

Bootstrapped and Means-Trimmed One-Way ANOVA and Multiple Comparisons in R

Another way to do a bootstrapped one-way ANOVA is to use Rand Wilcox's R libraries. Wilcox (2012) states that for one-way ANOVAs, both trimming means and bootstrapping cannot fix every problem but do work remarkably well to adjust for problems of heteroscedasticity (non-equal variances) and non-normality. Wilcox recommends a bootstrap-t method when comparing means or trimmed means with a small amount of trimming (20% is the default), and recommends a percentile bootstrap when comparing groups based on other measures of location that are not means (such as M-estimators).

To get started with Wilcox's commands, first install Wilcox's libraries (see Section 8.4.4 of the book for a complete list of what to install) and then open Wilcox's WRS package

```
(library(WRS)).
```

A command for the bootstrap-t method is Wilcox's `t1waybt()` command. Before using Wilcox's function the data will need to be massaged into the correct form. Fortunately, Wilcox has created a handy command that easily divides your data up into separate lists. Unfortunately, the nice labels that distinguish your groups will be lost, but you can retrieve the names of the groups by going back to your original dataset and looking at your group variable names (Wilcox's group #1 will correspond to the first group in your list, and so on).

Here are two different commands I used to subset the data for the syntaxvariety variable by each of the 3 groups in the Ellis & Yuan (2004) dataset, which have the same results. Please choose whichever syntax you find easiest to understand.

names(EllisYuan) <pre>[1] "group" "time" "words" "syllables" [5] "spm" "disfluencies" "syntaxcomplexity" "syntaxvariety" [9] "msttr" "errorfreeclauses" "nucorrectverbs"</pre>	Call for the names of your variables
ellis=fac2list(EllisYuan\$syntaxvariety, EllisYuan\$group)	Choice 1: Here we specify the actual name of the dependent variable that we will split by the second argument, the name of the group
ellis=fac2list(EllisYuan[,8], EllisYuan[,1])	Choice 2: Here we use the number of the columns of the variables we want
ellis <pre>[[1]] [1] 20 15 12 18 9 13 14 14 25 21 18 15 18 15 [[2]] [1] 21 18 22 20 19 19 20 26 24 24 31 25 19 21 [[3]] [1] 23 23 22 18 18 15 16 17 24 15 18 19 21 15</pre>	Type in the name of the list just to make sure your data is correctly subsetting! There are 3 groups with 14 pieces of data in each of them . . . check!

Now let's try out Wilcox's command for a bootstrap-t with 10% trimmed means, since we didn't see any outliers in the data.

t1waybt(ellis, tr=.1, alpha=.05, grp=NA, nboot=2000)	
t1waybt()	Performs a bootstrap-t (parametric bootstrap) comparison of trimmed means
ellis	The first argument is the data, which needs to be in list or matrix form, and which we just did
tr=.1	Default is that 20% of the means should be trimmed
alpha=.05	This is the default and we aren't changing it
grp=NA	We want to test all of the groups contained in the list right now, but if you wanted to test only specific groups, you could specify that here. For example, suppose you had 6 groups, but only wanted to perform a one-way ANOVA with groups 1-5, you would write <code>grp=c(1,2,3,4,5)</code>
nboot=2000	Specifies the number of bootstraps to perform, with 599 as the default; We will change it to 2000 as it does not take too long

The output does not give a confidence interval, only a p -value:

```
[1] "Taking bootstrap samples. Please wait."
[1] "Working on group 1"
[1] "Working on group 2"
[1] "Working on group 3"
$test
[1] 9.025117

$p.value
[1] 0.001
```

This tells us that the omnibus ANOVA test is statistical, and we can report that “A 10% means-trimmed, bootstrapped one-way ANOVA found $F = 9.03$, $p = .001$, meaning the group means were not all equal.” Notice that no degrees of freedom are given, and since they are not, I assume that we do not need to report them in a bootstrapped analysis.

