////////////////////////////////////////////////////////////////////////////////

/// \f$ \chi^{2} \f$ test for comparing weighted and unweighted histograms

///

/// Function: Returns p-value. Other return values are specified by the 3rd parameter

///

/// \param[in] h2 the second histogram

/// \param[in] option

/// - "UU" = experiment experiment comparison (unweighted-unweighted)

/// - "UW" = experiment MC comparison (unweighted-weighted). Note that

/// the first histogram should be unweighted

/// - "WW" = MC MC comparison (weighted-weighted)

/// - "NORM" = to be used when one or both of the histograms is scaled

/// but the histogram originally was unweighted

/// - by default underflows and overflows are not included:

/// \* "OF" = overflows included

/// \* "UF" = underflows included

/// - "P" = print chi2, ndf, p\_value, igood

/// - "CHI2" = returns chi2 instead of p-value

/// - "CHI2/NDF" = returns \f$ \chi^{2} \f$/ndf

/// \param[in] res not empty - computes normalized residuals and returns them in this array

///

/// The current implementation is based on the papers \f$ \chi^{2} \f$ test for comparison

/// of weighted and unweighted histograms" in Proceedings of PHYSTAT05 and

/// "Comparison weighted and unweighted histograms", arXiv:physics/0605123

/// by N.Gagunashvili. This function has been implemented by Daniel Haertl in August 2006.

///

/// #### Introduction:

///

/// A frequently used technique in data analysis is the comparison of

/// histograms. First suggested by Pearson [1] the \f$ \chi^{2} \f$ test of

/// homogeneity is used widely for comparing usual (unweighted) histograms.

/// This paper describes the implementation modified \f$ \chi^{2} \f$ tests

/// for comparison of weighted and unweighted histograms and two weighted

/// histograms [2] as well as usual Pearson's \f$ \chi^{2} \f$ test for

/// comparison two usual (unweighted) histograms.

///

/// #### Overview:

///

/// Comparison of two histograms expect hypotheses that two histograms

/// represent identical distributions. To make a decision p-value should

/// be calculated. The hypotheses of identity is rejected if the p-value is

/// lower then some significance level. Traditionally significance levels

/// 0.1, 0.05 and 0.01 are used. The comparison procedure should include an

/// analysis of the residuals which is often helpful in identifying the

/// bins of histograms responsible for a significant overall \f$ \chi^{2} \f$ value.

/// Residuals are the difference between bin contents and expected bin

/// contents. Most convenient for analysis are the normalized residuals. If

/// hypotheses of identity are valid then normalized residuals are

/// approximately independent and identically distributed random variables

/// having N(0,1) distribution. Analysis of residuals expect test of above

/// mentioned properties of residuals. Notice that indirectly the analysis

/// of residuals increase the power of \f$ \chi^{2} \f$ test.

///

/// #### Methods of comparison:

///

/// \f$ \chi^{2} \f$ test for comparison two (unweighted) histograms:

/// Let us consider two histograms with the same binning and the number

/// of bins equal to r. Let us denote the number of events in the ith bin

/// in the first histogram as ni and as mi in the second one. The total

/// number of events in the first histogram is equal to:

/// \f[

/// N = \sum\_{i=1}^{r} n\_{i}

/// \f]

/// and

/// \f[

/// M = \sum\_{i=1}^{r} m\_{i}

/// \f]

/// in the second histogram. The hypothesis of identity (homogeneity) [3]

/// is that the two histograms represent random values with identical

/// distributions. It is equivalent that there exist r constants p1,...,pr,

/// such that

/// \f[

///\sum\_{i=1}^{r} p\_{i}=1

/// \f]

/// and the probability of belonging to the ith bin for some measured value

/// in both experiments is equal to pi. The number of events in the ith

/// bin is a random variable with a distribution approximated by a Poisson

/// probability distribution

/// \f[

///\frac{e^{-Np\_{i}}(Np\_{i})^{n\_{i}}}{n\_{i}!}

/// \f]

///for the first histogram and with distribution

/// \f[

///\frac{e^{-Mp\_{i}}(Mp\_{i})^{m\_{i}}}{m\_{i}!}

/// \f]

