UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

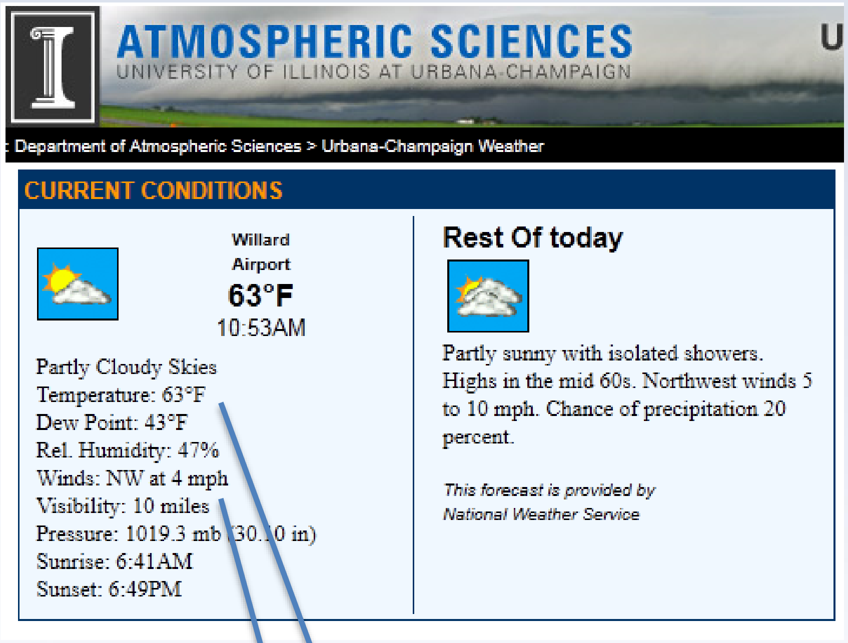
Basic Error Analysis

Physics 401
Spring 2019
Eugene V Colla

Agenda

- Errors and uncertainties
- The Reading Error
- Accuracy and precession
- Systematic and statistical errors
- Fitting errors
- Appendix. Working with oil drop data
  - Nonlinear fitting

What and when we need to know about errors. Everyday life.


ATMOSPHERIC SCIENCES
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Department of Atmospheric Sciences > Urbana-Champaign Weather

CURRENT CONDITIONS

Willard Airport
63°F
10:53AM

Partly Cloudy Skies
Temperature: 63°F
Dew Point: 43°F
Rel. Humidity: 47%
Winds: NW at 4 mph
Visibility: 10 miles
Pressure: 1019.3 mb (30.10 in)
Sunrise: 6:41AM
Sunset: 6:49PM

Rest Of today

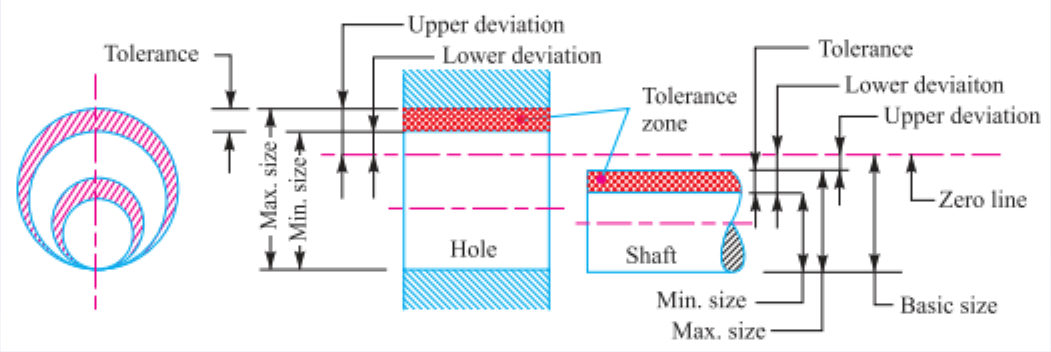
Partly sunny with isolated showers. Highs in the mid 60s. Northwest winds 5 to 10 mph. Chance of precipitation 20 percent.

This forecast is provided by National Weather Service

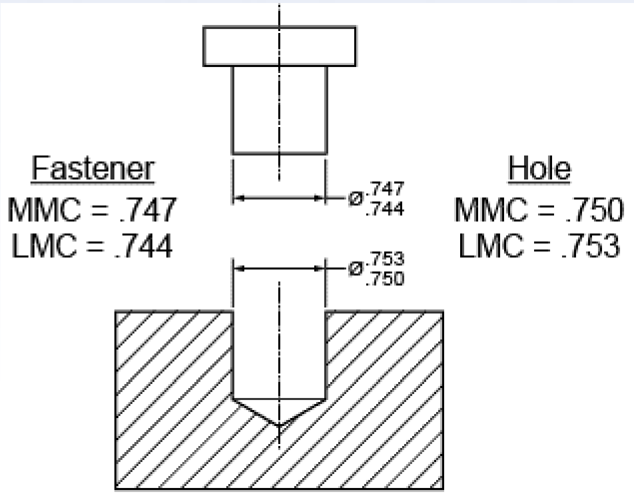
$T = 63°\text{F} \pm ?$ → **Best guess $\Delta T \sim 0.5°\text{F}$**

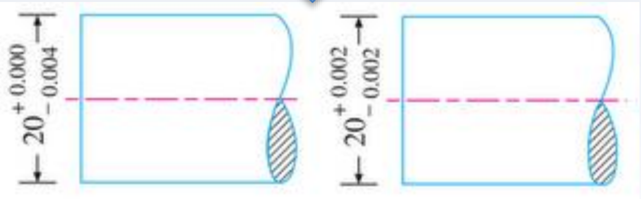
Wind speed 4mph$\pm$? → **Best guess $\pm 0.5 mph$**

What and when we need to know about errors. Industry.



Clearance fit





(a) Unilateral tolerance (b) Bilateral tolerance.

What and when we need to know about errors. Science.

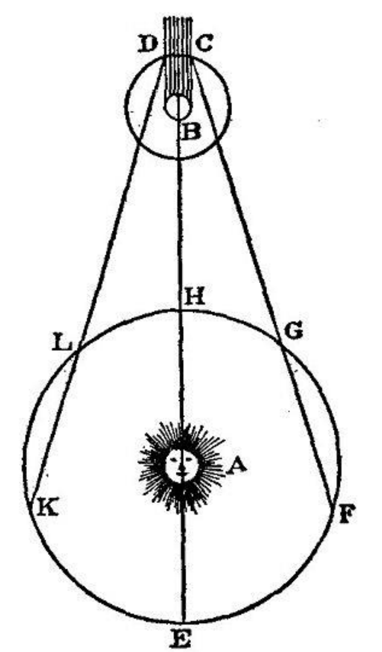


FIG. 70.

Measurement of the speed of the light

1675 Ole Roemer: 220,000 Km/sec

Ole Christensen Rømer
1644-1710

Does it make sense?
What is missing?

NIST Bolder Colorado c = 299,792,456.2±1.1 m/s.

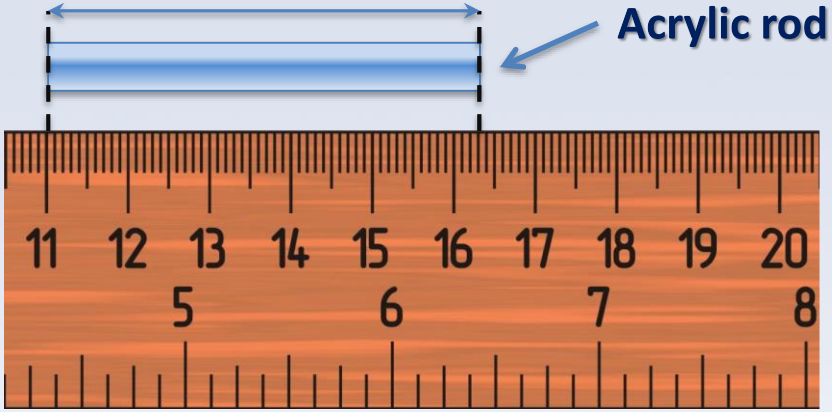
Reading error

$\Delta L \cong 0.5mm$

L=53mm±ΔL(?)

$\Delta L \cong 0.03mm$

Acrylic rod



How far we have to go in reducing the reading error?

