τ-ARGUS release notes version 4.1.6

Peter-Paul de Wolf
10 July 2017

1   Introduction

τ-ARGUS is a software package whose development started at Statistics Netherlands in the early 90s of the 20th century. Soon after, other institutes and universities joined forces to further develop the software. Moreover, additional funding from Eurostat was received in several project. Since December 2014 it is officially released as Open Source software.

τ-ARGUS is a software package to apply Statistical Disclosure Control methods to tabular data. The input for the program can be either micro data or tabulated data (with additional information on each table cell).

To contribute to further development and/or support, you can ask to be added to the GitHub organisation sdcTools by sending an email to pp.dewolf@cbs.nl and/or alexander.kowarik@statistik.gv.at.

Moreover, you can become member of the UserGroup sdcTools on https://joinup.ec.europa.eu/software/sdctools/home (click on "Actions" and select "Request to be a member of this software").

2   Installation

2.1   System requirements

The Windows binaries (32 bit) are tested on a Windows 7 machine (32 bit) and Windows 10 machine (64 bit, with 32 bit Java installed).

You will need to have Java 1.7 (or later) installed. Note that both jdk1.7.0 and jre7 are versions of Java 1.7.

Please note that a 32 bit installation of Java is needed for the precompiled Windows binaries. It is possible to have a 32 bit and a 64 bit version of Java installed on the same machine.

NB: Make sure when running τ-ARGUS that the 32 bit JVM is used! The distributed binary will not run in a 64 bit JVM!

We have provided a binary of τ-ARGUS with a bundled Java 7 as well. If you install that version, it will automatically use the bundled Java 7 that should work directly, also under 64 bit Windows.

2.2   Installation instructions

Unzip the file into a directory where you have read, write and execute permission. In case you own an Xpress license that makes use of an xpr-file, you should copy that file into the directory where τ-ARGUS is installed.

Then you can run the TauArgus.exe.

Experienced users know that τ-ARGUS stores various parameters in the registry. The new τ-ARGUS does the same, but the parameters are not copied from an installation of τ-ARGUS 3.5.0.

3 Licencing

The τ-ARGUS software is released under the EUPL licencing.

When using one of the included commercial solvers (Xpress and/or Cplex) an additional licence is needed. This licence can be purchased from FICO (Xpress) or IBM (Cplex).

In case of Xpress, a discount on the licence is possible when you indicate that you will use their software only in combination with τ-ARGUS. Contact support@fico.com and explain that you need a licence for use with τ-ARGUS only.

Since we had to change the licencing mechanism of Xpress when going to Open Source, existing users[1] should contact support@fico.com as well, to get their licences renewed in compliance with the new situation.

When using Soplex (the included Open Solver) a licence is included that grants to the European National Statistical Institutes a non-exclusive, non-transferable, non-sub-licensable, perpetual right to use Soplex version 2.0.x and SCIP version 3.1.x, linked to the τ-ARGUS software.

4 Structure of the software

The current version of τ-ARGUS is based on a Java user interface and C/C++ dlls for the computationally hard parts. Moreover, some functionality is provided by calling stand-alone executables like GHMiter (Fortran), Audit (Pascal), Network-Solution (C++) and CTA (C++).

For solving the optimality problems in Cell Suppression, CTA, Rounding and the Audit we used high quality commercial solvers like XPress and CPlex. Some (free) Open Source solvers are available now as well. We expect that in many cases for small and not too complex tables the free solvers will give satisfactorily results, but for larger complex tables the commercial solvers will still be needed (the implementation with the (free) Open Solver is too unstable).

5 The new GUI

As this version has been built with Java, the look and feel will be a bit different compared to version 3.5, but the structure is still the same. Also the progress info windows will have a different look. Due to the structure requirements of Java we had to design them differently.

5.1 The new main window

The new central window in τ-ARGUS is similar as the window under "View Table" in τ-ARGUS 3.5.

Most of the buttons are the same as in τ-ARGUS 3.5. However the "Save Table" and the "Select Table" buttons are no longer there, as these options are now directly available from the main menu.

