

Wegen, is dat nog wel van deze tijd?

Jaarlijkse NPSO lezingendag 19 mei 2014, Amsterdam

De jaarlijkse lezingendag van het Nederlandstalig Platform voor Survey Onderzoek (NPSO) staat ditmaal in het teken van correctie voor selectiviteit in de werving voor een survey onderzoek. De selectie van steekproefpersonen en – huishoudens via een kanssteekproef en de weging van respons naar populatietotalen achteraf zijn zo langzaam geen vanzelfsprekende uitgangspunten meer voor surveys. Respondenten worden ook via niet-kanssteekproeven verzameld, ontbrekende gegevens worden (meervoudig) geïmputeerd, modelmatige beschrijving van ontbrekende gegevens wordt in verscheidene disciplines toegepast en het ontwerp van een survey kan al tijdens data verzameling gericht zijn op het balanceren van respons naar achtergrondkenmerken. Deze ontwikkelingen rechtvaardigen de vraag of wegen nog wel van deze tijd is. Tijdens de lezingendag werpen zeven sprekers hun blik op deze vraag vanuit hun eigen expertise en ervaring.

Locatie:
KNAW Trippenhuys
Kloveniersburgwal 29
Amsterdam

De lezingendag is voor iedereen toegankelijk, maar registratie vooraf is noodzakelijk. Dit kan op de website van het NPSO, www.npsso.net.

Peter Lugtig, Vera Toepoel en Barry Schouten

Programma

10:00 – 10:25	Inloop en registratie met koffie en thee
10:25 – 10:30	Welkom – dagvoorzitter (Ineke Stoop)
10:30 – 11:15	Over wegen Jelke Bethlehem (CBS/Universiteit Leiden)
11:15 – 11:50	To weight or not to weight? Challenges to correct for sample biases in non-probability web surveys Stephanie Steinmetz (Universiteit van Amsterdam)
11:50 – 12:25	Uitdagingen en oplossingen bij het wegen van telefonische steekproeven o.b.v. mixed sampling frames Kim De Cuyper (GfK België)
12:25 – 13:15	Lunch
13:15 – 13:50	Multiple imputation als een alternatief voor wegen Laura Boeschoten en Gerko Vink (Universiteit Utrecht)
13:50 – 14:25	Wegen vanuit het design in plaats van achteraf? Barry Schouten (CBS/Universiteit Utrecht)
14:25 – 15:00	Over wegingen in marktonderzoek Harm Hartman (Ipsos)

15:00 – 15:25	Pauze
15:25 – 16:10	Incomplete data: analyse en gevoeligheidsanalyse Geert Molenberghs (Universiteit van Hasselt/KU Leuven)
16:10 – 16:30	Afsluitende discussie: Is wegen nog wel van deze tijd? Dagvoorzitter (Ineke Stoop)
16:30 – 17:30	Borrel

Korte samenvattingen

Over wegen

Jelke Bethlehem (CBS en Universiteit Leiden)

Er zijn allerlei redenen waarom de uitkomsten van een survey niet representatief zijn voor de populatie waaruit de steekproef is getrokken. Er kan sprake zijn van onderdekking, omdat het steekproefkader niet de hele populatie afdekt. Ook kan de steekproef niet door loting zijn getrokken, maar door zelfselectie. En dan is er natuurlijk non-respons waardoor een overigens nette steekproef kan worden aangetast. Om toch te kunnen generaliseren naar de populatie, moeten de uitkomsten van de survey worden gecorrigeerd. De aanbevolen techniek hiervoor is wegen. Maak de survey representatief met betrekking tot een serie hulpvariabelen. Dan wordt hij vanzelf ook wel representatief met betrekking tot de doelvariabelen van het onderzoek. Is het allemaal zo simpel? Is wegen een soort Haarlemmer Olie die alle representativiteitsproblemen kan oplossen? Wanneer werkt het wel en wanneer niet? De presentatie geeft een overzicht en maant tot enige voorzichtigheid wanneer u overweegt te wegen.

