A. Statistiek en steekproeven

A.1 Inleiding

Bij marktonderzoek worden grote hoeveelheden gegevens verzameld. De kunst is deze gegevens om te zetten naar waardevolle informatie die gebruikt kan worden om zinvolle beslissingen te nemen. Zo is een eerste stap de data te ordenen. Eenvoudigweg sorteren met behulp van een spreadsheetprogramma. Na het ordenen ontstaan er soms meteen al herkenbare patronen. Maar vaak moet er meer analyse worden gedaan om interessante informatie uit de data te krijgen. Bij veel online onderzoekprogramma's wordt de data al door het programma zelf geordend en gevisualiseerd met behulp van bijvoorbeeld grafieken.
In veel gevallen is ordenen niet genoeg om het juiste inzicht te krijgen. De volgende stap is dan om statistische technieken te gebruiken. Statistische technieken gebruikt men om systematisch verzamelde gegevens te bewerken en te analyseren om zo inzicht te krijgen in zogenaamde massaverschijnselen. Het uiteindelijke doel is er zo mogelijk waarschijnlijkheidsuitspraken over te doen. Waarschijnlijkheidsuitspraken zijn nodig omdat de verzamelde informatie vaak maar een deel vertegenwoordigt van een groter geheel. Als een ziekenhuis een klanttevredenheidsonderzoek doet onder 1.000 patiënten en uit de resultaten blijkt dat 70% van deze patiënten tevreden is, in hoeverre zegt dit onderzoek dan iets over de klanttevredenheid onder alle patiënten? Om de klanttevredenheid van alle patiënten te meten, zouden alle patiënten geënquêteerd moeten worden. Maar dat is in de praktijk vaak niet haalbaar. Dan wel kostentechnisch of uit praktische overwegingen. Als het onderzoek morgen herhaald zou worden, komt er dan weer precies hetzelfde antwoord uit? Waarschijnlijk niet. Dus wat is de waarheid?

Met behulp van op statistische basis onderbouwde regels en formules kan een onderzoeker op basis van onderzoek bij een deel van een groep toch uitspraken doen over de totale groep, maar wel met in het achterhoofd dat deze uitspraken een bepaalde waarschijnlijkheid in zich herbergen. Zo zal bij een onderzoek de uitkomst gepresenteerd moeten worden met daarbij een uitspraak over de betrouwbaarheid en de nauwkeurigheid van die uitspraak. Men zou bijvoorbeeld als charitatieve instelling kunnen onderzoeken hoeveel mensen in een bepaalde stad goede doelen ondersteunen met een donatie. Door 1.200 mensen in die stad te ondervragen, zou het antwoord kunnen zijn: 37%. Maar het onderzoeksbureau zal daarbij opmerken dat deze uitspraak gebaseerd is op een betrouwbaarheid van 95% en een nauwkeurigheidsinterval van 34 tot 40%. Dat betekent dat er aangenomen mag worden dat het aantal mensen dat in die stad goede doelen ondersteunt met een donatie met een zekerheid van 95% ligt tussen 34 en 40%. Het echte percentage kan 37% zijn of 40% of alles ertussenin. Er bestaat echter nog steeds een kans van 5% dat het echte percentage buiten het interval ligt. In de praktijk vergeet men dat allemaal en heeft het alleen maar over 37%. Maar dat is in feite niet correct!
In de rest van deze appendix wordt stap voor stap duidelijk hoe statistiek bijdraagt aan een goede onderbouwing van beslissingen en worden begrippen zoals betrouwbaarheid, nauwkeurigheid en representativiteit uitgelegd.

Er zijn twee soorten statistiek: de beschrijvende statistiek en de verklarende statistiek. De beschrijvende statistiek is dat deel van de statistiek dat zich bezighoudt met de verwerking en weergave van feiten, zodanig dat een goed overzicht van de gegevens ontstaat. De beschrijving van de gegevens vindt plaats door middel van statistische kengetallen of door tabellen.
De verklarende statistiek probeert de samenhang tussen bepaalde verschijnselen te verklaren of probeert voorspellingen te doen over toekomstige situaties.

A.2 Groeperen en presenteren van gegevens

Wil men een goed inzicht krijgen in de voor een onderzoek verzamelde gegevens, dan moeten deze gegroepeerd en geclassificeerd worden.

De frequentieverdeling

Een frequentieverdeling van een variabele geeft aan hoe vaak elke waarde (of klasse) van de variabele voorkomt.

Bijvoorbeeld: er wordt aan twintig mensen gevraagd hoeveel uur zij gemiddeld per week televisiekijken. Hun antwoorden zijn als volgt: 10, 23, 21, 8, 2, 12, 28, 28, 11, 23, 3, 15, 17, 0, 4, 14, 15, 18, 20 en 16 uur.

Zo'n rij met getallen zegt op zich nog niet veel. Er zijn nog geen conclusies aan te verbinden. Maar als ze in een frequentieverdeling worden gezet, wordt dat heel anders.
In dit voorbeeld is het aantal uren dat men televisie kijkt, ondergebracht in groepen van 0 tot 6, van 6 tot 12 enzovoort. Deze groepen worden klassen genoemd. De klassenbreedte in dit voorbeeld is 6. Het klassenmidden is 3, 9, 15, 21 en 27.

Figuur A.1 Frequenties

Klasse	Frequentie
0 - < 6	4
6 - < 12	3
12 - < 18	6
18 - < 24	5
24 - < 30	2
Totaal	20

De verdeling in figuur A.1 geeft de gegevens weer in absolute vorm. Maar ze kunnen ook in relatieve vorm weergeven worden. In dat geval deelt men de gevonden frequentie door het totaal en maakt er een percentage van:

Figuur A.2 Relatieve frequentie

Klasse	Frequentie	Relatieve frequentie
0 - < 6	4	20%
6 - < 12	3	15%
12 - < 18	6	30%
18 - < 24	5	25%
24 - < 30	2	10%
Totaal	20	100%

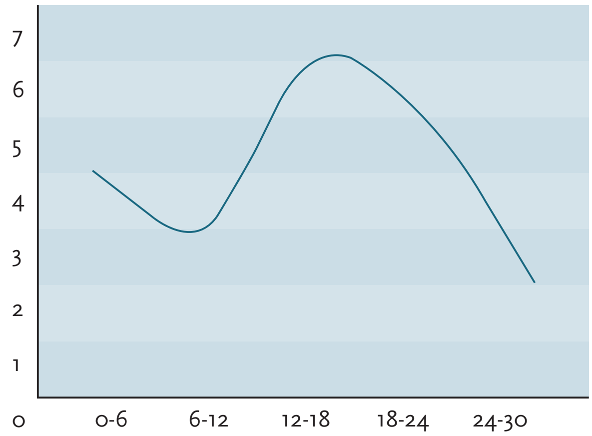
Bij de *cumulatieve frequentieverdeling* wordt het aantal waarnemingen per klasse bij elkaar opgeteld. Hiermee ontstaat inzicht in het aantal waarnemingen dat boven of onder een bepaalde grens ligt.

