Requirements for the Library statistics_lib

Conventions

Requirements listed in this document are constructed according to the following structure:

Requirement ID: REQ-UVW-XYZ

Title: Title / name of the requirement

Description: Description / definition of the requirement

Verification Method: I / A / T / D

The requirement ID starts with the fixed prefix 'REQ'. The prefix is followed by 3 letters abbreviation (in here 'UVW'), which defines the requiement type - e.g. 'FUN' for a functional and capability requirement, 'AWM' for an alarm, warnings and operator messages, etc. The last part of the ID is a 3-digits *hexadecimal* number (0..9|A..F), with the first digit identifing the module, the second digit identifing a class / function, and the last digit - the requirement ordering number for this object. E.g. 'REQ-FUN-112'. Each requirement type has its own counter, thus 'REQ-FUN-112' and 'REQ-AWN-112' requirements are different entities, but they refer to the same object (class or function) within the same module.

The verification method for a requirement is given by a single letter according to the table below:

Term	Definition
Inspection (I)	Control or visual verification
Analysis (A)	Verification based upon analytical evidences
Test (T)	Verification of quantitative characteristics with quantitative measurement
Demonstration (D)	Verification of operational characteristics without quantitative measurement

Functional and capability requirements

Requirement ID: REQ-FUN-000

Title: Purpose of the library

Description: The library should provide a consistent, simple but useful API for:

- Calculation of basic statistical properties of distribution of measured numerical data sequence(s)
- Analysis of the shape of the distribution of the empricial numerical data
- Performance of simple statistical tests concerning the shape of the distirbution of the empricial numerical data

The numerical data may come in a form of a generic sequence of either real numbers (integer or floating point numbers) or 'measurements with uncertainty' (coupled pair of tow real numbers representing the 'mean value' of a measurement and the associated measurement uncertainty)

Verification Method: A

Requirement ID: REQ-FUN-001

Title: Minimum level of the required functionality

Description: The library should provide the following minimal level of functionality:

- 1D statistical functions
  - Calculation of the arithmetic mean of a sample
  - Calculation of the variance and standard deviation, skewness and excess kurtosis treating the sample as the entire population (without Bessel correction) as well as slices of population (with Bessel correction)
  - Calculation of the generic central or non-central, normalized or not normalized moment of distribution treating the sample as the entire population
  - Calculation of the standard error of the mean of the sample disregarding the measurement uncertainties
  - Calculation of the total uncertainty of the mean of the sample including the measurement uncertainties
  - Calculation of the min, max, median, 1st and 3rd quartile of the sample distribution
  - Calculation of an arbitrary quantile of the sample distribution
  - Calculation of the mode(s) of the sample distribution
- 2D statistical functions
  - Calculation of the generic central or non-central, normalized or not normalized cross-moment of distributions treating the samples as the entire respective population
  - Calculation of the covariance of two samples (w/o Bessel correction)
  - Calculation of the Pearson`s r correlation coefficient
  - Calculation of the Spearman`s ρ rank correlation coefficient
  - Calculation of the Kendall τ rand correlation coefficient
- Classes providing probability distribution function , cummulative distribution function and quantile function (if possible) for the following common distributions:
  - Generic Gauss (normal) distribution $\mathbb{N}(\mu, \sigma)$
  - Z-distribution distribution $\mathbb{Z} = \mathbb{N}(0, 1)$
  - Student`s t-distribution
  - χ^2-distribution
  - F-distribution
  - Poisson distribution
  - Binomial distribution
- Basic test statistics functions:
  - Z-test for means comparison
  - t-test for means comparison
  - χ^2-test for variance
  - F-test for comparison of variances
  - Levene's test
  - Brown-Forsythe test
  - Barlett's test

Verification Method: A

Software system inputs and outputs

Requirement ID: REQ-SIO-000

Title: API input

Description: The API functions / methods should accept any generic sequences of numeric data. Generic sequence is either the Python built-in data types like **list**, **tuple** or any custom class instance, sub-classing **collections.abc.Sequence**, except for strings (**str**), byte-strings (**bytes**) or byte-arrays (**bytearray**). Numeric data is any real number (**int** or **float**) or an instance of a class compatible with **phyqus_lib.base_classes.MeasuredValue**, i.e. having *data attributes* (fields or properties) *Value* and *SE* - representing the 'mean' ('most probable') measured value and the associated measurement uncertainty / error.

Verification Method: A

Requirement ID: REQ-SIO-001

Title: API output

Description: The API functions / methods should return only native Python data types - **int**, **float**, **bool** or **list** / **tuple** of elements of the three mentioned scalar types.

Verification Method: A

Interfaces

Requirement ID: REQ-INT-000

Title: Reliable dependencies

Description: The library should be based either solely on the Standard Python Library, or it should use only widely accepted / used and well maintained libraries / packages.

Verification Method: I

Alarms, warnings and operator messages

Requirement ID: REQ-AWM-000

Title: Inappropriate data input

Description: An exception should be raised with the informative description if the input data is inappropriate, e.g. wrong data type or containing not acceptable value(s).

Verification Method: D

Usability requirements

Requirement ID: REQ-USE-000

Title: Intended users qualification

Description: The library is designed to be used by the persons with sufficient profiency with Python programming language and only basic skills with the statistical analysis.

Verification Method: D

Installation and acceptance requirements

Requirement ID: REQ-IAR-000

Title: Python interpreter version

Description: The library should be used with Python 3 interpreter. The minimum version requirement is Python v3.6.

Verification Method: D

Requirement ID: REQ-IAR-001

Title: Operational system

Description: The library should work, at least, under MS Windows and GNU Linux operational systems. Ideally, it should not utilize any platform-specific functionality, therefore it should work under any OS, for which Python 3 interpreter is available.

Verification Method: D

Requirement ID: REQ-IAR-002

Title: System requirements check

Description: The library should provide a module / script to check if all system requirements are met, i.e. the Python interpreter version, other required libraries / packages presence as well as their versions. This module / script should report the missing requirements.

Verification Method: D

User documentation requirements

Requirement ID: REQ-UDR-000

Title: The library is thoroughly documented.

Description: The library should be sufficiently documented, including:

- Design document
- Requirements documents
- Test reports
- User and API references

The reference documentation should provide sufficient data on the implementation for the future maintenance and modification as well as clear and comprehensive usage instructions and examples.

Verification Method: I