helemaal oneens' tot 'helemaal eens' te behandelen als een *intervalvariabele*. Dit is erg jammer, omdat veel krachtige analysetechnieken voor kwantitatieve variabelen niet kunnen worden gebruikt voor het bestuderen van *ordinale* variabelen. Om deze reden wordt

nogal eens verondersteld dat respondenten bij de beantwoording van enquêtevragen gebruikmaken van een impliciete afstandsfunctie, waardoor het als nog mogelijk is een ordinale variabele te behandelen als een intervalvariabele. Deze opwaardering wordt in de praktijk veel toegepast bij de verwerking van enquêtegegevens.

In bijzondere gevallen kan een nominale schaal worden behandeld als een ratioschaal. Wanneer een variabele slechts twee mogelijke uitkomsten heeft, zoals de variabele *Branche* in de voorbeeldcase, dan kan deze variabele – met enige welwillendheid – worden beschouwd als gemeten op een *quasi-ratioschaal*. Immers, er hoeven slechts twee uitkomsten vergeleken te worden, waartussen slechts een afstand bestaat. Er hoeven dus geen afstanden vergeleken te worden, zodat een ‘natuurlijke’ interpretatie niet van belang is. Er is zelfs sprake van een ‘natuurlijk’ nulpunt, waarbij we uiteraard vrij zijn in de keuze tussen de bestaande codering en de alternatieve codering, 0 = laag en 1 = hoog. Bedenk dat er een onzuiverheid binnensluipt wanneer een nominale tweepuntsschaal wordt geïnterpreteerd als quasi-ratioschaal. Het verdient aanbeveling om deze benadering in statistische analyses alleen te gebruiken voor het opdoen van indrukken. Voor het leveren van bewijskracht kan men beter terugvallen op de technieken voor nominale schalen.

Een nominale variabele met *meer* dan twee mogelijke uitkomsten mag *nooit* worden beschouwd als gemeten op quasi-ratioschaal. Stel dat in een enquête onder studenten wordt gevraagd naar de faculteit waaraan men studeert, met drie mogelijke uitkomsten: Rechten, Economie en Bedrijfskunde, die gecodeerd worden als 0, 1 en 2. Indien dit als quasi-ratioschaal wordt beschouwd, dan wordt ten onrechte een volgorde en een mogelijke afstandinterpretatie gesuggereerd. Hetzelfde geldt wanneer men in de voorbeeldcase de opleiding van vertegenwoordigers (*Opleid*) zou behandelen als een quasi-ratiovariabele.

1.5 Samenhang, afhankelijkheid en causaliteit

Eén van de doelstellingen van statistisch onderzoek is het onderzoeken van samenhang tussen variabelen. Hierbij is men geïnteresseerd in vragen als: is er samenhang tussen persoonskenmerken en koopgedrag van consumenten? Welke variabelen hangen samen met insolventie van bedrijven? Is er samenhang tussen arbeidsomstandigheden en ziekteverzuim? Is het inkomen afhankelijk van het geslacht van werknemers? Het aantal verschijnselen dat onderzocht kan worden op samenhang is vrijwel eindeloos. Het belang van onderzoek naar samenhangen is evident. Wanneer het bestaan van relaties tussen variabelen aannemelijk is, kan men deze kennis gebruiken om de uitkomsten van doelvariabelen te beïnvloeden.

Stel dat uit het onderzoek in de voorbeeldcase zou blijken dat verkoopprestaties sterk samenhangen met de intelligentie en sociale vaardigheden van vertegenwoordigers. Het management van het gefuseerde bedrijf zou dit kunnen aangrijpen om bij het werven en selecteren van vertegenwoordigers in de toe-

komst gebruik te maken van een *assessment centre* voor het beoordelen van de vaardigheden van beoogde medewerkers. Stel dat uit het onderzoek in het busvervoer bij bepaalde haltes een structureel verschil tussen feitelijke en gestelde aankomsttijden geconstateerd zou worden. De directie van de busmaatschappij zou hierin een aanleiding hebben om de gestelde aankomsttijden bij deze haltes aan te passen (*the easy way out*).

Het interpreteren van geconstateerde statistische samenhangen als oorzakelijke relaties, hetgeen door beide directies impliciet wordt gedaan, is methodologisch niet helemaal zuiver. Maar in de praktijk is het een sterke motivatie voor dit type statistisch onderzoek. Hierna wordt dieper ingegaan op het gebruik van de term samenhang in de statistiek, het verschil tussen statistische samenhang en causale relaties en verschillende benaderingen van causaliteitsveronderstellingen.

1.5.1 Statistische samenhang

Zowel in het dagelijks spraakgebruik als in de statistiek spreekt men van samenhang tussen twee variabelen, wanneer de uitkomsten van de ene variabele systematisch samengaan met uitkomsten van de andere variabele. Er is systematiek of een patroon in de gezamenlijke waarnemingsuitkomsten. Het begrip samenhang, dat betrekking heeft op waargenomen of waarneembare systematiek, heeft als tegenhanger in de statistische theorie het begrip afhankelijkheid.

Statistische onafhankelijkheid en afhankelijkheid

Twee kansvariabelen X en Y heten (*statistisch*) *onafhankelijk*, als de kansverdeling van de mogelijke uitkomsten van de ene variabele niet wordt beïnvloed door de uitkomsten van de andere variabele. Stel dat men twee kansvariabelen X en Y met geheeltallige uitkomsten beschouwt. Deze X en Y heten statistisch onafhankelijk als de kans dat Y een bepaalde uitkomst y heeft, *gegeven* dat X de waarde x heeft, gelijk is aan de (marginale) kans dat Y gelijk is aan y :

$$P(Y = y | X = x) = P(Y = y) \quad (1.1)$$

voor alle mogelijke uitkomsten x en y .

Het tentamencijfer wiskunde is vermoedelijk onafhankelijk van de lengte van de student: de verdeling van mogelijke tentamencijfers wiskunde wordt niet beïnvloed door het gegeven dat een student $1,65m$, $1,90m$ of om het even welke lengte heeft. Het tentamencijfer wiskunde is naar alle waarschijnlijkheid wel afhankelijk van het eindexamencijfer wiskunde op de middelbare school: de verdeling van mogelijke tentamencijfers wiskunde voor degenen met een 9 of een 10 als eindexamencijfer verschilt van de cijfers van degenen met een 5 of 6 als eindexamencijfer. Men spreekt van statistische afhankelijkheid tussen twee variabelen, zodra deze niet onafhankelijk zijn. In hoofdstuk 5 over de kruistabel en de correlatiecoëfficiënt wordt het begrip onafhankelijkheid verder uitgewerkt.