Figuur A.3 Cumulatieve frequentie

Klasse	Frequentie	Relatieve frequentie	Klasse	Cum. absoluut	Cum. relatief
0 - < 6	4	20%	Tot 6	4	20%
6 - < 12	3	15%	Tot 12	7	35%
12 - < 18	6	30%	Tot 18	13	65%
18 - < 24	5	25%	Tot 24	18	90%
24 - < 30	2	10%	Tot 30	20	100%
Totaal	20	100%			

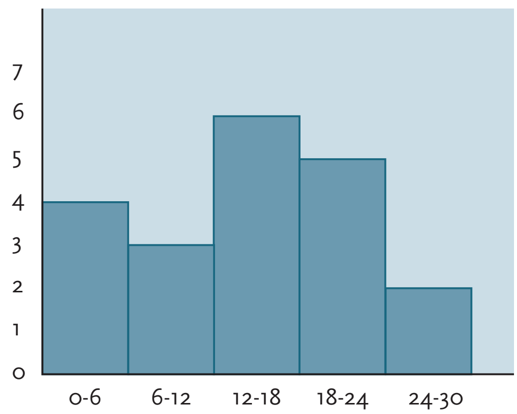
Als men de gegevens uitzet in een *frequentiepolygoon* of een *lijndiagram*, dan ontstaat de grafiek in figuur A.4.
De frequenties worden door punten aangegeven die boven de klassenmiddens staan. Daarna worden de punten door rechte lijnen met elkaar verbonden. Op de x-as staat het aantal uren televisiekijken en op de y-as het aantal respondenten dat een bepaald aantal uur televisie kijkt.

Figuur A.4 Lijndiagram



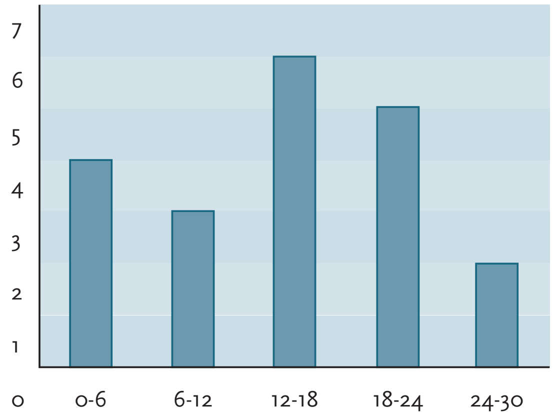
Een andere manier van presenteren is in de vorm van een *histogram* (kolommendiagram, figuur A.5). In een histogram worden kolommen getekend die op elkaar aansluiten. Bij klassenbreedtes die even groot zijn, is de frequentie evenredig aan de hoogte van de kolom. Bij ongelijke klassenbreedtes is de frequentiedichtheid evenredig aan de hoogte van de kolom. Frequentiedichtheid is de absolute frequentie van de klasse gedeeld door de klassenbreedte. Op de y-as staat in dat geval de frequentiedichtheid.

Figuur A.5 Histogram



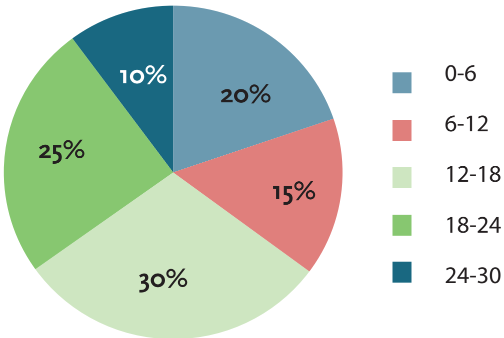
De grafiek in figuur A.6 is een voorbeeld van een *staafdiagram*. Bij een staafdiagram presenteert men bijvoorbeeld de resultaten over smaken of voorkeuren. Zaken die niet uit opeenvolgende klassen bestaan.

Figuur A.6 Staafdiagram



De frequentiegegevens gepresenteerd in de vorm van een *cirkeldiagram* of pie-chart.

Figuur A.7 Cirkeldiagram



De concentratiecurve

Een speciale curve is de concentratiecurve. Men komt deze tegen bij de 20-80-regel. Enkele voorbeelden van de 20-80-regel in de praktijk:

– in de groothandel komt het regelmatig voor dat men met 20% van het assortiment 80% van de omzet binnenhaalt;
– in de levensmiddelenbranche maakt 20% van de grootste bedrijven 80% van de totale winst;
– bij de facturering blijkt dat 20% van de debiteuren 80% van het uitstaande krediet voor zijn rekening neemt.

De 20-80-regel moet niet al te letterlijk genomen worden. De verhouding kan ook bijvoorbeeld 25-75 zijn. Deze regel is een aanwijzing dat men eraan moeten denken dat soms een beperkt aantal (cliënten, producten, facturen) heel veel invloed op het totaal (omzet, winst, kosten) kan hebben.

Voorbeeld
Een leverancier van kantoorartikelen vraagt zich af in welke mate de grote cliënten uit het bestand bijdragen aan de omzet van het bedrijf. Om dit soort vragen te beantwoorden, maakt men gebruik van een concentratiecurve. In een concentratiecurve worden twee variabelen voor hetzelfde bedrijf, in dezelfde periode en in dezelfde grootteklasse, tegen elkaar afgezet. De curve wordt gebaseerd op cumulatieve aantallen.

De boekhouding levert de volgende gegevens aan over het afgelopen jaar.

Figuur A.8 Bedrijfsgegevens

Ordergrootte	Aantal klanten	Relatief %	Totale omzet	Relatief %
0 -< 25	32	12,8 %	400	0,8
25 -< 50	88	35,2%	3.300	6,4
50 -< 100	49	19,6%	3.675	7,1
100 -< 250	31	12,4%	5.425	10,5
250 -< 750	28	11,2%	14.000	27,2
750 -< 1.500	22	8,8%	24.750	48,0
Totaal	250	100,0%	51.550	100,0

Uit figuur A.8 valt af te lezen dat de 8,8% grootste klanten 48% van de omzet genereert. Om deze gegevens grafisch weer te geven in een concentratiecurve gaat men als volgt te werk. In de curve wordt het gecumuleerde aantal klanten op de x-as uitgezet. Deze beschouwt men als de oorzaakvariabele. De gecumuleerde omzet zet men op de y-as (als gevolgvariabele). De concentratiecurve wordt getekend aan de hand van cumulatieve relatieve percentages. Deze plaatst men, gerangschikt van hoge naar lage klasse, in een tabel. De cumulatieve percentages lopen van 0% naar 100%.

Figuur A.9 Cumulatieve percentages

Ordergrootte	Cumulatief aantal klanten %	Gecumuleerde omzet %
> 1.500	0,0	0,0
> 750	8,8	48,0
> 250	20,0	75,2
> 100	32,4	85,7
> 50	52,0	92,8
> 25	87,2	99,2
> 0	100,0	100,0

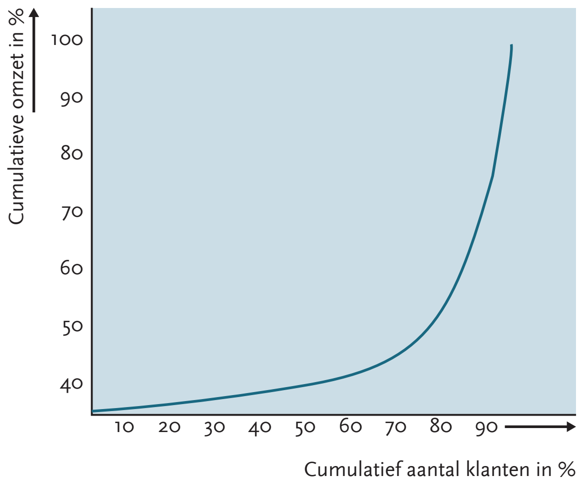
Op basis van de gegevens uit figuur A.9 valt te zien wat het vermoedden al was. Namelijk: 20% van het klantenbestand zorgt voor 75% van de totale omzet. Op basis hiervan kan het management beslissingen nemen, over bijvoorbeeld de minimale ordergrootte, de hoogte van de berekende administratiekosten, het in rekening brengen van bezorgkosten en dergelijke.
In figuur A.8 is de ordergrootte gerangschikt van hoog naar laag. Men begint dus met de grote klanten. Over het algemeen worden in een concentratiecurve de variabelen gerangschikt van klein naar groot. Dit zou in het voorbeeld leiden tot figuur A.10.

