Instructions for using the Excel Bootstrap Tool: Testing for significant differences

14 July 2021

Entering data

Data for two recording conditions is entered in Columns B and C (Condition 1) and in Columns F and G (Condition 2), following the same procedure described below for Condition 1. Your data goes into Columns B (Quantities) and C (Weights) starting in Row 5. A Quantity is the data value from a single recording. A Weight is the length of time of each recording. For example, if you want to compute the mean open time (MOT) across 10 recordings, fill B5 with the MOT from Recording 1 and C5 with the length of Recording 1. Fill in B6,…,B14 with the MOTs from the other recordings and C6,…C14 with the respective recording lengths. As you fill Column B, Column A will autopopulate with a running index of the record number. This is just to better keep track of how many recordings you have entered data for.

Note on weights. You can choose whatever weighting you want for the data (it does not have to be record length). If you want all data weighted equally, put 1 in each Weight cell. The weights can be in any units, as long as they are the same for all entries. The weights do not have add up to 1; the calculation takes care of that automatically.

There can be a lag of a few seconds between updating a cell and the recalculation of the bootstrapping. This can be annoying when entering data, so the easiest ways to enter data without this lag is either:

1. cut and paste the data from another application or Excel spreadsheet. You should see the bootstrap data update. If the bootstrap data does not compute automatically, then *either* select Formulas > Calculation Options > Automatic *or* click on Calculate Now in the Formulas ribbon.
2. turn off the automatic recalculation of all formulas (Formulas > Calculation Options > Manual). To update the bootstrap data after entering the data, click on Calculate Now in the Formulas ribbon.

Important: If you want to get rid of old data in Columns B and C, do so by selecting the existing data (click on a corner cell and drag the mouse until the rectangle contains the data) and then pressing Backspace or Delete (or right-clicking and choosing Clear Contents). **DO NOT** use the Excel Delete function located in the Home ribbon to get rid of old data. Selecting and pressing the Delete *key* just clears the numbers from the cells (GOOD), while using the Excel Delete *function* in the Home ribbon deletes underlying metadata that the worksheet needs (BAD). If other cells physically move after deletion or if you are asked which way cells should shift, you are using the Delete *function* that you shouldn't be using. If all the calculated cells become #REF!, then you probably deleted cell B3 in a bad way. Use Control-Z to undo the problem. (Followed by Calculate Now, if you're using that.)

Bootstrap output

The purpose of this tool is to compare the means of the two data sets and test the null hypothesis that they are the same. The associated *p*-value will be low if the null hypothesis is false (i.e., if the means are different).

All the output data you need is in Sheet 1 of the spreadsheet in the form of two histograms and a *p*-value.

The left plot shows the histograms of the bootstrap means for both recording conditions. This is just to give you a sense of how much overlap there is in the possible means of each data set.

The right plot show the histogram of the bootstrapped *t*-statistics. It is shifted so that 0 indicates the two data sets have the same mean (i.e., the null hypothesis is true). This is also just to show visually whether the means are the same (i.e., 0 is in the middle of the histogram) or different (i.e., 0 is on the edge or outside the histogram).

The bootstrapped *p*-value is in J28 with an interpretation of it displayed in K28. In I29 is the number of bootstrap samples that were created.

The bootstrap uses new random resamples every time it is run (e.g., when clicking on Calculate Now in the Formulas ribbon) and therefore will give a slightly different answer every time. If the *p*-value is close to 0.05 (e.g., between 0.04 and 0.06), you should run it multiple times and get a sense of the average *p*-value that is generated and use that average value for publications. This suggestion will display in K28.

We will keep an updated version of the program in the Harvard Dataverse repository at https://doi.org/10.7910/DVN/RTFGBG. Comments/questions: dirk_gillespie@rush.edu