Visual Averaging Library (V.AveLib) User Manual

Michael Birch

March, 2015

Contents

Introduction

The Visual Averaging Library (V.AveLib) program was developed in the Java programming language as a data analysis tool to be used in particular for determining the recommended value of a quantity from a series of measurements. The program includes eight different averaging methods, some of which have published descriptions in journal articles. These references are given in the "About" box which can be accessed from the **Top Menu**. It is expected that the user already has some familiarity with the averaging methods and the Chi-Square test which is used to check the validity of some of the results. Three prescriptions for the identification of outliers are also available to "clean up" discrepant data sets at the discretion of the user.

To give comments or report bugs please contact Michael Birch (birchmd@mcmaster.ca), the developer of this software.

Chapter 1

Installation and Running the Program

Since this code was developed in Java, a Java Runtime Environment (JRE) is required. In particular, this software was developed for use with the Oracle Java and may not be compatible with other Java Virtual Machines (e.g. OpenJDK). The latest Oracle JRE can be downloaded from the Oracle website. Ubuntu users can instead follow the tutorial found at http://www.webupd8.org/2012/09/install-oracle-java-8-in-ubuntu-via-ppa.html.

Once the appropriate JRE is installed, the executable jar file "VisualAveragingLibrary.jar" located in the "dist" folder of the main VisualAveragingLibrary directory will start the program. On some systems, simply double-click it to start. On other systems, it may be necessary to launch the program from the command line using the command `java -jar VisualAveragingLibrary.jar`.

The plotting functionality of this software requires "GNU plot" which is freely available for all platforms from http://www.gnuplot.info/. See subsection 2.1.2 for more details.

Chapter 2

The Interface

The interface of the V.AveLib has five main parts which be discussed in detail: the **Top Menu**, the **The Quick Instructions**, the **Data Set Tabs**, the **Data Entry Box** and the **Side Buttons**. There is also a window separate from the main interface which will be described: the **Check for Outliers dialog**.

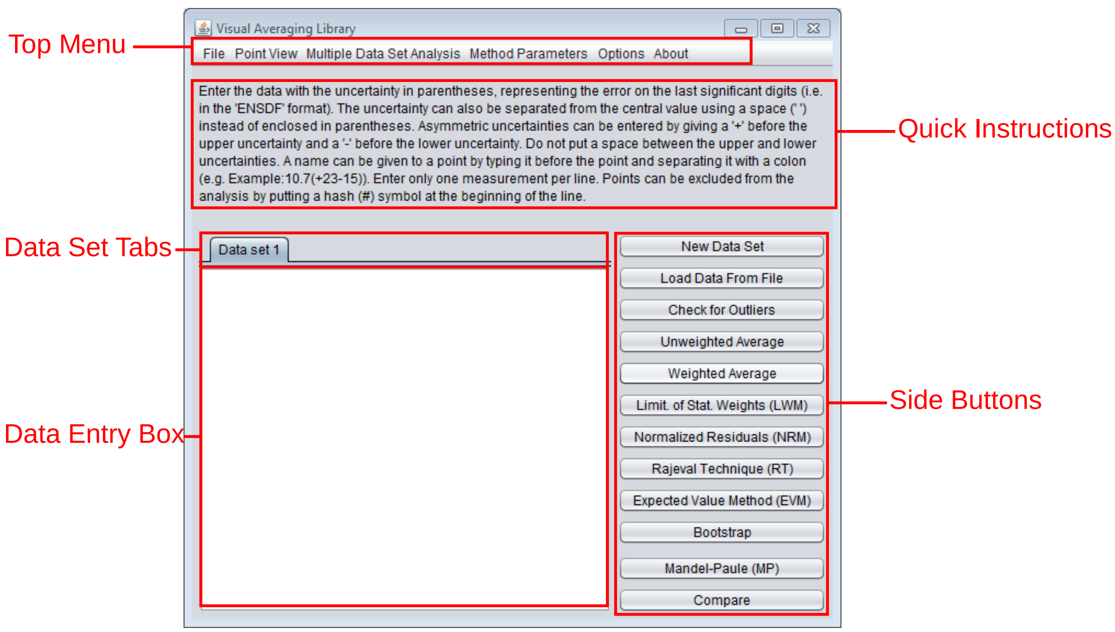

Figure 2.1: The main interface of V.AveLib

2.1 Top Menu

As implied by the name, the **Top Menu** is found along the top of the window and primarily consists of options which modify how the program runs. The following drop-down menu items are listed: File, Multiple Data Set Analysis, Method Parameters, options and About.