/// for the second histogram. If the hypothesis of homogeneity is valid,

/// then the maximum likelihood estimator of pi, i=1,...,r, is

/// \f[

///\hat{p}\_{i}= \frac{n\_{i}+m\_{i}}{N+M}

/// \f]

/// and then

/// \f[

/// X^{2} = \sum\_{i=1}^{r}\frac{(n\_{i}-N\hat{p}\_{i})^{2}}{N\hat{p}\_{i}} + \sum\_{i=1}^{r}\frac{(m\_{i}-M\hat{p}\_{i})^{2}}{M\hat{p}\_{i}} =\frac{1}{MN} \sum\_{i=1}^{r}\frac{(Mn\_{i}-Nm\_{i})^{2}}{n\_{i}+m\_{i}}

/// \f]

/// has approximately a \f$ \chi^{2}\_{(r-1)} \f$ distribution [3].

/// The comparison procedure can include an analysis of the residuals which

/// is often helpful in identifying the bins of histograms responsible for

/// a significant overall \f$ \chi^{2} \f$ value. Most convenient for

/// analysis are the adjusted (normalized) residuals [4]

/// \f[

/// r\_{i} = \frac{n\_{i}-N\hat{p}\_{i}}{\sqrt{N\hat{p}\_{i}}\sqrt{(1-N/(N+M))(1-(n\_{i}+m\_{i})/(N+M))}}

/// \f]

/// If hypotheses of homogeneity are valid then residuals ri are

/// approximately independent and identically distributed random variables

/// having N(0,1) distribution. The application of the \f$ \chi^{2} \f$ test has

/// restrictions related to the value of the expected frequencies Npi,

/// Mpi, i=1,...,r. A conservative rule formulated in [5] is that all the

/// expectations must be 1 or greater for both histograms. In practical

/// cases when expected frequencies are not known the estimated expected

/// frequencies \f$ M\hat{p}\_{i}, N\hat{p}\_{i}, i=1,...,r \f$ can be used.

///

/// #### Unweighted and weighted histograms comparison:

///

/// A simple modification of the ideas described above can be used for the

/// comparison of the usual (unweighted) and weighted histograms. Let us

/// denote the number of events in the ith bin in the unweighted

/// histogram as ni and the common weight of events in the ith bin of the

/// weighted histogram as wi. The total number of events in the

/// unweighted histogram is equal to

///\f[

/// N = \sum\_{i=1}^{r} n\_{i}

///\f]

/// and the total weight of events in the weighted histogram is equal to

///\f[

/// W = \sum\_{i=1}^{r} w\_{i}

///\f]

/// Let us formulate the hypothesis of identity of an unweighted histogram

/// to a weighted histogram so that there exist r constants p1,...,pr, such

/// that

///\f[

/// \sum\_{i=1}^{r} p\_{i} = 1

///\f]

/// for the unweighted histogram. The weight wi is a random variable with a

/// distribution approximated by the normal probability distribution

/// \f$ N(Wp\_{i},\sigma\_{i}^{2}) \f$ where \f$ \sigma\_{i}^{2} \f$ is the variance of the weight wi.

/// If we replace the variance \f$ \sigma\_{i}^{2} \f$

/// with estimate \f$ s\_{i}^{2} \f$ (sum of squares of weights of

/// events in the ith bin) and the hypothesis of identity is valid, then the

/// maximum likelihood estimator of pi,i=1,...,r, is

///\f[

/// \hat{p}\_{i} = \frac{Ww\_{i}-Ns\_{i}^{2}+\sqrt{(Ww\_{i}-Ns\_{i}^{2})^{2}+4W^{2}s\_{i}^{2}n\_{i}}}{2W^{2}}

///\f]

/// We may then use the test statistic

///\f[

/// X^{2} = \sum\_{i=1}^{r} \frac{(n\_{i}-N\hat{p}\_{i})^{2}}{N\hat{p}\_{i}} + \sum\_{i=1}^{r} \frac{(w\_{i}-W\hat{p}\_{i})^{2}}{s\_{i}^{2}}

///\f]

/// and it has approximately a \f$ \sigma^{2}\_{(r-1)} \f$ distribution [2]. This test, as well

/// as the original one [3], has a restriction on the expected frequencies. The

/// expected frequencies recommended for the weighted histogram is more than 25.

