

OVER WEGEN

Mogelijkheden en beperkingen van het wegen van steekproeven

arie den boon

(met dank aan martin koopman en harm 't hart)

Wat houdt het wegen van steekproeven in?

Het wegen van een steekproef is het aanpassen van een gevonden verdeling aan een gewenste verdeling (bijvoorbeeld populatie verdeling), door middel van het toekennen van gewichten aan de elementen van de steekproef. Een gewicht is groter dan 1 wanneer een element in de steekproef is ondervertegenwoordigd en daarom sterker moet meetellen. Een gewicht kleiner dan 1 wordt toegekend als een element is oververtegenwoordigd en minder zwaar moet meetellen.

De oorspronkelijke steekproef kan dan ook beschouwd worden als een gewogen steekproef waarvan alle gewichten gelijk zijn aan 1.

Redenen om te wegen

Er zijn verschillende redenen om steekproeven te wegen. De hieronder gegeven motieven zijn grotendeels overgenomen van 't Hart (1974, p. 64):

1. herstel van gelijke kansen om in de steekproef te worden betrokken
2. verandering van analyse eenheid
3. correctie voor non-response
4. post-stratifikatie
5. kwasi-experimenten
6. correctie voor ongelijke errorvariantie
7. het vergroten van de vergelijkbaarheid van onafhankelijk getrokken steekproeven

Aanvankelijk werden steekproeven gewogen door random elementen uit een subgroep te dupliceren of te elimineren. Een betere methode is die waarbij ieder element een wegingsfaktor krijgt ('t Hart, 1974, p. 69). Deze methode wordt tegenwoordig meestal toegepast.

Als de elementen van een subgroep bijvoorbeeld met een faktor twee zijn oververtegenwoordigd, krijgt ieder van die elementen na het wegen een wegingsfaktor van .5 toegekend.

Het effect van wegen op de steekproeven-variantie.

Bij het wegen van een steekproef treedt een zeker verlies op in nauwkeurigheid door de toename van de steekproeven-variantie. Kish geeft voor de gevallen van aselekte en gestratificeerde steekproeven afzonderlijke afleidingen (1965, pp. 424-433). Wat is die steekproeven-variantie nu eigenlijk? De steekproeven-variantie is de variantie van het gemiddelde of de variantie van een variabele. Er kunnen een heleboel steekproeven worden getrokken uit een bepaalde populatie. Omdat die steekproeven niet steeds gelijk zijn, kan het gemiddelde van een variabele de ene keer wat hoger zijn dan de andere keer. De variantie van het gemiddelde is afhankelijk van de omvang van de getrokken steekproeven. Hoe kleiner de omvang van de steekproeven, des te groter wordt de steekproeven-variantie van het gemiddelde. En daarmee hebben we een maat voor de nauwkeurigheid van een steekproef.

Bij een eenvoudige aselekte steekproef (e.a.s.) is de steekproeven-variantie van het gemiddelde:

$$\text{var}(\bar{y}) = \frac{S^2}{n} \cdot (1-f)$$

waarin:

S^2 = populatie variantie van Y

n = steekproef grootte

(1-f) = eindigheidskorrektie (bij benadering gelijk aan 1)

(Moors en Muilwijk, 1975, p. 25)

De eindigheidskorrektie kan in veel gevallen worden verwaarloosd omdat de populatie zeer groot is ten opzichte van de steekproef. De eindigheidskorrektie benadert dan de waarde 1, waardoor de steekproeven-variantie kan worden geschreven als:

$$\text{var}(\bar{y}) \doteq \frac{S^2}{n}$$

$\doteq$ betekent: bij benadering gelijk aan

Als uit verschillende strata in een populatie aselekte steekproeven worden getrokken met een omvang evenredig aan de aantallen elementen in de strata, spreekt men van een proportioneel gestratificeerde steekproef (p.s.s.).

