Short notes on error analysis

(1) Introduction

Everyone makes mistakes!

In doing a physics experiment, even though you do not make any mistake, your results are still subject to experimental errors.

Experimental error does not mean "mistake". It actually means **uncertainty**.

To avoid human uncertainty, we usually do an experiment with the help of suitable equipment* (ruler, multimeter, power supply, CRO, etc.). These equipment have limitations or capability & therefore no measurement is exactly correct.

* The terms: **equipment**, **instrument** & **apparatus** have roughly the same meaning.

Experimental error is inevitable!

(2) Sources of experimental errors

1. **Real mistake:** It may be due to malfunction of equipment, wrong design of experimental setup, careless operation of equipment, wrong procedure,..
   We should avoid real mistake by careful planning of experiment and proper operation of equipment.
2. **Instrument error**: It comes from the limitation **(precision, reliability & reproducibility)** of equipment and tools used in experiment.

   (a) Analog equipment:

   Depending on how careful you take the reading, the accuracy of the result can be anywhere from one-half to one-fifth of the smallest division.

   E.g. 1: For the analog scale shown in Fig. 1, the reading is recorded as 2.14 ± 0.05.

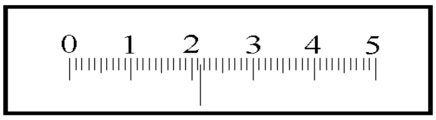


Fig. 1

   E.g. 2: A meter stick or measuring tape is accurate to 0.5 mm.

   (b) Digital equipment:

   An equipment usually comes with a **manual** in which its capability & precision are listed in the "**specifications**".

   Take the digital multimeter (model GDM-8135) as an example. According to its manual, the error of DC voltage reading is

   "±(0.1% of reading + 1 digit)".

   This includes two types of errors (to be explained in **4** & **5** below):

   (i) **Systematic error**: "±0.1% of reading"
   The manufacturer did a generic calibration to guarantee such a precision or accuracy.

   (ii) **Random error**: "+ 1 digit"
   It is due to fluctuation of electronics.
   "1 digit" actually means one **least significant digit or last digit** of display.

   E.g, if the reading on the display is "18.5 mV", then "5" is the least significant digit & the error of data is $(18.5\times 0.1\%)+0.1 \simeq 0.1$.

   If the reading on the display is 118.5 mV then its error is $(118.5\times 0.1\%)+0.1 \simeq 0.2$.

   - **The instrument errors are intrinsic.**
   - **Always use the most sensitive scale of equipment.**

3. **Human error**: It may come from the uncertainty in visual judgment and human response if we need to do so in the measurements.

   E.g. 1: In a free fall experiment, we use a **measuring tape** (拉尺) to measure the height (d) at which a ball is dropped. The tape has a precision of 0.5 mm. But the experimental error δd is larger than 0.5 mm because in the measurement process, we need to judge where the end points are. Furthermore, if we use a stopwatch to measure the falling time (t), our reaction time will significantly increase the error δt.

   E.g. 2: In a geometric optics experiment, it is difficult to judge the position of an image precisely.

   - To avoid human uncertainty, always try to design an experiment using only digital equipment for measurements.
   - **Human uncertainty can result in both systematic error and random error**.

4. **Systematic error**: It may come from
   - persistent biased human observation;
   - uncalibrated equipment;
   - faulty technique (e.g. misalignment in an optics experiment);
   - drifting of lab environment;
   - simplification in analysis; or ….

   E.g. 1: In doing a free fall experiment, your stopwatch may run constantly faster. Then your falling time data have a systematic error. In the analysis, we neglect the air friction. Your falling time data are systematically larger.

(For a good experiment, such a simplification in analysis should be **justified** by an order-of-magnitude estimation.)

E.g. 2: In Exp. 1 of PHY2811, you need to judge when the water drop falls across a mark. You should observe the drop & mark horizontally at the same time. If you persistently look at an angle, you introduce a systematic error. Furthermore, when you hit the timer, the lag due to the gap between your finger and the button or keyboard increases the actual time. Since you press the button twice to get a time difference, the actual experimental error is smaller.

E.g. 3: If a voltmeter does not read zero even when the inputs are shorted, then all measured values will have a constant offset.

It is difficult to detect systematic error, but you can eliminate most of them by proper calibration of equipment. At least you should check their consistency.

5. **Random error** or statistical error:

It comes from
- random fluctuation of equipment reading due to finite precision;
- random drifting of environment (electric power voltage, room temperature, air pressure, etc.); &
- random fluctuation in human judgment or reaction.

Random fluctuation may also be due to stochastic nature of the physical process involved in the experiment (e.g., counting of radiation due to radioactive decay).

Random errors can be determined by repeated measurements. (See **(3)** below.)

6. **Summary**:
- Equipment should be calibrated and checked (at least zero-checked).
- Avoid human error and systematic error.
- Your **measurement error** (the experimental error of measurement) is usually due to instrument error. If human judgment is involved, then the measurement error is mainly due to human uncertainty (usually human error > instrument error.) Always use the larger error as the overall measurement error.
- Random error can be determined & minimized by repeated measurements:

(3) Repeated measurement & distribution of data

Suppose you are doing an experiment to measure a physical quantity x. (E.g. in a free fall experiment, you measure the falling time t for a fixed height.) When repeating the measurement, very likely, you get slightly different results.

Suppose you repeat n times and obtain data:

$$x_1, x_2, x_3 \ldots, x_n$$

The average value is almost always the best result which should be close to the "**true value**". The data can be described in a histogram of frequency distribution:

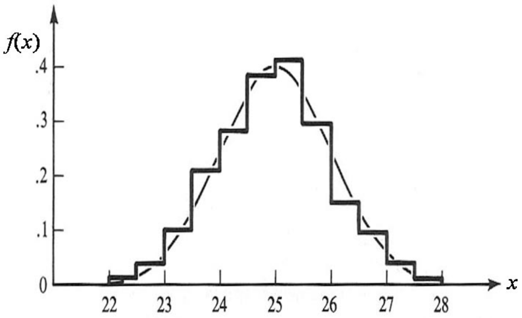


Fig. 2 A typical histogram.

To plot the histogram, the horizontal scale (for x) is divided into small intervals of equal size (called **bins**). The vertical scale is for the (normalized) number of observations with x falling inside each bin. In Appendix A, we show you how to plot histogram by Excel.

A smooth plot can be obtained only for large n.

Very often we can use a **normal distribution** (also called **Gaussian distribution**#) to fit the histogram (as shown in Fig. 2):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\bar{x})^2/2\sigma^2} \qquad \text{Eq. [1]}$$

where $\int_{-\infty}^{\infty} f(x)dx = 1$.

[# For counting radioactive decay, we need to use **Poisson distribution**.]

According to the Sampling Theory,

the mean $\bar{x} = (\sum_{i=1}^{n} x_i)/n$ Eq. [2]

is the best estimate of your measurement of x,

and $\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$

is the best estimate of the standard deviation (also called the true standard deviation) .

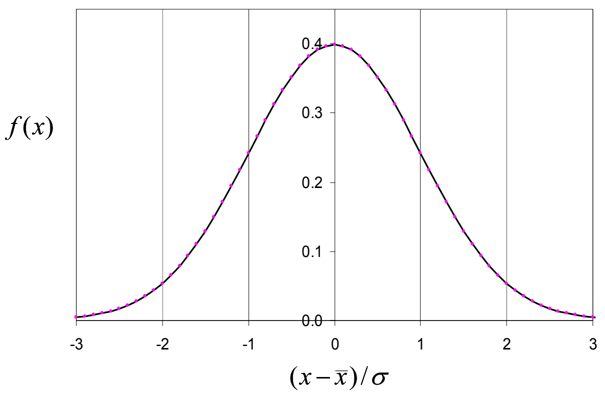


Fig. 3 Normal distribution.

The area under the curve in region $\bar{x}-\sigma < x < \bar{x}+\sigma$ is 0.68.
That means if you repeat one more measurement, you have 68% chance to get a new value of x falling within the interval: $\bar{x}\pm\sigma$.
For this first set of data, the results are labeled as $\bar{x}_a$ & σ_a.
Now repeat measurements of the whole data set, you get similar distribution with results: $\bar{x}_b$ & σ_b.
Keep repeating the measurements of the whole data set & thus obtain:
$\bar{x}_a$ & σ_a, $\bar{x}_b$ & σ_b, $\bar{x}_c$ & σ_c,...
The distribution of $\bar{x}$ is also normal (i.e., Gaussian) with a smaller standard deviation:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

(4) What is experimental error?

After measuring one data set ($x_1, x_2, x_3 \ldots, x_n$), you can present your result as $\bar{x}\pm\delta x$ where δx is called the **standard error**:

$$\delta x = \sigma_{\bar{x}} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i-\bar{x})^2} \qquad \text{Eq. [3]}$$

δx can be determined by Excel (Appendix A).

Meaning of standard error:

Assume that the data set can be described by a normal distribution. Then when you repeat the measurements & obtain another data set, you have 68% chance to get a new value of $\bar{x}$ falling within the interval: $\bar{x}\pm\delta x$.

[Note: Some people prefer using "**probable error**" δx_p defined as $\delta x_p \simeq 0.67\delta x$
Then when you repeat measurements, you have ~50% chance to get a new value of $\bar{x}$ falling within the interval: $\bar{x}\pm\delta x_p$.]

(5) How to present data & error?

Write your result as $\bar{x}\pm\delta x$.
If you repeat measurement only for a few times, then there is no point to calculate random error. In this case, just use the measurement error for the data.

1. **The error δx should be quoted to one (or at most two) significant figure(s) only.**
2. **The datum should have significant figures consistent with its error.**
3. **Round off data to the right number of significant figures.**
   E.g. 1: $3.3 + 0.22 = 3.5$
   E.g. 2: $3.1416\times 2.54^2 = 20.3$

E.g. the CODATA recommended value of elementary charge:

$$e = (1.602176487 \pm 0.000000040)\times 10^{-19}\,\text{C}$$

or in concise form,

$$e = 1.602176487(40)\times 10^{-19}\,\text{C}.$$

(http://physics.nist.gov/cuu/Constants/index.html)
You can also use the **percentage error**

$$e = 1.602176487\times 10^{-19}\,\text{C} \ \pm 2.5\times 10^{-6}\%,$$

but be careful when using this format.

