Creativity2.sas: Explanation of code

Goals of code:

- Save a copy of SAS output in a Word file

- Read a space delimited file using a data step

- Randomization test

Saving a copy of results in a rtf file: `ods rtf file='creativity.rtf';`
This line is optional. It tells SAS to save a copy of the output in the RTF format file called creativity.rtf. This file is easily readable by WORD, and makes it easy to copy selected results to another document.

The file=name indicates the filename, which must include .rtf at the end. The name must be in quotes. This file will go in your working directory (see creativity1 explanation) for how to set your working directory. The rtf file will include a copy of all SAS results between this line and a ods rtf close; line (here, at the end of creativity2.sas). When SAS executes the rtf close, it saves the rtf file, opens Word, and displays the file. You will be able to see two copies of the results: one in the SAS results viewer and a copy in the Word file.

Reading the data set: `data creativity;`
The data creativity; statement tells SAS that you want to create a SAS data set called creativity. All the statements from data ; to the run; are part of the data step. The infile command names the file in the working directory that contains the data. This file has a header line with variable names, so the real data starts on the second line. The `firstobs=2` option tells SAS to start reading data from the second line. The input line tells SAS how many values to read on each line, what names to give the variables, and how to read the values. This is described in the Introduction to SAS document. Treatment is a character string, so it is followed by a $; *score is numeric so it is not followed by a $.*

I generally check the line in the log that says "The data set WORK.CREATIVITY has 47 observations and 2 variables". I check that SAS read the number of observations I expected and produced the number of variables I expected.

Permutation test: `proc npar1way scores=data;`
The procedure name has the digit 1 in it. It is shorthand for NonPARametric ONE WAY. We'll talk later about nonparametric and one-way. The `scores=data` goes *before* the semi-colon. It is part of the proc statement and tells SAS to permute the data. In a few weeks we'll use proc npar1way for non-parametric tests using ranks.

Just as with proc univariate (described last week in creativity1.sas), the class statement specifies the grouping variable and the var statement specifies the response. The exact statement specifies that we want to enumerate all possible permutations of treatments to observations. For large data sets,

this can be very time consuming (and not necessary). The `/maxtime=10` option tells SAS to spend no more than 10 seconds on the problem. This is a failsafe in case you ask for a computation that will take months. You will see a note in the log that this is one of those "very long" computations. SAS will stop after 10 seconds and tell you that it's stopping. If you see this, change to using a random sample of permutations (next proc step).

Including the exact statement is important. If you omit it, SAS uses something called a normal approximation, which in this case is no different from the usual t-test.

Randomization test: `proc npar1way scores=data; exact /mc n=10000;`
If the problem is very large, the p-value can be estimated from a sample of permutations of group labels to observations. It isn't necessary to enumerate them all. The `/mc` option requests a Monte-Carlo analysis, i.e. a random sample of possible permutations. This is usually called a randomization test. The `n=10000` requests 10000 samples.

The most difficult part of this analysis is finding the number you really want. That is the two-sided p-value from the Exact test (or Monte Carlo approximation). That number is the Two-sided ..., Estimate number in the box labeled Monte Carlo Estimates for the Exact Test.

Closing the rtf output: `ods rtf close;`
This tells SAS to stop saving results in the RTF file. SAS will then open up WORD so you can look at the results.