One more thing to note is that if you a large number of groups, the `t1waybt()` function cannot be computed. If you get a warning message that:

```
[1] "Some bootstrap estimates of the test statistic could not be computed"
[1] "Effective number of bootstrap samples was"
[1] 597
```

this means a test statistic cannot be defined. For this particular warning, it means that the actual number of bootstrap samples was 597 instead of the default 599. Wilcox (2012) says you can avoid this problem by using a percentile bootstrap method, but does not recommend comparing means with the percentile bootstrap, but rather you should use other estimators of location such as M-estimators (command is `pbadept()`; see Wilcox, 2012, p. 461 for more information).

Another type of analysis you might want to use is a non-bootstrapped but means-trimmed one-way ANOVA with the command `t1wayv2()`. The command syntax is equal to `t1waybt()` except you cannot specify alpha or the number of bootstrap samples.

```
> tlwayv2(ellis, tr=.2, grp=NA)
```

```
$TEST
```

```
[1] 7.44622
```

```
$nu1
```

```
[1] 2
```

```
$nu2
```

```
[1] 17.85265
```

... data about the N of each group deleted ...

```
$p.value
```

```
[1] 0.004450589
```

```
$Var.Explained
```

```
[1] 0.3391496
```

```
$Effect.Size
```

```
[1] 0.5823655
```

The `$nu1` value is the degrees of freedom between groups (the hypothesis df) and the `$nu2` value is the degree of freedom between groups (the error df). By the way, this error df is not a whole integer because it uses Welch's method as an adjustment for comparing groups that do not have equal variances. In addition, the command `tlwayv2()` will return an effect size of the r-family with the label `$Var.Explained`, or of the d-family with the level `$Effect.Size`. Therefore, using this data we would report that the omnibus ANOVA for the `syntaxvariety` variable was statistical, $F_{2,17.9} = 7.45$, $p = .004$, $PV = 34\%$ (or $d = .58$).

Remember, these tests just give the omnibus F test. In some cases you may not care to even check the omnibus test, but instead just proceed directly to multiple comparisons. In order to perform multiple comparisons to determine which groups are different from each other, you can use the `linconb()` command. It compares all of the pairs of groups using the bootstrap-t method and can compare trimmed means. Wilcox (2012) says that the bootstrap-t is a good method when comparing all pairs of means, or when using a small amount of means trimming. With your data

in the correct format for the one-way ANOVA you are ready to use this command. Below, I give an analysis of it.

<code>linconb(ellis, con=0, tr=.2, alpha=.05)</code>	
<code>t1waybt()</code>	Performs a bootstrap-t (parametric bootstrap) comparison of trimmed means
<code>ellis</code>	The first argument is the data, which needs to be in list or matrix form
<code>con=0</code>	This argument can be used to specify that only some of the contrasts should be tested, and should be a matrix of numbers. <code>con=0</code> means that all contrasts will be tested. See Section 10.4.9 for information about how to use this argument to test only certain contrasts
<code>tr=.2</code>	Default is that 20% of the means should be trimmed
<code>alpha=.05</code>	This is the default and we aren't changing it

Here is the output from the command.

```
[1] "Note: confidence intervals are adjusted to control FWE"
[1] "But p-values are not adjusted to control FWE"
[1] "Taking bootstrap samples. Please wait."
$
      [,1] [,2] [,3]
[1,]   14   14   14
```

Note the warnings here. I am interested in the confidence intervals, and so understand that they have been adjusted for multiple comparisons. If you are interested in the p -values, take care of do some kind of adjustment of them if you have more than 3. I recommend the FDR method. Here's how you could adjust your p -values with the FDR method. Let's say your p -values were [.03, .0002, .43, .09].

```
p <- c(.03, .0002, .43, .09)
```

```
p.adjust(p, "fdr")
```

```
[1] 0.0600 0.0008 0.4300 0.1200
```

This shows that only the second p -value would be considered statistical because after adjustment it is the only one under $\alpha = .05$.