Hint: **always present data in scientific format**.

(6) Is your experiment satisfactory?

Discrepancy: If your experiment is new, then you just present the datum with estimated error. If a theory (a law) exists to predict an expected value x_T, then we consider discrepancy:

$$\Delta = |\bar{x} - x_T|$$

Very likely $\bar{x}\neq x_T$.

1. If $\Delta < \delta x$, we say *we have verified the theory within the experimental error.* To further test the theory, we need to improve the experiment for smaller δx.
2. If $\Delta > \delta x$, then the theory may be wrong. You need to check both theory & experiment carefully & look for an explanation of the discrepancy. New law or theory of physics may be discovered when one observes such discrepancy.

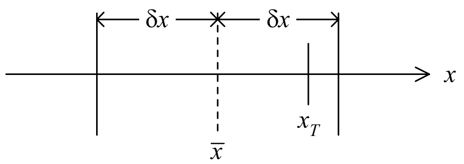


Fig. 4

Compare result with accepted value:

In the first-year labs, you will verify various laws or repeat some well-known experiments. In many cases, what you measure has an **accepted value** x_T (available in textbooks or webpage like http://physics.nist.gov/cuu/Constants/index.html).

1. If $\Delta < \delta x$ (Fig. 4), you can say
   - you have just repeated the experiment with satisfactory result, or
   - your experimental result agrees with the accepted value.

   Present Δ / x_T (in %) in your conclusion.

   In the first-year lab, your result is satisfactory if the discrepancy is < 5%.
2. If $\Delta > \delta x$ (Fig. 5), then the discrepancy is significant. This indicates that you made a mistake or there exist some systematic errors in equipment.

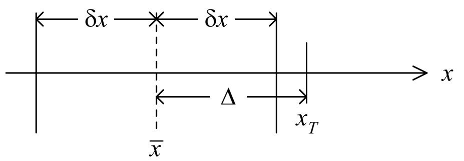


Fig. 5

(7) ***Accuracy vs precision***:

Accuracy: It measures how close the experimental result comes to the "true" value.
Precision: It measures how exactly the experimental result is determined without reference to any "true" value.
Compare **Fig. 4** and **Fig. 6**:

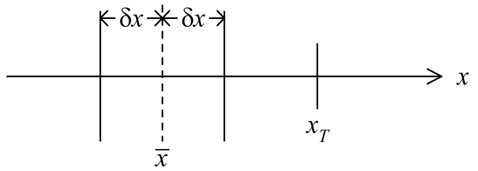


Fig. 6

Result in **Fig. 6** is more precise than that in Fig**. 4***,* probably due to better equipment (but not properly calibrated). Result in **Fig. 4** is more accurate than that in **Fig. 6**.

Good result should be both accurate and precise (i.e., $\Delta < \delta x \ll \bar{x}$).

(8) Handling more complicated experiment

Very often we need to check a theory by measuring a quantity y as a function of x & verify an expression:

$$y = f(x).$$

It is easier to handle & visualize a linear plot. So usually we try to convert the expression into a linear equation & plot a linear graph.

Least square fit to a linear graph:

Given: n data points: (x_i, y_i), $i = 1, 2, \ldots n$.
If you try to verify the linear equation:

$$y = ax + b,$$

then a and b can be obtained by minimizing the mean-square deviation

$$S = \frac{1}{n} \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

where $y_i - ax_i - b$ is the **deviation** of the point (x_i, y_i) from the straight line.
In Appendix B, we shows you how to get the fitting done by Excel and find the

slope $a \pm \delta a$ and intercept $b \pm \delta b$.

The formulae used in Excel (see also Ref. 1) were derived with the following assumptions:

1. The x-values are assumed to be exact or with negligible uncertainty $(\delta x \approx 0)$.
2. The y-values are subject to random errors only.
3. The y-values measured for a fixed x follow a normal distribution.
4. δy is the same for all x-values.

- You can find the derivation in any reference book listed in **(13)** & you will learn it in the 2nd-year lab.
- We show you here only the "**unweighted**" linear least square fit. If you repeat measurements of y many times for each x & obtained δy_i, then you better use the weighted least square analysis. (See Ref. 1 for detail.)

(9) ***Propagation of error***

Very often you need to determine a physical quantity indirectly.
E.g., in a free fall experiment, you need to

verify $d = \frac{1}{2} g t^2$, Eq. [4]

where d is the falling height & t is the falling time.

In experiment, you measure d & t and estimate their errors: $d \pm \delta d$ and $t \pm \delta t$.

Then you can get g by $g = \frac{2d}{t^2}$.

The error in g is due to error in d and error in t. We can determine δg in terms of δd and δt.

Suppose you need to calculate a physical quantity k by a given expression

$$k = f(p_1, p_2, ...).$$

In the experiment, you can measure the independent (uncorrelated) quantities $p_1, p_2, ...$ & estimate their errors:

$$p_1 \pm \delta p_1, p_2 \pm \delta p_2, ...$$

Assume all independent quantities are subject to random error and normal distributions. Then by Sampling Theory (see reference 1 for derivation), the (standard) error of k is given by

$$\delta k = \sqrt{\sum_{i=1}^{m} \left(\frac{\partial f}{\partial p_i}\right)^2 (\delta p_i)^2} \,. \qquad \text{Eq. [5]}$$

This tells you how the errors δp_i affect δk.

Special cases:

1. If $k = \prod_{i=1}^{m} (p_i)^{\ell_i}$ (with constant ℓ_i),

then $$\frac{\delta k}{k} = \sqrt{\sum_{i=1}^{m} (\ell_i)^2 \left(\frac{\delta p_i}{p_i}\right)^2} \,. \qquad \text{Eq. [6]}$$

E.g., in a free fall experiment, $g = \dfrac{2d}{t^2}$

Then the error of g is given by

$$\frac{\delta g}{g} = \sqrt{\left(\frac{\delta d}{d}\right)^2 + 2^2 \left(\frac{\delta t}{t}\right)^2} \,.$$

2. If $k = \sum_{i=1}^{m} c_i p_i$,

where c_i are constant coefficients,

then $$\delta k = \sqrt{\sum_{i=1}^{m} (c_i \delta p_i)^2} \,. \qquad \text{Eq. [7]}$$

- **Based on these formulae, we know that it is sufficient to consider only those parameters that have significant errors.**
- Try to identify major sources of errors for each experiment.

(10) *Accuracy of our equipment*

- Each equipment you use in the Teaching Labs was calibrated by the manufacturer and its instrument error is given in the manual that comes with the equipment. Better calibration can be done only if appropriate standards are available (e.g. you need a standard voltage source to calibrate a voltmeter).
- Refer to the **Equipment Notes in Lab** for the instrument error of our equipment.

(11) Repeated measurements?

Now we know that the random error can be determined & minimized by repeated measurements. But we have to compromise between data precision and time. Consider the free fall experiment again. Suppose you have time to make 20 measurements. How do you plan the experiment to obtain g?

Option 1: Fix d & measure t 20 times. Use the average falling time $\bar{t}$ to calculate g by Eq. [4].

Option 2: For each d, measure t & repeat once. Use average value $\bar{t}$ to calculate g by Eq. [4]. Repeat measurements for different d and use $\bar{g}$ as the final answer.

Option 3: Make measurements like option 2 but plot d against $\bar{t}^2$ & determine g from the slope of the straight line. Here the data are (d_i, t_{i1}, t_{i2}), i=1, 2,..,10. $\bar{t}_i = \frac{1}{2}(t_{i1} + t_{i2})$. For simplicity, take $\delta \bar{t}_i \approx \frac{1}{2}|t_{i1} - t_{i2}|$.

Option 3 is usually the best choice because
- the plot can help us to avoid systematic error (in d for this free fall experiment);
- you can verify the theory (Eq. [4]);
- it is easy to identify wrong data if they are far from the straight line. (But be careful in rejection of data.)

You can now use Eq. [6] to calculate the error of g:

$$\frac{\delta g}{g} = \frac{\delta(\text{slope})}{\text{slope}},$$

where slope & $\delta(\text{slope})$ are obtained by curve fitting (Appendix B).

(12) *Error bars*

Usually we illustrate the data errors graphically using "error bars". See **Appendix C** for an example to inset error bars in your Excel plot.

(13) *References* (all reserved at UL)

(1) **L. Kirkup, Data Analysis with Excel, (Q180.55S7K57 2002)** (highly recommended)

(2) John R. Taylor, "An introduction to error analysis: the study of uncertainties in physical measurements" (QA275.T38 1982)