De steekproeven-variantie van een p.s.s. is:

$$\begin{aligned} \text{var}(\bar{y}_{\text{prop}}) &= \frac{\sum W_h \cdot S_h^2}{n} \cdot (1-f) \\ &= \frac{\sum W_h \cdot S_h^2}{n} \end{aligned}$$

waarin:

W_h = populatiefractie van stratum h

S_h^2 = populatievariantie in stratum h

(Kish, p. 139)

Om deze varianties te vergelijken kunnen we de variantie van een e.a.s. in twee delen opsplitsen. Het ene deel is de "within groups variance" die overeenkomt met de variantie van een p.s.s. Het andere deel is de "between groups variance":

$$\begin{aligned} \text{var}(\bar{y}) &= \frac{\sum W_h \cdot S_h^2 + \sum W_h \cdot (\bar{y}_h - \bar{y})^2}{n} \cdot (1-f) \\ &= \frac{\sum W_h \cdot S_h^2}{n} + \frac{\sum W_h \cdot (\bar{y}_h - \bar{y})^2}{n} \end{aligned}$$

(Kish, p. 139)

Over het algemeen is de steekproeven-variantie van een proportioneel gestratificeerde steekproef kleiner dan van een aselekte steekproef van dezelfde omvang. Hoe groot dat verschil is hangt af van de grootten en de gemiddelden van de strata.

Zoals gezegd wordt door wegen de steekproeven-variantie van een steekproef groter. We zullen nu eerst het geval behandelen waarin aan de elementen van een enkelvoudige aselekte steekproef een random gekozen wegingsfaktor wordt toegekend. De wegingsfactoren zijn wel zodanig gekozen dat de omvang van de gewogen steekproef gelijk is aan de oorspronkelijke omvang.

De steekproeven-variantie van een e.a.s. wordt na weging gelijk aan:

$$\text{var}\left(\frac{\sum k_h \cdot y_h}{\sum k_h \cdot n_h}\right) = \frac{s^2}{n} \cdot \frac{\sum W_h \cdot k_h^2}{(\sum W_h \cdot k_h)^2}$$

waarin:

$$W_h = \frac{n_h}{n} = \text{fraktie in steekproef met wegingsfaktor } k_h$$

$$k_h = \text{wegingsfaktor in subgroep } h$$

$$\begin{aligned} \text{var}\left(\frac{\sum k_h \cdot y_h}{\sum k_h \cdot n_h}\right) &= \frac{s^2}{n} \cdot \frac{\sum \frac{n_h}{n} \cdot k_h^2}{\left(\sum \frac{n_h}{n} \cdot k_h\right)^2} \\ &= \frac{s^2}{n} \cdot \frac{\sum n_h \cdot k_h^2}{n} \end{aligned}$$

Wanneer de wegingsfactoren gelijk zijn aan 1, zoals bij een ongewogen steekproef, is de rechterterm gelijk aan 1. Maar wanneer de wegingsfactoren niet gelijk zijn aan 1 wordt de rechterterm groter dan 1. De steekproeven-variantie en dus de onnauwkeurigheid neemt toe. Het is alsof de steekproefgrootte (n) kleiner wordt met een faktor gelijk aan de rechterterm.

Die rechterterm wordt wel de Kishfaktor genoemd (K). Delen we de oorspronkelijke wegingsfactoren door deze Kishfaktor, dan neemt de steekproef in omvang af. En wel net zoveel als bij de kleinere nauwkeurigheid van de steekproef past. Deze nieuwe omvang wordt wel de effectieve steekproefgrootte genoemd (zie ook 't Hart, p. 69).

Ook gestratificeerde steekproeven kan men wegen. Meestal is het de bedoeling om in de steekproef de frakties van de strata gelijk te maken aan de frakties in de populatie. De steekproeven-variantie neemt volgens Kish toe en kan geschreven worden als:

$$\text{var}(\bar{y}_{\text{prop}}) = \left(\sum W_h \cdot k_h \cdot S_h^2 \right) \cdot \left(\sum \frac{W_h}{k_h} \right) \cdot \left(1 - \frac{f}{k_h} \right)$$

waarin:

$$W_h = \frac{N_h}{N} = \frac{n_h \cdot k_h}{n} = \text{populatiefraktie van stratum } h$$

