ABSTRACTS NPSO INNOVATIEDAG - 26 NOVEMBER 2019

PLENAIRE SESSIE I

Squats in surveys: the use of accelerometers for fitness tasks in surveys.

Anne Elevelt – Universiteit Utrecht

Smartphones are becoming increasingly important and widely-used in survey completion. Smartphones also offer many new possibilities for survey research, such as extending data collection by using sensor data (e.g., acceleration). Sensor data, for instance, can be used as a more objective supplement to health and physical fitness measures in mobile web surveys. In this study, we therefore investigate respondents' willingness to participate in fitness tasks during mobile web survey completion. In addition, we investigate the appropriateness of acceleration data to draw conclusions about respondents' health and fitness level. For this purpose, we use "SurveyMotion (SM)," a JavaScript-based tool for smartphones to gather the acceleration of smartphones during survey completion and additionally employ traditional health and physical fitness measures. We ask respondents if they would generally be willing to take part in a fitness task during mobile web survey completion and employ a subsequent fitness task in which we ask respondents to do squats (knee bends) for one minute. Thus, we investigate respondents' hypothetical as well as actual willingness and the general comparability of acceleration data with established health and physical fitness measures. We conduct an observational study by using a German nonprobability-based web panel with n = 1,500 respondents; the data collection takes place in September 2018. This study contributes to the development of more objective measures of respondents' health and fitness in mobile web surveys and could be extended by further physical activity tasks in future research.

Why is half my data missing? Identifying and assessing the impact of missingness mechanisms in always-on location data from smartphones

Danielle McCool, Universiteit Utrecht / Centraal Bureau voor de Statistiek

Abstract: Early results from a mobile travel diary app study conducted in late 2018 show promising results for response rate and general enthusiasm. Early results suggest that collecting sensor data with respondent's own smartphones does, in fact, reduce cost, burden, and some human error. However, researchers are faced with an unanticipated hurdle: are the technological issues severe enough to impact the data quality? And if so, what can be done?

Graph-based inference from non-probability road sensor data

Jonas Klingwort – Universiteit Duisburg / Centraal Bureau voor de Statistiek

Big data from sensors is receiving increasing attention in official statistics for its informational power. More specifically, big data from road sensors has the potential to report official statistics more frequently and on a higher level of detail compared with sample surveys. However, in the absence of a study design and knowledge about the unknown data-generating process, design-based inference methods are not suitable for such non-probability based data. Therefore, research on methods to use big data for population inference in official statistics is required. In this contribution, we aim to infer the truck traffic distribution in the Dutch road network from 16 sensors installed on a non-probability sample of road segments. Using web-scraping, we constructed a directed and weighted graph of the

Dutch transport network. The relationship between graph features and traffic counts on the graph edges is modeled using a GLM with logit link and binomial error distribution. The model is used to predict traffic counts on road segments without sensors. Preliminary results show that the proposed methodology is capable of inferring the traffic distribution. Ideas for improvements will be discussed in detail.

A Bayesian analysis of nonresponse for adaptive survey design accounting for time change in response propensities

Shiya Wu – Universiteit Utrecht

Motivated by the desire to adapt survey design, accurate estimation of main survey design parameters, such as response propensities, is crucial. However, the ability to accurately estimate such parameters depends on the time elapsed since fieldwork was last conducted. Ideally, in a time-invariant survey environment, the parameter values are constant and posteriors in a Bayesian analysis will ultimately converge to point masses as data come in. Clearly, the parameters change gradually in time. For example, response rates have been declining for many years. As a result, the Bayesian analysis mixes newer with older data arising from different mechanisms. At any given time the analysis, then, displays a time lag and may be too pessimistic or too optimistic. Earlier research has shown that such "misspecification" may have dramatic consequences to the utility of the framework. Of course, this complication occurs also in non-adaptive designs, but adaptive designs require more detailed information and may be more sensitive to such change when not recognized.