Returning to the bootstrapped comparisons for the Ellis & Yuan data, the groups are not labeled but I could type `EllisYuan$group` and the order I see the groups listed would be the same order as these numbers, which would be NP, PTP, and OLP.

```
$psihat
      con.num psihat  ci.lower  ci.upper
[1,]        1   -5.5 -9.246417 -1.753583
[2,]        2   -2.7 -6.897237  1.497237
[3,]        3    2.8 -1.218607  6.818607
```

The next part of the output gives the number of the contrast that is tested, which is given at the end of the output under the heading `$con`. Here is the `$con` for this output:

```
$con
      [,1] [,2] [,3]
[1,]     1     1     0
[2,]    -1     0     1
[3,]     0    -1    -1
```

`$con` specifies the contrasts, and this is read in columns, so that the first contrast is comparing groups 1 and 2, the second groups 1 and 3, and the third groups 2 and 3 (the places where the 1s are). Under `$psihat`, we see that the difference between groups 1 (NP) and 2 (PTP) is 5.5 points, with a bootstrap-t, 20% means-trimmed CI of [-9.25, -1.75]. This means that the difference in mean scores between the NP and PTP group could be as low as around 2 points or as high as 9 points (I discussed a similar result for the bootstrapped means in Section 9.4.5 so I won't do it again here). The difference between groups 1 (NP) and 3 (OLP) is 2.7 points, with a CI of [-6.90, 1.50]. This result goes through 0 with more overlap on the other side of 0 than some of our other analyses. The last comparison is between group 2 (PTP) and 3 (OLP) and the CI is [-1.22, 6.82].

This looks very similar to the comparison between NP and OLP, so we will conclude there is no real effect for the comparison between these groups.

The next part of the output gives the test value, standard error, and p -value, and then the critical value for the test (the value of the test must be larger than this to have a p -value smaller than .05).

Since I don't plan to use these I won't comment on them.

```
$test
      con.num      test      se      p.value
[1,]        1 -3.922699 1.402096 0.001669449
[2,]        2 -1.718853 1.570815 0.093489149
[3,]        3  1.861748 1.503963 0.086811352

$crit
[1] 2.672012
```

The reason to use this type of test then, is that you would want to both bootstrap and trim the means of your multiple comparisons. Two other commands you might be interested in for multiple comparisons are `lincon()`, which uses trimmed means only, or, `mcppb20()`, which uses trimmed means plus the percentile bootstrap. The basic syntax of the `lincon()` command is exactly the same as for `linconb()` but the output differs a bit in that you will get degrees of freedom listed and a more straightforward reporting in the lines of output which groups are tested against which other groups (for example, Group 1 vs. Group 2). For the `mcppb20()` command the alpha level is fixed at .05, trimming is fixed at 20%, and the number of boots used depends on the number of groups being compared and cannot be controlled. If you think you would like the percentile bootstrap, this command can be used by just putting in the name of the dataset as the only argument, like this: `mcppb20(ellis)`. The output it returns is quite similar to the `linconb()` command, and the p -value cutoff is determined by using the Bonferroni correction.

Bootstrapped and Means-trimmed One-way ANOVAs and Multiple Comparisons

- 1 Use Wilcox's commands (WRS library). First, however, you must change the structure of your data to fit Wilcox's commands, which use the wide form of the data (N.B. items in red should be replaced with your own data name).

`ellis=fac2list(EllisYuan$syntaxvariety, EllisYuan$group)` #the first argument is the #dependent variable, the second is the independent or group variable

- 2 You have a couple of options for a robust omnibus one-way ANOVA test:

`t1waybt(ellis, tr=.2, nboot=2000)` #bootstrap-t, means-trimmed

`t1wayv2(ellis, tr=.2)` #mean-trimmed

- 3 Perform robust post-hoc multiple comparisons, again with several options:

`linconb(ellis, con=0, tr=.2, alpha=.05)` #bootstrap-t, means trimming

`linconb(ellis, con=0, tr=.2, alpha=.05)` # means trimming

`mcppb20(ellis)` #percentile bootstrap, 20% means trimming

Bibliography

- Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, 26, 59–84.
- Wilcox, R. R. (2012). *Modern statistics for the social and behavioral sciences: A practical introduction*. New York: Chapman & Hall/CRC Press.