(Kish, p. 429)

Omdat de stratumvariantie allerlei verschillende variabelen kan betreffen en de verschillen tussen de strata vaak betrekkelijk klein zijn, kan S_h vervangen worden door S . Ook kan in de meeste gevallen de eindigheidskorrektie vervallen. Dit vereenvoudigt de formule tot:

$$\begin{aligned}\text{var}(\bar{y}_{\text{prop}}) &= \frac{S^2}{n} \cdot \left(\sum W_h \cdot k_h \right) \cdot \left(\sum \frac{W_h}{k_h} \right) \\ &= \frac{S^2}{n} \cdot \left(\sum \frac{n_h \cdot k_h}{n} \cdot k_h \right) \cdot \left(\sum \frac{n_h \cdot k_h}{n} \cdot \frac{1}{k_h} \right) \\ &= \frac{S^2}{n} \cdot \frac{\sum n_h \cdot k_h^2}{n}\end{aligned}$$

(Kish, p. 430)

Wanneer een eenvoudige aselekte steekproef wordt gewogen alsof het een gestratificeerde steekproef is (post-stratifikatie), krijgt men hetzelfde resultaat:

$$\begin{aligned}\text{var}(\bar{y}) &= \frac{S^2}{n} \cdot \left(\sum W_h \cdot k_h \right) \cdot \left(\sum \frac{W_h}{k_h} \right) \\ &= \frac{S^2}{n} \cdot \frac{\sum n_h \cdot k_h^2}{n}\end{aligned}$$

(Kish, p. 430)

In alledrie de gevallen neemt de steekproeven-variantie met een bij benadering gelijke faktor toe. Over de grootte van de faktor kunnen we twee opmerkingen maken:

1. De grootste toename van de variantie ontstaat wanneer er maar twee strata zijn. Dat pleit voor het verdelen van de steekproef in meerdere strata. Aan de andere kant moeten de strata ook weer niet te klein zijn om erg ongelijke gewichten te vermijden. Kish geeft als vuistregel een verwachte steekproef stratumgrootte $n \cdot W_h \geq 10$ (Kish, p. 90).
2. De variantie van een "gepost-stratificeerde" steekproef kan de variantie van een proportioneel gestratificeerde steekproef wel benaderen maar kan niet kleiner zijn. In het algemeen verdient het dan ook de voorkeur om vooraf te stratificeren.

Mogelijkheden en beperkingen van wegen.

Eén van de bezwaren die wel worden geopperd tegen het wegen van steekproeven is de soms sterke afname van de effectieve steekproefgrootte. In hoeverre is dat bezwaar terecht? Over het algemeen is een gestratificeerde steekproef te prefereren boven een aselekte, omdat de laatste een grotere steekproevenvariantie heeft. Door een steekproef te wegen kan men van een aselekte steekproef een gestratificeerde steekproef maken. De variantie van de aldus verkregen gestratificeerde steekproef is groter dan van een vooraf gestratificeerde steekproef. Daarom wordt de omvang van de steekproef verkleind tot die wel te vergelijken is met een gewone gestratificeerde steekproef.

Het komt er dus op neer dat men voor het veranderen van een aselekte steekproef in een gestratificeerde, moet betalen met steekproef elementen, ongeacht of men achteraf ook daadwerkelijk gebruik maakt van de mogelijkheden van een gestratificeerde steekproef. De voordelen van een gestratificeerde steekproef krijgen we alleen wanneer ook formules voor gestratificeerde steekproeven worden gebruikt. De effectieve steekproefgrootte is weliswaar afgenomen, maar dat wordt gecompenseerd door de lagere steekproevenvariantie van een gestratificeerde steekproef. De winst zit hem in de toegenomen precisie of validiteit. Onder validiteit verstaan we hier de overeenkomst tussen datgene wat we willen meten en datgene wat gemeten is. De representativiteit van de steekproef is na de weging voor de wegingsvariabelen gegarandeerd. Er gaat door de weging dan ook geen informatie verloren, de ene zekerheid wordt ingeruild voor een andere. In het geval de steekproef na weging nog steeds wordt behandeld als een eenvoudige aselekte steekproef, neemt door de kleinere steekproef omvang de nauwkeurigheid af, de betrouwbaarheidsgrenzen worden groter, maar de eventuele vertekening neemt af.