Figuur A.10 Gerangschikte cumulatieven

Ordergrootte	Cumulatief aantal klanten %	Gecumuleerde omzet %
< 0	0,0	0,0
< 25	12,8	12,8
< 50	48,0	48,0
< 100	67,6	67,6
< 250	80,0	80,0
< 750	91,2	91,2
< 1.500	100,0	100,0

In figuur A.11 is de uiteindelijke concentratiecurve te zien.

Figuur A.11 Concentratiecurve



Als men een concentratiecurve tekent, blijkt dat de curve aan de andere kant van de diagonaal loopt. De curve begint bij de kleine klanten. Een concentratiecurve geeft de verdeling aan. Hoe krommer de curve, hoe schever de verdeling. In dit voorbeeld kunnen we spreken van een scheve verdeling: 48% van de kleinste klanten draagt slechts bij aan 7,2% van de omzet. En 80% draagt bij aan 24,8% van de omzet.

A.3 Gemiddelden

Eerder in dit hoofdstuk is de frequentieverdeling behandeld. Daarbij wordt een aantal gegevens omgezet in groepen (klassen) en die groepen worden weergegeven in diverse vormen van grafieken. De gegevens zijn nog verder samen te vatten, waardoor bepaalde specifieke eigenschappen van een frequentieverdeling eruit springen. Enerzijds valt te bepalen wat het centrum is van de frequentieverdeling. Men zoekt dan naar de centrale tendentie (of centrummaat of locatiemaatstaf). Anderzijds valt te bepalen hoe de waarden verspreid zijn over de frequentieverdeling.
De gemiddelden geven de centrale tendentie aan.

Het rekenkundig gemiddelde

Bijna iedereen kan wel het gemiddelde van iets uitrekenen. Men berekent dan eigenlijk het ongewogen *rekenkundig gemiddelde*.
Neem bijvoorbeeld de waarden 4, 7, 3, 8, 1, 4, 4, 6, 9, 12, 40 en 27. Het ongewogen rekenkundig gemiddelde wordt berekend door alle getallen bij elkaar op te tellen en dat totaal te delen door het aantal waarden. In dit geval dus:

$$\frac{4+7+3+8+1+4+4+6+9+12+40+27}{12} = 10{,}4$$

Bij het ongewogen rekenkundig gemiddelde hebben alle waarden evenveel invloed op de uitkomst.
In dit voorbeeld zijn er twee waarden die erg afwijken van de rest en een grote invloed hebben op het gemiddelde. Daardoor kan er een verkeerd beeld ontstaan over de verdeling. Door ook andere locatiemaatstaven te berekenen, krijgt men een vollediger beeld van de frequentieverdeling.
Stel dat de bovenstaande getallen de prijzen zijn van twaalf verschillende producten, dan zou de gemiddelde prijs € 10,40 zijn geweest. Maar als er per product verschillende aantallen zijn verkocht, dan zal de gemiddelde prijs waarschijnlijk op een ander bedrag uitkomen.

Figuur A.12 Aantallen per product

Product	Prijs in €	Verkochte aantallen
A	4	12
B	7	3
C	3	4
D	8	40
E	1	5
F	4	6
G	4	3
H	6	7
I	9	3
J	12	1
K	40	2 +
L	27	91

Het gewogen rekenkundig gemiddelde wordt dan:

$$\frac{(4x12)+(7x3)+(3x5)+(8x4)+(1x40)+(4x5)+(4x6)+(6x3)+(9x7)+(12x3)+(40x1)+(27x2)}{91} = € 4{,}52$$

De verkochte aantallen heten de gewichten of wegingscoëfficiënten.

De mediaan

Ook een eenheid van centrale tendentie is de *mediaan*. Alle waarden worden op volgorde van grootte gezet en de middelste waarde is dan de mediaan.
Is het aantal waarden een even getal, dan is de mediaan het gemiddelde van de middelste twee waarden. De mediaan is de middelste waarde, dus 50% van de waarden ligt lager en 50% ligt hoger. Bij het eerste voorbeeld in dit hoofdstuk zijn er 12 waarden. Men neemt dus de zesde en de zevende waarde (toevallig 6 en 7) en de mediaan is dan 6,5 (1, 2, 4, 4, 4, 6, 7, 8, 9, 12, 27, 40).

In het tweede voorbeeld kennen we 91 waarden. Men bepaalt dan de zesenveertigste waarde en dat is in dit geval € 4.

Modus

De modus is de waarde die het vaakst voorkomt. In het voorbeeld in figuur A.13 is dat dus € 1.

Figuur A.13 Verkochte aantallen gesorteerd op prijs product.

Product	Prijs in €	Verkochte aantallen
E	1	40
C	3	5
A	4	12
F	4	5
G	4	6
H	6	3
B	7	3
D	8	4
I	9	7
J	12	3
L	27	2
K	40	1 +
		91

Met betrekking tot de frequentieverdeling heeft men het vaak over de modale klasse. Dat is de klasse met de hoogste frequentie, tenminste als de klassenbreedte van alle klassen even groot is. Zie bijvoorbeeld figuur A.14.

Figuur A.14 Modus

	A	B	
Leeftijd in jaren	Aantal waarnemingen	Verhouding klassenbreedte	A/B
0 < 10	20	2	10
10 < 20	24	2	12
20 < 25	15	1	15
25 < 30	18	1	(modus) 18
30 < 40	22	2	11
40 < 60	32	4	8
60 < 80	28	4	7

De klasse 40 < 60 lijkt in eerste instantie de grootste frequentie te hebben, maar nadat de klassenbreedte is gecorrigeerd, blijkt de modale klasse in dit voorbeeld klasse 25 < 30 te zijn.
In het dagelijkse spraakgebruik heeft men het ook vaak over het modale inkomen (Jan Modaal). Het gaat dan over de inkomensklasse waar de meeste Nederlanders in vallen. Vanuit de politiek is het belangrijk om met deze groep rekening te houden, omdat het er zo veel zijn. Een verandering in belastingwetgeving die vooral betrekking heeft op de modale klasse, kan grote gevolgen hebben op het kiesgedrag. Het modale inkomen (2011 € 33.000) ligt in Nederland lager dan het gemiddelde inkomen, omdat de verdeling van inkomens naar boven afwijkt.

A.4 Spreidingsmaatstaven

Om inzicht te krijgen in de manier waarop een populatie is opgebouwd, geeft alleen een gemiddelde weinig inzicht. Zo is het gemiddelde van 99 en 101 precies 100, maar het gemiddelde van 1 en 199 is ook 100. De spreiding is echter wel duidelijk verschillend. Spreiding wordt weergeven als variatiebreedte, gemiddelde afwijking en als standaarddeviatie.

Variatiebreedte

De *variatiebreedte* of range geeft aan wat het verschil is tussen de grootste waarde en de kleinste waarde in de populatie.
Zijn die niet bekend, omdat de waarnemingen zijn ondergebracht in klassen van een frequentieverdeling, dan is de variatiebreedte het verschil tussen de bovengrens van de hoogste klasse en de ondergrens van de laagste klasse. In het voorbeeld bij de behandeling over de modus is de range dus 80. De variatiebreedte zegt nog maar weinig over de spreiding, omdat één uitspringende waarde het beeld sterk kan beïnvloeden.