We do not care about accuracy better than 1mm

If ruler is not okay, we need to use digital caliper

Probably the natural limit of accuracy can be due to length uncertainty because of temperature expansion. For 53mm $\Delta L \cong 0.012mm/K$

Reading Error = $\pm\frac{1}{2}$ (least count or minimum gradation).

Reading error. Digital meters.

Fluke 8845A multimeter

Example Vdc (reading)=0.85V

$$\Delta V = 0.83 \times \left(1.8 \times 10^{-5}\right) + 1.0 \times \left(0.7 \times 10^{-5}\right) \cong 2.2 \times 10^{-5} = 22\mu V$$

8846A Accuracy

Accuracy is given as ± (% measurement + % of range)

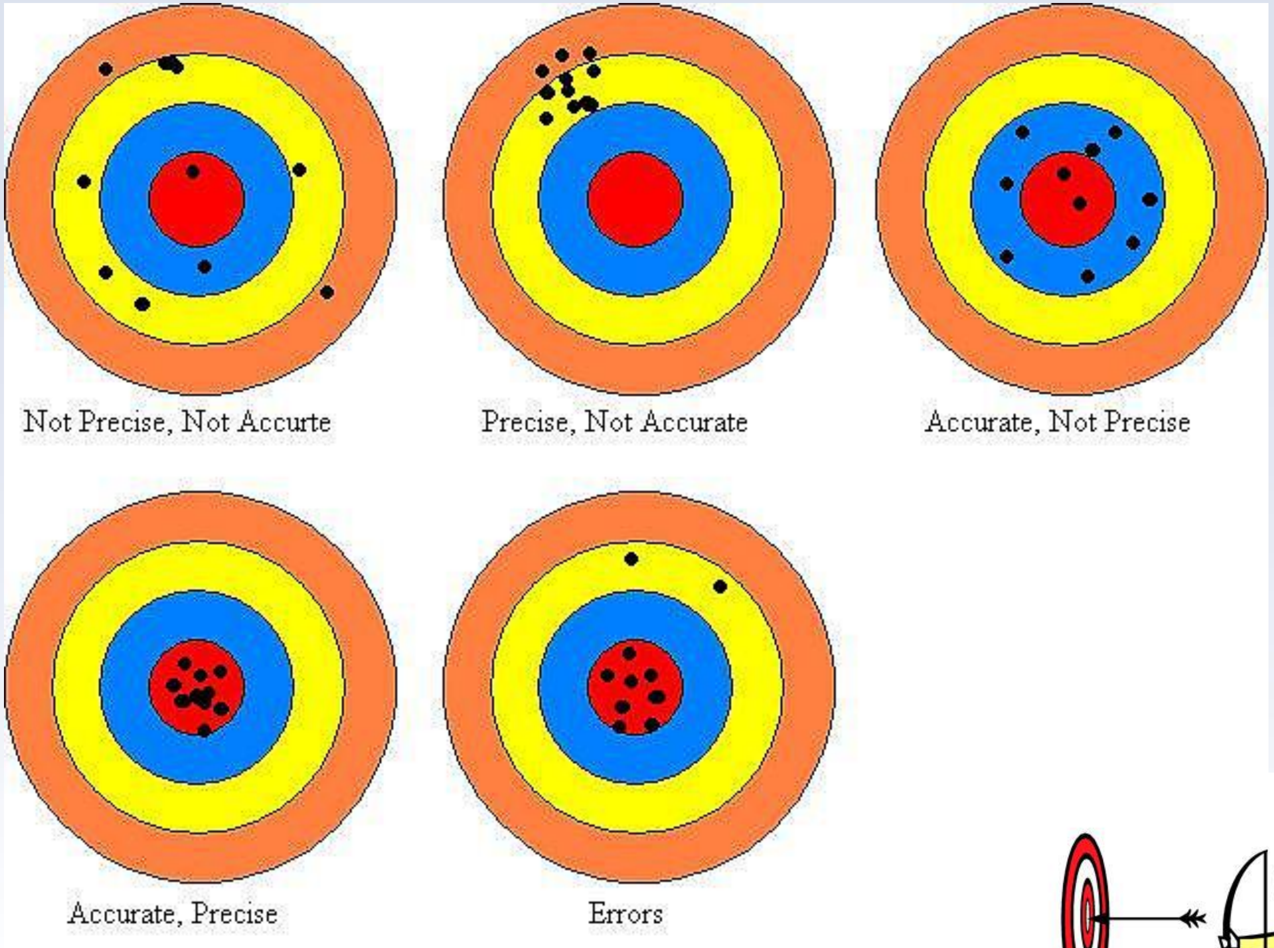
Range	24 Hour (23 ±1 °C)	90 Days (23 ±5 °C)	1 Year (23 ±5 °C)	Temperature Coefficient/ °C Outside 18 to 28 °C
100 mV	0.0025 + 0.003	0.0025 + 0.0035	0.0037 + 0.0035	0.0005 + 0.0005
1 V	0.0018 + 0.0006	0.0018 + 0.0007	0.0025 + 0.0007	0.0005 + 0.0001
10 V	0.0013 + 0.0004	0.0018 + 0.0005	0.0024 + 0.0005	0.0005 + 0.0001
100 V	0.0018 + 0.0006	0.0027 + 0.0006	0.0038 + 0.0006	0.0005 + 0.0001
1000 V	0.0018 + 0.0006	0.0031 + 0.001	0.0041 + 0.001	0.0005 + 0.0001

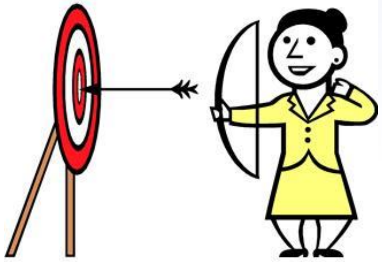
Accuracy and precession

The accuracy of an experiment is a measure of how close the result of the experiment comes to the true value

Precision refers to how closely individual measurements agree with each other

Accuracy and precession





Systematic and random errors

- ***Systematic Error:*** **reproducible inaccuracy introduced by faulty equipment, calibration or technique.**

- ***Random errors*****: Indefiniteness of results due to finite precision of experiment. Measure of fluctuation in result after repeatable experimentation.**

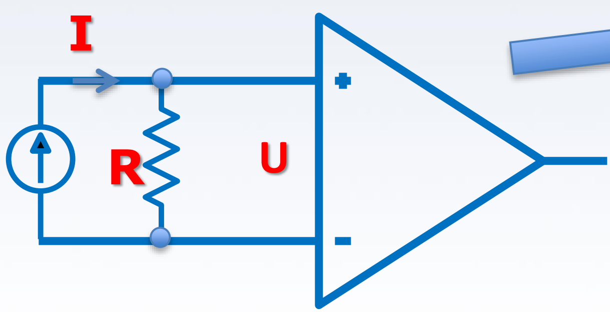
Philip R. Bevington "Data Reduction and Error Analysis for the Physical sciences", McGraw-Hill, 1969

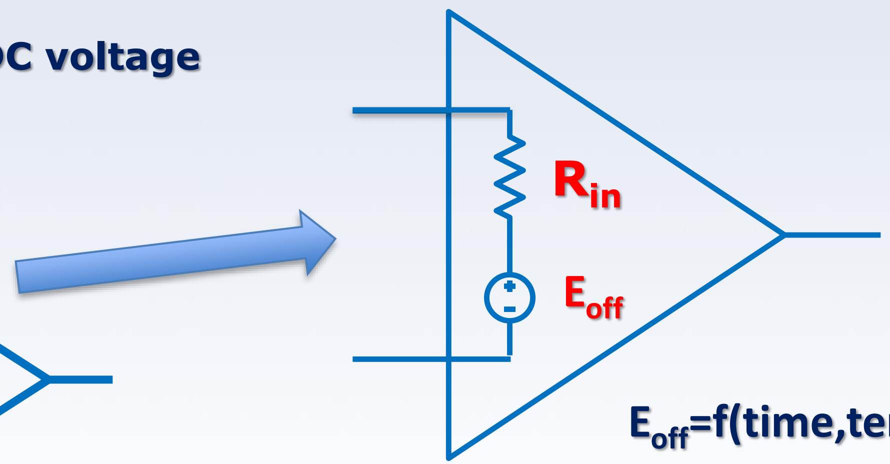
Systematic errors

Sources of systematic errors: **poor calibration of the equipment, changes of environmental conditions, imperfect method of observation, drift and some offset in readings etc.**

