

Peilen

doe je zo



Centraal Bureau
voor de Statistiek



Peilen

doe je zo

Colofon

Uitgever

Centraal Bureau voor de Statistiek
Henri Faasdreef 312, 2492 JP Den Haag
www.cbs.nl

Prepress: Centraal Bureau voor de Statistiek, Grafimedia
Ontwerp: Edenspiekermann

Inlichtingen

Tel. 088 570 70 70, fax 070 337 59 94
Via contactformulier: www.cbs.nl/infoservice

© Centraal Bureau voor de Statistiek, Den Haag/Heerlen, 2014.
Verveelvoudigen is toegestaan, mits het CBS als bron wordt vermeld.

Inhoud

1. Wat zijn peilingen? 5

- 1.1 Surveys, enquêtes, peilingen, polls ... 6
- 1.2 Een peiling in vogelvlucht 9

2. Een peiling opzetten 12

- 2.1 De doelpopulatie 13
- 2.2 De variabelen 14
- 2.3 Populatiekenmerken 15

3. De vragenlijst 18

- 3.1 De formulering van de vraag 19
- 3.2 Soorten vragen 22
- 3.3 De volgorde van de vragen 26
- 3.4 Testen van vragenlijsten 28

4. Het verzamelen van gegevens 31

- 4.1 Traditionele gegevensverzameling 32
- 4.2 Computergestuurd interviewen 34
- 4.3 Peilingen via het internet 36

5. Het trekken van een steekproef 39

- 5.1 Het steekproefkader 40
- 5.2 Hoe het niet moet 41
- 5.3 De steekproef 44

6. Schatten van populatiekenmerken 49

- 6.1 Schatters en schattingen 50
- 6.2 Het schatten van een percentage 51
- 6.3 De omvang van de steekproef 55

7. Het akelige probleem van de non-respons 60

- 7.1 Non-respons in peilingen 61
- 7.2 Analyse van de non-respons 64
- 7.3 Wat te doen aan non-respons? 66

8. Een checklist voor peilingen 69

- 8.1 Waarom een checklist? 70
- 8.2 De checklist 70
- 8.3 Een voorbeeld: Social Media Stress 72
- 8.4 Tips voor je eigen peiling 74

9. Bijlagen 75

- 9.I Aselecte getallen **76**
- 9.II Het luisteronderzoek van Hazenbeek **78**
- 9.III Het luisteronderzoek (met non-respons) **83**

Begrippenlijst **88**

Gebruikte bronnen **90**

Gebruikte illustraties **91**

Medewerkers **92**

1.

Wat zijn

peilingen?

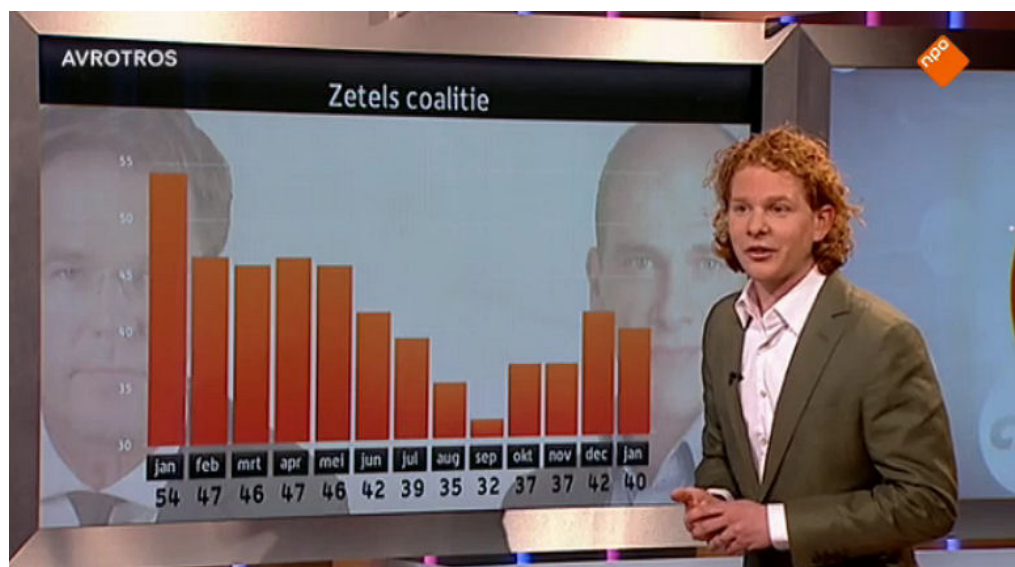
1.1 Surveys, enquêtes, peilingen, polls ...

Voel je je wel eens onveilig op straat? Heb je op dit moment een baantje om wat geld te verdienen? Heb je vanmorgen ontbeten? Welke sport beoefen je? Heb je een smartphone? Zit je op Facebook? Dit zijn allemaal voorbeelden van vragen die je voorgelegd zou kunnen krijgen als je meedoet aan een peiling.

Wat is een peiling? Een *peiling* is een vorm van onderzoek waarbij je gegevens verzamelt door het stellen van vragen aan een groep mensen. Dat doe je omdat je meer te weten wilt komen over het gedrag of de mening van een grote groep mensen. Bij een groep kan het bijvoorbeeld gaan om alle stemgerechtigden in een land, alle leerlingen op een school, of alle leraren in het voortgezet onderwijs. We noemen die groep ook wel de *doelpopulatie*.

Er zijn verschillende namen voor dit soort onderzoek in omloop. In wetenschappelijke kringen wordt wel gesproken over een *survey*. Een overheidsinstantie als het Centraal Bureau voor de Statistiek (CBS) noemt het ook wel een *enquête*. Een voorbeeld is de Enquête Beroepsbevolking (EBB). Daarin meet het CBS hoeveel mensen een baan hebben en hoeveel mensen werkloos zijn. Marktonderzoekbureaus hebben het vaak over een *peiling*. Ze doen veel politieke peilingen. Typische voorbeelden zijn de Politieke Barometer van Ipsos, de politieke peilingen van Maurice de Hond en het EenVandaag Opiniepanel van de TROS en de AVRO. In het Engels wordt een peiling een *poll* genoemd. Die term kom je soms ook wel in het Nederlands tegen.

1.1.1 Het EenVandaag Opiniepanel, met presentator Gijs Rademaker



Peilingen zijn niet echt anders dan surveys of enquêtes. Toch zijn er wel wat praktische verschillen. Peilingen zijn vaak klein en snel. Er doen hooguit een paar duizend mensen aan mee en het aantal vragen is beperkt. Peilingen zijn vooral bedoeld om snel een idee te krijgen van de publieke opinie over een actueel onderwerp. Peilingen volgen de actualiteit op de voet. Surveys en enquêtes zijn vaak groter en minder snel. Meer mensen krijgen een uitnodiging om aan het onderzoek deel te nemen. De vragenlijsten zijn groter en soms ook ingewikkelder. De onderzoekers zullen zich inspannen om toch vooral correcte uitspraken te doen. Dat kan meer tijd kosten. Daarom kan het wat langer duren voordat de uitkomsten beschikbaar komen.

Wij zullen het vooral hebben over peilingen, maar daarbij moet je dus bedenken dat alles wat we hier bespreken, ook van toepassing is op surveys en enquêtes.

Het voorleggen van een reeks vragen aan elk lid van de doelpopulatie kan een hele klus zijn, vooral als de groep groot is. Stel maar eens dat je een peiling wilt uitvoeren onder alle jongeren in Nederland in de leeftijd van 12 tot en met 18 jaar. Dat waren er 1 403 980 op 1 januari 2014. Het is een heel karwei om al die bijna anderhalf miljoen jongeren de vragen te laten beantwoorden. Om de werkzaamheden en de kosten binnen de perken houden, wordt meestal slechts een *steekproef* uit de doelpopulatie onderzocht. De vraag is dan wat je wel en niet kunt doen met de gegevens die je in die steekproef hebt verzameld. Kun je daarmee iets zinnigs zeggen over de doelpopulatie als geheel? Ja, dat kan, maar daar zitten wel een paar haken en ogen aan. De belangrijkste voorwaarde is dat die steekproef op een goede manier moet zijn getrokken. We zullen in deze lesmodule uitleggen wat een goede steekproef is en hoe je daaruit de juiste conclusies kunnen trekken.

Een goede peiling is niet alleen een kwestie van een goede steekproef trekken. Er komt nog wel wat meer bij kijken. Wat je met een peiling probeert te doen is het meten van allerlei zaken bij de deelnemers aan het onderzoek. Als je het gewicht van iemand wilt weten, dan zou je hem op een weegschaal kunnen zetten. En wil je weten hoe hoog de bloeddruk van iemand is, dan kun je die meten met een bloeddrukmeter. Voor het meten van meningen en gedragingen van mensen hebben we helaas geen instrumenten. Je moet het daarom doen met het stellen van vragen. Daarbij kan veel misgaan waardoor het gegeven antwoord niet het goede antwoord hoeft te zijn. En als je verkeerde antwoorden krijgt op je vragen, dan trek je vervolgens ook verkeerde conclusies uit je onderzoek. Daarom is het maken van een goede vragenlijst een zeer belangrijk onderdeel bij het uitvoeren van een peiling. Ook dat komt in deze lesmodule aan de orde.

Met een peiling kun je op een wetenschappelijk verantwoorde manier betrouwbare informatie verzamelen over een grote groep mensen, ook al hebben lang niet alle leden van die groep de vragenlijst ingevuld. Bij een goede peiling komt echter nog wel het nodige kijken. Je moet je aan allerlei richtlijnen houden. En die kunnen het onderzoek best wel eens wat ingewikkelder maken. Maar dat is de prijs die je betaalt voor een deugdelijke peiling. Het is wel eens verleidelijk om je niet aan die richtlijnen te houden. Ook dan krijg je uitkomsten, maar het is de vraag wat die waard zijn. Er wordt tegenwoordig heel wat onderzocht. Er zijn goede peilingen en er zijn slechte peilingen. Het kaf is soms moeilijk van het koren te scheiden. We hopen dat deze lesmodule je daarbij een handje kan helpen.

Er wordt steeds meer gepeild in Nederland. Dat kun je heel goed zien op het internet. Op veel websites zie je van die kleine 'polletjes' staan. Zie figuur 1.1.2 voor een voorbeeld.

1.1.2 Een poll op de website www.girlscene.nl



Uitroep . . .

Vier jij Valentijnsdag?

☐ Ja, ik heb wel iets leuks gepland!

☐ Geen idee, misschien wel

☐ Had ik maar een lover...

☐ Nee, ik heb een hekel aan Valentijnsdag

STEM [Check resultaten](#)

Vaak is er maar één vraag. De steekproef zit meestal niet goed in elkaar en ook worden de vragen niet goed gesteld. Daarom kun je niet zoveel met de uitkomsten van zo'n peiling. We zullen in deze lesmodule uitleggen wat er mis mee is.

Ook al zet je een peiling heel netjes op, en houd je je aan alle richtlijnen, toch kan er in de praktijk van alles misgaan. Eén van die problemen is non-respons. Dat is het verschijnsel dat de vragenlijst niet wordt ingevuld door mensen die in de steekproef zijn getrokken. Non-respons kan allerlei oorzaken hebben, zoals *geen contact* (je kunt iemand niet te pakken krijgen omdat hij voor langere tijd in het buitenland zit), *weigering* (iemand wil gewoon niet meedoen aan de peiling) of *niet in staat* (bijvoorbeeld omdat iemand geen Nederlands spreekt). Door non-respons kunnen specifieke groepen mensen ontbreken in de peiling. Daardoor kun je verkeerde conclusies trekken. We zullen ingaan op het probleem van de non-respons.

In deze lesmodule leggen we uit hoe je een goede peiling moet maken. Met 'goed' bedoelen we dat een peiling tegelijk betrouwbaar en valide moet zijn. *Betrouwbaar* betekent dat als je de peiling steeds weer zou herhalen (onder dezelfde omstandigheden), je ook steeds dezelfde uitkomsten moet krijgen. De peiling moet dus een *stabiel* meetinstrument zijn. *Valide* betekent dat de peiling werkelijk meet wat je wilt meten en niet iets anders. De uitkomsten moeten *geldig* zijn. Ze mogen je niet op het verkeerde been zetten.

Deze lesmodule is een handig hulpmiddel als je zelf een keer een peiling wilt opzetten. Je zult dan merken dat het lang niet altijd eenvoudig is om je te houden aan alle regels die we in deze lesmodule beschrijven. Helaas zijn al die regels nodig om conclusies uit je peiling te trekken die je kunt vertrouwen. Maar laat je niet ontmoedigen door al die regels. Helemaal aan het einde, in paragraaf 8.4 geven we nog wat tips voor je eigen onderzoek.

Je kunt deze lesmodule ook goed gebruiken als wilt controleren of een peiling van iemand anders wel goed in elkaar zit. Om je nog wat meer te helpen, hebben we een 'checklist voor peilingen' gemaakt. Die zit achteraan in deze lesmodule. Door het stellen en beantwoorden van de negen vragen in deze checklist kun je een aardig idee krijgen van de kwaliteit van een peiling. En misschien moet dan je conclusie wel zijn dat het beter is om maar geen aandacht te besteden aan de peiling.

De richtlijnen voor het uitvoeren van goede peilingen zijn gebaseerd op de uitkomsten van wetenschappelijk onderzoek. We noemen dit onderzoeksterrein ook wel *survey-methodologie*. Er zijn hierover dikke boeken volgeschreven met wiskundige formules. In deze handleiding houden we het simpel en beperken we ons tot de basisprincipes. Als je wat dieper in de stof wilt duiken, dan kunnen we je verwijzen naar Bethlehem (2013) (zie p.90). En voor alle statistisch-wiskundige details kun je bijvoorbeeld een kijkje nemen in het Engelstalige boek van Bethlehem (2009).

1.2 Een peiling in vogelvlucht

Bij het opzetten en uitvoeren van een peiling kun je een aantal stappen onderscheiden. Deze lesmodule neemt je mee door deze stappen. Bij elke stap leggen we uit waar je op moet letten en waar het fout kan gaan. Het zal duidelijk zijn dat deze aanpak je ook helpt om peilingen van anderen te beoordelen. In de rest van deze paragraaf lopen we alvast in vogelvlucht door deze stappen heen.

Het begint allemaal met de opzet van je onderzoek. Daarin moet je precies vastleggen welke groep mensen je gaat onderzoeken, wat je bij die groep gaat onderzoeken en hoe je het gaat onderzoeken. Daarna moet je de gegevens verzamelen. Iedereen in de steekproef moet de vragenlijsten invullen. We noemen dat wel het 'veldwerk'. Als je dan alle gegevens binnen hebt, moet je die gaan analyseren. Daaruit trek je dan conclusies over de doelpopulatie. En tenslotte schrijf je een onderzoeksverslag waarin je precies beschrijft hoe je peiling in elkaar zat en wat de uitkomsten van je peiling zijn.

Stap 1: Kies de doelpopulatie

De *doelpopulatie* is de hele groep van mensen die je wilt onderzoeken en waarover je uitspraken wilt doen. Het is dus ook de groep mensen waaruit je de steekproef trekt.

Stap 2: Kies de variabelen

De *variabelen* zijn de eigenschappen van de mensen in de doelpopulatie die je wilt meten. Die eigenschappen meet je door er vragen over te stellen. Als je bijvoorbeeld wilt weten hoeveel mensen Facebook gebruiken, dan vraag je aan alle personen in de steekproef of ze op Facebook zitten. De variabele is dan Facebookgebruik.

Stap 3: Kies de populatiekenmerken die je wilt schatten

Je gebruikt de verzamelde gegevens om meer te weten te komen over de doelpopulatie. Het komt erop neer dat je de doelpopulatie probeert te beschrijven in een aantal kerncijfers. Dergelijke cijfers noemen we *populatiekenmerken*. Het zijn grootheden die je kunt uitrekenen als je de waarden van een variabele kent voor iedereen in de

doelpopulatie. Dus als je de variabele Facebookgebruik meet, dan kun je uitrekenen hoeveel mensen in totaal in de populatie Facebook hebben. Andere voorbeelden van populatiekenmerken zijn: het *percentage* jongeren met een smartphone, het *gemiddelde* aantal uren per week dat ouderen surfen op het internet en het totaal aantal fietsen dat in het afgelopen jaar in huishoudens is gestolen.

Stap 4: Maak een vragenlijst

Je krijgt je gegevens door het stellen van vragen aan mensen. Je moet die vragen aan iedereen in de steekproef op precies dezelfde manier stellen. Daarvoor gebruik je een *vragenlijst*. Het maken van een goede vragenlijst is een heel werk. Daarbij kan veel mis gaan. Dan krijg je de verkeerde antwoorden en dan trek je ook de verkeerde conclusies uit je onderzoek.

Stap 5: De methode van gegevensverzameling

Alle mensen in de steekproef moeten de vragenlijst invullen. Dit kun je op allerlei manieren laten doen, en elke manier heeft zijn voor- en nadelen als je kijkt naar de kosten en kwaliteit. Je hebt de keuze uit een *mondelinge peiling* (bij de mensen thuis), een *telefonische peiling*, een *schriftelijke peiling* (per post) en een *online-peiling* (via het internet).

Stap 6: Kies een steekproefkader

Als je weet hoe je de gegevens gaat verzamelen, dan moet je nadenken over de manier waarop je de steekproef gaat trekken. Je hebt daarvoor een lijst nodig waarin de hele doelpopulatie staat opgesomd. We noemen dat een *steekproefkader*. Voorbeelden zijn een papieren kaartenbak (de ledenadministratie van een tennisclub), een elektronische databank (het bevolkingsregister van een gemeente), een lijst met telefoonnummers (het telefoonboek) en een lijst met e-mailadressen (van de leerlingen van een school).

Stap 7: De steekproef trekken

Het is een fundamenteel principe van peilingen dat je de steekproef moet *loten* uit het steekproefkader. En je moet bedenken hoe groot die steekproef moet zijn. Daarbij geldt dat de uitkomsten van je peiling nauwkeuriger zijn als je steekproef groter is.

Stap 8: Verzamel de gegevens

Heb je de methode van gegevensverzameling gekozen en heb je de steekproef getrokken, dan moet je de vragenlijsten ingevuld zien te krijgen. Bij een mondelinge peiling gaan de interviewers op pad. Bij een telefonische peiling gaan de interviewers bellen. Bij een schriftelijk peiling verstuur je de vragenlijsten per post naar de geselecteerde personen. En bij een online-peiling zet je de vragenlijst op het internet.

Stap 9: Probeer non-respons te voorkomen

Een vervelend verschijnsel bij peilingen is het optreden van *non-respons*. Er kunnen allerlei redenen zijn waarom mensen de vragenlijsten niet invullen: ze zijn niet thuis, ze weigeren hun medewerking, of ze zijn niet in staat de vragen te beantwoorden (bijvoorbeeld wegens taalproblemen). Vaak doen bepaalde groepen niet mee, bijvoorbeeld bejaarde mensen bij een online-peiling. Dat kan leiden tot een vertekend beeld, en dus tot verkeerde conclusies. Daarom moet je proberen alle personen in de steekproef de vragenlijst te laten invullen.

Stap 10: Schat de populatiekenmerken

Omdat je werkt met een steekproef, kun je de populatiekenmerken niet precies uitrekenen. Je kunt ze wel schatten. Zo is het percentage in de steekproef meestal wel een goede schatting voor het percentage in de populatie. Die schattingen kunnen dus een afwijking hebben. De *onzekerheidsmarge* geeft aan hoe groot die afwijking maximaal kan zijn. Die onzekerheidsmarge moet je uitrekenen. Om aan te geven hoe groot de afwijking kan zijn, moet je niet alleen schattingen publiceren, maar ook de onzekerheidsmarges.

Stap 11. Publiceer de uitkomsten

In de laatste stap van de peiling maak je een verslag van je peiling. Daarin vertel je niet alleen wat je allemaal voor interessante zaken hebt ontdekt in de verzamelde gegevens, maar ook leg je precies uit hoe je de peiling hebt opgezet en uitgevoerd (zodat anderen kunnen zien of je je aan de regels hebt gehouden).

2.

Een peiling opzetten

De allereerste stap in het opzetten van een peiling is nadenken over wat je precies wilt weten, van wie je dat wilt weten en wat je daarvoor moet meten. Je moet daarom antwoord geven op de volgende vragen:

- Wie ga ik onderzoeken? Met het antwoord op deze vraag leg je de doelpopulatie vast. Uit die doelpopulatie ga je de steekproef trekken. En de uitkomsten van je peiling hebben op die doelpopulatie dan ook betrekking.
- Wat ga ik meten? Met het antwoord op deze vraag kies je de variabelen die je gaat meten. Die variabelen zullen in de vorm van vragen terugkomen in de vragenlijst.
- Wat wil ik weten? Met het antwoord op deze vraag kies je de populatiekenmerken die je wilt gaan schatten.

In dit hoofdstuk leggen we uit hoe je de bovenstaande drie vragen moet beantwoorden.

2.1 De doelpopulatie

Welke groep mensen ga je onderzoeken in je peiling? Gaat het om iedereen die woont in Nederland? Of gaat het misschien alleen om inwoners van een stad? Gaat het om alle scholieren in een bepaalde leeftijdscategorie? Of gaat het misschien alleen om de leerlingen van een bepaalde school? De groep die je onderzoekt, noemen we de *doelpopulatie*. Het is de hele groep mensen waarover je in je peiling uitspraken wilt doen. Het is dus ook de groep waaruit je de steekproef trekt. En het is dus de groep waarop de conclusies van je peiling betrekking hebben.

Het is belangrijk om je doelpopulatie goed te beschrijven. Voor alle mensen die je 'in het veld' kunt tegenkomen, moet je kunnen vaststellen of ze wel of niet tot de doelpopulatie behoren. Als dat niet duidelijk is, dan kan dit bij het veldwerk leiden tot vergissingen en problemen. Het kan gebeuren dat je ten onrechte mensen ondervraagt die niet in de doelpopulatie thuishoren, of je slaat mensen over die er juist wel in thuishoren. Het is denkbaar dat je daardoor verkeerde conclusies trekt uit je peiling.

2.1.1 Studio van een lokale omroep



Als voorbeeld kijken we naar een luisteronderzoek. Nederland kent bijna 300 lokale omroepen. Die omroepen doen af en toe een luisteronderzoek. Met zo'n peiling kunnen ze antwoord krijgen op vragen als hoeveel mensen de omroep kennen, hoeveel mensen er

naar de radioprogramma's luisteren, naar welke programma's ze luisteren en wat ze van de omroep vinden.

Wat is de doelpopulatie bij een luisteronderzoek? De meeste lokale omroepen zenden uit voor één gemeente. Dan zal de doelpopulatie in principe bestaan uit de inwoners van die gemeente. Dat moet je nog wel wat nauwkeuriger vastleggen. Gaat het om alle mensen die permanent wonen in de gemeente? Of ook om mensen die er tijdelijk wonen? En wat te zeggen van mensen die niet in de gemeente wonen, maar er wel werken (en misschien op hun werk naar de omroep luisteren)? Verder zou je nog kunnen denken aan een minimum leeftijd. Immers, baby's zullen niet (bewust) naar de radio luisteren. Het is bovendien lastig om een vragenlijst voor te leggen aan heel jonge kinderen. Bij veel luisteronderzoek gaat het daarom alleen om mensen vanaf een zekere leeftijd. Bedenk daarbij wel dat veel lokale omroepen ook programma's voor kinderen uitzenden. Als de omroep meer wil weten over het luisteren naar deze programma's, dan zal je die leeftijdsgroep toch ook in het onderzoek moeten meenemen. Het is niet ongebruikelijk bij dit soort peilingen dat de doelpopulatie bestaat uit alle personen die permanent in de gemeente wonen en die minimaal 13 jaar oud zijn.

2.2 De variabelen

Van de mensen in de steekproef ga je allerlei kenmerken meten. Dat doe je door het stellen van vragen. Zo krijg je de gegevens die je nodig hebt voor je analyse en voor het uiteindelijk trekken van conclusies over de toestand van de doelpopulatie. Je kunt allerlei eigenschappen van mensen meten. Zo'n eigenschap noemen we een *variabele*. Die naam is zo gekozen omdat de eigenschap voor elke persoon anders kan zijn en dus kan *variëren*.

Voorbeelden van eigenschappen zijn de lengte van iemand, of iemand 's morgens wel of niet heeft ontbeten, de kleur van iemands ogen, het aantal glazen bier dat iemand het afgelopen weekend heeft gedronken en de mening van iemand over de kwaliteit van het voetbal in Nederland.

We maken onderscheid tussen kwalitatieve variabelen en kwantitatieve variabelen. Een *kwalitatieve variabele* wordt ook wel een *categorische variabele* genoemd. Hij verdeelt de doelpopulatie in groepen (categorieën). We kunnen niet rekenen met de waarden van zo'n variabele. Het zijn slechts etiketten voor de verschillende groepen. We kunnen alleen vaststellen of personen wel of niet tot dezelfde groep behoren. Voorbeelden van kwalitatieve variabelen zijn de kleur van iemands haar, de geloofsovertuiging van een persoon (Rooms-katholiek, Protestant, Islamitisch, Hindoe, enz.) en het vervoersmiddel waarmee een leerling naar school gaat (lopend, fiets, scooter, motor, bus, tram, lightrail, metro, enz.). Het enige wat we met een kwalitatieve variabele kunnen doen, is het tellen van de aantallen in elke groep en die aantallen eventueel omzetten in percentages.