Gemiddelde afwijking

Bij de *gemiddelde afwijking* wordt gekeken naar het verschil tussen een waarde en de gemiddelde waarde. Dat verschil drukt men uit in de absolute waarde (dus zonder de min). Die waarden telt men bij elkaar op en deelt men door het aantal waarden.

Figuur A.15 Gemiddelde afwijking

Naam	Lengte in cm	Afwijking absoluut
Jan	178	7
Peter	183	2
Paul	186	1
Gerard	175	10
Kees	194	9
Willem	189	4
Bert	192	7
Gemiddeld	185	

Als er in dit voorbeeld ook meisjes waren meegenomen, dan was de gemiddelde afwijking waarschijnlijk groter geweest, omdat meisjes in het algemeen kleiner zijn dan jongens.

Standaarddeviatie

De meest gebruikte spreidingsmaatstaf is de *standaarddeviatie*. De formule luidt als volgt:

$$\sigma = \sqrt{\frac{\Sigma\left(x - x_{gem}\right)^2}{n}}$$

met:
σ = standaarddeviatie (spreek uit: sigma)
x = de waarneming
x_{gem} = het gemiddelde
Σ = de som
n = het aantal waarnemingen
In het voorbeeld van de jongens krijgt men dan:

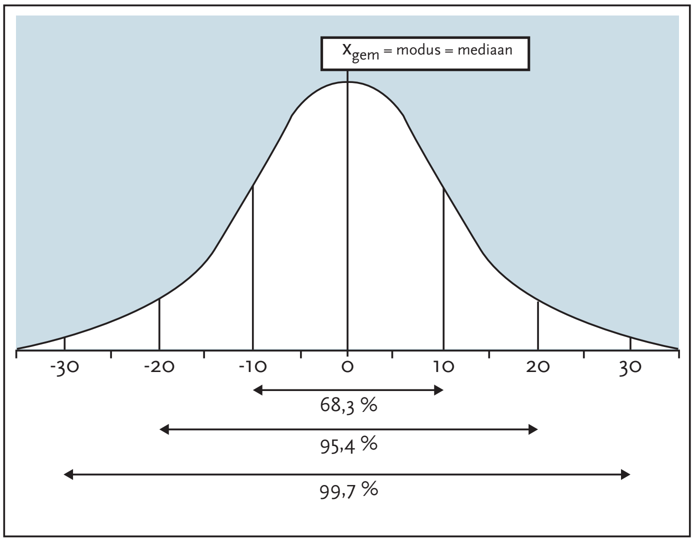
$$\sigma = \sqrt{\frac{\Sigma\left((7)^2 + (2)^2 + (1)^2 + (10)^2 + (9)^2 + (4)^2 + (7)^2\right)}{7}} = 6{,}547$$

De standaarddeviatie wordt veel gebruikt bij de berekening van de uitkomst van een steekproef. Men kan er een schatting mee maken hoeveel van de waarden uit de frequentieverdeling in de buurt van het gemiddelde liggen, uitgaande van een *normale verdeling*. Verderop wordt dit duidelijk.

Normale verdeling

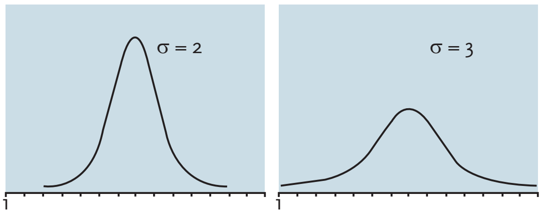
De wiskundigen Gauss en Laplace hebben proefondervindelijk vastgesteld dat bij een normale verdeling 68,3% van alle waarnemingen ligt tussen de grenzen die worden gevormd door het gemiddelde plus of min éénmaal de standaarddeviatie. Bij tweemaal de standaarddeviatie is dat 95,4%, bij driemaal de standaarddeviatie valt 99,7% van de waarnemingen binnen die grenzen.

Figuur A.16



De mate waarin de oppervlakte onder de verdeling verdeeld is, is afhankelijk van de standaarddeviatie.

Figuur A.17



Deze figuren hebben hetzelfde gemiddelde en dezelfde mediaan en modus, maar de spreiding is heel anders. Met de standaarddeviatie krijgt men dus een goed beeld van de spreiding. Het linker plaatje heeft een kleinere σ dan het rechter plaatje.

De volgende stap is het toepassen van deze statistische modellen bij steekproeven. Maar voordat we dat doen, zullen we eerst duidelijker maken wat steekproeven zijn en waarom deze zo belangrijk zijn.

Reden voor steekproefonderzoek

Als er een uitspraak gedaan moet worden over een grote groep mensen (de populatie), dan is het in de regel te duur of te tijdrovend om hen allemaal te ondervragen. Maar met behulp van statistische technieken kan men, door maar een klein deel van de populatie te ondervragen, toch veel zeggen over de hele groep. Als maar een deel van de gehele populatie wordt ondervraagd, heet dat een *steekproef*. De definitie luidt als volgt: ‘Een steekproef is een selectie van een subgroep van elementen uit een grotere verzameling van elementen, de populatie.’ Logischerwijs is dan de populatie de verzameling elementen waarover men in het kader van een steekproefonderzoek uitspraken wil doen.

Met een steekproef kan en mag men de uitkomsten van die steekproef generaliseren voor de hele populatie. Voorwaarde is wel dat de steekproef niet te klein is, de steekproefelementen aselect gekozen zijn en de steekproef representatief is.
De te ondervragen mensen of bedrijven worden gekozen uit het *steekproefkader*. Het steekproefkader is een fysieke lijst van *elementen* (personen, huishoudens, bedrijven enzovoort) waaruit een steekproef wordt getrokken. Als er bijvoorbeeld een steekproef bepaald moet worden om een onderzoek te doen onder alle huishoudens in Bussum, dan zou men de telefoongids kunnen pakken als steekproefkader (deze manier wordt onbetrouwbaarder nu er steeds meer huishoudens zijn die alleen een mobiele telefoon hebben). Voor bedrijven zou de Gouden Gids gebruikt kunnen worden. Maar er zijn allerlei bedrijven en instanties die adressen en lijsten kunnen leveren die te gebruiken zijn als steekproefkader. Belangrijk is dat er goed op wordt gelet dat de lijst representatief is voor de populatie die onderzocht moet worden. Als men dus een steekproef wil houden over de consumptie van vis bij huishoudens in Nederland, dan kan het telefoonboek van Volendam wel eens heel andere resultaten opleveren dan dat van Oldenzaal.

Waarde van steekproefresultaten

Om goede uitspraken te kunnen doen over de totale populatie, moet men wel een paar zaken goed in de gaten houden.

Representativiteit
Met representativiteit wordt bedoeld dat de elementen binnen een steekproef een getrouwe afspiegeling moeten vormen van de totale populatie. Onderzoekt men bijvoorbeeld de fileproblemen in Nederland, maar ondervraagt men alleen mensen in de Randstad, dan is dat geen representatieve steekproef.
Ook non-respons beïnvloedt de representativiteit van een steekproef. In dat geval zijn er mensen op de steekproeflijst die niet mee willen of kunnen werken of niet bereikbaar zijn op het moment dat het onderzoek wordt gedaan. Als deze non-respondenten een gemeenschappelijk kenmerk hebben (ze zijn bijvoorbeeld overdag aan het werk) en men wil overdag een onderzoek doen bij werkende en niet werkende mensen, dan zullen de uitkomsten van dat onderzoek niet representatief zijn voor de hele populatie.