Afhankelijke en onafhankelijke variabelen

Wanneer bij onderzoek naar samenhang de ene variabele verklarend wordt geacht voor de andere variabele, wordt ook gesproken van afhankelijkheid van de twee variabelen. Hierbij is de verklarende variabele de *onafhankelijke* variabele en de verklaarde of de te verklaren variabele de *afhankelijke* variabele. In het voorbeeld van de samenhang tussen het tentamencijfer wiskunde en het eindexamencijfer wiskunde, ligt het vanwege de tijd tussen de meetmomenten voor de hand het tentamencijfer op te vatten als de afhankelijke variabele en het eindexamencijfer als de onafhankelijke variabele.

Aard van de afhankelijkheid

De samenhang tussen twee variabelen wordt *positief* genoemd, wanneer hogere waarden van de ene variabele systematisch samengaan met hogere waarden van de andere variabele. De samenhang tussen de twee wiskundecijfers is hier een voorbeeld van. De samenhang heet *negatief*, wanneer hogere waarden van de ene variabele gepaard gaan met lagere waarden van de andere variabele.

Merk op dat het kwalificeren van de aard van de samenhang als positief of negatief alleen mogelijk is, indien de uitkomsten van de betrokken variabelen een natuurlijke ordening hebben. In het geval van nominale variabelen kan dus nooit gesproken worden van positieve of negatieve samenhang. Verder heeft de kwalificering uitsluitend betrekking op de feitelijke uitkomsten van de betrokken variabelen en niet op de aantrekkelijkheid of de wenselijkheid van het gesuggereerde verband. Stel dat men een positieve samenhang constateert tussen de omvang van het begrotingstekort en het werkloosheidspercentage, dan betekent dit geenszins dat dit verband positief is, in de zin van maatschappelijk of economisch gewenst. Stel dat men bij onderzoek naar samenhang tussen de hoogte van het inkomen en de arbeidstevredenheid, de tweede variabele heeft gemeten op een aflopende in plaats van een oplopende vijfpuntsschaal. Men zal dan een negatieve samenhang waarnemen in plaats van de verwachte positieve samenhang. Dit laatste is weliswaar slechts een kwestie van codering, maar geeft in de praktijk nogal eens aanleiding tot verwarring.

1.5.2 Statistische samenhang versus causaliteit

Kenmerkend voor empirisch onderzoek is dat de oorzakelijkheid van relaties tussen grootheden principieel niet kan worden waargenomen. Het enige dat kan worden waargenomen is het herhaald voorkomen van specifieke combinaties van uitkomsten van variabelen. De interpretatie van deze systematiek als het resultaat van oorzakelijkheid is een subjectieve betekenisgeving door de onderzoeker. Empirisch onderzoek verschilt in dit opzicht wezenlijk van een technisch onderzoek naar de werking van een apparaat. Dit laatste is in beginsel mogelijk door het oplichten van het deksel en een deskundige blik op de verbindingen tussen de samenstellende delen van het apparaat. Bij onderzoek naar empirische verschijnselen is een vergelijkbare operatie niet mogelijk.

Empirisch onderzoek wordt ten diepste gemotiveerd door de vooronderstelling dat er causale relaties bestaan en dat deze gekend kunnen worden op basis van repeterende patronen in de uitkomsten.

De praktische consequentie van deze vooronderstellingen is dat statistische samenhang weliswaar objectief is vast te stellen, maar dat de interpretatie van deze samenhang in termen van causaliteit inherent subjectief van aard is. De verwarring die ontstaat wanneer men de begrippen statistische samenhang en causaliteit onvoldoende onderscheidt of onvoldoende in de gaten houdt dat causaliteit voortkomt uit het eigen, subjectieve interpretatiekader, wordt toegelicht met enkele voorbeelden.

- Er blijkt een positief verband te zijn tussen het aantal brandweermannen dat uitrukt om branden te blussen en de ontstane brandschade. De statistische samenhang tussen deze twee variabelen is objectief meetbaar. De onderliggende causaliteit komt echter voort uit het eigen interpretatiekader. Men kan de samenhang interpreteren als 'brandweerlieden veroorzaken schade' of als 'bij grote branden rukken veel brandweerlieden uit'. Men *geloof*t de tweede interpretatie en doet de eerste interpretatie af als kennelijke dwaasheid.
- Er blijkt een negatief verband te zijn tussen de arbeidsproductiviteit van medewerkers en de mate van autoritair optreden van de chef. Dit statistische verband is objectief meetbaar. De onderliggende causaliteit kan geïnterpreteerd worden als 'autoritair gedrag veroorzaakt lage arbeidsproductiviteit', maar ook als 'bij lage arbeidsproductiviteit vluchten chefs in autoritair gedrag'. Sommigen *gelo*ven in de eerste interpretatie en doen smalend over de tweede uitleg. Anderen *gelo*ven in de tweede interpretatie en doen de eerste af als onzin. Geen statistiekboek kan helpen bij het kiezen tussen de twee interpretaties.
- In Noordwest-Europa is gedurende de afgelopen 50 jaar het aantal ooievaarsnesten geteld en het geboortecijfer gemeten. Er blijkt een objectief meetbaar positief verband tussen de twee variabelen te bestaan: minder nesten gaat systematisch samen met minder geboorten. Slechts een enkeling zal nog *gelo*ven in een oorzakelijke relatie tussen de omvang van de ooievaarspopulatie en het geboortecijfer. Kennelijk is er sprake van een *schijn*verband: in vroegere jaren waren er meer ooievaarsnesten in Noordwest-Europa en was ook het geboortecijfer hoger.
- In serieus wetenschappelijk onderzoek is aangetoond dat na het drinken van een paar borrels minder ongelukken optreden bij het skiën dan in nuchtere toestand. Deze objectief gemeten samenhang tussen alcoholgebruik en ski-ongelukken kan echter op allerlei manieren worden geïnterpreteerd: 'na een paar borrels durft men meer en skiet dus beter' of 'men valt minder krampachtig en krijgt daardoor minder gauw letsel'. Het is ook mogelijk dat sprake is van een *schijn*verband: 'jonge mensen drinken meer en skiën beter dan ouderen'; of 'zwarte drinkers besteden minder tijd aan skiën en krijgen dus minder ongelukken'.

