# DATA PROCESSING CALCULATOR

## Abstract

After an experiment, the most difficult part is to process the data, during which, evaluating the data is vital. In order to show the reliability of your data, calculating the uncertainties of the data is of necessity.

This process is miscellaneous and toilsome, in the other hand, however, this process is done in a fixed pattern, which means that it is suitable for computer to finish. We intended to exploit such a program to do this fussy work and it is believed that this program will be of great help in physical experiment.

What’s more, linear data is the most common kind to outcome in physical experiment, based on which, the functions of processing linear data were added in our program.

## Introduction/Problem Statement

**Introduction of Uncertainties**

1. For the direct measurement quantities, there are four types of uncertainties to evaluate them.

|  |  |
| --- | --- |
| Type-A | The evaluation of the experiment operation and the deviation of the original data. |
| Type-B | The evaluation of the experiment operation and the deviation of the original data. |
| Type-C | The combination of type-A and type-B uncertainties mentioned above. |
| Extensive | The uncertainties which is the function of confidence probabilities. |

2. For indirect measurement quantities.

Researchers are supposed to calculate out every type of uncertainties of each variable and then calculate the uncertainties of indirect quantities by uncertainty transfer formulas.

**Functions of the Program**

The diagram above shows the main structure of our programs. Using our program, users are available to calculate four types of uncertainties for direct measurement quantities, and the uncertainties of indirect measurement quantities can also be calculated out by easily operations.

As for the linear data processing, there are two ways of processing and evaluation supplying in our program. Using our program, they can not only figure out the slopes and intercepts, the questions that how linear-like the data is and what the deviation is can also be answered.

The diagram below shows the main structure and functions of our program.

**Improvement**

Compared with other such programs, our program is not limited to one specific experiment. The users can choose how many measurement quantities they want to analyze and uncertainties of the quantities will be calculated. Nine the most common formulas are also provided in the program, so that the users are able to calculate more indirect uncertainties in not only one experiment. It is not supposed to neglected that in our physics labs, the computers there are only equipped only one program which can only calculate the uncertainties for one specific.

What’s more, our program is attached to the functions of evaluating the linear data, which is more convenient that other this kind of programs.

## **Group Division**

|  |  |  |
| --- | --- | --- |
| Name | student ID | College |
| A |  | p |
| B |  | P |
| C |  | P |
| D |  | P |

### A

He is mainly in charge of designing the algorithm and coding of numerical calculation. The functions of calculation of each kind of uncertainties are finished by him and the ways of evaluation of the data is designed by him and he implemented the whole part of numerical calculation.

### B

He is mainly in charge of integrating the program with window frameworks. The window interface is designed and coded by him using C#. What’s more, cooperating with Wu Rui, he also took charge of the improvement of the program.

### C

He is mainly in charge of testing the program using some illegal inputs. By taking the position as a user, he gave some feedbacks to improve the program. Many problems were picked up by him. He also made comparations with the outcomes calculated out in traditional ways.

