

SESSION 1.4: FEBRUARY 8, 2022

A practical example: types of manuscripts

Denise Silva

TYPES OF MANUSCRIPTS

Journals publish manuscripts of different formats and types:

- **Expository manuscripts**
- **Review manuscripts**
- **Research paper /Original work**
- **Technical notes**

Also,

- Abstracts
- Tutorials
- Interview papers
- Book Reviews
- Discussion papers

TYPES OF MANUSCRIPTS

A manuscript may address:

- **Theoretical work /methods**

- Presents new method, test or procedure. The work may be completely new or may demonstrate an improvement of an existing method
- Describes a verifiable advance on what is currently available

- **Empirical work/applications**

- Presents novel analyses of data not considered or addressed in the literature
- Reports analyses and interpretation based on data (from observation, study, experimentation, surveys, etc.) collected/organized to conduct an original study
- Provides data analysis with the goal of generating knowledge

<https://www.ilovephd.com/5-types-of-manuscripts-journal-publications/>

https://www.jmu.edu/uwc/files/link-library/empirical/empirical_research_article_overview.pdf

https://libguides.csmd.edu/empirical_research

TYPES OF MANUSCRIPTS - EXPOSITORY

- Describes a subject/case/example or a method in a certain detail, without much interest in alternative approaches, to inform the audience about this specific issue
- Focus can be on a subject, region, a certain period or a certain group of households/businesses
- The context for the situation in the specific case is relevant
- The novelty is the report of the method or a specific analysis for a region/period, giving deeper insight to the subject, usefulness of certain data, techniques, etc.

<https://www.ilovephd.com/5-types-of-manuscripts-journal-publications/>

https://www.jmu.edu/uwc/files/link-library/empirical/empirical_research_article_overview.pdf

https://libguides.csmd.edu/empirical_research

TYPES OF MANUSCRIPTS - EXPOSITORY

Expository papers seeks to explain or 'expose' a subject or a case

International Statistical Review (ISR): expository papers on emerging areas of research or application

The Survey Statistician: Ask the Experts or New and Emerging Methods sections

The Survey Statistician. 2021. Vol. 84. 12-17.



Ask the Experts

Machine Learning from the Perspective of Official Statistic

Marco J. H. Puts¹ and Piet J. H. Daas²

¹Statistics Netherlands, The Netherlands, m.puts@cbs.nl

²Statistics Netherlands, The Netherlands, pjh.daas@cbs.nl

Abstract

Artificial Intelligence methods enable the extraction of information from large amounts of data. Machine learning is a sub field of Artificial Intelligence. In this paper, Artificial Intelligence and Machine Learning are introduced and discussed in the context of applying them to produce official statistics. The five quality dimensions of statistical output are used to identify the challenges for their application. The paper ends with a list of the most important research topics that need to be studied to enable the successful application of those methods for official statistics.

Keywords: Artificial Intelligence, Big data, Data science, Output quality.

http://isi-iass.org/home/wp-content/uploads/Survey_Statistician_2021_July_N84_02.pdf

TYPES OF MANUSCRIPTS - EXPOSITORY

- **Machine Learning from the Perspective of Official Statistic**
- **The Survey Statistician, 2021, Vol. 84, 12–17.**

http://isi-iass.org/home/wp-content/uploads/Survey_Statistician_2021_July_N84_02.pdf

1 Introduction

2 Quality of Statistics

2.1 Accessibility and Clarity

2.2 Coherence and Comparability

2.3 Accuracy and Reliability

3 Conclusions

“However, to fully enable the use of ML algorithm in official statistics a number of challenges need to be solved. According to us, the following topics need to be studied within the realm of official statistics:”

References

TYPES OF MANUSCRIPTS - EXPOSITORY

1.Introduction

2.The Traditional Typology of Data Uncertainties

2.1.Sampling Errors

2.2.Nonsampling Errors

3.Manski's (2015) Typology of Data Uncertainties

3.1.A Bayesian Approach

3.2.Measures of Transitory, Permanent and Conceptual Errors

4.Consequences of Uncertainty

4.1.Revisions: Real-Time Data Analysis

4.2.Case-study on GDP: Cross-country Comparisons Measuring Data Revisions

4.3.Trade-Off Between Timeliness and Accuracy

5.Conclusion

6.References

Journal of Official Statistics, Vol. 37, No. 2, 2021, pp. 289–316 <http://dx.doi.org/10.2478/JOS-2021-0013>

