

Iterative Proportional Fitting

Iterative proportional fitting (IPF), also known as raking, is a very useful tool once a survey has been conducted. This example shows a very simple IPF algorithm than can be used to adjust survey weights. This technique is usually done when you know the true population values that your survey should match. An example of this would be gender, assuming the population of the United States, would be (approximately) .5 male and .5 female. If, in a survey, the observed proportion ends up being .6 male and .4 female then IPF can be used to adjust the weight to match the known values. The output creates a dataset with a new variable called final.weight that contains the weights that match the targets. With just a little bit of work this example can be extended to handle multiple variables. However, when multiple variables are used then the process needs to be iterated multiple time because after each variable the previous variable loses its convergence. The details of how IPF works and the many issues related is a topic of a longer discussion.

```
library(e1071); #used for the rdiscrete function
n = 500;

raw = as.data.frame(matrix(NA, nrow=n, ncol=7));
names(raw) = c("weight", "x", "y", "tar.wq", "wq", "adj", "final.weight");

raw$weight = rep(1,n);

#raw observed values
raw$x = rdiscrete(n,c(.60,.40),values=1:2);
raw$y = rdiscrete(n,c(.4,.4,.2),values=1:3);

#raking targets
raw$target[raw$x==1] = .5;
raw$target[raw$x==2] = .5;

raw$final.weight = raw$weight;
raw$wq[raw$x==1] = sum(raw$final.weight[raw$x==1]);
raw$wq[raw$x==2] = sum(raw$final.weight[raw$x==2]);

raw$tar.wq = raw$target*sum(raw$final.weight);
raw$adj = raw$tar.wq/raw$wq;
raw$final.weight = raw$adj*raw$final.weight;
```