To weight or not to weight? Challenges to correct for sample biases in non-probability web surveys

Collecting data on wages is central to socio-economic research. However, besides high rates of people not answering wage-related questions, measurement issues are also relevant. Most data from official statistics are too aggregated to allow for detailed individual-level analyses which are crucial for encouraging innovative political-economic solutions in the long run. In this context, web surveys seem to offer various advantages, such as worldwide coverage, cost benefits and a fast data collection process. In particular for sensitive questions, like income, they might provide more reliable results because social desirability effects can be eliminated. While web surveys could represent a good supplement to official statistics data, they pose many methodological challenges. A core problem concerns the representativeness of the data as the sub-population with Internet access might be quite specific.

Against this background, the lecture will provide an overview of the application of existing and new calibration methods to enhance the representativeness of different types of web surveys. The characteristics of a probability-based (LISS panel) and a non-probability based (WageIndicator) web survey on labour market-related topics for the Netherlands are compared with population data from Statistics Netherlands providing a detailed bias description of core variables related to wages for both samples. For a selection of core variables, different adjustment models are applied, such as simple weighting and propensity score adjustment. Finally the properties and theoretical advantages and disadvantages

of the methods are discussed exploring the potentials and constraints of different adjustment methods for probability and non-probability web surveys.

Uitdagingen en oplossingen bij het wegen van telefonische steekproeven o.b.v. mixed sampling frames (RDD, mobile, landline)

Kim De Cuyper (GfK België)

De mobiele telefoon/gsm is uitgegroeid tot een van de meest wijdverspreide communicatieverschijnselen in het afgelopen decennium. De "Eurobarometer Flash E - communication Household survey" bevestigt de trend dat het aandeel mobiele telefonie in Europa blijft toenemen, en voor vele huishoudens ten koste van de vaste lijn. Deze tendens heeft niet alleen een impact op de steekproeftrekking bij CATI onderzoek, maar ook het post-survey aspect, met name de weging.

De trekkingskans (p) wordt negatief of positief beïnvloed al naargelang een van volgende scenario's:

- Telefoonnummers gekoppeld aan meer dan een persoon
- Overlap in de steekproef mobiele en vaste telefoon
- Personen met meer dan een vast / mobiel telefoonnummer

Designgewichten als $1/n_{\text{eligible}}$ en n_{eligible} om de ongelijkheid in de selectiekans te corrigeren worden toegelicht op de NPSO dag 19 mei.

Multiple imputation als een alternatief voor wegen

Laura Boeschoten en Gerko Vink (Universiteit Utrecht)

Van een gewogen dataset weten we dat hij representatief is met betrekking tot de achtergrondvariabelen waarmee hij is gewogen. Ook weten we dat de relatie tussen de respondenten en non-respondenten met betrekking tot deze achtergrondvariabelen MAR is (als hij MCAR was geweest zouden we niet hoeven wegen). Echter, de relatie tussen de respondenten en non-respondenten met betrekking tot de uitkomstvariabelen wordt als MCAR gezien. Non-respondenten krijgen namelijk exact dezelfde responspatronen toebedeeld als de respondenten (indien ze gelijk zijn op basis van de achtergrondkenmerken).

Met wegen wordt de assumptie gemaakt dat een non-respondent met dezelfde achtergrondkenmerken als een respondent, ook hetzelfde responspatroon heeft. Deze assumptie mag je eigenlijk niet maken, en daarom zoeken we een alternatief voor wegen dat wel rekening houdt met dit probleem.

We weten al van multiple imputation dat zij goed met MAR missing data om kan gaan in item non-response. In een simulatiestudie onderzoeken wij wat er gebeurt wanneer multiple imputation wordt toegepast op een situatie waar er normaal gesproken gewogen zou worden.