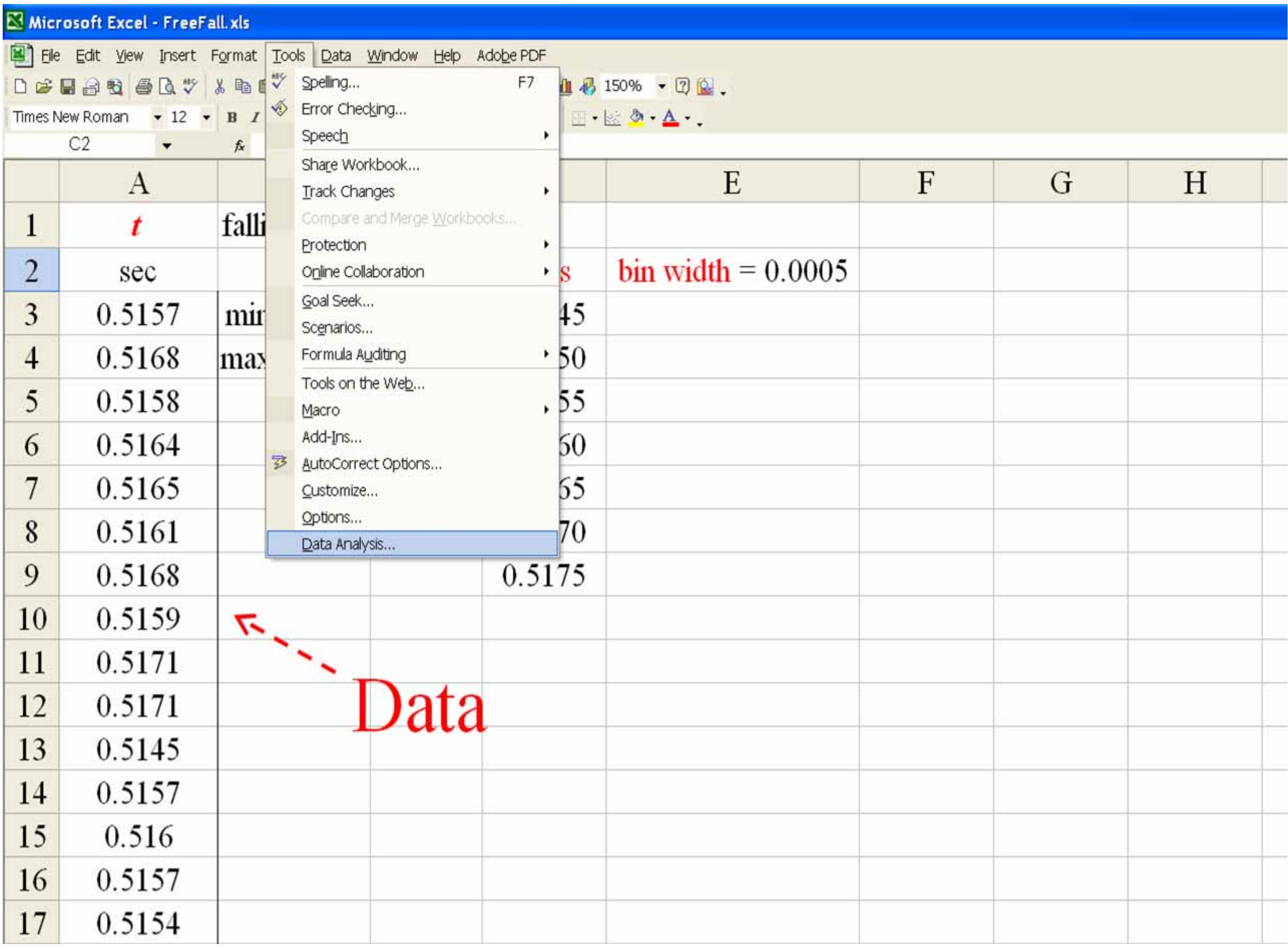
(3) L. Lyons, "A practical guide to Data analysis for physical science students" (QC33L9 1991)

(4) D.C. Baird, "Experimentation" (QC39B17 1995)

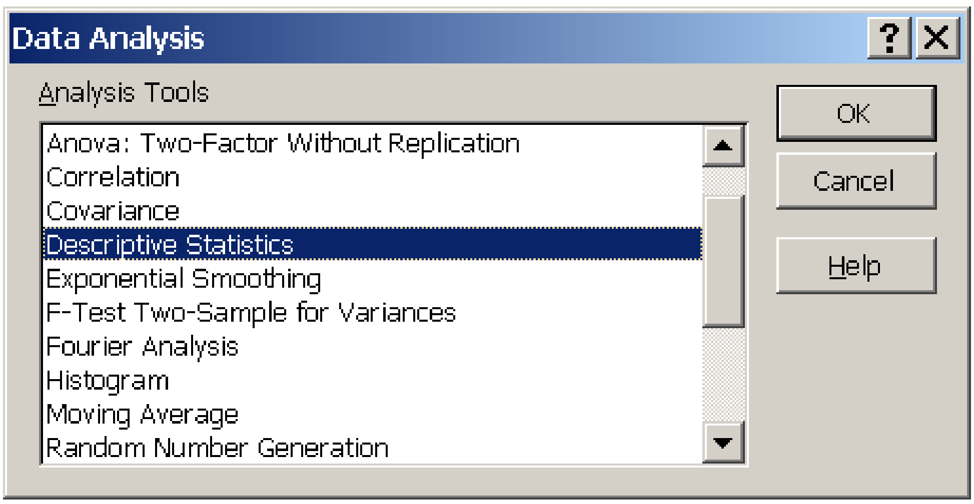
(5) F.C. Chen's notes, "Principles of data analysis" (available in Course Web Page)

Appendix A:
Error analysis & histogram of data (x_i, i = 1, 2,…, n) by Excel

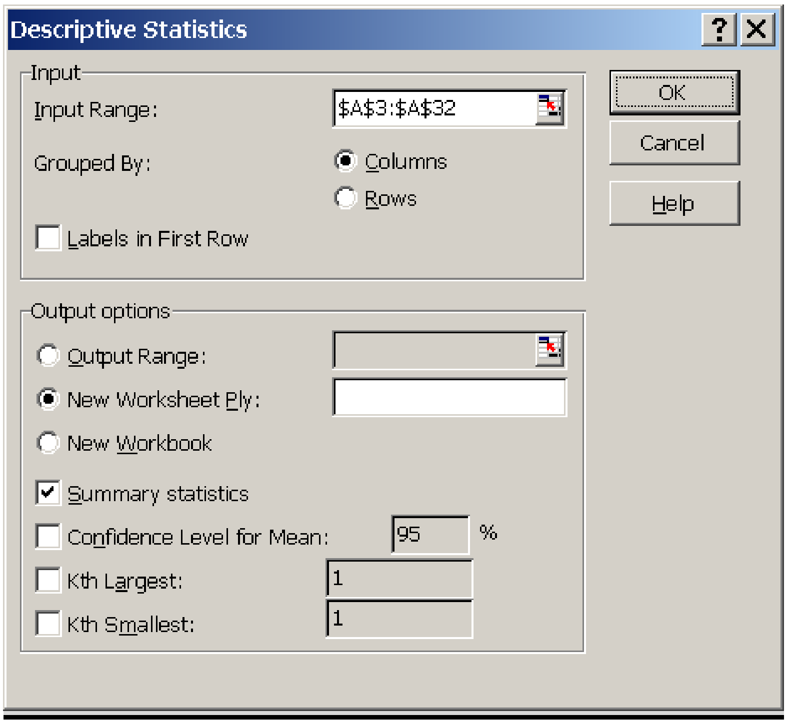
(1) The data file (FreeFall.xls) is available for download in the course page.
(2) We use **Free fall Experiment** as an example. The falling height (d) was 1.3 m.
(3) Input data (the falling time t) in one column, e.g. cell A3 to A32 in **Column A**.
(4) To get standard error δt, select **Tools: Data Analysis** (工具: 資料分析).



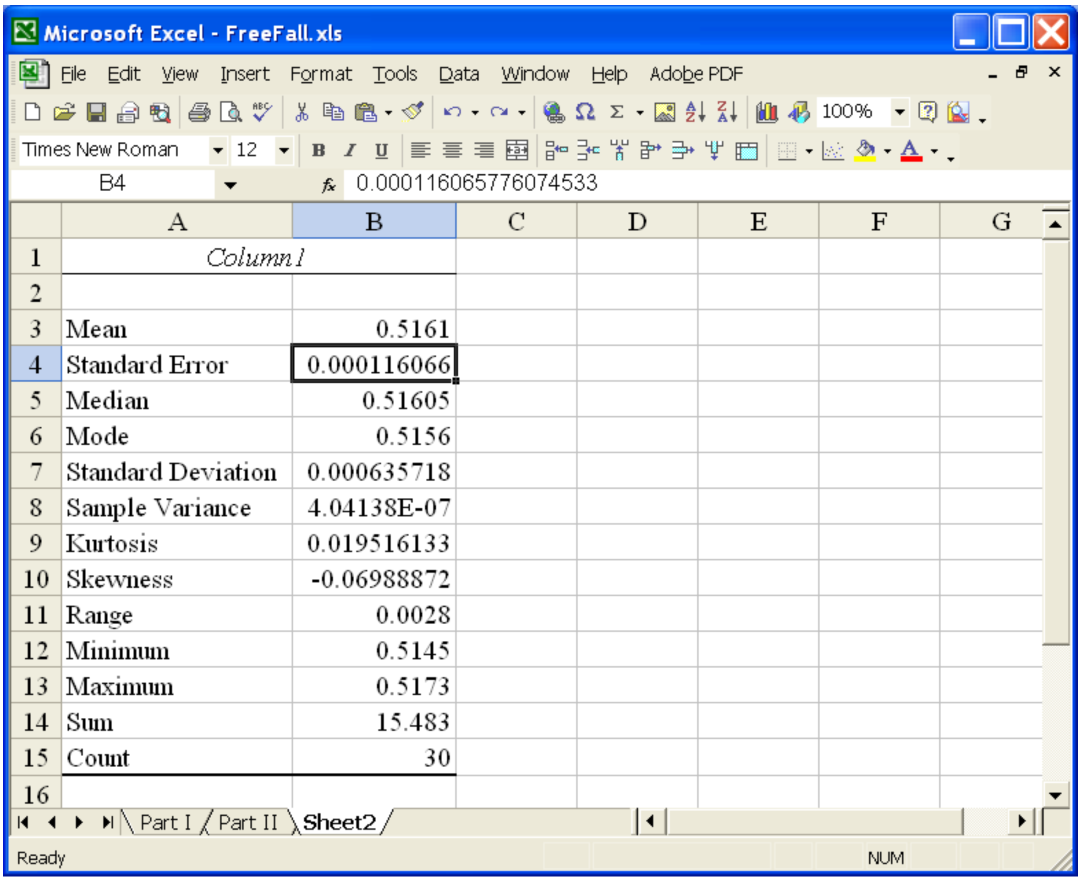
(5) In the **Data Analysis** dialog box, select **Descriptive Statistics** (敘述統計) and click **OK**.



(6) In the **Descriptive Statistics** dialog box, Fill **Input Range** with your mouse.
Select **Summary statistics (摘要統計)**
and then click **OK**.