Example #1: measuring of the DC voltage

Current source





E_{off}=f(time,temperature)

expectation

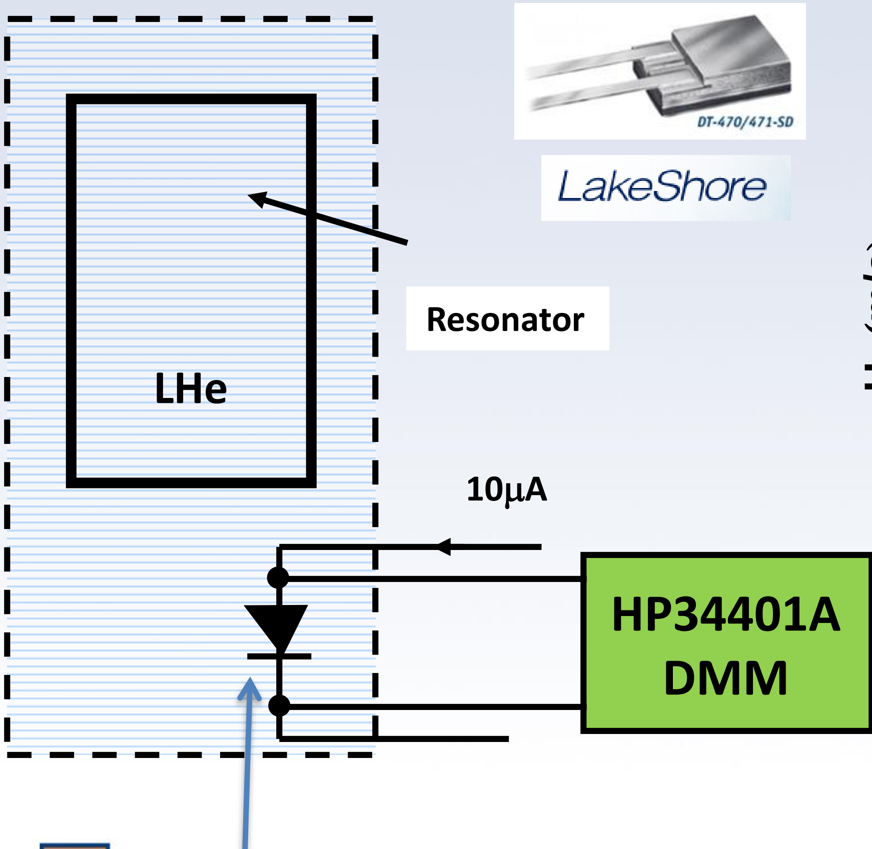
U=R*I

actual result

$$U = \frac{R * I - \left(\frac{R}{R_{in}}\right) E_{off}}{\left(1 + \frac{R}{R_{in}}\right)}$$

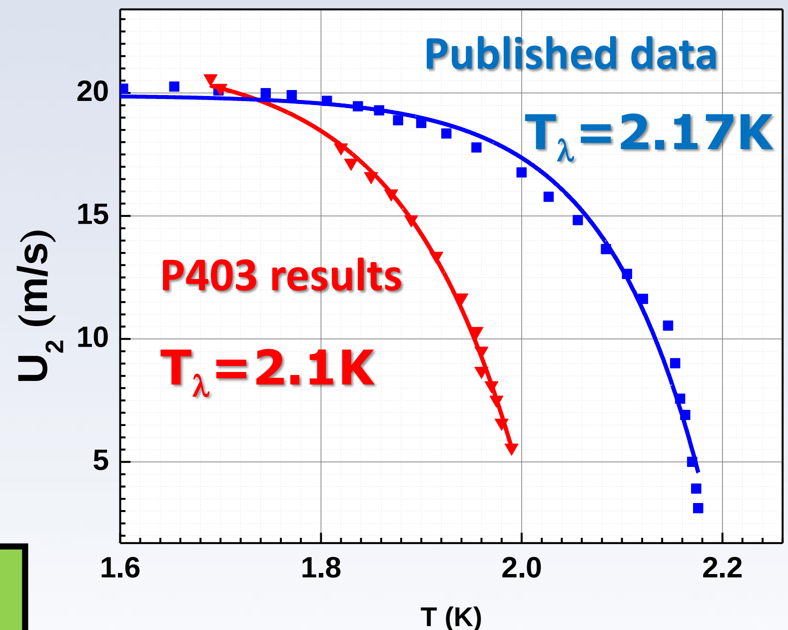
Systematic errors

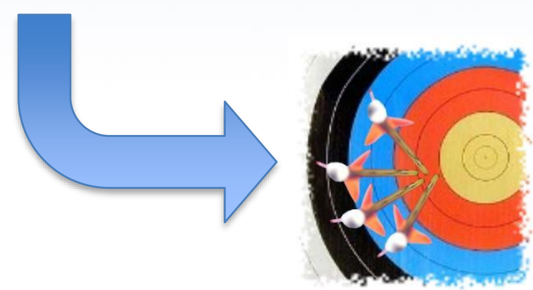
Example #3: poor calibration



Temperature sensor

Measuring of the speed of the second sound in superfluid He4





Random errors

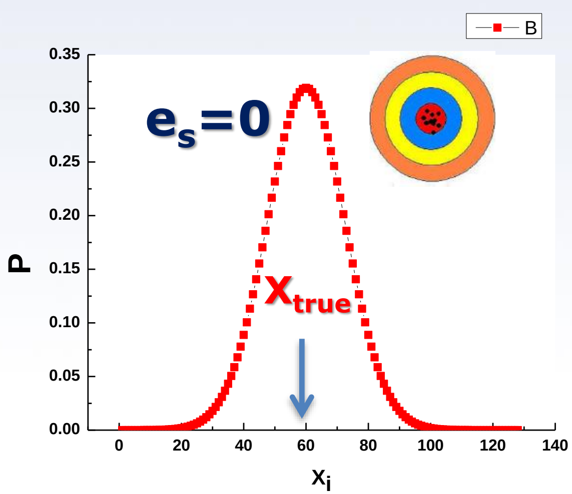
Result of measurement

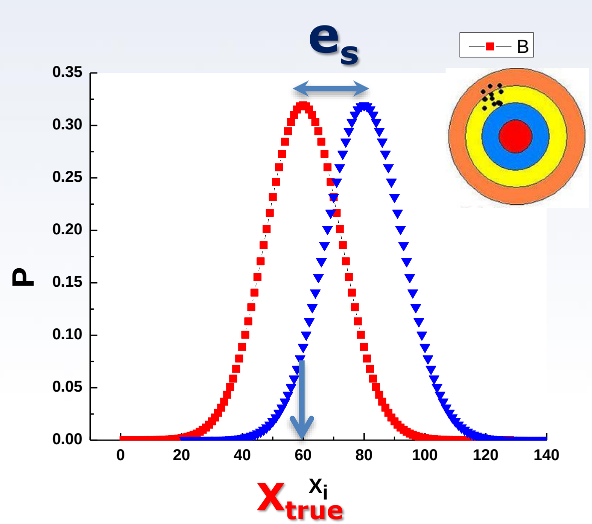
Systematic error

$$x_{meas} = x_{true} + e_s + e_r$$

Correct value

Random error





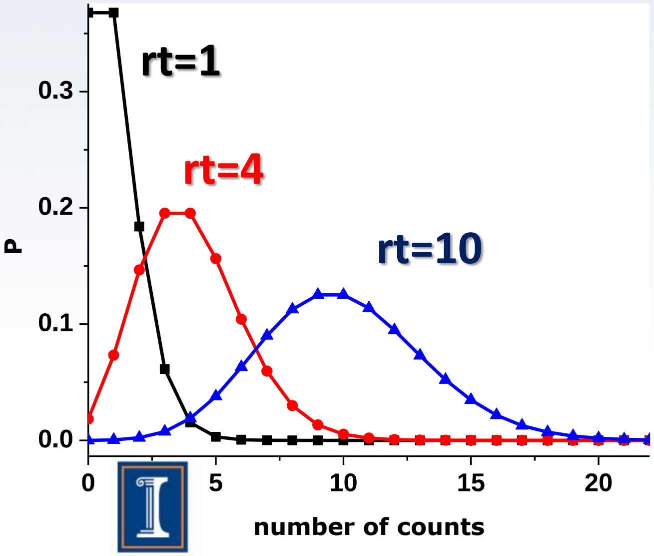
Random errors. Poisson distribution

Siméon Denis Poisson
(1781-1840)

$$P_n(t) = \frac{(rt)^n}{n!} e^{-rt} \quad n = 0,1,2,\ldots$$

r: decay rate [counts/s] *t*: time interval [s]