Wegen vanuit het design in plaats van wegen achteraf?

Barry Schouten (CBS en Universiteit Utrecht)

Gedifferentieerde waarneemstrategieën (in Engels: adaptive survey design) beogen proxy maten voor nonresponsvertekening te optimaliseren gegeven het beschikbare budget van een survey. Die proxy maten zijn gebaseerd op relevante, externe informatie bij een survey afkomstig uit de steekproefkaders, andere gekoppelde registraties, eerdere golven van de survey of uit paradata. Gedifferentieerde waarneemstrategieën zijn daarmee een vorm van wegen

vanuit het design omdat ze al tijdens de dataverzameling respons balanceren op relevante variabelen. De keuze van variabelen en de stratificering in dergelijke designs lijkt sterk op die bij nonresponscorrectie achteraf. De terechte vraag is of dat balanceren dan niet beter achteraf kan gebeuren. In de presentatie bespreek ik de condities waaronder gedifferentieerde waarneemstrategieën zinvol zijn.

Over wegingen in marktonderzoek
Harm Hartman (Ipsos, Nederland)

Al tien jaar geldt de MOA Gouden Standaard, een succesvolle samenwerking tussen marktonderzoekbranche en CBS, als leverancier van weeginput voor marktonderzoek. Even zo lang is er discussie hoe online surveys gewogen moeten worden, omdat specifieke weeginformatie over de online populatie geen onderdeel uitmaakt van de Gouden Standaard. Er zijn grofweg drie strategieën om met de selectiviteit van online surveys om te gaan: negeren, voorkomen of corrigeren. De meest gebruikte strategie is het negeren van het probleem en online onderzoek gewoon te wegen op een set socio-demografische kenmerken. Aangezien de beste weging geen weging is, kun je weging ook voorkomen door de quoterings aan te passen bij de uitzet van het sample. Propensity weighting is een manier om te corrigeren voor de specifieke selectiviteit van online surveys, maar deze techniek wordt weinig geaccepteerd. De afnemende bereidheid van respondenten om deel te nemen aan (lange) surveys, de slechte reputatie van access panels, nopen tot nieuwe strategieën. Tot slot zal een toekomstvisie worden gegeven waarin datafusie de rol van wegingen min of meer overneemt.

Incomplete data: Analysis and sensitivity analysis
Geert Molenberghs – Universiteit van Hasselt en Katholieke Universiteit Leuven

Over the last decade a variety of models to analyze incomplete multivariate and longitudinal data have been proposed, many of which allowing for the missingness to be not at random (MNAR), in the sense that the unobserved measurements influence the process governing missingness, in addition to influences coming from observed measurements and/or covariates. The fundamental problems implied by such models, to which we refer as sensitivity to unverifiable modeling assumptions, has, in turn, sparked off various strands of research in what is now termed sensitivity analysis. The nature of sensitivity originates from the fact that an MNAR model is not fully verifiable from the data, rendering the empirical distinction between MNAR and random missingness (MAR), where only covariates and observed outcomes influence missingness, hard or even impossible, unless one is prepared to accept the posited MNAR model in an unquestioning way. In this paper, we show that the empirical distinction between MAR and MNAR is not possible, in the sense that each MNAR model fit to a set of observed data can be reproduced exactly by an MAR counterpart. Of course, such a pair of models will produce different predictions of the unobserved outcomes, given the observed ones. Theoretical considerations are supplemented with an illustration based on the Slovenian Public Opinion survey, analyzed before in the context of sensitivity analysis.

Missing data can be seen as latent variables. Such a view allows extension of our results to other forms of coarsening, such as grouping and censoring. In addition,

the technology applies to random effects models, where a parametric form for the random effects can be replaced by certain other parametric (and non-parametric) form, without distorting the model's fit, latent classes, latent variables, etc.

The concepts developed are illustrated using the relatively simple but illustrative data from the Slovenian Public Opinion survey.