Een *kwantitatieve variabele* meet de omvang, het gewicht, de waarde of de duur van iets. Met die waarden kunnen we wel op zinvolle wijze rekenen. Er is ook altijd sprake van een meeteenheid. Voorbeelden van kwantitatieve variabelen zijn het gewicht van iemand (in kilogram), de leeftijd van iemand (in jaren) of de hoeveelheid zakgeld van een scholier (in euro's). Met de waarden van kwantitatieve variabelen kunnen we berekeningen uitvoeren

zoals het bepalen van de totale of de gemiddelde waarde in de steekproef. Een speciaal type variabele is ook nog de *indicatorvariabele*. Die meet de aan- of afwezigheid van een bepaalde eigenschap. Heeft iemand die eigenschap, dan geven we dit aan met de waarde 1. Heeft iemand die eigenschap niet, dan geven we dit aan met de waarde 0. Voorbeelden zijn het wel of niet hebben van internet, het wel of niet hebben van een betaalde baan, en het wel of niet hebben van een smartphone. De indicatorvariabele is een soort kwalitatieve variabele, want hij verdeelt de doelpopulatie in twee groepen: personen die de eigenschap wel hebben en personen die de eigenschap niet hebben. Maar de indicatorvariabele is ook een soort kwantitatieve variabele, want we kunnen rekenen met een indicatorvariabele. We kunnen bijvoorbeeld uitrekenen welk percentage mensen in de steekproef een smartphone heeft en welk percentage niet.

In een peiling onderscheiden we doelvariabelen en hulpvariabelen. De *doelvariabelen* zijn de belangrijke variabelen die meten waar we op uit zijn in onze peiling. Ze meten allerlei aspecten van het verschijnsel dat we onderzoeken. In een vakantiepeiling zouden de doelvariabelen bijvoorbeeld kunnen meten of je op vakantie gaat, waar je naar toe gaat, hoe lang je gaat, met welk vervoermiddel je op vakantie gaat en hoeveel geld je uitgeeft.

De *doelvariabelen* meten dus allerlei aspecten van het verschijnsel dat we onderzoeken. Vaak meten we echter nog meer variabelen die zo op het oog weinig of niets met het doel van je onderzoek te maken hebben. Deze variabelen noemen we *hulpvariabelen*. Het gaat hier meestal om achtergrondkenmerken van de personen, zoals geslacht, leeftijd, burgerlijke staat en provincie.

Hulpvariabelen bieden je de mogelijkheid om verschillende groepen met elkaar te vergelijken. Het zou kunnen zijn dat mannen zich anders gedragen dan vrouwen. Of dat ouderen anders scoren op een variabele dan jongeren. Of dat personen in stedelijke gebieden een andere mening hebben dan personen op het platteland. Kortom, de analyse van je gegevens zal meer informatie opleveren als je er hulpvariabelen bij gebruikt.

2.3 Populatiekenmerken

In je peiling verzamel je gegevens bij een steekproef van personen. Met die gegevens wil je uitspraken doen over de doelpopulatie. Het komt erop neer dat je de toestand in de populatie probeert te beschrijven in een aantal kerncijfers. Zulke cijfers noemen we *populatiekenmerken*. Het zijn grootheden die je exact zou kunnen uitrekenen als je alle waarden van de doelvariabelen in de doelpopulatie kende.

Voor een kwantitatieve doelvariabele kun je twee populatiekenmerken definiëren: het *totaal* en het *gemiddelde* van de waarden van de doelvariabele. Stel dat de doelpopulatie bestaat uit jongeren in de leeftijd van 12 tot en met 18 jaar en dat de doelvariabele het aantal berichten is dat iemand op een bepaalde dag heeft verstuurd via zijn of haar Facebook-account. Het totaal van de doelvariabele is dan het totaal aantal berichten dat door jongeren op een dag via Facebook is verstuurd. Als je dezelfde vraag voor een andere dag stelt, kun je bijvoorbeeld zien of het aantal Facebook-berichten is toegenomen of afgenomen.

Stel weer dat de doelpopulatie bestaat uit jongeren in de leeftijd van 12 tot en met 18 jaar en dat de doelvariabele het aantal uren is dat een jongere op internet heeft gezeten op een bepaalde dag. Het gemiddelde van de doelvariabele is dan het gemiddeld aantal uren dat jongeren op een dag op internet zitten. Je zou dit populatiekenmerk ook voor jongens en meisjes apart kunnen uitrekenen. Dan kun je bijvoorbeeld zien of jongens meer op internet zitten dan meisjes, of dat het juist andersom is.

Met de waarden van een kwalitatieve variabele kun je niet echt rekenen. Wat je wel kunt doen is het percentage personen in elke groep (categorie) uitrekenen. Je telt daarvoor eerst hoeveel mensen er in een groep zitten. Dat aantal deel je door het totaal aantal mensen in de doelpopulatie. Het resultaat vermenigvuldig je dan weer met 100.

Stel dat de doelpopulatie bestaat uit alle Nederlanders op 1 januari 2014. Dat waren op die dag in totaal 16 827 775 personen. We gebruiken de kwalitatieve variabele geslacht met de twee categorieën: man en vrouw. Er bleken 8 333 654 mannen en 8 494 121 vrouwen te zijn. Als we het aantal mannen delen door het totaal aantal personen, is de uitkomst 0,495. En vermenigvuldigen met 100 geeft dat een percentage van 49,5 procent. Op dezelfde manier kun je uitrekenen dat het percentage vrouwen gelijk is aan 50,5 procent. Om te controleren dat je geen fouten in de berekeningen hebt gemaakt, kun alle percentages optellen. Dan moet je precies op 100 procent uitkomen. Dat is hier het geval: $49,5 + 50,5 = 100$.

Voor een indicatorvariabele kun je twee populatiekenmerken berekenen: het totaal aantal personen in de doelpopulatie met een specifieke eigenschap en het percentage objecten met die eigenschap.

Stel dat de doelpopulatie gelijk is aan alle 2300 leerlingen van een school en de doelvariabele is het wel of niet hebben van een smartphone. Stel eens dat 1357 leerlingen een smartphone hebben, dan is het populatietotaal dus gelijk aan 1357. Voor het populatiepercentage delen we 1357 door 2300. Dit geeft als uitkomst 0,59. Vermenigvuldigen met 100 levert dan een percentage van 59 procent.

Vraag 2.1

Een gemeentebestuur wil weten of de inwoners tevreden zijn met het enige zwembad in de gemeente. Wat vinden de inwoners goed aan het zwembad en wat zou er moeten worden verbeterd? Het gemeentebestuur besluit een peiling te doen. De vraag is nu wat de doelpopulatie van deze peiling moet zijn? Hieronder staan een aantal mogelijkheden. Maar misschien bedenk je nog wel een betere doelpopulatie.

- a. Alle inwoners van de gemeente.
- b. Alle inwoners vanaf een zekere leeftijd, bijvoorbeeld vanaf 12 jaar.
- c. Alle inwoners die de afgelopen jaar in het zwembad zijn geweest.
- d. Alle inwoners die ooit wel eens een keer in het zwembad zijn geweest.
- e. Alle inwoners met een jaarkaart voor het zwembad.
- f. Alle mensen die ooit wel eens een keer in het zwembad zijn geweest.
- g. Alle mensen met een jaarkaart voor het zwembad.

Leg voor elke mogelijke doelpopulatie uit waarom je er hier wel of niet voor zou kiezen.

3.

De vragenlijst

In je peiling verzamel je gegevens door het stellen van vragen aan de personen in de steekproef. Dat moet je heel zorgvuldig doen. De antwoorden op de vragen moeten een juist beeld geven van de respondenten. Als twee mensen hetzelfde over iets denken, moeten ze hetzelfde antwoord geven. En als ze er verschillend over denken, moet dat ook aan de antwoorden te zien zijn. Daarom is er behoefte aan een goed meetinstrument. Dat meetinstrument is de vragenlijst.

Bij het maken van een vragenlijst moet je de tekst van elke vraag zorgvuldig formuleren. En ook moet je goed bedenken wat voor soort antwoord je wilt hebben. Doe je dat niet, dan stel je verkeerde vragen en krijg je verkeerde antwoorden. Dat werkt door in het hele onderzoek. Uiteindelijk trek je ook verkeerde conclusies uit je peiling.

Er wordt wel eens gezegd dat het ontwerpen van vragenlijsten meer een kunst is dan een kunde. Het is inderdaad lastig om harde regels te geven voor het maken van goede vragenlijsten. Toch heeft wereldwijde en jarenlange ervaring wel wat vuistregels opgeleverd. We zullen een aantal van die vuistregels in dit hoofdstuk doornemen. In het eerste deel van dit hoofdstuk gaat het vooral om de formulering van de vragen. Het tweede deel gaat over de vragenlijst als geheel, zoals de volgorde van de vragen en de voorwaarden waaronder je ze moet stellen. In het derde deel gaan we nog even in op het testen van vragenlijsten.

3.1 De formulering van de vraag

Een van de belangrijkste aspecten van de vragenlijst is de formulering van de tekst van de vraag. Iedereen moet de vragen kunnen begrijpen en kunnen beantwoorden. We geven een aantal regels waar je je aan moet houden en we leggen uit wat er misgaat als je je er niet aan houdt.

Stel begrijpelijke vragen

De vragen moeten begrijpelijk zijn voor alle personen in de steekproef. Soms worden vragenlijsten gemaakt door deskundigen op een bepaald gebied. Die weten wel veel van het onderwerp van de peiling, maar het gevaar bestaat dat ze teveel moeilijke woorden en teveel vakjargon gebruiken.

Een voorbeeld van een begrip dat voor wetenschappers relatief eenvoudig is te begrijpen, maar toch geen alledaagse kost is voor de gemiddelde Nederlander, is de snelheid waarmee veranderingen plaatsvinden. De volgende vraag is een voorbeeld hiervan:

Vindt u dat de prijzen van levensmiddelen op het ogenblik sneller stijgen, even snel stijgen of langzamer stijgen dan een jaar geleden?

Lang niet elke respondent heeft door wat je hier vraagt. Het gaat niet om het stijgen of dalen van de prijzen, maar om de snelheid waarmee de prijzen stijgen. Het is een lastige

vraag en daarom ligt het gevaar op de loer dat hij verkeerd wordt begrepen en dus verkeerd wordt beantwoord.

Om de tekst van een vraag voor iedereen begrijpelijk te houden, moet je die tekst niet te lang maken en simpele woorden gebruiken. Probeer de tekst op te schrijven in 'gesproken taal' en niet in 'geschreven taal'. Stel eens dat je de mensen in je peiling de volgende vraag voorlegt:

Bent u tevreden over de recreatieve voorzieningen in uw woonplaats?

De kans is dan heel groot dat de respondent niet precies weet wat je bedoelt met 'recreatieve voorzieningen'. Als het mee zit, dan zal hij denken aan een zwembad, maar of hij ook zal denken aan zaken als de bibliotheek en het stadspark, valt te betwijfelen. In deze situatie is het beter om precies uit te leggen wat je bedoelt.

Het algemene advies is om de tekst van de vraag kort te houden. Er zijn echter situaties denkbaar waarin de tekst wel wat langer zou kunnen zijn. Het idee daarbij is om bij een mondeling of telefonisch interview de respondent wat meer tijd te geven om na te denken over het antwoord op de vraag.

Stel ondubbelzinnige vragen

Zelfs als een vraag in simpele taal is gesteld, dan betekent dat nog niet dat daarmee alle problemen zijn opgelost. De vraag moet ook ondubbelzinnig zijn. Hij moet maar voor één uitleg vatbaar zijn. Een vraag als

Wanneer bent u van school afgegaan?

is zonder nadere toelichting voor allerlei uitleg vatbaar. Antwoorden als 'na mijn eindexamen', 'in 1974', 'toen ik ging trouwen', 'toen ik 18 was' en 'toen ik zwanger was' zouden allemaal kunnen. Als je, bijvoorbeeld, een datum wilt weten, vraag dan ook naar die datum.

Over het algemeen is het voor respondenten makkelijker om op feitelijke vragen antwoord te geven dan op opinie vragen. Maar zelfs feitelijke vragen kunnen onduidelijk en verwarrend zijn. Als je aan een werkloze vraagt of hij op zoek is naar werk, dan kan dat van alles betekenen. Op zoek zijn naar werk kan neerkomen op zelf initiatief nemen en allerlei bedrijven aanschrijven, maar het kan ook betekenen dat iemand de hele dag alleen maar uit het raam zit te staren. Het is zelfs denkbaar dat de vraag suggestief is; de werkloze zoekt eigenlijk helemaal niet naar werk, maar hij antwoordt dat hij dat wel doet omdat dit sociaalwenselijk gedrag is.

Bij het opstellen van een vraag moet je beseffen dat een respondent allerlei gebruikte termen op een eigen manier kan interpreteren. Die interpretatie kan voor iedereen anders zijn. Zo kan een simpele term als 'gezin' door iemand in enge zin worden begrepen als de partners met hun kinderen, maar het kan ook heel breed worden gezien als alle personen in het huishouden, dus inclusief inwonende opa's en oma's. En als je vraagt hoeveel kamers er in het huis zijn, wat moet iemand dan allemaal meetellen? De bijkeuken? De badkamer? De hal?

Stel geen suggestieve vragen

De formulering van de vraag moet neutraal zijn. De respondent moet een kans krijgen zijn eigen mening te geven, zonder dat je probeert hem een bepaalde kant op te sturen. Je moet dus suggestieve (sturende) vragen vermijden. Met een vraag als

U hebt toch zeker ook wel een iPad?

wek je de indruk dat het eigenlijk abnormaal is om geen iPad te hebben, en daarom zal een respondent aarzelen dat toe te geven. Woorden als 'ook', 'toch' en 'wel' moet je daarom zoveel mogelijk vermijden. Ook is het niet goed om deskundigen aan te halen, omdat niet iedereen altijd even makkelijk een deskundige zal durven tegen te spreken.

Stel geen dubbele vragen

Een veel gemaakte fout is in een vraag meer dan één onderwerp aan de orde te stellen, terwijl er maar één antwoord kan worden gegeven. De vraag

Is het u bekend dat de lokale omroep een website heeft, en dat daarop een samenvatting is te vinden van al het lokale nieuws?

bevat in feite twee vragen. Als iemand dan antwoord met 'ja', waarop geeft hij dan antwoord? Op de eerste vraag (*Is het u bekend dat de lokale omroep een website heeft?*) of op de tweede vraag (*Is het u bekend dat de website van de lokale omroep een samenvatting is te vinden van al het lokale nieuws?*) of op beide vragen? Om dit soort problemen te voorkomen, moet je een vraag als deze splitsen in twee aparte vragen.

Vermijd (dubbele) ontkenningen

Vragen zijn moeilijker te begrijpen als ze een ontkenning bevatten, of nog erger, een dubbele ontkenning. Bij de vraag

Gaat u liever niet naar een café waar niet gerookt mag worden?

zult je even heel diep moeten nadenken wat het betekent als je 'nee' zou antwoorden. Meestal kun je een vraag met een of meer ontkenningen ombouwen tot een vraag zonder ontkenningen:

Gaat u liever naar een café waar gerookt mag worden?

Pas op met vragen over gevoelige onderwerpen

Ook al heb je een vraag duidelijk en begrijpelijk geformuleerd, dan nog kan het zo zijn dat je niet het juiste antwoord krijgt. Een reden kan zijn dat een respondent een *sociaalwenselijk* antwoord geeft. Als een onderwerp nogal gevoelig ligt (abortus, crimineel gedrag, gebruik van drugs, seksueel gedrag), dan zal de respondent eerder een algemeen

geaccepteerd antwoord geven dan zijn 'eigen' antwoord. Daarom is het twijfelachtig of de antwoorden op een vraag als

Bezoekt u wel eens porno-websites op het internet?

een goed beeld zullen geven van het surf-gedrag van de respondenten.

Pas op met vragen die een beroep doen op de herinnering

De mens onthoudt sommige gebeurtenissen beter dan andere gebeurtenissen. Daarmee moet je rekening houden als je een vraag stelt over gebeurtenissen die in het verleden hebben plaatsgevonden. Bij een vraag als

Hoeveel keer bent u de afgelopen drie maanden naar uw huisarts geweest?

zal een aantal bezoeken worden vergeten. Dit kan in sommige situaties oplopen tot wel 30 procent. Wat helpt is de periode korter maken waarover iets moet worden gezegd. In bovenstaande vraag zou je, bijvoorbeeld, 'drie maanden' kunnen vervangen door 'week'. Dat verhoogt wel de betrouwbaarheid van de antwoorden, maar daar staat tegenover dat je minder informatie krijgt (over een week in plaats van over drie maanden).

Vermijd hypothetische vragen

Nog moeilijker dan vragen over het verleden zijn *hypothetische vragen*. Als je een hypothetische vraag stelt, dan krijg je ook een hypothetisch antwoord. Als je mensen vraagt wat ze zouden hebben gedaan als zich in hun leven bepaalde gebeurtenissen wel of niet hadden voorgedaan, dan geef je een heel moeilijke opdracht. Je vraagt veel verbeeldingskracht van een respondent die op een dergelijke vraag een serieus antwoord wil geven. Die respondent zal flink wat tijd nodig hebben om alle aspecten van zo'n hypothetische situatie te overdenken. Hier is een voorbeeld van zo'n hypothetische vraag:

Als u minister-president was, wat zou u dan doen om de misdaad in Nederland beter te bestrijden?

Hypothetische vragen worden vaak gesteld in een poging iets meer te weten te komen over de houding en mening van de respondent over allerlei zaken. Er is echter zeer weinig bekend over wat zich afspeelt in de geest bij de beantwoording van zulke vragen. Het is dus maar zeer de vraag of de vraag meet wat je wilt meten.

3.2 Soorten vragen

Tot nog toe is alleen de tekst van de vraag aan de orde geweest. Het is echter ook belangrijk om na te denken over het soort antwoord dat de respondent moet geven op de vraag. We kunnen verschillende soorten vragen onderscheiden. We beperken ons hier tot open vragen, gesloten vragen en numerieke vragen.

Open vraag

Het meest voor de hand liggende soort vraag is de *open vraag*. Daarbij accepteert je elke willekeurige tekst als antwoord. Open vragen hebben het voordeel dat de respondent spontaan en in zijn eigen woorden kan reageren. Er zijn echter ook nadelen. Zo is het niet uitgesloten dat de respondent bepaalde antwoorden over het hoofd ziet. Het volgende voorbeeld illustreert dit probleem. In een peiling werd de volgende open vraag gesteld:

Welke weekbladen leest u?

Als je de vraag zo stelt, dan blijken veel mensen allerlei bladen over het hoofd te zien. Dat bleek in dit voorbeeld vooral het geval te zijn voor omroepbladen. Dit zijn toch echt ook weekbladen. Alleen veel mensen dachten er niet aan. Als je er een lijst van bladen bij vermeldt, en daar zitten ook de omroepbladen bij, dan blijken ineens veel meer mensen omroepbladen te lezen.

Het stellen van een open vraag kan ook leiden tot onduidelijke antwoorden. In een luisteronderzoek van een lokale omroep kwam de volgende vraag voor:

Wat vindt u het belangrijkste aspect van de lokale omroep?

Veel respondenten gaven als antwoord 'het lokale karakter' of iets dergelijks. Daarbij werd het niet duidelijk wat ze daar precies mee bedoelden. Bij verder vragen bleek dat het bij sommigen ging om de lokale nieuwsvoorziening, terwijl anderen vooral dachten aan de omroep als lokale vrijwilligersorganisatie.

Wat doe je met al die antwoorden op een open vraag? Analyse van dit soort teksten is erg lastig. Daarom is het beter om open vragen zoveel mogelijk te vermijden. Je moet alleen open vragen gebruiken als het echt niet anders kan. Een voorbeeld is de vraag naar het dorp of de stad waar de respondent zijn laatste zomervakantie heeft doorgebracht:

In welk dorp of welke stad heeft u uw laatste zomervakantie doorgebracht?

Je zou natuurlijk de respondenten kunnen laten kiezen uit een lijst van alle mogelijke vakantiebestemmingen. Die lijst zou wel eens erg lang kunnen worden. Het is vrijwel ondoenlijk om daarin de juiste bestemming te vinden. Er zijn misschien ook wel mensen die niet meer weten in welke plaats ze geweest zijn. Dan kunnen ze bij een open vraag in ieder geval de streek of het land invullen. Lastig is ook dat de er schrijffouten in de antwoorden kunnen voorkomen. Hopelijk levert dat geen verwarring op.

Gesloten vraag, één antwoord

Bij een *gesloten vraag* leg je aan de respondent een lijst met mogelijke antwoorden voor. Die moet dan één van de mogelijkheden in de lijst uitkiezen en aankruisen. Dit kan alleen als alle mogelijke antwoorden in de lijst staan, zodat elke respondent altijd een passend antwoord kan vinden. Toch moet je vermijden dat de respondent een erg lange lijst met

mogelijkheden krijgt voorgelegd. De keuze is moeilijker voor de respondent naarmate de lijst langer is:

Welk vervoermiddel gebruik je het meest bij verplaatsingen binnen je dorp of stad?

- ☐ Lopen
- ☐ Fiets
- ☐ Elektrische fiets
- ☐ Bromfiets
- ☐ Bromscooter
- ☐ Motor
- ☐ Auto
- ☐ Bus
- ☐ Tram
- ☐ Metro / lightrail
- ☐ Ander vervoermiddel

Bij mondelinge en telefonische peilingen leest de enquêteur alle mogelijke antwoorden op. De respondent kan die nooit allemaal onthouden. De eerste antwoorden is hij al weer snel vergeten en daarom kiest hij vaak een antwoord ergens achteraan in de lijst. Deze voorkeur van een antwoord achteraan in de lijst noemen we het *recency effect*. Bij schriftelijke peilingen en online-peilingen moet de respondent zelf de lijst met mogelijke antwoorden lezen. Dat gebeurt niet altijd even goed. De respondent verliest al gauw de aandacht. Dat leidt daarom juist tot een voorkeur voor een antwoord vooraan in de lijst. Dit noemen we het *primacy effect*. Om dit soort problemen te voorkomen, is het verstandig om een zo kort mogelijke lijst van mogelijke antwoorden aan te bieden.

Je kunt gesloten vragen goed gebruiken voor het peilen van de mening van personen. De respondenten kunnen dan kiezen uit een beperkt aantal antwoorden. Het is niet ongebruikelijk hiervoor een zogenaamde 5-puntsschaal te hanteren:

Hoe tevreden of ontevreden bent u over het huidige kabinet?

- ☐ Zeer tevreden
- ☐ Tevreden
- ☐ Niet tevreden en niet ontevreden
- ☐ Ontevreden
- ☐ Zeer ontevreden
- ☐ Weet niet

Merk op dat bij deze vraag nog een zesde optie 'Weet niet' is opgenomen. Die is uiteraard bedoeld voor mensen die echt helemaal geen mening hebben. Helaas wordt dit antwoord ook vaak gekozen door mensen die wel een mening hebben, maar die niet willen geven. Het is daarmee een soort 'nooduitgang'.

Het blijft lastig om te kiezen of je 'Weet niet' wel of niet als mogelijk antwoord moet opnemen in de lijst. Is 'Weet niet' aanwezig als een van de mogelijke antwoorden, dan zullen veel respondenten dit antwoord kiezen om te voorkomen dat ze moeten nadenken over een 'echt' antwoord. Weglaten van 'Weet niet' dwingt respondenten echter tot een mening die ze misschien niet hebben. Dan vluchten ze vaak naar een neutrale middencategorie ('niet tevreden en niet ontevreden').

Een mogelijke oplossing is de toevoeging van een *filtervraag*. Daarin vraag je de respondenten eerst of ze een mening hebben over een bepaalde kwestie. En alleen als ze een mening hebben, kom je met een vervolgvraag waarin je naar die mening vraagt.

Gesloten vraag, meer antwoorden

Meestal moet de respondent bij een gesloten vraag maar één antwoord uit de lijst kiezen. Soms komt het echter voor dat meer dan één antwoord in aanmerking komt. Dan kun je de respondenten de mogelijkheid bieden om meer antwoorden aan te kruisen:

Over welke aspecten van het verkeer in uw buurt bent u ontevreden?

(Meer antwoorden zijn mogelijk)

- ☐ De parkeergelegenheid
- ☐ De afstand tot de bushalte
- ☐ De busverbindingen
- ☐ De genomen verkeersmaatregelen
- ☐ De straatverlichting
- ☐ De veiligheid voor fietsers
- ☐ De veiligheid voor voetgangers
- ☐ Andere aspecten

Het kan heel goed zijn dat iemand ontevreden is over meer dan één aspect van het verkeer. Dan is het verstandiger meer dan één antwoord toe te staan. Merk ook op dat dit voorbeeld zich indekt tegen de situatie dat iemand met een antwoord komt dat niet in de lijst staat. Daarvoor is de categorie 'Andere aspecten' opgenomen.

Als dit soort vragen in een schriftelijke peiling of een online-peiling zit, dan kan er een *primacy effect* optreden. De respondent kruist vooraan in de lijst een paar antwoorden aan en houdt het dan voor gezien.

Numerieke vraag

Een ander type vraag dat veel voorkomt, is de *numerieke vraag*. Een dergelijke vraag gebruik je om een kwantitatieve variabele te meten. Je kunt daarbij denken aan vragen over geld (hoeveel geld verdient je met je bijbaantje?), tijdsduur (hoeveel uur heb je op het internet gezeten?), omvang (wat is je gewicht?) of aantallen (hoeveel glazen bier heb je gisteren gedronken?). Een veel voorkomende vraag in een peiling is:

Uit hoeveel personen bestaat uw huishouden? _ _

Met een aantal vakjes of lijntjes kun je aangegeven uit hoeveel cijfers het antwoord mag bestaan. Dit geeft wat extra aansporing tot het geven van een juist antwoord.

Bedenk wel dat bij veel numerieke vragen het voor de respondent niet altijd makkelijk is om het exacte antwoord te geven, omdat hij dat exacte antwoord niet weet en het veel werk is om het op te zoeken of te bepalen. Voorbeelden daarvan zijn vragen naar het netto jaarinkomen en het aantal uren en minuten dat iemand gisteren op het internet bezig was. Om dit soort problemen te vermijden, kun je overwegen om van de numerieke vraag een gesloten vraag te maken:

Hoe lang was je gisteren actief op het internet?

