

Die 1985-Steekproefbevolkingsopname in die Republiek van Transkei

D.J. Stoker

Raad vir Geesteswetenskaplike Navorsing, Pretoria

1. Inleiding

'n Sensus is 'n volledige bevolkingsopname en is veronderstel om akkurate gegewens te verskaf. In die praktyk is dit gewoonlik nie die geval nie en aansuiwering moet gereeld in die verkreeë resultate aangebring word. Dit geld nie alleen vir Derde-wêreldlande nie, maar selfs ook vir Eerstewêreldlande. Hierdie sydighe in die resultate spruit uit sowel nie-dekkingsprobleme as opnamesydighede (metingsydighede) voort. So byvoorbeeld was daar in die VSA nie minder nie as 50 hofsake teen die Bureau of the Census na die 1980-sensus wat almal met onvolledighede (en dus onakkuraathede) in die 1980-sensus te make gehad het. In die geval van 'n sensus in 'n Derde-wêreldland is daar soveel meer faktore wat onakkurate syfers tot gevolg kan hê. Benewens nie-dekkingsprobleme, is daar onder andere ook die volgende belemmerende faktore:

- ongeletterde bevolking
- min bevoegde opnemers
- onvoldoende opleiding en kontrole van opname
- die lang periode waarvoor die sensus noodgedwonge moet strek.

Dit is seker algemeen bekend dat daar ten tye van die 1980-sensus in Suid-Afrika 'n ondertelling in menige gebiede was en dat die gepubliseerde syfers vir 1980 vanuit vroeëre sensus-en ander gegewens aangesuiwer moes word.

'n Steekproefbevolkingsopname lewer dikwels akkurrater resultate as 'n sensus en is bowendien baie goedkoper as 'n sensus. Tyd, geld en voldoende bevoegde opnemers vir 'n steekproefbevolkingsopname bring mee dat die faktore wat tot sydighe sensussyfers aanleiding kan gee, nie tot dieselfde mate in die geval van 'n steekproefopname geld nie. Intensiewe kontrole is ook deurentyd moontlik. As byvoorbeeld Rx miljoen vir 'n sensus beskikbaar is en $R(x/3)$ miljoen vir 'n steekproefbevolkingsopname gebruik word, beteken dit dat 3 keer soveel geld per huishouding in die geval van 'n 10% steekproefbevolkingsopname in vergelyking met 'n sensus beskikbaar is. Die resultate verkry in 'n steekproefbevolkingsopname kan dus baie minder aan sydighe onderhewig wees as dié van 'n sensus. Solank 2 x die standaardafwyking van die beraamde bevolkingsgetal kleiner as die sydigheid in die sensussyfer is, sal die steekproefbevolkingsopname akkurrater syfers as 'n sensus oplewer.

Die belangrikste tekortkoming van 'n steekproefbevolkingsopname is dat bevolkingsramings vir klein domeine, byvoorbeeld vir 'n bepaalde distrik (kyk par. 4.1), relatief onpresies is (dit is die koëffisiënt van relatiewe variasie is te groot), sodat die ramings vir klein domeine van betreklik min waarde in die praktyk is.

Na oorweging van al die voor- en nadele verbonde aan 'n sensus enersyds en 'n steekproefbevolkingsopname andersyds, het die Transkeise Regering besluit om in 1985 nie 'n sensus in daardie land te onderneem nie, maar wel 'n steekproefbevolkingsopname.

2. Die Steekproefnemingsraamwerk

Om 'n steekproefbevolkingsopname te kan uitvoer, moet een of ander raamwerk van die populasie-elemente of groepe van populasie-elemente (byvoorbeeld huishoudings) beskikbaar wees. Hoe vollediger die raamwerk is, hoe beter kan die beraamings van die populasiegetalle wees. In die geval van die 1985-bevolkingsopname in Transkei kon lugfoto's van die totale gebied, wat teen die einde van 1982 geneem is, vir die landelike gebiede gebruik word. Hierby was dit nodig dat 'n huishouding (HH) eenduidig op so 'n lugfoto identifiseerbaar moes wees. Heelwat opleiding en ervaring was vir hierdie doel nodig. Die hele Transkei is deur 'n 1400-tal lugfoto's van ongeveer 1m^2 gedek. Alle nie-stedelike huishoudings op die lugfoto's is eers *tweemaal* deur 'n aantal persone getel en wel op so 'n wyse dat die twee tellings van HHs op enige bepaalde foto nie deur dieselfde persoon gedoen is nie, maar deur verskillende persone. Relatief groot verskille in die tellings wat gevind is, is intensief ontleed om sodoende die tellers beter in die telling van HHs op die lugfoto's te skool. Hierna is die aantal HHs op die lugfoto's vir die derde en finale maal getel en daarna deur 'n 10% steekproef gekontroleer ten einde 'n beraaming van die moontlike afwyking in die tellings te verkry. Voor elke telling is alle merke op die lugfoto uitgevee alvorens 'n volgende persoon die aantal HHs op die lugfoto getel het. Op hierdie wyse is getrag om die telfout so klein as moontlik te hou.

