Centering Examples: SPSS and R[1]

The HLM package makes centering (either group- or grand-mean centering) very convenient and self-explanatory. Below, I show the steps I use in SPSS and R to center variables. Grand-mean centering in either package is relatively simple and only requires a couple lines of code (comment lines designated by * are ignored by SPSS and # are ignored in R). Be careful that the means you center on are derived from the same cases used in your analysis. It is always a good idea to check your data at several steps along the way, as I have shown here.

SPSS

Grand-Mean Centering

```
get file='c:\jason\spsswin\mlrclass\hsbmerged.sav'.

* replace [filepath] with location of your data file, for example
    get file='c:\mlrclass\hsbmerged.sav'.

compute tses=(ses*10) + 50.

*grand mean centering.

*routine for listwise deletion to get correct sample size for variables used in the model.
count nummiss=mathach tses meanses sector (missing).
select if nummiss eq 0.

descriptives vars=mathach tses meanses sector.

*The aggregate command obtains mean for the full sample.
AGGREGATE
    /meantses=MEAN(tses).

*computes the grand-mean centered variable gdcses.
compute gdcses=tses - meantses.
```

Group-Mean Centering
Group-mean centering in SPSS proceeds in the same way, but specifying the grouping ID variable is needed on the /BREAK subcommand get group means. I provide some example of SPSS syntax to illustrate centering computations with the HSB data set.

```
get file='c:\jason\spsswin\sqms\hsbmerged.sav'.

* replace [filepath] with location of your data file, for example
    get file='c:\mlrclass\hsbmerged.sav'.

*This computation is irrelevant, but is how I created the TSES var.
compute tses=(ses*10) + 50.

* AGGREGATE creates group means for tses variable.

AGGREGATE
    /BREAK schoolid
    /meantses=MEAN(tses).

*compute the group-mean centered ses using meantses.
compute gpcses=tses - meantses.

descriptives vars=all.
```

---

[1] I strongly recommend thorough consideration of the implications for interpretation before making centering decisions (in particular, the decision not to center). See my other handout on centering issues and other sources, such as Enders and Tofighi (2007) and Algina and Swaminathan (2011).

R Code

There are several ways to center in R, but here is one fairly simple way. The mean() function takes the mean across all cases in the data set (as opposed to across a list of variables in other programs). The ave() function is a handy way to compute the means of the variable within each group. Within the parenthesis, the first variable is the variable to be centered and the second variable is the grouping variable.

```
#recompute ses to T-score values
mydata$tses=(mydata$ses*10) + 50
SummaryStats(tses)
```

Grand-Mean Centering

```
#grand mean centering of tses
mydata$gdcses <- mydata$tses - mean(mydata$tses)
#good idea to double check with some descriptive statistics
SummaryStats(gdcses)
```

Group-Mean Centering

```
#group mean centering tses
mydata$gpcses <- mydata$tses - ave(mydata$tses,mydata$schoolid)
SummaryStats(gpcses)
```

Other Resources

There is a macro for SPSS available that makes this a little bit easier. You could potentially store the macro code in a file and then access it with a single line that gives the location. Be careful in using these that you are using the correct sample size in your mean computations.

       https://stats.idre.ucla.edu/spss/faq/how-can-i-create-multiple-grand-mean-centered-or-group-mean-centered-variables/

       http://myweb.fsu.edu/bcox2/spss.html