Het methodologische verschil tussen objectief waarneembare samenhang en subjectief veronderstelde causaliteit wordt aangeduid met de symbolische formule:

$$\text{Samenhang} \neq \text{Causaliteit}$$

1.5.3 A priori en a posteriori causaliteit

In wetenschappelijk onderzoek wordt op twee verschillende manieren omgegaan met het probleem van de causale interpretatie van samenhang. Men kan *a priori* causaliteit of *a posteriori* causaliteit veronderstellen. Bij *a priori* veronderstelde causaliteit maakt men de eigen ideeën over een causale relatie expliciet en gaat men vervolgens na of deze ideeën worden ondersteund door een gemeten samenhang. Bijvoorbeeld, vanuit theoretische overwegingen *stelt* men dat autoritair gedrag van managers leidt tot lagere arbeidsproductiviteit. Een geconstateerde samenhang wordt vervolgens geïnterpreteerd als steun voor de veronderstelling. In het geval van *a posteriori* veronderstelde causaliteit meet men samenhang en bedenkt men achteraf hoe deze samenhang een causale interpretatie zou kunnen hebben. Bijvoorbeeld, men constateert een negatieve statistische samenhang tussen autoritair gedrag van chefs en arbeidsproductiviteit en men voorziet de geconstateerde samenhang achteraf van de interpretatie dat autoritair gedrag leidt tot lage productiviteit. Deze interpretatie is een direct gevolg van de overtuiging van de onderzoeker dat lage arbeidsproductiviteit een uiting is van arbeidsdissatisfactie. Het verschil tussen *a priori* en *a posteriori* veronderstelde causaliteit vindt men overigens terug in statistische methoden: regressiemodellen zijn gebaseerd op *a priori* veronderstelde causaliteit, terwijl (exploratieve) factoranalyse een typisch voorbeeld is van *a posteriori* veronderstelde causaliteit.

De vraag is nu natuurlijk welke van de twee benaderingen de voorkeur verdient. Sommigen zijn bevreesd dat de methode van *a priori* causaliteit de onderliggende ideologie steeds meer bevestigt. Bijvoorbeeld, men *geloofd* in rasverschillen, bij meting vindt men ook wat men denkt te zullen vinden en dus *zijn* er rasverschillen. Anderen zijn bevreesd dat de methode van *a posteriori* causaliteit de impliciete vooroordelen onzichtbaar maakt, terwijl deze wel degelijk sturend zijn geweest voor de richting van het onderzoek. Men meet de samenhang tussen arbeidsproductiviteit en hiërarchisering in ondernemingen zogenaamd waarde vrij, waarbij men doet alsof de causaliteit achteraf geconstateerd wordt, maar waarom heeft men arbeidsproductiviteit en hiërarchisering überhaupt tegelijkertijd gemeten?

In dit boek wordt als uitgangspunt gekozen: wees expliciet in de eigen vooroordelen, want niemand begint blanco aan een onderzoek. Ga bij belanghebbenden na in hoeverre andere opvattingen bestaan. Gebruik zowel eigen vooroordelen als de vooroordelen van belanghebbenden voor het *a priori* vaststellen van causaliteit. Voer het onderzoek uit en meet de veronderstelde verbanden. Confronteer de belanghebbenden met de resultaten en vraag naar hun oordeel over de implicaties van de resultaten voor de onderliggende causaliteiten. Geef ten slotte een eigen oordeel over de onderliggende causaliteit. De rol van vooroordelen in onderzoek wordt dus geaccepteerd. Maar door

het aanbrengen van variëteit in opvattingen wordt geprobeerd het gevaar van eenzijdigheid te vermijden.

1.6 Het causale relatieschema: netwerken met variabelen

Het aantal mogelijk te onderzoeken samenhangen tussen variabelen loopt bij statistisch onderzoek al snel op. In het verhoudingsgewijs beperkte onderzoek van de voorbeeldcase, waarbij slechts zes kenmerken van de vertegenwoordigers zijn verzameld, kunnen al 15 ($= 6 \times 5 / 2$) paarsgewijze samenhangen worden onderzocht. Wanneer ook de relaties tussen meer dan twee variabelen worden onderzocht, dan blijkt zelfs dit eenvoudige onderzoek naar verkoopprestaties al redelijk complex te zijn. Het causale relatieschema is een hulpmiddel voor het structureren van deze onderzoeksinspanningen. In deze paragraaf wordt de achtergrond van het relatieschema toegelicht, worden de conventies voor het opstellen van relatieschema's beschreven en wordt het relatieschema toegelicht aan de hand van de voorbeeldcase.

1.6.1 Achtergrond van het relatieschema

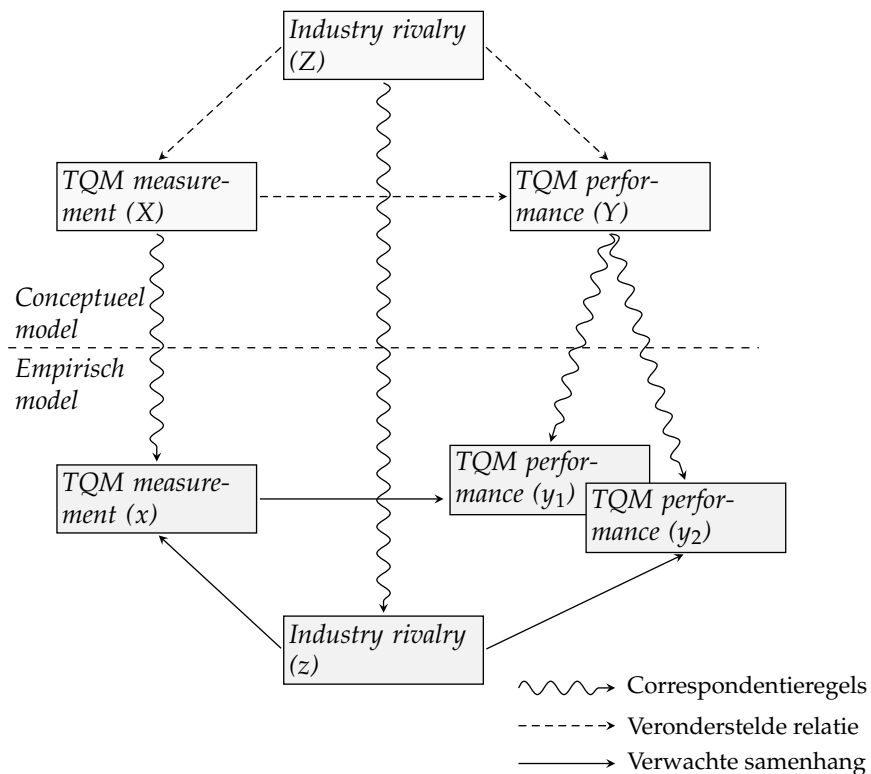
Een relatieschema is een grafische weergave van alle veronderstelde oorzakelijke relaties tussen variabelen. De variabelen die in het schema worden opgenomen, kunnen verwijzen naar theoretische concepten die nog helemaal niet gemeten zijn, of naar de waargenomen variabelen die feitelijk beschikbaar zijn in een databestand. Relatieschema's op basis van theoretische concepten zijn vooral zinvol bij het begrijpen van een theorie of bij het vertalen van een probleemstelling naar een onderzoeksopzet. Relatieschema's op basis van de waargenomen variabelen zijn vooral zinvol als hulpmiddel bij de statistische verwerking van informatie in een gegeven databestand. In dit boek worden relatieschema's gemaakt op basis van de waargenomen variabelen. Dit is hoofdzakelijk het gevolg van het feit dat het traject tussen probleemformulering en dataverzameling als gegeven wordt beschouwd.

