

NOTITIE

Van : Mervin van Veen

Aan : SK123

d.d. : 16-04-2018

CC :

Betr. : Korte beschrijving weging

Aanleiding

Vanuit SK123 en verschillende instellingen is de vraag gekomen op welke wijze de data van de NSE 2018 is gewogen en wanneer de weegfactor wel en niet gebruikt zou kunnen worden.

Weging

In enquêteonderzoek is het normaal dat de behaalde respons niet voor alle groepen die aan het onderzoek meedoen even groot is. Stel dat van onderwijsinstelling A de helft van de studenten meedoet met de NSE, en van instelling B 30 procent, dan zou instelling A te zwaar meetellen wanneer we de totaalcijfers van instelling A en B tezamen zouden berekenen. Weging is een manier om achteraf voor deze "scheefheid" te corrigeren: instelling B laten we wat zwaarder meetellen en instelling A wat minder zwaar, zodat de verhoudingen weer zijn zoals in de studentenpopulatie als geheel.

Hetzelfde geldt voor het berekenen van cijfers op lagere schaalniveaus: wanneer we een totaalcijfer voor een faculteit willen berekenen en de respons tussen de verschillende opleidingen binnen die faculteit sterk uiteenloopt, is het raadzaam de verhoudingen weer "recht te trekken" door middel van weging, zodat de opleidingen niet sterker of minder sterk meetellen dan eigenlijk zou moeten, gezien hun daadwerkelijke studentenaantallen.

Methode

Voor de NSE wordt rekening gehouden met de volgende "groepen" voor de weging:

- Instelling (BrinActueel)
- Opleiding (CrohoActueel)
- BrinVolgNummerActueel

Meer precies: we kijken naar de studentaantallen per combinatie van BrinActueel * CrohoActueel * BrinVolgNummerActueel, zowel in de studentenpopulatie als geheel (de studenten die zijn uitgenodigd voor het onderzoek) als in de respons (de studenten die de vragenlijst van de NSE hebben ingevuld).

Stel nu dat bij een onderwijsinstelling bij een specifieke opleiding op een bepaalde vestiging minder studenten meedoen aan het onderzoek dan gemiddeld binnen die instelling het geval is. Deze groep krijgt dan een weegfactor hoger dan 1, zodanig dat de groep precies zo meetelt in de totaalcijfers, als volgens het werkelijke studentaantal zou moeten. Het omgekeerde geldt ook: als in een combinatie van BrinActueel * CrohoActueel * BrinVolgNummerActueel naar verhouding meer studenten hebben meegedaan dan gemiddeld, dan wordt de weegfactor lager dan 1 zodat de resultaten niet onevenredig sterk door deze groep worden beïnvloed.

Dit valt beter te illustreren op basis van een fictief rekenvoorbeeld (zie tabel 1).

Stel: er zijn 2 instellingen met elk slechts 1 opleiding, die op slechts 1 vestiging wordt gegeven. De instellingen hebben in werkelijkheid exact evenveel studenten, alleen de studenten van 11AA hebben een twee keer zo hoge respons als die van 22BB. Als we nu een cijfer zouden willen uitrekenen voor 'alle instellingen', dan telt 11AA voor twee derde mee (10 van de 15 respondenten in totaal) en 22BB voor een derde (5 van de 15 respondenten). In dit voorbeeld leidt dat tot een gemiddeld oordeel op de themascore voor inhoud van de studie van 3,64. De gemiddelde score voor 22BB (3,70) is hoger dan van 11AA (3,61), dus door de lagere respons van 22BB wordt het gemiddelde van beide instellingen tezamen te laag ingeschat¹. De werkelijke verhoudingen zijn echter 50/50 procent.