In die stedelike gebiede is van straatkaarte, waarop persele aangedui is, gebruik gemaak. Alle onbeboude en besigheids-erwe is op die kaarte gemerk voor die steekproef ontwerp is.

3. Opleiding van opnemers

Die dienste van die persone wat as opnemers aangestel is, is reeds 'n geruime tyd voor sensusdag (6 Maart) op 'n voltydse basis bekom. Gedurende die eerste week is hulle geleer om kaarte te lees en te teken en is hulle ook in onderhoudvoering opgelei. Daarna is elke opnemer na sy opnemersgebied gestuur en moes hy/sy 'n sketsplan van sy/haar opnemersgebied teken. Ook moes hy/sy 'n aantal onderhoude in 'n nie-getrekte dorpie of gebied voer. Na 'n week moes hulle dan na Umtata terugkeer waar (a) die probleme wat met die onderhoudvoering ervaar is, bespreek is en (b) die sketsplan met die lugfoto van die dorpie of gebied vergelyk is. Indien die sketsplan onvolledig of swak was, moes hy/sy na die gebied teruggaan om die sketsplan oor te teken. (Kyk par. 4.3 vir verdere besonderhede.)

Die doel van die teken van die sketsplanne was (a) om seker te maak dat die opnemer by die korrekte dorpie uitgekóm het en (b) om hom/haar met sy/haar opnemersgebied vertrou te maak. In alle gevalle waar die sketsplan nie hutte of HHs getoon het nie waar daar volgens die lugfoto wel HHs of hutte behoort te gewees het, is na die dorpie of gebied

gery om bevestiging te kry dat daar wel vroeër 'n hut of hutte was wat sedert die neem van die lugfoto afgebreek is.

In die week voor die aanvang van die steekproefbevolkingsopname is die opnemers weer vir 'n kort periode opgelei. Op hierdie wyse is getrag om opnamefoute aan die kant van die opnemers tot 'n minimum te beperk.

4. Die Steekproefontwerp

4.1 Beskrywing van Transkei

Transkei is verdeel in 28 distrikte, elk met sy dorpsgebied (stedelike gebied genoem). Elke distrik is op sy beurt in administratiewe gebiede of areas (AAs) onderverdeel en binne elkeen van die AAs is die bevolking in dorpie ("villages") woonagtig. Elke dorpie het sy eie hoofman. Enige moontlike onsekerheid oor die grense van 'n dorpie kon met die hoofman uitgeklar word. Die twee grootste stedelike gebiede is Umtata en Butterworth. Om hierdie twee stedelike gebiede, sowel as om Libode en Ngqeleni, kom omstedelike gebiede (of buitestedelike gebiede) voor. Die omstedelike gebiede, hoewel dit bepaalde ooreenkomste met die landelike gebiede het, moet tog daarvan onderskei word. 'n Groot deel van die landelike gebiede bestaan uit sogenaamde hervestigde gebiede waar elke huishouding (HH) op die rand van sy stukkie grond gevestig is. (Let wel; 'n HH bestaan in die algemeen uit meerdere hutte.) Om 'n HH in 'n hervestigde gebied op 'n lugfoto te identifiseer, is redelik maklik. Dit is egter moeiliker in die tradisionele (nie-hervestigde) gebiede waar die hutte al op die toppe van die heuwels gebou is, somtyds verwyderd van hul landjies en van waterbronne.

4.2 Beskrywing van die steekproefontwerp¹

Met die oog op die trek van die steekproef, is die volgende indelings gemaak:

- Instellings soos hostels, koshuise, hotelle en woonstelle
- Groter stedelike gebiede (Umtata en Butterworth) (uitgeslote instellings)
- Omstedelike gebiede (om Umtata, Butterworth, Libode en Ngqeleni)
- Kleiner stedelike gebiede (uitgeslote instellings)
- Landelike gebiede.