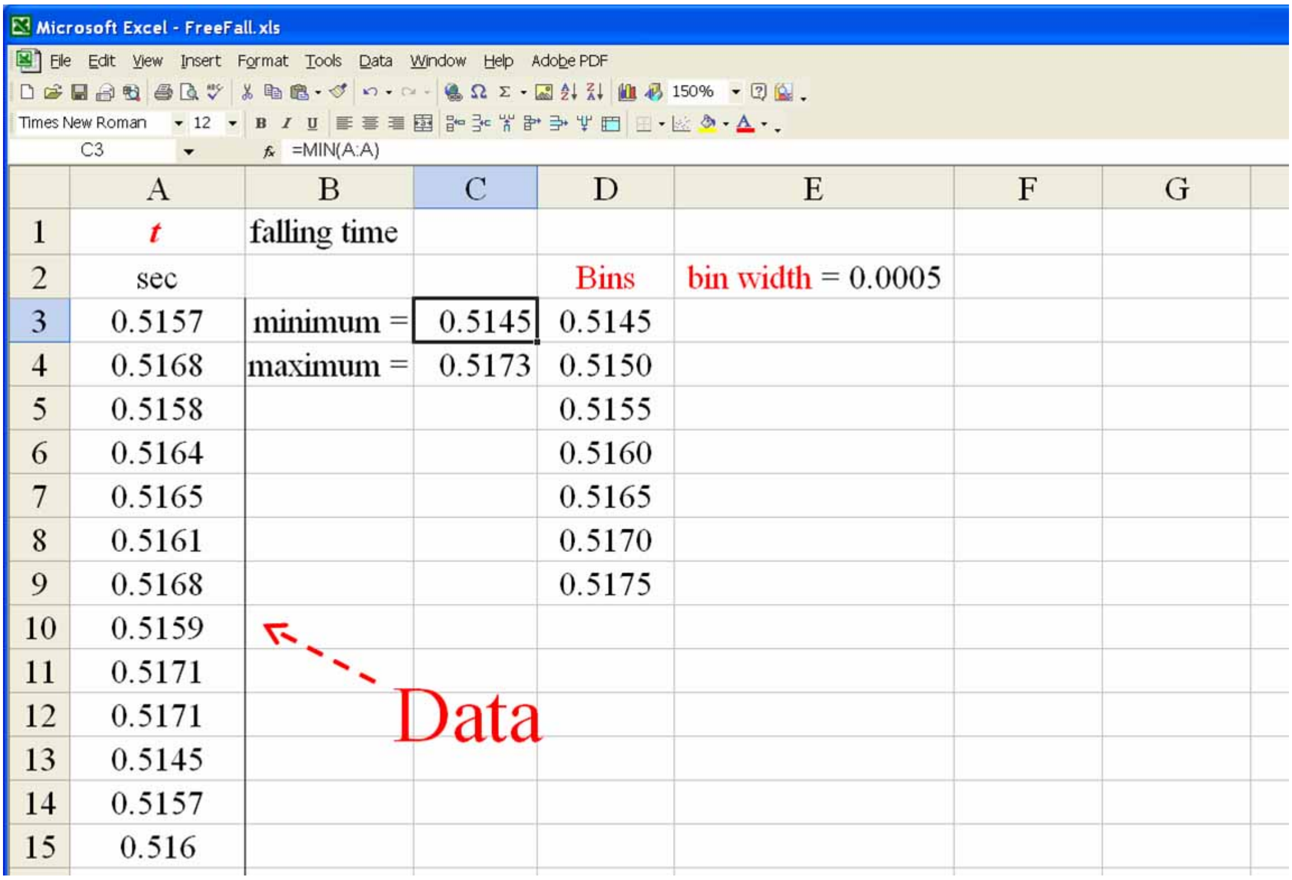
(7) A new worksheet will pop out and show the mean in cell b3 and δt in cell B4.
In this example, $\delta t = 0.0001$. It is just equal to the instrument error of the Smart Timer.
Skip other details.



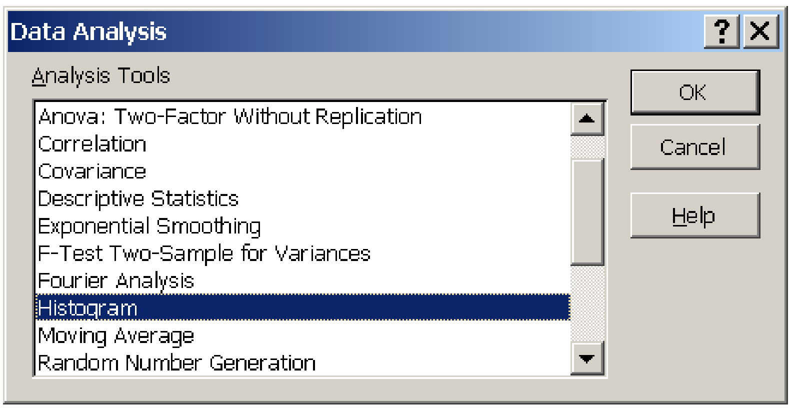
(7) Plot the distribution of data in a histogram:

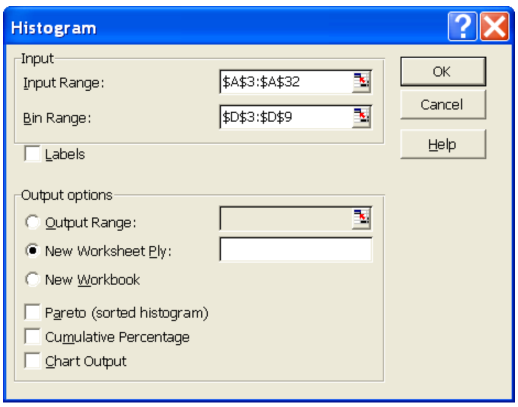
Step 1: Find minimum data by typing =min(a:a) in cell C3.
Find maximum data by typing =max(a:a) in cell C4.

Step 2: Take cell width = 0.0005 s
and declare the bins for the histogram on column D.
The first bin 0.5145 is for data ranging from 0.5141 to 0.5145.
The second bin 0.5150 is for data ranging from 0.5146 to 0.5150.
The third bin 0.5155 is …

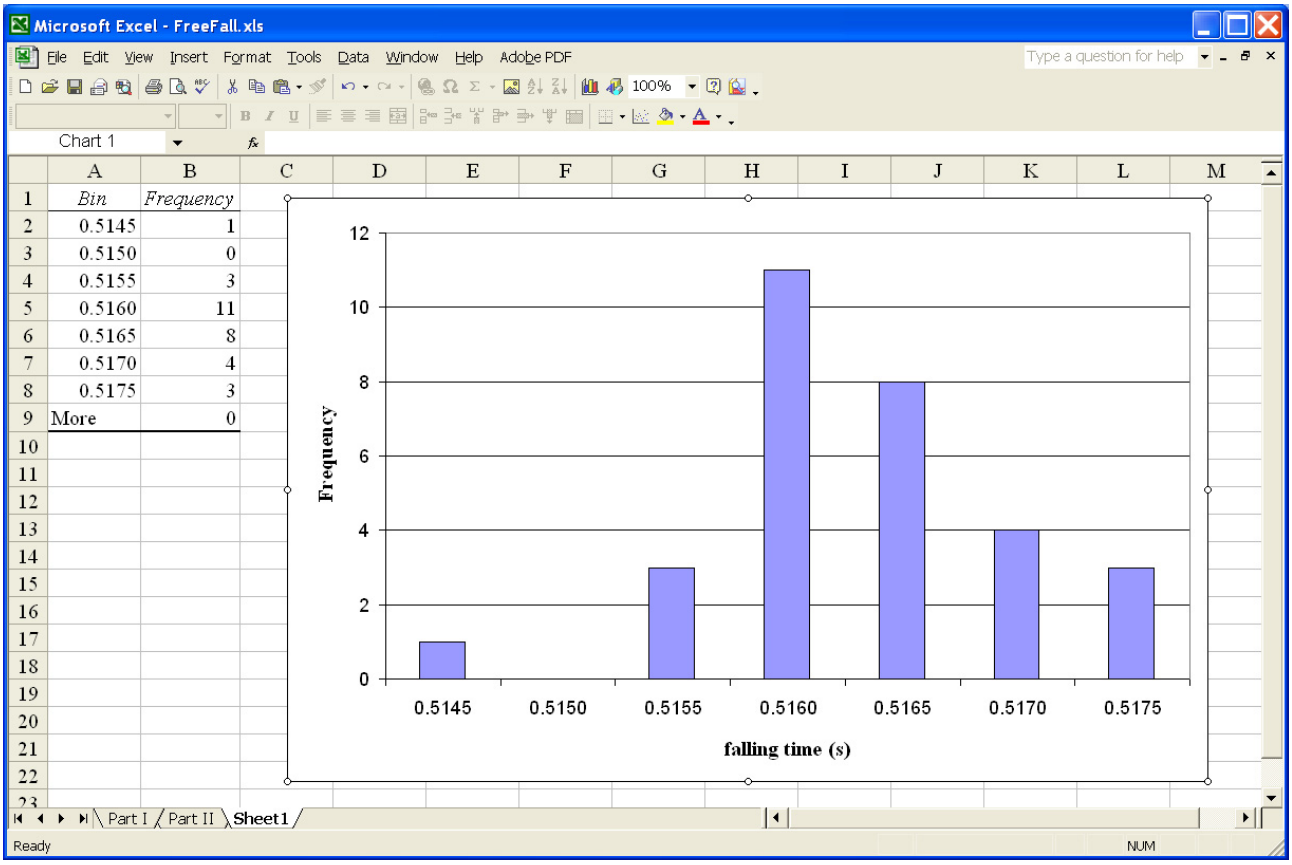


Step 3: Select **Tools: Data Analysis** (工具: 資料分析) and **Histogram** (直方圖)
Fill **Input Range** & **Bin Range** with your mouse. Select **Chart Output**.