/// The value of the minimal expected frequency can be decreased down to 10 for

/// the case when the weights of the events are close to constant. In the case

/// of a weighted histogram if the number of events is unknown, then we can

/// apply this recommendation for the equivalent number of events as

///\f[

/// n\_{i}^{equiv} = \frac{ w\_{i}^{2} }{ s\_{i}^{2} }

///\f]

/// The minimal expected frequency for an unweighted histogram must be 1. Notice

/// that any usual (unweighted) histogram can be considered as a weighted

/// histogram with events that have constant weights equal to 1.

/// The variance \f$ z\_{i}^{2} \f$ of the difference between the weight wi

/// and the estimated expectation value of the weight is approximately equal to:

///\f[

/// z\_{i}^{2} = Var(w\_{i}-W\hat{p}\_{i}) = N\hat{p}\_{i}(1-N\hat{p}\_{i})\left(\frac{Ws\_{i}^{2}}{\sqrt{(Ns\_{i}^{2}-w\_{i}W)^{2}+4W^{2}s\_{i}^{2}n\_{i}}}\right)^{2}+\frac{s\_{i}^{2}}{4}\left(1+\frac{Ns\_{i}^{2}-w\_{i}W}{\sqrt{(Ns\_{i}^{2}-w\_{i}W)^{2}+4W^{2}s\_{i}^{2}n\_{i}}}\right)^{2}

///\f]

/// The residuals

///\f[

/// r\_{i} = \frac{w\_{i}-W\hat{p}\_{i}}{z\_{i}}

///\f]

/// have approximately a normal distribution with mean equal to 0 and standard

/// deviation equal to 1.

///

/// #### Two weighted histograms comparison:

///

/// Let us denote the common weight of events of the ith bin in the first

/// histogram as w1i and as w2i in the second one. The total weight of events

/// in the first histogram is equal to

///\f[

/// W\_{1} = \sum\_{i=1}^{r} w\_{1i}

///\f]

/// and

///\f[

/// W\_{2} = \sum\_{i=1}^{r} w\_{2i}

///\f]

/// in the second histogram. Let us formulate the hypothesis of identity of

/// weighted histograms so that there exist r constants p1,...,pr, such that

///\f[

/// \sum\_{i=1}^{r} p\_{i} = 1

///\f]

/// and also expectation value of weight w1i equal to W1pi and expectation value

/// of weight w2i equal to W2pi. Weights in both the histograms are random

/// variables with distributions which can be approximated by a normal

/// probability distribution \f$ N(W\_{1}p\_{i},\sigma\_{1i}^{2}) \f$ for the first histogram

/// and by a distribution \f$ N(W\_{2}p\_{i},\sigma\_{2i}^{2}) \f$ for the second.

/// Here \f$ \sigma\_{1i}^{2} \f$ and \f$ \sigma\_{2i}^{2} \f$ are the variances

/// of w1i and w2i with estimators \f$ s\_{1i}^{2} \f$ and \f$ s\_{2i}^{2} \f$ respectively.

/// If the hypothesis of identity is valid, then the maximum likelihood and

/// Least Square Method estimator of pi,i=1,...,r, is

///\f[

/// \hat{p}\_{i} = \frac{w\_{1i}W\_{1}/s\_{1i}^{2}+w\_{2i}W\_{2} /s\_{2i}^{2}}{W\_{1}^{2}/s\_{1i}^{2}+W\_{2}^{2}/s\_{2i}^{2}}

///\f]

/// We may then use the test statistic

///\f[

/// X^{2} = \sum\_{i=1}^{r} \frac{(w\_{1i}-W\_{1}\hat{p}\_{i})^{2}}{s\_{1i}^{2}} + \sum\_{i=1}^{r} \frac{(w\_{2i}-W\_{2}\hat{p}\_{i})^{2}}{s\_{2i}^{2}} = \sum\_{i=1}^{r} \frac{(W\_{1}w\_{2i}-W\_{2}w\_{1i})^{2}}{W\_{1}^{2}s\_{2i}^{2}+W\_{2}^{2}s\_{1i}^{2}}

///\f]

/// and it has approximately a \f$ \chi^{2}\_{(r-1)} \f$ distribution [2].