Elkeen van bogenoemde indelings is as 'n stratum beskou. Meer bepaald is die 28 stedelike gebiede, die 4 omstedelike gebiede sowel as die landelike gebiede van die 28 distrikte as afsonderlike strata geneem.

Die steekproefontwerp in elkeen van bogenoemde 5 indelings word vervolgens afsonderlik bespreek.

Instellings: Alle persone in instellings is opgeneem. Gevolglik is geen steekproef uit instellings getrek nie.

Groter stedelike gebiede (Umtata en Butterworth): In Umtata is woongebiede as afsonderlike substrata geneem. In Butterworth sowel as in elke woongebied (voorstad) van Umtata is eers primêre steekproefnemingseenhede (PSEs) gevorm. Elke PSE het uit een of meer straatblokke bestaan sodanig dat die PSEs so na moontlik uit ewe veel woonpersele bestaan het en 'n bepaalde minimum aantal woonpersele bevat het. (Die PSEs is sodanig gevorm dat die grootte (in terme van die aantal woonpersele) van die PSEs met minder as 'n faktor van twee varieer het.)

Nadat die PSEs gevorm is, is hulle in substrata (dit is sekondêre substrata in die geval van Umtata) gegroepeer sodanig dat elke (sekondêre) substratum uit vier (soms uit drie of vyf) soortgelyke (naburige of homogene) PSEs bestaan het. Hierna is uit *elkeen* van hierdie (sekondêre) substrate *een* PSE ewekansig getrek. Hierdie getrekte PSEs is dan volledig opgeneem. Gevolglik is in die groter stedelike gebiede 'n nagenoeg 25 % steekproef van HHs, dit is van persone, getrek.

Omstedelike gebiede (om Umtata, Butterworth, Libode en Ngqeleni): AAs om bogenoemde stedelike gebiede is as omstedelike AAs geklassifiseer en elkeen van hierdie vier omstedelike gebiede is as 'n afsonderlike stratum geneem. Uit elkeen van hierdie strata is 'n relatief groot proporsie van AAs ewekansig getrek. Vervolgens is uit elke getrekte AA groepe van huishoudings getrek waar 'n groep in wese 'n hoofmansgebied was, dit is daardie HHs waaroor die hoofman jurisdiksie het. Soveel van hierdie hoofmansgroepe van HHs binne 'n getrekte AA is getrek om, sover doenlik, ten minste 200 HHs in die steekproef uit elke getrekte AA in te sluit. Getrekte hoofmansgroepe van HHs is dan volledig opgeneem.

Kleiner stedelike gebied: Kleiner stedelike gebiede met minder as ongeveer 40 tot 50 beboude woonpersele is volledig opgeneem. Waar kleiner stedelike gebiede meer beboude woonpersele bevat het, is HHs ewekansig getrek en alle persone in die getrekte HHs opgeneem. Die aantal HHs wat getrek is, was nooit minder as 30 nie en het in die algemeen toegeneem met die aantal beboude woonpersele in die stedelike gebied.

Landelike gebiede: Binne elkeen van die landelike gebiede van die 28 distrikte is twee substrata van AAs gevorm sodanig dat die twee substrata nagenoeg eweveel AAs bevat het en verder sodanig dat die AAs binne elke substratum so homogeen moontlik met betrekking tot bepaalde stratifikasieveranderlikes (byvoorbeeld die topografie van die gebied) was. Vervolgens is uit elk van die 56 substrata twee AAs ewekansig getrek. Die rede waarom die AAs nie proporsioneel aan grootte getrek was nie, was omdat die telling van HHs op die lugfoto's nog in sy beginstadium was toe die steekproef getrek moes word. Nadat die 112 AAs getrek was, is die aantal HHs in elk van die dorpie binne die getrekte AAs (op die lugfoto's) getel, waarna uit elk van hierdie AAs een of meer dorpie met waarskynlikheid proporsioneel aan grootte getrek is sodanig dat, sover doenlik, minstens 200 HHs in die steekproef uit elke getrekte AA ingesluit was. Elke getrekte dorpie is volledig opgeneem.