- ☐ Helemaal niet
- ☐ Minder dan 1 uur
- ☐ 1 tot 2 uur
- ☐ 2 tot 3 uur
- ☐ 3 tot 4 uur
- ☐ 4 uur of langer

3.3 De volgorde van de vragen

Zijn alle vragen op een correcte manier geformuleerd, dan moet je nadenken over de volgorde waarin je de vragen gaat stellen. Het is in ieder geval goed om vragen over hetzelfde onderwerp bij elkaar te houden. Dus al je iemand vraagt over zijn hobby's (wat is je hobby? Hoeveel tijd besteed je eraan? Wat kost het?), dan zet je alle vragen hierover bij elkaar. Verder is het handig om de vragenlijst met een paar leuke en interessante vragen te beginnen. Bewaar saaie, oninteressante vragen (bijvoorbeeld over opleiding of inkomen) maar voor het laatste deel van de vragenlijst. Zo houd je de aandacht van de respondent beter vast.

Verder moet je proberen te voorkomen dat iemands antwoord op een vraag teveel wordt beïnvloed door vragen die je daarvoor al hebt gesteld. Die eerdere vragen kunnen als het ware de toon zetten voor de respondent. De objectiviteit van de vragen zou zo wel eens kunnen worden aangetast.

Will Tiemeijer vergelijkt in zijn boek 'Wat 93,7 procent van de Nederlanders moet weten over opiniepeilingen' twee verschillende peilingen die op het zelfde moment in 2007 zijn uitgevoerd en die beide vragen naar de mening van de Nederlanders over de Europese Unie. De ene peiling was van Maurice de Hond. Daaruit bleek dat 43 procent van de Nederlanders de voordelen van de Europese Unie groter vond dan de nadelen, terwijl 35 procent de nadelen groter vond dan de voordelen. De rest wist het niet. Een niet zo erg positief beeld dus.

3.3.1 De Eurobarometer



De andere peiling was de *Eurobarometer*. Dit is een peiling die in opdracht van de Europese Commissie regelmatig wordt uitgevoerd. Daaruit bleek dat 69 procent van de Nederlanders het lidmaatschap van de EU een voordeel vond. En 77 procent vond het een goede zaak dat Nederland lid is van de EU. Dit is een veel positiever beeld.

Bij nader onderzoek blijkt dat in de peiling van Maurice de Hond de mensen in de voorafgaande vraag moesten aangeven wat ze de grootste nadelen van de EU vonden. Daarbij konden ze kiezen uit mogelijke antwoorden als de snelle uitbreiding, de mogelijke toetreding van Turkije, de invoering van de euro, het bemoeien met zaken die beter door de landen zelf geregeld kunnen worden, de verspilling van geld in Brussel, het tekort aan democratische rechten voor de burgers, het verlies aan eigen identiteit, enzovoorts. Kortom, respondenten werden heel erg bewust gemaakt van de mogelijk negatieve aspecten van de EU. Positieve aspecten kwamen niet aan de orde. Dat had gevolgen voor de beantwoording van de volgende vraag. De Eurobarometer bevatte geen vragen met een dergelijke negatieve lading. Daar werd de vraag op een neutralere manier gesteld.

Je moet je ook bedenken dat een vraag vooraan in een peiling een heel ander effect kan hebben dan een vraag achteraan. Dit verschijnsel is bijvoorbeeld aangetoond in onderzoek naar woningbehoeften van mensen. Mensen konden eerst heel tevreden zijn over de eigen woning, maar als na allerlei vragen (Heeft u een ligbad? Heeft u een tuin op het zuiden? Is uw huis goed geïsoleerd? Heeft u wel eens lekkage?) bleek dat er toch nog wel het een en ander ontbrak, dan kon die tevredenheid wel eens omslaan in ontevredenheid. Probleem is dat een peiling de respondenten veel meer bewust kan maken van een probleem dat aan de orde komt. Daardoor worden ze aan het denken gezet en dat kan hun mening wijzigen.

Een tweede aspect van de volgorde van de vragen is dat niet elke vraag relevant hoeft te zijn voor elke respondent. In een peiling over werk en werkloosheid ondervraag je mensen met en zonder werk. Werkenden vraag je dan niet hoe ze naar werk zoeken en werklozen vraag je niet naar hun werkomstandigheden. Het stellen van overbodige vragen kan leiden tot irritatie bij de respondent en uiteindelijk zelfs tot een weigering om verder mee te werken. Bovendien zal het voor een respondent niet eenvoudig zijn om overbodige vragen te beantwoorden. Om deze problemen te vermijden, moet je, waar nodig, *sprongopdrachten* opnemen in de vragenlijst. Dat zijn instructies die aangeven dat een respondent bepaalde vragen moet overslaan, en de beantwoording verderop in de vragenlijst moet voortzetten.

Voorbeeld 3.3.2 laat zien hoe je sprongopdrachten in een vragenlijst kunt opnemen. Het is een verkorte versie van een denkbeeldig luisteronderzoek van een lokale omroep. Als respondenten de lokale omroep niet kennen, dan hoeven ze ook geen vragen over de omroep te beantwoorden. Daarom springen deze respondenten naar de algemene vragen aan het einde van de vragenlijst. Hetzelfde geldt voor respondenten die de omroep wel kennen, maar er niet naar luisteren. Deze respondenten moeten echter wel eerst uitleggen waarom ze niet luisteren.

3.3.2 Voorbeeld van een korte vragenlijst voor een luisteronderzoek

1. Kent u Lommerdal FM, de lokale omroep van de gemeente Lommerdal?

- ☐ Nee → Vraag 5
- ☐ Ja → Vraag 2

2. Luistert u wel eens naar Lommerdal FM?

- ☐ Nee → Vraag 3
- ☐ Ja → Vraag 4

3. Waarom luistert u nooit naar Lommerdal FM?

- ☐ Luister nooit naar de radio
- ☐ Geen belangstelling voor lokale gebeurtenissen
- ☐ Kan de lokale omroep niet goed ontvangen
- ☐ Andere reden

→Vraag 5

4. Naar welke programma's van de lokale omroep luistert u regelmatig? (meer antwoorden zijn mogelijk)

- ☐ Programma's met lokaal nieuws en informatie
- ☐ Sportprogramma's
- ☐ Programma's over kunst en cultuur
- ☐ Kerkelijke programma's
- ☐ Muziekprogramma's
- ☐ Andere programma's

5. Wat is uw geslacht?

- ☐ Man
- ☐ Vrouw

6. Wat is uw leeftijd?

- ☐ 13-19 jaar
- ☐ 20-39 jaar
- ☐ 40-59 jaar
- ☐ 60 jaar of ouder

Sprongopdrachten hebben nog het extra voordeel dat ze ook het aantal vragen beperken dat de respondenten moeten beantwoorden. Daardoor kan het interview sneller worden afgerond. Wel is het zo dat het maken van een vragenlijst met sprongopdrachten ingewikkelder is. Je moet heel goed controleren of elke respondent wel het juiste pad door de vragenlijst neemt.

3.4 Testen van vragenlijsten

Je moet een vragenlijst altijd eerst testen voordat hij het veld in gaat. Elke onderzoeker zal het hiermee eens zijn, maar dat betekent niet dat het in de praktijk ook altijd gebeurt. Het komt regelmatig voor dat er niet genoeg tijd en geld is om een goede test uit te voeren. Daarom wordt aan dit aspect wel eens te weinig aandacht besteed.

Wat houdt een testprocedure precies in? Concreet betekent het dat je de vragenlijst in de praktijk uitprobeert. Dan kan op twee manieren gebeuren. In de eerste plaats kun je respondenten benaderen en de echte peiling naspelen. De respondenten weten dan dus niet dat het slechts om een test gaat. Dat heeft het voordeel dat ze zich net zo gedragen als bij een echt interview. Een tweede manier om de test aan te pakken is de respondenten vertellen dat het om een test gaat. Dat heeft het voordeel dat je aan de respondenten kunt vragen of ze de vragen hebben begrepen, wat er onduidelijk was in een vraag, en waarom ze een bepaald antwoord hebben gegeven. We noemen dit ook wel *cognitief interviewen*.

Onderzoekers willen nog wel eens vergeten dat niet elke respondent evenveel enthousiasme kan opbrengen voor de vragen in een peiling. Als de vragen geen belangstelling oproepen bij de respondenten, zal de kwaliteit van de antwoorden te wensen overlaten. Daarom moeten interviewers in een test goed registreren welke vragen interessant gevonden worden en welke niet. Het zal duidelijk zijn dat een vragenlijst vol saaie, oninteressante vragen niet erg succesvol zal zijn.

Ook al zijn de vragen redelijk interessant, toch kan er tegen het einde van het interview een soort *vragenmoeheid* optreden. Dat zal natuurlijk eerder bij lange vragenlijsten het geval zijn. Daarom moet je altijd de vragenlijst zo kort mogelijk houden. Een mondelinge peiling zou niet langer dan 50 minuten moeten duren (een lesuur). Bij een schriftelijke peiling en een online-peiling moeten de respondenten het helemaal zelf doen. Ze houden het meestal geen 50 minuten vol. Vragenlijsten moeten dan veel korter zijn. Denk hierbij aan, pakweg, 15 minuten. Het is verstandig om tijdens een test te controleren of de vragenlijst niet te lang is. Misschien is de conclusie dan wel dat het beter is om een paar vragen te schrappen.

Een goede test van een vragenlijst verloopt in twee fasen. In de eerste fase worden zo'n 25 tot 75 interviews afgenomen. Het is een goed idee om de ervaringen van de interviewer(s) in deze fase vast te leggen met een klein vragenlijstje. Vragen die zouden kunnen worden gesteld, zijn bijvoorbeeld:

- Waren er vragen waarbij de respondent zich ongemakkelijk voelde?
- Waren er vragen die moesten worden herhaald?
- Waren er vragen die de respondent verkeerd begreep?
- Welke vragen waren het moeilijkst of vervelendst om te stellen?
- Waren er onderdelen in de vragenlijst waarin de respondenten ongemotiveerd en ongeïnteresseerd reageerden?
- Waren er onderdelen in de vragenlijst waarbij het gevoel opkwam dat de respondent meer had willen zeggen?

De eerste fase moet een grondig onderzoek van de vragenlijst zijn dat de essentiële fouten in de vragenlijst aan het licht brengt. De tweede fase van de test is een soort generale repetitie. Daarin gaat het niet meer om het repareren van grote fouten, of het proberen van een geheel nieuwe aanpak. Het gaat om de 'fine tuning' en 'finishing touch', het wegwerken van de laatste loshangende rafeltjes. In de tweede fase simuleer je het echte interview. Je vertelt de respondenten niet dat het om een test gaat.

Grote onderzoeksorganisaties hebben voor het testen van vragenlijsten vaak een vragenlaboratorium ingericht. In een huiskamersituatie wordt op video vastgelegd wat er gebeurt tijdens het invullen van de vragenlijst. Eventueel kunnen onderzoekers door eenzijdige spiegels meekijken. Ook al is een dergelijke professionele testomgeving niet beschikbaar, dan is het nog steeds erg zinvol om de vragenlijst te testen op een beperkt aantal willekeurige personen. Problemen zullen dan al snel aan het licht komen.

Ook voor je eigen peiling is het belangrijk om de vragenlijst te testen. Al probeer je de vragenlijst alleen maar uit bij je klasgenoten of je familie, dan zullen toch allerlei problemen al gauw aan het licht komen.

Vraag 3.1

Leg uit wat er niet goed is aan de volgende vragen. Probeer ook een verbeterde versie van de vragen te maken:

- a. Hoe vaak ben je in de afgelopen twee jaar naar de bioscoop geweest?
 - b. Ben je tegen een verbod op alcohol voor jongeren onder de 18 jaar?
 - c. Heb je wel eens een coffeeshop bezocht en daar softdrugs gekocht?
 - d. Je hebt toch zeker ook wel een smartphone?
 - e. Hoeveel geheugen heeft je computer?
 - f. Word je vaak gepest op school?
-

4.

Het verzamelen van gegevens

Elke persoon in de steekproef zou de vragenlijst moeten invullen. Hoe ga je dat doen? Dat kan op verschillende manieren:

- **Mondeling:** interviewers gaan bij de geselecteerde personen thuis op bezoek. De interviewers stellen de vragen, de personen geven de antwoorden, en de interviewers registreren de antwoorden (op papier of in een laptop of tablet);
- **Telefonisch:** interviewers bellen de geselecteerde personen op. De interviewers stellen telefonisch de vragen, personen geven de antwoorden, en de interviewers registreren de antwoorden (op papier of in een computer);
- **Schriftelijk:** Je stuurt de vragenlijst in een brief naar de geselecteerde personen. Die personen vullen het formulier zelf in en sturen het daarna ingevuld terug;
- **Online:** Je stuurt de selecteerde personen een link naar een website met de elektronische vragenlijst. De personen vullen de vragenlijst zelf in op hun computer.

Bij je keuze voor een van deze vier methoden van gegevensverzameling zijn twee zaken van belang: hoeveel kost het en hoe goed zijn de antwoorden die je krijgt? Als je goed getrainde interviewers de gegevens laat verzamelen, dan levert dat meestal goede gegevens op, maar de kosten zijn hoog. Die interviewers moeten immers worden betaald. Als je het zonder interviewers doet (zoals bij een schriftelijke peiling of een online peiling), dan is het goedkoper. De kwaliteit van de antwoorden is dan veel minder. Er zitten veel meer fouten in de antwoorden.

Het verzamelen van de gegevens noemen we soms het *veldwerk*. Die term verwijst eigenlijk naar peilingen waarbij interviewers op pad gaan ('het veld in gaan') om bij de respondenten thuis de vragenlijsten in te vullen. We gebruiken de term hier ook voor andere manieren van het verzamelen van de gegevens.

In dit hoofdstuk beschrijven we verschillende methoden van gegevensverzameling. We beginnen met drie traditionele manieren van het verzamelen van gegevens waarbij je gebruik maakt van vragenlijsten op papier. Dan beschrijven we enkele manieren van computergestuurd interviewen. Hierbij is de papieren vragenlijst vervangen door een digitale vragenlijst in een computerprogramma. Ten slotte gaan we nog wat dieper in op de steeds populairdere online-peilingen.

4.1 Traditionele gegevensverzameling

Bij een *schriftelijke peiling* verstuurt je de (nog lege) vragenlijsten per post naar de personen in de steekproef. Daarbij zit uiteraard het verzoek de vragenlijsten ingevuld weer terug te sturen. Je hebt geen interviewers nodig. Daarom is dit een goedkope manier van gegevensverzameling. Een bijkomend voordeel is dat respondenten eerlijker antwoord geven op gevoelige vragen als er geen onbekende (de interviewer) bij aanwezig is. De afwezigheid van interviewers heeft echter ook nadelen. Er is niemand die de persoon kan overtuigen van het belang van het invullen van de vragenlijst. En ook kan de interviewer niet helpen bij het invullen van de vragenlijst.

De (onervaren) respondenten moeten de vragenlijst zonder problemen kunnen invullen. Dat stelt hoge eisen aan de manier waarop je de vragen stelt en aan de opmaak van de vragenlijst. Een schriftelijke peiling kan de indruk wekken dat het niet zo belangrijk is om mee te doen. Veel mensen zullen de vragenlijst niet invullen, maar bij het oud papier gooien.

Als je de nadelen van een schriftelijke peiling te groot vindt, kun je ook kiezen voor een *mondelinge peiling*. Hierbij bezoeken interviewers de geselecteerde personen thuis (of op een andere locatie) en proberen ze in een persoonlijk gesprek de vragen beantwoord te krijgen. Een punt van overweging is of je het bezoek van de interviewer van tevoren moet aankondigen. Je kunt dan uitleggen wat het doel van de peiling is, dat het belangrijk is om mee te doen, en dat je de gegevens vertrouwelijk zult behandelen. Je kunt dat bijvoorbeeld doen in een *aanschrijfbrief*. Voor een grotere peiling zou je ook een persbericht naar de media kunnen sturen. Hopelijk besteden die er dan ook de nodige aandacht aan.

4.4.1 Een mondelinge peiling



„OPINIEONDERZOEK!
BEANTWOORDEN!“

Een nadeel van een mondelinge peiling is dat het duur is. Je moet een team interviewers trainen en betalen. Daarbij kunnen nog veel reiskosten komen (als de respondenten ver uit elkaar wonen). Een mondelinge peiling levert echter wel meestal een hoge respons op en verzamelde gegevens van goede kwaliteit.

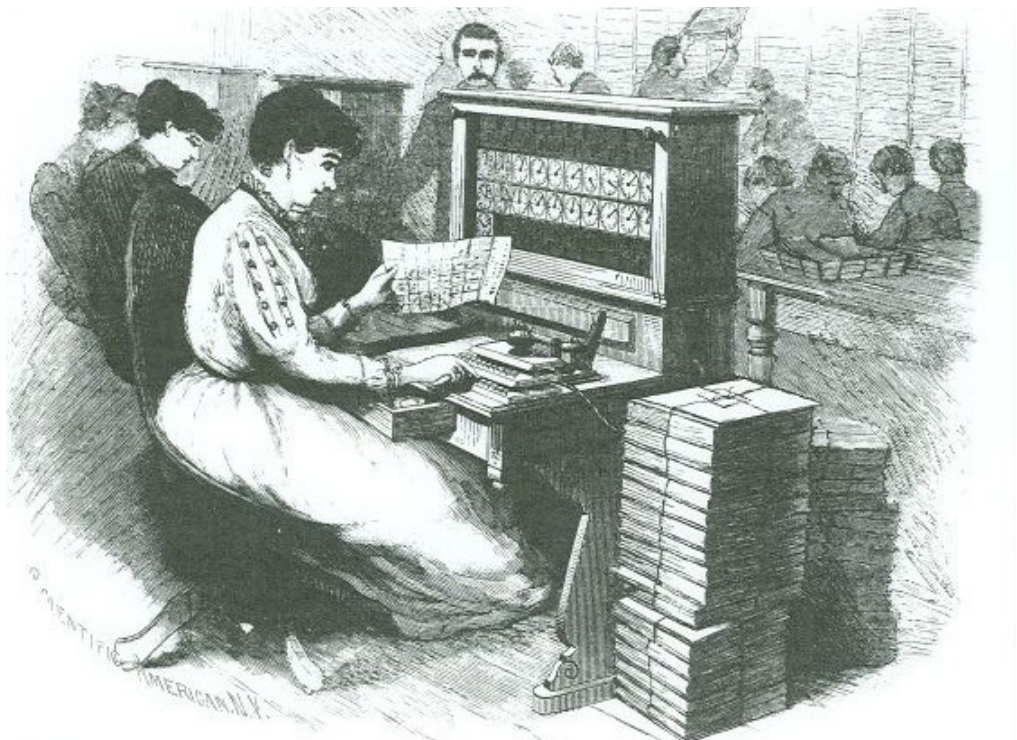
Een derde traditionele methode van gegevensverzameling is een *telefonische peiling*. Als je kijkt naar de kosten en de kwaliteit van de gegevens, dan ligt een telefonische peiling in tussen een schriftelijke en een mondeling peiling. Telefonisch interviewen is minder duur dan schriftelijk interviewen. Je hebt nog wel interviewers nodig, maar minder dan bij een mondelinge peiling. Omdat interviewers niet hoeven te reizen, kunnen ze meer interviews op een dag doen. Dat levert een behoorlijke kostenbesparing op. Daar staat dan wel weer tegenover dat de gestelde vragen niet te ingewikkeld mogen zijn. En als het gesprek te lang duurt, dan haken de respondenten af.

De keuze van de methode van gegevensverzameling blijft een lastige zaak. Vaak zal de keuze een compromis zijn, waarbij je aan de ene kant probeert de kosten zo laag mogelijk te houden, en andere kant de kwaliteit zo hoog mogelijk.

4.2 Computergestuurd interviewen

De computer heeft altijd een grote rol gespeeld bij de verwerking van gegevens die met peilingen zijn verzameld. Al in 1890 gebruikte het Amerikaanse Census Bureau zogenaamde Hollerith-machines voor het verwerken van de formulieren van de volkstelling.

4.2.1 Een Hollerith-machine



Pas in de laatste 40 jaar heeft de computer ook zijn intrede gedaan bij het interviewen zelf. Dat werd mogelijk door de opkomst van de relatief goedkope en gebruikersvriendelijke microcomputers. Bij computergestuurd enquêteren staat de vragenlijst niet op papier, maar zit in digitale vorm in een computerprogramma. Dit programma bepaalt welke vragen wanneer moeten worden gesteld en het controleert ook de antwoorden.

Het gebruik van computers bij het interviewen heeft drie belangrijke voordelen. In de eerste plaats ontlast je hiermee de interviewers. Die hoeven niet steeds de volgende vraag op te zoeken. Dat doet de computer al voor hen. Dat lijkt misschien een simpele activiteit, maar soms hangt het stellen van een vraag af van de antwoorden op al eerder gestelde vragen. Je moet dan naar een ander deel van de vragenlijst springen. Een fout is hierbij snel gemaakt.

In de tweede plaats kan de computer tijdens het gesprek allerlei controles uitvoeren. Fouten die je tijdens het gesprek ontdekt, kun je ook tijdens het gesprek verbeteren. Dat gaat beter tijdens het interview dan achteraf. De respondent zit dan nog aan tafel en kan je helpen bij het verbeteren van de fout.

Het derde voordeel van het gebruik van een computer bij het interview is dat na afloop van het gesprek de gegevens meteen al in de computer zitten. Dat hoeft je dus niet meer achteraf te doen. Dat levert een aanzienlijke versnelling op bij het verwerken van de gegevens.

De computer werd het eerst ingezet bij telefonische peilingen. Dat gebeurde in de jaren zeventig van de vorige eeuw. De interviewers krijgen de beschikking over een telefoon en een computer. Ze nemen telefonisch contact op met de respondenten en starten vervolgens het interview. De computer stuurt het gesprek en controleert de antwoorden. Deze vorm van interviewen noemen we *CATI* (*Computer Assisted Telephone Interviewing*).

In de jaren tachtig van de vorige eeuw begon ook het gebruik van de computer bij mondeling interviewen op gang te komen. Dat werd mogelijk toen kleine, handzame computers op de markt kwamen die door interviewers makkelijk konden worden meegenomen. Toen was het een laptop, nu steeds meer een tablet. De interviewers starten bij de mensen thuis het interviewprogramma op hun apparaat. De computer neemt weer het heft in handen bij het stellen van de vragen en het controleren van de antwoorden. Mondeling interviewen met een computer duiden we meestal aan met *CAPI* (*Computer Assisted Personal Interviewing*).

4.2.2 Een van de eerste interview-computers: de Epson PX-4 notebook (1987)



4.3 Peilingen via het internet

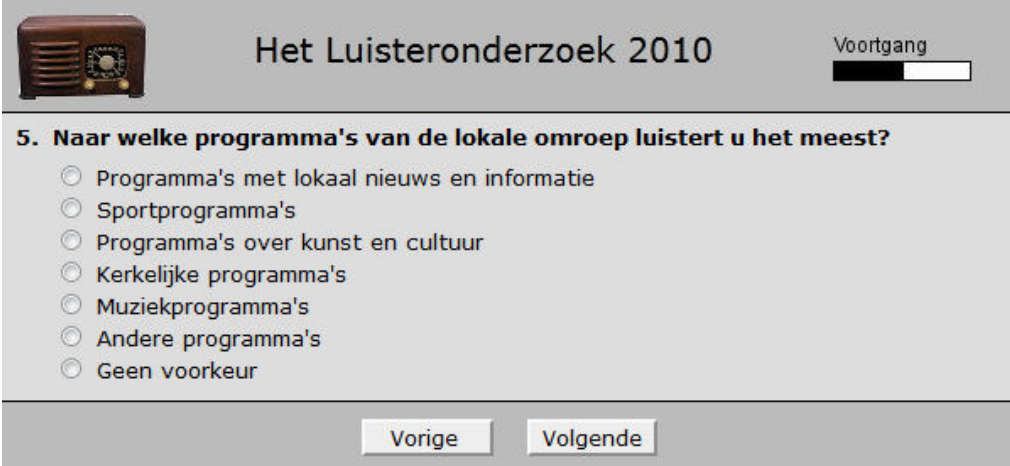
Met de razendsnelle opkomst van het internet heeft een nieuw type gegevensverzameling zijn intrede gedaan: peilingen via het internet. Dit wordt soms ook wel *CAWI* (*Computer Assisted Web Interviewing*) genoemd. Hierbij zet je de vragenlijst op het internet en de respondent beantwoordt die vragen ook op het internet. Er is geen interviewer.

Op het eerste gezicht heeft een peiling via het internet een aantal aantrekkelijke eigenschappen:

- Je krijgt op een eenvoudige manier toegang tot een zeer grote groep potentiële respondenten, namelijk iedereen met internet;
- Het kost niet veel geld om de vragenlijsten aan te bieden. Je hoeft immers geen interviewers in te schakelen, je hebt geen drukkosten (zoals bij papieren vragenlijsten), en je hebt ook geen verzendkosten;
- Je kunt een peiling heel snel uitvoeren. Er hoeft maar weinig tijd verloren te gaan tussen opstellen en aanbieden van de vragenlijst;
- Het internet biedt aantrekkelijke extra mogelijkheden om zaken als beeld (foto, video, animatie) en geluid in de vragenlijst op te nemen.

Een peiling via het internet lijkt een snelle, goedkope en aantrekkelijke manier om veel gegevens te verzamelen. Het is echter niet allemaal rozengeur en maneschijn. Het is erg eenvoudig geworden om een online-peiling op te zetten. Kijk maar eens rond op het internet. Je komt de polls overal tegen, vaak bijvoorbeeld tijdens een verkiezingscampagne. Er zijn allerlei websites (zoals die van *SurveyMonkey*, *SurveyGizmo*, *QuestionPro*) waarmee je in korte tijd een peiling in de lucht kunt brengen, ook al heb je geen enkel verstand van het maken van peilingen. Veel van die peilingen zijn niet goed opgezet. En door het grote aanbod is het moeilijk de goede peilingen van de slechte te onderscheiden.

