Chemistry 722
Homework # 2

1. [30 points] You will receive the source code (hw2.2.c) for a C program that will read in two ascii files, each containing a series of replicate measurements, into arrays \( x \) and \( h \), respectively. Your subroutines \( \text{mean()} \) and \( \text{variance()} \), from Homework One will be called for each of these arrays and the results reported. You will then be asked for a common bin width to histogram the two data sets into the arrays \( g \) and \( h \), respectively, with common bin centers contained in the array \( b \). Then the following subroutine will be called:

   \[
   \text{float chiSquaredTest2(float g[], float h[], int nbin)} \quad \text{This subroutine returns the } \chi^2 \text{ value for the comparison}
   \]

   \[
   \chi^2 = \sum_{i=1}^{\text{\#bins}} \frac{[\sqrt{\frac{H}{G}} \times g(x_i) - \sqrt{\frac{G}{H}} \times h(x_i)]^2}{g(x_i) + h(x_i)}
   \]

   where \( H = \sum_i h(x_i) \) and \( G = \sum_i g(x_i) \). Given as input are the histograms \( g(x_i) \) contained in the array \( g \), and \( h(x_i) \) contained in the array \( h \), and the number of bins given in the integer variable \( \text{nbin} \).

   After this routine is called, the resulting \( \chi^2 \) value will be reported. You will then be asked for the number of degrees of freedom so that the probability that other distributions would give worst agreement, that is,

   \[
   Q(\chi^2|\nu) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_{\chi^2}^{\infty} (\chi^2)^{(\nu-2)/2} e^{-\chi^2/2} d\chi^2.
   \]

   can be calculated and reported.

   Your homework assignment is to write the \( \text{chiSquaredTest2()} \) subroutine [15 points] and then using your Homework One baseline values use the \( \text{hw2.2.c} \) program to calculate the \( \chi^2 \), \( \chi^2 \) reduced, and the probability that the 50 point baseline data set and the 1000 point baseline data set come from the same parent distribution. (Tip: If a given bin has no data for either data set, then skip this bin in the sum, and reduce the number of degrees of freedom by one).

2. [50 points] You will receive the source code (hw2.3.c) for a C program that will read an ascii file containing a series of replicate measurements into an array \( x \). Your subroutines \( \text{mean()} \) and \( \text{variance()} \), from Homework One will be called using this array and the results reported. You will then be asked for a bin width to histogram the data set into the array \( h \) with bin centers contained in the array \( b \). Then the following routines will be called:

   \[
   \text{void gaussian(float mean, float variance, int nbin, int N, float g[], float b[])} \quad \text{This subroutine returns \text{nbin} Gaussian histogram intensities in the array } g \text{ given the bin centers in the array } b \text{ and the parameters } \text{mean} \text{ and } \text{variance}. \text{ The Gaussian histogram is normalized to } N \text{ measurements.}
   \]

   \[
   \text{float chiSquaredTest3(float g[], float h[], int nbin)} \quad \text{This subroutine returns the } \chi^2 \text{ value for the comparison}
   \]

   \[
   \chi^2 = \sum_{i=1}^{\text{\#bins}} \frac{[g(x_i) - h(x_i)]^2}{h(x_i)}
   \]

   Given as input are the theoretical histogram \( g(x_i) \) contained in the array \( g \), and the experimental histogram \( h(x_i) \) contained in the array \( h \), and the number of bins given in the integer variable \( \text{nbin} \).

   After this routine \( \text{gaussian()} \) subroutine is called, the \( \text{chiSquaredTest3()} \) subroutine will be called to compare your experimental histogram in the array \( h \) with the Gaussian histogram in the array \( g \) and the resulting \( \chi^2 \) value will be reported. You will then be asked for the number of degrees of freedom so that the probability that other distributions would give worst agreement can be calculated and reported.

   Your homework assignment is to write the \( \text{gaussian()} \) [10 points] and \( \text{chiSquaredTest3()} \) [10 points] subroutines, and then use the \( \text{hw2.3.c} \) program to complete the following assignments using your Homework One baseline values:

   (a) [15 points] Calculate the probability that the parent distribution of your 50 point baseline is Gaussian.

   (b) [15 points] Calculate the probability that the parent distribution of your 1000 point baseline is Gaussian.