A new development cycle of the Statistical Toolkit

Statistical tests in the ST

The ST now also comes with an extensive set of unit tests, which are meant to test the consistency of the statistical test for each new version of the ST.

New development cycle

The new development cycle of the ST has been moved to a SVN repository. To be as self-consistent as possible, the number of dependencies on external software systems has been minimized. The only proper requirement is the GNU Scientific Library [3]. In order to SVN repository. To be as self-consistent as possible, the number of dependencies on external software systems has been minimized. The only proper requirement is the GNU Scientific Library [3].

The new development cycle of the ST has been moved to a SVN repository. To be as self-consistent as possible, the number of dependencies on external software systems has been minimized. The only proper requirement is the GNU Scientific Library [3].

The new development cycle of the ST has been moved to a SVN repository. To be as self-consistent as possible, the number of dependencies on external software systems has been minimized. The only proper requirement is the GNU Scientific Library [3].

The new development cycle of the ST has been moved to a SVN repository. To be as self-consistent as possible, the number of dependencies on external software systems has been minimized. The only proper requirement is the GNU Scientific Library [3].

The new development cycle of the ST has been moved to a SVN repository. To be as self-consistent as possible, the number of dependencies on external software systems has been minimized. The only proper requirement is the GNU Scientific Library [3].

Example

Is the number of earthquakes per year a random number?

We have two data samples. If the differences between the two were due only to statistical fluctuations, one would expect the differences to be a rather large number of short sequences (runs) of consecutive positive and negative values. Using Wald-Wolfowitz runs test [8] one can infer some systematic effects between the two samples. One can use the runs test to test the randomness of the numerical data around some given value (or function).

Construct OneDimData object, and make it binary around mean value.

```
OnedimData nundata = new OneDimData(); //construct one
dimens:dimensional data
nundata.setName("earthquakes.csv");
```

Run the Wald-Wolfowitz runs test and get the results

```
WaldWolfowitzQualityChecker qc =
new WaldWolfowitzQualityChecker();
vector<double> v =
new vector<double>(20);
```

```
cq.R = 31
```

```
cq.run = 7
```

```
cq.n = 32
```

```
cq.m = 32
```

```
num쿰 것은
```

```
cq.w = 32
```

```
cq.ww = 32
```

```
cq.x = 32
```

```
cq.y = 32
```

```
cq.w = 32
```

```
cq.ww = 32
```

```
cq.x = 32
```

```
cq.y = 32
```

Calculate the p-value from given test results

```
WaldWolfowitzQualityChecker qc =
new WaldWolfowitzQualityChecker();
```

```
ComparisonResult result =
ComparisonResult.get("quality.csv",
mockup("<result:quality|<result>"));
```

```
result = 0.000196720
```

The calculated p-value is 0.000196720. So: can we answer the question? No. But we can say that the exact p-value is small enough to have strong statistical evidence to reject the null hypothesis that in the 20th century the number of earthquakes with magnitude above 7.0 is not randomly distributed around their mean value.

Bibliography


See also the associated contribution to the conference proceedings.