Robustness to time change, may be achieved by inclusion of an additional hierarchical level for time in the Bayesian analysis framework. All or part of the regression parameters in models for response propensities are modelled as time dependent, a latent time series model is introduced for time change, and the parameters in the time series models receive prior distributions. These parameters are included in the joint posterior distribution of interest and MCMC methods need to make draws from these conditional distributions. Although the option is most intuitive and flexible, it is also more complex and may require longer initial periods of survey data collection to provide useful posterior distributions.
The main research question is whether predictions based on models accounting for time change outperform predictions from time-invariant models.

The methodology is applied and evaluated to survey data that stretch several years of data collection. The Dutch Health survey for the years 2015-2019 is used as a case study to explore the methodology and evaluate the Bayesian framework.

A general framework for multiple - recapture estimation that incorporates linkage error correction

Daan Zult – Universiteit Utrecht / Centraal Bureau voor de Statistiek

The size of a partly observed population is often estimated with the capture – recapture (for two sources) or multiple – recapture (for multiple sources) estimation method. An important assumption for these models is that records in different sources can be identified such that it is known whether these records belong to the same unit or not, i.e. records can be perfectly linked between sources. This assumption of perfect linkage is of particular relevance if identification is not obtained by some perfect identifier (like a tag or id-code) but by indirect identifiers (like name and address or animal's skin patterns). In that case the perfect linkage assumption is often violated, which in general leads to biased population size estimates. A solution to this problem was provided by Ding and Fienberg (1994), Di Consiglio and

Tuoto (2015) and De Wolf et al. (2018). These authors show how to use linkage probabilities to correct the capture - recapture estimator for linkage errors. Recently, Di Consiglio and Tuoto (2018) extended their method to three sources. In this paper we provide a general framework that allows us to extend this work further in two ways. First, we extend this work further to any number of sources. Second, our framework allows to incorporate covariates in a better way. We do this by generalising the standard log - linear modelling approach used in multiple - recapture estimation such that it incorporates linkage error correction. We show how the method performs in a simulation study with data that resemble real data.

PLENAIRE SESSIE II

Gezinsrelaties tijdens detentie: Een verkenning van de Five Minute Speech Sample

Simon Venema – Rijksuniversiteit Groningen / Hanzehogeschool Groningen

Het onderhouden van positieve gezinsrelaties tijdens detentie is belangrijk voor het welzijn van gedetineerden en hun gezinsleden (Dyer, Pleck, & McBride, 2012; Poehlmann-Tynan & Arditti, 2018). De kwaliteit van gezinsrelaties is echter lastig in kaart te brengen vanwege de complexe aard van relaties tussen mensen. Vragenlijsten en zelfrapportages zijn niet altijd voldoende in staat om deze complexiteit volledig te vangen. Observationeel onderzoek naar interacties tussen familieleden, wat vaak wordt gezien als de gouden standaard om gezinsrelaties in kaart te brengen, is in een detentiesetting vaak moeilijk te realiseren.

Een recente methodologische innovatie in het meten van de kwaliteit van gezinsrelaties is de Five Minute Speech Sample (FMSS). Hierbij wordt de respondent gevraagd om voor 5 minuten te spreken over een gezinslid. De FMSS wordt opgenomen en gecodeerd. De FMSS lijkt een veelbelovend instrument om de kwaliteit van gezinsrelaties te meten vanwege de samenhang met observationele metingen van ouder-kindinteracties (Weston, Hawes, & Pasalich, 2017) en met het welzijn en gedrag van kinderen (Bullock & Dishion, 2007; Psychogiou, Daley, Thompson, & Sonuga-Barke, 2007).

De FMSS is nog niet eerder toegepast in een detentiesetting. In ons lopende onderzoek hebben wij de FMSS bij 20 gedetineerde vaders in de penitentiaire inrichting in Veenhuizen afgenomen. Tijdens deze presentatie wordt ingegaan op de eerste resultaten van deze gegevensverzameling en worden nieuwe vraagstukken rondom het meten van gezinsrelaties binnen en buiten detentie verkend.