Betrouwbaarheid
Een steekproef is betrouwbaar als men bij herhaalde meting dezelfde resultaten krijgt. Natuurlijk wel onder de voorwaarde dat de omstandigheden vergelijkbaar zijn. Als er wordt bepaald dat de betrouwbaarheid van een steekproef 96% is, dan betekent het dat als men die steekproef zou herhalen, in 96 van de 100 gevallen dezelfde uitkomst uit de bus zal rollen.

Nauwkeurigheid
Bij nauwkeurigheid kijken we naar in hoeverre het werkelijke resultaat af kan wijken van het gemeten resultaat. Meestal wordt dat weergegeven in een percentage. Dus als de gevonden waarde 56,2% is en de nauwkeurigheid 2,2%, dan ligt de echte waarde ergens tussen de 54% en 58,4%.
Naarmate iets betrouwbaarder is, is het minder nauwkeurig en andersom. Dit wordt aangeduid als een omgekeerd evenredig verband. De uitspraak ‘Volgende week zal er tussen de 0 en 100 millimeter regen vallen’, is erg betrouwbaar, maar niet erg nauwkeu-

rig. Maar de uitspraak 'Volgende week valt er 4,75 millimeter regen' is zeer nauwkeurig, maar erg onbetrouwbaar.

Nog een opmerking over de steekproefgrootte. Hoe kleiner de steekproef, hoe kleiner de nauwkeurigheid en de betrouwbaarheid van de uitkomst. Maar andersom gaat dat niet helemaal op. Met statistiek is te bepalen hoe groot een steekproef moet zijn om een betrouwbaarheid van bijvoorbeeld 97% en een nauwkeurigheid van 2% te krijgen. Maakt men de steekproef veel groter, dan neemt de betrouwbaarheid niet noodzakelijk toe. Deze kan zelfs afnemen, omdat er bij grote aantallen te verwerken informatie ook eerder fouten gemaakt worden. Er is dus zoiets als een optimale steekproefgrootte.

Validiteit
Als men willen weten hoeveel mannen het een feest vinden om te winkelen en men telt alle mannen die samen met hun vrouw het winkelcentrum inlopen, dan is er een hele grote kans dat men niet meet wat men denkt te meten. De validiteit is dus de mate waarin wat men meet, overeenkomt met wat men beoogt te meten.

Standaardfout bij steekproeven

Het kunnen berekenen van een standaarddeviatie is een opstapje naar het kunnen interpreteren van de uitkomsten van steekproeven.
Eerder in dit hoofdstuk werd gesproken over de betrouwbaarheid en de nauwkeurigheid van de uitkomst van een steekproef. Het is duidelijk dat als men een deel van een totale populatie (zeg tweehonderd van de tweeduizend) vraagt of zij in het bezit zijn van een auto, de uitkomst van die steekproef (zeg 34%) nooit precies dezelfde uitkomst zal zijn als die van de totale populatie. Elke keer als men opnieuw tweehonderd mensen uit de populatie ondervraagt, zal er een ander percentage uitkomen. Toch valt te bepalen met welke zekerheid het percentage van de hele populatie binnen een bepaald interval rond die 34% valt.

Ervan uitgaande dat de uitkomsten van een steekproef verdeeld zijn volgens een normale verdeling, kan de onderzoeker uitspraken doen op basis van de standaarddeviatie. Men kan bijvoorbeeld zeggen: 'Als men een betrouwbaarheid van de uitkomsten van een steekproef wil hebben van 99,7%, dan ligt de echte waarde van de steekproef ergens tussen driemaal de standaarddeviatie naar links en driemaal de standaarddeviatie naar rechts van de gevonden waarde.'

Specifiek bij steekproeven waarbij het antwoord op een vraag ja of nee is, mag men de volgende vereenvoudigde formule gebruiken:

$$s = \sqrt{\frac{a \times b}{n}}$$

a is het percentage dat voldoet aan een bepaald kenmerk (bijvoorbeeld: merk X) en b is gelijk aan (100 – a). n is de omvang van de steekproef. Bij de steekproef heet de uitkomst van deze formule de *standaardfout.*

Het aantal keren dat de standaardfout afwijkt van de gevonden waarde noemt men de Z-waarde. Dus bij een Z-waarde van drie zal de kans 99,7% zijn dat de echte waarde binnen driemaal de standaardfout naar links en naar rechts van de gemeten waarde zal liggen. Bij verschillende Z-waarden gelden de volgende betrouwbaarheidsniveaus:

Figuur A.18

Betrouwbaarheidsniveau	Z-waarde
99,7 %	Z = 3
99,0 %	Z = 2,58
95,4 %	Z = 2
95,0 %	Z = 1,96
90,0 %	Z = 1,65
68,3 %	Z = 1

Met behulp van deze kennis en de formule om de standaardfout te berekenen, is het nu ook mogelijk om te berekenen wat het zekerheidsinterval is bij een bepaalde betrouwbaarheid. De formule voor het berekenen van de standaardfout wordt nu uitgebreid tot:

Waarin:

$$I = a \pm Z \sqrt{\frac{a \times b}{n}}$$

I = het interval van het te schatten percentage in de populatie
a = het gevonden percentage in de steekproef als getal (30% is dan 30)
b = 100 – a
n = de grootte van de steekproef
Z = de constante die samenhangt met de gewenste mate van betrouwbaarheid van de uitspraken die men over de uitkomst betreffende de populatie wil doen.

In de praktijk gebruikt men bij marktonderzoek altijd een betrouwbaarheidsniveau van 95% of 90% (dus een Z-waarde van respectievelijk 1,96 of 1,65).

Voorbeeld
Er wordt aan vijfhonderd aselect gekozen inwoners van Zwolle gevraagd of zij een auto bezitten. Alle respondenten geven antwoord. Nadat alle vijfhonderd respondenten hebben geantwoord, blijkt dat 60% van hen een auto bezit. Maar, zoals eerder uitgelegd, het echte aantal autobezitters in Zwolle zal waarschijnlijk niet precies 60% zijn. Het zal wel ergens in de buurt van dit percentage liggen. Om nu een uitspraak te kunnen doen over het verwachte percentage autobezitters in Zwolle, met als voorwaarde dat de uitspraak een betrouwbaarheid van 95,4% moet hebben (Z-waarde is 2), wordt de berekening als volgt:

$$60 \pm 2 \times \sqrt{\frac{60 \times 40}{500}} = 60 \pm 4{,}4$$

dus de grenzen van het interval zijn 55,6% en 64,4%.

Met andere woorden: het percentage autobezitters in Zwolle ligt tussen de 55,6% en de 64,4%, en dat valt met een zekerheid van 95,4% te beweren. De kans dat het percentage autobezitters buiten dit interval zal liggen, is 100% minus de betrouwbaarheid, dus 100% – 95,4% = 4,6%. Dit noemt men de overschrijdingskans.