➔ $P_n(rt)$: Probability to have *n* decays in time interval *t*



illinois.edu

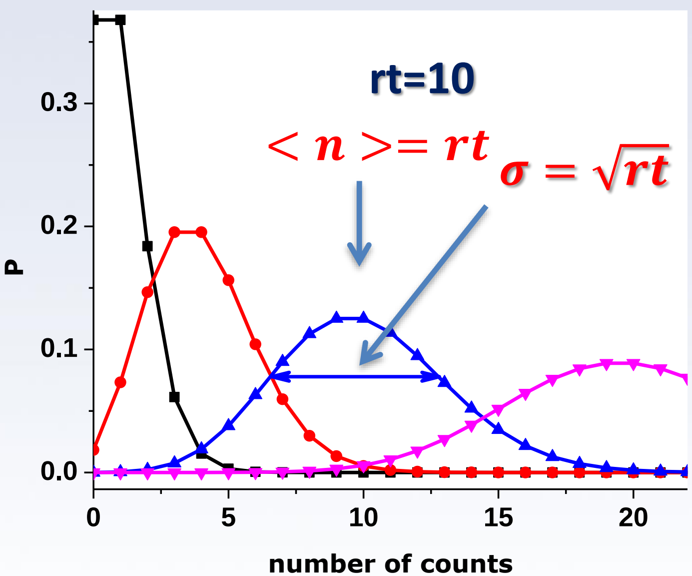
A statistical process is described through a Poisson Distribution if:

- **random process** → for a given nucleus probability for a decay to occur is the same in each time interval.
- **universal probability** → the probability to decay in a given time interval is same for all nuclei.
- **no correlation between two instances** (the decay of on nucleus does not change the probability for a second nucleus to decay.

Poisson distribution

$$P_n(t) = \frac{(rt)^n}{n!} e^{-rt} \qquad n = 0,1,2,\ldots$$

r: decay rate [counts/s] *t*: time interval [s]
➔ $P_n(rt)$: Probability to have n decays in time interval *t*



Properties of the Poisson distribution:

$$\sum_{n=0}^{\infty} P_n(rt) = 1$$, probabilities sum to 1

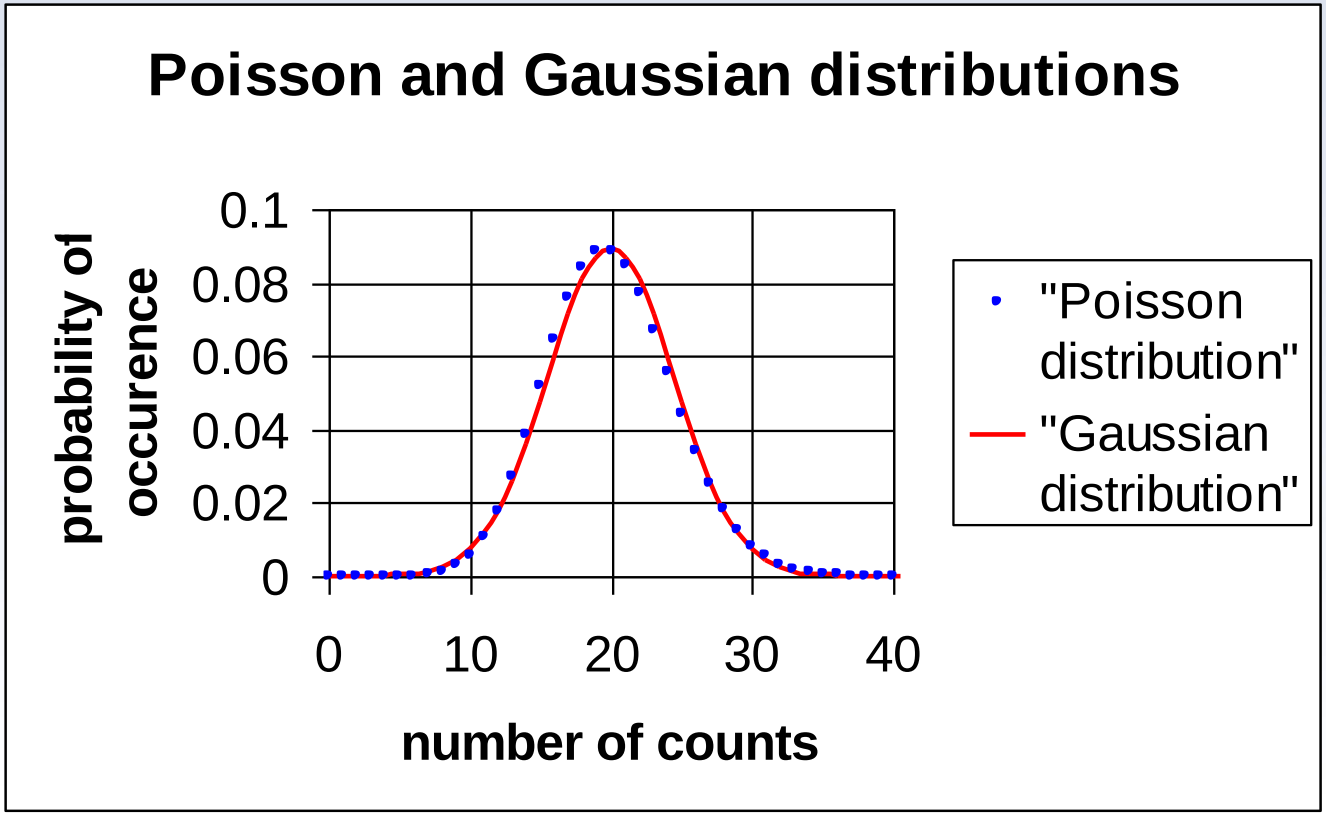
$$<n> = \sum_{n=0}^{\infty} n \cdot P_n(rt) = rt$$, the mean

$$\sigma = \sqrt{\sum_{n=0}^{\infty} (n - <n>)^2 P_n(rt)} = \sqrt{rt}$$, standard deviation