5.2 Files menu

Under Files you will see the menu-items for "Open microdata", "Open Table", "Open Table set", "Open Batch process" and "Exit". After "Open Table set" (for linked tables) you will see the "Specify Table Meta" only once, but it will be applied to all tables.

---

[1] with restricted OEM-licences

5.3 Specify Menu

Under "Specify" we have two menu-items "Metadata" and "Tables".

"Metadata" will open the window to specify and adapt the metadata.

The "Tables" window is different depending on the type of input (micro data versus tabular data).

5.4 Modify Menu

Only the menu items "Select Table" and "Linked Tables" are available as "View table" is no longer needed. "Select tables" is obvious. The linked tables option now always automatically builds the 'cover table' so the window has been simplified. Only the buttons to start the protection (via Modular or Hypercube) are available.

5.5 Output menu

The menu items "Save Table", "View report", "Generate Apriory", "Write Batchfile" are available.

5.6 Help menu

The menu items "Content", "News", "Options" and "About" are available.

"Content" will open the manual. In certain windows it is possible to get content sensitive help by pressing F1.

"News" opens a file with information about the recent developments in τ-ARGUS.

"Options" opens a window to specify the solver to be used, the name of the logfile to be used and the maximum time allowed to find an optimal suppression pattern in a sub-table in the Modular approach.

"About" opens a simple about box.

6 Changes with respect to τ-ARGUS version 3.5.0

While rewriting the source code, some changes were made:

- The <APRIORI> syntax in a batch-file (*.arb) is now:

  ```
  <APRIORI> ("Filename", TableNumber, "Separator", Ignore, Expand)
  ```

  That is, the Separator should now be enclosed in double-quotes and two parameters are added:
  "`Ignore`" to control "Ignore incorrect lines" (1 = ignore, 0 = do not ignore)
  "`Expand`" to control "Expand for trivial levels" (1 = ignore, 0 = do not ignore)

  Example:
  ```
  <APRIORI> ("D:\Argus\Test\Exmaple.hst",1,";",1,0)
  ```

- The <WRITETABLE> keyword in a batch-file (*.arb) has different options:

  ```
  <WRITETABLE> (TableNumber, OutputType, ParameterString, "Filename")
  ```

  with

```
OutputType
```
1. CSV
2. CSV for pivot table
3. Code-Value
4. SBS
5. Intermediate
6. JJ-format
```
ParameterString
```
Corresponds to the check-boxes in the "Save Table" window, with "+" for on and "-" for off

AR: Add Audit Results (default -)

AS: Add Status (default -)

FL: Variable names on First Line (default + for OutputType 2, - for other OutputTypes)

HI: Use holding info (default -)

HL: Add hierarchical levels to the SBS file (default -)

QU: Embed spanning variables in Quotes (default +)

SE: Suppress empty cells (default -)

SO: Status Only in intermediate format (default -)

TR: Remove trivial levels in JJ format (default -)

Example:
```
<WRITETABLE> (1,2,AS+QU-,"D:\Argus\Test\Example.csv")
```

7   Feedback

Any feed-back on this Open Source version is highly appreciated. At
https://github.com/sdcTools/UserSupport issues can be filed with your remarks, comments or whishes.

 We are especially interested in

- Bug reports

- Workarounds for bugs/strange behavior

- Wishes for future development

Your feedback will be used to improve the Open Source version of τ-ARGUS. It may lead either to bug-fixes in the next release or your remarks will be added to a "wish-list", containing suggestions for improvements.

Whenever reporting bugs, please provide a working example, so we can more easily try to replay your issue.

Please note that those issues can be viewed by anybody on the internet, so do not include any sensitive data!