4.3 Vergelyking tussen lugfoto's en sketsplanne van getrekte dorpie

Soos reeds gestel, moes die opnemers voor die opname 'n sketsplan van elke getrekte dorpie in die veld teken. Hierdie sketsplanne is met die lugfoto's van die dorpie vergelyk. Die doel hiervan was onder andere om

- seker te maak dat die opnemer die korrekte dorpie(s) geïdentifiseer het, en
- 'n beraming van die groei koers vanaf die datum waarop die lugfoto's geneem is (einde 1982) tot die datum van opname te verkry.

Ten einde 'n betroubare beraming van die groei koers te verkry, is uit elke subdistrik 'n verdere twee AAs ewekansig getrek en uit elke getrekte AA, soos hierbo beskrywe, is een of meer dorpie getrek. Sodoende is uit elke landelike substratum ten opsigte van dorpie (wat met waarskynlikheid proporsioneel aan grootte getrek is) in vier AAs lugfototellings van die aantal HHs (einde 1982) sowel as van die werklike aantal HHs op sensusdag (Maart 1985) verkry. Van hierdie gegewens is gebruik gemaak om die toename in die aantal HHs vanaf die datum waarop die lugfoto's geneem is tot en met sensusdag te beraam.

5. Beraming van die Bevolkingstotaal en van die gepaardgaande standaardfout

5.1 Inleiding

In par. 4.2 is genoem dat die Republiek van Transkei met die oog op die steekproefontwerp in 5 dele ingedeel is. Elkeen van die 5 dele word vervolgens afsonderlik beskou en die wyse

waarop die bevolkingsgrootte sowel as die standaardfout van die beraamde waarde van die bevolkingsgrootte beraam is, kortliks bespreek. Hierna word die wyse waarop die beramings vir al die strata en substrata saamgestel is om die beraamde waarde van die bevolkingsgrootte vir die hele Transkei, sowel as van die standaardfout daarvan, te verkry, aangedui. 'n Aanduiding van die omvang van die variasie in die beraamde waarde van die standaardfout van die beraamde bevolkingstotaal en die 95 % vertrouensinterval vir die bevolkingstotal van die Republiek van Transkei op 6 Maart 1985, word ook aangebied.

Die gewigsfaktor geassosieer met 'n rekord om die bevolkingstotaal te beraam is die resiproke waarde van die seleksie-waarskynlikheid. Weens praktiese omstandighede was dit nie moontlik om AAs in landelike en omstedelike gebiede en straatblokke in groter stedelike gebiede met waarskynlikheid eweredig aan bevolkingsgrootte (of eweredig aan die aantal HHs) te trek nie. Gevolglik moes die trekking noodgedwonge ewekansig geskied. (In die groter stedelike gebiede is weliswaar eers getrag om die straatblokke so te groepeer dat PSEs van so na moontlik dieselfde "grootte" verkry is.) Ten einde akkurrater of presieser beramings te verkry, is die gewigte bereken asof die AAs in landelike en in omstedelike gebiede en die straatblokke in groter stedelike gebiede met waarskynlikheid eweredig aan die aantal HHs (volgens die 1982 lugfoto's in die geval van die landelike en omstedelike gebiede) getrek is.² Dit kom op dieselfde neer as om die verhouding van die populasiegetal HHs tot z die steekproefgetal HHs as gewigte te neem, waar z die aantal AAs is wat uit 'n substratum getrek is. Met ander woorde, vanuit elke getrekte AA is die $(1/z)$ ste van die bevolkingsgrootte van die substratum beraam. In wese kom hierdie prosedure op agterna-stratifikasie neer. Die vraag ontstaan nou of dié agterna-stratifikasie nie die standaardfout van die beraamde bevolkingstotaal sal vergroot nie. Volgens Cochran (1977:134-135) is die omvang van die vergroting van die standaardfout verwaarloosbaar indien daar minstens 20 waarnemings (HHs in hierdie geval) per "agterna-gevormde stratum" voorkom. Dit was deurgaans die geval in die huidige steekproefbevolkingsopname.