illinois.edu

Poisson distribution at large *rt*

$$P_n(t) = \frac{(rt)^n}{n!} e^{-rt} \quad n = 0,1,2,\ldots$$

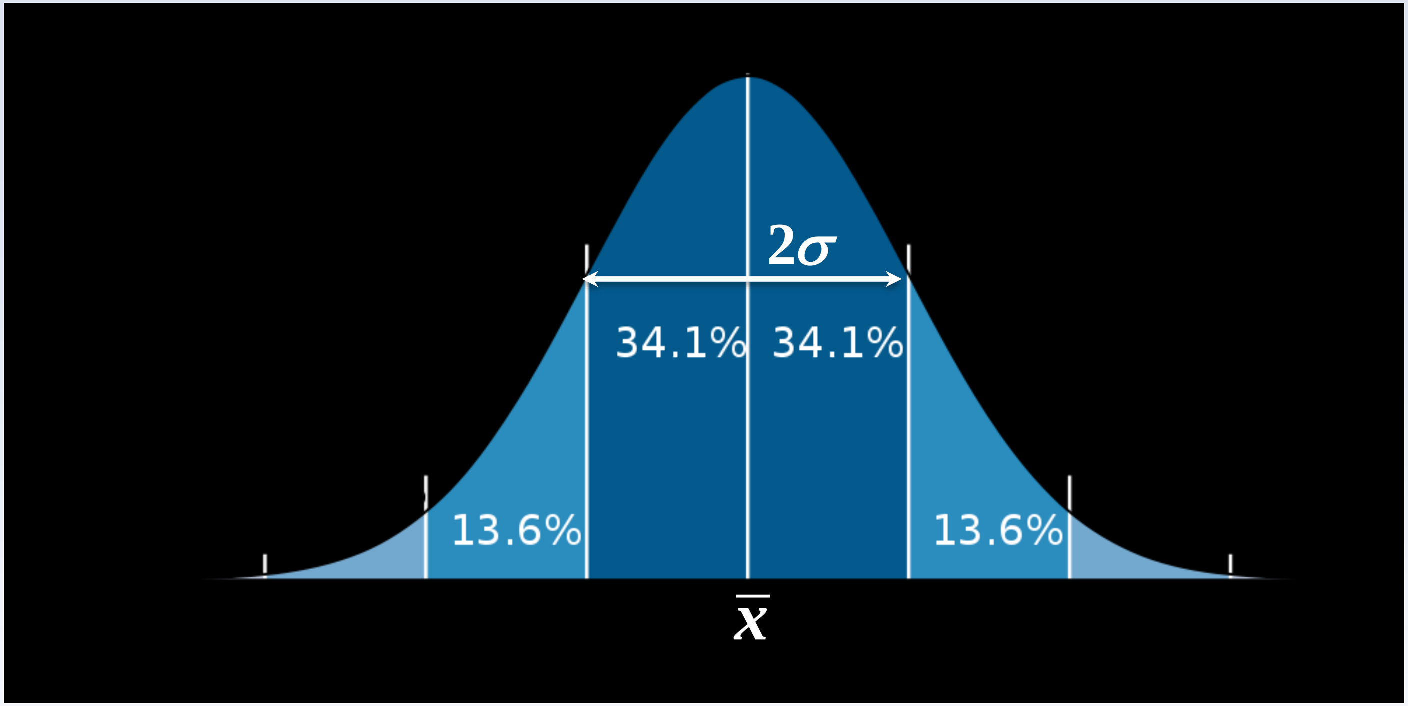


Carl Friedrich Gauss
(1777–1855)

$$P_n(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

Gaussian distribution: continuous

Normal (Gaussian) distribution



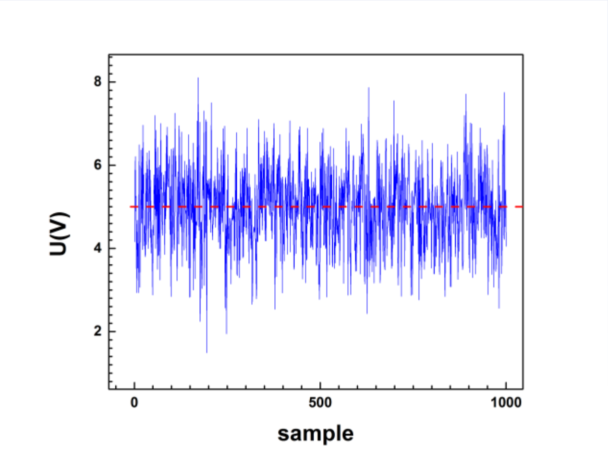
$$P_n(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