/// The normalized or studentised residuals [6]

///\f[

/// r\_{i} = \frac{w\_{1i}-W\_{1}\hat{p}\_{i}}{s\_{1i}\sqrt{1 - \frac{1}{(1+W\_{2}^{2}s\_{1i}^{2}/W\_{1}^{2}s\_{2i}^{2})}}}

///\f]

/// have approximately a normal distribution with mean equal to 0 and standard

/// deviation 1. A recommended minimal expected frequency is equal to 10 for

/// the proposed test.

///

/// #### Numerical examples:

///

/// The method described herein is now illustrated with an example.

/// We take a distribution

///\f[

/// \phi(x) = \frac{2}{(x-10)^{2}+1} + \frac{1}{(x-14)^{2}+1} (1)

///\f]

/// defined on the interval [4,16]. Events distributed according to the formula

/// (1) are simulated to create the unweighted histogram. Uniformly distributed

/// events are simulated for the weighted histogram with weights calculated by

/// formula (1). Each histogram has the same number of bins: 20. Fig.1 shows

/// the result of comparison of the unweighted histogram with 200 events

/// (minimal expected frequency equal to one) and the weighted histogram with

/// 500 events (minimal expected frequency equal to 25)

/// Begin\_Macro

/// ../../../tutorials/math/chi2test.C

/// End\_Macro

/// Fig 1. An example of comparison of the unweighted histogram with 200 events

/// and the weighted histogram with 500 events:

/// 1. unweighted histogram;

/// 2. weighted histogram;

/// 3. normalized residuals plot;

/// 4. normal Q-Q plot of residuals.

///

/// The value of the test statistic \f$ \chi^{2} \f$ is equal to

/// 21.09 with p-value equal to 0.33, therefore the hypothesis of identity of

/// the two histograms can be accepted for 0.05 significant level. The behavior

/// of the normalized residuals plot (see Fig. 1c) and the normal Q-Q plot

/// (see Fig. 1d) of residuals are regular and we cannot identify the outliers

/// or bins with a big influence on \f$ \chi^{2} \f$.

///

/// The second example presents the same two histograms but 17 events was added

/// to content of bin number 15 in unweighted histogram. Fig.2 shows the result

/// of comparison of the unweighted histogram with 217 events (minimal expected

/// frequency equal to one) and the weighted histogram with 500 events (minimal

/// expected frequency equal to 25)

/// Begin\_Macro

/// ../../../tutorials/math/chi2test.C(17)

/// End\_Macro

/// Fig 2. An example of comparison of the unweighted histogram with 217 events

/// and the weighted histogram with 500 events:

/// 1. unweighted histogram;

/// 2. weighted histogram;

/// 3. normalized residuals plot;

/// 4. normal Q-Q plot of residuals.

///

/// The value of the test statistic \f$ \chi^{2} \f$ is equal to

/// 32.33 with p-value equal to 0.029, therefore the hypothesis of identity of

/// the two histograms is rejected for 0.05 significant level. The behavior of

/// the normalized residuals plot (see Fig. 2c) and the normal Q-Q plot (see

/// Fig. 2d) of residuals are not regular and we can identify the outlier or

/// bin with a big influence on \f$ \chi^{2} \f$.

///

/// #### References:

///

/// - [1] Pearson, K., 1904. On the Theory of Contingency and Its Relation to

/// Association and Normal Correlation. Drapers' Co. Memoirs, Biometric

/// Series No. 1, London.

/// - [2] Gagunashvili, N., 2006. \f$ \sigma^{2} \f$ test for comparison

/// of weighted and unweighted histograms. Statistical Problems in Particle

/// Physics, Astrophysics and Cosmology, Proceedings of PHYSTAT05,

/// Oxford, UK, 12-15 September 2005, Imperial College Press, London, 43-44.

/// Gagunashvili,N., Comparison of weighted and unweighted histograms,

/// arXiv:physics/0605123, 2006.

/// - [3] Cramer, H., 1946. Mathematical methods of statistics.

/// Princeton University Press, Princeton.

/// - [4] Haberman, S.J., 1973. The analysis of residuals in cross-classified tables.

/// Biometrics 29, 205-220.

/// - [5] Lewontin, R.C. and Felsenstein, J., 1965. The robustness of homogeneity

/// test in 2xN tables. Biometrics 21, 19-33.

/// - [6] Seber, G.A.F., Lee, A.J., 2003, Linear Regression Analysis.

/// John Wiley & Sons Inc., New York.