Conceptuele en empirische relatieschema's

De genoemde relatieschema's op basis van theoretische concepten en waargenomen variabelen staan niet los van elkaar, maar liggen in elkaars verlengde of zijn elkaars tegenhangers. Een voorbeeld van de verwantschap van beide relatieschema's wordt getoond in figuur 1.2. Het voorbeeld is gebaseerd op een studie van Powell naar de invloed van kwaliteitsmanagement (*total quality management, TQM*) op het resultaat van ondernemingen.¹ In deze studie

¹ Powell, T.C., 1995, Total Quality Management as Competitive Advantage: A Review and Empirical Study, *Strategic management Journal* 16(1), 15-37.

wordt eerst de nodige theorie behandeld om op theoretische gronden aannemelijk te maken dat de kwaliteit van meetsystemen (*TQM measurement, X*) invloed heeft op de resultaten die met kwaliteitsmanagement worden geboekt (*TQM performance, Y*). Beide theoretische grootheden worden op hun beurt verondersteld afhankelijk te zijn van het mededingingsklimaat in de bedrijfstak (*industry rivalry, Z*). De veronderstelde relaties tussen deze theoretische grootheden worden tot uitdrukking gebracht in het conceptuele relatieschema in de bovenste helft van de figuur.



Figuur 1.2 Correspondentie tussen conceptuele en empirische modellen

Het onderzoek vervolgt met het ontwikkelen van indicatoren om de theoretische grootheden te kunnen meten. Dit gebeurt in Powells onderzoek met een vragenlijst uitgezet onder een aantal bedrijven in Noord-Amerika. De ontwikkelde en vervolgens waargenomen indicatoren kunnen eveneens met elkaar in verband gebracht worden geheel parallel aan de veronderstelde relaties tussen de theoretische grootheden. Dit leidt tot het empirische relatieschema in de onderste helft van de figuur.

De beide relatieschema's lijken natuurlijk sterk op elkaar, maar toch zijn er verschillen. Het conceptuele schema is opgebouwd uit abstracte, theoretische grootheden, terwijl het empirische relatieschema bestaat uit feitelijk waargenomen indicatoren. De theoretische grootheden en empirische indicatoren zijn 'verbonden' door zogenaamde correspondentieregels, waarmee de overeenstemming in interpretatie van beide typen grootheden wordt geregeld. Een aardig detail is dat een theoretische grootheid zeer helder geformuleerd kan zijn, maar dat men toch de nodige vrijheid heeft in de keuze of het ontwerp van empirische indicatoren. Dit wordt geïllustreerd met de afhankelijke variabele *TQM performance* die op twee verschillende manieren wordt gemeten, op basis van percepties van respondenten en feitelijke winstpercentages. Onderzoek moet verder uitmaken welk van beide indicatoren het meest geschikt is.

Verschillen tussen waargenomen samenhang en veronderstelde relaties

Een causaal relatieschema wordt opgesteld *voordat* de waarnemingen geanalyseerd worden. Dit sluit aan bij de eerdergenoemde voorkeur om causaliteit *a priori* expliciet te maken. Tijdens de statistische analyse worden de objectief waargenomen samenhangen telkens geconfronteerd met de subjectief veronderstelde oorzakelijke relaties in het schema. Wanneer de empirische resultaten ondersteuning bieden aan de veronderstelde causaliteit, dan wordt men gesterkt in de intuïtieve ideeën. Wanneer empirisch resultaat en veronderstelde causaliteit verschillen, dan is er huiswerk te doen. Het kan zijn dat de veronderstelling juist is en dat de meetresultaten niet kloppen door fouten in definities, metingen, toegepaste methoden en modellen. Een voorbeeld van definitiefouten doet zich voor als men in plaats van de effecten van prijs op omzet had moeten kijken naar de effecten van de relatieve prijs ten opzichte van concurrenten. Meetfouten kunnen zich voordoen als gevolg van een slechte vraagstelling in de enquête. Methodische fouten doen zich voor als men bijvoorbeeld een nominale variabele ten onrechte behandelt als gemeten op een intervallschaal. Een voorbeeld van modelfouten is als men de effecten van een doorkruisende variabele over het hoofd ziet. Het kan natuurlijk ook zijn dat de veronderstellingen moeten worden aangepast.

Het vergt enige zelfdiscipline om vooraf de intuïtieve ideeën vast te leggen in een relatieschema en pas daarna naar de waarnemingen te kijken. Wanneer men echter regelrecht het waarnemingsmateriaal analyseert, dreigt het gevaar dat men naar het empirisch materiaal toe praat. Het is juist de kunst om subjectieve veronderstellingen en harde feiten met elkaar te confronteren en de oorzaken van discrepanties op te sporen.

Het causale relatieschema is een hulpmiddel bij het uitvoeren van de analyse. In het algemeen zal men in de eindrapportage alleen de slotconclusies van de analyse vermelden en niet alle tussenstappen die tot deze conclusies geleid hebben. Relatieschema's worden dan ook zelden gerapporteerd. Daarvoor in de plaats wordt soms een lijst van hypothesen weergegeven, met commentaar per hypothese of het empirisch materiaal tot verwerping of tot acceptatie van de hypothese geleid heeft. Het voordeel hiervan is dat uitgeschreven hypothe-

sen meer aanspreken dan abstracte diagrammen. Een nadeel is dat mogelijke doorkruisendheden minder duidelijk zichtbaar worden.

1.6.2 Het opstellen van een relatieschema

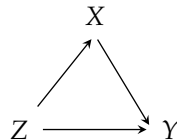
Voor het weergeven van oorzakelijke relaties in een relatieschema geldt een aantal regels. Wanneer men voor een veronderstelde relatie tussen twee variabelen X en Y wil aangeven dat X wordt gezien als de onafhankelijke, verklaarende variabele en Y als de afhankelijke, te verklaren variabele, dan geeft men dit aan met een enkele pijl:

$$X \rightarrow Y$$

Om in een gegeven situatie te achterhalen welke van de variabelen X en Y de afhankelijke en welke de onafhankelijke variabele is, kan het helpen om na te gaan of veranderingen in X ook veranderingen in Y veroorzaken en andersom. Neem als voorbeeld inkomen en leeftijd: een hogere leeftijd kan wel tot een hoger inkomen leiden, maar een hoger inkomen kan niet leiden tot een hogere leeftijd. Kennelijk is het inkomen de afhankelijke variabele en leeftijd de onafhankelijke variabele. Bij gedragsonderzoek bestaan de onafhankelijke variabelen veelal uit (i) karakterkenmerken van de respondent; (ii) persoonskenmerken, zoals leeftijd en geslacht; en (iii) kenmerken van de sociale klasse waartoe de respondent behoort, zoals inkomen, woonomgeving, enzovoort.