Measuring and Communicating the Uncertainty in Official Economic Statistics

Gian Luigi Mazzi¹, James Mitchell², and Florabela Carausu³

Official economic statistics are uncertain even if not always interpreted or treated as such. From a historical perspective, this article reviews different categorisations of data uncertainty, specifically the traditional typology that distinguishes sampling from nonsampling errors and a newer typology of Manski (2015). Throughout, the importance of measuring and communicating these uncertainties is emphasised, as hard as it can prove to measure some sources of data uncertainty, especially those relevant to administrative and big data sets.

Accordingly, this article both seeks to encourage further work into the measurement and communication of data uncertainty in general and to introduce the **Comunikos** (COMmunicating UNcertainty In Key Official Statistics) project at Eurostat. **Comunikos** is designed to evaluate alternative ways of measuring and communicating data uncertainty specifically in contexts relevant to official economic statistics.

Key words: Measurement error; data revisions; official economic statistics; data communication.

<https://sciendo.com/article/10.2478/jos-2021-0013>

TYPES OF MANUSCRIPTS – REVIEW

Summarize the existing literature on a topic to describe the current state of understanding on the subject

Focus is on a specific research field, set of theories, methods or applications

Review should be exhaustive to provide a fair and clear overview

Could bring together several approaches/theories under one umbrella as a novelty

This paper cites 99 references



The screenshot shows the top portion of a research article page. At the top left is the ISI logo and the journal title 'International Statistical Review'. Below this, it indicates 'Original Article' and 'Full Access'. The main title of the article is 'A Review of Spatial Causal Inference Methods for Environmental and Epidemiological Applications'. The authors listed are Brian J. Reich, Shu Yang, Yawen Guan, Andrew B. Giffin, Matthew J. Miller, and Ana Rappold. The publication date is 'First published: 31 May 2021' and the DOI is 'https://doi.org/10.1111/insr.12452'. There is one citation listed. Navigation options include 'SECTIONS', 'PDF', 'TOOLS', and 'SHARE'. The 'Summary' section begins with the text: 'The scientific rigor and computational methods of causal inference have had great impacts on many disciplines but have only recently begun to take hold in spatial applications. Spatial causal inference poses analytic challenges due to complex correlation structures and interference between the treatment at one location and the outcomes at others. In this paper, we review the current literature on spatial causal inference and identify areas of future work. We first discuss methods that exploit spatial structure to account for unmeasured confounding variables. We then discuss causal analysis in the presence of spatial interference including several common assumptions used to reduce the complexity of the interference patterns under consideration. These methods are extended to the spatiotemporal case where we compare and contrast the potential outcomes framework with Granger causality and to geostatistical analyses involving spatial random fields of treatments and responses. The methods are introduced in the context of observational environmental and epidemiological studies and are compared using both a simulation study and analysis of the effect of ambient air pollution on COVID-19 mortality rate. Code to implement many of the methods using the popular Bayesian software OpenBUGS is provided.'

<https://onlinelibrary.wiley.com/doi/10.1111/insr.12452>

TYPES OF MANUSCRIPTS – REVIEW



Original Article | Full Access

A Review of Spatial Causal Inference Methods for Environmental and Epidemiological Applications

Brian J. Reich Shu Yang, Yawen Guan, Andrew B. Giffin, Matthew J. Miller, Ana Rappold

First published: 31 May 2021 | <https://doi.org/10.1111/insr.12452> | Citations: 1