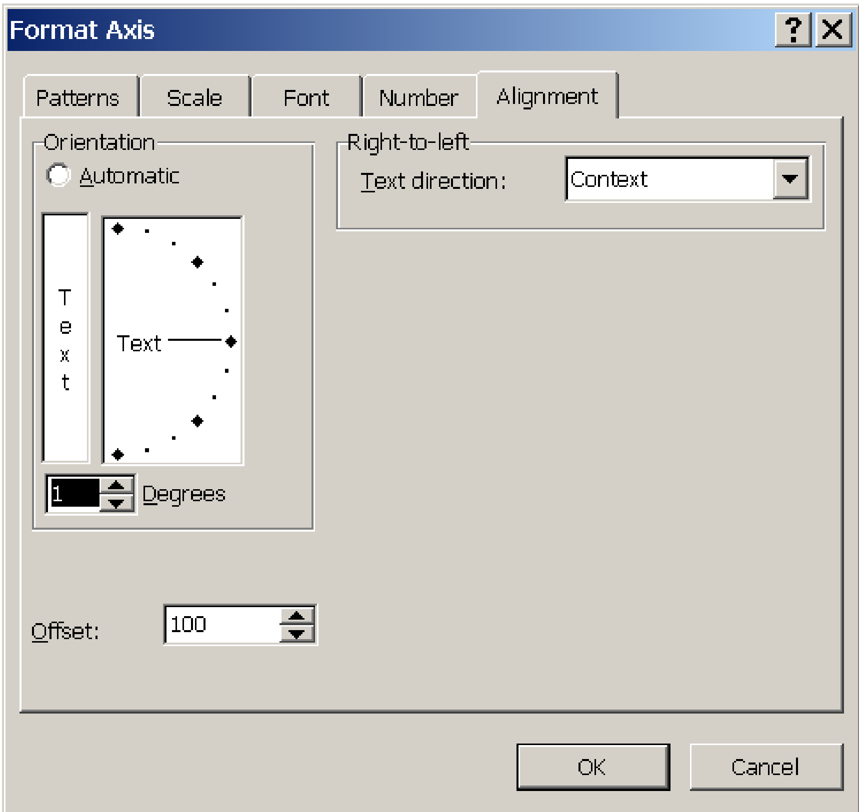




Step 4: Then a histogram is popped out in a new worksheet.



Use your mouse to point at any number of the x-axis and right-click mouse.
Select **Format Axis…**
Select **Font** size or adjust **Alignment** to get a proper scale of the x-axis.



Appendix B: Linear least-square fit by Excel

Initial setup

(1) Input data in two columns, e.g. **Column A** for x and **Column B** for y.

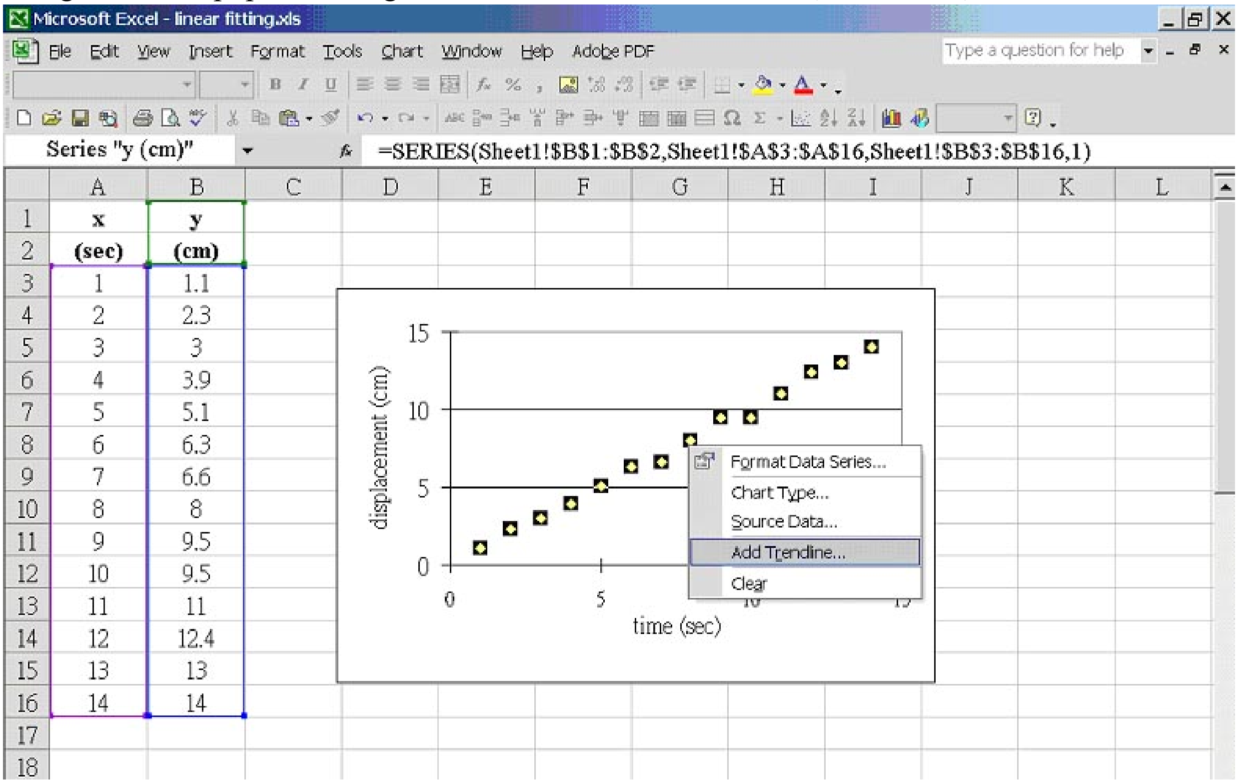
(2) Plot y vs x:

1. Highlight **Column A and Column B**.
2. Click the **chart (圖表精靈)** icon once.
3. Click **XY Scatter Plot (XY 散佈圖)**.
4. Click **Next (下一步)** and fill **Chart Title** and **Axial Titles**.
5. Click **Finish (完成)**.

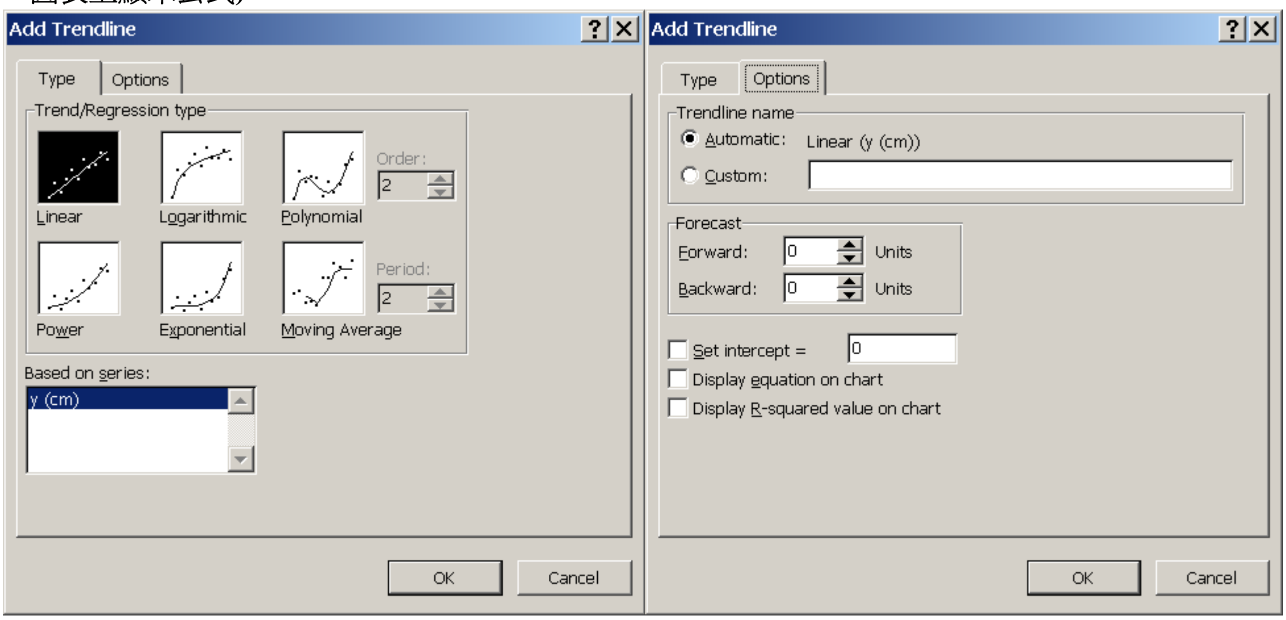
Linear least-square fit

(1) Point at any data point on the graph and click to highlight all data.

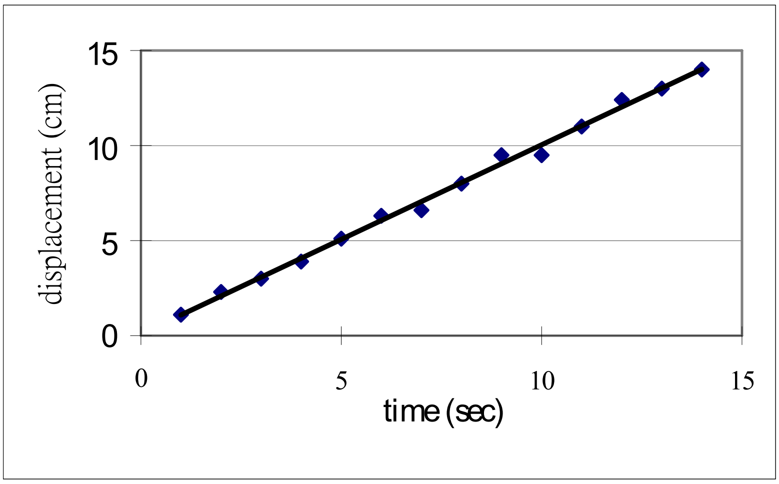
(2) Then right-click to pop out dialog box and select **Add Trendline (加上趨勢線)**.