4.3.1 Een voorbeeld van een online-vragenlijst



The screenshot shows a web-based survey interface. At the top left is a small icon of a vintage radio. The title 'Het Luisteronderzoek 2010' is centered at the top. To the right of the title is a progress bar labeled 'Voortgang' with a black bar indicating the current progress. Below the title bar, the survey question is displayed: '5. Naar welke programma's van de lokale omroep luistert u het meest?'. There are seven radio button options listed below the question: 'Programma's met lokaal nieuws en informatie', 'Sportprogramma's', 'Programma's over kunst en cultuur', 'Kerkelijke programma's', 'Muziekprogramma's', 'Andere programma's', and 'Geen voorkeur'. At the bottom of the survey area, there are two buttons: 'Vorige' (Previous) and 'Volgende' (Next).

De grote hoeveelheid online-peilingen leidt tot steeds meer non-respons. Mensen worden overvoerd en haken af. In feite treedt hetzelfde effect op als dat waaraan telefonische peilingen lijden: de vele telefonische verkoopactiviteiten waarmee mensen worden lastig gevallen (vooral 's avonds tijdens het eten), vermindert de bereidheid om aan echte telefonische enquêtes mee te werken.

Samenvattend kunnen we zeggen dat het bij veel online-peilingen mogelijk is om goedkoop en in korte tijd een grote hoeveelheid gegevens te verzamelen. Maar dat wil niet zeggen dat het goede peilingen zijn. Vaak zitten die peilingen verkeerd in elkaar, waardoor verkeerde conclusies uit de gegevens worden getrokken.

Als je erover denkt om een online-peiling op te zetten, moet je goed letten op een aantal zaken. Een van de belangrijkste problemen van dit soort peilingen is misschien wel de *onderdekking*. Er zijn steeds meer mensen die toegang hebben tot het internet, maar het is zeker niet zo dat iedereen dat heeft. Vooral ouderen, laagopgeleiden en allochtonen zijn vaak onvoldoende vertegenwoordigd. De peiling is dan niet *representatief*. Hij is geen goede afspiegeling van de doelpopulatie. De verhoudingen kloppen niet. Er zitten teveel mensen van de ene groep in de peiling en te weinig van een andere groep. Het gevolg daarvan is dat je verkeerde conclusies trekt uit je peiling.

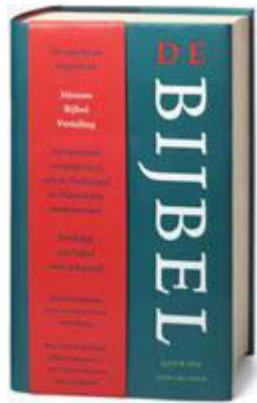
Een ander probleem wordt gevormd door het trekken van de steekproef. Hoe trek je een steekproef van internetgebruikers? Er is geen lijst met alle e-mailadressen beschikbaar. Veel internetpeilers omzeilen dit probleem door de internetgebruikers zichzelf te laten selecteren voor de peiling. We noemen dit *zelfselectie*. Via 'banners', 'popup windows' en andere vormen van reclame maken ze internetgebruikers attent op de peiling. De peiler hoopt dan dat deze gebruikers zich hierdoor laten overhalen en besluiten de vragenlijst in te vullen.

Typische voorbeelden van zelfselectie zijn te vinden bij veel politieke peilingen in Nederland. In deze panels zitten vooral mensen die het leuk vinden om aan peilingen mee te doen en die geïnteresseerd zijn in politiek. Daardoor kun je grote vraagtekens zetten bij de representativiteit van dit soort peilingen.

Je kunt het probleem van de zelfselectie oplossen door de selectie van respondenten niet via het web te doen. Je kunt, bijvoorbeeld, eerst een steekproef van namen en adressen van personen trekken. Die stuur je vervolgens een brief waarin je ze uitnodigt naar de website te gaan waar de vragenlijst is te vinden. In die brief moet dan ook een unieke code staan waarmee de geselecteerde personen toegang krijgen tot de vragenlijst. Daarmee kun je voorkomen dat iemand meer dan één keer de vragenlijst invult, en ook dat iemand van buiten de doelpopulatie de vragenlijst invult.

Een aardig voorbeeld van de effecten van zelfselectie was de verkiezing van de NS Publieksprijs in 2005. Dat is een literaire prijs die elk jaar wordt toegekend. Voor deze prijs wordt geen speciale jury ingesteld. De lezers kunnen zelf stemmen. In 2005 werden zes boeken genomineerd. In een online-peiling konden de lezers stemmen op een van deze zes boeken, maar ze konden ook zelf een boek opgeven als hun favoriete boek niet in de lijst stond. In totaal brachten 92 duizend mensen hun stem uit. Tot verbazing van iedereen werd niet een van de genomineerde boeken tot winnaar gekozen. Zo'n 72 procent van de stemmers koos voor de Nieuwe Bijbelvertaling.

4.3.2 De winnaar van de NS Publieksprijs 2005



Deze verpletterende uitslag was het resultaat van een campagne gevoerd door onder anderen het dagblad Trouw, de Evangelische Omroep, het Nederlands Bijbelgenootschap, de Katholieke Bijbelstichting en de Protestantse Kerk om te stemmen op de nieuwe Bijbelvertaling. Deze gang van zaken was niet tegen de regels van deze literaire prijs, maar je kunt je natuurlijk wel afvragen of deze uitslag representatief is voor de Nederlandse bevolking.

Opgave 4.1

De directie van een school wilt weten of er leerlingen op school worden gepest. Er wordt een peiling opgezet om aan een steekproef van leerlingen vragen hierover te stellen. Pesten op school is een gevoelig onderwerp. Er is een risico dat leerlingen niet allemaal eerlijk zullen antwoorden.

Welke methode van gegevensverzameling zou je kiezen voor deze peiling? Je kunt kiezen uit een mondelinge peiling, een telefonische peiling, een schriftelijke peiling of een online-peiling. Leg voor elke methode van gegevensverzameling uit waarom die in aanmerking komt of niet.

5.

Het trekken

van een

steekproef

Met je peiling wil je conclusies trekken over de hele doelpopulatie, terwijl toch maar een steekproef van personen de vragenlijst heeft ingevuld. Kan dat? Ja, dat kan, maar alleen als je de steekproef op een correcte manier hebt getrokken. Correct betekent dat je de steekproef moet loten uit de doelpopulatie. Als je het toeval laat bepalen wie er wel en niet in de steekproef komt, dan weet je zeker dat je niemand systematisch bevoordeelt of benadeelt. Iedereen heeft dezelfde kans om in de steekproef te komen.

Hoe trek je in de praktijk een steekproef die een goede afspiegeling is van de doelpopulatie? Daarvoor zijn twee ingrediënten nodig: een steekproefkader en een procedure voor het trekken van een steekproef uit dat steekproefkader. We zullen beide zaken behandelen in dit hoofdstuk.

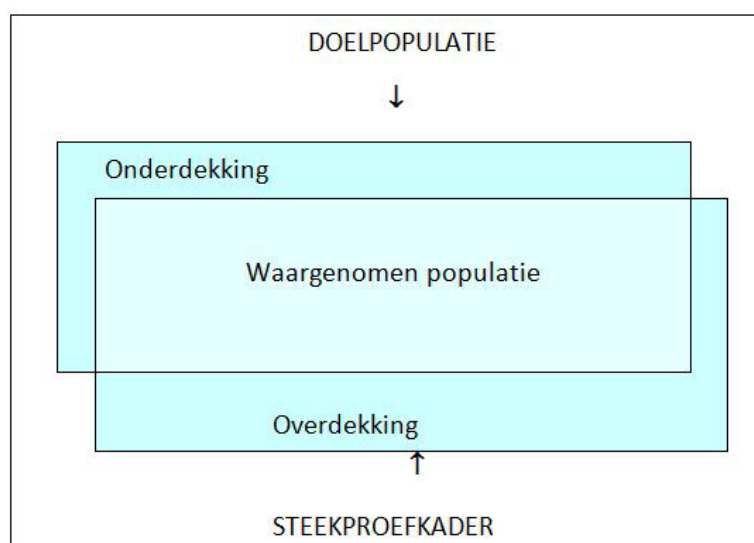
5.1 Het steekproefkader

Om een steekproef te trekken heb je een lijst nodig met namen, adressen of andere contactgegevens van iedereen in de doelpopulatie. Die lijst noemen we het *steekproefkader*.

Een *steekproefkader* is dus soort een administratieve weergave van de doelpopulatie. Het is een lijst waarin elke persoon uit de doelpopulatie staat vermeld. Ook zal duidelijk moeten zijn hoe je contact kunt opnemen met die personen.

Voor een schriftelijke peiling of een mondelinge peiling is een lijst met adressen erg handig. Voor een telefonische peiling gaat de voorkeur natuurlijk uit naar een lijst met telefoonnummers. En voor een online-peiling denk je in de eerste plaats aan een lijst met e-mailadressen.

5.1.1 Doelpopulatie en steekproefkader



Het is van groot belang dat het steekproefkader precies de hele doelpopulatie bevat. Is dat niet het geval dan trek je de steekproef in feite uit een andere populatie en doe je ook uitspraken over een andere populatie. Wat gebeurt er als je, door praktische problemen gedwongen, een steekproef trekt uit een steekproefkader dat niet overeenkomt met de doelpopulatie? Die situatie is weergegeven in figuur 5.1.1.

In de eerste plaats kunnen er mensen in de doelpopulatie zitten die niet zijn vermeld in het steekproefkader. Deze situatie heet *onderdekking*. Dit verschijnsel doet zich bijvoorbeeld voor als je onderzoek wilt doen onder mensen die hier illegaal verblijven, terwijl je een steekproef trekt uit het bevolkingsregister.

De gevolgen van onderdekking kunnen ernstig zijn. Als een speciale groep uit de doelpopulatie daardoor ontbreekt in je peiling, kunnen de resultaten een vertekend beeld opleveren. Daarbij komt bovendien dat onderdekking niet eenvoudig is waar te nemen als je er niet echt goed naar zoekt.

Het kan ook gebeuren dat er mensen in het steekproefkader zitten die niet thuishoren in de doelpopulatie. Dat verschijnsel duiden we aan met *overdekking*. Je hebt dan mensen in je onderzoek die er helemaal niet in thuishoren. Ook dat kan tot verkeerde conclusies leiden. Overdekking kun je vrij eenvoudig ontdekken in het veld, door voor alle personen in de steekproef eerst te controleren of ze wel voldoen aan de definitie van de doelpopulatie. Een interviewer kan dat controleren, maar dit kan ook met een paar gerichte vragen aan het begin van de vragenlijst.

Stel dat je een telefonische peiling wilt doen onder de gezinnen in een gemeente. Daarvoor zou je het telefoonboek als steekproefkader kunnen gebruiken. In de eerste plaats kan er dan onderdekking optreden omdat niet alle nummers in het telefoonboek staan. Pakweg 30 tot 40 procent van de mensen staat tegenwoordig niet meer in het telefoonboek. Daarbij zitten veel mensen die alleen maar een mobieltje hebben. Het gaat hier vooral om jongeren. Dus die groep zou wel eens zwaar ondervertegenwoordigd kunnen zijn in je onderzoek. In de tweede plaats staan er nummers in het telefoonboek die niet bij gezinnen horen. Denk hierbij aan nummers van winkels en bedrijven. Daardoor kan het gebeuren dat je mensen belt en interviewt die niet in de populatie thuishoren. Dit is overdekking.

5.2 Hoe het niet moet

Het is in de praktijk lang niet altijd eenvoudig om een geschikt steekproefkader te vinden. Daarom zou je kunnen overwegen om je steekproef op een andere manier te trekken uit de doelpopulatie. Er zijn allerlei twijfelachtige manieren om een steekproef te trekken die niet zijn gebaseerd op loting uit een steekproefkader. Daarbij loop je een groot risico op onbetrouwbare uitkomsten. Immers, je steekproef hoeft dan geen goede afspiegeling meer te zijn van de populatie waaruit hij is getrokken. Schattingen van populatiekenmerken kunnen de plank dus volledig misslaan.

Ook al heb je geen steekproefkader tot je beschikking, dan moet je toch je best doen om een zo representatief mogelijke steekproef te krijgen. Stel, bijvoorbeeld, dat je een peiling wilt doen onder de klanten van een supermarkt. Je hebt geen lijst van alle klanten. Dus moet je de steekproef selecteren door de klanten in de winkel direct aan te spreken. Zorg er dan in ieder geval voor dat je allerlei soorten klanten in je steekproef krijgt. Kies je klanten verspreid over de hele dag, en misschien ook wel verspreid over de hele week. Kies je klanten zonder 'aanziens des persoons'. Neem bijvoorbeeld elke derde klant die de winkel uitkomt.

We geven in deze paragraaf een paar voorbeelden van peilingen waarbij minder aandacht voor de representativiteit van de peiling kan leiden tot onjuiste uitkomsten.

Peilen in het winkelcentrum

Een lokale omroep wilde een luisteronderzoek uitvoeren. Om toch maar snel een heleboel gegevens te verzamelen, besloot de onderzoeker om een paar medewerkers als interviewers op zaterdagmiddag naar het plaatselijke winkelcentrum te sturen. Het was daar gezellig druk zodat veel mensen konden worden aangesproken. In korte tijd was er een al een hele stapel vragenlijsten ingevuld.

Bij de analyse van de gegevens kwam de onderzoeker tot een onverwachte conclusie: bijna niemand luisterde naar het sportprogramma dat op zaterdagmiddag werd uitgezonden. Bij nader inzien is die conclusie natuurlijk helemaal niet zo verrassend. Er werden alleen maar mensen ondervraagd die op zaterdagmiddag in het winkelcentrum aanwezig waren. Daarmee werd de doelpopulatie in feite gereduceerd van alle inwoners tot het zaterdagse winkelpubliek. En het is niet zo verbazingwekkend dat die op zaterdagmiddag niet naar de radio luisterden. Ze waren immers aan het winkelen. De mensen in de steekproef waren geen goede afspiegeling van alle inwoners van de gemeente.

De peiling van de ANWB over het rekeningrijden

In de eerste maanden van 2010 was er in Nederland een uitgebreide discussie over het invoeren van rekeningrijden. Automobilisten zouden dan op bepaalde tijden op bepaalde wegen extra belasting moeten betalen. Het idee was om hiermee de files terug te dringen.

De ANWB besloot de mening van de leden hierover te peilen. Op de website van de ANWB werd een uitgebreide vragenlijst gezet. Er werd geen steekproef getrokken. Iedereen kon die vragenlijst invullen, dus ook bijvoorbeeld buitenlanders (Belgen) en niet-leden. En als je zin had, kon je de vragenlijst zelfs meer dan één keer invullen.

Door deze opzet kun je vraagtekens zetten bij de representativiteit van de peiling. Om te beginnen kun je je afvragen wat eigenlijk de doelpopulatie is. De ANWB-leden? Alle autorijders? Alle Nederlanders?

De ANWB gaf aan dat het om de leden ging. Gelukkig werd in de peiling de vraag gesteld of de respondent lid was van de ANWB. Het is echter nog maar de vraag of iedereen die zijn mening kwijt wilde, die vraag ook correct heeft beantwoord. Niet-leden zouden zich kunnen voordoen als leden en zo de uitkomsten van de peiling kunnen beïnvloeden.

5.2.1 De peiling van de ANWB over rekeningrijden



Meerderheid ANWB-leden vóór betalen naar gebruik

68% van de ondervraagde ANWB-leden is het met dat principe eens, waarmee er ook in 2010 nog steeds draagvlak voor bestaat...>

Uiteindelijk werd de vragenlijst ingevuld door 350 duizend mensen die zeiden lid te zijn. Daarvan was 68 procent voorstander van een vorm van rekeningrijden. In hun onderzoeksrapport vermeldden de onderzoekers zelf ook dat de peiling niet representatief was. Ze gaven aan dat voor hen belangrijker was dat veel mensen meededen aan het onderzoek en hun mening gaven.

Een lezersonderzoek van een tijdschrift

De redactie van een gratis huis-aan-huisblad wilde weten of het blad goed werd gelezen. Daarom werd besloten tot het uitvoeren van een lezersonderzoek. In het blad werd een vragenlijst opgenomen. En uiteraard was een van de vragen of men het blad wel eens las. De respondenten moesten de ingevulde vragenlijst uitknippen en opsturen naar de redactie.

Tot de niet geringe verbazing en blijdschap van de redactie bleken bijna alle respondenten het blad wel eens te lezen. De redactie trok dan ook de conclusie dat het blad goed werd gelezen.

De adder die hier onder het gras zat, was dat mensen die het blad niet lazen, het onmiddellijk ongezien weggooiden en dus ook de vragenlijst helemaal niet tegenkwamen. In feite werd de doelpopulatie beperkt tot alleen mensen die op zijn minst het blad wel eens doorbladerden. De redactie hield zichzelf voor de gek. Jammer, want zo werd het bijvoorbeeld ook niet duidelijk waarom al die andere mensen het huis-aan-huisblad niet lazen.

5.3 De steekproef

Je kunt op allerlei manieren een steekproef trekken uit een populatie, maar er is slechts één manier om het goed te doen en dat is door het loten van de steekproef. Ieder persoon moet dezelfde positieve kans hebben om in de steekproef terecht te komen. En al die kansen moeten bekend zijn. Alleen dan hebben je schattingen geen systematische afwijking. En alleen dan kan je de onzekerheidsmarges van de uitkomsten uitrekenen. Die onzekerheidsmarges geven aan hoe ver je schatting maximaal van de werkelijkheid kan afliggen. Kortom, als je met je peiling uitspraken wilt doen over een populatie als geheel, dan zal je de steekproef door loting moeten trekken.

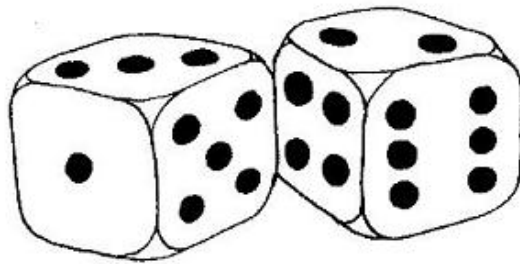
Loten met een aselector

Je hebt dus een soort lotingsmechanisme nodig dat iedereen in de populatie een even grote kans geeft om in de steekproef te komen. In de statistiek noemen we zo'n eerlijke lotingmechanisme een *aselector*. Het is een 'apparaat' met de volgende eigenschappen:

- je kunt het apparaat steeds weer opnieuw gebruiken;
- elke keer dat je het apparaat in werking stelt, produceert het één van de getallen 1 t/m N als uitkomst. Hierin is N de omvang van de populatie. We nemen aan dat de waarde van N bekend is;
- elke keer opnieuw hebben alle N mogelijke uitkomsten dezelfde kans. Kennis over eerdere uitkomsten helpt je niet bij het beter voorspellen van de volgende uitkomst. Kortom, elk voorspellingssysteem faalt.

De getallen die de aselector produceert, noemen we *aselecte getallen*. Een simpel voorbeeld van een aselector is een munt. Daarbij hebben 'kop' en 'munt' een even grote kans ($N = 2$). Een ander voorbeeld van een aselector is een dobbelsteen, waarbij de cijfers één tot en met zes ($N = 6$) elk een even grote kans hebben (als de dobbelsteen tenminste 'eerlijk' is).

5.3.1 Dobbelstenen



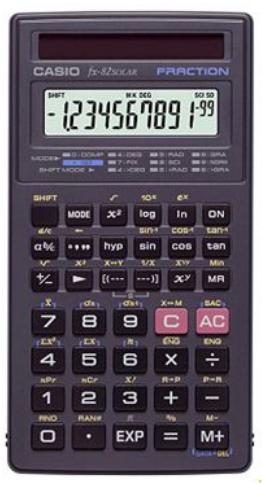
De munt is alleen te gebruiken als de populatie uit twee elementen bestaat en de dobbelsteen in principe alleen voor populaties van zes elementen. Meestal wil je een steekproef trekken uit een grotere doelpopulatie. Stel eens dat je een steekproef moet trekken van 1.000 personen uit een gemeente van 18.000 inwoners. Dobbelstenen kun je dan niet gebruiken, want er bestaan geen 18.000-zijdige dobbelstenen. Hoe moet je

het dan wel doen? We beschrijven hieronder drie technieken om 'echte' steekproeven te trekken: de rekenmachine, de tabel met aselechte getallen en een web app.

Aselechte gevallen met een zakrekenmachine

In de praktijk trek je een steekproef met een rekenmachine of een computer-programma. Die hebben vaak de mogelijkheid om willekeurige getallen te genereren uit het interval $[0, 1)$. Elke willekeurige waarde tussen 0 en 1 is mogelijk. De waarde 0 kan wel voorkomen, maar de waarde 1 net niet. Figuur 5.3.2 toont een voorbeeld van een rekenmachine met een knop voor aselechte getallen. Het is de CASIO FX-82.

5.3.2 Een zakrekenmachine met een knop voor aselechte getallen



Elke keer dat je op de knop RAN# drukt (SHIFT .), verschijnt er een willekeurige waarde uit het interval $[0, 1)$. Het op een gegeven moment twintig keer drukken op deze knop leverde de volgende waarden op:

0,360 0,319 0,778 0,753 0,521 0,652 0,609 0,812 0,057 0,756
0,205 0,465 0,023 0,128 0,394 0,381 0,802 0,031 0,415 0,065

Om een steekproef te trekken heb je willekeurige getallen nodig uit de reeks 1 t/m N , waarbij N de omvang van de populatie is. Dan heb je dus niet onmiddellijk wat aan de willekeurige waarden uit het interval $[0, 1)$. Je moet waarden omzetten in gehele getallen. Dat gaat als volgt:

1. vermenigvuldig de willekeurige waarde met de omvang van de populatie. Dat levert een waarde op in het interval $[0, N)$. De waarde 0 kan wel voorkomen en de waarde N net niet;
2. rond de verkregen waarde af naar beneden (door de cijfers achter de komma weg te laten). Dat levert een geheel getal op in de reeks van 0 tot $N-1$;
3. tel 1 op bij het verkregen getal. Dat levert een geheel getal op in de reeks van 1 t/m N .

Stel dat je een steekproef van leerlingen wilt trekken in een bepaalde school. De school heeft 642 leerlingen. Met de rekenmachine trek je dan een willekeurige waarde. Zeg dat

die gelijk is aan 0,360. Vermenigvuldigen met 642 levert dan 231,12. Afronden geeft 231. En 1 hierbij optellen geeft 232. Dus komt de leerling met volgnummer 232 in je steekproef.

Het zou kunnen gebeuren dat je zo twee of meer keren hetzelfde volgnummer krijgt. In dat geval gooi het volgnummer gewoon weg en maak je een nieuw volgnummer.

Aselecte gevallen met een tabel van aselechte getallen

Vroeger, toen er nog geen computers of rekenmachines waren, werden ook al steekproeven geloot uit populaties. Het was toen lastig om aselechte getallen te maken. Daarom waren er boeken waarin alleen maar aselechte getallen stonden. Tabel 5.3.3 bevat een stukje uit zo'n boek. In veel grotere tabel is achterin deze lesmodule te vinden.

5.3.3 Tabel met aselechte getallen

00822	63134	04080	29373	68731	34282	41827	94880	11505	07677
52659	69262	79771	19758	62062	81259	11215	42167	70001	78364
74388	10001	62523	83366	58614	41056	09869	27746	12931	93018
56160	39534	93340	87194	18428	51695	71287	49101	03330	45468
52358	62658	33674	26879	17227	49102	83879	23802	12073	76580
28601	14410	57528	04036	28540	91001	89127	94058	95697	78977
94634	76365	91906	35686	94492	14670	68416	09784	02580	49887
92168	10854	57730	62042	06833	49740	82038	84243	64678	45632
47619	03845	86451	94136	17232	09606	06571	88647	71395	87251
33459	67930	42923	05552	14955	33796	04093	78836	28134	09011

Je maakt een serie aselechte getallen door ergens een startpunt te kiezen en daarna volgens een bepaald pad door de tabel te wandelen. Je neemt dan steeds groepjes van voldoende cijfers om een getal te kunnen maken in de reeks van 1 t/m N .

Stel dat je een steekproef wilt trekken uit een school met 642 leerlingen. Dan heb je dus aselechte getallen nodig uit de reeks van 1 t/m 642. Daarvoor neem je steeds groepjes van drie cijfers. We beginnen in rij 6 en nemen steeds van elke groep van vijf cijfers de eerste groep van drie cijfers. De levert de volgende getallen op:

286, 144, 575, 040, 285, 910, 891, 940, 956, 789, 946, 763, 919, 356, 944, 146, 684, 097, 025, ...

Getallen groter dan 642 zijn niet bruikbaar. Dus die kan je weggooien. Je houdt dan de volgende getallen over:

286, 144, 575, 040, 285, 356, 146, 097, 025, ...

Mocht je zo nog nummers tegenkomen die je al eerder hebt gehad, dan gooi je die ook weg.

En zo ga je door totdat je de gewenste omvang van de steekproef hebt bereikt. Hoe groot je steekproef moet zijn wordt behandeld in paragraaf 6.3.

Aselecte getallen met een web app

Op het internet staan allerlei websites waarmee je aselechte getallen kunt maken of een steekproef kunt trekken. Eén van die websites is www.aselector.nl. Er staat een app op waarmee je aselechte getallen kunt maken.