1 Introduction

2 Adjusting for Spatial Confounders

2.1 Review of Spatial Confounding

2.2 Potential Outcomes Framework

2.3 Case-Control Matching Methods

2.4 Neighbourhood Adjustments

2.5 Propensity Score Methods

2.6 Instrumental Variables

2.7 Structural Equation Modelling

2.8 Simulation Study

2.9 Effect of PM_{2.5} Exposure on COVID-19 Mortality

3 Methods for Spatial Interference/Spillover

3.1 Potential Outcomes Framework

3.2 Partial Interference

3.3 Spatial Network Interference

3.4 Process-Based Spillover Models

4 Spatiotemporal Methods

4.1 Testing for Missing Spatial Confounders

4.2 Difference in Difference Methods

4.3 Granger Causality

5 Methods for Point-Referenced Data

5.1 Potential Outcomes Framework

5.2 Matching Methods

5.3 Regression Discontinuity

5.4 Neighbourhood Adjustments

5.5 Spillover/Interference Methods

6 Summary and Future Work

Acknowledgments and References

<https://onlinelibrary.wiley.com/doi/10.1111/insr.12452>

TYPES OF MANUSCRIPTS

RESEARCH PAPERS/ORIGINAL WORK

- Papers to promote the understanding and advancement of an area (official/survey statistics) and to foster its development
- Usually present original work, but may not contain strictly novel material
- Researchers and practitioners report first-hand about their new research/work
- Report a full analysis in a new domain, method or (empirical) application of data
- Present novel analysis of data not considered before, provide data analysis to generate knowledge
- Focus is on the specific (new) theory/method/data/analysis used and the corresponding results
- It is important that the audience can follow and reproduce the analysis
- The novelty is the conceptual, methodological or analytical approach

<https://guides.library.utoronto.ca/bioa01-02/bioa01-types>

RESEARCH PAPER - EMPIRICAL WORK

Empirical Articles

Empirical articles are based on an experiment or study. The authors will report the purpose of the study, the research methodology, and results. This is a familiar structure for empirical articles (IMRAD):

- introduction
- methods
- results
- discussion

In describing the purpose of their study, authors will present a mini literature review to discuss how previous research has led up to their original research project.

Also called:

- primary research article/source
- primary literature article
- original research article

Example: The prevalence of sleep disorders in college students: Impact on academic performance

Application of statistical methods motivated by real-world problems and examples

Statistical Journal of the IAOS 37 (2021) 123–135
DOI 10.3233/SJI-200676
IOS Press

SJIAOS

123

Household consumption allocation and the collective household model: Children share of household resources in The Gambia

Madi Mangan

Statistician, National Accounts, Economic Statistics Division, The Gambia Bureau of Statistics (GBoS), Kanifing Institutional Layout, P.O. Box. 3504, Serekunda, The Gambia
E-mail: manganmadi@gmail.com

Abstract. This paper applies the collective household model to allocate household resources among household members. With a Collective Quadratic Almost Ideal Demand System (CQUAIDS) estimated by a Feasible Generalized Nonlinear Least Squares (FGNLS) method, it studies the household demand for six categories of household goods using household income and expenditure survey data from The Gambia, directed to studying the allocation of resources among young and adult members of households in The Gambia. It establishes the sharing rule for children and adult members of the household and shows the effect of demographic, distributive factor, price and income elasticities on the shares of household resources. The results establish that a higher share of resources goes for children while the sharing rule varies for different household types. Also, the findings show significant effects of demographic, distributive factor, price and income on the allocation of the household resources of consumption goods by the household.

https://libguides.ucmerced.edu/wri101_downey/articles

<https://content.iospress.com/download/statistical-journal-of-the-iaos/sji200676?id=statistical-journal-of-the-iaos%2Fsji200676>

RESEARCH PAPER - EMPIRICAL WORK

1. Introduction

2. Methodology

3. Data, findings and discussion of results

3.1 Consumption expenditure and the compilation of the consumption data

3.2 Sharing rule

3.3 Demographic effects

3.4 Income/expenditure effect

3.5 Price and income elasticities

4. Conclusion and recommendation

Notes

Acknowledgments

References

Statistical Journal of the IAOS 37 (2021) 123–135
DOI 10.3233/SJI-200676
IOS Press

123

Household consumption allocation and the collective household model: Children share of household resources in The Gambia

Madi Mangan

Statistician, National Accounts, Economic Statistics Division, The Gambia Bureau of Statistics (GBoS), Kanifing Institutional Layout, P.O. Box. 3504, Serekunda, The Gambia
E-mail: manganmadi@gmail.com

In this work, the collective household framework is employed to define and estimate the sharing rule with the household demand system. The collective model is built on the fundamentals of methodological individualism.