Overigens moet worden gezegd dat de mogelijkheden van gestratificeerde steekproeven meestal niet worden gebruikt in de sociale wetenschappen. Bij mijn weten bestaan er in de gangbare statistische pakketten geen faciliteiten voor. Eén van de redenen hiervoor is mogelijk dat er naast stratifikatie meestal klustering plaatsvindt. Dat wil zeggen dat in de onmiddellijke omgeving van een aselekt getrokken element meerdere gesprekken plaatsvinden. Daardoor kan de variantie met een meestal onbekende faktor toenemen. Veelal neemt men dan maar aan dat het stratifikatie effect en het cluster effectelkaar ongeveer zullen opheffen. Daarvoor zijn echter geen gronden aan te voeren ('t Hart, p.68).

Een belangrijke reden om een steekproef te wegen is de vergelijkbaarheid met andere al dan niet gewogen steekproeven (punt 7 onder de redenen om te wegen). Bij longitudinaal onderzoek maar ook bij onderzoek op door anderen verzameld materiaal kan men de resultaten gedeeltelijk ontdoen van toevallige en systematische fouten. Veelal wordt bij dergelijk onderzoek gezocht naar veranderingen in de tijd. Daarbij maakt men impliciet of expliciet gebruik van verschillskores en die zijn buitengewoon gevoelig voor afwijkingen van de steekproef. Het wegen van de steekproeven kan dan weliswaar niet helpen bij het verkleinen van de betrouwbaarheidsgrenzen om een gevonden waarde, maar de vertekening (bias) is wel kleiner geworden. Voor een voorbeeld van de vergelijking van twee enquêtes wordt verwezen naar Koopman (1978).

Als twee onafhankelijk getrokken steekproeven worden gekombineerd tot één grote, is het beter om pas achteraf te wegen. De afwijkingen van steekproeven kunnen namelijk worden verdeeld in toevallige afwijkingen (error) en systematische (bias). De toevallige afwijkingen hebben de neiging om elkaar op te heffen, waardoor de totale fout kleiner wordt. Daarom zal de Kishfaktor van de gekombineerde steekproef in het algemeen kleiner zijn dan wanneer hij is gebaseerd op de afzonderlijk gewogen steekproeven.

Een mogelijk nadelig effect van het wegen van een steekproef is het kapitaliseren op toeval. Als om de een of andere reden een subgroep in de steekproef is ondervertegenwoordigd, zal door de weging iedere ondervraagde uit die subgroep meer dan één keer meetellen. Naarmate de subgroepen kleiner zijn is de kans op extreme wegingsfactoren groter. Wanneer bepaalde groepen ondervraagden meer dan één keer meetellen kan een vertekend beeld van de subgroep ontstaan. Het is immers niet denkbeeldig dat de elementen die wel in de subgroep aanwezig zijn, afwijken van niet aanwezige elementen. Dit is een bekend probleem dat wordt onderzocht in non-response onderzoek. Uit diverse van dergelijke onderzoeken blijkt dat sommige bevolkingsgroepen nu eenmaal minder kans hebben om in een steekproef terecht te komen dan andere.

Een verwant probleem doet zich voor wanneer men uitspraken wil doen over subgroepen. Als er uitspraken moeten worden gedaan over een gedeelte van de steekproef, moet men er eerst voor zorgen dat het aantal elementen in dat gedeelte voor en na de weging gelijk blijft. Er geldt op het nivo van de subgroepen hetzelfde als voor de hele steekproef: men kan niet een steekproef trekken van 100 elementen,

alle elementen tweemaal laten meetellen en uitspraken doen op grond van een steekproef van 200 elementen. Als er dus een uitspraak moet worden gedaan over een gedeelte van de steekproef, moet in dat gedeelte de omvang gelijk blijven en op grond van die wegingsfactoren moet ook de Kishfaktor berekend worden.