Remark 2.1.1. If not all these options are visible, resize the window to make it wider (i.e. click on the right edge of the window and drag the mouse to the right).

2.1.1 File Drop-down

The File drop-down menu allows the user to save a detailed report of the operations performed. The contents of this report are be given in the " Output of the Detailed Report" section. To enable the saving of a report simply click the File drop-down, hover the mouse over the "Save report?" button and click "Yes" when the option appears. A location for the report must then be chosen by clicking the "Choose File" button which has now been enabled inside the File drop-down. A "Save As" window will then appear and the user may browse to the desired directory and type the name of the file in which the report will be saved.

Remark 2.1.2. The contents of the report file will be overridden each time the program performs an operation (e.g. executing one of the averaging methods). Therefore, be careful when choosing a report file and a new file should be chosen if the contents of a particular report wish to be kept for future reference.

Once a file has been chosen, clicking on the File drop-down reveals a "Current File" button. Clicking this button will allow the user to change which file is being used to write the report to.

The last button under the File drop-down menu is the "Export Data" button. Pressing this button will open a file save window where the user chooses a file to write the *numeric* values of the data points entered in the currently selected dataset. This feature is handy for converting ENSDF format uncertainties to numeric uncertainties.

2.1.2 Point View Drop-Down

This menu access the plotting functionality of V.AveLib. The V.AveLib program only creates scripts and data files which can then be used with GNU Plot (downloadable from http://www.gnuplot.info/) to create plots. Plots created by GNU plot are highly customizable and can be saved to a variety of formats, allowing the user to tailor visualization to his or her own needs. However, no knowledge of GNU plot is required since the default script created by V.AveLib should be sufficient for the needs of most users. For more information about how to use GNU plot see http://www.gnuplot.info/documentation.html.

There are three buttons on this drop-down menu: "Generate GNU Plot Script", "Change Plotting Directory", and "Average to display on plot". The first button will create a the GNU plot script (a '.p' file) and data file (a '.dat' file) in the directory chosen by the user. This directory can be changed after it is chosen using the second button. Once a script is generated the user produces a plot by using the `load` command within GNU plot. A message box appears after the script is generated telling the user exactly what to type into the GNU plot terminal.

The final button on this drop-down menu is a sub-menu where the user can select which averaging method should be used to calculate the average which is shown on the plot. The average is indicated by a coloured solid line and shaded region behind the data points. This allows the user to visually inspect how well the average describes the data and if any particular data points are outliers.

2.1.3 Multiple Data Set Analysis Drop-down

The Multiple Data Set Analysis drop-down has a "Run On All Data Sets" option then a list of all the data sets. The data sets which are checked are those which the averaging methods will be applied to. Checking and unchecking data sets is done by clicking on their names in the Multiple Data Set Analysis drop-down menu. If "Run On All Data Sets" is checked then the list is active and all the datasets will appear checked because the averaging methods will be run on all the data sets associated with tabs at that time.

Remark 2.1.3. Running the averaging methods on more than one data set means that the results will be given subsequently for each data set. It does not mean that all the data will be averaged at once. Running on more than one data set can be thought of as a "batch mode", where several inputs are processed one after the other.

2.1.4 Method Parameters

The "Show Parameters Window" button under the Method parameters drop-down menu will open a new window which may be used to modify the parameters of each method. The parameters which may be modified are as follows:

- "Critical Chi**2 Confidence Level" - the confidence level with which one may reject the hypothesis of the weighted average (and those methods derived from it: Limitation of Statistical Weights, Normalized Residuals Method, and Rajeval Technique) if the reduced Chi-Squared exceeds the critical value.

- "Limitation of Statistical Weights; maximum weight" - sets the highest relative weighting which a data point can have in the Limitation of Statistical Weights Method (LWM), where the relative weight on the ith point is $w_i / \sum_{j=1}^{n} w_j$. Since the weights are those of the normal weighted average, any point with too much weight has its uncertainty increased until the weighting is lowered to the maximum value.