Indirecte effecten en gezamenlijke oorzaken

Veronderstelde relaties tussen verschillende variabelen X , Y en Z kunnen worden gecombineerd tot een groter schema:



In een dergelijk schema worden de drie paarsgewijze relaties $X \rightarrow Y$, $Z \rightarrow X$ en $Z \rightarrow Y$ directe relaties of directe effecten genoemd. De uitgebreidere relatie $Z \rightarrow X \rightarrow Y$ wordt een indirecte relatie of indirect effect van Z op Y via X genoemd. De X wordt in deze indirecte relatie de interveniërende grootheid genoemd. Daarnaast bevat dit schema de relatie $X \leftarrow Z \rightarrow Y$. Hierin wordt Z aangeduid als de gezamenlijke oorzaak (*common cause*), omdat ze gelijktijdig variatie in X en Y veroorzaakt. Het herkennen van indirecte effecten en gezamenlijke oorzaken is van belang voor de interpretatie van de onderzoeksresultaten, zoals verder wordt uitgewerkt in hoofdstuk 7.

Simultane relaties

Soms beïnvloeden twee variabelen elkaar over en weer. Er is dan een wederkerige of simultane relatie tussen X en Y . Een bekend voorbeeld van een simultane relatie is die tussen marktprijs en verhandelde hoeveelheid. Een hoge marktprijs is voor aanbieders een aanleiding om de aangeboden hoeveelheid te vergroten, terwijl een groot aanbod zal leiden tot een neerwaartse druk op de marktprijs. Een dergelijke simultane relatie wordt weergegeven met een dubbele pijl:

$$X \rightleftharpoons Y$$

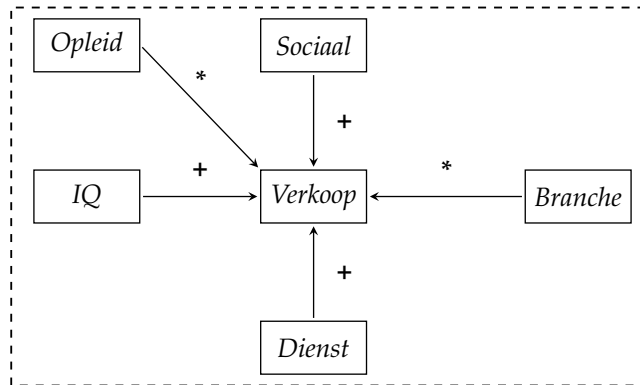
Overigens laten werkelijk simultane relaties, in de zin dat X en Y elkaar op exact hetzelfde moment beïnvloeden, zich slecht voorstellen. Het is daarom nuttig om bij simultane relaties altijd een tijdspanne op de achtergrond te veronderstellen: een verandering in X leidt even later tot een verandering in Y , welke op haar beurt een fractie later leidt tot een verandering in X : $X_t \rightarrow Y_{t+1} \rightarrow X_{t+2}$, waarbij t een tijdsindex is. Aangezien de reactiemomenten t , $t+1$ en $t+2$ binnen het ruimere tijdsbestek vallen waarin de steekproefgegevens zijn verzameld, is dit onderscheid in perioden niet functioneel. In plaats daarvan wordt de dubbele pijl tussen variabelen zonder tijdsindices gebruikt.

Verwachte aard van relaties

In het relatieschema geeft men tevens de aard van de verwachte samenhang aan. Indien men een positieve samenhang verwacht, zet men een '+' bij de pijl; en indien men een negatieve samenhang verwacht, een '-'. In situaties waar men niet zeker weet of het verband positief of negatief zal zijn, plaatst men een vraagteken '?' bij de pijl. Het uitbeelden van de verwachte aard van samenhang als positief of negatief is echter alleen van toepassing wanneer de betrokken variabelen ten minste op ordinaal niveau gemeten zijn. In situaties waarin een van de variabelen nominaal is, plaatst men een asterisk '*'. De aard van de verwachte samenhang wordt altijd toegelicht in de tekst.

1.6.3 Het relatieschema voor de voorbeeldcase

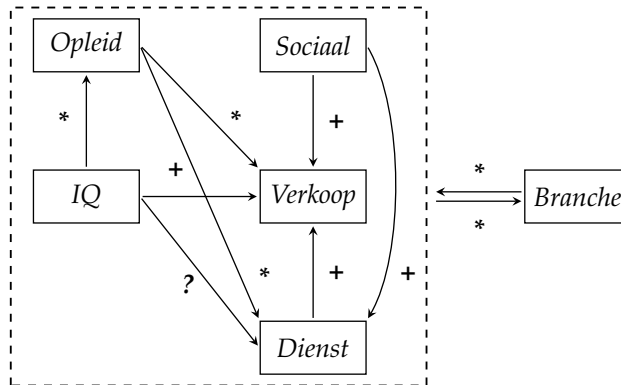
Wanneer hierna het schema voor de voorbeeldcase wordt uitgewerkt, dan betekent dit niet dat dit schema beter zou zijn dan andere schema's. Over de te verklaren variabele hoeft echter weinig verschil van inzicht te bestaan. Dit is de verkoopprestatie (*Verkoop*). Verder lijkt het redelijk te veronderstellen dat de verkoopprestatie positief wordt beïnvloed door de sociale vaardigheden (*Sociaal*), de diensttijd (*Dienst*) en de intelligentie (*IQ*) van vertegenwoordigers. Daarnaast zullen de opleiding (*Opleid*) en de afdeling waar men werkt (*Branche*) van invloed zijn op de verkoopprestatie. Voor deze variabelen kan de aard van de verwachte samenhang echter niet met '+' of '-' worden aangegeven, aangezien het nominale variabelen betreft. In het relatieschema wordt dit weergegeven met een asterisk '*'. Uitgaande van deze gedachten is het voorlopige relatieschema in figuur 1.3 opgesteld.



Figuur 1.3 Voorlopig relatieschema met directe effecten op de verkoopprestaties

Naast de directe effecten van de verklarende variabelen op de verkoopprestaties, bestaan er allerlei relaties tussen de verklarende variabelen onderling. Deze onderlinge relaties veroorzaken indirecte effecten van de verklarende variabelen op de te verklaren variabele. Deze onderlinge relaties hebben invloed op de meting van de directe effecten op de verkoopprestaties. In hoofdstuk 7 over doorkruisende variabelen wordt hier dieper op ingegaan. In de voorbeeldcase wordt verondersteld dat intelligentie verklarend is voor de opleiding die vertegenwoordigers hebben gevolgd. Let op dat de causaliteitsrichting loopt van *IQ* naar *Opleid*, en niet andersom: *IQ* meet namelijk de intelligentie en niet de kennis van vertegenwoordigers. Verder wordt verondersteld dat de persoonsgebonden kenmerken, met name intelligentie en sociaal gedrag, van invloed zijn op de duur van de dienstbetrekking. Afhankelijk van de arbeidsomstandigheden zouden gekwalificeerde vertegenwoordigers lang of kort in dienst kunnen zijn. In het eerste geval is er een goed werkend selectieproces, terwijl in het tweede geval de betere vertegenwoordigers zich makkelijk weg laten kopen door andere bedrijven. Het is dus onzeker of het verband positief of negatief zal zijn: in het schema wordt een vraagteken gezet. Een relatie tussen het sociaal gedrag en de opleiding van vertegenwoordigers ligt niet voor de hand en wordt om die reden niet in het schema opgenomen. Een verband tussen sociaal gedrag en intelligentie lijkt evenmin vanzelfsprekend. Als dit wel zo zou zijn, moet een wederkerige relatie verondersteld worden. Samenhang tussen opleiding en dienstduur lijkt plausibel: vooral interne cursussen kunnen samenhangen met langere dienstduur. De relatie wordt gemarkeerd met een '*', vanwege het nominale karakter van *Opleid*.