Men doet er goed aan om niet op al teveel variabelen te wegen, of beter, het totaal aantal subgroepen niet te groot te laten worden. Als de groeps grootte te klein zou worden, kunnen eventueel groepen worden samengenomen. Het is dan natuurlijk wel van belang dat de groepen een logisch verband met elkaar hebben. Bij niet nominale variabelen moeten daarom in het algemeen aansluitende categorieën worden samengenomen. Een en ander hangt natuurlijk in eerste instantie af van de variabelen die moeten worden onderzocht en de relatie met de wegingsvariabelen.

In sommige gevallen levert een gewogen steekproef problemen op bij de analyse. Bij het verwerken van ordinale gegevens zijn de analyse methoden meestal niet geschikt voor "gebroken" elementen. In dat geval moet worden gewerkt met de ongewogen steekproef, of de wegingsfactoren moeten worden afgerond op gehele getallen. Een andere methode, zoals bijvoorbeeld gebruikt in het pakket SPSS, gaat als volgt. Een element wordt even vaak gemultipliceerd als het gehele gedeelte van de wegingsfaktor aangeeft. Daarnaast krijgt het element nogmaals een kans om de analyse te worden betrokken, deze kans is gelijk aan de resterende breuk van de wegingsfaktor. Deze laatste methode is zeker bij grote steekproeven te preferen boven afronden, maar heeft als nadeel dat de analyses iedere keer enigszins verschillende resultaten opleveren. De wegingsfactoren geven bij deze methode een grotere steekproeven-variantie dan bij gebroken elementen (zie bijvoorbeeld Kish, pp. 424-433).

Veel statistische programmatuur die in principe wel geschikt zou zijn voor het werken met gebroken wegingsfactoren, is daar niet voor uitgerust. In die gevallen moet men gebruik maken van random duplicatie en eliminatie van elementen.

Het effect van wegen op onderzoeksresultaten.

Effecten van wegingsfactoren zijn er alleen op de "wegings" variabelen zelf en op de variabelen die daarmee samenhangen. Het meest duidelijk zijn de veranderingen op de één-dimensionale verdelingen van die variabelen. Op de niet gewogen variabelen zijn de effecten meestal betrekkelijk klein, hoewel ze in geval van sterke wegingsfactoren en een sterke samenhang met de wegingsvariabelen vrij sterk kunnen zijn. Het effect van wegen op de associaties tussen variabelen is meestal zeer gering, tenzij weer één of meer variabelen sterk samenhangen met de wegingsfactoren ('t Hart, pp. 78-80). Een en ander hangt ook af van de gevoeligheid van de associatiemaat van de marginale verdelingen.

Een probleem dat zich vaak voor doet bij het wegen van steekproeven is de onbekendheid van sommige populatiegegevens. Meestal zijn wel de marginale verdelingen, zoals bijvoorbeeld leeftijd en geslacht bekend, maar niet hun simultane verdeling (leeftijd maal geslacht). Een oplossing voor dat probleem is om eerst naar de ene en vervolgens naar de andere variabele te wegen. Als de eerste variabele dan weer afwijkt, moet het proces worden herhaald. Dit gaat net zolang door tot de marginale verdelingen van steekproef en populatie overeenkomen. Deze procedure heet "iterative proportional fitting" en convergeert meestal zeer snel. Meestal zijn er niet meer dan drie tot zeven iteraties nodig. Deming en Stephan hebben deze procedure geïntroduceerd (zie bijvoorbeeld Feinberg, 1978, p. 34). Het voordeel van deze methode is dat het verband tussen de variabelen zoals dat is gevonden in de steekproef, zoveel mogelijk intact blijft. En bij gebrek aan exakte populatiegegevens is dat de beste schatting van de samenhang die er is.

Beperkte weging.

Omdat erg grote en kleine wegingsfactoren nadelig zijn, kan een steekproef ook beperkt gewogen worden. De afwijking ten opzichte van de populatie wordt dan wel verminderd, maar niet volledig opgeheven, om tegemoet te komen aan het bezwaar van extreme wegingsfactoren. Hiervoor kunnen verschillende criteria worden gehanteerd.