Error in the mean is given as $\frac{\sigma}{\sqrt{N}}$,

N – number od events

Measurement in presence of noise

Source of noisy signal

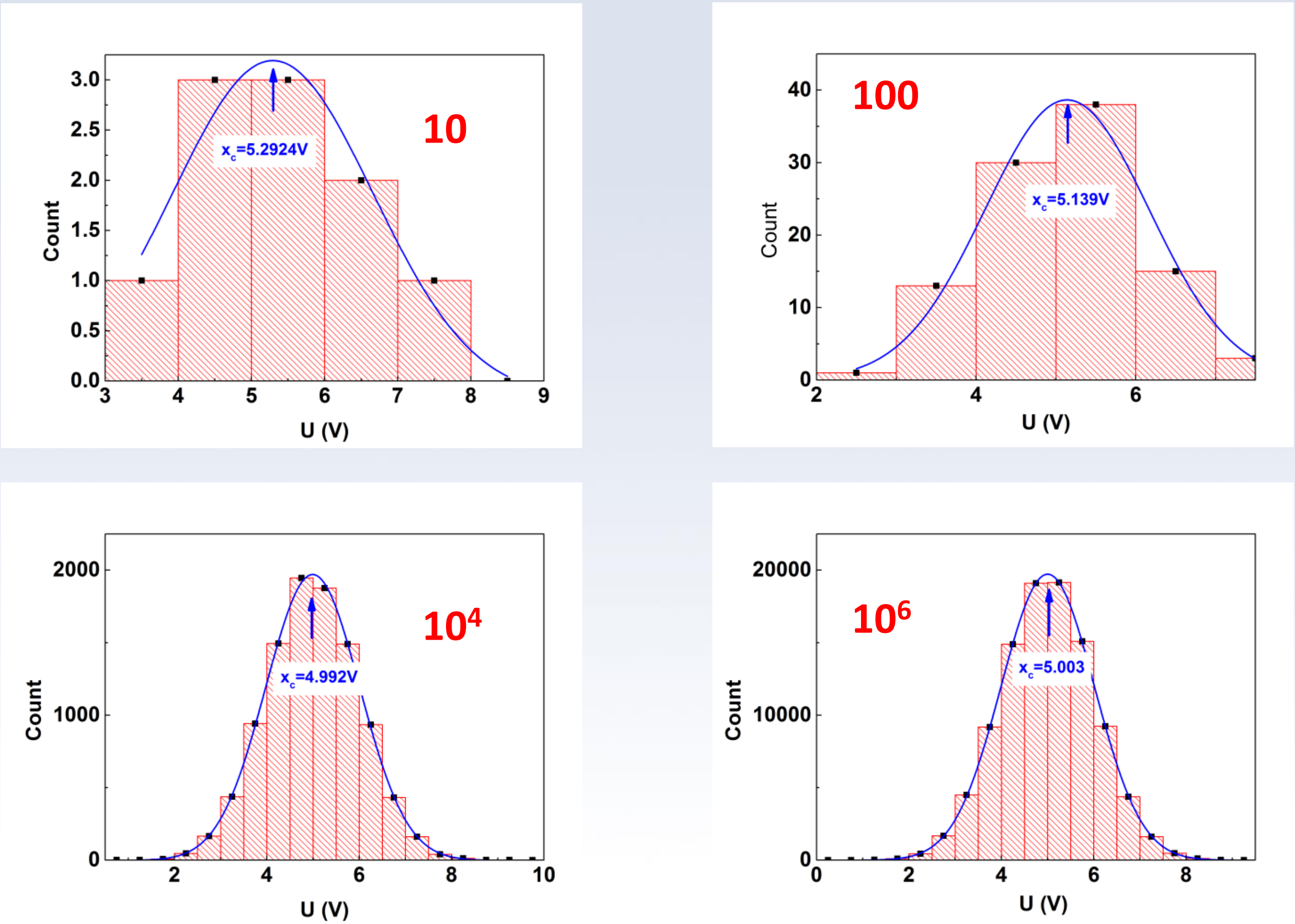


4.89855
5.25111
2.93382
4.31753
4.67903
3.52626
4.12001
2.93411

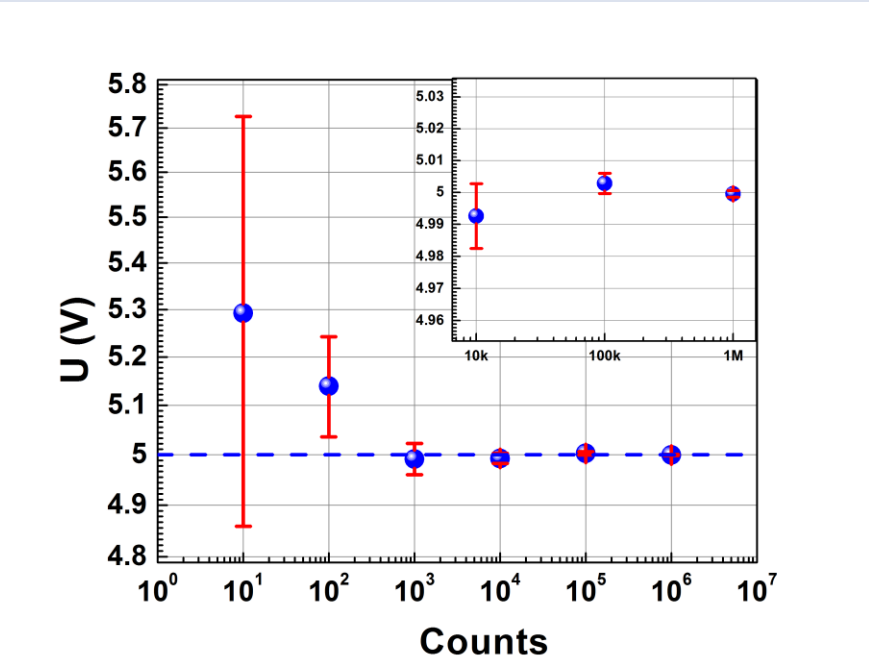
Expected value 5V

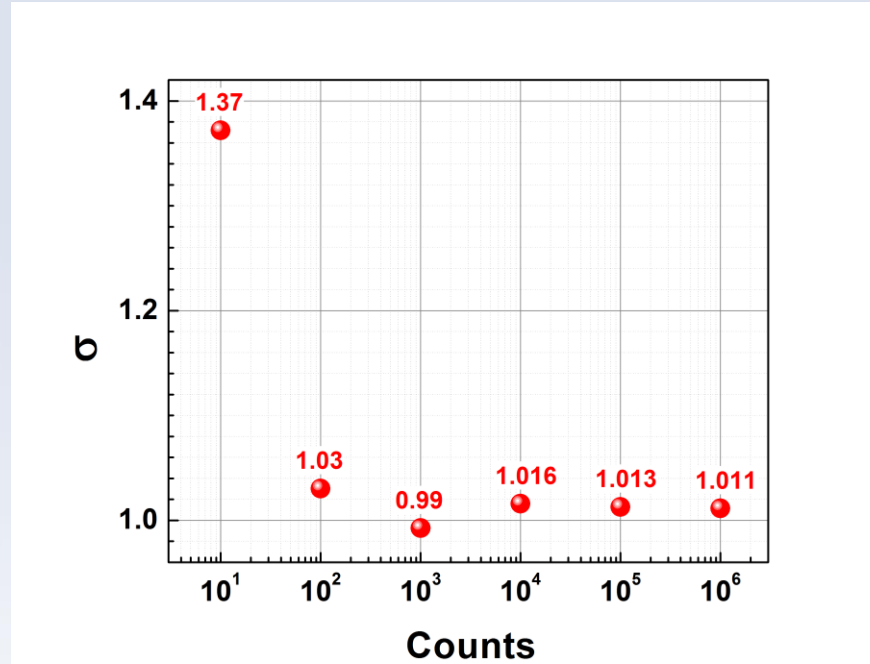
Actual measured values

Measurement in presence of noise



Measurement in presence of noise





Result → $U = x_c \pm \frac{\sigma}{\sqrt{N}}$

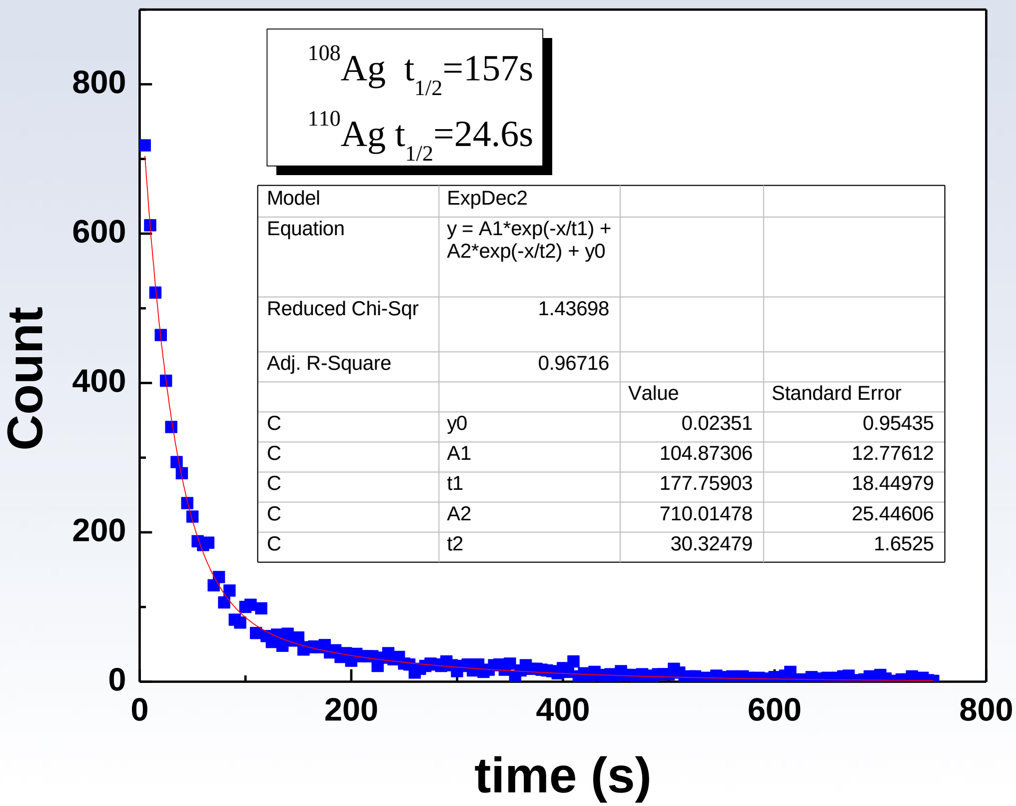
σ - standard deviation
N – number of samples

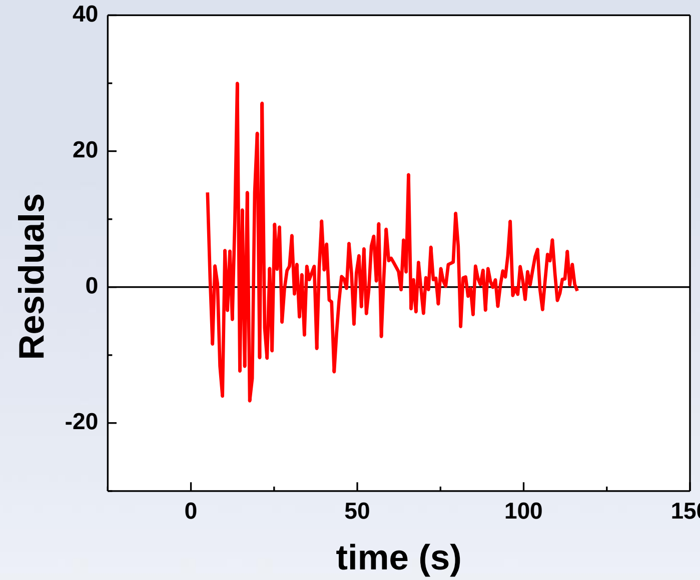
For $N=10^6$ U=4.999±0.001 0.02% accuracy