- "Limitation of Statistical Weights; outlier criterion" - Selects the prescription with which to check for outliers before using LWM. Chauvenet's, Peirce's and Birch's criteria have brief descriptions given in the Check for Outliers dialog. The Modified Peirce's criterion is not recommended for use as it is too strict.

- "Normalized Residuals Method; outlier confidence level" - sets the maximum absolute value of the normalized residuals according to the formula given in the original paper for the method, $R = \sqrt{1.8(n/p) + 2.6}$, where n is the number of data points in the data set and p is the probability (in percent) of a normalized residual exceeding the maximum. $p = 100 -$ (the value of this parameter). The uncertainty in increased for data points which exceed the maximum normalized residual until the normalized residual in equal to the maximum value. For more details see James et al. *Nuclear Instruments and Methods in Physics Research A*, vol. 313, p. 277, 1992.

- "Rajeval Technique; outlier confidence level" - The confidence level with which outliers are detected using the statistic for the Rajeval Technique. 95% corresponds to the maximum value of this statistic being 1.96, 99% corresponds to 2×1.96 and for 99.99% 3×1.96. For the definition of the statistic see Rajput et al. *Nuclear Instruments and Methods in Physics Research A*, vol. 312, p. 289, 1992. The user is asked to reject outliers (points for which their statistic exceeds the maximum value).

- "Bootstrap; number of sub-sample medians to take" - "sub-samples" are chosen with replacement from the input data points, which are of size n (the number of input data points). The median of each sub-sample is recorded and the final result of the bootstrap is the unweighted average of these medians. This parameter determines how many sub-samples are chosen and thus how many medians are taken.

- "Mandel-Paule; Numeric Tolerance" - The Mandel-Paule method is implemented using a bisection algorithm to solve an equation. This parameter gives how close the left hand side must be to the right hand side before a solution is returned.

- "Maximum number of iterations" - This parameter gives the maximum number of iterations for the Mandel-Paule bisection algorithm. A warning is given when the algorithm terminates if this number of iterations is exceeded. This prevents a potentially infinite loop from occurring if a data set results a particularly poorly behaved equation.

Once any changes have been made hit "OK" to apply them or "Cancel" to disregard them. The "Restore Default Values" button changes all the parameters to their values when the program is first started.

2.1.5 Options

The "Round results to be no more precise than input data" option under this menu will cause the results generated by the averaging methods to not have more decimal places than the inputs when it is checked.

2.1.6 About

The "Show About Window" button under the About drop-down menu opens a dialogue box which includes the version number, contact information for the developer and journal references for some of the averaging methods.

2.2 The Quick Instructions

The text located just below the **Top Menu** gives a brief description of how to use the program. More detailed information is found in this manual.

2.3 Data Set Tabs

Below the The Quick Instructions on the left side of the window are the **Data Set Tabs**. These allow the user to select the active data set, which can then be edited in the **Data Entry Box** (to be discussed in the next section). These tabs also enable the user to edit the title of a data set (i.e. the name which appears in the text of the tab) and remove the data set. Right-click on a tab to show the menu which gives access to these features. The "Rename" button opens a window to change the title of the data set corresponding to the tab which was right-clicked. The "Remove" button will remove the tab which was right-clicked from the **Data Set Tabs** as well as its associated data set from memory.

2.4 Data Entry Box

The text box below the **Data Set Tabs** is the **Data Entry Box**. It is where a data set can be input and edited. The proper format for data entry is discussed in chapter 3, Data Input Format.

2.5 Side Buttons

Below the **The Quick Instructions** on the right side of the window are the **Side Buttons**. These activate the main functions of the program.

