

Stata programs for differential item functioning (DIF) and PARSCALE

We have assembled a group of Stata programs with the primary purpose of obtaining ability/trait scores that account for DIF from a group of test items. Adjustment for one covariate (group) can be accomplished with **difwithpar**, which identifies differential item functioning, creates group-specific items to be used to adjust ability (trait) estimates in PARSCALE, writes the code and data file needed to process the updated item list in PARSCALE, runs PARSCALE, and merges the ability estimates and standard errors into the original data set. If you want to adjust for a second covariate, **mergevirtual** merges the group-specific items into your data set for use in the next round.

Programs containing some of the components of **difwithpar** are useful for other purposes as well:

- **difd** evaluates test items for DIF using an ability score you provide, and does not require PARSCALE. **difd** can test for DIF in continuous ‘grouping’ variables.
- **prepar** writes the input code and data file for PARSCALE, and is a real time-saver if you aren’t familiar with PARSCALE. You also have the option to merge the PARSCALE theta estimates and their standard errors back into the original data set, or to estimate ability using fixed parameters.
- **runparscale** also merges the PARSCALE theta estimates and their standard errors back into the original data set, and outputs PARSCALE results to the Stata log.

Specifics and options can be found in the help files for each program.

Here is a brief outline of how to use difwithpar to obtain adjusted scores:

1. Run difwithpar. Assume here that your initial run name is “test0”. thetatest0 and sethetatest0 are now in your data set.
2. Look at the log file for lack of convergence, and other errors.
3. Run difwithpar with ab(thetatest0) ru(test1)
4. Repeat steps 1-3 until the same items come up with DIF. You can look at the log file or examine vars_testN.txt.
5. If you want to test another covariate, first run mergevirtual, which merges the group-specific items into original data set.
6. Repeat steps 1-3 with the next group, with the new list of items (which you can copy from vars_testN.txt).

For more details, see:

Crane PK, Hart DL, Gibbons LE, Cook KF. A 37-item shoulder functional status item pool had negligible differential item functioning. J Clin Epidemiol 2006;59:478-484.

Gibbons LE, Feldman B, Crane HM, Mugavero M, Willig JH, Patrick D, Schumacher J, Saag M, Kitahata MM, Crane PK. Migrating from a legacy fixed format measure to CAT administration: calibrating the PHQ-9 to the PROMIS depression measures. Qual Life Res 2011 Nov;20(9):1349-57.

Contact: Laura Gibbons <gibbonsl@u.washington.edu>