Figuur 1: rekenvoorbeeld weging

BRIN	CROHO	BrinVolg Nummer Actueel	WEEGFACOR	INHOUD	INHOUD (X WEEGFACOR)
11AA	22222	1	0,750	3,44	2,58
11AA	22222	1	0,750	3,57	2,68
11AA	22222	1	0,750	4,02	3,02
11AA	22222	1	0,750	3,98	2,99
11AA	22222	1	0,750	3,32	2,49
11AA	22222	1	0,750	3,56	2,67
11AA	22222	1	0,750	3,44	2,58
11AA	22222	1	0,750	4,78	3,59
11AA	22222	1	0,750	2,45	1,84
11AA	22222	1	0,750	3,55	2,66
22BB	11111	1	1,500	3,47	5,21
22BB	11111	1	1,500	3,61	5,42
22BB	11111	1	1,500	4,06	6,09
22BB	11111	1	1,500	4,02	6,03
22BB	11111	1	1,500	3,35	5,03
Som			15	54,62	54,85
Gemiddelde (=54,62/aantal waarnemingen)				3,64	
Gewogen gemiddelde (=54,85/som van de weegfactoren)					3,66

Om hiervoor te corrigeren, zullen we dus de respondenten van 11AA minder sterk moeten laten meetellen en die van 22BB juist sterker.

¹ Uitgaande van niet-selectieve non-respons

Dit gaat als volgt in zijn werk:

1. Eerst bepalen we de weegfactoren. In dit vereenvoudigde voorbeeld doen we dat door de proportie van de instelling in de steekproef te delen door die in de populatie ($50\%=0,5$ voor elke instelling), dus voor 11AA wordt dat $0,5/(10/15)=0,75$ en voor 22BB $0,5/(5/15)=1,50$.
2. Vervolgens vermenigvuldigen we per respondent de themascore met de weegfactor.
3. Daarna sommeren we de gewogen themascores en delen we de somscore door de somscore van alle weegfactoren (in dit geval is die gelijk aan het aantal waarnemingen, 15).

Zoals verwacht is de gewogen themascore nu hoger (3,66 in plaats van 3,64), omdat rekening gehouden is met de lagere respons bij instelling 22BB (in vergelijking tot 11AA).

Overigens hoeven punt 2 en 3 niet handmatig uitgevoerd te worden, SPSS doet dat automatisch op de achtergrond. Om de cijfers te wegen hoeft u alleen naar "Data" > "Weight Cases" te gaan en daar bij "Weight cases by" de variabele neer te zetten waarop moet worden gewogen (bij de NSE is dat de variabele "Weegfactor"). Als u vervolgens op OK klikt, staat de weging "aan" en worden alle cijfers gewogen berekend.

De weging van de cijfers van NSE is natuurlijk iets ingewikkelder dan in het fictieve voorbeeld, omdat ook naast op BrinActueel ook gewogen wordt op CrohoActueel en BrinVolgNummerActueel, maar het principe blijft hetzelfde. De weging vindt daarbij dusdanig plaats, dat de weegfactoren niet hoger dan 4 kunnen worden en de gemiddelde weegfactor over alle instellingen 1 is.

Wanneer wegen?

Wegen is met name zinvol wanneer gegevens op geaggregeerd niveau worden berekend, bijvoorbeeld totalen over meerdere instellingen, meerdere opleidingen binnen een instelling of meerdere vestigingen binnen één opleiding. Dat zal vaak het geval zijn, het is daarom raadzaam de weging veiligheids-halve altijd "aan te zetten". Dat geldt zeker wanneer sprake is van sterke responsverschillen tussen onderzoekseenheden (vestigingen, opleidingen, faculteiten, etc.) waarover totalen moeten worden berekend.

In slechts een enkel geval is het verstandiger om niet te wegen, bijvoorbeeld als significantietoetsingen worden berekend. SPSS dient dan te rekenen met de echte netto steekproefomvang, in plaats van met de gewogen netto responsaantallen. Omdat binnen de NSE de gehele studentpopulatie wordt benaderd is toetsing echter veelal niet direct nodig en wenselijk, en daardoor blijft het meest verstandige om de weegfactor altijd te gebruiken.

Kort samengevat:

- De weegfactor liefst altijd "aanzetten" (in SPSS: "Data" >> "Weight Cases")
- Soms kan het wel wenselijk zijn om de weegfactor niet te gebruiken, bijvoorbeeld bij significantietoetsing.

Mervin van Veen
Amstelveen, 14 april 2018