- The "Add New Data Set" button will open a new tab and associate it with a blank data set in memory, which can then be edited using the **Data Entry Box**.
- The "Load Data from File" button will open a file browsing window to choose a data set text file which will be loaded into memory. A new tab will be opened for the loaded data set automatically. The data set files which may be loaded are those which follow the correct format discussed in chapter 3, Data Input Format. Such files can be produced using a text editor.
- The "Check for Outliers" button opens the **Check for Outliers dialog** box which will enable the user to use three different outlier identification prescriptions to find possible outliers in the data set associated with the tab which was selected when the button was pressed. The details of this interface are given in section 2.6.
- The next eight buttons will execute the averaging procedure named on the button on the data sets which are selected in the Multiple Data Set Analysis Drop-down (see section 2.1.3). If the Save Report option is active in the File Drop-down (see section 2.1.1) then a report corresponding to the executed averaging procedure will also be output to the chosen file. The contents of the various reports are detailed in chapter 4. Once the procedure is complete a message box will appear with a summary of the result.
- The "Compare" button will execute all the averaging procedures on the data sets which are selected in the Multiple Data Set Analysis Drop-down (see section 2.1.3). If the Save Report option is active in the File Drop-down (see section 2.1.1), then a report for each method will be output to the chosen file. Once all the procedures have completed, a message box will appear with a summary of all the results.

  Remark 2.5.1. Due to the nature of the Bootstrap method much computation time is needed to apply this procedure. While the Bootstrap is running the interface may not be responsive. Please wait for the completion message box to appear before attempting to interact with the program again.

2.6 Check for Outliers dialog

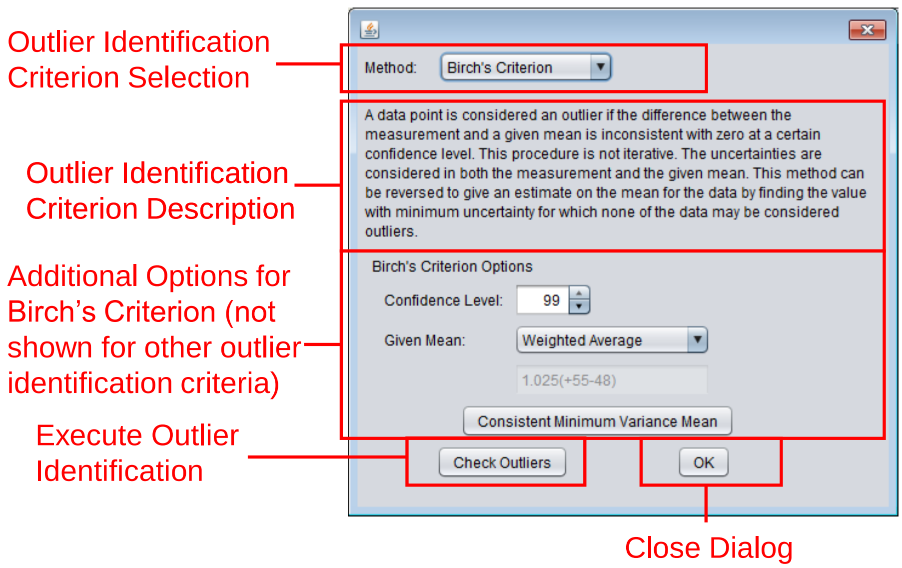

Figure 2.2: The Check for Outliers dialog

The **Check for Outliers** dialog box is opened using the "Check for Outliers" button in the **Side Buttons** section of the window. It enables the user to use three different outlier identification prescriptions to find possible outliers in the data set associated with the tab which was selected when the button was pressed. To select the desired outlier identification criterion use the drop-down menu at the top of the dialogue. Immediately below this drop-down is a brief description of how each criterion works.

For Birch's Criterion there are additional options which appear below the description. The confidence level determines how accepting of outliers the criterion is, a higher confidence level means it is more likely that the identified points are outliers with respect to the selected mean. Approximately speaking, at a confidence level of 68% outliers are points for which their uncertainty range does not overlap the uncertainty range of the mean within one standard deviation, 95% means outliers do not overlap within two standard deviations, and 99.7% mean they do not overlap within three standard deviations. Which mean is used may be chosen using the drop-down box below the confidence level. The means refer to those taken using the data set as it appears in the **Data Entry Box** in the main window and the corresponding value is shown in the textbox below the mean drop-down box. The user may also choose the "Custom" option from the drop-down and enter any mean which the evaluator believes is reasonable. This feature allows the evaluator's mean to be checked for consistency with the data. The "Consistent Minimum Variance Mean" button returns the mean with the smallest uncertainty such that no data point may be considered an outlier at the given confidence level.