Zoals eerder besproken, heeft elk steekproefresultaat te maken met een bepaalde betrouwbaarheid in combinatie met een daarbij behorende nauwkeurigheid. De betrouwbaarheid in het hiervoor genoemde voorbeeld is 95,4% en de nauwkeurigheid is 4,4%. Tweemaal de nauwkeurigheid, in dit geval 8,8%, bepaalt het interval en wordt soms ook de *onnauwkeurigheidsmarge* genoemd.
Neemt men genoegen met een lagere betrouwbaarheid van bijvoorbeeld 90%, dan wordt de standaardfout (2,19) vermenigvuldigd met 1,65 (de Z-waarde bij 90% betrouwbaarheid). Het interval wordt dan 60 plus en min 3,6% (56,4 – 63,6). Het antwoord is dus nauwkeuriger (onnauwkeurigheidsmarge gedaald van 8,8% naar 7,2%), maar minder betrouwbaar (gedaald van 95,4% naar 90%).

Met behulp van de hiervoor genoemde formules wordt het nu eenvoudig om te bepalen hoe groot een steekproef moet zijn om met een bepaalde betrouwbaarheid een gevraagd maximum interval te krijgen. Bijvoorbeeld uitgaande van een betrouwbaarheid van 95,4% (in vakjargon heet dat een betrouwbaarheid op twee-sigmaniveau) en een maximum nauwkeurigheid van 3%, blijkt dat er 1.067 mensen ondervraagd moeten worden:

$$3 = 2\sqrt{\frac{60 \text{ x } 40}{n}} \Rightarrow \frac{3}{2} = \sqrt{\frac{2.400}{n}} \Rightarrow 1{,}5 = \sqrt{\frac{2.400}{n}} \Rightarrow 2{,}25 = \frac{2.400}{n} \Rightarrow n = \frac{2.400}{2{,}25} = 1.067$$

Probleem is dat de waarde a nog niet bekend is als de steekproef nog niet is uitgevoerd. In de praktijk wordt dan gewerkt met de waarde 50%. Bij 50% wordt de steekproefgrootte 1.112.
In figuur A.19 wordt duidelijk zichtbaar dat bij het vergroten van de steekproef de nauwkeurigheid groter wordt, maar wel in een afnemende mate.

Figuur A.19

Steekproefgrootte	Betrouwbaarheid	Nauwkeurigheid
500	95,4%	4,4%
1.000	95,4%	3,1%
1.500	95,4%	2,5%
2.000	95,4%	2,2%

Naarmate de steekproef groter is, zijn de kosten om het onderzoek te doen ook groter. Dus zal men een afweging moeten maken hoeveel extra geld men wil uitgeven om een nauwkeuriger antwoord te krijgen. In de praktijk komt men met een steekproefgrootte van 1.500 al een heel eind.
Opmerkelijk is dat in alle hiervoor genoemde formules de populatie géén rol speelt! Dus deze formules zijn te gebruiken om een uitspraak te doen over een populatie van 50.000 mensen, maar ook een populatie van 50 miljoen mensen.

Significantie

Stel dat er bij een steekproef (achthonderd) in Bussum gemeten wordt dat daar 66% van de inwoners een auto bezit. Men kan zich dan afvragen of er echt een verschil is tussen het autobezit in Zwolle (60%) en Bussum (66%), of dat het gemeten verschil alleen maar berust op toeval. Immers: het juiste percentage autobezitters in Zwolle ligt hoogstwaarschijnlijk tussen de 55,6% en de 64,4% en kan dus 64,4% zijn. Bij Bussum (ervan uitgaande dat de steekproefgrootte achthonderd en de vereiste betrouwbaarheid 95,4% is) is dat interval 62,7% – 69,3% en kan het juiste percentage 62,7% zijn. In dat geval is het aantal autobezitters in Zwolle dus groter dan in Bussum.
Om het antwoord op de vraag te krijgen of men mag stellen dat er een significant verschil zit tussen het aantal autobezitters in beide plaatsen, gebruikt men de significantietoets.

De formule luidt:

$$v \geq z \times \sqrt{\frac{a1 \times b1}{n1} + \frac{a2 \times b2}{n2}}$$

Waarin v het absolute verschil is tussen de gevonden percentages.
Als er een betrouwbaarheid van 95,4% wordt aangehouden, dan ontstaat de volgende berekening:

$$6 \geq 2 \times \sqrt{\frac{60 \times 40}{500} + \frac{66 \times 34}{800}} \Rightarrow 6 \geq 5{,}5$$

Omdat de gevonden waarde kleiner is dan 6, mag aangenomen worden dat het verschil significant is. Men mag er dus van uitgaan dat er in Bussum en Zwolle verschillende percentages zijn van autobezit.

Soorten steekproeven

Een steekproef kan aselect of niet-aselect zijn. Bij aselecte steekproeven worden de te ondervragen mensen of bedrijven willekeurig geselecteerd uit het steekproefkader. Of, zoals het *NIMA Marketing Lexicon* het definieert: 'Een aselecte steekproef is een steekproef waarbij de elementen uit het steekproefkader volgens toeval worden getrokken en waarbij elk element een van tevoren bekende kans heeft om in de steekproef terecht te komen.'
Bij een niet-aselecte steekproef stelt men juist wel eisen aan de respondenten.

Er zijn verschillende soorten aselecte steekproeven (ook toeval- of kanssteekproeven genoemd) mogelijk:
1. enkelvoudige steekproef;
2. systematische of intervalsteekproef;
3. gestratificeerde steekproef;
4. random-walk steekproef.

Enkelvoudige steekproef
De enkelvoudige steekproef is de meest eenvoudige manier van steekproeftrekking. Als men bijvoorbeeld een up-to-date bestand van de leden van een vereniging heeft, kan men alle lidnummers van die leden op briefjes schrijven en in een hoge hoed stoppen. Als men dan vervolgens blind een aantal briefjes uit de hoge hoed trekt, dan is dat een enkelvoudige steekproefselectie (simple random sample).

Systematische of intervalsteekproef
Bij 10.000 leden is het natuurlijk veel werk om al die briefjes te maken en de hoed moet ook wel heel groot zijn. In dat geval is het beter om een systematische steekproef te gebruiken. Stel, er moeten 500 leden geïnterviewd worden, dan deelt men 10.000 door 500 en krijgt het getal 20 (1 op de 20 leden moet geïnterviewd worden). Iemand kiest een getal onder de twintig (bijvoorbeeld 7) en men begint op de ledenlijst bij het zevende lid. Deze wordt aangevinkt, dan een vinkje bij 27, 47, 67 en zo verder tot alle 500 leden zijn geselecteerd.

De definitie voor de systematische steekproef is:

> *'Een systematische steekproef is een aselecte steekproef waarbij ieder k'de element uit het steekproefkader wordt getrokken en waarbij verondersteld wordt, dat er geen voor het onderzoek relevante ordening van de elementen in het steekproefkader bestaat.'*

Een variant op de systematische steekproef is de clustersteekproef (trossteekproef) of getrapte steekproef. Het is een steekproefselectiemethode waarbij het steekproefkader wordt gesplitst in groepen. Uit deze clusters wordt aselect een keuze gemaakt. Als dan iedereen in een gekozen cluster wordt geïnterviewd, heet het een *eentrapsclustersteekproef* en als er per cluster ook weer aselect een groep respondenten wordt getrokken, heet dat een *tweetrapsclustersteekproef*.

Bijvoorbeeld, er moet een steekproef worden gedaan onder studenten. Men neemt dan eerst een lijst met universiteiten. Daar wordt aselect een aantal universiteiten uit getrokken. Als men nu van de geselecteerde universiteiten *alle* studenten ondervraagt, heet dat een eentrapsclustersteekproef, maar als men per geselecteerde universiteit eerst weer aselect een aantal studenten selecteert, dan wordt dat een tweetrapsclustersteekproef genoemd.

