Stratified Sampling Designs - Exercise 2

Stefan Zins and Matthias Sand

February, 25th 2017

Exercise 2

Estimation under a stratified design

1. Download the ESS for Sweden and Denmark (round 5)
2. Import the data to R and combine the two datasets
3. Define a `survey` object (stratified design)
4. Estimate the empirical distribution of TV consumption (`tvtot`) in Sweden and Denmark (separately) and the mean and total of those that watch more than 3 hours of TV
5. Estimate the joint empirical distribution of tv consumption in Sweden and Denmark and the mean and total of those that watch more than 3 hours of TV
6. Calculate the poststratification weights for both countries by

$$postweight = pspwght * pwght * 10000$$

7. Conduct a weighted linear regression for the income decile (`hinctnta`) of the joined data set using the `lm` function. Use age (`agea`), household size (`hhmmb`), the integer of ISCED levels (`eisced`) and the total working hours (`wkhtot`) as independent variables
8. Do the same with the generalised linear model (`svyglm`) that is incuded in the `survey` package and compare the results

Codebook ESS round 5