Weight Calibration across Packages

Stas Kolenikov
9/23/2019
Weight calibration

- Last step in creating analysis weights in survey data files
- Adjusting the weights so that they sum to known population totals in different subgroups (age, sex, race, ethnicity, geography, etc.)
- Desirable to minimize changes from the input weights (probability of selection, nonresponse adjustments, frame integration, etc.)

Deville & Sarndal (1992)
Contenders

Stata

- *ipfraking* (Kolenikov 2014, 2019)
- *svyca1* (official Stata)
- *svywgt* (Winter 2002)
- *sreweight* (Pacifico 2014)

R

- *survey::calibrate()* (Lumley 2010)

SAS

- *rake_and_trim()* (Izrael, Battaglia, Hoaglin, Frankel, Ball, 2017)
Out of scope

- SUDAAN PROC WGTADJ, PROC WGTADJX
- Stata ipfweight (Bergmann 2011)
- R library(ReGenesees) (Zardetto 2015)
- R library(ipfr) (Ward, Macfarlane 2019)
Expectations

- Produce usable results
- Provide weight diagnostics
- Speed
- Fool proof
In the original 2019 presentation, there was a picture of a truck with a bumper sticker that said, "Don't CA my TX", but I was asked to remove that picture by a copyright law firm representing the image holder.
https://www.google.com/search?q=don't+ca+my+tx
CPS 2018 March ASEC data

- estimate control totals based on 13353 adults in CA
- calibrate 8403 adults in TX on
  - sex
  - age (14 categories)
  - race/ethnicity (6 categories)
  - education (5 categories)
  - HH income (9 categories)
  - nativity (3 categories)
  - marital status (6 categories)
  - own vs. rent
  - metro area of TX (23 categories)
Tasks and tests

1. Straight raking
2. Raking with divergent population control totals
3. Raking with bounded weight adjustment ratios [0.3,3]
4. Raking with bounded weight values (2nd and 98th percentile of unrestricted distribution)
5. Linear calibration
6. Linear calibration with trimming
7. (Informative error expected) incorrect specification of control totals
## Performance summary

<table>
<thead>
<tr>
<th>package</th>
<th>TOT</th>
<th>DIV</th>
<th>REL</th>
<th>ABS</th>
<th>LIN</th>
<th>LIN+TR</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipfraking</td>
<td>name</td>
<td>+W</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>N</td>
<td>7.14 sec</td>
</tr>
<tr>
<td>svycal</td>
<td>name/=</td>
<td>F</td>
<td>F</td>
<td>N</td>
<td>+</td>
<td>+</td>
<td>0.18 sec</td>
</tr>
<tr>
<td>survwgt</td>
<td>order</td>
<td>NW</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>0.80 sec</td>
</tr>
<tr>
<td>sreweight</td>
<td>order</td>
<td>F</td>
<td>F</td>
<td>N</td>
<td>+</td>
<td>N</td>
<td>0.19 sec</td>
</tr>
<tr>
<td>calibrate</td>
<td>name</td>
<td>-W</td>
<td>..</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0.35 sec</td>
</tr>
<tr>
<td>rake_and_trim</td>
<td>name+magic</td>
<td>-W</td>
<td>+</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>61 sec</td>
</tr>
</tbody>
</table>

**N**: no documented functionality exists

**W**: issued reasonable warnings

**F**: failed with cryptic error message / no message
Stata ipfraking

The output from the ipfraking command shows the number of missing values generated and the iteration details. The summary of weight changes includes the mean, standard deviation, minimum, maximum, and coefficient of variation for the original and trimmed weights. The graphs illustrate the frequency distribution of the calibration weights and the adjustment factor.
Stata svyca1