5.3.4 Een app voor aselechte getallen

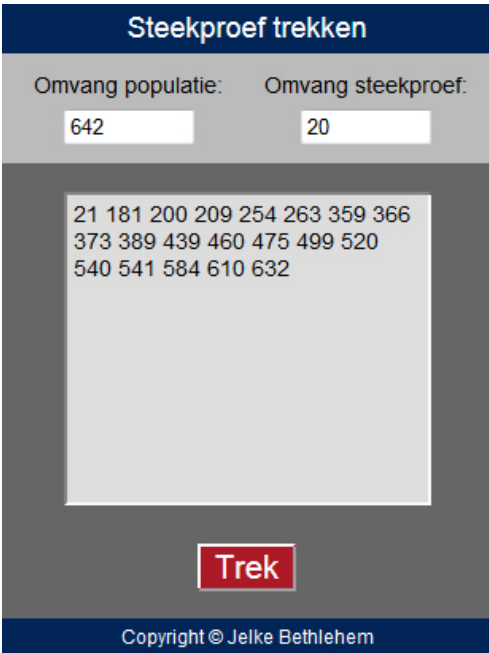


The screenshot shows a web application titled "Aselecte getallen". It has two input fields at the top: "Ondergrens:" with the value "1" and "Bovengrens:" with the value "642". In the center, a large box displays the number "451". Below this box is a red button labeled "Trek". At the bottom, there is a copyright notice: "Copyright © Jelke Bethlehem".

Het werkt heel simpel. Je voert eerst een ondergrens en een bovengrens in. Vervolgens klik je op de knop **Trek**. Elke keer dat je dat doet, verschijnt er weer een nieuw aselechte getal (zie figuur 5.3.4). De ondergrens is hier op 1 gezet en de bovengrens is 642 (de omvang van de populatie). De eerste persoon in de steekproef is persoon 451.

Op die website staat ook nog een app waarmee je in één keer een hele steekproef kunt trekken, zoals te zien in figuur 5.3.5. De omvang van de populatie is gezet op 642.

5.3.5 Een app voor het trekken van een steekproef



The screenshot shows a web application titled "Steekproef trekken". It has two input fields at the top: "Omvang populatie:" with the value "642" and "Omvang steekproef:" with the value "20". Below these fields is a large box containing a list of 20 numbers: 21 181 200 209 254 263 359 366, 373 389 439 460 475 499 520, 540 541 584 610 632. Below this box is a red button labeled "Trek". At the bottom, there is a copyright notice: "Copyright © Jelke Bethlehem".

Gevraagd is om een steekproef van 20 personen. Het onderste scherm bevat de volgnummers van de getrokken personen. Er zitten geen dubbele nummers bij en voor het gemak zijn de nummers ook nog eens in oplopende volgorde gezet.

Opgave 5.1

Een klas bestaat uit 30 leerlingen. De leerlingen zijn genummerd van 1 tot en met 30. Uit die klas moet je een steekproef van 10 leerlingen trekken. Elke leerling mag hooguit maar één keer in de steekproef komen.

Gebruik tabel 5.3.3 voor het trekken van de steekproef. Werk de tabel rij voor rij af. Neem van elk blokje van 5 cijfers steeds de laatste twee cijfers. Vind je een getal in de reeks van 1 tot en met 30, zet het dan in de tabel. Ga net zo lang door tot je 10 verschillende getallen hebt.

--	--	--	--	--	--	--	--	--	--	--	--	--	--

6.

Schatten

van

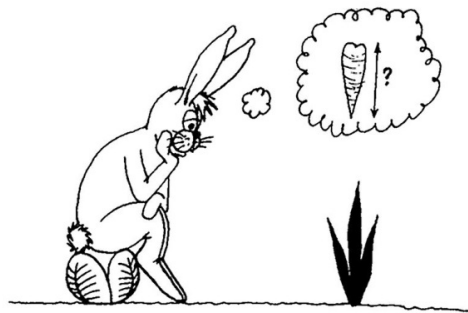
populatiekenmerken

6.1 Schatters en schattingen

De vragenlijst voor een peiling laat je invullen door alle personen in de steekproef. Als alles verder goed gaat, dan vullen die personen de vragenlijst allemaal keurig in. Zo krijg je dus de waarden van de doelvariabelen en de hulpvariabelen. Uiteraard komen de waarden van de doelvariabelen alleen beschikbaar voor de personen in de steekproef. Toch kun je op basis van deze gegevens uitspraken doen over de hele doelpopulatie. Dat is mogelijk als je de steekproef netjes door loting hebt getrokken. Die uitspraken nemen de vorm aan van schattingen van populatiekenmerken.

Voor het berekenen van een schatting gebruik je een *schat*ter. Een schatter is een *recept*. Dit recept beschrijft welke berekeningen je moet uitvoeren om een schatting te krijgen. Het recept maakt ook duidelijk welke ingrediënten nodig zijn voor de berekening. Uiteraard zijn dat de steekproefgegevens. Soms is het ook mogelijk om aanvullende informatie te gebruiken voor het berekenen van betere schattingen.

6.1.1 Schatten



Een schatter is alleen bruikbaar als hij een schatting oplevert die dicht in de buurt ligt van de waarde die je wilt schatten. Daarvoor moet een schatter aan twee eisen voldoen:

- de schatter moet *zuiver* zijn. Stel eens dat je het trekken van de steekproef een groot aantal keren zou herhalen. Dat levert elke keer een andere steekproef op. Immers, het toeval bepaalt wie er in de steekproef komt, dus levert de berekening van de schatting ook steeds weer een andere waarde op. Er is sprake van een zuivere schatter als het gemiddelde van alle mogelijke uitkomsten precies gelijk is aan de waarde die je wilt schatten. Of anders gezegd: herhaald trekken van een steekproef leidt niet tot een systematische onderschatting of overschatting. Gemiddeld genomen zal de schatting de correcte waarde moeten opleveren;
- de schatter moet ook *precies* zijn. Elke nieuwe steekproef levert een andere uitkomst voor de schatter op. Bij voorkeur moet de variatie van die schattingen zo klein mogelijk zijn. Alle mogelijke schattingen moeten zo dicht mogelijk bij elkaar in de buurt liggen.

Een schatter moet zuiver en precies zijn. Een schatter die wel zuiver is, maar niet precies, kan toevallig een waarde opleveren die ver uit de buurt ligt van de werkelijke waarde. Een schatter die wel precies is, maar niet zuiver, levert systematisch verkeerde waarden op. Als de schatter zuiver en precies is, dan noemen we dat een *nauwkeurige schatter*. Voor een

nauwkeurige schatter ligt de schatting met een zeer grote waarschijnlijkheid dicht bij de te schatten waarde van het populatiekenmerk.

In dit hoofdstuk bespreken we alleen de schatter voor een populatiepercentage; hoe groot moet je steekproef zijn om juiste conclusies te trekken over de gehele doelpopulatie? We leggen uit hoe je met die schatter een schatting moet uitrekenen. En ook leggen we uit hoe je de precisie van de schatter bepaalt.

De precisie van een schatter hangt af van de omvang van de steekproef. Als je een grotere steekproef trekt, is de schatting preciezer. We leggen uit hoe je de omvang van de steekproef berekent uitgaande van een gewenste precisie.

6.2 Het schatten van een percentage

Eerst behandelen we het schatten van een populatiepercentage. Voorbeelden hiervan zijn het percentage mensen dat in een gemeente naar de lokale omroep luistert en het percentage scholieren op een school dat wel eens wordt gepest.

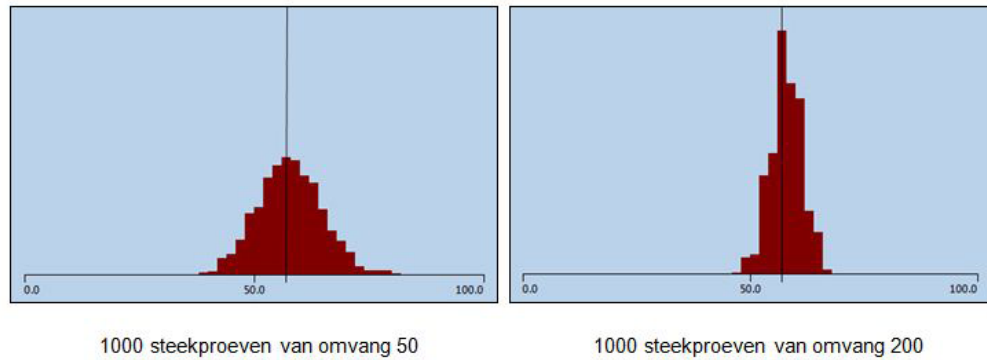
Je gaat het populatiepercentage schatten op basis van een steekproef. Je moet die steekproef hebben geloot uit de doelpopulatie. In dit geval is het *steekproefpercentage* een goede schatter voor het populatiepercentage. Wil je het percentage mensen schatten dat naar de lokale omroep luistert, dan neem je als schatter daarvoor het percentage in de steekproef dat luistert.

Het steekproefpercentage is een zuivere schatter voor het populatiepercentage. Dat kunnen we wiskundig aantonen. Er is ook nog een andere manier om dat te laten zien, en dat is door het naspelen van het trekken van een steekproef. We noemen dat een simulatie. Eerst hebben we met de computer een denkbeeldige gemeente Rijnveld gemaakt. Daarin wonen 15 000 mensen met een leeftijd van 12 jaar of ouder. Er is een populatiebestand gemaakt voor al deze 15 000 mensen. Bij 8 535 personen is vastgelegd dat ze wel eens naar de lokale omroep luisteren. De overige 6 465 personen luisteren nooit naar de lokale omroep. Dus het percentage luisteraars in de populatie is $100 \times 8\,535 / 15\,000 = 56,9$ procent. Dit percentage gaan we proberen te schatten met een steekproef.

We hebben de computer een groot aantal keren een steekproef laten trekken uit deze populatie. Voor elke steekproef hebben we geteld hoeveel mensen er naar de lokale omroep luisteren. En dit aantal hebben we steeds omgezet in een percentage. Zo hebben we dus een groot aantal schattingen gekregen voor het percentage luisteraars in de populatie. Al die schattingen kunnen we grafisch weergeven in de vorm van een histogram. In figuur 6.2.1 is dat gedaan.

De linker grafiek in figuur 6.2.1 is ontstaan door 1 000 steekproeven van omvang 50 te trekken. De verticale lijn in de grafiek geeft de populatiewaarde (56,9%) aan die we proberen te schatten. Alle blokjes liggen keurig om deze lijn verspreid. De meeste schattingen liggen in de buurt van het populatiepercentage. Soms is een schatting wat te laag, soms wat te hoog, maar gemiddeld is hij goed. Daarom kunnen we zeggen dat we hier een zuivere schatter hebben.

6.2.1 Simulatie van een enkelvoudige aselechte steekproef



De rechter grafiek laat zien wat er gebeurt als je de omvang van de steekproef verhoogt van 50 naar 200. Het verschil met de linker grafiek is dat nu de schattingen nog veel dichter in de buurt van de verticale lijn liggen. Er zit veel minder variatie in de mogelijke uitkomsten. Op basis van een steekproef van omvang 200 kun je dus een veel nauwkeuriger schatting maken. Dit is een algemeen principe bij steekproeven: hoe groter de steekproef, des te nauwkeuriger de schatter.

Het is meestal niet zo dat de schatting op grond van de steekproef precies gelijk is aan het populatiepercentage. Wel is het zo dat die schatting in de buurt ligt. Er is dus altijd een beetje onzekerheid over de echte waarde. Dat kan niet anders als je met een kleine steekproef van personen iets over alle personen in de populatie wilt zeggen. Als je de steekproef netjes hebt geloot, kun je wel uitrekenen hoe groot de onzekerheid in de schatting is. Daarvoor reken je de *onzekerheidsmarge* uit. Die onzekerheidsmarge geeft aan hoe groot de afwijking tussen je schatting en het (onbekende) populatiepercentage maximaal kan zijn. De formule voor de onzekerheidsmarge is:

$$m = 2 \times \sqrt{\frac{p \times q}{n-1}}$$

Hierin is p het steekproefpercentage, q het complementaire percentage ($100-p$) en n de omvang van de steekproef. Tabel 6.2.2 legt uit hoe je de onzekerheidsmarge stap voor stap kunt berekenen.

De berekeningen in de tabel zijn wat vereenvoudigd. Zo zou de 2 in de formule en stap 6 eigenlijk 1,96 moeten zijn. In de praktijk maakt dit weinig verschil. De laatste kolom bevat een voorbeeld. In dit voorbeeld is de omvang van de steekproef gelijk aan 300. Van die 300 personen hebben 180 personen aangegeven dat ze wel eens naar de lokale omroep luisteren. Dit is 60 procent. De laatste regel van de tabel bevat de onzekerheidsmarge. Die is aangegeven met de letter m . In het voorbeeld is de onzekerheidsmarge gelijk aan 5,6. Dit betekent dat het percentage luisteraars in de populatie nooit meer kan afwijken van de schatting dan 5,6. Dus het percentage luisteraars in de populatie is minimaal gelijk aan $60 - 5,6 = 54,4$ procent en het kan niet hoger zijn dan $60 + 5,6 = 65,6$ procent. Ofwel: tussen 54,4 procent en 65,6 procent van de inwoners van Rijnveld luistert wel eens naar de lokale omroep.

Je kunt dus een interval uitrekenen waarin het populatiepercentage moet liggen. De ondergrens van dit interval krijg je door de onzekerheidsmarge af te trekken van je

6.2.2 Berekening van de onzekerheidsmarge

Stap	Ingrediënten: n: omvang van de steekproef a: aantal mensen in de steekproef met een bepaalde eigenschap	Formule	Voorbeeld: n = 300 a = 180
Stap 1	Bereken percentage in steekproef	$p = 100 \times a / n$	$p = 100 \times 180 / 300 = 60$
Stap 2	Bereken complementair percentage	$q = 100 - p$	$q = 100 - 60 = 40$
Stap 3	Vermenigvuldig beide percentages	$r = p \times q$	$r = 60 \times 40 = 2400$
Stap 4	Deel het resultaat door de steekproefomvang minus 1	$v = r / (n - 1)$	$v = 2400 / 299 = 8$
Stap 5	Trek de wortel uit het resultaat	$s = \sqrt{v}$	$s = \sqrt{8} = 2,8$
Stap 6	Vermenigvuldig het resultaat met 2.	$m = 2 \times s$	$m = 2 \times 2,8 = 5,6$

schatting. En de bovengrens krijg je door onzekerheidsmarge op te tellen bij je schatting. In tabel 6.2.3 is dat nog eens op een rij gezet. De eerste zes stappen in deze tabel zijn hetzelfde als in tabel 6.2.2. Stappen zeven en acht zijn erbij gekomen.

6.2.3 Berekening van het betrouwbaarheidsinterval

Stap	Ingrediënten: n: omvang van de steekproef a: aantal mensen in de steekproef met een bepaalde eigenschap	Formule	Voorbeeld: n = 300 a = 180
Stap 1	Bereken percentage in steekproef	$p = 100 \times a / n$	$p = 100 \times 180 / 300 = 60$
Stap 2	Bereken complementair percentage	$q = 100 - p$	$q = 100 - 60 = 40$
Stap 3	Vermenigvuldig beide percentages	$r = p \times q$	$r = 60 \times 40 = 2400$
Stap 4	Deel het resultaat door de steekproefomvang minus 1	$v = r / (n - 1)$	$v = 2400 / 299 = 8$
Stap 5	Trek de wortel uit het resultaat	$s = \sqrt{v}$	$s = \sqrt{8} = 2,8$
Stap 6	Vermenigvuldig het resultaat met 2.	$m = 2 \times s$	$m = 2 \times 2,8 = 5,6$
Stap 7	Bereken de ondergrens van het betrouwbaarheidsinterval	$og = p - m$	$og = 60 - 5,6 = 54,4$
Stap 8	Bereken de bovengrens van het betrouwbaarheidsinterval	$bg = p + m$	$bg = 60 + 5,6 = 65,6$

Er is nog een andere manier om het steekproefpercentage en de bijbehorende onzekerheidsmarge uit te rekenen en dat is gebruik maken van een web app. Op de website www.aselector.nl staat een web app die dit voor je doet. Zie figuur 6.2.4.

6.2.4 Een app voor het berekenen van het betrouwbaarheidsinterval

The screenshot shows a web application interface with a dark blue header and footer. The main area has a light grey background. It contains three input fields for population size, sample size, and sample percentage, followed by a red 'Bereken' button. Below the button, the calculated confidence interval is displayed, including the margin of error and the lower and upper bounds.

Onzekerheidsmarge	
Omvang van de populatie:	
Omvang van de steekproef:	300
Percentage in de steekproef:	60
Bereken	
Onzekerheidsmarge:	5.5
Ondergrens betrouwbaarheidsinterval:	54.5
Bovengrens betrouwbaarheidsinterval:	65.5
Copyright © Jelke Bethlehem	

Merk op dat je bij de app eventueel ook nog de omvang van de populatie kunt invullen. De berekeningen zijn dan iets nauwkeuriger. Je mag de omvang van de populatie ook weglaten. Merk ook op dat de uitkomst in tabel 6.2.3 (een marge van 5,6%) ietsje afwijkt van de uitkomst van de app (een marge van 5,5%). Dit is het gevolg van het afronden van de uitkomsten in de tabel. In de praktijk maakt het allemaal niet zoveel uit voor de conclusies die je trekt.

We hebben gezegd dat het populatiepercentage niet meer kan afwijken van je schatting dan de onzekerheidsmarge aangeeft. En we hebben ook gezegd dat het populatiepercentage in het betrouwbaarheidsinterval moet liggen. Maar eigenlijk is dit niet helemaal waar. Als je heel toevallig een keer per ongeluk een hele rare steekproef trekt, dan zouden beide uitspraken wel eens niet waar kunnen zijn. Zo'n rare steekproef komt niet zo vaak voor. Om helemaal correct te zijn is het toch beter om te zeggen dat *met een hele grote waarschijnlijkheid* het populatiepercentage in het betrouwbaarheidsinterval ligt.

Het is belangrijk dat je bij de uitkomsten van je peiling altijd vermeldt hoe groot de onzekerheidsmarges zijn. Als je alleen maar de schattingen geeft, dan zou je kunnen denken dat de uitkomsten exact goed zijn en geen onzekerheidsmarges hebben. Je ziet het nog wel eens gebeuren bij politieke peilingen. Een partij is in de peilingen met 1 procent gestegen en legt dan op TV uit hoe het komt dat het zo goed gaat. Maar de onzekerheidsmarges van die peilingen zijn vaak 3 procent. Dat betekent dat verschillen kleiner dan 3 procent net zo goed door het toeval veroorzaakt kunnen zijn. Er is toevallig een net even andere steekproef getrokken. Er is alleen maar sprake van 'ruis'. Je kunt dus alleen maar spreken van 'echte' veranderingen als deze groter zijn dan de onzekerheidsmarge.

6.3 De omvang van de steekproef

Als je een peiling wilt uitvoeren, dan zal je van te voren moeten bedenken hoe groot de steekproef moet zijn. Dat is een belangrijke beslissing. Immers, als je de steekproef groter neemt dan echt nodig is, dan verkwist je misschien wel veel tijd en geld. En trek je een te kleine steekproef, dan zullen je schattingen minder nauwkeurig zijn dan je had gehoopt.

Het vaststellen van de omvang van de steekproef is lastig. Er is geen simpele regel die zegt hoe groot je steekproef moet zijn. Al eerder hebben we aangegeven dat er een verband bestaat tussen de omvang van de steekproef en de precisie van de schatting: hoe groter de steekproef, des te preciezer de schatting. Daarom kun je de vraag naar de omvang van de steekproef pas beantwoorden als duidelijk is welke precisie je wilt hebben voor je schattingen. Dat is dan ook de procedure die je het beste kunt volgen. Eerst stel je de precisie vast die je wilt hebben. Vervolgens kun je dan uitrekenen welke steekproefomvang hiervoor nodig is.

We beperken ons hier tot de situatie waarin we een populatiepercentage willen schatten op grond van een aselechte steekproef met gelijke kansen. De formule voor de omvang van de steekproef is

$$n = \frac{4 \times p \times q}{m \times m}$$

Je moet de uitkomst naar boven afronden op een geheel getal. In de formule is p het steekproefpercentage, q het complementaire percentage ($100-p$) en m is de onzekerheidsmarge die niet mag worden overschreden. Tabel 6.3.1 geeft aan hoe

6.3.1 Berekening van de omvang van de steekproef

Stap	Ingrediënten: m: maximaal toegestane onzekerheidsmarge. p: globale indicatie van het te schatten percentage.	Formule	Voorbeeld: m = 3 p = 50
Stap 1	Bereken complementair percentage	$q = 100 - p$	$q = 100 - 50 = 50$
Stap 2	Vermenigvuldig beide percentages	$r1 = p \times q$	$r1 = 50 \times 50 = 2500$
Stap 3	Vermenigvuldig het resultaat met 4.	$r2 = 4 \times r1$	$r2 = 4 \times 2500 = 10000$
Stap 4	Deel het resultaat door de maximale marge	$r3 = r2 / m$	$r3 = 10000 / 3 = 3333,3$
Stap 5	Deel het resultaat nog een keer door de maximale marge.	$r4 = r3 / m$	$r4 = 3333,3 / 3 = 1111,1$
Stap 6	Rond het resultaat naar boven af naar een geheel getal.	$n = \text{rondaf}(r4)$	$n = \text{rondaf}(1111,1) = 1112$

je de steekproefomvang stap voor stap kunt uitrekenen als je weet hoe groot de onzekerheidsmarge maximaal mag zijn.

De laatste kolom van tabel 6.3.1 bevat een voorbeeld. Uitgangspunt is hier dat de onzekerheidsmarge hooguit plus of min 3 procent mag zijn. Verder vermoed je dat het te schatten percentage ergens in de buurt van de 50 procent ligt. Dan leiden de berekeningen tot een omvang van de steekproef van 1 112 personen.

Een probleem bij deze berekening is dat je een globaal idee moet hebben van het populatiepercentage dat je wilt schatten. Als je echt helemaal geen enkel idee hebt van dat percentage, gebruik dan de waarde 50. De steekproefomvang voor percentages in de buurt van de 50 procent is zeker ook voldoende voor andere percentages. Je zou hooguit iets teveel personen in je steekproef kunnen trekken. Maar daar wordt de precisie alleen maar groter van.

De exacte formules voor de berekening van de steekproefomvang zijn wat ingewikkelder. Ze zijn daarom niet zo handig om met de hand uit te rekenen. De versie in tabel 6.3.1 gebruikt daarom een vereenvoudigde versie van deze formules. Op de website www.selector.nl staat een web app die de omvang uitrekent met de exacte formules (zie figuur 6.3.2).

6.3.2 Een app voor het berekenen van de omvang van de steekproef

Omvang van de steekproef

Omvang van de populatie:

Maximale onzekerheidsmarge:

Raming populatiepercentage:

Bereken

Omvang steekproef: 1068

Copyright © Jelke Bethlehem

Als je de gegevens van het voorbeeld in tabel 6.3.1 invult bij de web app, dan komt daaruit dat de steekproef minimaal een omvang van 1.068 moet hebben. Dat is dus een iets andere omvang dan de 1 112 in de tabel, maar de orde van grootte is wel ongeveer hetzelfde.

Merk op dat je bij de app eventueel nog de omvang van de populatie kunt invullen. Weet je die, vul die dan in. Dan gebruikt de app een iets nauwkeuriger formule. Weet je de omvang van de populatie niet, laat dat veld dan leeg.

Een derde manier om de omvang van de steekproef te bepalen is het gebruiken van tabel 6.3.3. Je stelt eerst vast hoe groot de onzekerheidsmarge maximaal mag zijn. Je zoekt de bijbehorende kolom op in de tabel. Vervolgens bedenkt je welke orde van grootte het te

schatten percentage ongeveer zal hebben. Je zoekt de bijbehorende rij op in de tabel. En als je het echt niet weet, dan houd je een waarde van 50 procent aan. Op het kruispunt van de gekozen rij en kolom vind je dan de omvang van de steekproef. Wil je, bijvoorbeeld, weten wat de omvang van de steekproef is voor een maximale onzekerheidsmarge van 3 procent en een populatiepercentage van 50 procent, dan vind je een waarde van 1 068. Dat is dezelfde uitkomst als die van de web app.

6.3.3 Benodigde steekproefomvang voor het schatten van een percentage

Populatie- percentage	Maximale onzekerheidsmarge (%)				
	1	2	3	4	5
5	1825	457	203	115	73
10	3458	865	385	217	139
15	4899	1225	545	307	196
20	6147	1537	683	385	246
25	7204	1801	801	451	289
30	8068	2017	897	505	323
35	8740	2185	972	547	350
40	9220	2305	1025	577	369
45	9508	2377	1057	595	381
50	9605	2402	1068	601	385
55	9508	2377	1057	595	381
60	9220	2305	1025	577	369
65	8740	2185	972	547	350
70	8068	2017	897	505	323
75	7204	1801	801	451	289
80	6147	1537	683	385	246
85	4899	1225	545	307	196
90	3458	865	385	217	139
95	1825	457	203	115	73

Marktonderzoekbureaus werken vaak met een steekproefomvang in de buurt van de 1 000 personen. Voor een dergelijke steekproef is de onzekerheidsmarge nooit meer dan 3 procent. Dan weet je dus dat de werkelijke percentages in de populatie met een grote waarschijnlijkheid niet meer dan plus of min 3 procent zullen afwijken van de schattingen in de peiling.

Opgave 6.1

In het kleine dorpje Hazenbeek is een lokale omroep. Je krijgt opdracht om uit te zoeken hoeveel procent van de inwoners vanaf 12 jaar luistert naar die omroep. De doelpopulatie bestaat uit 1 000 personen. Die hele populatie staat in bijlage 2 (Het luisteronderzoek van Hazenbeek). Voor iedere persoon in de populatie is aangegeven of deze wel of niet naar de lokale omroep luistert.