1. De wegingsfactoren moeten kleiner zijn dan een bepaald getal, bijvoorbeeld $kh \leq 5$ en/of groter zijn dan een ander getal, bijvoorbeeld $kh \geq .33$.

2. De stratumgrootte mag niet te klein zijn, bijvoorbeeld $nh \leq 10$.

Aan de andere kant is natuurlijk de bedoeling om de steekproef zo goed mogelijk naar de populatie toe te wegen. Als criterium voor de afwijking van de steekproef wordt vaak de chi-kwadraat toets voor één steekproef gebruikt. Hieruit volgen nog een aantal criteria:

3. De afwijking van de populatie slechts zover verminderen tot die op grond van toeval zou kunnen zijn verkregen met een waarschijnlijkheidskans van 1 of 5 of 20 %.
4. Men kan de aanpassing bij belangrijke variabelen vollediger laten zijn dan bij minder belangrijke.
5. Tenslotte kan men ook eerst die cellen wegen die de grootste bijdrage leveren aan de chi-kwadraat (zie bijvoorbeeld Verzijden, 1967).

Kombinaties van deze criteria zijn mogelijk. Door de beperking van de range van de wegingsfactoren kan de steekproefgrootte al afnemen voordat er gedeeld is door de Kishfaktor. Daarvoor kan worden gecorrigeerd door een aangepaste formule voor de Kishfaktor te gebruiken.

$$K = n \cdot \frac{(\sum n_h \cdot k_h^2)}{(\sum n_h \cdot k_h)^2}$$

Een voorbeeld: omroepverkiezingen.

Al geruime tijd worden er in Nederland diskussies gehouden over de wenselijkheid van omroepverkiezingen. Daarbij is het voor alle partijen natuurlijk interessant om te weten welke gevolgen dat zou hebben voor de onderlinge verhoudingen tussen de omroepen. Daarom is in de Kwartaalenquête gevraagd naar het stemgedrag van Nederlanders van 17 jaar en ouder bij eventuele omroepverkiezingen. De Kwartaal-enquête is een (overigens slechts twee of driemaal per jaar) gehouden onderzoek naar politieke en andere vraagstukken, georganiseerd door de Werkgroep "Kwartaal onderzoek" van de subfaculteit der Algemene Politieke en Sociale Wetenschappen van de Universiteit van Amsterdam in samenwerking met het Steinmetzarchief. De vraag luidt:

"Er wordt wel eens over gesproken om omroepverkiezingen te houden. De hoeveelheid zendtijd voor een omroepvereniging zou dan afhangen van het aantal stemmen dat die omroepvereniging krijgt. Als er zulke verkiezingen zouden zijn, op welke omroepvereniging zou u dan stemmen?"

De data waarop de onderzoeken zijn gehouden zijn:

1. Golf 15, november 1977
2. Golf 16, januari 1978
3. Golf 17, juni 1978
4. Golf 18, november 1978

Het aantal vraaggesprekken waarnaar bij ieder onderzoek wordt gestreefd bedraagt 600. Na de dataverzameling is iedere steekproef apart gewogen, waarbij de steekproeven 2, 3 en 4 zijn teruggewogen naar de populatiegegevens behorend bij steekproef 1. De reden hiervoor is dat men meer was geïnteresseerd in de verandering van de populariteit van de programma's op zich dan in veranderingen door populatie wijziging. Met andere woorden, de populatie is gedurende de analyse konstant gehouden. De variabelen waarop de steekproeven zijn gewogen zijn gezinsgrootte, leeftijd, geslacht, burgerlijke staat en provincie. Omdat het wegen op al deze variabelen tegelijkertijd teveel subgroepen zou opleveren zijn diverse subgroepen samengenomen. Zie voor de beschrijving van de weging van de Kwartaalenquête Bakker (1978). In tabel 1 staan de aantallen elementen in de steekproef voor en na de weging vermeld.