(3) Select **Type: Linear (類型: 線性)**, and in **Options,** <u>do not select</u> **Display equation on chart (選項: 圖表上顯示公式)** and click **OK**:

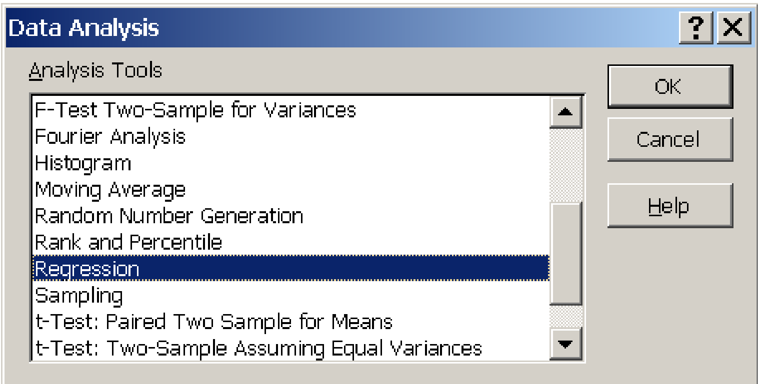


Then we obtain:

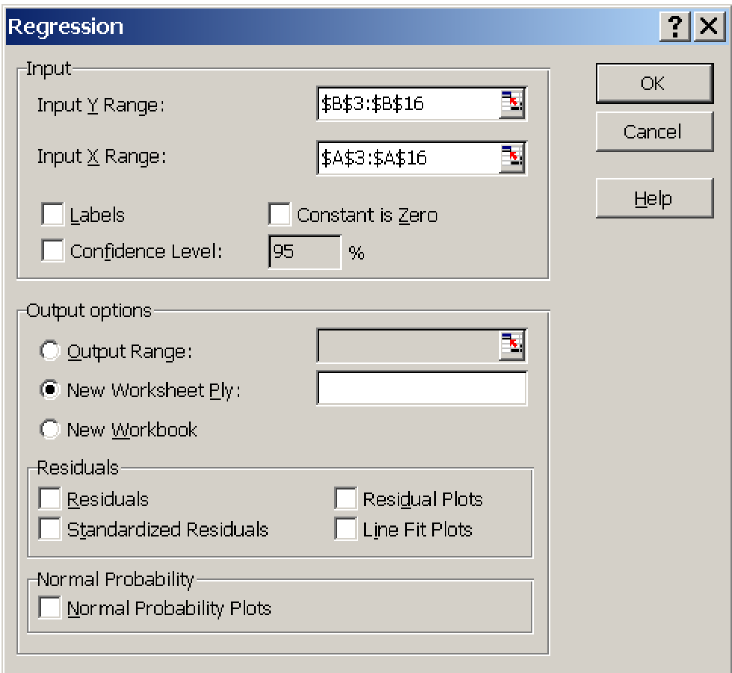


(4) To get $a \pm \delta a$ and $b \pm \delta b$, select **Tools: Data Analysis (工具: 資料分析)**.

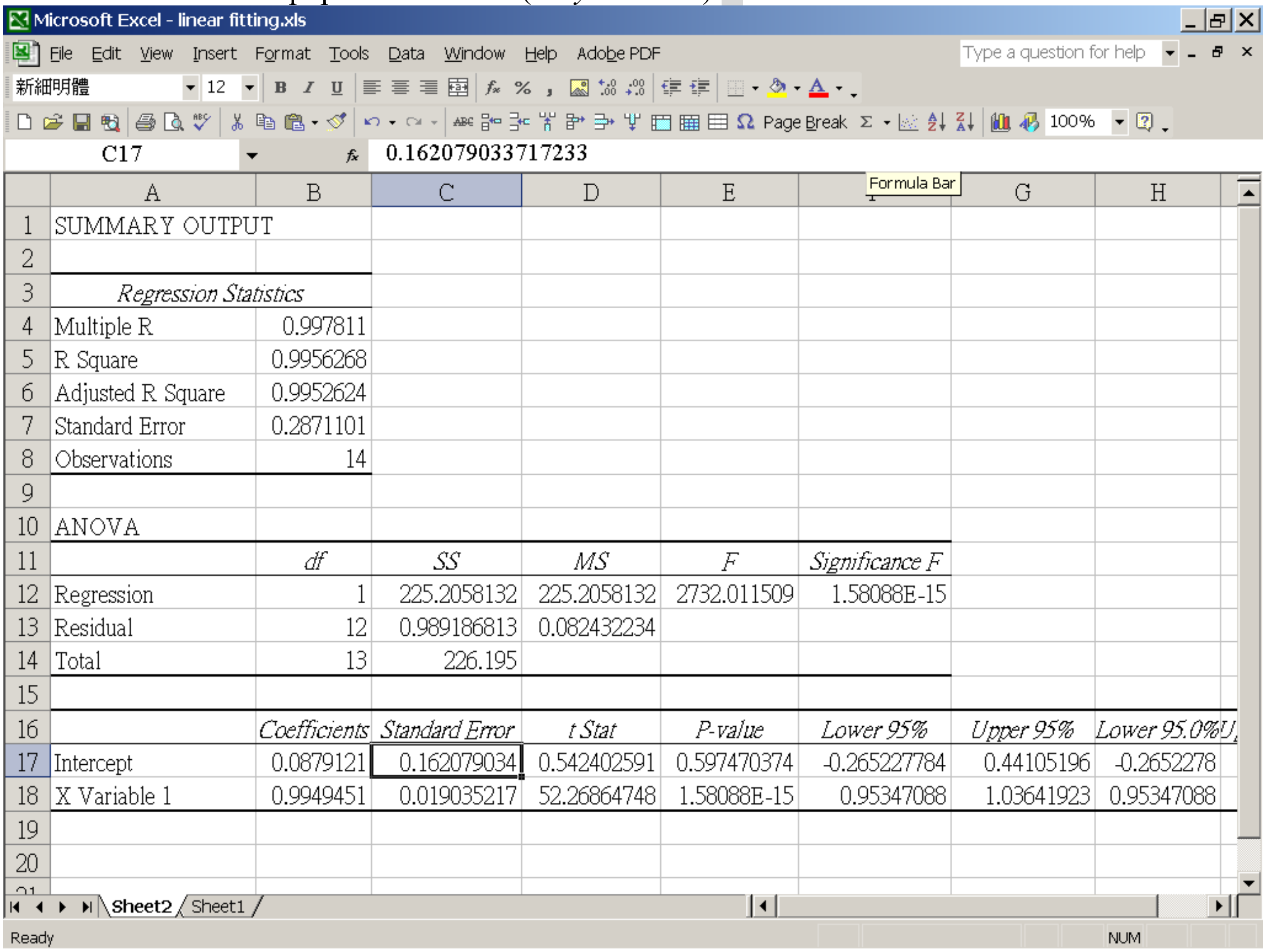
(5) In the **Data Analysis** dialog box,
select **Regression (迴歸)** and click **OK**.



(6) In the **Regression** dialog box, **Input Y range and Input X range** with your mouse.
Then click **OK**.



(7) A new worksheet will pop out and show (for $y = ax + b$):



SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.997811
R Square	0.9956268
Adjusted R Square	0.9952624
Standard Error	0.2871101
Observations	14

ANOVA

	df	SS	MS	F	Significance F
Regression	1	225.2058132	225.2058132	2732.011509	1.58088E-15
Residual	12	0.989186813	0.082432234		
Total	13	226.195			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%
Intercept	0.0879121	0.162079034	0.542402591	0.597470374	-0.265227784	0.44105196	-0.2652278
X Variable 1	0.9949451	0.019035217	52.26864748	1.58088E-15	0.95347088	1.03641923	0.95347088

The results:

1. slope a in cell B18,
2. intercept b in cell B17,
3. standard error δa in cell C18,
4. standard error δb in cell C17.

Skip other details. (See p. 374 of Ref. 1 for explanation.)
(ANOVA = analysis of variance; see p. 354 of Ref. 1))

Appendix C: Error Bars

Now we know how to estimate errors in experiments.
Usually we illustrate the data errors graphically using **"error bars".**
Let's take Exp. 2 of PHY2811 as an example.
In part II, we plot delay time (t) against distance (d):

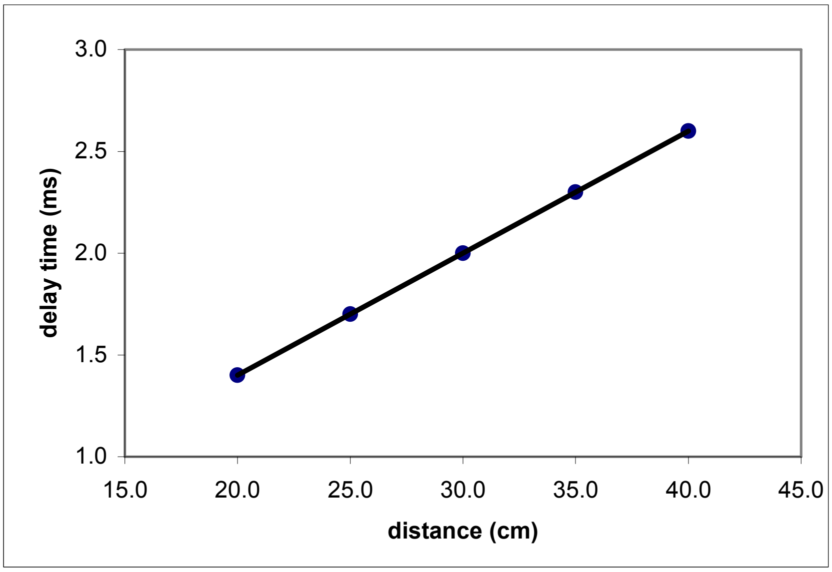


Fig.1 Calibration in ultrasonic experiment

In general, different data have different errors.
For PHY2811/2822 lab, in order to save time, we just consider the error of one data point.
Take the first data point:
$d = 20.0 \pm 0.5$ cm
& $t = 1.4 \pm 0.1$ ms
We can insert error bar in Excel graph.
First input errors δd & δt in Excel worksheet (Fig. 2)

	A	B	C	D	E	F	G
1	PHY2811 Exp. 2						
2							
3	***d***	***t***	***δd***	***δt***			
4	(cm)	(msec)	(cm)	(msec)			
5	20.0	1.4	0.5	0.1			
6	25.0	1.7					
7	30.0	2.0					
8	35.0	2.3					
9	40.0	2.6					
10							
11							
12							
13							

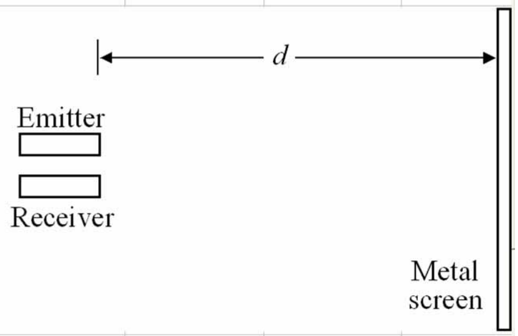


Fig. 2 Data sheet

Remove the trendline.