De vraag die overblijft is in hoeverre de afdeling waar men werkt, samenhangt met de andere variabelen. In principe kunnen alle variabelen verschillen voor de twee divisies van het gefuseerde bedrijf. Om dit aan te geven wordt een groot vierkant geplaatst om het gedeelte van het relatieschema dat betrek-



Figuur 1.4 Uiteindelijk relatieschema voor de voorbeeldcase

king heeft op de andere variabelen. Met een dubbele pijl tussen *Branche* en het blok overige variabelen wordt aangegeven dat de relaties een simultaan karakter hebben. Bijvoorbeeld, meer intelligente mensen zullen terecht komen in de technisch geavanceerde elektronica-afdeling, terwijl omgekeerd de leiding van de elektronica-afdeling bij voorkeur personen met een hoger IQ in dienst neemt. Het uiteindelijke resultaat van alle opgestelde hypothesen is weergegeven in figuur 1.4. Dit schema dient als leidraad voor de analyse in de navolgende hoofdstukken.

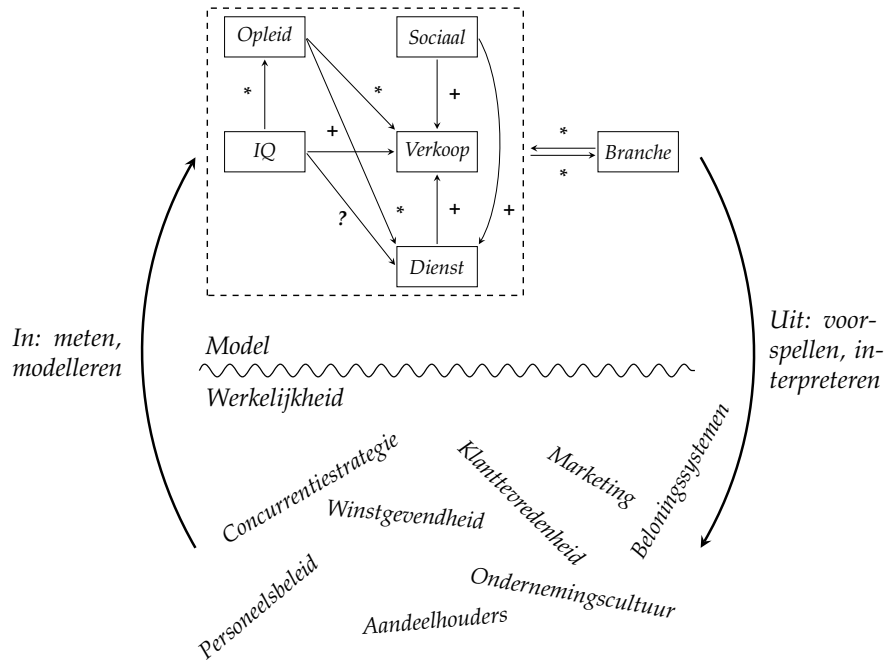
Voor grote aantallen variabelen wordt het relatieschema onoverzichtelijk. In dit geval biedt het combineren van variabelen tot groepen soms uitkomst. Bijvoorbeeld, een groep 'persoonskenmerken', een groep 'koopgedrag' en een groep 'productinformatie'. Met deze groepen wordt vervolgens een versimpeld relatieschema gemaakt.

1.6.4 Model en werkelijkheid: *In & Uit*

Het is van belang kort stil te staan bij enkele implicaties van het bouwen van modellen of causale relatieschema's zoals hiervoor geïllustreerd.

Het startpunt van het onderzoek was de vraag van het management van het gefuseerde bedrijf in de voorbeeldcase om inzicht te geven in de prestaties van medewerkers. Deze vraag komt niet uit het luchtledige, maar hangt samen met allerlei processen, gebeurtenissen en omstandigheden waar het bedrijf dagelijks mee te maken heeft. In figuur 1.5 worden deze onderdelen van de werkelijkheid als thema's genoemd: personeelsbeleid, ondernemingscultuur, klanttevredenheid, winstgevendheid, et cetera.

Wanneer we in een dergelijke complexe werkelijkheid een model opzetten, dan is het onvermijdelijk dat impliciet of expliciet allerlei keuzen worden gemaakt. De keuzen hebben betrekking op welke grootheden wel of niet van belang zijn voor de onderzoeksvraag, wat deze precies betekenen en hoe ze



Figuur 1.5 Bouwen van modellen voor processen in de werkelijkheid

aan elkaar gerelateerd zijn. Uitkomst van deze inspanningen is een model, zoals bijvoorbeeld het causale relatieschema in de bovenste helft van de figuur. Kenmerkend voor modellen is dat het abstracte, formele en subjectieve verzamelingen van grootheden en relaties tussen grootheden zijn. Modellen zijn *abstract* omdat ze lang niet alle aspecten van de werkelijkheid weergeven. Alleen de meest belangrijk geachte aspecten krijgen een plek in het model. Zo wordt in het model in figuur 1.5 wel aandacht besteed aan de verkoopprestatie, maar niet aan de inspanningen die nodig zijn geweest om tot de verkopen te komen. Kennelijk heeft men deze benodigde inspanningen over het hoofd gezien of niet relevant geacht, maar de zaak is dat ze niet in het model zitten. Modellen heten *formeel* omdat binnen het model geen enkele vaagheid bestaat over de betekenis van de gemodelleerde grootheden en hoe deze onderling gerelateerd zijn. Het causale relatieschema in de figuur legt de logische structuur van deze relaties ondubbelzinnig vast. Modellen zijn *subjectief* omdat ze het resultaat zijn van tal van grotere en kleinere beslissingen van de onderzoeker. Uiteraard gaat men bij deze keuzen zorgvuldig te werk, maar dat poetst de rol van deze beslissingen niet weg.