To execute the chosen outlier identification method press the "Check Outliers" button in the bottom left of the dialogue. The user will be asked to *exclude* data points which are identified as outliers and may select yes or no based on their own discretion. If no outliers are identified or if the user chooses not to exclude any point a "No outliers chosen" message box will appear and no modifications will be made to the data set. However, if at least one outlier is identified and excluded by the user then a "Outliers indicated in data set" message box will appear and the data set will be modified such that the outliers will not be included in the averaging methods and the reason for the exclusion will be noted. Details about excluding points and data point comments are given in chapter 3.

If the user is finished checking for outliers the "OK" button in the bottom right of the dialog will close it.

Chapter 3

Data Input Format

A data set consists of a series of data points as well as a title. A data point is a measurement which has both a central value and an uncertainty estimate. The format to input a data point is as follows:

Name: CentralValue(Uncertainty) #Comment

where the uncertainty in parentheses is applied to the same number of significant figures as there are digits in the uncertainty and the Name as well as the Comment are optional. The uncertainty in parentheses should never contain a decimal point. Consider the following examples:

Measurement	Input format
1.05 ± 0.03	1.05(3)
9.60 ± 0.4	9.60(40)
10.6 ± 1.8	10.6(18)
57432 ± 386	57432(386)
$22.6^{+0.5}_{-2.7}$	22.6(+5-27)
$378.0^{+10.1}_{-5.9}$	378.0(+101-59)

The central value and uncertainty may also be separated by a space rather than parentheses. For example, "1.05(3)" and "1.05 3" are both correct formats for the first measurement in the previous table. For asymmetric uncertainties the order of the upper and lower uncertainties may be inter-changed as long as the correct symbol precedes it. For example, $22.6^{+0.5}_{-2.7}$ may be entered as "22.6(+5-27)" or "22.6(-27+5)." Again a space may also be used instead of parentheses.

In a data set, only one data point may be on a line. To add a comment to a data point use a hash symbol (#) after the data point. For example, "9.60(40) #This is an example comment" is a valid data point and comment which would constitute a single line in a data set. Lines which begin with a hash symbol will be ignored in by the averaging methods. Each data point may also be given a name by typing it before the data point and separating it from the data point by a colon. For example, "Data1: 10.6(18) #The first measurement." The default name of a data point is its number in the data set. Names can include spaces, but not hash symbols since these are used to denote comments.

A data set may be input directly into the V.AveLib program using the **Data Entry Box** or loaded from a text file using the "Load Data from File" button in the **Side Buttons**. The title of a data set may be changed in the V.AveLib program using the "Rename" button in the right-click menu of the **Data Set Tabs**. To have a title included in a text file which will be loaded into the program, start it with a line beginning with "Title=" and followed by the desired title. The "Title=" command is not case sensitive, however do not separate the equals sign from the word "Title". A new data set can be added within the V.AveLib program using the "Add New Data Set" button in the **Side Buttons**. More than one data set may be included in a text file by separating them using the line "*new". The example on the following page is for a text file containing two data sets with three data points that will be used for averaging in each.

```
Title=First example data set
#This is a general comment about the measurements
#in this data set
FirstName: 1.00(10) #The first data point
1.05 9 #The name of this point will be "2"
#BadPoint: 7.9 50 #This point will be ignored
NoComment: 1.02 +10-6
*new
Title=Second Example
#This is a new data set
12.7 8 #The name of this point will be "1"
Asymmetric: 12.2(-3+10)
Name with spaces: 13 2
```

Chapter 4

Output of the Detailed Report

A detailed report of the operations performed by the V.AveLib program can be saved to a file using the File Drop-down in the **Top Menu** (section 2.1.1). Each averaging method gives different information in its report, however all reports have similar headers. First the title of the data set upon which the analysis was performed is given within a line of asterisks (*). A solid line of asterisks is used to end the current data set and if multiple data sets were being analyzed subsequently (see the Multiple Data Set Analysis Drop-down in the **Top Menu**, section 2.1.3), the next one would begin with its own title in a line of asterisks. In this way the asterisks separate the report into blocks according to data set. Similarly, within a data set block the report is split into the information blocks given by individual averaging procedures using dashes (-). The name of the method is given in a line of dashes, then the report information follows, and it ends with a solid line of dashes. The details of the information given by each method follows.