Referenties

Bullock, B. M., & Dishion, T. J. (2007). Family processes and adolescent problem behavior: Integrating relationship narratives into understanding development and change. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*(3), 396–407. https://doi.org/10.1097/chi.0b013e31802d0b27

Dyer, W. J., Pleck, J. H., & McBride, B. A. (2012). Imprisoned fathers and their family relationships: A 40-year review from a multi-theory view. *Journal of Family Theory & Review*, *4*(1), 20–47.

Poehlmann-Tynan, J., & Arditti, J. A. (2018). Developmental and family perspectives on parental incarceration. In C. Wildeman, A. R. Haskins, & J. Poehlmann-Tynan (Eds.), *When parents are incarcerated: Interdisciplinary research and interventions to support children.* (pp. 53–81). https://doi.org/http://dx.doi.org/10.1037/0000062-004

Psychogiou, L., Daley, D. M., Thompson, M. J., & Sonuga-Barke, E. J. S. (2007). Mothers' expressed emotion toward their school-aged sons: Associations with child and maternal symptoms of psychopathology. *European Child and Adolescent Psychiatry*, *16*(7), 458–464. https://doi.org/10.1007/s00787-007-0619-y

Weston, S., Hawes, D. J., & Pasalich, D. S. (2017). The Five Minute Speech Sample as a Measure of Parent–Child Dynamics: Evidence from Observational Research. *Journal of Child and Family Studies*, *26*(1), 118–136. https://doi.org/10.1007/s10826-016-0549-8

Zeggen foto's meer dan woorden? Een studie aan de hand van foto-elicitatie technieken om de perceptie van ouderen met een lage SES jegens fysieke, sociale en mentale gezondheid te onderzoeken.

Feline Platzer – Rijksuniversiteit Groningen

Achtergrond: In onderzoek naar gezondheid worden vaak fysieke, sociale en mentale aspecten onderscheiden waarbij zowel vragenlijsten als interviews worden gebruikt. Maar schriftelijke of mondelinge communicatie met ouderen met een lage sociaaleconomische status kan beperkte informatie opleveren. Het gebruik van visueel materiaal in onderzoek kan zorgen voor andere inzichten. In de huidige studie wordt onderzocht wat de percepties zijn van ouderen met een lage sociaaleconomische status over fysieke, mentale en sociale gezondheid met gebruik van foto-elictatie interviews.
Methode: Online zijn 48 foto's verzameld gericht op fysieke, mentale of sociale gezondheid. Een definitieve selectie van de foto's is gekozen in samenwerking met zes vertegenwoordigers van de doelgroep. Uiteindelijk zijn er tien foto's gekozen die de meeste associaties met de drie aspecten van gezondheid opriepen. De 10 foto's zijn vervolgens getest in twee focusgroepen en vier individuele interviews. Na de testfase zijn vervolgens 17 foto-elicitatie interviews afgenomen. De interviews worden momenteel geanalyseerd aan de hand van een thematische analyse.
Voorlopige resultaten: Participanten gaven aan dat het gebruik van foto's een leuke manier is om met elkaar te praten over gezondheid en dat het interview hun aan het denken heeft gezet over de eigen gezondheid. Sommige participanten hadden moeite om de situatie afgebeeld op de foto op de eigen gezondheid te betrekken.
Conclusie: Deze studie laat zien hoe een foto-elictatie studie opgezet kan worden bij ouderen met een lage sociaaleconomische status. Resultaten van de thematische analyse moeten nog uitwijzen of de foto-elicitatie interviews meer inzicht geven in de percepties van gezondheid in deze doelgroep.