5.2 Groter stedelike gebiede

Soos in par. 4.2 gestel, is die groter stedelike gebiede (Umtata en Butterworth) in substrata onderverdeel. Uit elke substratum is 'n onafhanklike steekproef getrek. Meer bepaald het elke substratum uit ongeveer 4 (met 3 as minimum en 5 as maksimum) nagenoeg ewegroot (in terme van aantal beboude woonpersele) PSEs bestaan waarby elke PSE uit een of meer straatblokke bestaan het. Een van hierdie PSEs is dan ewekansig getrek. Vir beramingsdoeleindes is die volgende formules gebruik:

Laat \hat{Y}_h die beraamde waarde van die bevolkingstotaal Y_h in die h -de substratum aandui. As M_{hi} die aantal beboude woonpersele in die i -de PSE van die h -de substratum aandui, dan is

$$M_h = \sum_i M_{hi}$$

die totale aantal beboude woonpersele in die h -de substratum. (Let op dat $M_h/M_{hi} \approx 3$ of 4 of 5 vir alle i .) Gestel die i -de PSE in die h -de substratum word getrek en laat $y_{hi'j}$ = aantal persone woonagtig op die j -de perseel van die i -de PSE in die h -de substratum.

Die substrata is so homogeen moontlik saamgestel. Indien aanvaar kan word dat $ROH=0^3$, dit is dat die variasie in die aantal persone per woonperseel binne PSEs en tussen PSEs gelyk is, benader die steekproefnemingsprosedure dié van 'n

eenvoudige ewekansige steekproef, in welke geval 'n onsydige beramer van Y_h gegee word deur

$$\hat{Y}_h = (M_h/M_{hi'}) \sum_j y_{hi'j}, \quad j=1,2,\dots,M_{hi'}$$

met beraamde variansie gegee deur (Cochran, 1977)

$$\text{Var}(\hat{Y}_h) = (1-f_{hi'}) (M_h/M_{hi'})^2 \cdot \frac{M_{hi'}}{M_{hi'}-1} \sum_j (y_{hi'j} - \bar{y}_{hi'})^2$$

waarby

$$f_{hi'} = M_{hi'}/M_h \text{ en } \bar{y}_{hi'} = \sum_j y_{hi'j}/M_{hi'}, \quad j=1,\dots,M_{hi'}$$

Voorbeeld: Ikwezi A

$$M_{hi'} = \text{aantal beboude woonpersele in getrekte PSE} = 41$$

$$M_h = \text{totale aantal woonpersele in substratum} = 219$$

$$\sum_j y_{hi'j} = \text{aantal persone in getrekte PSE} = 228$$

$$\bar{y}_{hi'} = \text{gemiddelde aantal persone in die getrekte PSEs}$$

$$\sum_j (y_{hi'j} - \bar{y}_{hi'})^2 = \text{som van die kwadrate van die afwykings vanaf die gemiddelde} = 312,10$$

Dan is die beraamde totale bevolking in Ikwezi A

$$\hat{Y}_h = (219/41) \times 228 = 1\,218$$

met beraamde variansie

$$\begin{aligned} \text{Var}(\hat{Y}_h) &= \left(1 - \frac{41}{219}\right) \left(\frac{219}{41}\right)^2 \left(\frac{41}{41-1}\right) (312,10) \\ &= 7\,418 \end{aligned}$$

Die beraamde waarde van die standaardfout van die beraamde bevolkingstotaal in Ikwezi A word dan

$$S\hat{F}(\hat{Y}_h) = \sqrt{7\,418} = 86,1$$

met beraamde waarde van die koëffisiënt van relatiewe variansie

$$K\hat{V}(\hat{Y}_h) = (86,1/1218) \times 100 = 7,07 \%$$

As $\hat{Y} = \sum_h \hat{Y}_h$ die beraamde bevolkingstotaal vir 'n groter stedelike gebied, dan geld vir die beraamde variansie van die beraamde totaal

$$\text{Var}(\hat{Y}) = \sum_h \text{Var}(\hat{Y}_h)$$

By hierdie beraamde waarde van die totale bevolking in 'n groter stedelike gebied is die aantal persone bygevoeg wat in instellings getel is. Die waarde van die variansie bly onveranderd aangesien instellings volledig opgeneem is.

5.3 Kleiner stedelike gebiede

In kleiner stedelike gebiede is 'n eenvoudige ewekansige steekproef van beboude woonpersele getrek, indien die gebied nie volledig opgeneem is nie. In laasgenoemde geval is die bevolkingsgrootte eksak bekend met standaardfout gelyk aan nul. In eersgenoemde geval waar wel 'n steekproef getrek is, is die volgende formules vir beramingsdoeleindes gebruik:

Laat N = totale aantal beboude woonpersele
en n = getrekte aantal beboude woonpersele.

