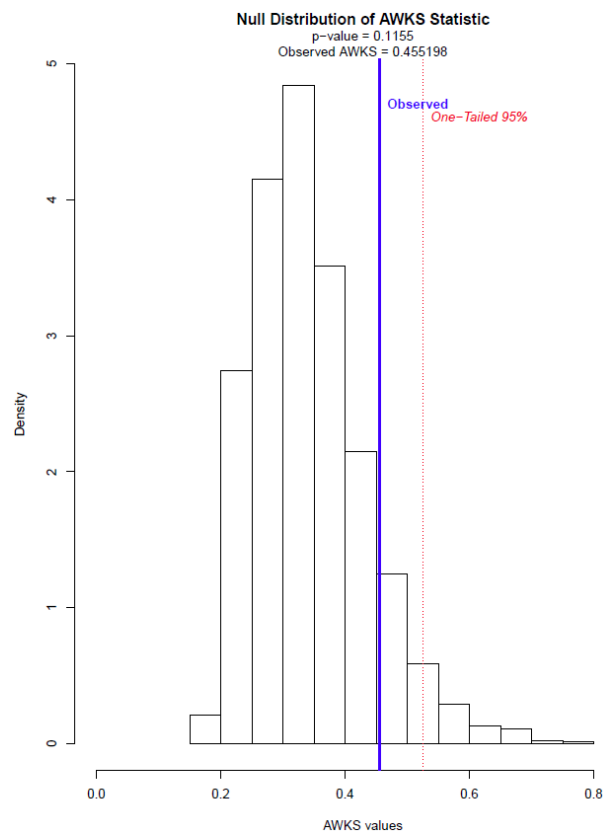


## How to Calculate the AWKS Statistic (Two Cohort Analysis)

In this example, we measure the difference between the taxonomic profiles of two cohorts using the Abundance-Weighted Kolmogorov-Smirnov (AWKS) statistic. This test statistic provides the user with the magnitude of the difference in ubiquities between the two cohorts. The larger this value, the greater the difference between the two cohorts.

Along with the calculation of the AWKS statistic, a p-value is estimated to provide the statistical significance of the AWKS statistic. If there is no difference between the two cohorts analyzed, i.e. the AWKS value is 0. The null distribution is generated by computing the AWKS statistic by resampling with replacement between the two cohorts assuming they are not different. In other words, any sample can be found in either cohort. The p-value is the proportion of the null distribution that is greater than the AWKS statistic calculated between the observed profiles of the 1<sup>st</sup> cohort and the 2<sup>nd</sup> cohort.

In the following example, the two cohorts compared were not statistically significant at an alpha of .05.



### Direct command line invocation

cd into the “*lib*” directory and then execute

```
lib/Compare_wAWKS.r -a ../example_data/cohortA.phylum.summary_table.tsv -b  
../example_data/cohortB.phylum.summary_table.tsv -o example_two_cohort_output/two_cohort
```

### OUTPUT:

*two\_cohort.awks\_test.tsv*

*two\_cohort.awks\_test.pdf*