Web versus mondelinge en schriftelijke vragenlijstgegevens voor gezondheidsenquêtes: de impact op de prevalentie van gezondheidsindicatoren

Elise Braekman – Universiteit Antwerpen

Introductie: In het kader van de Belgische Gezondheidsenquête (BHIS) werd, parallel aan de face-to-face (F2F) survey inclusief schriftelijke vragenlijst voor de meest gevoelige onderwerpen, een web survey georganiseerd. In functie van deze verschillende dataverzamelingsmodi varieerden ook andere elementen van de dataverzameling (bv. de rekrutering van de geselecteerde individuen, incentives, wijze van steekproefname). Er werd onderzocht of de prevalentie van indicatoren rond gezondheidsstatus, consumptie van gezondheidszorg en levensstijl verschilden naargelang de wijze van gegevensverzameling.
Methoden: De gegevens werden verkregen door middel van twee cross-sectionele studies georganiseerd in de algemene bevolking: de BHISWEB (web survey, n=1.010) en de BHIS2018 (F2F survey met schriftelijke vragenlijst, n=2.748). Logistische regressieanalyses werden gebruikt om na te gaan of verschillen in prevalenties te maken hadden met de wijze van gegevensverzameling – onder controle van demografische kenmerken.
Resultaten: Met betrekking tot de schriftelijke versus web vragenlijst, werden er significante verschillen gevonden voor de prevalentie van 2 van de 9 onderzochte gezondheidsindicatoren. Van de indicatoren die via de F2F versus web vragenlijst werden verzameld, vertoonden 9 van de 18 indicatoren significante verschillen die terug gebracht kunnen worden tot de verschillende wijze van gegevensverzameling.

Conclusies: Indicatoren die werden verzameld via de web en schriftelijke vragenlijst waren over het algemeen minder verschillend dan indicatoren die werden verzameld via de web en F2F modus. Bovendien werden minder verschillen vastgesteld voor indicatoren die gebaseerd zijn op eenvoudige en feitelijke vragen dan voor indicatoren die gebaseerd zijn op subjectieve of complexe vragen.

Capturing dynamics of psychopathology using experience sampling methods

IJsbrand Leertouwer – Universiteit van Tilburg

Experience sampling methods (ESM) are a hot topic in psychological research. In an ESM data collection effort, participants fill in multiple questionnaires per day at semi-random timepoints (i.e., random within certain time blocks), about their current experiences and environment. A prominent application of these methods is to provide patients with personalized feedback about (dynamics in) these experiences. Quantification of such feedback however, is in development. One of the directions to explore, pertains to specific measurement issues that psychopathological variables come with. For example, in order to provide the fearful patient with relevant feedback, reports on their most fearful experiences are crucial, yet hard to come by using semi-random sampling of experiences in the exact moment. This presentation will provide suggestions to deal with such challenges that arise when the goal is to provide patients with relevant feedback about their moment-to-moment experiences.

PLENAIRE SESSIE III

The effect of measurement error on clustering algorithms: a sensitivity analysis

Paulina Pankowska – Vrije Universiteit Amsterdam

Social scientists, biostatisticians, data scientists and many others often employ a variety of clustering techniques in order to separate survey data into interesting groups for further analysis or interpretation. Examples of well-known algorithms in the social and medical sciences as well as the data mining literature are K-means, DBSCAN, PAM, Ward, and Gaussian mixture models.

Surveys, however, are well-known to contain measurement errors. Such errors may adversely affect clustering - for instance, by producing spurious clusters, or by obscuring clusters that would have been detectable without errors. Furthermore, measurement error might reduce intra-cluster homogeneity as well as inter-cluster separation. Yet, to date, little work has examined the effect that such errors may exert on commonly used clustering techniques and how any potential bias could be corrected for.

While there is some literature on the topic and a few adaptations to specific clustering algorithms exist to make them "error-aware", the research available focuses predominantly on random measurement error and there is virtually no mention of systematic errors. In addition, the studies available often assume that extent of measurement error is known a priori, an assumption which is rarely fulfilled in practice.

Therefore, in this paper, we investigate the sensitivity of commonly used model- and density-based clustering algorithms (i.e. Gaussian Mixture Models- GMMs- and DBSCAN, respectively) to differing magnitudes and types of random and systematic measurement errors, through a simulation study. In doing so we look both at the effects on the number of clusters found and the similarity of these clusters to the ones obtained in the absence of measurement error.