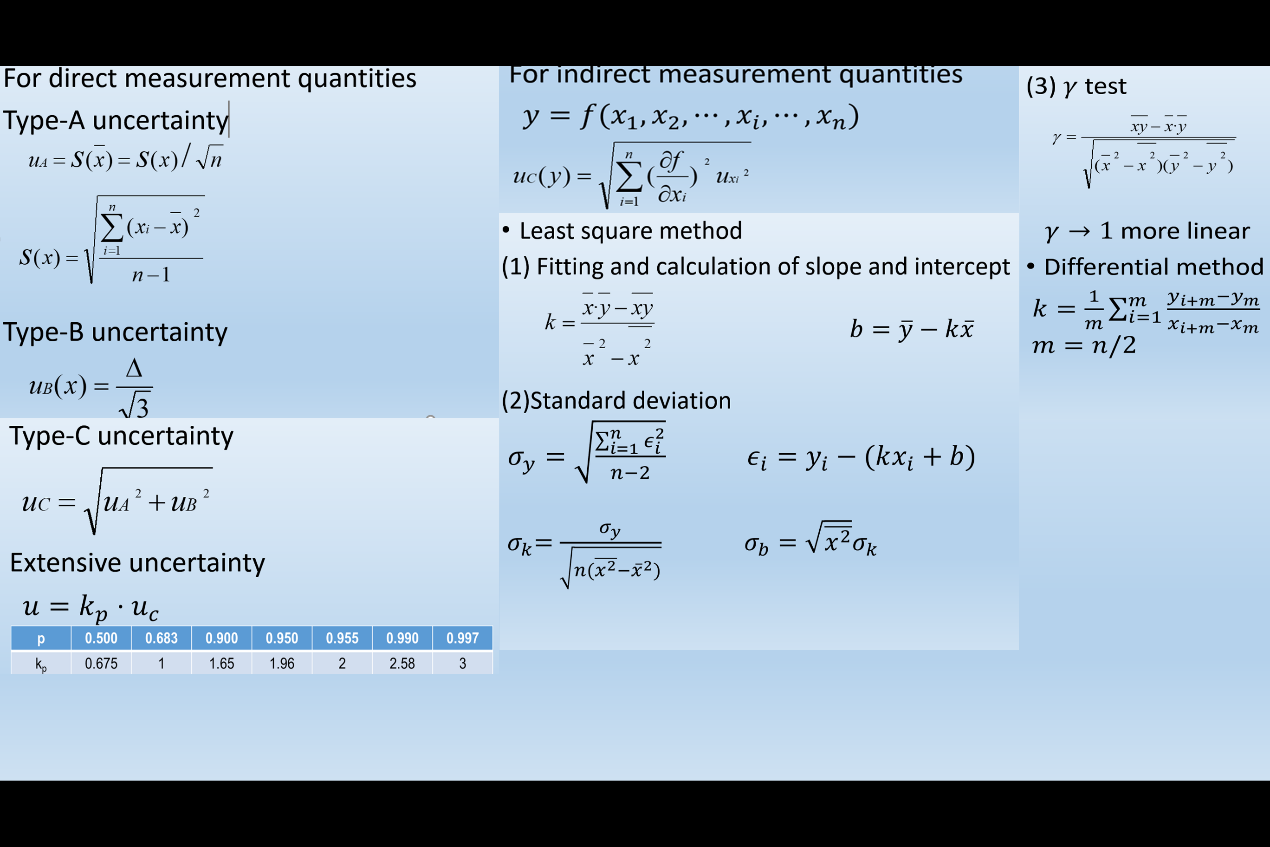
### D

She is mainly in charge of gathering relevant information and helping to solve the problem in the implementation. What’s more, the general design and PPTs are finished by her.

## Analysis

The formula form below shows the specific calculation process of each uncertainty and explains the design of the part of linear data processing. The whole program is actually to finish the calculations in the form and these formulas are actually the core of our program.

In order to finish our goal and functions we designed, we made the analysis and several problems we have to solve were listed as follows:



1. **For uncertainties of direct measurement quantities**:

First of all, we don’t know previously how many uncertainties the users want to calculate, which means we don’t know previously how many times the program has to run.

As for the confidence probability, the functions or the mapping between the confidence probability and the extensive uncertainties, which can be obtained by calculating out the confidence interval, follows the normal distribution. This process is too complicated for computer to calculate.

2. **For uncertainties of indirect measurement quantities**:

Same as before, we don’t know previously what the transfer formulas are, and we don’t know how many variables there are in the formulas.

Even if we let users to input the formulas, since the formulas are various and functions such as “sin”，“tan” and“ln”are included, how can we let the computer to transfer these character strings into arithmetic operations? After all, C is not very suitable for numerical calculation.

What’s more, this processing involves differentiate operations. Since computer can only process discrete values, we have to use infinitesimal values to do the differentiate operations. However, how small the values should be? The quantities in physical experiments are various. For example, 0.001 is a small number in this experiment, maybe it will be a large number in another experiments. How can we balance infinitesimals and the least precision values of instruments?

Besides, due to the wide ranges of experiment quantities, the problem that small number might be “eaten” by a much bigger number should be noticed.

As for the algorithm design, from the formula shown above, in order to calculate the uncertainties for indirect measurement quantities, four types of uncertainties of each variable should be calculated out first. How to save the intermediate variables and uncertainties so that these outcomes can be used in a more efficient way?

