Weighting in SPSS Tutorial

First, open up the Washington Post/ABC News Pre-election Poll from September 2004 available on the course website. Now, find the unweighted predicted vote percentages. Select Analyze, then Descriptive Statistics, then Frequencies. Select the variable Q3 and add it to the list by pressing the arrow button. Press OK and now take a look at the output.

We see that the unweighted output suggests that 50.7 of those respondents interviewed claimed they would vote for Bush and 41.0 percent claimed they would vote for Kerry.

Next, find the proportions of the sample that identify as Democratic, Republican, and Independent. Select Analyze, then Descriptive Statistics, then Frequencies. Select the variable Q901 and add it to the list by pressing the arrow button. Press OK and examine the output. We see that the sample proportion Democrat is 32.1% and the sample proportion Republican is 31.1%. We categorize all respondents who do not report themselves as Republicans or Democrats as Independents. Thus, we have an Independent sample proportion of 35.8%.

Now, create a new variable consisting only of these three categories for party identification with no missing values. Select Transform, then Recode into Different Variables. Select Q901 from the list of variables and press the arrow button. Enter ‘PID’ in the Output Variable name field, and ‘3 Category PID’ in the Output Variable label field. Click Change.
Now, click the Old and New Values button. Click the Value radio button in the Old Value box, and enter in ‘1’. Click the Value radio button in the New Value box, and enter in ‘1’, and click Add. Now, click the Value radio button in the Old Value box, and enter in ‘2’. Click the Value radio button in the New Value box, and enter in ‘2’, and click Add. Finally, click the All Other Values radio button in the Old Value box, and enter in ‘3’. Click Continue, and OK.

Now, visit [http://www.cnn.com/ELECTION/2004/pages/results/states/US/P/00/epolls.0.html](http://www.cnn.com/ELECTION/2004/pages/results/states/US/P/00/epolls.0.html) and obtain population proportions for party identification based on exit polls in the 2004 election. We see that the population is 37 percent Republican, 37 percent Democratic, and 26 percent Independent.

Now, create a weight for each of the three categories of party ID. For any given observation of category j, the weight of that observation is \( \frac{\pi_{x=j}}{\hat{p}_{x=j}} \). Thus, the weight of any Democratic observation is the proportion of the population that is Democratic divided by the proportion of the sample that is Democratic. This is \( .37/.321 \), or 1.15. We may do the same to find the weight of any Republican observation, which is \( .37/.311 \), or 1.19. Finally we can compute the weight for any independent observation, which is \( .26/.367 \), or .71.

Now, create a new variable consisting of the weights. Click Transform, then Compute Variable. Input ‘pweight’ into the field for Target Variable. Enter the following into the numeric expression box text box,

\[
(pid=1) \times .37/.321 + (pid=2) \times .37/.311 + (pid=3) \times .26/.367
\]
This expression simply set the weight equal to \(0.37/0.321\) for Democrats, \(0.37/0.311\) for Republicans, and \(0.26/0.367\) for independents. Now select OK.

Now, we will check our answer. Select Data, then Weight Cases. Select ‘Weight cases by’ then scroll to the bottom, select variable ‘pweight’ and click the arrow. Select Analyze, then Descriptive Statistics, then Frequencies. Select the variable Q901 and add it to the list by pressing the arrow button. Press OK and now take a look at the output. If you have done things correctly, you should find that the sample now consists of 37% Democrats and 37% Republicans (i.e. the weighting targets).

Now, we will find the new weighted predicted vote percentage. Click OK. Select Analyze, then Descriptive Statistics, then Frequencies. Select the variable Q3 and add it to the list by pressing the arrow button. Press OK and now take a look at the output. We see that the output weighted by party suggests that 51.7 of those respondents interviewed claimed they would vote for Bush and 41.0 percent claimed they would vote for Kerry.