User's guide for *IND_TEST* Ver20220309

Tae-Kun Seo,
Division of Life Sciences,
Korea Polar Research Institute,
26 Songdomirae-ro, Yeonsu-gu,
Incheon 406-840,
Republic of Korea,
E-mail) seo.taekun@gmail.com

March 09, 2022

This is a user's guide for the *IND_TEST* (pronounced as 'independence test') program. A detail information can be found in the Seo et al.'s manuscript.

FORMATS AND OPTIONS

Sequence file : The sequence data format is the phylip (sequential) format as shown in Figure 1.

Option file : Three options should be represented as a separate line within the "ind_test_option.txt" file (Figure 2). In each line that specifies an option, only the first item after the '>' character is recognized as an input. Do not change the string between the '<' and '>' characters.

(1) < seq file >

The sequence file is set here. Gap character should be one of '-' or '?' and should not be 'N'

(2) < boot iter >

Number of bootstrap iterations. This corresponds to B of Eq (2) of the Seo et al.'s manuscript.

(3) < rand seed >A random seed value is required for the bootstrap procedure. Any integer between -2^{63} and $2^{63} - 1$ can be selected for random seed number.

HOW TO RUN

```
 12 1000
T1    TTCCGAGGAGCCCTATCACATGGTCGTTTCTAAGATGCCTCCGA...
T2    TTCCGAGGGACTCTCTCACATGATTGTTTCTAAAATGCCTCCGA...
T3    CTCTGAGGGGCCCTGACACATGATTGTTTGTAAAATGCACTCAA...
T4    CTCCGGGAGGCCCTTACATATGACCGTTTGTAAAATTCCTTCAA...
T5    TCCCGAGGGGTACCAGGACATGATAATTTCTAAAATGCCCCTGA...
T6    TCCCGGAGGGTATTGTGACATGATTATTTCTAGAGCGCCTCTGA...
T7    TCCTGGGGGGCCCTACTACATGATAACTCCTAAATTGCCCCTGA...
T8    TCCTGGGGGACCCTATTACATGATAACTCCTAAACTGCCTCTGA...
T9    TTCAGGGGGGCCCCAAGGTATAATTGCTTCCAAAACGCCCTTAG...
T10   TCCAGAAAGGCCCTAAGGCATAATCGCTTCCAAAACGCCCCTAG...
T11   TCCAGGGGAGCCCTGGTTTATAATTGCTTCTGAGGTGCCCCAAG...
T12   CCCAAGGGGGCTCAAGTATATAGTCGCGTCTAAGGTGCCCCCAG...
```

Figure 1: testdata.txt

```
< seq file >  testdata.txt
< boot iter >  1000
< rand seed >  1
```

Figure 2: ind_test_option.txt

You must install JAVA runtime environment (JRE; https://java.com/en/download/) to run *IND_TEST*. *IND_TEST* can be run on any OS as long as JRE is pre-installed.

Once JRE is installed, copy the IND_TEST folder to your favourite location. Put the sequence and option files in the bin folder. If your OS is windows, open command prompt by running 'cmd.exe'. Go to the bin folder and type the commands of Figure 3 ('TestMain' is case-sensitive) and hit ENTER. Because the results are displayed in the console, you need to capture them by redirecting messages from screen to the appropriate file. In the example of Figure 3, results are redirected from the console to the scr.txt file.

INTERPRETATION OF THE RESULTS

In scr.txt, after list of input options, proportions of uninformative columns (n_3/n), fewer-gap columns

```
C:\ind_test\bin>  java ind_test.TestMain   >  scr.txt   ↵
```

Figure 3: How to run

```
 Ind_test Version:  20220131

### Input Options ###
< seq file >  testdata.txt
< boot iter >  1000
< rand seed >  1

uninformativeGappyColRatio=0.000000
lessGappyColRatio=0.500000
moreGappyColRatio=0.500000
SeqLength=1000
T-Stat=-0.006400
F(T)=0.350000
```

Figure 4: scr.txt

$(n_1/(n_1+n_2))$ and more-gap columns $(n_1/(n_1+n_2))$ are represented. After these ratios, sequence length, T statistic and $F(T)$ values are represented.