Dan word 'n onsydige beramer van Y, die totale bevolkingsgrootte, gegee deur

$$\hat{Y} = (N/n) \sum_{i=1}^n y_i$$

met y_i = aantal persone woonagtig op i-de getrekte perseel, en met variansie van \hat{Y} beraam deur (Cochran, 1977)

$$\text{Var}(\hat{Y}) = \left(1 - \frac{n}{N}\right) \left(\frac{N}{n}\right)^2 \cdot \frac{n}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

$$\text{waarby } \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i/n.$$

Voorbeeld: Cofimvaba stedelike gebied

$N = 112$, $n = 30$ en y_i = totale aantal persone getel op die getrekte persele = 248,

sodat die beraamde bevolking $\hat{Y} = (112/30) \times 248 = 926$ met beraamde variansie

$$\text{Var}(\hat{Y}) = \left(1 - \frac{30}{112}\right) \left(\frac{112}{30}\right)^2 \left(\frac{30}{30-1}\right) (1815,87) = 19\,169$$

Die beraamde waarde van die standaardfout word dan

$$\text{SF}(\hat{Y}) = \sqrt{19\,169} = 138,5$$

en die beraamde waarde van die koëffisiënt van relatiewe variansie

$$\text{KV}(\hat{Y}) = (138,5/926) \times 100 = 15,0 \%$$

Om 'n beraming van die totale bevolking in 'n kleiner stedelike gebied te verkry, is by die beraamde totaal \hat{Y} die aantal persone in instellings getel. Aangesien in instellings 'n volledige telling plaasgevind het, het dit nie die waarde van die variansie verander nie.

5.4 Landelike gebiede

In par. 4.2 is genoem dat die landelike gebied van elke distrik, deur die groepering van AAs, in twee meer homogene substrata as die distrik as geheel verdeel is. Uit elk van hierdie 56 substrata is twee AAs ewekansig getrek. Ten einde die bevolkingsgrootte van 'n substratum of van 'n distrik te beraam, moet die geassosieerde gewig van elke getrekte AA eers bereken word. Hierdie gewig het uit drie faktore bestaan, naamlik

- 'n deflasiefaktor van die betrokke (getrekte) AA,
- 'n inflasiefaktor vir die distrik waartoe die AA behoort en
- die korrekte bevolking: steekproefverhouding. (Soos in par. 5.1 genoem, moes die resiproke van die seleksiewaarskynlikheid hier geneem gewees het, maar om beter beramings te verkry, is die bevolking: steekproefverhouding eerder gebruik.)

Aangesien daar 'n tydsverloop van meer as twee jaar tussen die datum waarop die lugfoto's geneem is en die bevolkingsopnamedatum was, was dit nodig om die bevolkingstelling van die getrekte dorpie in 'n AA eers te deflasioneer tot die van die "lugfotodatum". Hierdie gedeflasioneerde bevolkingstelling moes daarna geïnflasioneer word tot die van die bevolkingsopnamedatum deur gebruikmaking van 'n gemeenskaplike inflasiefaktor vir die hele distrik waartoe die dorpie/AA behoort het (waardeur 'n meer stabiele beraamde groeifaktor verkry is). Ten slotte, aangesien twee AAs uit elke substratum van AAs getrek is, is die bevolking: steekproefverhouding sodanig bereken dat elke getrekte AA in 'n substratum die helfte van die bevolkingsgrootte in daardie substratum beraam het.

Laat

IF = gemeenskaplike inflasiefaktor vir die distrik

n_p = lugfototelling van HHs in die getrekte dorpie in 'n getrekte AA

n_c = getelde aantal HHs in die getrekte dorpie in 'n getrekte AA op die bevolkingsopnamedatum (dit is die sketsplantelling)

N_p = lugfototelling van al die HHs in die substratum w = gewig geassosieerd met 'n getrekte AA.

Dan geld vir die getrekte AA:

$$w = (n_p/n_c) \times \text{IF} \times N_p / (2n_p)$$

Opmerking: In par. 4.3 is genoem dat, met die oog op die verkryging van 'n meer betroubare beraming van die groeifaktor (IF) in 'n distrik, uit elke substratum 4 AAs getrek is sodat 8 AAs uit elke distrik getrek is. Uit elk van hierdie 8 AAs is een of meer dorpie proporsioneel aan die aantal HHs getrek. Sowel die lugfototelling (einde 1982) as die aantal HHs in Maart 1985 in hierdie dorpie is verkry. As inflasiefaktor vir die distrik is dan geneem die verhouding van die som van die 1985-tellings tot die som van die lugfototellings vir hierdie 8 gebiede.