3. **For differential method**:

Due to the algorithm, each group of data must be coupled with another group of data, however, what if the group number is an odd number? Users have to delete one group of data if they input the data whose number is an odd number. How to implement this without any influences on other data, especially the intermediate variables.

4. For least square method:

Since outcomes will always be obtained, if the users put a group of number. Evaluations will be needed to describe the data. What aspects should the evaluation part describe?

5.For languages we used on forms

It is quite complicated and difficult to create a form in C language. Every specific details should be figured out by code, like adding a button, editing the size and position of a button. But C# solve these problems in a visible and efficient way. Therefore, using C# to edit forms may be a better choice.

6. For sequence of forms

The program aims to solve several different kinds of data processing problems. Therefore, it must be contained a lot of different forms. We have to consider the logical sequence of each form to lead users to solve their problems to make it more user-friendly.

7.For input and output form

A large quantity of numbers need to be input and processed in the program. It is a problem of how to input a group of related numbers in a form and do some calculations on these numbers. Besides, how to output the outcomes should also be concerned.

8.For quantities of data that input

The most troublesome problem is that it is unknown of how many numbers will users input. If we list lots of textboxes in a form and let users to input only one number in each textbox, we do not know how many textboxes is enough and it is knotty to handle those empty textboxes.

9.For restrictions on pressing keys in a textbox

Besides, If we use textbox to have users input numbers and to display different outcomes, it should be considered that only numbers and decimal point can be input in those textboxes for input and nothing can be input in those textboxes for outputting the results of calculation.

**Solution:**

The solutions to the problem mentioned above are as follows.

1. We used “for” loop to control the times the program will run in.

2. We installed nine the most common transfer formulas, with their corresponding program bodies to calculate the indirect measurement quantities.

3. We use arrays to save every intermediate variables, which could be easily process later.

## Design

**4.1 The solutions to the problem mentioned above** are as follows.

1. We used “for” loop to control the times the program will run in.

2. We installed nine the most common transfer formulas, with their corresponding program bodies to calculate the indirect measurement quantities.

3. We use arrays to save every intermediate variable, which could be easily process later.

4. We decided to use C# to code the window framework.

**4.2 general design**

1. The calculation of uncertainties:

(1) For the data from direct measurement:

a. The users should input how many uncertainties they want to calculate. --- Decide how many times the program will run in and the array length. (The arrays are supposed to save the intermediate variables)

b. Input the data --- Put the data into the array.

c. Calculate the type-A uncertainties.

d. Input the precision values of the instruments. --- Calculate the type-B uncertainties.

e. Using type-A and type-B uncertainties to obtain type-C uncertainties.

f. Choose the confidence probabilities --- obtain the extensive uncertainties.

(2) For the data from indirect measurement.

a. Choose the formula --- Decide the variables number and transfer formula.

b. Input the data --- using the program body in the calculation for direct measurement data to get every uncertainty for every variable.