Een bijzondere vorm van de clustersteekproef is de zogenaamde area-steekproef (gebiedssteekproef). Het steekproefkader bestaat hierbij uit een kaart of een plattegrond van een gebied, waarbij de clusters gevormd worden door blokken op de kaart. Dat kunnen stadswijken, provincies of regio's zijn.

Gestratificeerde steekproef
De gestratificeerde steekproef noemt men ook wel de gelaagde of gelede steekproef. Net als bij de clustersteekproef wordt de populatie in groepen verdeeld. Maar deze keer maakt men groepen die een bepaald kenmerk hebben. Als een televisiezender bijvoorbeeld een onderzoek wil doen naar het aantal uren dat mensen televisiekijken, kan zij de populatie opdelen in: kinderen tot 12 jaar, jeugd van 13 t/m 20, mensen van 21 t/m 40, van 41 t/m 60 en mensen ouder dan 60. Er worden dan aselect per groep een aantal respondenten getrokken. Zouden zij hetzelfde onderzoek doen met een clustersteekproef, dan kan het zijn dat er van een bepaalde groep (bijvoorbeeld de kinderen) te weinig respondenten in het onderzoek zitten, waardoor het onderzoek niet representatief is. Voorwaarde voor de representativiteit van de gestratificeerde steekproef is wel dat de groepen proportioneel vertegen-woordigd zijn in de steekproef. Het gaat dus om een proportioneel gestratificeerde steekproef. Is volgens het CBS 7% van de bevolking jonger dan 12 jaar, en uit de steekproef blijkt dat zij gemiddeld 10 uur per week televisiekijken, dan moet men die 10 uur maal 7% nemen als waarde voor de kinderen. Voor alle groepen moet men deze waarde op dezelfde manier bepalen. Worden deze waarden bij elkaar opgeteld, dan krijgt men een proportioneel gemeten aantal uren dat mensen televisiekijken. Worden de waarden niet proportioneel genomen, dan spreken we van een disproportioneel gestratificeerde steekproef. Het zal duidelijk zijn dat deze minder representatief is.

Random-walk steekproef
Deze manier om een steekproef te houden wordt ook wel de start-adressenmethode genoemd. Eerst wordt uit een lijst met adressen, via een intervaltrekkingsmethode, aselect een aantal adressen getrokken. Iedere enquêteur gaat dan naar een geselecteerd adres en van daaruit volgt hij een bepaalde instructie om de volgende respondent te vinden. Bijvoorbeeld: 'Begin bij nummer 62 op de Melkweg en ga daarna de eerste straat links en neem dan het tweede huis dat je tegenkomt. Wordt er niet opengedaan, neem dan het volgende huis. Ga dan de eerste straat links en neem het vierde huis.'

Niet-aselecte steekproeven

Zoals al eerder in dit hoofdstuk is vermeld, wordt bij de niet-aselecte steekproef het toeval niet gebruikt. Er wordt bewust gezocht naar bepaalde mensen. Een voorbeeld van een niet-aselecte steekproef is de quotasteekproef. Hierbij krijgt de onderzoeker een opdracht mee in de vorm van: 'Ondervraag minstens vijftien mensen die een brommer

hebben en veertig mensen die alleen maar een fiets hebben.' De interviewer mag dus zelf de respondenten uitkiezen.
Het mogelijke nadeel van deze manier van selecteren is dat de onderzoeker mensen kiest (bijvoorbeeld uit zijn eigen omgeving) die niet representatief zijn voor de te onderzoeken zaken.

Steekproeffouten

Het zal duidelijk zijn dat het resultaat dat men vindt in een steekproef, lang niet altijd exact gelijk is aan de werkelijke waarde in de beoogde populatie. Het verschil tussen het steekproefresultaat en de werkelijke waarde noemen we de steekproeffout.
Er zijn twee soorten steekproeffouten: toevallige fouten en systematische fouten. Als de fouten berusten op toeval nemen we daar genoegen mee. Als het systematische fouten betreft, dan moeten we daar wat aan doen. Zo kan bijvoorbeeld de selectie van clusters niet representatief zijn. Door systematische fouten kunnen de gemeten waarden veel te hoog of veel te laag zijn, met alle gevolgen van dien.

A.5 Vragenlijst ontwerp

Om bij enquêtes goede resultaten te krijgen, is het uiterst belangrijk veel aandacht te besteden aan het formuleren van de vragen.
Ook bij een enquête gaat het om communiceren. En dus moet de boodschap (de vraag) goed gecodeerd worden en de respondent moet de vraag zo objectief mogelijk kunnen decoderen en beantwoorden. Daarom is er een aantal punten waar men op moet letten.

- De taal moet begrijpelijk zijn voor de doelgroep en onafhankelijk zijn van bijvoorbeeld emoties, cultuur en attitudes.
- Verder moeten vragen kort en bondig zijn en het aantal vragen mag niet te groot zijn.
- Vragen moeten eenduidig zijn. Ze moeten voor maar één uitleg vatbaar zijn.
- Men mag geen vragen stellen die een bepaald gewenst antwoord uitlokken. Dat zijn suggestieve vragen en die vervormen het beeld dat een onderzoek oplevert.
- Vragen waarvan men van tevoren kan verwachten dat de respondent er een belang bij heeft om ze onjuist te beantwoorden, zijn uit den boze.
- Vermijdt dubbelvragen of vragen die elkaars tegengestelde zijn. Bijvoorbeeld: 'Wilt u drie of vier keer per jaar een donatie doen?' Als daar alleen 'Ja' of 'Nee' op geantwoord kan worden, heeft men niets aan de uitkomst.
- Vragen over inkomen en leeftijd kan men beter bewaren tot het laatst. Zij kunnen irritatie opwekken, waardoor de respondent of meteen stopt met de vragenlijst of met een verkeerde mindset begint.

Bij het opstellen van de vragenlijst zijn er verschillende soorten vragen mogelijk. Men kan enerzijds open vragen gebruiken en anderzijds gesloten vragen. Gesloten vragen kan men dan nog weer opdelen in dichotome vragen, meerkeuzevragen (multiple choice) en meervoudige vragen.
Daarnaast kan men directe en indirecte vragen stellen en het is mogelijk om gebruik te maken van verschillende schalen, die in paragraaf A.6 worden behandeld.

Open vragen

Bij een open vraag heeft een respondent alle vrijheid om te antwoorden. Typisch een vraag die men gebruikt om een mening te vragen of als er een veelvoud van antwoorden mogelijk is.
Nadeel van open vragen is de verwerking. Wil men open vragen kwantitatief gaan verwerken met een computer, dan kan men eventueel de antwoorden nog in antwoordcategorieën of klassen indelen. Dit heet coderen.
Een voorbeeld van een open vraag is: 'Kunt u uitleggen waarom u dit product hebt gekocht?'

Gesloten vragen

Bij een gesloten vraag moet de respondent kiezen uit twee of meer vooraf gedefinieerde mogelijkheden. Voordeel van een gesloten vraag is dat deze snel te beantwoorden en eenvoudig te verwerken is.

– Dichotome vragen
  Bij een dichotome vraag zijn er maar twee mogelijke antwoorden. Bijvoorbeeld: 'Vindt u dat het koningshuis moet blijven?', met als mogelijke antwoorden 'Ja' en 'Nee'.
– Meerkeuzevragen (multiple choice)
  Algemeen bekend. Er is maar één antwoord mogelijk.
– Meervoudige vragen
  Bij een meervoudige vraag zijn er meer antwoorden mogelijk. Bijvoorbeeld: 'Welke automerken kent u?', met als mogelijke antwoorden: Alfa Romeo, Audi, BMW, Mercedes, Volvo enzovoort.