- Trek een steekproef van 50 personen uit deze populatie. Je moet daarvoor dus 50 aselechte getallen trekken uit de reeks van 1 tot en met 1 000. Je kunt dat doen met een rekenmachine, met een tabel met aselechte getallen, of met de web app. Vul de volgnummers van de getrokken personen in de tabel op de volgende pagina.

- b. Zoek elke getrokken persoon op in bijlage 2. Als een persoon naar de lokale omroep luistert, kleur dan het bijbehorende vakje in de tabel hierboven. Luistert iemand niet naar de lokale omroep, laat het vakje dan wit.
 - c. Tel het aantal gekleurde vakjes. Bereken dan het percentage personen in de steekproef dat luistert naar de lokale omroep. Let op: in dit speciale geval van een steekproef van omvang 50 is het percentage gelijk aan het aantal gekleurde vakjes vermenigvuldigd met 2.
 - d. Bereken de onzekerheidsmarge. Gebruik hiervoor het recept in tabel 6.2.2.
 - e. Bereken het betrouwbaarheidsinterval. Gebruik hiervoor het recept in tabel 6.2.3. Welke conclusie kun je trekken uit het betrouwbaarheidsinterval?
 - f. Als we je vertellen dat het werkelijke percentage luisteraars in de hele populatie gelijk was aan 57 procent, wat kun je dan zeggen als je dit percentage vergelijkt met de schatting uit je peiling?
-

7.

Het akelige probleem van de non-respons

In de voorgaande hoofdstukken hebben we beschreven hoe je een peiling moet opzetten en uitvoeren. Als je deze regels volgt, zou je een betrouwbare peiling moeten krijgen, waarmee je nauwkeurige uitspraken kunt doen over de populatie die je onderzoekt. Helaas is de praktijk vaak wat lastiger. Bij het uitvoeren van een peiling loop je vrijwel altijd tegen problemen aan. Een van de belangrijkste problemen is het optreden van non-respons. Omdat we in Nederland nogal vaak te maken hebben met non-respons en non-respons kan leiden tot verkeerde conclusies, gaan we in dit hoofdstuk wat dieper in op dit akelige probleem. We leggen uit wat non-respons is, hoe het ontstaat en wat de effecten kunnen zijn.

7.1 Non-respons in peilingen

We spreken van *non-respons* als je in een peiling de gewenste informatie niet krijgt van een persoon, terwijl die persoon toch tot de doelpopulatie van het onderzoek behoort en in de steekproef is getrokken (en dus de gegevens had moeten verstrekken).

Een eerste, voor de hand liggend, effect van non-respons is dat de gerealiseerde steekproef minder groot is dan je had gepland. Als je gegevens wilt verzamelen met een steekproef van 1 000 personen, en de helft doet niet mee, dan houd je uiteindelijk maar gegevens van 500 personen over. In principe hoeft dat niet tot onjuiste conclusies te leiden. Het betekent wel dat je steekproef kleiner is en dus dat de onzekerheidsmarge, en dus ook het betrouwbaarheidsinterval, groter is.

Je kunt al van te voren rekening houden met non-respons. Als je de gegevens van 1 000 personen wilt hebben en je verwacht dat de helft niet meedoet aan de peiling, trek dan in eerste instantie een (bruto) steekproef van 2 000 personen.

Helaas heeft non-respons meestal ook nog een ander, veel ernstiger, effect. Non-respons kan namelijk selectief zijn. Dit verschijnsel doet zich voor als, ten gevolge van non-respons, bepaalde groepen onder- of oververtegenwoordigd zijn in de peiling. Als een ondervertegenwoordigde groep zich duidelijk anders gedraagt met betrekking tot de te onderzoeken variabelen dan de oververtegenwoordigde groep, dan leidt dit tot een vertekening in de uitkomsten. Anders gezegd: een schatting valt systematisch te hoog of te laag uit.

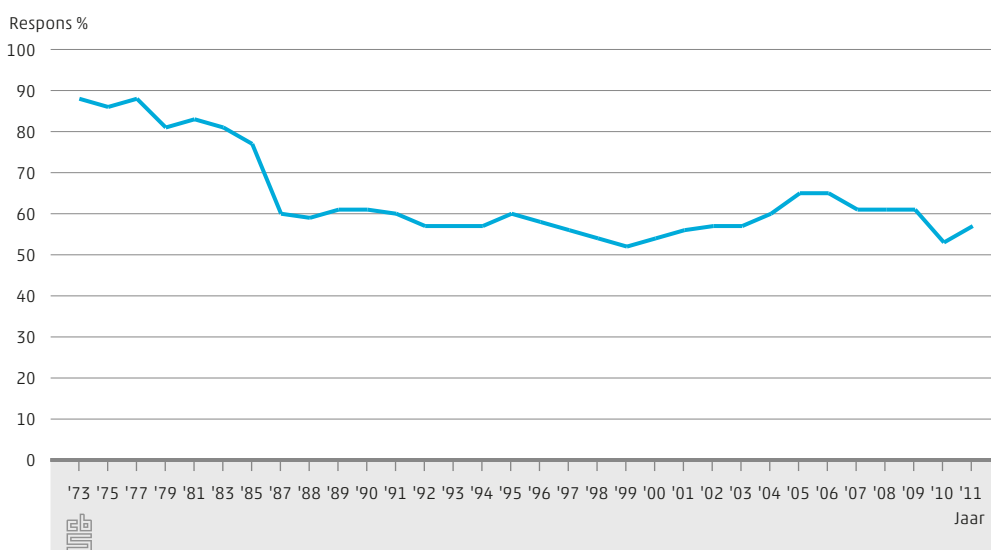
Keer op keer blijkt dat non-respons selectief is. Dat was bijvoorbeeld het geval bij een aantal onderzoeken van het Centraal Bureau voor de Statistiek (CBS):

- bij de Enquête Slachtoffers Misdrijven weigerden bange mensen vaker om mee te doen aan het onderzoek. Ze deden de deur niet open als de interviewer langs kwam. Daarmee onderschat je dus het percentage bange mensen;
- aan onderzoeken over woningbehoeften deden vooral mensen mee die niet tevreden waren met hun huidige woning. Mensen die tevreden waren met hun huis, hadden minder belangstelling voor zo'n onderzoek. De non-respons leidt hier dus tot een overschatting van ontevreden mensen;
- bij het Onderzoek Verplaatsingsgedrag bleek dat mobielere mensen (mensen die vaak onderweg zijn) ondervertegenwoordigd waren. Het was lastig om contact met ze te leggen. Ze waren immers minder vaak thuis;

- in verkiezingsonderzoek zijn de mensen die gaan stemmen altijd oververtegenwoordigd. De opkomst voor de verkiezingen wordt dan dus te hoog geschat.

Non-respons in peilingen neemt eerder toe dan af. Figuur 7.1.1 toont het verloop van het percentage respons door de jaren heen bij de Enquête Beroepsbevolking (EBB) van het CBS. Bij ander onderzoek zijn dezelfde patronen te zien. In de jaren zeventig van de vorige eeuw was de respons nog hoog (bijna 90%), maar in de loop van de tijd is de respons afgekald. Nu mogen we blij zijn met 60% respons. Daarvoor is meestal dan ook nog heel wat inspanning nodig.

7.1.1 Het percentage respons in de Enquête Beroepsbevolking (EBB)



De omvang en het effect van non-respons hangt van heel wat factoren af. Een eerste belangrijke factor is het onderwerp van de peiling. Als de benaderde personen belangstelling hebben voor het onderwerp van de peiling, zullen ze vaker geneigd zijn om mee te doen. Bij een saaie, oninteressante peiling is de kans op succes veel kleiner.

Bij peilingen onder huishoudens is het vaak zo dat één persoon in het huishouden de vragen beantwoordt. Als het er niet zoveel toe doet wie dat is, dan maakt dit de kans op respons in een groot huishouden een stuk hoger.

Ook de periode waarin je de peiling doet, speelt een rol. Je kunt maar beter de vakantieperiodes in de zomer en rondom de kerst mijden. Dan zijn veel mensen niet thuis. En als ze wel thuis zijn, hebben ze het misschien wel te druk. Het verzenden van schriftelijke vragenlijsten in de kerstperiode is ook niet zo handig. De brief kan dan makkelijker zoekraken tussen alle andere kerstpost.

De inzet van interviewers heeft heel veel invloed op het percentage respons. Bij mondelinge en telefonische peilingen liggen de responspercentages veel hoger dan bij schriftelijke peilingen of peilingen via het internet.

Non-respons kan verschillende oorzaken hebben. Het is goed de non-respons op basis van deze oorzaken in groepen te verdelen. Uit onderzoek van non-respondenten is gebleken dat de diverse groepen non-respondenten nogal kunnen verschillen. Elk type non-respons kan aanleiding geven tot een ander soort vertekening. Een goede indeling van de non-respons is dus belangrijk. Dit geldt niet alleen voor de analyse van de non-respons. Ook voor een goede verantwoording van het veldwerk is een duidelijke classificatie nuttig. Als je weet welke oorzaak van non-respons belangrijk is in je peiling, kun je er een volgende keer misschien gericht wat aan doen. De belangrijkste drie oorzaken van non-respons zijn meestal 'geen contact', 'weigering' en 'niet in staat'.

7.1.2 Oorzaken van non-respons



De allereerste stap bij het verkrijgen van respons is het contact leggen met de personen die in de steekproef zijn getrokken. Dat kan om allerlei redenen misgaan. We spreken dan van non-respons door *geen contact*. Iemand kan niet thuis zijn als de interviewer aan de deur komt. Dat kan voor korte tijd zijn (even een boodschap aan het doen), voor langere tijd (overwinteren in Spanje), of zelfs permanent (verhuisd naar een onbekend adres). Ook kan een portier of huismeester de toegang weigeren, of er lopen gemene waakhonden in de voortuin. Ook bij een telefonische peiling kan iemand niet thuis zijn. Maar de contactpoging kan ook falen omdat iemand de telefoon gewoonweg niet opneemt of in gesprek is. En bij een schriftelijke peiling kan non-respons ontstaan als het adres onjuist is of de bewoners de post meteen ongezien weggooien. Hetzelfde kan gebeuren bij online-enquêtes als het e-mailadres niet juist is of als de e-mail niet door een spamfilter komt.

Dit soort non-respons kunnen we proberen te verminderen door meer contactpogingen te doen. Bij het CBS is het niet ongebruikelijk dat bij een mondelinge peiling een interviewer zes contactpogingen doet voordat de desbetreffende persoon als non-respons wordt afgeboekt.

Is het gelukt om contact te maken met iemand, dan moeten we vaststellen of deze behoort tot de doelpopulatie. Is dat niet het geval, dan zijn we klaar. De persoon hoort dan niet thuis in de steekproef en er hoeft daarom geen vragenlijst te worden ingevuld. We kunnen deze persoon negeren als een geval van *overdekking*. Behoort de persoon wel tot de doelpopulatie, dan moeten we deze overhalen om mee te werken aan het onderzoek. Lukt dat niet dan is er sprake van non-respons als gevolg van *weigering*.

Bij weigering is het zinvol om onderscheid te maken tussen tijdelijke weigeraars en permanente weigeraars. Bij tijdelijke weigeraars komt het tijdstip van het gesprek ongelegen. De benaderde persoon heeft op dat moment geen zin of geen tijd (bezig met het klaarmaken van het eten, de baby is ziek, er is een voetbalwedstrijd op tv), maar er is een goede kans dat na het maken van een nieuwe afspraak het gesprek toch nog plaatsvindt. Bij permanente weigeraars zal het nooit tot beantwoording van de vragen komen. Permanente weigering kan zich bijvoorbeeld voordoen als de persoon het onderwerp niet aan staat, of als hij vindt dat het onderzoek teveel inbreuk doet op zijn privacy.

Een derde oorzaak van non-respons betreft personen die wel mee willen doen met het onderzoek maar het niet kunnen. Dit is de categorie 'niet in staat'. Mogelijke oorzaken zijn ziekte, dronkenschap, dementie, doofheid, blindheid of een geestelijke handicap. Ook een taalbarrière tussen de interviewer en de respondent kan de reden zijn.

7.2 Analyse van de non-respons

We moeten altijd op onze hoede zijn als we te maken krijgen met non-respons in een peiling. We moeten eerst nagaan of de non-respons zo ernstig is dat die tot een vertekening kan leiden. En als dat het geval is, dan moeten we een methode vinden om voor deze vertekening te corrigeren. Dit corrigeren is technisch nogal ingewikkeld en komt daarom in deze lesbrief niet aan de orde.

Hoe kun je in de praktijk nagaan of de non-respons selectief is? Met de beschikbare informatie over de doelvariabele (de antwoorden op de vragen aan de respondenten) kom je niet verder. Je kent immers de antwoorden van de non-respondenten niet. Dus kun je de antwoorden van de respondenten niet vergelijken met die van de non-respondenten. Toch is er een uitweg en dat is gebruik maken van hulpvariabelen. Dat zijn variabelen die je hebt gemeten in je peiling en waarover je nog extra informatie hebt, bijvoorbeeld de verdeling in de doelpopulatie. Denk aan een variabele als geslacht. Je kunt van elke respondent het geslacht registreren. Als je dan ziet dat de verhouding man/vrouw in de respons anders is dan in de populatie, dan weet je dat er iets mis is. Als het percentage mannen in de peiling bijvoorbeeld 60 procent is, dan zitten er dus teveel mannen en te weinig vrouwen in de peiling. Kennelijk responderen mannen dan beter dan vrouwen. Er is dan sprake van een verband tussen responsgedrag en geslacht. Dat leidt tot een selectieve respons.

Je moet dus op zoek naar hulpvariabelen die een verband hebben met het responsgedrag. Tref je zulke variabelen aan, dan is de respons selectief en loop je het risico dat je verkeerde conclusies trekt uit je peiling.

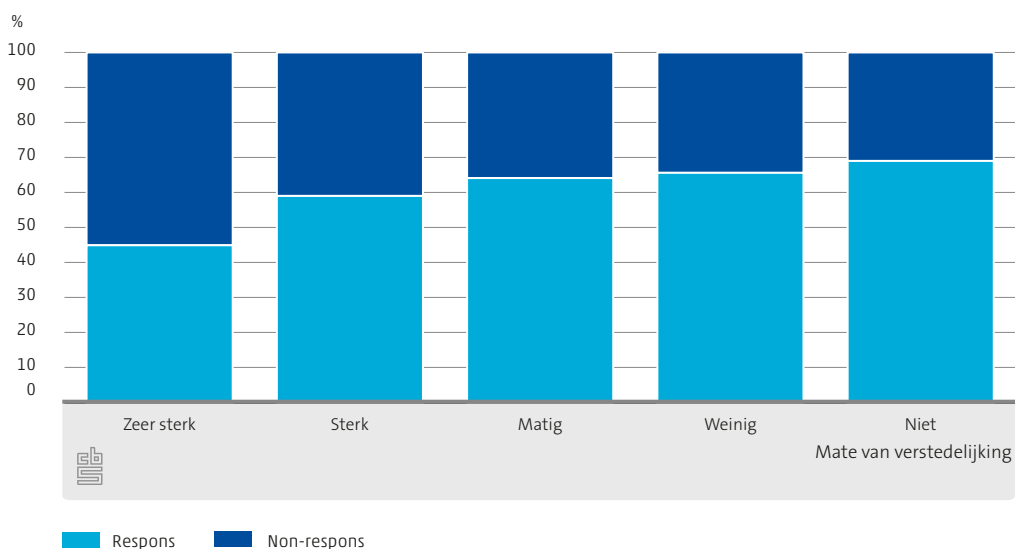
Waar haal je die hulpvariabelen vandaan? Het gaat om variabelen die je meet in je peiling en waarvoor je ook de verdeling in de populatie (of in de hele steekproef) beschikbaar hebt. Hier zijn wat mogelijke bronnen van hulpvariabelen:

- **het steekproefkader.** Soms bevat het steekproefkader allerlei variabelen. Een voorbeeld is het bevolkingsregister. Daarin zitten variabelen als geslacht, leeftijd (af te leiden uit de geboortedatum), burgerlijke staat en land van geboorte;

- **het statistisch bureau.** Het CBS kan de verdeling van heel wat variabelen leveren. Dat is natuurlijk alleen zinvol voor variabelen die over dezelfde doelpopulatie gaan. En als je een peiling op je school doet, dan heeft de administratie van de school vast wel informatie zoals het aantal jongens en meisjes op school en de leeftijdsverdeling van de leerlingen;
- **waarnemingen van interviewers.** Je kunt hierbij bijvoorbeeld denken aan een variabele als het type woning en de ouderdom van de woning. Interviewers kunnen die variabelen registreren voor zowel respondenten als non-respondenten.

Figuur 7.2.1 bevat een voorbeeld van een grafiek waarmee je kunt vaststellen of er verband is tussen responsgedrag en een hulpvariabele. De gegevens zijn afkomstig uit een peiling van het CBS. Het gaat om het Permanent Onderzoek naar de Leefsituatie (POLS) uit 1998. De hulpvariabele is de mate van verstedelijking. Die geeft aan hoe verstedelijkt de plaats is waar de respondent woont. De schaal loopt van niet verstedelijkt (het platteland) tot zeer sterk verstedelijkt (de grote steden).

7.2.1 Verband tussen responsgedrag en mate van verstedelijking

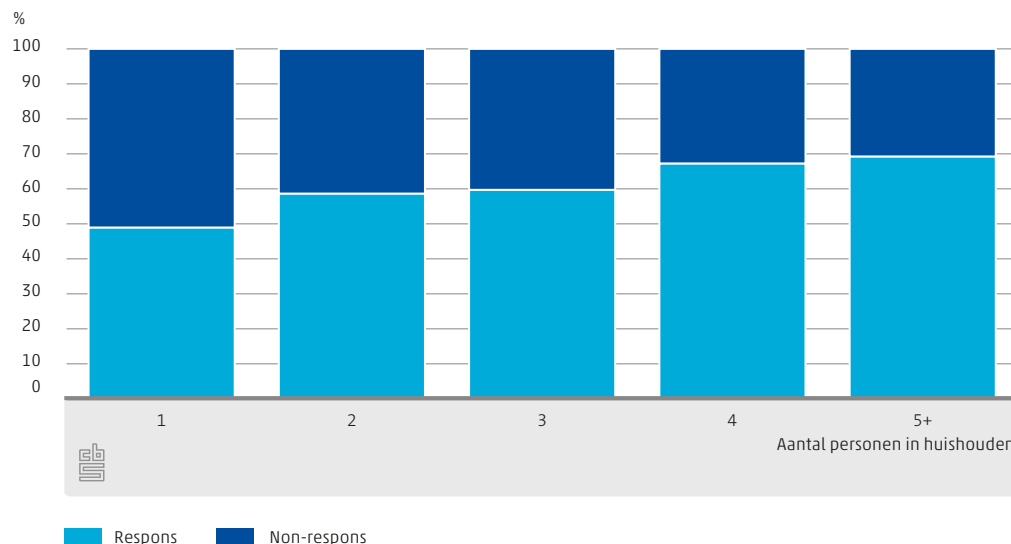


In de grafiek is duidelijk te zien dat de respons erg laag is in de grote steden en erg hoog op het platteland. Een dergelijk patroon kom je bijna wereldwijd tegen. Het is lastig om een hoge respons te halen in grote steden. Dat komt vooral omdat het moeilijk is om contact te leggen met de mensen die in de steekproef zijn getrokken.

Figuur 7.2.2 toont nog een andere hulpvariabele. Het gaat om de omvang van het huishouden waarvan de respondent deel uitmaakt.

Ook hier is een duidelijk patroon te zien. De respons neemt toe met de omvang van het huishouden. De belangrijkste oorzaak is dat het vooral bij alleenstaanden moeilijk is om contact te leggen. Ze zijn vaak niet thuis en dan kun je geen contact maken. Bij grotere huishoudens is veel vaker iemand thuis. Verder is er wat vaker non-respons bij kleinere huishoudens omdat de personen niet in staat zijn om mee te doen. Het gaat hier vooral om oudere alleenstaanden en oudere echtparen. En ook zijn er wat meer weigeringen bij kleine huishoudens.

7.2.2 Verband tussen responspercentage en omvang van het huishouden



Uit de analyse hierboven blijkt dat mensen in de grote steden en mensen in kleine huishoudens vaak ondervertegenwoordigd zijn in een peiling. Als je iets onderzoekt wat hiermee verband houdt, dan kunt je verwachten dat je schattingen een vertekend beeld geven. Er zijn veel meer hulpvariabelen die een verband hebben met responsgedrag (zie hiervoor bijvoorbeeld het boek van Bethlehem, Cobben en Schouten (2011)).

7.3 Wat te doen aan non-respons?

Non-respons is een vervelend verschijnsel dat ertoe kan leiden dat je de verkeerde conclusies uit je peiling trekt. Je peiling is niet representatief meer. Het is dus belangrijk om de non-respons in je peiling zo klein mogelijk te houden. Anders gezegd: het percentage respons moet zo hoog mogelijk zijn. Mocht het percentage respons onder de 50 procent zakken, dan moet je met ernstige problemen rekening houden.

Je moet er dus voor zorgen dat zoveel mogelijk personen in de steekproef meedoen aan je peiling. Een paar acties kunnen daarbij helpen:

- je moet uitleggen aan iedereen in de steekproef hoe belangrijk het is dat ze meedoen. Als ze niet meedoen, dan is de steekproef immers niet representatief meer;
- je moet uitleggen dat de informatie van iedereen in de steekproef vertrouwelijk is. Je maakt alleen statistieken. Je gaat geen individuele gegevens van de respondenten openbaar maken. Wat ze zeggen blijft dus geheim;
- je zou de respondenten een beloning kunnen geven voor hun meedoen. Ze krijgen allemaal een klein beetje geld of een aardig aandenken, of ze doen mee met de verloting van een prijs;
- je doet minstens zes pogingen om contact te maken met de personen in de steekproef, op verschillende dagen en tijden;
- als mensen weigeren en je hebt de indruk dat ze misschien wel om te praten zijn, probeer je het op een ander moment nog een keer;

- als mensen in de steekproef de Nederlandse taal niet machtig zijn, zou je kunnen overwegen de vragenlijsten naar andere talen te vertalen. Je kunt ook interviewers inschakelen die andere talen spreken.

Helaas zal het in de praktijk vaak zo zijn dat er, ondanks al je inspanningen, toch een behoorlijke hoeveelheid non-respons overblijft. Om dan te voorkomen dat je verkeerde conclusies trekt uit je onderzoek, moet je een correctie uitvoeren. Dat heet *wegen*. Wat je in feite doet is ondervertegenwoordigde groepen een groter gewicht geven en oververtegenwoordigde groepen een kleiner gewicht. Dat uitrekenen van die gewichten is een heel gedoe. Het voert te ver om dat in deze lesmodule allemaal uit te leggen. In de publicatie 'Deugdelijke peilingen' van Bethlehem (2013) worden de basisprincipes van wegen uitgelegd. Een uitgebreide en diepgaande behandeling van wegen staat in het (Engelstalige) boek 'Handbook of Nonresponse in Household Surveys' van Bethlehem, Cobben en Schouten (2011).

Opgave 7.1

Deze opgave is voor een flink deel hetzelfde als opgave 6.1. Er is echter een belangrijk verschil: nu treedt er non-respons op in de peiling.

In het kleine dorpje Hazenbeek is een lokale omroep. Je krijgt de opdracht om uit te zoeken hoeveel procent van de inwoners vanaf 12 jaar luistert naar de omroep. Deze populatie bestaat uit 1 000 personen. Die hele populatie staat in bijlage 3 (het luisteronderzoek met non-respons). Voor iedere persoon in de populatie is aangegeven of deze wel of niet respondeert. Non-respons is aangegeven met een vraagteken. Als iemand wel meedoet aan de peiling dan staat er of hij/zij wel of niet naar de lokale omroep luistert.

- Kopieer de steekproef van opgave 6.1 in de tabel hieronder.

- Zoek elke getrokken persoon op in bijlage 3. Als iemand niet respondeert, gum dan het nummer weg of kras het door. Als iemand wel respondeert en naar de lokale omroep luistert, kleur dan het vakje.
 - Tel het aantal respondenten. Bereken vervolgens het percentage respons in deze peiling.
 - Tel het aantal gekleurde vakjes. Bereken dan het percentage personen in de respons dat luistert naar de lokale omroep. Let op: hierbij deel je dus niet meer door 50 (de steekproef waar je mee begon), maar door het aantal respondenten (de uiteindelijke omvang van de steekproef).

- e. Bereken de onzekerheidsmarge. Gebruik hiervoor het recept in tabel 6.2.2.
 - f. Bereken het betrouwbaarheidsinterval. Gebruik hiervoor het recept in tabel 6.2.3. Welke conclusie kun je trekken uit het betrouwbaarheidsinterval?
 - g. Als we je vertellen dat het werkelijke percentage luisteraars in de hele populatie gelijk was aan 57 procent, wat kun je dan zeggen als je de schatting uit je peiling daarmee vergelijkt?
-

8.

Een checklist voor peilingen

8.1 Waarom een checklist?

Er wordt in Nederland heel veel gepeild. Dat is vooral te merken als er weer eens verkiezingen aankomen. In die campagnetijd volgen de politieke peilingen elkaar in hoog tempo op. Maar ook buiten de verkiezingen om wordt de mening van 'de Nederlander' steeds vaker gevraagd over allerlei onderwerpen.

In die peilingen wordt aan een steekproef van personen een vragenlijst voorgelegd. Het is dan de bedoeling om met de antwoorden op de vragen uitspraken te doen over de gehele populatie. Dat kan, maar dan moet het onderzoek wel op een goede manier worden opgezet en uitgevoerd.

Voorals het internet is er de oorzaak van dat het aantal peilingen sterk is toegenomen. Internet maakt het mogelijk eenvoudig, snel en goedkoop bij heel veel mensen informatie te verzamelen. De vraag is echter of al die peilingen wel een goed beeld geven van de werkelijkheid.