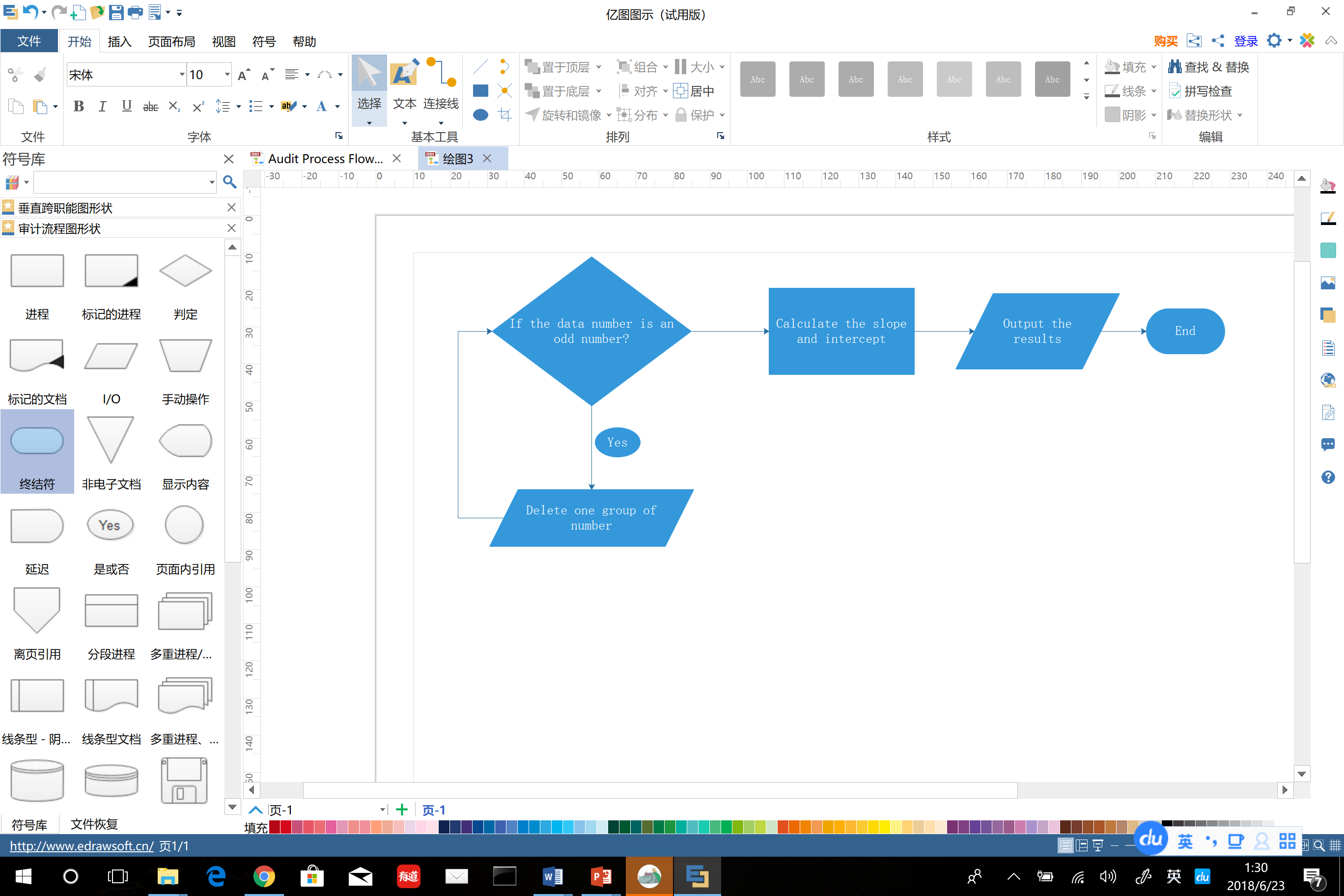
c. Get the results --- using the transfer formula which was decided before.

2. linear data processing:

(1) least square method.

a. Input the data

b. Analyze the data.

 (2) differential method:

3. window framework:

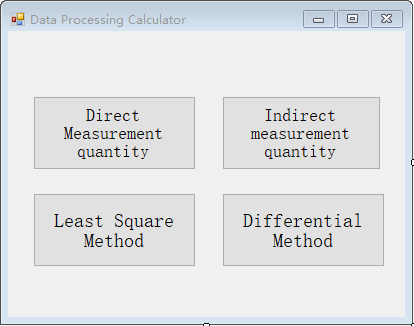
We designed some forms in our program to make it looks simple and clear. Because forms can be edited directly in C# language, we actually designed the forms in C#. It is easy to put the code of calculation part in the form code because syntax of C# resembles C language.

According to introduction part, users may have two general problems:

(1). Process linear data by least square method or differential method.

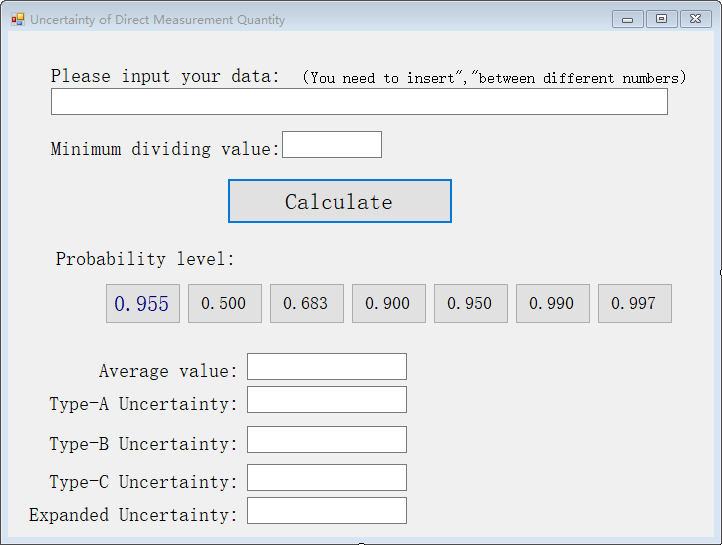
(2). Calculation of uncertainty of particular quantities. Quantities may be measured directly or indirectly.