Na het identificeren van de relevante grootheden zijn er de vragen hoe deze gemeten moeten worden, welke indicatoren gebruikt worden en hoe deze gemeten moeten worden. Ook hierbij worden tal van keuzen gemaakt. Neem

bijvoorbeeld de verkoopprestatie die gemeten wordt met een variabele *Verkoop*. Volgens tabel 1.1 is er kennelijk voor gekozen om deze grootheid te meten met een soort van index die uitkomsten rond de 100 heeft. Maar ergens in het proces dat tot deze variabele heeft geleid zijn beslissingen genomen over welke informatie in de berekening van de verkoopprestatie moest worden betrokken en hoe deze informatie bewerkt moest worden. Ook met betrekking tot de ontwikkeling van de empirische indicatoren voor de conceptuele grootheden in het model worden dus tal van keuzen gemaakt door de onderzoeker of andere betrokkenen bij het onderzoek.

Wanneer het model eenmaal is geschat op basis van waargenomen gegevens, wordt onderzoek gedaan naar welke grootheden op welke manier en in welke mate samenhangen met andere grootheden. Deze inzichten kunnen vervolgens worden gebruikt om uitspraken te doen over de situatie in de werkelijkheid. De resultaten van de analyses kunnen bijvoorbeeld worden gebruikt om wervingsbeleid voor nieuwe medewerkers aan te scherpen, de opzet van interne cursussen opnieuw te bezien, of om extreem slecht of juist excellent presterende vertegenwoordigers te identificeren. Met het toepassen van de modelresultaten op de werkelijkheid gaat men als het ware het model weer uit. Het is hierbij van belang te realiseren dat het model niet de werkelijkheid is en dat er in het voortraject tal van keuzen zijn gemaakt die invloed hebben op de interpretatie en de reikwijdte van de verkregen resultaten. Om een voor de hand liggend voorbeeld te noemen: het model is geschat op informatie over vertegenwoordigers, zodat toepassen van de resultaten op alle medewerkers van het bedrijf tot grote vergissingen kan leiden. Of, bij de opzet van het model is concurrentie buiten beschouwing gelaten, zodat het gebruik van de resultaten voor personeelsbeoordeling geen rekening houdt met het feit dat de invloed van concurrentie voor vertegenwoordigers in de ene regio veel sterker is dan voor vertegenwoordigers in een andere regio.

Bij empirisch onderzoek zullen we op deze manier vaak het model in- en uitgaan. Dit is een even onvermijdelijk als boeiend aspect van de activiteit van onderzoek doen. Het een eigenschap van onderzoek die noopt tot een zekere bescheidenheid in de stelligheid waarmee men uitspraken doet over de werkelijkheid, en die aanzet tot een gezonde kritische houding bij de interpretatie van onderzoeksgegevens.

1.7 Structuur van het boek

De hoofdstukken in dit boek zijn als volgt gestructureerd. In hoofdstuk 2 volgt een inleiding *SPSS for Windows*. Hierin wordt aandacht besteed aan het inlezen van gegevens, het maken van aanpassingen in de ingelezen gegevens, het uitvoeren van allerlei analyses en het wegschrijven van resultaten. Allemaal essentiële vaardigheden voor het verwerken van statistische gegevens met behulp van een computer. Hoofdstuk 3 behandelt in vogelvlucht de meest relevante statistische concepten, zoals het verschil tussen eindige frequenties en

kansen, de overstap van kansen naar eindige frequenties en de overstap van eindige frequenties naar kansen. Voor wat betreft het laatste staat de theorie van schatten en hypothesetoetsen centraal. Aangezien de doelstelling van hoofdstuk 3 niet verder gaat dan het aanbrenge van helderheid in statistische basisconcepten is gekozen voor simpele, weinig realistische voorbeelden zoals het gooien met een munt of een dobbelsteen. De werkelijke betekenis van statistiek ligt natuurlijk in een geheel ander toepassingsdomein, zoals de verwerking van enquêtes, het maken van kwantitatieve voorspellingen op basis van waargenomen tijdreeksen, en het meten van parameters wanneer sprake is van kansverschijnselen. In de latere hoofdstukken zullen dergelijke meer realistische gevallen aan de orde worden gesteld.

Nadat in de hoofdstukken 2 en 3 een basis is gelegd met een uitleg van eenvoudige statistische concepten en een inleiding *SPSS for Windows*, worden in de volgende hoofdstukken analysemethoden behandeld aan de hand van de voorbeeldcase. In hoofdstuk 4 wordt ingegaan op univariate methoden en technieken. Dit zijn statistische methoden en technieken voor het analyseren van de uitkomsten van afzonderlijke variabelen. Wanneer bijvoorbeeld een steekproefonderzoek is gedaan waarin de maandelijks aangekochte hoeveelheid wasmiddelen is gemeten, dan worden deze univariate technieken ingezet om vragen te beantwoorden als:

- Welke kengetallen zijn geschikt om de steekproefuitkomsten te beschrijven? (gemiddelde, standaarddeviatie, enzovoort)
- Hoe kan de verdeling van de steekproefresultaten visueel zichtbaar gemaakt worden? (frequentiediagram, enzovoort)
- Gegeven het steekproefgemiddelde, wat kan men dan zeggen over de gemiddelde aangekochte hoeveelheid voor de gehele populatie? (schattingstheorie)
- Gegeven de gemiddelde aankoop in het vorige jaar en gegeven een afwijkend gemiddeld resultaat in de huidige steekproef, in hoeverre kan men dan stellen dat een meer dan toevallige verschuiving heeft plaatsgevonden? (toetsingstheorie)
- Wanneer men verwacht dat aangekochte hoeveelheden min of meer normaal verdeeld zijn en wanneer de verdeling van de steekproefresultaten daarvan nogal afwijkt, in hoeverre kan dit dan toegeschreven worden aan toeval? (*goodness-of-fit*-toets)
- In hoeverre zijn er in de steekproef extreme waarnemingsuitkomsten aanwezig; en waardoor worden deze veroorzaakt? Moeten zulke waarnemingsuitkomsten in de steekproef worden gehandhaafd?
- Kan men aan de hand van de gemeten variabelen een indruk krijgen van de representativiteit van de steekproef? Indien er geen sprake is van representativiteit, hoe kunnen we dan nog steeds uitspraken doen over het koopgedrag in de populatie?

Natuurlijk zal men zich bij een enquête naar koopgedrag ten aanzien van wasmiddelen niet beperken tot alleen de aangekochte hoeveelheden. Men zal allerlei andere variabelen in de enquête meenemen die verklarend kunnen zijn voor verschillen in koopgedrag, zoals gezinsomvang, leeftijd, sociale klasse, inkomen, enzovoort. Een van de meest belangrijke functies van statistiek is het verkrijgen van inzicht in mogelijke verklarende factoren voor een bepaald verschijnsel. In de hoofdstukken 5 en 6 wordt ingegaan op het meten van samenhang tussen telkens twee variabelen: de bivariate methoden en technieken. Door onderzoek naar verklarende factoren voor het koopgedrag ten aanzien van wasmiddelen hoopt men antwoord te krijgen op vragen als:

- Hangt de gekochte hoeveelheid wasmiddelen samen met de gezinsomvang? Vast wel, en nadat statistische samenhang is vastgesteld, zal men niet schromen om te spreken van een causaal verband.
- Hangt de gekochte hoeveelheid wasmiddelen samen met leeftijd? Waarschijnlijk niet. Wanneer men dus tegen de verwachtingen in toch een statistisch verband vindt, zal men moeten uitzoeken wat er aan de hand is. Misschien waren de verwachtingen niet terecht. Misschien was er iets fout met de onderzoeksopzet of de meetmethode. Misschien ook is sprake van een indirect verband met leeftijd, bijvoorbeeld vanwege een verband tussen leeftijd en gezinsomvang.