Voorbeeld: Bizana landelike gebied

Gemeenskaplike IF = 1,0734

Substratum	AA nr.	Gewig
1	2	$\frac{250}{301} \times 1,0734 \times \frac{10430}{2 \times 250} = 18,597$
	12	$\frac{212}{227} \times 1,0734 \times \frac{10430}{2 \times 212} = 24,660$
2	10	$\frac{240}{250} \times 1,0734 \times \frac{11720}{2 \times 240} = 25,160$
	23	$\frac{141}{156} \times 1,0734 \times \frac{11720}{2 \times 141} = 40,321$

Substratum	AA nr	Gewig	Getelde bevolking in getrekte dorpie	Beraamde bevolkingsgetalle
1	2	18,597	1 643	30 555
	12	24,660	1 489	36 719
	Subtotaal			
2	10	25,160	1 523	38 319
	23	40,321	845	34 071
	Subtotaal			
TOTAAL				139 664

Beraamde variansie van die beraamde bevolkingstotaal vir Bizana word dan (Kish, 1965; Verma, 1982; Stoker, 1983)

$$\begin{aligned} \text{Var}(\text{Bizana}) &= (30\,555 - 36\,719)^2 + (38\,319 - 34\,071)^2 \\ &= 56\,040\,400 \end{aligned}$$

met beraamde standaardfout

$$\text{SF}(\text{Bizana}) = \sqrt{56\,040\,400} = 7\,486$$

en beraamde koëffisiënt van relatiewe variasie

$$\text{KV}(\text{Bizana}) = (7\,486/139\,664) \times 100 = 5,36\%$$

5.5 Omstedelike gebiede

In die steekproefontwerp is vier omstedelike gebiede gedefinieer, naamlik AAs om elk van Umtata, Butterworth, Libode en Ngqeleni. Die steekproefnemingsprosedures wat hier gevolg is, is identies aan die van die landelike gebiede behalwe in soverre dat uit elke omstedelike gebied meer as twee AAs getrek is.

Voorbeeld: Butterworth omstedelike gebied

Uit hierdie gebied is 3 AAs getrek.

Gemeenskaplike IF = 1,1496

AA nr	Gewig			
1	$\frac{192}{238}$	x	1,1496	x $\frac{1659}{3 \times 192} = 2,671$
25	$\frac{147}{202}$	x	1,1496	x $\frac{1659}{3 \times 147} = 3,147$
9	$\frac{283}{339}$	x	1,1496	x $\frac{1659}{3 \times 283} = 1,875$

Let wel: Aangesien 3 AAs getrek is, is uit elke getrekte AA 'n derde van die bevolkingsgrootte beraam. Dit verklaar die faktor "3" in bostaande berekeningsformules.

AA nr.	Gewig	Getelde bevolking in getrekte gebiede	Beraamde bevolkingstotal
1	2,671	1 402	3 745
25	3,147	1 026	3 229
9	1,875	2 173	4 074
T O T A A L			11 048

Beraamde variansie van die beraamde bevolkingstotaal (Kish, 1965; Verma, 1982; Stoker, 1983)

$$\begin{aligned} &= \frac{3}{3-1} [(3745)^2 + (3229)^2 + (4074)^2 - (11048)^2/3] \\ &= 544\,261 \end{aligned}$$

Beraamde standaardfout = $\sqrt{544\,261} = 737,7$

Beraamde koëffisiënt van relatiewe variasie =

$$(737,7/11048) \times 100 = 6,68\%$$

5.6 Die Republiek van Transkei as geheel

Die bevolkingstotal vir die hele Transkei is soos volg beraam:—
Beraamde totale bevolking

= (die som van die beraamde bevolkingsgroottes vir alle substrata van Umtata en Butterworth, met die insluiting van die getelde aantal persone in instellings)
PLUS
(die som van die beraamde bevolkingsgroottes vir alle

kleiner stedelike gebiede, met die insluiting van die getelde aantal persone in instellings)

PLUS

(die som van die beraamde bevolkingsgroottes vir die landelike gebiede van alle distrikte)

PLUS

(die som van die beraamde bevolkingsgroottes vir die vier omstedelike gebiede)

Die *variensie* vir die beraamde bevolkingsgrootte vir die Transkei is verkry deur die variensies van die beraamde bevolkingsgroottes vir alle substrata waaruit 'n steekproef van HHs getrek is (dit is gebiede wat nie volledig opgeneem is nie), bymekaar te tel.