Er worden dus steeds meer peilingen gedaan. Daar zitten goede en slechte peilingen bij. Het is niet eenvoudig voor gebruikers van de uitkomsten van peilingen (journalisten, bestuurders, beleidsmakers) om op simpele wijze het kaf van het koren te scheiden. Daarom is een checklist gemaakt. Door de negen vragen in de checklist één voor één af te lopen, kun je een eerste indruk krijgen van hoe goed of slecht een peiling is. Als het een goede peiling is, dan kun je vertrouwen hebben in de uitkomsten ervan. Roept het doorlopen van de checklist veel vragen op, dan kun je maar beter besluiten geen aandacht te besteden aan de uitkomsten van de peiling.

Hieronder volgt een eenvoudige versie van de checklist. Hij heeft ook een uitgesproken karakter gekregen: een aanpak is goed of fout. De werkelijkheid is vaak wat ingewikkelder en dan zal je eindoordeel misschien ook wat minder hard moeten zijn.

De checklist is een gezamenlijk initiatief van het Nederlandstalig Platform voor Survey-onderzoek (NPSO), het Centraal Bureau voor de Statistiek en de Vereniging voor Onderzoeksjournalisten (VVOJ).

8.2 De checklist

1. Is duidelijk dat de opdrachtgever en/of sponsor van het onderzoek geen belang heeft bij bepaalde uitkomsten van de peiling?
 - ☐ Ja: **Ga door naar 2.**
 - ☐ Nee: **Let op!** Het kan zijn dat de uitkomsten zijn gemanipuleerd. Misschien was het wel een nep-peiling. Misschien was het persbericht over de peiling wel onderdeel van een reclamecampagne om een product of dienst te promoten ('Wij van Wc-eend adviseren Wc-eend').

2. Is er een onderzoeksrapport waarin precies staat beschreven hoe de peiling is opgezet en uitgevoerd?
- ☐ **Ja: Ga door naar 3.**
 - ☐ **Nee: Let op!** Je kunt niet vaststellen of het wel een goed onderzoek is, of dat er misschien maar wat aangerommeld is. Dus dan weet je ook niet of de conclusies van de peiling wel ergens op slaan.
3. Is duidelijk wat de doelpopulatie is? Dit is de groep personen die is onderzocht en waarop de conclusies van het onderzoek betrekking hebben.
- ☐ **Ja: Ga door naar 4.**
 - ☐ **Nee: Let op!** Je weet niet op wie de uitkomsten van toepassing zijn.
4. Is de vragenlijst opgenomen in het onderzoeksrapport?
- ☐ Ja, en de vragenlijst ziet er goed uit: **Ga door naar 5.**
 - ☐ **Ja, en de vragenlijst ziet er slecht uit: Let op!** Een vragenlijst met verkeerde of sturende vragen kan makkelijk leiden tot foute uitkomsten.
 - ☐ **Nee: Let op!** Als je niet weet hoe de vragenlijst eruit ziet, weet je ook niet of de uitkomsten wel betrouwbaar zijn.
5. Hoe is de steekproef getrokken?
- ☐ **Geloot uit de hele doelpopulatie: Ga door naar 6.**
 - ☐ **Geloot uit een deel van de doelpopulatie** (bijvoorbeeld alleen uit de internetbezitters of alleen uit personen die in het telefoonboek staan): **Ga door naar 6**, maar besef dat de uitkomsten alleen betrekking hebben op dat deel van de doelpopulatie.
 - ☐ **Zelfselectie via Internet: Let op!** Er zijn geen betrouwbare uitspraken over de doelpopulatie mogelijk.
 - ☐ **Een andere vorm van selectie van de steekproef: Let op!** Er zijn geen betrouwbare uitspraken over de doelpopulatie mogelijk.
6. Is bekend hoeveel respondenten uiteindelijk de vragenlijst hebben ingevuld?
- ☐ **Ja: Ga door naar 7.**
 - ☐ **Nee: Let op!** De onzekerheidsmarges van de uitkomsten kunnen niet worden berekend. Je weet dan dus niet hoe precies de uitkomsten zijn.
7. Is het percentage respons voldoende hoog, zeg hoger dan 50%?
- ☐ **Ja: Ga door naar 8.**
 - ☐ **Nee: Let op!** Een lage respons kan leiden tot een vertekening in de uitkomsten. De uitkomsten zijn dan dus fout.
8. Is een correctie (weging) uitgevoerd voor de opgetreden non-respons?
- ☐ **Ja: Ga door naar 9.**
 - ☐ **Nee: Let op!** Non-respons leidt vaak tot foute uitkomsten.
9. Worden onzekerheidsmarges van de uitkomsten vermeld?
- ☐ **Ja. OK.** Merk wel op dat non-respons en eventuele andere effecten (bijvoorbeeld geheugeneffecten) de onzekerheid nog kunnen vergroten.
 - ☐ **Nee. Let op!** Het is lastig om de uitkomsten op hun juiste waarde te schatten. Je kunt echte effecten niet onderscheiden van de 'ruis' van de steekproef.

8.3 Een voorbeeld: Social Media Stress

Op maandag 7 mei 2012 meldden verschillende media (NOS, Trouw, De Telegraaf) dat jongeren tussen 13 en 18 jaar lijden aan een serieuze vorm van Social Media Stress (SMS). De sociale media blijken met hun subtiele stimuli zoals geluiden, pushberichten, aandacht en beloningen jongeren in hun greep te houden. Jongeren geven aan niet meer zelfstandig te kunnen stoppen, omdat zij bang zijn buitengesloten te raken. Wanneer deze angst ernstige vormen aanneemt kan men spreken van FOMO – Fear Of Missing Out. De Nationale Academie voor Media & Maatschappij pleit voor een nieuw bewustzijn en betere begeleiding van jongeren.

8.3.1 Social Media Stress



Dit lijkt op het eerste gezicht een ernstige zaak. Maar voordat je in actie komt, is het natuurlijk wel van belang om vast te stellen waarop deze beweringen zijn gebaseerd. Dan blijkt dat het de uitkomsten zijn van een onderzoek dat is uitgevoerd door de Nationale Academie voor Media en Maatschappij. Hoe goed is dit onderzoek geweest en hoe betrouwbaar zijn de uitkomsten? We gaan dit na aan de hand van de checklist.

De opdrachtgever en uitvoerder van het onderzoek is de Nationale Academie voor Media en Maatschappij. De twee directeurs van deze stichting zijn ook de opstellers van het rapport. Ze zijn ook de hoofddocenten van deze Academie. Die academie geeft cursussen op het terrein van de sociale media, zoals bijvoorbeeld een cursus 'Social Media Professional'. De opdrachtgever van het onderzoek heeft dus belang bij de negatieve uitkomsten ervan. Daarmee is deze peiling al erg twijfelachtig.

Op de website van de Academie staat het onderzoeksrapport. De beschrijving van de opzet en uitvoering van het onderzoek is heel erg kort. De informatie is onvoldoende om te kunnen vaststellen of het om een goed uitgevoerd onderzoek gaat. Na een verzoek

daartoe per email werd meer informatie gestuurd. Toen werd ook duidelijk dat er van alles rammelde aan het onderzoek.

Er is wat onduidelijkheid over de doelpopulatie. In eerste instantie lijkt het onderzoek betrekking te hebben op jongeren tussen de 13 en 18 jaar. Het onderzoeksrapport beperkt zich echter tot jongeren in die leeftijdsgroep die de beschikking hebben over een smartphone met internetverbinding. In het rapport staat dat de helft van de jongeren zich gestrest voelt als ze merken dat ze niet alles kunnen bijhouden in de sociale media. Het maakt nogal wat uit of die 50 procent betrekking heeft op jongeren met een smartphone of op alle jongeren.

De vragenlijst is niet opgenomen in het onderzoeksrapport. Dus daarover valt weinig te zeggen. Als je de analyse van de uitkomsten leest, dan krijg je wel het gevoel dat sommige vragen erop gericht zijn om aan te tonen dat jongeren leiden aan Social Media Stress. Die vragen zijn allemaal negatief geformuleerd. Ze hebben bovendien meer negatieve categorieën dan positieve categorieën om uit te kiezen.

Uit nadere informatie van de Academie blijkt dat de steekproef niet netjes via loting is getrokken. De Academie beschikt volgens eigen zeggen over een groot netwerk van meer dan 850 jeugdprofessionals die, over het hele land verspreid, werken in het onderwijs, biblio- en mediatheek, jeugdhulpverlening of buurtwerk. Aan een selectie van deze jeugdprofessionals is gevraagd in hun eigen werkomgeving jongeren in de leeftijd 13 tot en met 17 jaar uit te nodigen om aan het onderzoek mee te werken. Er werd vooraf niet verteld waar het onderzoek over ging. Dit is een merkwaardige manier van trekken van de steekproef. Je kunt je afvragen voor welke populatie die steekproef representatief zou moeten zijn. Het netwerk van jongeren rondom jeugdprofessionals? Hoe selecteerden de jeugdprofessionals die jongeren? Waren het vooral jongeren die welwillend waren om mee te doen? Of jongeren die zo op het oog veel met sociale media bezig waren?

De steekproef bestond uit 493 jongeren (240 jongens en 253 meisjes). Dat is dus een vrij kleine steekproef. Je moet daarom rekening houden met ruime onzekerheidsmarges (zelfs als de steekproef netjes was geloot).

Het hele probleem van de non-respons is genegeerd. Wat is er gedaan met de jongeren die wel waren uitgenodigd om mee te doen aan het onderzoek, maar dat niet wilden? Is er sprake van selectieve non-respons? Waren die jongeren anders? De onderzoekers hebben de non-respons niet in beeld gebracht er al helemaal niet voor gecorrigeerd.

De onderzoekers hebben de uitkomsten wel gewogen. Ze dachten dat het dan wel representatief zou zijn. Bij navraag blijkt echter dat er alleen gewogen is naar geslacht. Dat was niet nodig, want er zaten al ongeveer evenveel jongens en meisjes in de steekproef. Dus de weging heeft geen enkel effect. Andere variabelen zouden veel meer in aanmerking komen voor wegen, zoals leeftijd en opleidingsniveau.

De onderzoekers van de Nationale Academie voor Media & Maatschappij vermelden geen onzekerheidsmarges bij hun uitkomsten. Dat kan ook niet vanwege hun krakkemikkige steekproeftrekking. Als eerste indicatie zouden minstens de marges vermeld kunnen worden die gebaseerd zijn op een aselechte steekproef.

Samenvattend zal het duidelijk zijn dat dit een slecht onderzoek is. Het is daarom jammer dat de media er zo bovenop zijn gesprongen en niet eerst wat kritischer naar het onderzoek hebben gekeken. Maar ja, daarvoor was het onderwerp natuurlijk te leuk.

8.4 Tips voor je eigen peiling

Er komt heel wat kijken bij het uitvoeren van een goede peiling. Je moet je aan allerlei regels houden. Dat is lang niet altijd eenvoudig. Ze maken een peiling tot een ingewikkelde zaak. Dat is echter toch de prijs die je moet betalen voor nauwkeurige uitkomsten die je kunt vertrouwen.

Het is soms niet mogelijk om je aan alle regels te houden. Zo komt het voor dat je geen geschikt steekproefkader hebt om de steekproef uit trekken. Of je hebt geen professionele, getrainde interviewers. Wat dan? Probeer dan toch om de regels zoveel mogelijk te volgen. We geven hier wat tips:

- Geef een duidelijke beschrijving van je doelpopulatie. Iedereen moet kunnen zien wie er wel en niet tot de doelpopulatie behoort;
- Als je geen steekproefkader hebt, probeer dan op een andere manier een representatieve steekproef te krijgen. De personen in de steekproef moeten een goed beeld geven van de hele doelpopulatie. Zorg voor steekproefpersonen uit alle hoeken en gaten van de doelpopulatie;
- Voor je peiling maak je een vragenlijst. Test die vragenlijst. Laat je klasgenoten of je familie de vragenlijst invullen en kijk of ze daarmee problemen hebben;
- Als je, bijvoorbeeld, een peiling wilt doen onder de klanten van een supermarkt, ondervraag dan elke zoveelste klant (zonder "aanzien des persoons"). Ondervraag ook personen op verschillende tijdstippen op de dag en op verschillende dagen van de week;
- Als je de inwoners van een stad wilt peilen, zorg er dan voor dat je personen uit verschillende wijken ondervraagt, van verschillende leeftijden en zowel mannen als vrouwen. Bij voorkeur moeten die allemaal in de juiste verhoudingen gepeild worden;
- Als je een peiling op je school wilt doen, zorg er dan voor dat leerlingen uit verschillende klassen in de peiling zitten. Dat moet wel in de juiste verhoudingen. Dus meer leerlingen uit grotere klassen en minder leerlingen uit kleinere klassen. Ook moet je zorgen voor jongens en meisjes in de juiste verhouding;
- Probeer de non-respons zo laag mogelijk te krijgen. Probeer alle mensen in de steekproef ervan te overtuigen dat het belangrijk is mee te doen. Vertel ook dat hun antwoorden anoniem blijven. Probeer eventueel de respons nog wat op te krikken door een prijs te verloten onder de deelnemers;
- Maak een verslag van je peiling. Leg precies uit hoe je je peiling hebt opgezet en uitgevoerd. Wees daarin helder, zodat iedereen kan zien wat de sterke en zwakke punten van je peiling zijn. Beschrijf ook duidelijk welke conclusies je hebt getrokken en wat ze waard zijn (in termen van onzekerheidsmarges).

Veel succes!

Bijlagen

9.I Aselecte getallen

00822	63134	04080	29373	68731	34282	41827	94880	11505	07677
52659	69262	79771	19758	62062	81259	11215	42167	70001	78364
74388	10001	62523	83366	58614	41056	09869	27746	12931	93018
56160	39534	93340	87194	18428	51695	71287	49101	03330	45468
52358	62658	33674	26879	17227	49102	83879	23802	12073	76580
28601	14410	57528	04036	28540	91001	89127	94058	95697	78977
94634	76365	91906	35686	94492	14670	68416	09784	02580	49887
92168	10854	57730	62042	06833	49740	82038	84243	64678	45632
47619	03845	86451	94136	17232	09606	06571	88647	71395	87251
33459	67930	42923	05552	14955	33796	04093	78836	28134	09011
06281	64451	93927	84527	34493	57413	44340	19548	60488	29672
92438	60424	52563	57719	47555	78568	40907	92697	00382	95973
19131	44132	20282	35023	07407	54700	25862	84406	64253	08580
66838	89025	50406	11036	45364	07928	47940	71568	67368	22030
38108	31249	41163	35881	08801	52655	47172	18802	86144	47848
50853	99563	10959	24640	07182	51669	26533	70680	39447	06821
98029	38791	02724	16069	76536	53330	27550	65325	63554	15772
49784	09544	58006	10788	01809	01929	00023	73879	37293	92395
71777	66623	22474	22625	78799	71411	90188	82068	87075	20740
37953	29594	11785	03936	90573	83520	45701	85588	19066	50816
18504	08794	81452	28436	00102	65663	39447	73598	44224	45996
23612	24602	97138	29350	31168	30939	32407	49658	71995	26774
17193	57903	64158	46524	68818	08583	68444	87024	54201	39109
79095	57300	38872	46020	19395	19175	42385	80834	43589	14255
06566	08437	60776	68350	06396	68017	61677	21175	28756	13754
44555	23351	33529	47794	72067	04740	61314	15531	95178	43845
63740	66617	98615	19991	82164	74064	14188	10267	67660	83166
91908	11565	31911	85953	32196	84843	62812	04801	42953	54868
92619	48338	55943	41633	38239	60828	39286	48375	32620	65542
81688	33483	83047	36563	69160	28642	12667	94792	68226	25024
48672	16112	15120	30212	06284	64246	71738	98923	98221	39368
05875	51358	95279	60078	47265	58158	48225	78500	16595	29678
74443	07361	15059	85275	79330	66264	97585	54097	17787	16872
70447	43895	46075	28628	67858	58225	51275	41234	34273	83857
84956	33743	84312	04061	15131	38032	15660	37734	36724	17422
38415	18984	74443	35103	53815	87198	16476	04135	08966	63268
98944	85367	64782	57402	51765	72364	51345	96129	97103	77778
63177	37122	34134	51062	40008	06181	72740	77599	68957	73399
11346	35310	94215	14841	08109	46889	23726	48225	97772	41375
10358	54716	08371	30751	85136	50208	61603	03168	64328	93780

22178	35777	30641	65389	91429	75867	08165	67028	53559	97963
48663	32423	69286	25864	41683	65484	40820	31868	56796	74874
08930	74628	77048	87102	75438	35606	60520	27343	76704	10328
40467	84230	28354	05374	93512	98596	74896	49773	11660	96835
33211	26107	75456	28030	01377	36545	34007	50655	58943	34207
90181	00740	37251	70239	55901	45044	70797	88513	10814	66666
05306	62410	58407	15063	37963	65107	19049	28104	26057	52707
95516	05410	98229	96625	62989	96755	26529	17719	58515	67926
56761	90183	55772	97746	78434	66889	93599	46030	82826	41083
03140	10741	12476	34408	17715	54206	84809	52726	51257	18358
42665	66658	12361	07389	03699	96423	29738	94948	59746	05307
36934	43189	73786	67705	59958	16501	22343	62698	24978	03482
57775	57213	58612	67660	81883	10143	15333	97351	40189	69167
24086	19119	92169	98851	64790	69308	43740	62466	85736	07521
30384	02060	23678	44862	10790	39888	72241	26293	15353	25621
00037	93300	99054	03855	67710	69378	56491	85867	43337	71509
37473	07650	25854	49410	85460	04800	87357	38867	77900	04561
82677	82268	16381	66469	39934	52537	98846	54728	75855	18536
52556	75997	93902	91032	39356	23335	93643	55788	02777	41250

9.II Het luisteronderzoek van Hazenbeek

De doelpopulatie van het luisteronderzoek bestaat uit 1000 personen. Achter elk volgnummer staat of die persoon wel (Ja) of niet (Nee) naar de lokale omroep luistert.

1 Ja	43 Ja	85 Nee	127 Ja	169 Ja
2 Ja	44 Nee	86 Nee	128 Nee	170 Ja
3 Ja	45 Nee	87 Nee	129 Nee	171 Nee
4 Nee	46 Ja	88 Ja	130 Nee	172 Ja
5 Ja	47 Ja	89 Nee	131 Ja	173 Ja
6 Ja	48 Nee	90 Nee	132 Nee	174 Nee
7 Ja	49 Nee	91 Ja	133 Nee	175 Nee
8 Ja	50 Ja	92 Ja	134 Nee	176 Ja
9 Nee	51 Ja	93 Ja	135 Ja	177 Nee
10 Ja	52 Nee	94 Nee	136 Nee	178 Nee
11 Ja	53 Nee	95 Nee	137 Nee	179 Nee
12 Nee	54 Ja	96 Ja	138 Ja	180 Ja
13 Ja	55 Ja	97 Ja	139 Ja	181 Ja
14 Ja	56 Ja	98 Ja	140 Ja	182 Nee
15 Ja	57 Ja	99 Nee	141 Ja	183 Nee
16 Ja	58 Nee	100 Nee	142 Ja	184 Ja
17 Ja	59 Ja	101 Ja	143 Nee	185 Ja
18 Nee	60 Nee	102 Nee	144 Nee	186 Ja
19 Ja	61 Nee	103 Ja	145 Ja	187 Ja
20 Ja	62 Nee	104 Nee	146 Ja	188 Nee
21 Ja	63 Nee	105 Nee	147 Nee	189 Nee
22 Ja	64 Ja	106 Nee	148 Ja	190 Nee
23 Ja	65 Nee	107 Nee	149 Nee	191 Ja
24 Ja	66 Nee	108 Nee	150 Nee	192 Nee
25 Ja	67 Nee	109 Nee	151 Ja	193 Ja
26 Ja	68 Nee	110 Nee	152 Ja	194 Nee
27 Nee	69 Ja	111 Ja	153 Ja	195 Ja
28 Ja	70 Nee	112 Ja	154 Nee	196 Nee
29 Ja	71 Nee	113 Ja	155 Nee	197 Ja
30 Ja	72 Ja	114 Ja	156 Nee	198 Ja
31 Ja	73 Ja	115 Nee	157 Ja	199 Nee
32 Ja	74 Ja	116 Nee	158 Ja	200 Ja
33 Nee	75 Ja	117 Nee	159 Nee	201 Nee
34 Ja	76 Ja	118 Ja	160 Nee	202 Ja
35 Nee	77 Ja	119 Nee	161 Ja	203 Ja
36 Nee	78 Nee	120 Nee	162 Ja	204 Nee
37 Ja	79 Ja	121 Nee	163 Ja	205 Ja
38 Ja	80 Ja	122 Nee	164 Nee	206 Nee
39 Ja	81 Ja	123 Ja	165 Ja	207 Nee
40 Ja	82 Ja	124 Ja	166 Nee	208 Ja
41 Nee	83 Ja	125 Nee	167 Ja	209 Ja
42 Ja	84 Ja	126 Ja	168 Nee	210 Nee

211 Ja	261 Nee	311 Ja	361 Ja	411 Nee
212 Ja	262 Nee	312 Nee	362 Ja	412 Ja
213 Nee	263 Nee	313 Ja	363 Nee	413 Ja
214 Nee	264 Ja	314 Ja	364 Ja	414 Ja
215 Nee	265 Ja	315 Ja	365 Ja	415 Nee
216 Nee	266 Nee	316 Nee	366 Ja	416 Ja
217 Ja	267 Ja	317 Ja	367 Nee	417 Nee
218 Ja	268 Ja	318 Nee	368 Ja	418 Nee
219 Nee	269 Ja	319 Nee	369 Ja	419 Ja
220 Ja	270 Nee	320 Ja	370 Ja	420 Nee
221 Nee	271 Ja	321 Nee	371 Nee	421 Ja
222 Ja	272 Ja	322 Nee	372 Ja	422 Ja
223 Ja	273 Ja	323 Nee	373 Nee	423 Ja
224 Ja	274 Ja	324 Nee	374 Ja	424 Ja
225 Nee	275 Ja	325 Nee	375 Nee	425 Ja
226 Ja	276 Nee	326 Ja	376 Nee	426 Nee
227 Ja	277 Ja	327 Nee	377 Nee	427 Nee
228 Ja	278 Ja	328 Ja	378 Ja	428 Nee
229 Ja	279 Nee	329 Nee	379 Ja	429 Nee
230 Nee	280 Ja	330 Nee	380 Ja	430 Ja
231 Ja	281 Nee	331 Ja	381 Ja	431 Ja
232 Ja	282 Ja	332 Nee	382 Nee	432 Ja
233 Ja	283 Nee	333 Nee	383 Nee	433 Nee
234 Ja	284 Ja	334 Ja	384 Nee	434 Ja
235 Nee	285 Ja	335 Ja	385 Ja	435 Nee
236 Nee	286 Ja	336 Ja	386 Nee	436 Nee
237 Ja	287 Nee	337 Ja	387 Ja	437 Nee
238 Nee	288 Nee	338 Ja	388 Ja	438 Ja
239 Ja	289 Ja	339 Ja	389 Ja	439 Ja
240 Ja	290 Ja	340 Ja	390 Nee	440 Ja
241 Nee	291 Ja	341 Ja	391 Nee	441 Ja
242 Ja	292 Nee	342 Ja	392 Nee	442 Nee
243 Nee	293 Nee	343 Ja	393 Ja	443 Nee
244 Nee	294 Ja	344 Ja	394 Ja	444 Ja
245 Ja	295 Ja	345 Ja	395 Ja	445 Nee
246 Ja	296 Nee	346 Ja	396 Ja	446 Nee
247 Ja	297 Ja	347 Nee	397 Ja	447 Nee
248 Ja	298 Ja	348 Nee	398 Ja	448 Nee
249 Nee	299 Ja	349 Ja	399 Nee	449 Nee
250 Nee	300 Ja	350 Ja	400 Nee	450 Ja
251 Ja	301 Ja	351 Ja	401 Nee	451 Ja
252 Nee	302 Ja	352 Ja	402 Nee	452 Ja
253 Nee	303 Nee	353 Ja	403 Ja	453 Ja
254 Ja	304 Nee	354 Nee	404 Nee	454 Nee
255 Ja	305 Nee	355 Nee	405 Ja	455 Nee
256 Nee	306 Ja	356 Nee	406 Ja	456 Ja
257 Nee	307 Nee	357 Ja	407 Ja	457 Nee
258 Ja	308 Ja	358 Ja	408 Nee	458 Ja
259 Ja	309 Ja	359 Ja	409 Nee	459 Nee
260 Ja	310 Nee	360 Ja	410 Nee	460 Ja