It addresses the question of how individual preferences lead to a collective choice and how household members reconcile different preference, introduces children to this framework (collective household) as agents and not as a public good in order to be able to identify children's share of household resources.

RESEARCH PAPER - EMPIRICAL WORK

“In this article, **we focus on arguably the most important problem** hindering the application of mobile phone data in official statistics: detecting home locations.

We argue that current efforts to detect home locations suffer from a blind deployment of criteria to define a place of residence and from limited validation possibilities.

We support our argument by analysing the performance of five home detection algorithms (HDAs) that have been applied to a large, French, Call Detailed Record (CDR) data set (,18 million users, five months).”

Journal of Official Statistics, Vol. 34, No. 4, 2018, pp. 935–960. <http://dx.doi.org/10.2478/JOS-2018-0046>

Assessing the Quality of Home Detection from Mobile Phone Data for Official Statistics

Maarten Vanhoof¹, Fernando Reis², Thomas Ploetz³, and Zbigniew Smoreda⁴

Mobile phone data are an interesting new data source for official statistics. However, multiple problems and uncertainties need to be solved before these data can inform, support or even become an integral part of statistical production processes. In this article, we focus on arguably the most important problem hindering the application of mobile phone data in official statistics: detecting home locations. We argue that current efforts to detect home locations suffer from a blind deployment of criteria to define a place of residence and from limited validation possibilities. We support our argument by analysing the performance of five home detection algorithms (HDAs) that have been applied to a large, French, Call Detailed Record (CDR) data set (~ 18 million users, five months). Our results show that criteria choice in HDAs influences the detection of home locations for up to about 40% of users, that HDAs perform poorly when compared with a validation data set (resulting in 35°-gap), and that their performance is sensitive to the time period and the duration of observation. Based on our findings and experiences, we offer several recommendations for official statistics. If adopted, our recommendations would help ensure more reliable use of mobile phone data vis-à-vis official statistics.

Key words: Mobile phone data; home location; home detection algorithms; official statistics; big data.

<https://sciendo.com/article/10.2478/jos-2018-0046>

EMPIRICAL WORK

1. Introduction

2. Mobile Phone Data, Official Statistics and the Role of Home Detection Methods

2.1. Mobile Phone Data and Official Statistics

2.2. The Role and Method of Home Location

3. Identifying Homes from Large-Scale Location Traces

3.1. Identifying Meaningful Places from Continuous Location Traces

3.2. Identifying Meaningful Places from Noncontinuous Traces

3.3. Defining Decision Rules for Single-Step Home Identification

3.4. Validating Large-Scale Home Detection Methods

3.5. Current Deficits of Home Detection Using Noncontinuous Location Traces

4. Investigating Home Detection Algorithms for French CDR Data

4.1. The French CDR Data Set

4.2. Applying Five HDAs to the French CDR Data

4.3. Comparison of HDAs at Individual Level

4.4. High Level Validation of Home Detection Algorithms

5. Discussion

5.1. Differences at Individual Level and the Absence of Ground Truth Data

5.2. Sensitivity of Performance Considering Time and Decision Rule Choice

5.3. The 358-gap in High-Level Validation

6. Recommendations

7. Conclusions

8. References

RESEARCH PAPER THEORETICAL WORK/METHODS

- Section 1. Introduction
- Section 2. Double hot-deck boundary information matching proportioned residual draw
- Section 3. Simulation
- Section 4. Empirical analysis
- Section 5. Conclusion
- Acknowledgements
- Appendix
- References

Statistics Canada Statistique Canada

Search website

Subjects Data Analysis Reference Geography Census Surveys and statistical programs About StatCan Canada.ca

Home > 12-001-X > Survey Methodology > Main page

Survey Methodology

A new double hot-deck imputation method for missing values under boundary conditions

by Yousung Park and Tae Yeon Kwon ¹

Release date: June 30, 2020

More information PDF version

Next →

Abstract

In surveys, logical boundaries among variables or among waves of surveys make imputation of missing values complicated. We propose a new regression-based multiple imputation method to deal with survey nonresponses with two-sided logical boundaries. This imputation method automatically satisfies the boundary conditions without an additional acceptance/rejection procedure and utilizes the boundary information to derive an imputed value and to determine the suitability of the imputed value. Simulation results show that our new imputation method outperforms the existing imputation methods for both mean and quantile estimations regardless of missing rates, error distributions, and missing-mechanisms. We apply our method to impute the self-reported variable “years of smoking” in successive health screenings of Koreans.