4.1 The Unweighted Average

The unweighted average report begins with a table listing the input data points and their squared deviations from the mean ($\overline{x}$), $(x-\overline{x})^2$. It then gives the number of input data points and the final result.

4.2 The Weighted Average

The weighted average report begins with a table listing the input data points, their relative weighting, $\frac{w_i}{\sum w_i}$, and contribution towards the total Chi-Square, $\left(\frac{x_i-\mu}{\sigma_i}\right)^2$, where μ is the weighted average. It then gives the number of input data points, the reduced Chi-Square, $\frac{\chi^2}{n-1}$, the critical reduced Chi-Square and the result with both internal and external uncertainties.

4.3 The Limitation of Relative Statistical Weights Method (LWM)

The LWM report gives the same information as the weighted average, however if the Chi-Square is too large and the most precise value had its uncertainty increased to limit its weight then that will marked by a double asterisk (**) next to the relative weighting. The adjusted uncertainty and corresponding weighting is what is listed in the table. If the uncertainty of the final result was increased to overlap the most precise value then this is also listed in the report. If the user decided to adopt the unweighted mean because the data is highly discrepant then the report will give the number of input values and the final result only. If any data points were marked as outliers using the criterion selected in the Method Parameters dialog accessible from the **Top Menu** (section 2.1.4), then these are listed after the table. Note that the "Number of input values" number given after the table refers to the number of entries which contributed to the final result (i.e. the ones which are listed in the table) plus the outliers (i.e. it includes ALL the input entries).

4.4 The Normalized Residuals Methods (NRM)

The NRM report gives the same information as the weighted average with an extra column in the table for the value of each normalized residual. If any uncertainties were adjusted these are indicated with a double asterisk (**) next to the relative weighting. It is then the modified uncertainty and associated weighting which is listed in the table.

4.5 The Rajeval Technique (RT)

The RT report gives the same information as the weighted average. If any uncertainties were adjusted these are indicated with a double asterisk (**) next to the relative weighting. It is then the modified uncertainty and associated weighting which is listed in the table. If any points were rejected as outliers using the criterion built into the method then these are listed after the table. Note that the "Number of input values" number given after the table refers to the number of entries which contributed to the final result (i.e. the ones which are listed in the table) plus the outliers (i.e. it includes ALL the input entries).

4.6 The Expected Value Method (EVM)

The EVM report begins with a table which lists the input data points and their relative weightings. This is followed by a summary of the confidence level calculation which is based on correctly predicting how many points fall above and below the mean. It then lists the number of input values, the confidence level for accepting the final result, and the final result with both internal and external uncertainties.

4.7 The Bootstrap Method

The Bootstrap report lists the number of input values, the two sampling parameters which can be set in the Method Parameters dialog accessible from the **Top Menu** (section 2.1.4), the reduced Chi-Square for the final result and the final result.

4.8 The Mandel-Paule (MP) Method

The MP report gives the same table as the EVM report then lists the number of input data points, the reduced Chi-Square for the final result and the final result.

Chapter 5

Mathematical Description of Each Method

Throughout this chapter we adopt the following notation:

- n will be the number of input data points
- the central value of the i^{th} measurement will be denoted by x_i
- the upper and lower uncertainties of the i^{th} measurement will be denoted by σ_i^+ and σ_i^-, respectively
- the variance of the the i^{th} measurement will be denoted as σ_i^2 and calculated from the upper and lower uncertainties according to the formula

$$\sigma_i^2 = \left(1 - \frac{2}{\pi}\right)\left(\sigma_i^+ - \sigma_i^-\right)^2 + \sigma_i^+ \sigma_i^- . \tag{5.1}$$