Assessing response quality by using multivariate control charts for numerical and categorical response quality indicators

Jiayun Jin – Katholieke Universiteit Leuven

When assessing interview response quality and identifying potentially low-quality interviews, both numerical and categorical response quality indicators (mixed indicators) are usually available. Research on how to use them simultaneously, however, is very rare. In this article, we extend the applications of conventional multivariate control charts to response quality indicators that are of mixed types. The data we analyze is from the eighth round of the European Social Survey in Belgium, characterized by six numerical and two categorical response quality indicators. First, we employ a Principal Component Analysis Mix procedure (PCA Mix) to transform the mixed quality indicators into principal components. These principal component scores are subsequently used to construct a Hotelling T^2 statistic. In order to deal with the non-multivariate normal nature of the principal component scores obtained from the PCA Mix, a non-parametric bootstrap method is then applied to calculate the control limit of the T^2 statistic. Second, we suggest tools to interpret an identified outlier in terms of finding the responsible original indicator(s). Third, we present a cyclic procedure for determining the in-control data, by iteratively removing the outliers until the process is considered as in control. Lastly, we identify the most important indicators that discriminate the outliers from the in-control data. The results of this study imply that multivariate control charts based on multivariate projection tools such as PCA Mix in combination with the bootstrap technique have great potential in evaluating interview response quality and identifying outliers.

Estimation of time-varying state correlation in state space models

Caterina Schiavoni – Universiteit Maastricht

Statistics Netherlands uses a state space model to estimate the Dutch unemployment by using monthly series about the labour force surveys (LFS). More accurate estimates of this variable can be obtained by including auxiliary information in the model, such as the univariate administrative series of claimant counts and high-dimensional Google searches related to the unemployment. The former series is observed on a monthly basis and is subject to a one-month publication delay, as the LFS, whereas the Google searches are available in real-time and therefore allow to nowcast the unemployment before LFS data becomes available. Legislative changes may affect the relation between unemployment and claimant counts. Additionally, the relevance of both specific search terms as well as internet search behaviour might change over time. We therefore propose a generalized autoregressive score (GAS), a cubic splines and a kernel density estimation approach to estimate the relations between the unemployment and the above-mentioned auxiliary series as time-varying, while preserving the linearity of the state space model. Each estimation method comes with several tests for the null hypothesis of constant relations. We conduct a simulation study in order to assess the performance of all tests and estimation methods.

Multivariate density estimation by artificial neural networks

Dewi Peerlings – Universiteit Maastricht

In this paper, we focus on obtaining a probability density function (PDF) to assess the properties of the data generating process (DGP) underlying a certain data set using artificial neural networks. This knowledge about the distributional characteristics of the underlying DGP can be used in filtering to acquire better predictions of the signal of interest [Arulampalam et al., 2002]. These predictions solve

the issue of high volatility, missing observations and sample selectivity that characterize the data set which is generated as a by-product of processes not related to statistical production purposes. In this way, these cleaned data sets can be used for official statistics. The proposed artificial neural networks have additional advantages compared to well-known parametric and nonparametric density estimators. Our approach builds on the literature on cumulative distribution function (CDF) estimation using neural networks [Magdon-Ismail and Atiya, 2002]. We extend this literature by providing the analytical derivatives of the obtained CDF from the artificial neural network. Our approach hence removes the approximation error in the second step of obtaining the PDF from the CDF output, leading to more accurate PDF estimates. We show that the proposed solution to obtain the PDF from the CDF output of an artificial neural network holds in a multivariate setting and for a neural network with several hidden layers. We illustrate the accuracy gains from our proposed method using several simulation examples, where the real data generating process is known, hence the improvements in accuracy compared to existing methods can be assessed. More specifically, we follow the illustrations in continuous data cases of [Magdon-Ismail and Atiya, 2002] and [Trentin et al., 2018], a discrete data application and a multivariate data application.