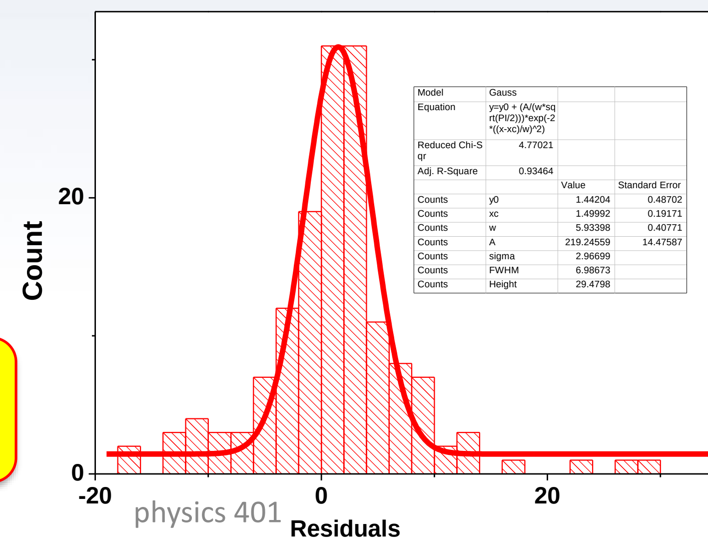
illinois.edu

Fitting errors

Ag β decay





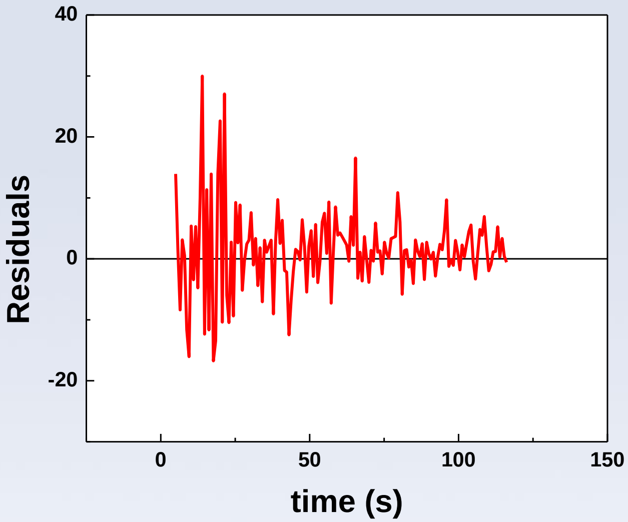


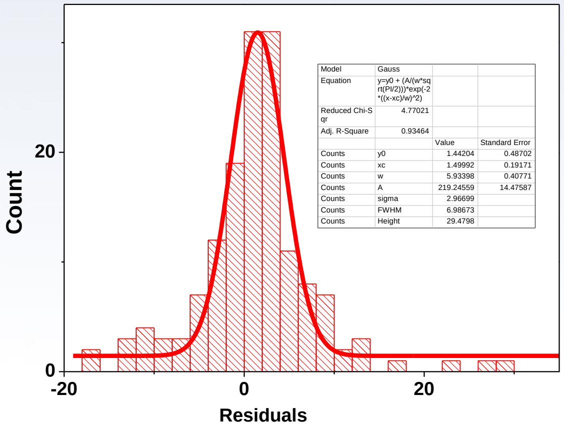
$$y = A1 \bullet \exp\left(\frac{-t}{t_1}\right) + A2 \bullet \exp\left(\frac{-t}{t_2}\right) + y_0$$

illinois.edu

Fitting. Analysis of the residuals

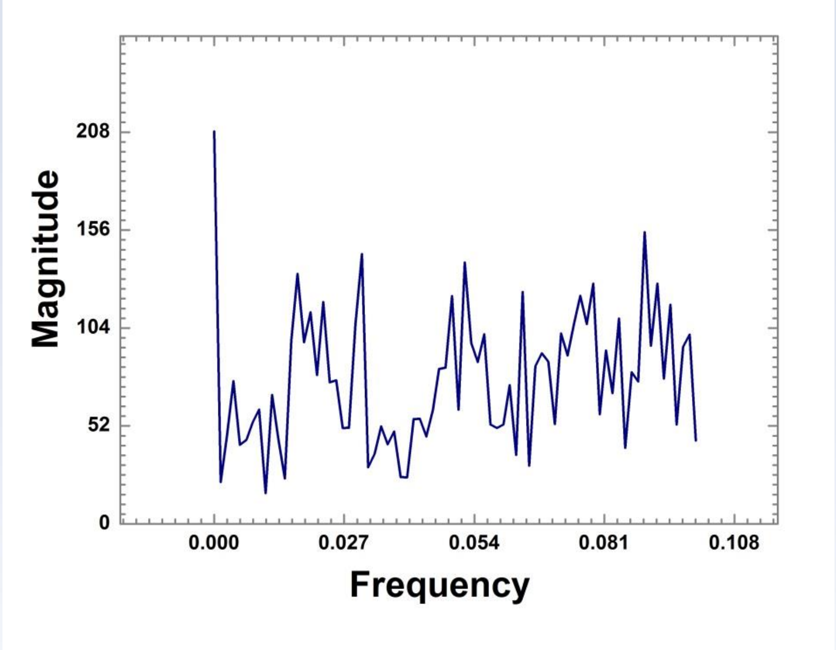
Ag β decay





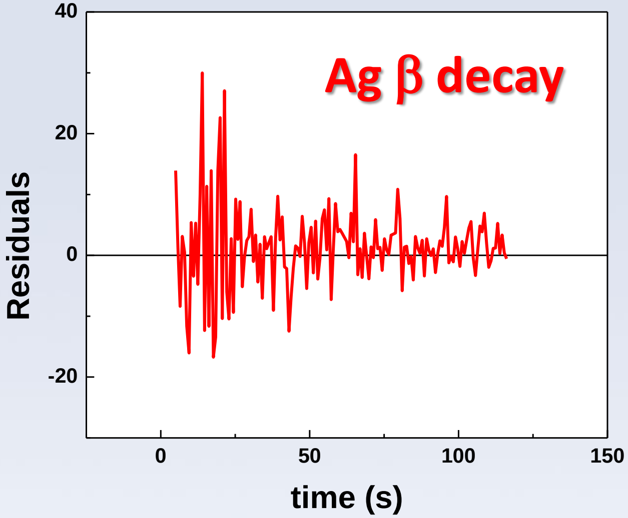
Model	Gauss		
Equation	y=y0 + (A/(w*sqrt(PI/2)))*exp(-2*((x-xc)/w)^2)		
Reduced Chi-Sqr	4.77021		
Adj. R-Square	0.93464		
		Value	Standard Error
Counts	y0	1.44204	0.48702
Counts	xc	1.49992	0.19171
Counts	w	5.93398	0.40771
Counts	A	219.24559	14.47587
Counts	sigma	2.96699	
Counts	FWHM	6.98673	
Counts	Height	29.4798	

Test 1. Fourier analysis

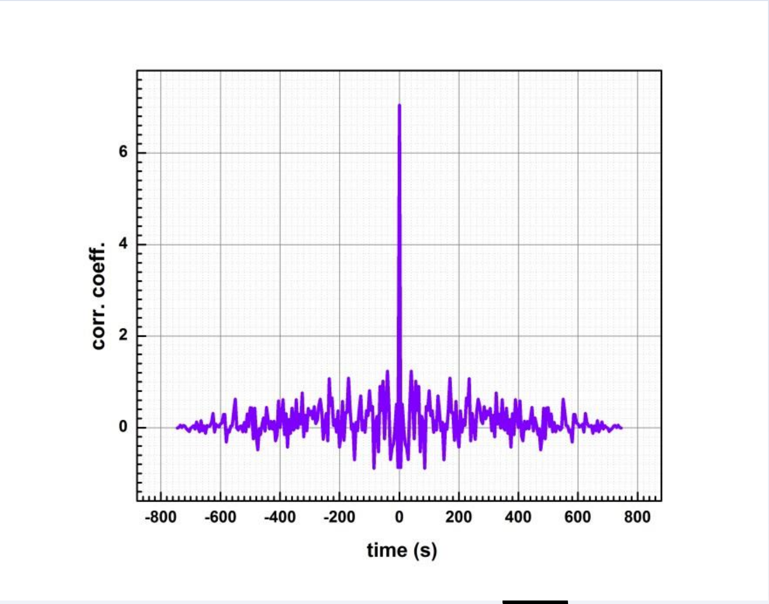


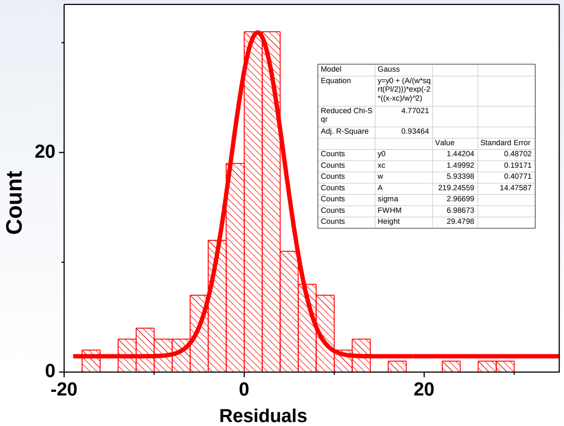
No pronounced frequencies found

Fitting. Analysis of the residuals



Test 1. Autocorrelation function





Model	Gauss		
Equation	y=y0 + (A/(w*sqrt(PI/2)))*exp(-2*((x-xc)/w)^2)		
Reduced Chi-Sqr	4.77021		
Adj. R-Square	0.93464		
		Value	Standard Error
Counts	y0	1.44204	0.48702
Counts	xc	1.49992	0.19171
Counts	w	5.93398	0.40771
Counts	A	219.24559	14.47587
Counts	sigma	2.96699	
Counts	FWHM	6.98673	
Counts	Height	29.4798	