Users should first choose which kind of process in the following form.



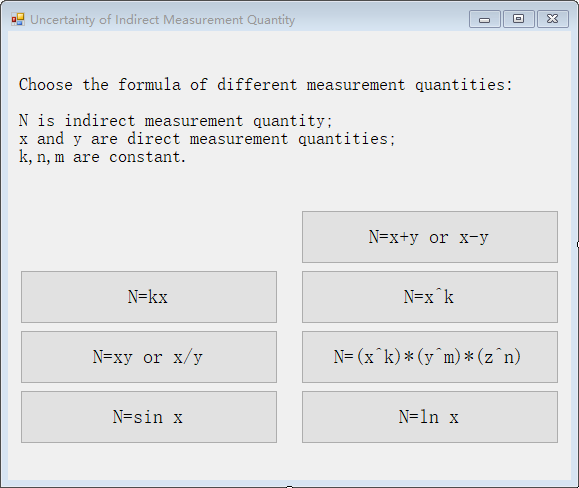
If the measurement quantity is direct, click the corresponding button and

users will see the following form to input the data.

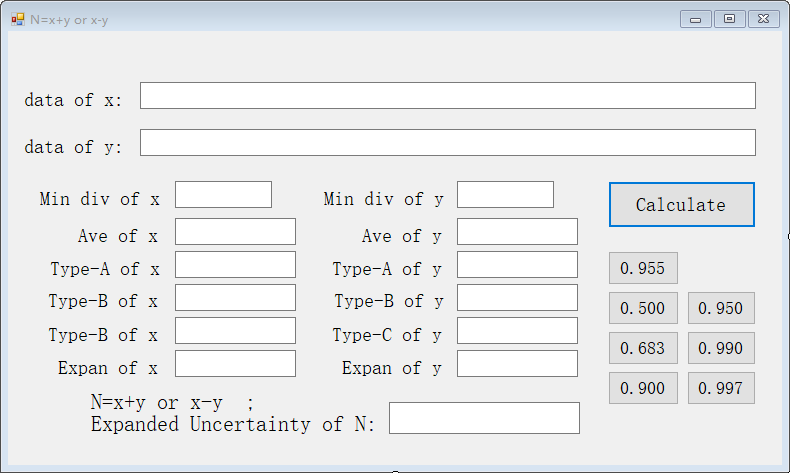


Input the data and minimum dividing value, then click “Calculate” and users will get Average value, type-A,B,C uncertainty. Choose one probability level and the corresponding uncertainty will be shown in the textbox.

If the measurement quantity is indirect, click “In direct measurement quantity”. Then users should choose the formula of different measurement quantities in the following form.

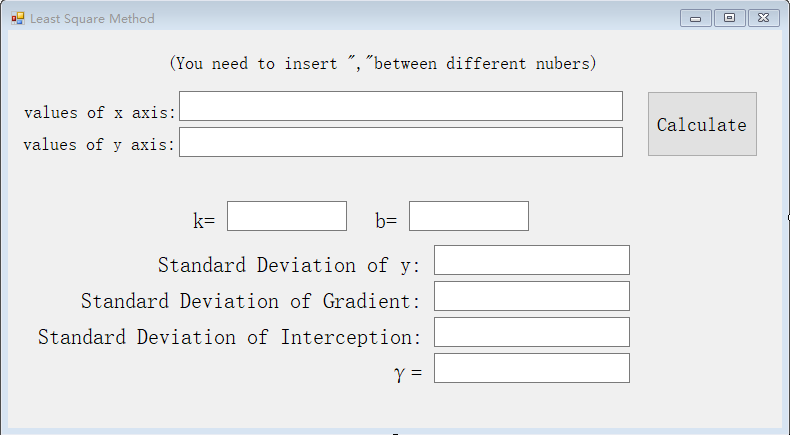


For example, click “N=x+y or x-y”, then input data of x and y in the follow form.