.svycal rake ibn,xex ibn.necethn6 ibn.edu5 ibn.age_cat ibn.origin3 ibn.hhinc9 ibn.own_rent ibn.marst ///
[pw=svycalwt] if stafip==48, generate(svycalwt) nocons totals( ///
1.xex = 14813330.08 2.xex = 15111779.64 ///
1.necethn6 = 12400052.47 2.necethn6 = 16556499.68 3.necethn6 = 4731138.97 ///
1.own_rent = 8356418.83 2.own_rent = 1430711.25 3.own_rent = 10635383.12 ///
1.edu5 = 4227166.98 2.edu5 = 7862681.95 3.edu5 = 8761781.29 ///
1.age_cat = 13478132.21 2.age_cat = 2092454.42 3.age_cat = 1671947.71 ///
1.age_cat = 2824235.95 2.age_cat = 2665726.94 3.age_cat = 2557485.76 ///
1.origin3 = 29085731.32 2.origin3 = 2435747.59 3.origin3 = 2194585.24 ///
1.hhinc9 = 2124598.55 2.hhinc9 = 1189250.39 ///
1.hhinc9 = 13970105.67 2.hhinc9 = 1976545.81 ///
1.marst = 14215424.56 2.marst = 5584485.04 3.marst = 108425749.7 ///
1.marst = 2918473.86 2.marst = 3957861.36 3.marst = 4113413.54 ///
1.marst = 2005284.07 2.marst = 3082911.69 3.marst = 3963594.29 ///
1.marst = 2929345.18 2.marst = 3082397.77 3.marst = 3872652.96 ///
1.own_rent = 17592993.52 2.own_rent = 1203935.77 ///
1.own_rent = 14840907.15 2.own_rent = 1853577.77 3.own_rent = 667192.81 ///
1.marst = 2621031.03 2.marst = 1020660.57 3.marst = 9979457.98 ///
)

end of do-file
Stata command for weight calculation:

```
foreach x of varlist sex race eth6 educ5 age_cat origin3 hhinc9 own_rent marst metarea {  
    qui gen double t_x' = .
    qui levelso r'
    foreach c of numlist `r(levels)' {  
        if "c' == "metarea" {  
            local where : colnumb cpsi8_metarea_r'[1',where'=] if 'x' == 'c'
            qui replace t_x' = cpsi8_metarea_r'[1',where'=] if 'x' == 'c'
        }  
        else {  
            local where : colnumb cpsi8_x' _one:='c'
            qui replace t_x' = cpsi8_x'[1',where'=] if 'x' == 'c'
        }
    }
}
```

The `svy: regress` command with the `svyset` command:

```
svyset weight_variable, by(sex race eth6 educ5 age_cat origin3 hhinc9 own_rent marst)  
svy: regress outcome_variable independent_variables
```

Warning: `svy: regress` reached maximum iterations before convergence.
R survey::calibrate()
**** Program terminated at iteration 11 because raking converged ****

**The FREQ Procedure**

Weighted Distribution **After Raking**

<table>
<thead>
<tr>
<th>Sex</th>
<th>Output Weight Sum of Weights</th>
<th>Target Total</th>
<th>Sum of Weights Difference</th>
<th>% of Output Weights</th>
<th>Target % of Weights</th>
<th>Difference in %</th>
<th>Marginal Category Difference in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14813715.70</td>
<td>14813331</td>
<td>385.02</td>
<td>49.012</td>
<td>49.011</td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td>2</td>
<td>15410993.62</td>
<td>15411379</td>
<td>-385.02</td>
<td>50.988</td>
<td>50.989</td>
<td>-0.001</td>
<td>-0.002</td>
</tr>
</tbody>
</table>
SAS `rake_and_trim()`
Weight trimming ≠ methodology
### Misspecified control totals

<table>
<thead>
<tr>
<th>package</th>
<th>Extra in pop</th>
<th>Extra in data</th>
<th>Wrong order</th>
</tr>
</thead>
<tbody>
<tr>
<td>ipfraking</td>
<td>E</td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>svycal</td>
<td>E</td>
<td>!!!</td>
<td></td>
</tr>
<tr>
<td>survwgt</td>
<td>N/A</td>
<td>N/A</td>
<td>!!!</td>
</tr>
<tr>
<td>sreweight</td>
<td>E</td>
<td>E</td>
<td>!!!</td>
</tr>
<tr>
<td>calibrate</td>
<td>E</td>
<td>E</td>
<td></td>
</tr>
<tr>
<td>rake_and_trim</td>
<td>!?!?</td>
<td></td>
<td>!!!</td>
</tr>
</tbody>
</table>

E: issued an error and stopped

!!!: did not issue an error – results highly suspect!
Thanks and out

Questions?

- stas_kolenikov@abtassoc.com
- @StatStas on Twitter