De geëigende statistische techniek voor het onderzoeken van paarsgewijze samenhangen is afhankelijk van de meetniveaus van de betrokken grootheden. Wanneer twee nominale variabelen onderzocht worden op samenhang, gebruikt men daarvoor kruistabellen. Wanneer twee variabelen op interval-schaal worden onderzocht, gebruikt men de correlatiecoëfficiënt. Wanneer er sprake is van twee ordinale variabelen, kan men de rangcorrelatie gebruiken. Wanneer sprake is van een nominale variabele en een variabele op interval-schaal, dan past men de *t*-toets of variantieanalyse toe. De technieken verschillen, maar de onderliggende gedachtegang is steeds dezelfde.

In de hoofdstukken 7, 8 en 9 wordt de aandacht verlegd naar het gelijktijdig analyseren van meer dan twee variabelen. Hoofdstuk 7 gaat in op doorkruisende variabelen en op de vraag hoe men rekentechnisch met doorkruisendheid omgaat. Hierbij kunnen de volgende vragen aan de orde komen:

- Stel dat er samenhang gemeten is tussen sociale klasse en aangekochte hoeveelheid wasmiddel. Is deze samenhang dan werkelijk betekenisvol, of zou het soms zo kunnen zijn dat in lagere sociale klassen de gezinsomvang groter is en dat daardoor de aangekochte hoeveelheid wasmiddelen groter is? Dit zou betekenen dat sociale klasse als zodanig geen effect heeft op de aangekochte hoeveelheid.
- Om na te gaan of het verband tussen sociale klasse en gezinsomvang werkelijk betekenis heeft, moet het effect van gezinsomvang worden verwijderd. Dit zogenaamde 'opschonen' is opnieuw een van de basisconcepten van toegepaste statistiek. Denk bijvoorbeeld aan het 'voor seizoen gecorrigeerde werkloosheidscijfer', dat wil zeggen het werkloosheidscijfer dat is opgeschoond voor seizoensinvloeden.

In hoofdstuk 8 wordt de techniek van variantieanalyse uitgebreid tot situaties waarin meer dan twee variabelen simultaan in beschouwing worden genomen.

In hoofdstuk 9 wordt aandacht geschonken aan regressieanalyse. Deze techniek kenmerkt zich door het feit dat er sprake is van een te verklaren variabele en verschillende verklarende variabelen. Met behulp van regressieanalyse kunnen verklarende modellen gebouwd worden. Vragen die hierbij aan de orde komen, kunnen zijn:

- Wat zijn verklarende factoren voor de gekochte hoeveelheid wasmiddelen?
- Hoe groot is de invloed van afzonderlijke factoren als inkomen, sociale klasse, gezinsgrootte, enzovoort?
- In welke mate kan de aangekochte hoeveelheid worden voorspeld, als men de waarden van de verklarende factoren kent?
- Waaraan herkent men dat een verklarend model nog onvoldoende kwaliteit bezit? Hoe kan men zo'n model verbeteren?

Met deze onderwerpen is zo ongeveer het basispakket van statistiek aan de orde gekomen. De analyse van situaties met (veel) meer variabelen die simultaan onderzocht worden, wordt wel aangeduid als multivariate analyse. Onderdelen hiervan zijn, naast de al genoemde regressieanalyse: factoranalyse, clusteranalyse, multivariate variantieanalyse, discriminantanalyse en nog veel meer. Deze technieken maken echter geen onderdeel uit van het voorliggende boek.

1.8 Literatuur

Hoewel dit boek veel aandacht besteedt aan de interpretatie van praktische computeruitvoer en daarnaast verschillende rekenvoorbeelden uitwerkt, bevat het geen oefenopgaven. Dit heeft te maken met de opzet van het boek, waarin voorrang wordt gegeven aan de betekenisgeving van onderzoeksresultaten. Dit neemt niet weg dat het voor de kennisverwerving van de behandelde methoden en technieken van belang is om oefenopgaven beschikbaar te hebben. Om deze reden is parallel aan dit boek het werkboek van Van Dalen (2004) verschenen. Hierin vindt men uiteenlopende opgaven voor een verder begrip van zowel de theoretische achtergrond als de toepassing van statistische methoden en technieken.

Verder zijn in de loop der jaren talloze statistiekboeken verschenen, die inhoudelijk vaak weinig verschillen. Voorbeelden zijn Berenson et al. (2004), Newbold et al. (2003), Kohler (2002), Aczel (1995), en Aczel en Sounderpan-dian (2002). Alle vijf zijn ze zeer bruikbaar, waarbij Newbold et al. (2003) zich onderscheidt met een relatief 'sobere' opzet, Kohler (2002) zeer gedegen is,

terwijl Aczel (1995) aardige anekdotes opvoert ter illustratie van de achtergrond van verschillende technieken. Andere boeken met een vergelijkbare opzet zijn Anderson et al. (2005), Anderson et al. (2006), Black (2004), Bowerman en O'Connell (1990), Cavanos en Miller (1993), Kitchens (1998), Kvanli et al. (1998) en Owen en Jones (1994). Nederlandse studenten zijn vaak te spreken over Abrahamse et al. (1987), hetgeen recht-toe-recht-aan, zwart-wit en zonder overbodige voorbeelden is uitgevoerd.

Specifieke uitleg over statistische theorie vindt men bijvoorbeeld bij Craig en Hogg (2005), Lindgren (1993) en Mood et al. (1982). Meer methodologisch getinte onderwerpen van steekproeftrekken worden behandeld in bijvoorbeeld Muilwijk et al. (1992) en Segers en Hagens (1983). Voor toepassingen van econometrische modelbouw kan men onder andere terecht bij Maddala (1992) en Judge et al. (1988).

Voorts bestaat er een grote hoeveelheid boeken over de werking van de statistische programmatuur. De verschillende handboeken van Norušis (1992), Norušis (2005a), Norušis (2005b), en Norušis (2006) geven een goed overzicht van de beschikbare statistische procedures in *SPSS for Windows* en gaan tevens in op de theoretische achtergrond van deze procedures. Tegenwoordig kan men overigens goed terecht op internet voor online oplossingen voor specifieke vragen. Huizingh (2006) is een goede Nederlandse handleiding voor *SPSS for Windows*. Buitengewoon aardig is Schreuder (1991), dat de verschillende procedures van *SPSS for Windows* systematisch weergeeft via een soort 'expert-systeem'.