Onderzoek:	1	2	3	4
steekproefomvang:	583	594	606	599
Kishfaktor:	1.49	1.58	1.55	1.57
steekproefomvang na weging:	390	377	392	382
kategorie weet niet geen antwoord:	67	46	80	55
blijft over:	323	331	312	327

Tabel 1, steekproefgrootte voor en na weging.

In tabel 3 staan de resultaten van de vier onderzoeken vermeld. Met een chi-kwadraat toets voor één steekproef kunnen we berekenen hoe groot de afwijking is tussen de verdelingen voor en na weging. In tabel 2 staan de chi-kwadraat waarden vermeld. Bij steekproef 2 zijn er maar 8 vrijheidsgraden omdat na weging één van de kategoriën geen cases meer bevatte.

Onderzoek:	1	2	3	4
Chi-kwadraat	8.13	9.52	4.80	4.91
vrijheidsgraden:	9	8	9	9

Tabel 2, verschil tussen steekproeven voor en na weging.

Onderzoek:	1	2	3	4
Omroep:				
AVRO	16.7 15.5	18.8 17.7	16.7 15.1	15.5 16.3
EO	5.2 6.5	3.3 4.6	2.7 2.8	4.3 3.4
KRO	10.1 8.8	13.5 11.5	9.6 8.0	10.1 9.5
NCRV	15.3 15.9	12.3 12.0	10.6 10.9	10.4 12.2
NOS	.8 .9	1.3 2.4	1.7 1.1	.6 .5
TROS	22.9 20.0	23.8 21.3	29.2 32.4	29.0 28.2
VARA	16.7 18.0	13.7 15.4	13.3 12.2	14.8 15.2
VOO	6.2 8.8	8.1 10.3	9.0 10.1	9.7 10.6
VPRO	5.4 5.4	4.8 4.6	6.7 7.0	5.2 3.3
Anders	.6 .4	.4 .1	.6 .3	.4 .7
Totaal:	100.0 100.0	100.0 100.0	100.0 100.0	100.0 100.0

Tabel 3, populariteit omroepen uitgedrukt in percentages;
bovenste getal: vóór weging, het onderste ná weging.

De afwijking tussen de steekproeven voor en na weging is niet groot. Er lijkt geen reden tot ongerustheid over het gebruik van de ongewogen datasets. Aan de Kishfaktor zien we weliswaar dat de afwijking van de wegingsvariabelen aanzienlijk is, maar doordat er blijkbaar geen sterk verband is tussen de wegingsvariabele en de omroepvoorkeur, komen die verschillen niet tot uitdrukking in deze resultaten. Meestal worden resultaten zoals hier gevonden echter op een andere manier geïnterpreteerd. Men bepaalt dan rij voor rij of er omroepen zijn die winst hebben geboekt of verlies hebben geleden, of er trends zijn, etc. Hoewel dit soort vergelijkingen tussen steekproeven riskant zijn geven ze wel de meest interessante gegevens. Bij een dergelijke beschouwingswijze blijken er wel degelijk belangrijke verschillen te zijn tussen de resultaten voor en na weging. Zo zien we bijvoorbeeld bij de TROS een verschil van meer dan 3 %, dat is een niet gering verschil. Verder lijkt het bij de resultaten voor de weging alsof de NCRV en de AVRO in populariteit dalen, na weging

lijken ze toch weer enig herstel te vertonen in de laatste peiling. Overigens moeten dergelijke resultaten eigenlijk worden geïnterpreteerd aan de hand van betrouwbaarheidsgrenzen.

Het wegingsprogramma WEGING

Op het Baschwitz Instituut voor Massapsychologie en Openbare Mening is enige ervaring opgedaan met het wegen van steekproeven (zie ook Bakker, 1978). Voor het wegen is het komputerprogramma WEGING geschreven, een eenvoudig interaktief programma voor gevallen waarin de populatieverdeling bekend is en voor gevallen met een onbekende simultane verdeling van maximaal twee variabelen. De invoer bestaat uit een kruistabel van variabelen met steekproef gegevens en een kruistabel of de marginale verdelingen van populatie gegevens. De uitvoer bestaat uit de wegingsfactoren zonder en met correctie voor de Kishfaktor, de afwijking tussen steekproef en populatie in de vorm van een chi-kwadraat waarde, het verband tussen de wegingsvariabelen voor en na weging, etc.