8   Version history

Changes with respect to version 4.1.5:

- Explicit empty cells in tabular input are accepted again.
- Tabular data input can only be in free format. Changing the format in the metadata window is prevented and a warning will be shown.
- Large cell values in 1-dimensional tables caused τ-argus to crash. This is fixed.
- In the batch file you can now specify the LP-solver to be used, including the licence file in case of Cplex.
- In case Modular is used on a table with more than 4 dimensions, a warning is given.
- The <CLEAR> command in the batch-file is improved.
- A fourth parameter is added to the command line call of the batch-version of τ-argus. This parameter specifies the batch data directory.
- The <READTABLE> command in the batch-file now accepts as parameter:
  - 0: (default) an additive table is required
  - 1: compute sub-totals in case non-additive
  - 2: allow non-additivity
- A <VERSIONINFO> command is added to the batch-file commands. This command has one parameter that specifies the file where to write the version info.
- All dialog windows where you can search for files now remember the last place they searched.
- The maximum number of hierarchical levels is now set to 10 for all methods.
- The <FREQ> command in batch-file is now working as expected.
- The directory of the CDL-file is now added to the .rda-file when saving in intermediate format.
- Some minor GUI adjustments (e.g., new icons).

Changes with respect to version 4.1.4:

- Modular and Optimal now use the same method to deal with the scaling of the cost-variable.
- Some memory-issues in TauArgusJava.dll fixed.
- Some issues related to the number of significant digits are now fixed.
- Partitioning in Rounding should work again.
- Some issues concerning saving tables (e.g., options) are fixed.
- Added many checks when reading tabular input.
- Some small bugs in interface fixed.

Changes with respect to version 4.1.3:

- In "Generate Apriori" it is now possible to only adjust the cost of a cell, without changing the status (combining "omit" with "cost").
- Bug fix: minimum frequency rule was used even when not selected. This is corrected.
- Length of code names in Hypercube is no longer limited.
- Zero unsafe is now correctly used in Calculate Table.

- Some small bugs in csv output fixed.
- Overlapping variables in fixed format allowed.
- <LOGBOOK> command in batch-job fixed ()quotes can be used again).
- Workaround for use of dots in code in hierarchies (note: still not allowed to *start* with a dot).
- Some small bugs in interface fixed.

Changes with respect to version 4.1.2:

- Linked tables now works up to 7 tables.
- Limiting time in subtables (modular approach) now also works in batch-mode.
- Bug fixes writing csv format (o.a. rounded figures).
- Some minor bug fixes in interface.
- When called from a Windows .bat-file the TauArgus.exe does not close the calling process.

Changes with respect to version 4.1.1:

- Possible to round cell values larger than maxint.
- Rounding shows a progress window.
- Rounding shows a stop/continue window after pre-set time limit.
- Holdings correctly detected in batch-mode.
- Settings of the safety rules are correctly saved after a batch-run.
- Some minor bug fixes in interface.

Changes with respect to version 4.1.0:

- Availability of SPSS is detected at runtime.

- Bug fix in Distance Cost functionality.

- Bug fix in Output format (spurious comma's removed).

- Bug fix in Apriori functionality (now possible to produce apriori info for use in tables with higher dimension than original table).

- Bug fix in Linked tables (bogus levels in cover table).

- Some minor bug fixes in interface.

Changes with respect to version 4.0.3 (beta):

- Added Linux support (all approaches should work under Linux).

- Included possibility to use SPSS datafile as microdata input.

- Included help functionality (use of F1-key).

- Some bug fixes.

Changes with respect to version 4.0.2 (beta):

- Improved CTA support.

- Distance and Frequency available as costs in batch.

- Improved error handling and error messages.

- Some bug fixes.

Changes with respect to version 4.0.1 (beta):

- Cell suppression with network-approach is now possible.

- Rounding is now possible with Xpress, Cplex [2]and Open Solver.

- One dimensional tables are now possible.

- It is now possible to use <freq> as response variable in a batch-file (*.arb).

- The command <APRIORY> in a batch-file (*.arb) now has two additional parameters.

- When no missing-codes are specified in the meta-data, there will be no missing-codes in the displayed table.

- Files mentioned in batch/meta-files are assumed to reside in the same directory of the batch/meta-file. If not, the complete directory-path should be specified.

- Colour of "manual unsafe" is changed.

- Some minor bug fixes.

---

[2] Note that the current Cplex license restricted to τ-ARGUS use is not fit for the rounding procedure.