Click “Calculate” and choose probability level and users will get all the uncertainties.

If a group of linear data need to be calculated and processed, choose “Least Square Method” or “Differential Method” in the first figure. For example, if users click “Least Square Method”, then they will get the following form.



In put the value of horizontal ordinate and vertical ordinate, click “Calculate” and users will get the value of gradient k interception b, and so on.

## Implementation

The implementation is based on the design mentioned before.

1. we coded the numerical calculation part of the program.

At the beginning, we coded the part of calculation of uncertainties for one group of data, which is from direct measurement uncertainties.

Then, we use a “for” loop and make the body mentioned above be the main body of this loop. If the users input how many uncertainties they want to calculate, this loop will run in the times the users input accordingly.

Given that the program might run in many times, we change to use array to save every intermediate variable, which is more efficient.

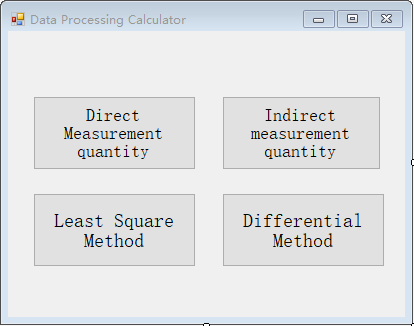
Next, it comes to the part of uncertainties for indirect measurement. We installed previously nine formulas for the indirect measurement quantities, and the transfer formulas were also put into the program. If the users choose one formula, the program will lead the users to input every necessary value and output the results in the end.

2. we coded the part of linear data processing.

The implementation is just followed the design illustrated above.

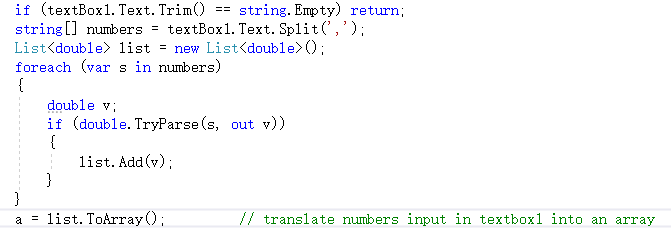
3. We integrated the program and added window framework.

Base on the forms designed, we developed the program.

 In first step, when users click a button to choose a certain processing method in the following form, it will trigger click event of the button, which aims at opening a new form that is needed.

When users input a group of numbers, which are divided by “,” ,in a textbox and the quantities of those numbers are unknown, we used the following code to translate these numbers to an array to make it easy to process later.

There is always a “Calculate” button in processing forms. Almost all the codes about data processing and results outputting are in the click event of “calculate” button. While those buttons of different probability level contain processing of related data and output the value of expanded uncertainty.



## Testing and Debugging

The accomplished project is made of four kinds of functions, including the calculation of one and various uncertainty and the calculation of slope of the fitting line by means of differential method and the least square method.

At first, all the codes are read in principle to testify the exactness of the code in grammar. The spelling mistakes and the wrong usage of coding language are checked and then corrected before inputting data to test.

About the first function, the calculation of uncertainty of one kinds of data, we look for some data of general physical experiment we have calculated by bare hand before. Then we input these data into the square frame according to the tips. It is noted that these data is separated by comma. After pressing the computing button, the average value of these data and their certainty of category A, B, and C have already been output into the corresponding output frame. Then the probability level should be chosen among the given common value. In this way, the final outcome can be output. We get a satisfying outcome of all kinds of uncertainty compared with the previous calculation by hand. The relevant outcome are listed in the following form.

Now, the first function has achieved our plan about one kinds of data. Then we should check what reaction the proje0ct will get when we input some illegal data. When we input the English letters completely, the error tips of “NAN” are displayed in the output form, which is reasonable. It explains that the project can resist the illegal data of pure English letters. When we input the number and English letter at the same time, the normal outcome can still be output without any tips of error. Therefore, we should make some improvements to avoid the input of English letter and relevant illegal data.