Om vragen eenvoudig te kunnen verwerken, kan men meervoudige vragen omzetten naar dichotome vragen. Er wordt dan per antwoord gecodeerd. Dus het antwoord op de meervoudige vraag: 'Welke automerken kent u?', kan zijn: 'Alfa Romeo en Mercedes'. In het verwerkingssysteem wordt dat dan omgezet naar: 'Kent Alfa Romeo ⇒ Ja', 'Kent Audi ⇒ Nee', 'Kent Mercedes ⇒ Ja' enzovoort.

Directe versus indirecte vraagstelling

Bij de directe vraagstelling vraagt men op de man af: 'Hoe denkt u over uw algemeen directeur?'
Dit soort vragen kan leiden tot geforceerde antwoorden. De respondent is wellicht voorzichtig met het geven van een eerlijk antwoord op zo'n vraag, zeker als hij niet erg positief is over de baas. Speciaal bij gevoelige onderwerpen kan het dan beter zijn om de vraag indirect te maken. Bijvoorbeeld: 'Hoe denkt u dat er gedacht wordt over de algemeen directeur?'

Ten slotte is het belangrijk om een goede opbouw van de vragenlijst te hebben. Begin met een gemakkelijke inleidende vraag. Ga dan langzaam naar de moeilijkere vragen. Begin ook met het onderwerp algemeen te benaderen en ga vervolgens wat specifieker op bepaalde zaken in. Op deze manier kan men ook voorkomen dat respondenten vragen moeten beantwoorden die niet voor hen van toepassing zijn. Als ze bijvoorbeeld geantwoord hebben dat ze geen eigen huis bezitten, dan kunnen ze het onderdeel over hypotheken overslaan en verder gaan met andere vragen in het onderzoek. Pas helemaal tegen het einde van de vragenlijst zal men vragen stellen van persoonlijke aard, zoals inkomen, geslacht en dergelijke. Alles bij elkaar mag het aantal vragen niet te groot zijn. Vermijd daarom zo veel mogelijk overbodige c.q. onnodige vragen.

A.6 Schaaltypen

In een kwantitatief onderzoek is het verstandig om zo veel mogelijk vragen te coderen. Daar zijn verschillende schalen voor te gebruiken. Maar er gelden beperkingen voor wat betreft het doen van berekeningen met de verschillende schaaltypen.
Er zijn vier veelgebruikte schaaltypen (te onthouden door het ezelsbruggetje NOIR):
– de nominale schaal;
– de ordinale schaal;
– de intervalschaal;
– de ratioschaal.

Nominale schaal

Bij een nominale schaal hebben de variabelen geen zelfstandige betekenis. Ze zijn in verschillende volgorde te zetten zonder dat het consequenties heeft. Het zijn slechts labels. Man, Vrouw, Kind, Volwassene. Daaraan worden aantallen gekoppeld. Met de labels valt niet te rekenen, want er is geen gemiddelde te bepalen van man, vrouw en kind. De modus is wel te bepalen.

Ordinale schaal

Bij een ordinale schaal geven de kenmerken een rangorde aan. Een voorbeeld is de sterrenindeling van hotels.

5 sterren	=	Zeer luxe
4 sterren	=	Luxe
3 sterren	=	Middenklasse
2 sterren	=	Toeristenklasse
1 ster	=	Backpackersklasse

Er is een logische volgorde, maar de onderlinge afstand tussen de verschillende klassen is niet duidelijk. Men kan niet stellen dat een vijfsterrenhotel 20% beter is dan een viersterrenhotel of zoiets. Het is uit den boze om met de uitkomsten van dit schaaltype rekenkundige bewerkingen uit te voeren. Wel kan men de modus en mediaan bepalen.
Een ordinale schaal die bekend zal voorkomen, is:

1	=	Helemaal mee eens
2	=	Mee eens
3	=	Geen mening
4	=	Mee oneens
5	=	Helemaal mee oneens

Ook hier zit geen rekenkundige afstand tussen de kenmerken. Dus ook hier geen berekeningen proberen uit te voeren.

De nominale en ordinale schalen worden vooral gebruikt in de gedragswetenschappen.

Intervalschaal

Bij deze schaal staan de kenmerken in een bepaalde volgorde en zijn de afstanden tussen de opeenvolgende kenmerken (de intervallen) gelijk. Het enige dat ze niet hebben, is een natuurlijk nulpunt. Ze hebben vaak wel een nulpunt, maar dat is zelf gekozen, zoals bij een kalender. Men kan bij de intervalschaal gemiddelde, modus en mediaan bepalen.

Ratioschaal

Als er wel een natuurlijk nulpunt aanwezig is, dan betreft het een ratioschaal. Voorbeelden zijn gewicht, lengte, afstand, bedragen, hoeveelheden en prijzen. Ze worden vooral gebruikt in natuurwetenschappen en er kunnen allerlei rekenkundige bewerkingen op losgelaten worden.

Continue en discontinue variabelen

Bij het gebruik van schalen kan men werken met continue en discontinue variabelen. Een discontinue variabele is bijvoorbeeld het aantal mensen dat rookt in een klas: 5, 6, 8 of 9. Het antwoord kan nooit 2,3 zijn omdat het om getelde mensen gaat. Maar als het gaat om een hoeveelheid zand, dan kan het 2,345 kilogram zijn (of nog preciezer met meer cijfers achter de komma). Dat zijn continue variabelen. Andere voorbeelden van continue variabelen zijn het gemiddelde inkomen, de tijd dat iemand moet wachten in een rij enzovoort.
Nominale en ordinale variabelen zijn altijd discontinu. Interval- of ratiovariabelen kunnen zowel continu als discontinu zijn.

A.7 Schaaltechnieken

Schaaltechnieken zijn methoden om een schaal te construeren. Onderzoekers proberen door middel van bepaalde schaalmethodieken nominale gegevens op een dusdanige manier te groeperen dat er rekenkundige bewerkingen mee kunnen worden uitgevoerd. De twee meest gebruikte schalen zijn:
– de Likert-schaal;
– de Osgood-schaal.

Likert-schaal

De Likert-schaal is een instrument dat bestaat uit een aantal verschillende uitspraken waarvan de respondent op een vijfpuntenschaal kan aangeven of hij het er wel of niet mee eens is.

Volledig mee eens	1	2	3	4	5	Volledig mee oneens

Het meetniveau is ordinaal. De Likert-schaal is bij uitstek geschikt voor het meten van attitudes en images en wordt ook wel de *vijfpuntenschaal* of de *unipolaire schaal* genoemd.

Osgood-schaal

Bij de Osgood-schaal zet men een aantal tegengestelde woordparen tegenover elkaar. Bijvoorbeeld mooi en lelijk. De respondent moet dan op een vijf- of zevenpuntenschaalverdeling zijn mening geven. Omdat de Osgood-schaal gebruikmaakt van tegengestelde woordparen, wordt het ook wel een *semantische differentiaal* genoemd.

'Ik vind hagelslag':

Lekker	1	2	3	4	5	Vies

Men mag met beide schalen geen gemiddelden berekenen, omdat het ordinale schalen zijn. In de praktijk gebeurt dit toch regelmatig.
Door de schaalitems onder elkaar te zetten en een lijn te trekken door de gekozen waarden, ontstaat er wel een beeld, een profiel, dat gebruikt kan worden om inzicht te krijgen in bepaalde zaken.