461 Ja	511 Nee	561 Ja	611 Ja	661 Ja
462 Nee	512 Nee	562 Ja	612 Ja	662 Ja
463 Ja	513 Ja	563 Ja	613 Ja	663 Ja
464 Nee	514 Ja	564 Nee	614 Ja	664 Nee
465 Ja	515 Ja	565 Nee	615 Nee	665 Ja
466 Ja	516 Ja	566 Ja	616 Ja	666 Ja
467 Nee	517 Nee	567 Nee	617 Nee	667 Ja
468 Ja	518 Ja	568 Nee	618 Ja	668 Ja
469 Ja	519 Nee	569 Nee	619 Ja	669 Ja
470 Nee	520 Ja	570 Ja	620 Nee	670 Nee
471 Ja	521 Ja	571 Ja	621 Nee	671 Ja
472 Nee	522 Ja	572 Ja	622 Ja	672 Ja
473 Ja	523 Ja	573 Ja	623 Ja	673 Ja
474 Ja	524 Ja	574 Nee	624 Ja	674 Nee
475 Ja	525 Nee	575 Ja	625 Nee	675 Ja
476 Nee	526 Nee	576 Nee	626 Ja	676 Ja
477 Nee	527 Ja	577 Nee	627 Nee	677 Nee
478 Ja	528 Nee	578 Ja	628 Nee	678 Nee
479 Ja	529 Nee	579 Nee	629 Nee	679 Ja
480 Ja	530 Nee	580 Nee	630 Ja	680 Nee
481 Nee	531 Nee	581 Ja	631 Ja	681 Nee
482 Ja	532 Nee	582 Nee	632 Nee	682 Ja
483 Ja	533 Ja	583 Nee	633 Nee	683 Nee
484 Nee	534 Nee	584 Nee	634 Ja	684 Nee
485 Ja	535 Nee	585 Ja	635 Nee	685 Nee
486 Ja	536 Ja	586 Nee	636 Nee	686 Nee
487 Ja	537 Nee	587 Nee	637 Ja	687 Ja
488 Ja	538 Ja	588 Ja	638 Nee	688 Nee
489 Nee	539 Ja	589 Nee	639 Nee	689 Ja
490 Ja	540 Nee	590 Nee	640 Nee	690 Nee
491 Ja	541 Ja	591 Nee	641 Ja	691 Ja
492 Nee	542 Nee	592 Ja	642 Ja	692 Nee
493 Nee	543 Nee	593 Nee	643 Nee	693 Ja
494 Nee	544 Ja	594 Ja	644 Nee	694 Ja
495 Nee	545 Nee	595 Ja	645 Nee	695 Ja
496 Ja	546 Nee	596 Ja	646 Nee	696 Ja
497 Nee	547 Nee	597 Ja	647 Nee	697 Nee
498 Ja	548 Nee	598 Nee	648 Ja	698 Ja
499 Ja	549 Nee	599 Nee	649 Nee	699 Ja
500 Ja	550 Nee	600 Ja	650 Nee	700 Nee
501 Ja	551 Nee	601 Ja	651 Nee	701 Nee
502 Ja	552 Nee	602 Ja	652 Nee	702 Nee
503 Ja	553 Ja	603 Ja	653 Nee	703 Ja
504 Nee	554 Ja	604 Nee	654 Ja	704 Ja
505 Ja	555 Nee	605 Ja	655 Nee	705 Nee
506 Ja	556 Ja	606 Ja	656 Nee	706 Nee
507 Ja	557 Ja	607 Ja	657 Nee	707 Ja
508 Nee	558 Nee	608 Ja	658 Ja	708 Nee
509 Ja	559 Nee	609 Ja	659 Ja	709 Ja
510 Ja	560 Nee	610 Ja	660 Nee	710 Ja

711 Ja	761 Ja	811 Nee	861 Ja	911 Ja
712 Ja	762 Ja	812 Ja	862 Nee	912 Ja
713 Nee	763 Nee	813 Ja	863 Nee	913 Ja
714 Ja	764 Ja	814 Ja	864 Nee	914 Nee
715 Ja	765 Nee	815 Nee	865 Nee	915 Nee
716 Ja	766 Ja	816 Ja	866 Nee	916 Nee
717 Nee	767 Ja	817 Nee	867 Nee	917 Ja
718 Ja	768 Ja	818 Ja	868 Nee	918 Ja
719 Ja	769 Ja	819 Nee	869 Ja	919 Ja
720 Nee	770 Nee	820 Nee	870 Ja	920 Nee
721 Nee	771 Nee	821 Nee	871 Nee	921 Ja
722 Ja	772 Ja	822 Ja	872 Ja	922 Ja
723 Ja	773 Nee	823 Ja	873 Nee	923 Ja
724 Ja	774 Ja	824 Ja	874 Nee	924 Ja
725 Nee	775 Ja	825 Ja	875 Nee	925 Ja
726 Ja	776 Ja	826 Nee	876 Nee	926 Ja
727 Ja	777 Ja	827 Nee	877 Ja	927 Ja
728 Nee	778 Ja	828 Nee	878 Ja	928 Nee
729 Nee	779 Ja	829 Ja	879 Ja	929 Nee
730 Nee	780 Nee	830 Ja	880 Ja	930 Ja
731 Ja	781 Nee	831 Ja	881 Ja	931 Nee
732 Nee	782 Ja	832 Ja	882 Ja	932 Ja
733 Ja	783 Ja	833 Nee	883 Nee	933 Ja
734 Ja	784 Nee	834 Ja	884 Nee	934 Nee
735 Nee	785 Ja	835 Nee	885 Ja	935 Ja
736 Nee	786 Ja	836 Ja	886 Nee	936 Nee
737 Ja	787 Ja	837 Nee	887 Ja	937 Ja
738 Ja	788 Ja	838 Ja	888 Nee	938 Ja
739 Ja	789 Ja	839 Ja	889 Ja	939 Ja
740 Ja	790 Ja	840 Nee	890 Ja	940 Ja
741 Nee	791 Ja	841 Ja	891 Nee	941 Ja
742 Nee	792 Nee	842 Ja	892 Ja	942 Ja
743 Ja	793 Nee	843 Ja	893 Ja	943 Nee
744 Ja	794 Nee	844 Ja	894 Ja	944 Ja
745 Nee	795 Ja	845 Ja	895 Nee	945 Nee
746 Nee	796 Ja	846 Ja	896 Ja	946 Nee
747 Ja	797 Ja	847 Ja	897 Ja	947 Nee
748 Ja	798 Nee	848 Ja	898 Nee	948 Nee
749 Nee	799 Ja	849 Ja	899 Nee	949 Ja
750 Ja	800 Ja	850 Nee	900 Nee	950 Nee
751 Ja	801 Ja	851 Ja	901 Ja	951 Ja
752 Nee	802 Nee	852 Nee	902 Nee	952 Nee
753 Nee	803 Ja	853 Nee	903 Ja	953 Ja
754 Nee	804 Nee	854 Ja	904 Ja	954 Ja
755 Ja	805 Ja	855 Ja	905 Ja	955 Ja
756 Nee	806 Ja	856 Ja	906 Nee	956 Ja
757 Ja	807 Nee	857 Nee	907 Ja	957 Ja
758 Ja	808 Nee	858 Ja	908 Nee	958 Nee
759 Nee	809 Nee	859 Ja	909 Ja	959 Ja
760 Ja	810 Ja	860 Nee	910 Nee	960 Ja

961 Ja	969 Ja	977 Ja	985 Ja	993 Nee
962 Ja	970 Nee	978 Nee	986 Nee	994 Ja
963 Ja	971 Nee	979 Nee	987 Ja	995 Ja
964 Ja	972 Ja	980 Nee	988 Ja	996 Ja
965 Ja	973 Ja	981 Nee	989 Nee	997 Ja
966 Nee	974 Nee	982 Ja	990 Nee	998 Ja
967 Nee	975 Ja	983 Ja	991 Ja	999 Nee
968 Nee	976 Nee	984 Ja	992 Nee	1000 Nee

9.III Het luisteronderzoek (met non-respons)

De doelpopulatie van het luisteronderzoek bestaat uit 1000 personen. Achter elk volgnummer staat of die persoon wel (Ja) of niet (Nee) naar de lokale omroep luistert. Er zijn heel wat personen die niet responderen. Dan staat er een vraagteken (?) achter het volgnummer.

1 ?	41 ?	81 Ja	121 ?	161 Ja
2 Ja	42 ?	82 Ja	122 ?	162 ?
3 Ja	43 Ja	83 ?	123 Ja	163 Ja
4 ?	44 ?	84 Ja	124 ?	164 Nee
5 Ja	45 ?	85 ?	125 ?	165 Ja
6 Ja	46 Ja	86 ?	126 Ja	166 Nee
7 Ja	47 ?	87 Nee	127 Ja	167 Ja
8 ?	48 ?	88 ?	128 Nee	168 ?
9 ?	49 Nee	89 ?	129 ?	169 Ja
10 ?	50 Ja	90 Nee	130 ?	170 ?
11 ?	51 Ja	91 Ja	131 ?	171 ?
12 ?	52 Nee	92 Ja	132 ?	172 Ja
13 Ja	53 ?	93 Ja	133 Nee	173 Ja
14 Ja	54 Ja	94 ?	134 ?	174 Nee
15 ?	55 Ja	95 Nee	135 Ja	175 Nee
16 Ja	56 ?	96 Ja	136 ?	176 ?
17 Ja	57 Ja	97 ?	137 Nee	177 ?
18 ?	58 Nee	98 ?	138 Ja	178 Nee
19 Ja	59 Ja	99 Nee	139 Ja	179 Nee
20 ?	60 ?	100 ?	140 ?	180 ?
21 ?	61 ?	101 Ja	141 Ja	181 ?
22 ?	62 ?	102 ?	142 Ja	182 ?
23 Ja	63 ?	103 Ja	143 Nee	183 ?
24 Ja	64 Ja	104 Nee	144 Nee	184 ?
25 Ja	65 ?	105 ?	145 Ja	185 Ja
26 Ja	66 ?	106 Nee	146 ?	186 ?
27 Nee	67 Nee	107 Nee	147 ?	187 Ja
28 ?	68 ?	108 Nee	148 Ja	188 ?
29 Ja	69 ?	109 ?	149 Nee	189 ?
30 Ja	70 ?	110 Nee	150 ?	190 Nee
31 Ja	71 Nee	111 Ja	151 Ja	191 ?
32 Ja	72 Ja	112 Ja	152 Ja	192 ?
33 Nee	73 Ja	113 Ja	153 Ja	193 Ja
34 ?	74 ?	114 ?	154 Nee	194 Nee
35 ?	75 Ja	115 Nee	155 Nee	195 Ja
36 ?	76 ?	116 Nee	156 ?	196 ?
37 Ja	77 Ja	117 ?	157 Ja	197 ?
38 ?	78 ?	118 Ja	158 Ja	198 Ja
39 ?	79 ?	119 ?	159 ?	199 Nee
40 Ja	80 ?	120 ?	160 Nee	200 Ja

201 Nee	251 Ja	301 ?	350 Ja	400 ?
202 Ja	252 ?	302 ?	351 Ja	401 ?
203 Ja	253 ?	303 ?	352 ?	402 Nee
204 Nee	254 Ja	304 Nee	353 Ja	403 Ja
205 ?	255 ?	305 ?	354 ?	404 ?
206 ?	256 ?	306 ?	355 ?	405 Ja
207 ?	257 Nee	307 ?	356 Nee	406 Ja
208 Ja	258 Ja	308 ?	357 ?	407 ?
209 Ja	259 Ja	309 ?	358 Ja	408 Nee
210 ?	260 ?	310 ?	359 Ja	409 ?
211 Ja	261 ?	311 Ja	360 ?	410 ?
212 Ja	262 ?	312 ?	361 ?	411 ?
213 ?	263 ?	313 Ja	362 ?	412 ?
214 Nee	264 Ja	314 Ja	363 Nee	413 Ja
215 Nee	265 Ja	315 Ja	364 ?	414 Ja
216 Nee	266 Nee	316 Nee	365 Ja	415 Nee
217 Ja	267 ?	317 ?	366 ?	416 Ja
218 Ja	268 Ja	318 ?	367 Nee	417 ?
219 Nee	269 Ja	319 ?	368 ?	418 Nee
220 Ja	270 ?	320 Ja	369 ?	419 ?
221 Nee	271 Ja	321 Nee	370 Ja	420 ?
222 ?	272 Ja	322 ?	371 ?	421 ?
223 Ja	273 ?	323 Nee	372 ?	422 Ja
224 ?	274 ?	324 Nee	373 Nee	423 ?
225 ?	275 ?	325 ?	374 Ja	424 ?
226 Ja	276 ?	326 ?	375 ?	425 ?
227 Ja	277 Ja	327 ?	376 ?	426 ?
228 Ja	278 ?	328 Ja	377 ?	427 ?
229 ?	279 ?	329 ?	378 Ja	428 ?
230 ?	280 Ja	330 ?	379 ?	429 ?
231 Ja	281 ?	331 ?	380 ?	430 Ja
232 Ja	282 Ja	332 ?	381 Ja	431 Ja
233 Ja	283 ?	333 ?	382 Nee	432 ?
234 ?	284 Ja	334 ?	383 Nee	433 Nee
235 ?	285 ?	335 Ja	384 ?	434 ?
236 ?	286 ?	336 Ja	385 ?	435 Nee
237 Ja	287 ?	337 Ja	386 ?	436 ?
238 ?	288 ?	338 Ja	387 Ja	437 ?
239 ?	289 Ja	339 Ja	388 Ja	438 ?
240 Ja	290 Ja	340 Ja	389 ?	439 Ja
241 ?	291 Ja		390 ?	440 Ja
242 ?	292 ?	341 Ja	391 ?	441 Ja
243 ?	293 ?	342 ?	392 Nee	442 Nee
244 Nee	294 Ja	343 Ja	393 Ja	443 Nee
245 ?	295 Ja	344 Ja	394 Ja	444 Ja
246 Ja	296 ?	345 Ja	395 ?	445 ?
247 ?	297 ?	346 ?	396 Ja	446 ?
248 ?	298 Ja	347 ?	397 ?	447 Nee
249 ?	299 Ja	348 ?	398 Ja	448 ?
250 Nee	300 Ja	349 Ja	399 ?	449 Nee

450 ?	500 Ja	550 Nee	600 ?	650 ?
451 Ja	501 Ja	551 ?	601 ?	651 Nee
452 Ja	502 Ja	552 ?	602 Ja	652 ?
453 ?	503 Ja	553 ?	603 ?	653 ?
454 ?	504 ?	554 ?	604 Nee	654 ?
455 ?	505 Ja	555 ?	605 ?	655 ?
456 ?	506 Ja	556 Ja	606 Ja	656 ?
457 ?	507 ?	557 Ja	607 ?	657 Nee
458 Ja	508 Nee	558 ?	608 Ja	658 Ja
459 ?	509 Ja	559 ?	609 Ja	659 ?
460 Ja	510 Ja	560 Nee	610 Ja	660 ?
461 Ja	511 Nee	561 Ja	611 ?	661 Ja
462 ?	512 Nee	562 ?	612 Ja	662 Ja
463 ?	513 ?	563 Ja	613 Ja	663 Ja
464 ?	514 Ja	564 ?	614 Ja	664 ?
465 ?	515 Ja	565 Nee	615 Nee	665 ?
466 Ja	516 ?	566 Ja	616 Ja	666 Ja
467 ?	517 ?	567 ?	617 ?	667 Ja
468 ?	518 Ja	568 ?	618 ?	668 ?
469 Ja	519 ?	569 ?	619 ?	669 ?
470 ?	520 Ja	570 Ja	620 Nee	670 Nee
471 Ja	521 Ja	571 ?	621 ?	671 ?
472 Nee	522 Ja	572 Ja	622 Ja	672 ?
473 ?	523 ?	573 Ja	623 Ja	673 Ja
474 Ja	524 Ja	574 ?	624 Ja	674 Nee
475 Ja	525 Nee	575 Ja	625 ?	675 Ja
476 ?	526 ?	576 ?	626 ?	676 ?
477 ?	527 Ja	577 Nee	627 ?	677 Nee
478 Ja	528 ?	578 ?	628 Nee	678 ?
479 Ja	529 ?	579 Nee	629 ?	679 Ja
480 Ja	530 Nee	580 ?	630 Ja	680 Nee
481 ?	531 Nee	581 Ja	631 Ja	681 ?
482 Ja	532 ?	582 ?	632 Nee	682 Ja
483 Ja	533 ?	583 ?	633 ?	683 Nee
484 ?	534 ?	584 ?	634 Ja	684 ?
485 Ja	535 ?	585 ?	635 ?	685 ?
486 Ja	536 Ja	586 Nee	636 ?	686 ?
487 Ja	537 ?	587 ?	637 Ja	687 Ja
488 Ja	538 ?	588 Ja	638 ?	688 Nee
489 Nee	539 ?	589 Nee	639 ?	689 ?
490 Ja	540 ?	590 ?	640 ?	690 ?
491 ?	541 ?	591 ?	641 ?	691 ?
492 Nee	542 Nee	592 Ja	642 ?	692 Nee
493 ?	543 Nee	593 ?	643 ?	693 Ja
494 ?	544 ?	594 ?	644 ?	694 Ja
495 ?	545 Nee	595 ?	645 Nee	695 Ja
496 Ja	546 ?	596 ?	646 ?	696 ?
497 ?	547 ?	597 Ja	647 Nee	697 ?
498 Ja	548 ?	598 Nee	648 Ja	698 Ja
499 ?	549 ?	599 Nee	649 ?	699 Ja

700 ?	750 Ja	800 ?	850 ?	900 ?
701 ?	751 Ja	801 ?	851 Ja	901 ?
702 Nee	752 Nee	802 ?	852 ?	902 ?
703 Ja	753 ?	803 ?	853 ?	903 Ja
704 Ja	754 ?	804 Nee	854 Ja	904 Ja
705 ?	755 ?	805 Ja	855 Ja	905 Ja
706 ?	756 ?	806 Ja	856 Ja	906 ?
707 ?	757 Ja	807 Nee	857 ?	907 Ja
708 ?	758 Ja	808 Nee	858 ?	908 ?
709 ?	759 ?	809 Nee	859 Ja	909 ?
710 Ja	760 Ja	810 Ja	860 ?	910 ?
711 Ja	761 ?	811 ?	861 ?	911 Ja
712 ?	762 Ja	812 ?	862 Nee	912 ?
713 ?	763 Nee	813 Ja	863 ?	913 Ja
714 Ja	764 Ja	814 ?	864 ?	914 ?
715 ?	765 ?	815 ?	865 Nee	915 ?
716 Ja	766 ?	816 ?	866 Nee	916 Nee
717 ?	767 Ja	817 ?	867 ?	917 ?
718 Ja	768 ?	818 ?	868 ?	918 Ja
719 ?	769 Ja	819 ?	869 ?	919 Ja
720 ?	770 Nee	820 ?	870 Ja	920 ?
721 Nee	771 Nee	821 ?	871 Nee	921 Ja
722 Ja	772 ?	822 Ja	872 ?	922 Ja
723 Ja	773 ?	823 ?	873 Nee	923 Ja
724 ?	774 ?	824 Ja	874 ?	924 Ja
725 ?	775 Ja	825 Ja	875 ?	925 Ja
726 Ja	776 Ja	826 ?	876 ?	926 ?
727 Ja	777 ?	827 Nee	877 ?	927 Ja
728 ?	778 ?	828 ?	878 Ja	928 Nee
729 ?	779 Ja	829 ?	879 Ja	929 ?
730 ?	780 ?	830 ?	880 Ja	930 Ja
731 Ja	781 ?	831 Ja	881 ?	931 Nee
732 ?	782 ?	832 ?	882 Ja	932 ?
733 Ja	783 Ja	833 Nee	883 Nee	933 Ja
734 Ja	784 ?	834 Ja	884 Nee	934 ?
735 ?	785 Ja	835 Nee	885 Ja	935 Ja
736 ?	786 Ja	836 Ja	886 Nee	936 ?
737 ?	787 Ja	837 ?	887 Ja	937 Ja
738 Ja	788 ?	838 Ja	888 Nee	938 Ja
739 Ja	789 Ja	839 Ja	889 ?	939 Ja
740 ?	790 ?	840 ?	890 Ja	940 Ja
741 Nee	791 ?	841 Ja	891 ?	941 ?
742 ?	792 ?	842 ?	892 ?	942 Ja
743 ?	793 Nee	843 ?	893 Ja	943 ?
744 Ja	794 Nee	844 Ja	894 Ja	944 Ja
745 ?	795 ?	845 Ja	895 Nee	945 ?
746 Nee	796 Ja	846 Ja	896 Ja	946 ?
747 ?	797 Ja	847 ?	897 Ja	947 ?
748 Ja	798 Nee	848 Ja	898 ?	948 ?
749 Nee	799 Ja	849 Ja	899 Nee	949 Ja

950 ?	961 Ja	972 ?	983 ?	994 Ja
951 Ja	962 ?	973 Ja	984 Ja	995 ?
952 ?	963 ?	974 Nee	985 Ja	996 Ja
953 Ja	964 Ja	975 Ja	986 Nee	997 Ja
954 Ja	965 Ja	976 ?	987 ?	998 Ja
955 Ja	966 Nee	977 ?	988 ?	999 ?
956 Ja	967 Nee	978 ?	989 ?	1000 Nee
957 ?	968 ?	979 ?	990 ?	
958 ?	969 Ja	980 Nee	991 ?	
959 ?	970 ?	981 ?	992 Nee	
960 Ja	971 ?	982 ?	993 ?	

Begrippenlijst

Betrouwbaar

Een schatter is betrouwbaar als herhaald gebruik in dezelfde situatie ongeveer dezelfde schattingen oplevert. De verschillende schattingen liggen dicht bij elkaar.

Doelpopulatie

De doelpopulatie is de groep mensen die je onderzoekt. Het is de groep waaruit je de steekproef trekt en het is ook de groep waarop de uitkomsten betrekking hebben.

Doelvariabele

Een doelvariabele is een belangrijke variabele die meet waar we op uit zijn in een peiling. De doelvariabelen samen meten allerlei aspecten van het verschijnsel dat we onderzoeken.

Hulpvariabele

Een hulpvariabele meet een achtergrondkenmerk van personen. Voorbeelden zijn geslacht, leeftijd, burgerlijke staat en provincie.

Nauwkeurig

Een schatter is nauwkeurig als hij zowel valide als betrouwbaar is. Een nauwkeurige schatter levert een schatting op die in de buurt ligt van het te schatten populatiekenmerk.

Non-respons

Non-respons is het verschijnsel dat de vragenlijst van een persoon in de steekproef leeg blijft. Er komen geen antwoorden op vragen. Oorzaken kunnen zijn: geen contact, weigering of niet in staat zijn om de vragenlijst te beantwoorden.

Onzekerheidsmarge

De onzekerheidsmarge geeft aan hoeveel de schatting maximaal (met grote waarschijnlijkheid) kan afwijken van het te schatten populatiekenmerk.

Populatiekenmerk

Een populatiekenmerk is een grootte die je kunt uitrekenen als je de waarden van een variabele kent voor iedereen in de doelpopulatie. Voorbeelden van populatiekenmerken zijn totalen, gemiddelden, aantallen en percentages.

Precies

Een schatter is precies als er weinig variatie zit in de mogelijke schattingen. Een betrouwbare schatter is precies.

Primacy effect

Het primacy effect komt voor bij peilingen waarbij de respondenten zelf (zonder interviewer) de vragenlijst invullen. Het is de neiging te kiezen voor een antwoord ergens vooraan in de lijst van mogelijke antwoorden.

Recency effect

Het recency effect komt voor bij peilingen waarbij interviewers de vragen stellen. Het is de neiging te kiezen voor een antwoord ergens achteraan in de lijst van mogelijke antwoorden.

Representatief

Een peiling is representatief als de verdeling van alle variabelen in de respons overeenkomt met de verdeling in de doelpopulatie. Bijvoorbeeld, de verhouding man/vrouw in de peiling moet gelijk zijn aan de verhouding man/vrouw in de doelpopulatie.

Sociaalwenselijk

Respondenten kunnen een sociaalwenselijk antwoord op een vraag geven. Dat gebeurt vooral bij vragen over gevoelige onderwerpen. De respondent geeft een algemeen geaccepteerd antwoord in plaats van zijn 'eigen' antwoord.

Valide

Een schatter is valide als hij schat wat hij zou moeten schatten. Er is geen vertekening. Een valide schatter is een zuivere schatter.

Variabele

Een variabele is een eigenschap van personen die je in een peiling meet. Het heet een variabele omdat de eigenschap voor elke persoon anders kan zijn en dus kan variëren.

Zuiver

Een zuivere schatter is een schatter die gemiddeld genomen de juiste waarde (het te schatten populatiekenmerk) oplevert. De schatter heeft geen systematische afwijking. Een zuivere schatter is valide.

Gebruikte bronnen

Bethlehem, J.G. (2009), *Applied Survey Methods – A Statistical Perspective*. John Wiley & Sons, Hoboken, NJ.

Bethlehem, J.G. (2013), *Deugdelijke peilingen*. Discussion Paper 2030|10, Centraal Bureau voor de Statistiek, Den Haag/Heerlen.

Bethlehem, J.G. (2012), *Peilingen beoordelen – Een checklist*. Centraal Bureau voor de Statistiek, Den Haag/Heerlen.

Bethlehem, J.G., Cobben, F. & en Schouten, B. (2011). *Handbook of Nonresponse in Household Surveys*. John Wiley & Sons, Hoboken, NJ.

Gebruikte illustraties

Figuur 1.1.1: Het EenVandaag Opiniepanel, met presentator Gijs Rademaker, www.eenvandaag.nl, januari 2014.

Figuur 1.1.2: Een poll op de website Girlscene.nl, www.girlscene.nl, februari 2014.
Het EenVandaag Opiniepanel, met presentator Gijs Rademaker

Figuur 2.1.1: Studio van een lokale omroep, Jelke Bethlehem, 2012.

Figuur 3.3.1. De Eurobarometer, Europese Commissie, www.ec.europa.eu.

Figuur 4.1.1: Een mondelinge peiling, Rondje non-response van de inter/view groep, inter/view b.v., 1978.

Figuur 4.2.1: Een Hollerith-machine, United States Census Bureau, Washington, DC.

Figuur 4.2.2. Een van de eerste interview-computers: de Epson PX-4 notebook, Jelke Bethlehem, 1987.

Figuur 4.3.1: Een voorbeeld van een online-vragenlijst, Jelke Bethlehem, 2014.

Figuur 4.3.2: De winnaar van de NS Publieksprijs 2005, Nederlands Bijbelgenootschap, 2005

Figuur 5.2.1. De peiling van de ANWB over rekeningrijden, ANWB, maart 2010.

Figuur 5.3.2. Een zakrekenmachine met een knop voor aselechte getallen, Jelke Bethlehem, 2014.

Figuur 8.3.1: Social Media Stress, NOS, www.nos.nl, mei 2012.

Medewerkers

Jelke Bethlehem



**Centraal Bureau
voor de Statistiek**

2014 Editie