In order to solve the problem, we have a correction on the property of all the textbox, including other three function. In the property of key press, we add some codes to prohibit the input of some characters except the number, decimal point and “Backspace”， which are necessary input for data processing. By means of the code, the users can only input pure numbers to carry on normal calculation.

For the second function for the uncertainty of various kinds of data, we have a similar test. We check the expression of every formula on the interface and have a test on them one by one. In the calculation of indirect uncertainty, we input the designed data and compared the outcome with the expected outcome. Similar to the first function, we get a relatively exact calculation of extensive uncertainty of N. What’s more, illegal date except for number is prohibited.

We also have a test on some accidental occasions. If the minimum indexing value are left out carelessly, the users will get the tips to input the value after you press the computing button. In addition, the uncertainty of category A can be only achieved when there is more than one data. So when we input only one data for one kinds of data, the corresponding frame of category A will show “NaN”to remind the users the deficiency of the data they want to input.

For the three kinds of function of the project, the least square method to get the slope of fitting line, we still take advantage of previous calculated data to have a test. After inputting the given data accurately, we can get the slope and nodal increment of the fitting line along with their standard deviations. Besides, users can get a extremely important parameter γ, which can give a good expression to the liner effect of the fitting line. That’s if the absolute value of γ is closer to 1, it indicates that the liner effect is better.

It is noted that the number of the horizontal and vertical coordinates should be same in this function. Therefore, we input the unequal data of x and y to have a test. That’s users will get a tip that the number of x and y should be identical. According to the tip, the users can have a adjustment on data to carry on the function.

For the last part of the project, differential methods to calculate the slope of fitting line, we have a similar test to the the third function. We get an accurate slope of fitting line only when the number of input data is even number. When users input the odd number, they will get the instructions to delete a set of data.

In this way, we have a test on each function and modules with support from given data appearing in the previous general physical experiment.

We have some timely and effective improvement on the interface, formula and relevant codes to achieve the accurate and reasonable outcome at last. The testing forms are as follows.

The first function:

|  |  |  |
| --- | --- | --- |
| Input | Expected result | Real result(Extensive uncertainty) |
| (2.23,2.24,2.25) | 0.114 | 0.113725049037503 |
| (10.65,10.62,10.67) | 0.438 | 0.439140574942666 |

The section function:

|  |  |  |  |
| --- | --- | --- | --- |
| Formula | Input(data &Minimum indexing value) | Expired  result | Output  result |
| N=x+y | (6.51,6.52,6.54)(0.1)  (3.45,3.46,3.48)(0.1) | 0.166 | 0.16519348995027 |
| N=k\*x(k=3) | (6.36,6.38,6.39)(0.1) | 0.3464 | 0.346410163027551 |
| N=In x | (7.25,7.23,7.26)(0.1) | 0.1582 | 0.0159342301300621 |
| N=x/y or N=x\*y | (5.21,5.23,5.26)(0.1)  (6.52,6.53,6.60)(0.1) | 0.2981 | 0.0297909194968379 |

The third function：

|  |  |  |
| --- | --- | --- |
| Input | Expected Result | Real Result( the slope K) |
| (1.2,1.3,1.5)  (6.5,6.4,6.8) | 1.143 | 1.14285714285721 |
| (3.2,3.1,3.4)  (5.4,5.6,5.2) | -1.29 | -1.2857142857145 |

The last function

|  |  |  |
| --- | --- | --- |
| Input | Expected Result | Real Result( the slope K) |
| (1.2,1.3,1.5)  (6.5,6.4,6.8) | -1.001 | -0.999999999999996 |
| (3.2,3.1,3.4)  (5.4,5.6,5.2) | -2.001 | -1.99999999999999 |

## Result & Conclusion

Based on the test and application in some experiments, the goal of exploiting a program was successfully completed. With the help of our program, researchers could readily finish the job of processing the experiment data and put their focuses mainly on the analysis of the data.

As for the problems in our program, it is unsatisfying that the choices of transfer formulas are limited. If the formula is uncommon or complicated, this program may not work out the uncertainties for the indirect measurement quantities. This problem includes calculation of differentiation of any given formulas, which is difficult and unsuitable for C to finish.

All in all, the program’s behavior in common experiments reaches the demand of researchers, which proved, in practice, that our program is successful.