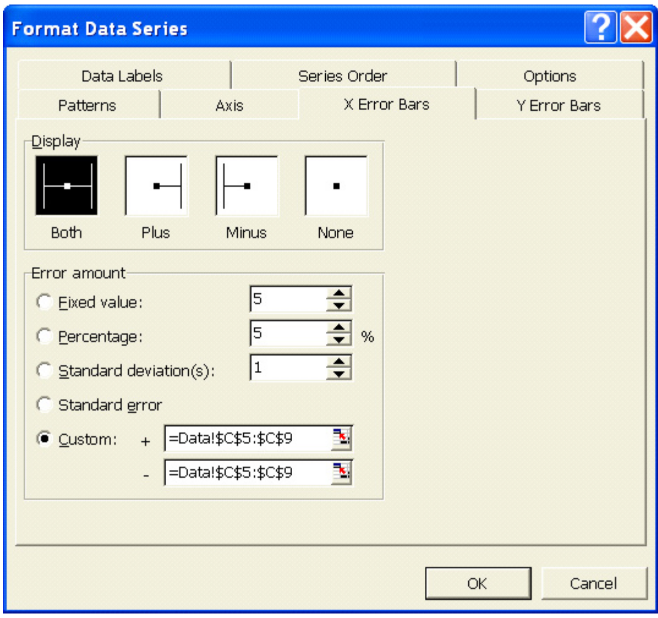
Use your mouse to point at any data point on the graph (Fig. 1) and double-click left button (or click right button & select "**Format Data Series**").

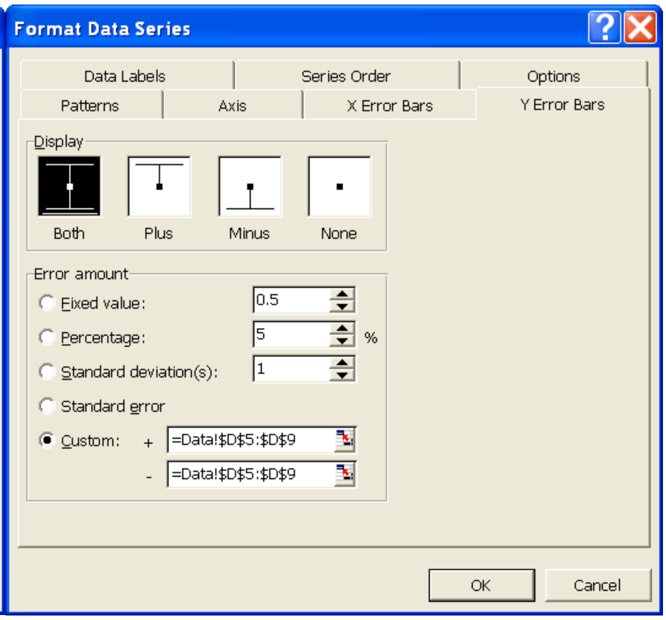
In the pop-out dialog box, select "**X Error Bars**" manual.

Select "**Both**" in "**Display**" and "**Custom**" in "**Error amount**".

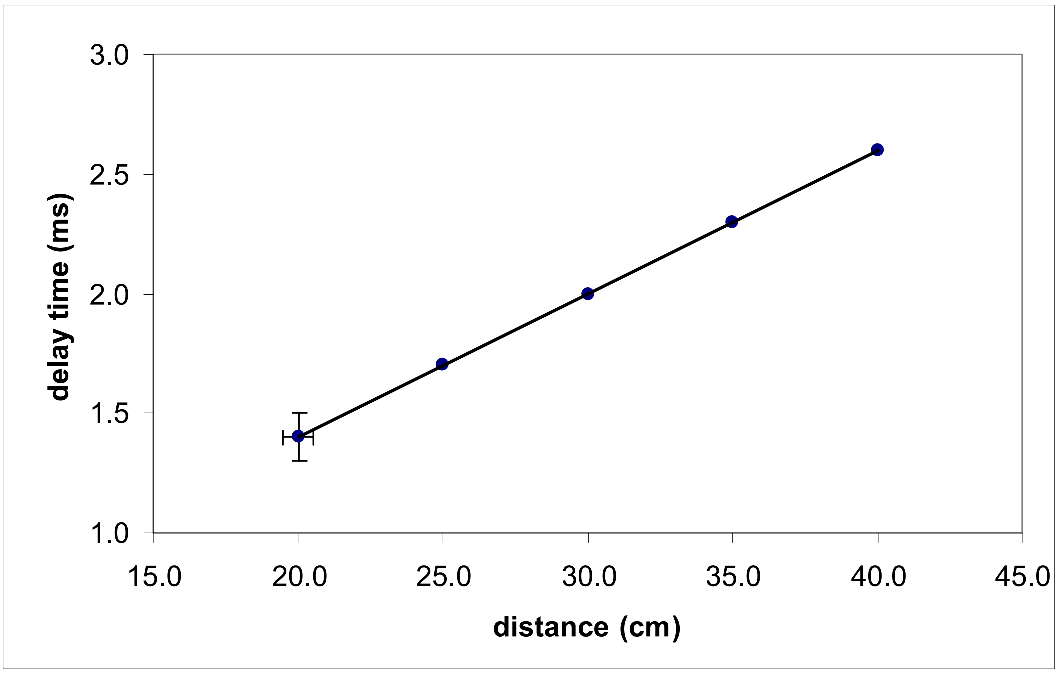
Enter the same error data for "+" and "–" boxes.

Do the same for "**Y Error Bars**":





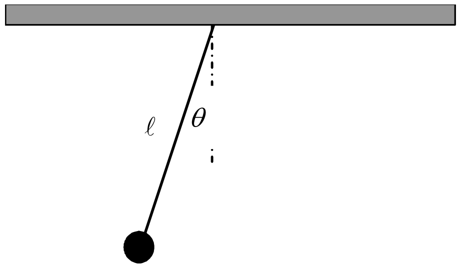
Click OK & add the trendline.   Then we have the final plot with error bars:



Important note:

1. All data points should have their error bars in the graph. To minimize our work load we require error bar of only one data point in each plot.
2. In most linear least square fit ($y = ax + b$) including Excel, it is assumed that $\delta x \approx 0$ or negligible. (See p. 4). So we usually need vertical error bar δy .

Appendix D: Multiple least square fit for $y = a_0 + a_1x_1 + a_2x_2 + ..$



We consider an experiment on simple pendulum with large amplitude. The period of the pendulum can be approximated by a series, for which the first three terms are

$$T = 2\pi\sqrt{\frac{\ell}{g}}\left(1 + \frac{1}{2^2}\sin^2\frac{\theta_0}{2} + \frac{3^2}{2^2 4^2}\sin^4\frac{\theta_0}{2} + \cdots\right) \qquad [1]$$

where θ_0 is the amplitude of the oscillation.

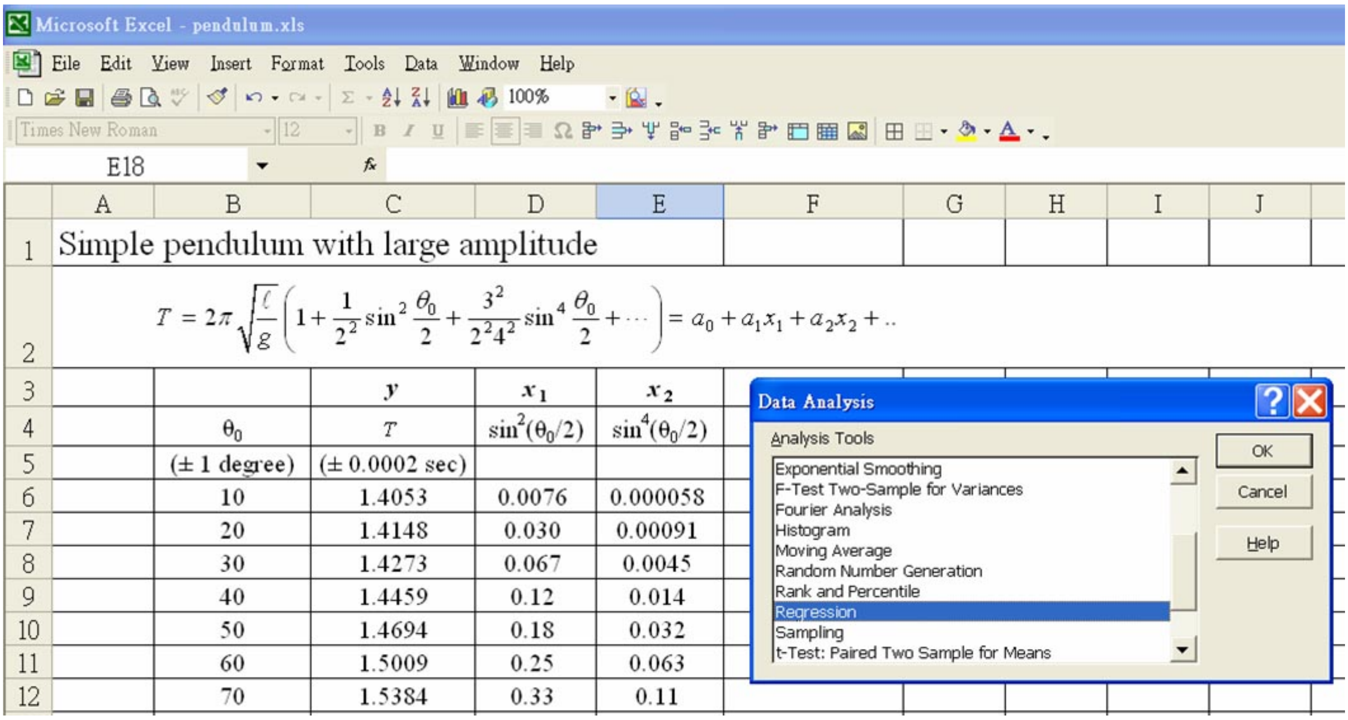
[References:
(1) D. Kleppner & R.J. Kolenkow, An introduction to mechanics (1973), p. 276.
(2) D. Halliday, R. Resnick and K. S. Krane, ***Physics***, Vol. 1, 5th ed., 2002, Chap. 17, 381-382.
(3) R.B. Kidd and S.L. Fogg, "A simple formula for the large-angle pendulum period", Physics Teacher **40**, 81 (2002).
(4) L. E. Millet, "The large-angle pendulum period", Physics Teacher 41, 162 (2003).]