Key Words: Hot-deck; Two-sided boundary conditions; Multiple imputation; Item nonresponse.

Appendix

Proof of Theorem 1

It suffices to show that $Y_{jU}^* \leq C_{jU}$ and $Y_{jL}^* \geq C_{jL}$ because of the constraints in the imputation step.

1. If $\hat{Y}_j^{\text{miss}} \in S^0$, then \tilde{r}_{jU}^* is sampled from R_U^0 whose element $\tilde{r}_{iU} \leq 1$ for all i because $Y_i \leq C_{iU}$ and $C_{iU} - \hat{Y}_i > 0$ for $\tilde{r}_{iU} \in R_U^0$. Since \tilde{r}_{jU}^* is one of such \tilde{r}_{iU} 's, we have $\tilde{r}_{jU}^* \leq 1$. Furthermore $C_{jU} - \hat{Y}_j^{\text{miss}} > 0$ gives

$$Y_{j,U}^* = \hat{Y}_j^{\text{miss}} + \tilde{r}_{j,U}^* (C_{j,U} - \hat{Y}_j^{\text{miss}}) \leq \hat{Y}_j^{\text{miss}} + 1 \times (C_{j,U} - \hat{Y}_j^{\text{miss}}) = C_{jU}. \quad (\text{A.1})$$

RESEARCH PAPER

THEORETICAL WORK/METHODS

As an alternative to integrated weighting, **we propose a two-level generalised regression estimator that is capable of both ensuring consistent person and household estimates and allowing for different weights for persons within a household.**

A Monte Carlo simulation supports the superiority of our two-level generalised regression estimator compared with integrated weighting.

 International Statistical Review

International Statistical Review (2021), 89, 3, 635–656 doi: 10.1111/insr.12460

A Two-level GREG Estimator for Consistent Estimation in Household Surveys

Anne Konrad , Jan Pablo Burgard and Ralf Münnich

Economic and Social Statistics Department, University of Trier, Trier, Germany
E-mail: konrada@uni-trier.de

Summary

Household surveys provide information on both person-level and household-level characteristics. To ensure consistent estimates between both levels, statistical offices often use integrated weights that are equal for all persons within a household and the household itself. However, these integrated weights ignore the individual patterns of the persons, and the heterogeneity within a household is no longer reflected. As an alternative to integrated weighting, we propose a two-level generalised regression estimator that is capable of both ensuring consistent person and household estimates and allowing for different weights for persons within a household. A Monte Carlo simulation supports the superiority of our two-level generalised regression estimator compared with integrated weighting.

Key words: Integrated weighting; household surveys; generalised regression estimator; calibration; consistent estimation.

<https://onlinelibrary.wiley.com/doi/10.1111/insr.12460>

RESEARCH PAPER

THEORETICAL WORK/METHODS

A Two-level GREG Estimator for Consistent Estimation in Household Surveys

Anne Konrad ✉, Jan Pablo Burgard, Ralf Münnich

ISR

1 Introduction

2 Basic Framework

3 Integrated Weighting

3.1 Point and Variance Estimator

3.2 Consequences of Integrated Weighting

4 Proposed Two-level Generalised Regression Estimator

4.1 Core Idea

4.2 Naïve Two-level Generalised Regression Estimator

4.3 Extended Two-level Generalised Regression Estimator

4.4 Comparison of the Naïve and Extended Two-level Generalised Regression Estimator

4.5 Further Remarks

5 Further Weighting Approaches

5.1 Generalised Least Squares Adjustment Algorithm According to Zieschang (1990)

5.2 Generalised Least Squares Adjustment Algorithm According to Merkouris (2004)

5.3 Comparison of the Two-level Generalised Regression Estimator and the Generalised Least Squares Adjustment Algorithm

6 Simulation Study

7 Conclusion

Acknowledgements

References

<https://onlinelibrary.wiley.com/doi/10.1111/insr.12460>

TYPES OF MANUSCRIPTS – TECHNICAL NOTE

Short article providing a brief description of novel aspects of a specific development, technique, procedure or computational method in the scope of the journal

Survey Methodology: publishes innovative theoretical or applied research papers, and sometimes review papers, that provide new insights on statistical methods relevant to National Statistical Offices and other statistical organizations.

