Journal of Statistical Software

MMMMMM YYYY, Volume VV, Issue II. *doi: 10.18637/jss.v000.i00*

nonprobsvy – An R package for modern methods for non-probability surveys

Łukasz Chrostowski
Adam Mickiewicz University

Piotr Chlebicki
Stockholm University

Maciej Beręsewicz
Poznań University of Economics and Business

Abstract

The abstract of the article.

Keywords: keywords, not capitalized, R.

1. Introduction

This template demonstrates some of the basic LaTeX that you need to know to create a JSS article.

1.1. Code formatting

In general, don't use Markdown, but use the more precise LaTeX commands instead:

- Java
- **plyr**

One exception is inline code, which can be written inside a pair of backticks (i.e., using the Markdown syntax).

If you want to use LaTeX commands in headers, you need to provide a `short-title` attribute. You can also provide a custom identifier if necessary. See the header of Section 2 for example.

2. Methods for non-probability samples R code

2.1. Basic setup

Let $U = \{1, ..., N\}$ denote the target population consisting of N labelled units. Each unit i has an associated vector of auxiliary variables $\boldsymbol{x}_i$ (a realisation of the random vector $\boldsymbol{X}_i$ in the super-population) and the study variable y_i (a realisation of the random variable Y_i in the super-population). Let $\{(y_i, \boldsymbol{x}_i), i \in S_A\}$ be a dataset of a non-probability sample of size n_A and let $\{(\boldsymbol{x}_i, \pi_i), i \in S_B\}$ be a dataset of a probability sample of size n_B, where only information about variables $\boldsymbol{X}$ and inclusion probabilities π (which in the super population model are also considered to be random variables) are available. Let δ be an indicator of inclusion into non-probability sample. Each unit in the sample S_B has been assigned a~design-based weight given by $d_i = 1/\pi_i$. The setting is summarised in Table ..

The goal is to estimate a~finite population mean $\mu_y = \frac{1}{N}\sum_{i=1}^{N} y_i$ of the target variable Y. As values of y_i are not observed in the probability sample, it cannot be used to estimate the target quantity. Instead, one could try combining the non-probability and probability samples to estimate μ_y. In this paper we do not consider modifications for the possibly occurring overlap.

2.2. Mass Imputation estimators

2.3. Inverse Probability Weighting estimators

2.4. Doubly Robust estimators

Can be inserted in regular R markdown blocks.

```
R> x <- 1:10
R> x
```

```
 [1]  1  2  3  4  5  6  7  8  9 10
```

2.5. Features specific to rticles

- Adding short titles to section headers is a feature specific to **rticles** (implemented via a Pandoc Lua filter). This feature is currently not supported by Pandoc and we will update this template if it is officially supported in the future.
- Using the `\AND` syntax in the `author` field to add authors on a new line. This is a specific to the `rticles::jss_article` format.

3. Package contents and implementation

4. Practical examples

Affiliation:

Łukasz Chrostowski
Pearson
First line
Second line
E-mail: lukchr@st.amu.edu.pl
URL: https://posit.co

Piotr Chlebicki
Stockholm University
Department of Statistics and Mathematics,
Faculty of Biosciences,
Universitat Autònoma de Barcelona
E-mail: piotr.chlebicki@math.su.se

Maciej Beręsewicz
Statistical Office in Poznań
Department of Statistics,
Institute of Informatics and Electronic Economy,
Poznań University of Economics and Business
E-mail: maciej.beresewicz@ue.poznan.pl
URL: https://posit.co

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
MMMMMM YYYY, Volume VV, Issue II *Submitted:* yyyy-mm-dd
doi:10.18637/jss.v000.i00 *Accepted:* yyyy-mm-dd