Die *standaardfout* vir die beraamde bevolkingsgrootte vir die Transkei is die vierkantswortel van die variensie.

Die koëffisiënt van relatiewe variensie

$$= (\text{Standaardfout/beraamde totale bevolking}) \times 100\%$$

6. Resultaat bereik

Vir die steekproefbevolkingsopname in die Republiek van Transkei is 'n koëffisiënt van relatiewe variasie van die beraamde bevolkingstotaal van 0,73% verkry.

As die beraamde bevolking van Transkei *argumentsonthalt* 3 miljoen is, dan impliseer dit 'n standaardfout van $3\,000\,000 \times 0,0073 = 21\,900$. Die 95% vertrouensinterval vir die werklike totale bevolking van Transkei sou dan

$$(2\,956\,200 ; 3\,043\,800)$$

wees.

Die koëffisiënt van relatiewe variasie van 'n beraamde variensie word in die geval van nagenoeg ewegroot PSEs benaderd gegee deur (Kish, 1965; Verma, 1982)

$\text{KV}(\text{beraamde variensie}) = \sqrt{2/\text{aantal vryheidsgrade}}$
waarby $\text{aantal vryheidsgrade} \approx \text{aantal PSEs} - \text{aantal strata}$.

Vir 2 PSEs per stratum geld dan dat

$$\text{KV}(\text{beraamde variensie}) \approx \sqrt{2/\text{aantal strata}}$$

In alleen die landelike gebiede is al 56 strata gevorm uit elk waarvan 2 PSEs getrek is, sodat

$$\text{KV}(\text{beraamde variensie in landelike gebiede}) \approx \sqrt{2/56} = 0,19.$$

Die groot aantal strata wat in die steekproefontwerp gevorm is, bring dus mee dat die variensie in die standaardfoutberaming vir alle praktiese doeleindes buite rekening gelaat kan word in die intervalberaming van die totale bevolking van die Transkei.

Opmerkings

(a) Hierdie beraamde syfers is *nie* die werklike syfers nie (behalwe die koëffisiënt van relatiewe variasie van 0,73%) en gee slegs 'n aanduiding van die presisie wat in die steekproefbevolkingsopname in Transkei bereik is.

(b) Die voorgaande beskrywing van hoe die steekproefbevolkingsopname van 1985 in die Republiek van Transkei uitgevoer is, moet nie as volledig beskou word nie. Die doel was slegs om die prosedure in breë trekke aan te dui. Die Institute for Management and Development Studies (IMDS) van die Universiteit van Transkei stel 'n volledige verslag op wat te geleentyd gepubliseer sal word en waarin 'n volledige beskrywing van die steekproefbevolkingsopname gegee sal word.

Notas

1. Kyk Stoker (1983) vir 'n algemene bespreking van die prosedure wat hier gevolg is.
2. Kyk Jagers et al (1985) vir die onderliggende rasionaal van hierdie prosedure.
3. ROH staan vir "rate of homogeneity". Kyk Kish et al (1976) en Verma (1982) vir meer besonderhede.

Verwysings

- Cochran, W.G. 1977. *Sampling Techniques*. New York: John Wiley & Sons. Third edition.
- Jagers, P., Oden, A., Trulsson, L. 1985. 'Post-stratification and ratio estimation: Usage of auxiliary information in survey sampling and opinion polls. - *International Statistical Review*, 53: 221-238.
- Kish, L. 1965. *Survey Sampling*. New York: John Wiley & Sons.
- Kish, L., Groves, R.M., Krotki, K.P. 1976. *Sampling Errors for Fertility*

- Surveys. World Fertility Survey Occasional Paper No. 17*, Voorburg, Nederland: International Institute.
- Stoker, D.J. 1983. *Steekproefneming in die praktyk*. Pretoria: RGN. Geleentheidspublikasie nr. 4.
- Verma, V. 1982. *The Estimation and Presentation of Sampling Errors. World Fertility Survey Technical Bulletin No. 11*, Voorburg, Nederland: International Statistical Institute.