  This equation follows from calculating the variance of the asymmetric Gaussian distribution associated with the i^{th} measurement, which is defined by the probability density function

$$\mathcal{A}\left(t; x_i, \sigma_i^+, \sigma_i^-\right) = \begin{cases} \sqrt{\frac{2}{\pi\left(\sigma_i^+ + \sigma_i^-\right)^2}}\, e^{-\frac{(t-x_i)^2}{2\left(\sigma_i^-\right)^2}} & , t \leq x_i \\ \sqrt{\frac{2}{\pi\left(\sigma_i^+ + \sigma_i^-\right)^2}}\, e^{-\frac{(t-x_i)^2}{2\left(\sigma_i^+\right)^2}} & , t > x_i \end{cases} . \tag{5.2}$$

5.1 The Unweighted Average

The central value of the unweighted average is given by

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i .$$

The uncertainty on this result is given by

$$\sigma_{\text{uwa}} = \max \left\{ \sqrt{\frac{\sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2}{n(n-1)}}, \frac{1}{\sqrt{\sum_{i=1}^{n} 1/\sigma_i^2}} \right\} .$$

5.2 The Weighted Average

To generalize the standard weighted average to handle asymmetric uncertainties we use an approach based on that described by R. Barlow (e.g. arXiv:physics/0406120 [physics.data-an]).

In the standard weighted average one assumes the data under consideration were sampled from Gaussian distributions with known standard deviations (the experimental uncertainties) and a common, unknown mean (the "true value" of the quantity being measured). The formula for the weighted mean then follows by computing the maximum likelihood estimate of the unknown mean. By analogy, we will assume that the data were sampled from *asymmetric* Gaussian distributions (c.f. Equation 5.2) with known upper and lower standard deviations and common, unknown central value. We will then compute the maximum likelihood estimate of this unknown central value.

Given our above assumption, the log-likelihood as a function of the common central value, $\hat{x}$, is

$$\ln L(\hat{x}) = -\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i - \hat{x}}{\sigma_i(x_i;\hat{x})}\right)^2 + \frac{n}{2}\ln\left(\frac{2}{\pi}\right) - \sum_{i=1}^{n}\ln(\sigma_i^+ + \sigma_i^-),$$

where $\sigma_i(x;\hat{x}) = \begin{cases}\sigma_i^- & x \le \hat{x} \\ \sigma_i^+ & x > \hat{x}\end{cases}$ is the varying width of the asymmetric Gaussian for the ith measurement (which also depends on the unknown central value!). Differentiating $\ln L$ with respect to $\hat{x}$, we obtain the following equation for the maximum likelihood value of the common central value

$$\hat{x}\sum_{i=1}^{n}\frac{1}{[\sigma_i(x_i;\hat{x})]^2} = \sum_{i=1}^{n}\left(\frac{x_i}{[\sigma_i(x_i;\hat{x})]^2} + \frac{(x_i-\hat{x})^2}{[\sigma_i(x_i;\hat{x})]^3}\frac{\partial\sigma_i}{\partial\hat{x}}\right).$$

Since $\sigma_i(x;\hat{x})$ is a step function, its derivative with respect to $\hat{x}$ is zero except when $\hat{x} = x_i$, and in the latter case, the factor in front of the derivative is zero. Hence, we can write

$$\hat{x} = \frac{\sum_{i=1}^{n} x_i/[\sigma_i(x_i;\hat{x})]^2}{\sum_{i=1}^{n} 1/[\sigma_i(x_i;\hat{x})]^2}, \tag{5.3}$$

which has the same form as the standard weighted average equation, however the right-hand side also depends on $\hat{x}$. We can therefore think of the maximum likelihood estimator as the fixed point of the function

$$f(\hat{x}) = \frac{\sum_{i=1}^{n} x_i/[\sigma_i(x_i;\hat{x})]^2}{\sum_{i=1}^{n} 1/[\sigma_i(x_i;\hat{x})]^2},$$

or equivalently, the root of the function

$$g(\hat{x}) = \hat{x} - \frac{\sum_{i=1}^{n} x_i/[\sigma_i(x_i;\hat{x})]^2}{\sum_{i=1}^{n} 1/[\sigma_i(x_i;\hat{x})]^2}.$$

A root finding algorithm (Brent's Method in the case of V.AveLib) can therefore be used to compute $\hat{x}$.