For this example, we rewrite Eq. [1] as

$$T = a_0 + a_1x_1 + a_2x_2 + ..$$

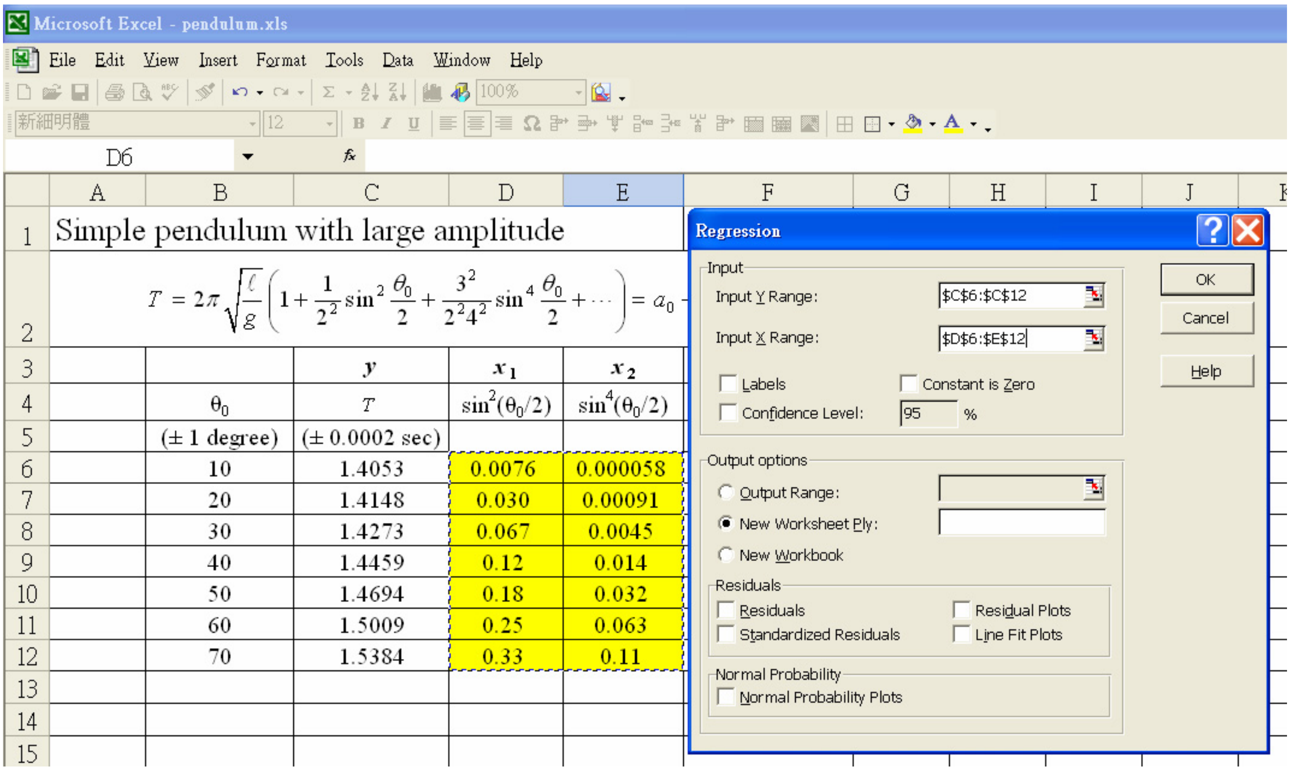
where $x_1 = \sin^2\frac{\theta_0}{2}$, $x_2 = \sin^4\frac{\theta_0}{2}$, $a_0 = 2\pi\sqrt{\frac{\ell}{g}}$, $a_1 = 2\pi\sqrt{\frac{\ell}{g}}\frac{1}{2^2}$, & $a_0 = 2\pi\sqrt{\frac{\ell}{g}}\frac{3^2}{2^2 4^2}$

Then prepare the Excel worksheet: (available for download in my course page)
& click Tools/Data Analysis/Regression as we did in Appendix B:



	y	x_1	x_2
θ_0	T	$\sin^2(\theta_0/2)$	$\sin^4(\theta_0/2)$
(± 1 degree)	(± 0.0002 sec)		
10	1.4053	0.0076	0.000058
20	1.4148	0.030	0.00091
30	1.4273	0.067	0.0045
40	1.4459	0.12	0.014
50	1.4694	0.18	0.032
60	1.5009	0.25	0.063
70	1.5384	0.33	0.11

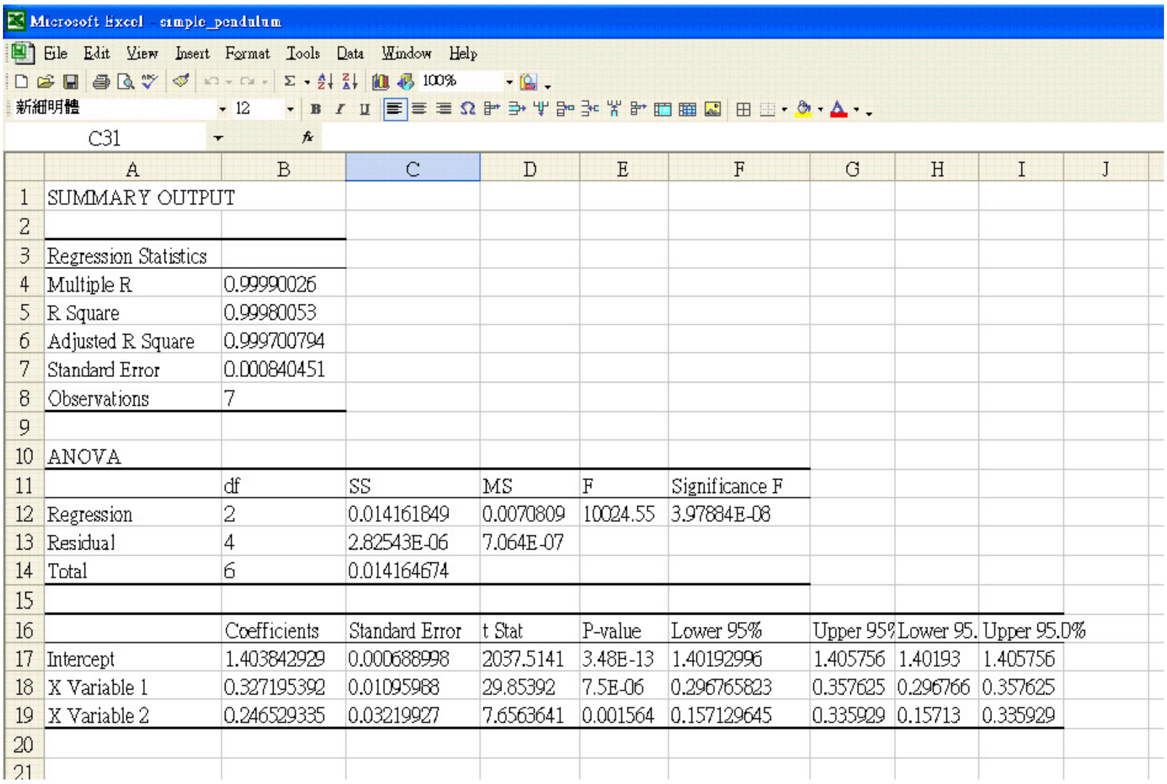
Input Y Range with T data.
Input X Range with all x_1 and x_2 data:



Microsoft Excel - pendulum.xls

	A	B	C	D	E
1	Simple pendulum with large amplitude				
2	$T = 2\pi\sqrt{\frac{\ell}{g}}\left(1+\frac{1}{2^2}\sin^2\frac{\theta_0}{2}+\frac{3^2}{2^2 4^2}\sin^4\frac{\theta_0}{2}+\cdots\right) = a_0$				
3			y	x_1	x_2
4		θ_0	T	$\sin^2(\theta_0/2)$	$\sin^4(\theta_0/2)$
5		(± 1 degree)	(± 0.0002 sec)		
6		10	1.4053	0.0076	0.000058
7		20	1.4148	0.030	0.00091
8		30	1.4273	0.067	0.0045
9		40	1.4459	0.12	0.014
10		50	1.4694	0.18	0.032
11		60	1.5009	0.25	0.063
12		70	1.5384	0.33	0.11

Then Click OK. The regression result is shown in a new worksheet:



Microsoft Excel - simple_pendulum

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	Regression Statistics								
4	Multiple R	0.99990026							
5	R Square	0.99980053							
6	Adjusted R Square	0.999700794							
7	Standard Error	0.000840451							
8	Observations	7							
9									
10	ANOVA								
11		df	SS	MS	F	Significance F			
12	Regression	2	0.014161849	0.0070809	10024.55	3.97884E-08			
13	Residual	4	2.82543E-06	7.064E-07					
14	Total	6	0.014164674						
15									
16		Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.	Upper 95.0%
17	Intercept	1.403842929	0.000688998	2037.5141	3.48E-13	1.40192996	1.405756	1.40193	1.405756
18	X Variable 1	0.327195392	0.01095988	29.85392	7.5E-06	0.296765823	0.357625	0.296766	0.357625
19	X Variable 2	0.246529335	0.03219927	7.6563641	0.001564	0.157129645	0.335929	0.15713	0.335929

The fitting results are $a_0 = 1.4038 \pm 0.0007$

$a_1 = 0.33 \pm 0.01$

$a_2 = 0.25 \pm 0.03$

For this experiment, we can plot $(\sin^2\frac{\theta_0}{2} + \frac{3^2}{4^2}\sin^4\frac{\theta_0}{2})$ vs. T

& use the slope of the linear graph to determine g as shown in the Excel file (pendulum.xls).