Voor het wegen van steekproeven op meer dan twee variabelen met onbekende simultane verdeling is het programma niet geschikt. In principe is het wel mogelijk andere programma's daarvoor te gebruiken, zoals programma's voor log-lineaire modellen. De invoer is dan echter bewerkelijk. Het zou dan ook zeer nuttig zijn als er voor het wegen van steekproeven een nieuw programma zou worden geschreven.

Voorwaarden waaraan een wegingsprogramma moet voldoen.

Een eerste belangrijke algemene eis die men mag stellen is de gebruikersvriendelijkheid. Zeer veel onderzoekers maken gebruik van steekproef materiaal en het programma moet zijn aangepast aan hen die niet zo'n uitgebreide kennis van dataverwerking hebben. Vreemd genoeg is er ondanks de populariteit van steekproeven maar weinig aandacht voor de theorie van het steekproef trekken. De meeste literatuur is ook nogal saai. Toch heeft met name de interpretatie en de generaliseerbaarheid van de resultaten direkt te maken met de kwaliteit van de steekproef.

De mogelijkheid om verantwoord van analyse element te veranderen is bij velen onbekend.

Voor de invoer van de gegevens zijn de volgende punten van belang:

1. De invoer van steekproefgegevens moet mogelijk zijn in de vorm van ruwe data, een kruistabel en via een (SPSS)interface. Minimaal zijn er drie dimensies nodig waarop kan worden gewogen.
2. De invoer van populatiegegevens moet mogelijk zijn in de vorm van een kruistabel en/of in de vorm van marginale verdelingen.
3. Criteria volgens welke er moet kunnen worden gewogen:
 - volledige of beperkte weging (1, 5 of 20 % chi-kwadraat overschrijdingskans)
 - grenzen aan de wegingsfactoren (ondergrens, bovengrens)
 - cellen die sterker bijdragen aan de chi-kwadraat sterker terugbrengen dan andere
 - prioriteit van bepaalde wegingsvariabelen

De uitvoer moet tenminste bevatten:

1. De wegingsfactoren, ofwel via een (SPSS)interface, ofwel in de vorm van een datamatrix of kruistabel.
2. Verslag van de analyse op papier met:
 - invoerparameters
 - chi-kwadraat voor verband tussen steekproef en populatie
 - chi-kwadraat voor verband tussen de variabelen voor en na weging en na correctie voor de Kishfactor
 - histogram van grootten van de subgroepen voor en na weging
 - histogram van de wegingsfactoren voor en na weging
 - Kishfactor en steekproef omvang bij volledige en beperkte weging.

Literatuur

- Bakker, T.P.V.: "De weging van het kontinu-onderzoek" in: "Röntgenfoto van enquête onderzoek. Een ongeneeslijk ziekte?" Amsterdam, Werkgroep Kwartaal onderzoek, UvA, Steinmetzarchief, 1978.
- Feinberg, S.E.: "The analysis of cross-classified categorical data", Massachusetts, The Massachusetts Institute of Technology, 2nd print, 1978.
- Hart, H. 't: "Selektie en zelf-selektie van informanten in enquêtes." Amsterdam, proefschrift, 1974.
- Kish, L.: "Survey sampling" New York, etc., 1965.
- Koopman, M.J.: "Verschillen en overeenkomsten tussen twee vergelijkbare tezelfdertijd afgenomen enquêtes; is er sprake van een bureau-effekt op survey-data? Eerste bevindingen uit een vergelijkend onderzoek." Amsterdam, Werkgroep Kwartaal onderzoek, UvA, Steinmetzarchief, 1978.
- Moors, J.J.A. en J. Mulwijk: "Steekproeven, een inleiding tot de praktijk" Amsterdam, Agon Elsevier, 1975.
- Verzijden, D.: "Homofilie-projekt. Controle en weging van de steekproef" Amsterdam, UvA, Baschwitz Instituut, 1967.