✓ **SHORT NOTES** Documents submitted for the short notes section must have a maximum of 3,000 words, including tables, figures and references.

Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

A note on multiply robust predictive mean matching imputation with complex survey data

by Sixia Chen, David Haziza and Alexander Stubblefield

Release date: June 24, 2021



 Statistics Canada Statistique Canada



<https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2021001/article/00009-eng.pdf?st=GzoQ8cP->

TYPES OF MANUSCRIPTS – TECHNICAL NOTE

1. Introduction
 2. Basic setup
 3. Proposed methods
 4. Simulation study
- Acknowledgements
- References

“In this note, we propose a novel predictive mean matching procedure that allows the user to specify multiple outcome regression models”

Survey Methodology, June 2021
Vol. 47, No. 1, pp. 215-222
Statistics Canada, Catalogue No. 12-001-X

A note on multiply robust predictive mean matching imputation with complex survey data

Sixia Chen, David Haziza and Alexander Stubblefield¹

Abstract

Predictive mean matching is a commonly used imputation procedure for addressing the problem of item nonresponse in surveys. The customary approach relies upon the specification of a single outcome regression model. In this note, we propose a novel predictive mean matching procedure that allows the user to specify multiple outcome regression models. The resulting estimator is multiply robust in the sense that it remains consistent if one of the specified outcome regression models is correctly specified. The results from a simulation study suggest that the proposed method performs well in terms of bias and efficiency.

Key Words: Multiple robustness; Nearest-neighbour imputation; Survey data; Variance estimation.

<https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2021001/article/00009-eng.pdf?st=GzoQ8cP->

TYPES OF MANUSCRIPTS

Journals publish manuscripts of different formats and types:

- Expository manuscripts
- Review manuscripts
- Research paper /Original work
- Technical notes

Also,

- Abstracts
- Tutorials
- Interview papers
- Book Reviews
- Discussion papers

OTHER TYPES OF MANUSCRIPTS

- **Abstracts:** summary article of a longer report/research publication, explaining the main issues (background, analysis, results) in brief
- **Tutorials/Overview:** a short review manuscript directed to an application and supporting readers to apply a method/theory/technique
- **Interviews:** On one person/group of persons, can also deal with a specific topic
- **Book reviews:** Provide a brief description of a recently-written book and analyse its contents, highlighting the strengths and weaknesses of the work
- **Discussion papers:** Aim to be provocative on a topic, presenting ideas that form the basis of a discussion on a particular topic

TYPES OF MANUSCRIPTS – INTERVIEWS

- **International Statistical Review (ISR):** Publish interviews with statisticians who have made prominent contributions to applications, research, and to the development of the profession.
- **Statistical Journal of IAOS** also publishes interviews

Interview with Stephen Penneck

SJIAOS

Cite

Article type: Research Article

Authors: West, Kirsten

Affiliations: Interview Editor, Statistical Journal of the IAOS. E-mail: kwestiaos@gmail.com

DOI: 10.3233/SJI-160998

Journal: Statistical Journal of the IAOS, vol. 32, no. 2, pp. 141-148, 2016

Published: 20 May 2016

Get PDF 




Interview with Stephen

Kirsten West
Interview Editor, Statistical Journal of the IAOS
E-mail: kwestiaos@gmail.com



President of ISI

 International Statistical Review

Interview |  Open Access |  

An Interview with Chris Skinner

David Haziza , Paul A. Smith

First published: 26 November 2019 | <https://doi.org/10.1111/insr.12356>

[The copyright line for this article was changed on 27 February 2020 after original online publication]



Chris Skinner receiving CBE (Commander of the Most Excellent Order of the British Empire) from the Queen, Windsor Castle, 2010.