Following Barlow's prescription, the internal uncertainty interval in associated with the maximum likelihood estimate is given by the points where the likelihood is half a unit lower than at the maximum, i.e. the values of x such that

$$\Delta(\ln L) = \ln L(\hat{x}) - \ln L(x) = -\frac{1}{2}.$$

Hence, the internal uncertainty is found by using the root finding algorithm again to solve $\ln L(\hat{x}) - \ln L(x) + \frac{1}{2} = 0$. The external interval uncertainty is then found by multiplying the internal uncertainty interval by $\sqrt{\chi^2/(n-1)}$, where

$$\chi^2 = \sum_{i=1}^{n}\left(\frac{x_i - \hat{x}}{\sigma_i(x_i;\hat{x})}\right)^2.$$

5.3 The Limitation of Relative Statistical Weights Method

See LWEIGHT documentation. Note: the weighted average portion of the method is as in section 5.2.

5.4 The Normalized Residuals Method

See M.F. James, R.W. Mills, D.R. Weaver, *Nucl. Instr. and Meth. in Phys. Res.* **A313**, 277 (1992). Note: σ_i^2 in the linked article is implemented in this program using Equation 5.1.

5.5 The Rajeval Technique

See M.U. Rajput and T.D. MacMahon, *Nucl. Instr. and Meth. in Phys. Res.* **A312**, 289 (1992). Note: σ_i^2 in the linked article is implemented in this program using Equation 5.1.

5.6 The Expected Value Method

See M. Birch, B. Singh, *Nucl. Data Sheets* **120**, 106 (2014).

5.7 The Bootstrap Method

The bootstrap method creates new datasets of size n points by sampling from the collection of probability distributions associated with the dataset. The median of each generated dataset is taken and the final result is the unweighted average (as in section 5.1) of these medians. The algorithm is as follows:

```
For k in 1:(number of sub-samples){
  For j in 1:n{
    i = (random integer between 1 and n)
    newdataset[j] = (random number sampled from asym. Gaussian
                            distribution associated with measurement i)
  }
  medians[k] = (the median of newdataset)
}
Return (the unweighted average of medians)
```

5.8 The Mandel-Paule Method

See A.L. Rukhin and M.G. Vangel, *J. Am. Stat. Assoc.* **93** 303 (1998). Note: σ_i^2 in the linked article is implemented in this program using Equation 5.1.

5.9 Chauvenet's Outlier Criterion

See "Cleaning Data the Chauvenet Way".

5.10 Peirce's Outlier Criterion

See B. Peirce, *Astronomical Journal* vol. 2, iss. 45 161 (1852), also "Peirce's criterion for the elimination of suspect experimental data".

5.11 Birch's Outlier Criterion

Let X be a normally distributed random variable with mean μ_1 and standard deviation σ_1, denoted $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and let $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$. Without loss of generality, assume that $\mu_1 > \mu_2$. Define $Z = X - Y$. One can prove that $Z \sim \mathcal{N}(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$, and hence it follows that

$$\Pr(Z > 0) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\mu_1 - \mu_2}{\sqrt{2(\sigma_1^2 + \sigma_2^2)}}\right)\right].$$

Motivated by this result, we have the following definition: given two measurements, $x_1 \pm \sigma_1$ and $x_2 \pm \sigma_2$, we say they are inconsistent with probability p if

$$\frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{|x_1 - x_2|}{\sqrt{2(\sigma_1^2 + \sigma_2^2)}}\right)\right] > p. \tag{5.4}$$

Note that the absolute value of the difference is used in Equation 5.4 since we assumed that $\mu_1 > \mu_2$ in the above derivation.

Birch's outlier criterion identifies measurements as outliers at a given probability, p, and with respect to a given mean, $\mu \pm \sigma$. This is done by simply checking which measurements are inconsistent with the mean in the sense of the above definition. By default, the numeric implementation of Birch's criterion uses the weighted average as the given mean, however the user has the option to provide a different mean.

This idea can also be reversed to find the mean which in some sense is most consistent with the dataset. This "consistent minimum variance mean" is computed by finding the mean such that the associated variance in order to be consistent with the entire dataset (in the above sense) is minimized. The user can perform this computation by clicking the "Consistent Minimum Variance Mean" button in the Birch's Criterion Options within the Check for Outliers dialog.