Correlation function

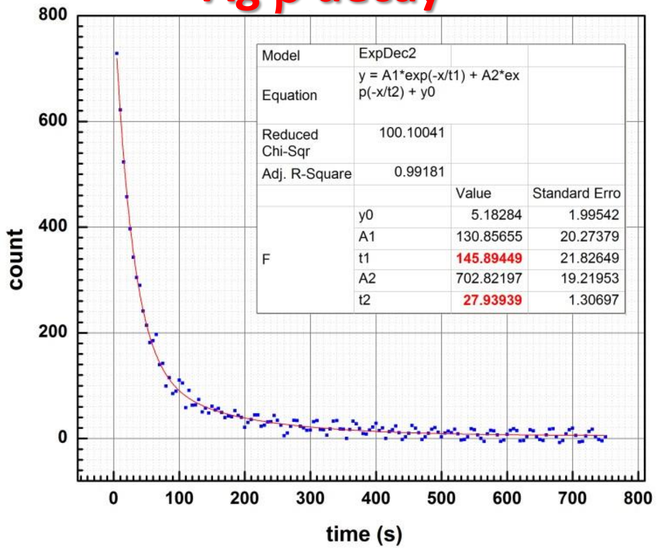
$$y(m) = \sum_{n=0} f(n)g(n-m)$$

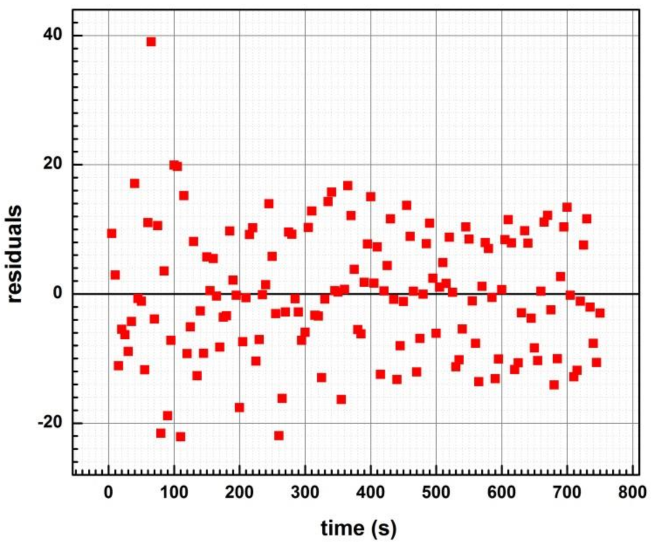
autocorrelation function

$$y(m) = \sum_{n=0}^{M-1} f(n)f(n-m)$$

Fitting. Analysis of the residuals. Non “ideal” case

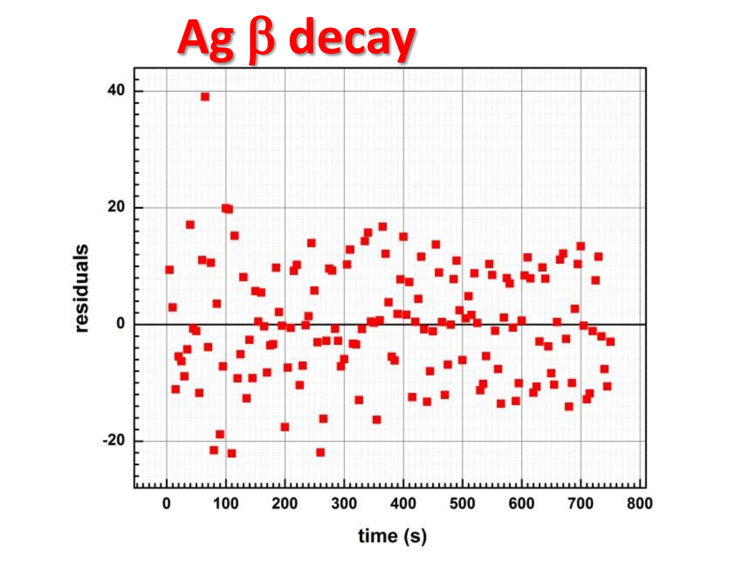
Ag β decay

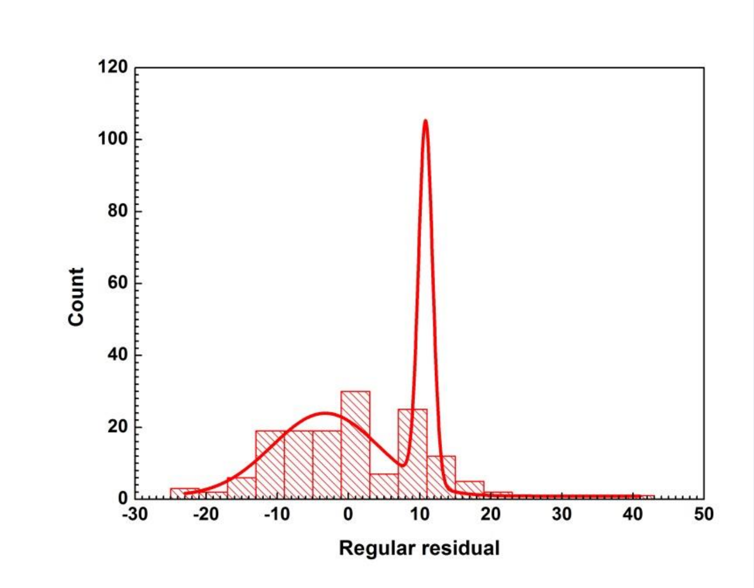


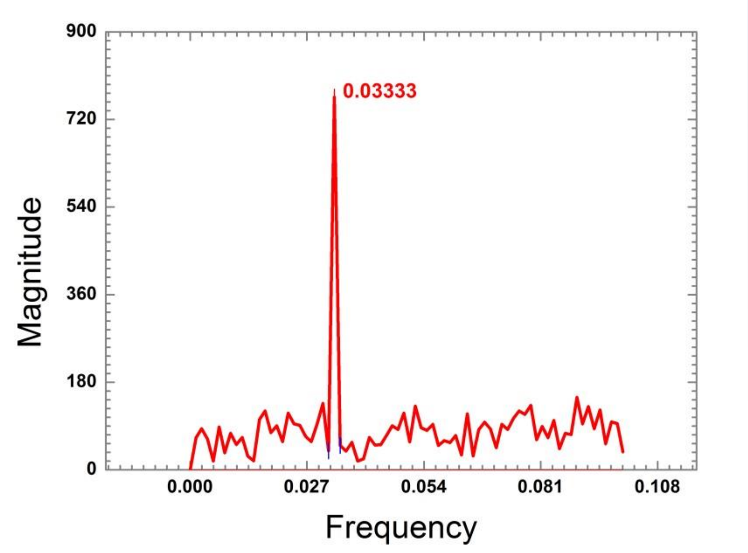


	Clear experiment	Data + “noise”
t_1(s)	177.76	145.89
t_2(s)	30.32	27.94

Fitting. Analysis of the residuals. Non "ideal" case

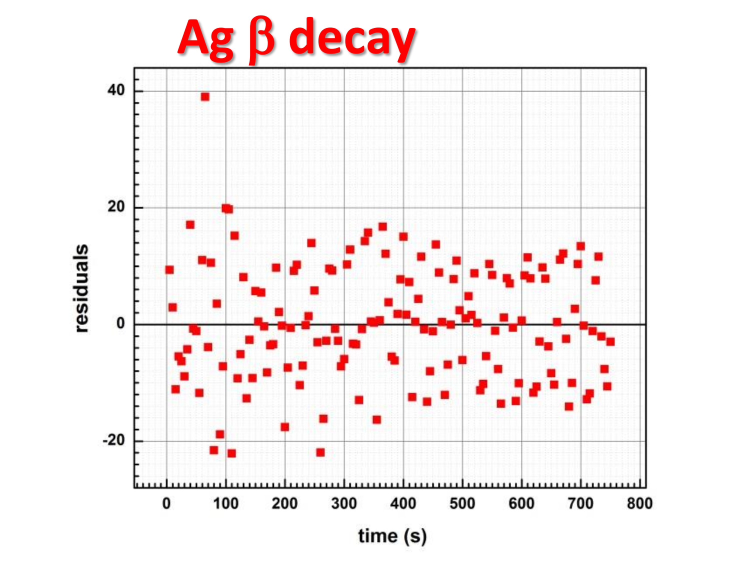


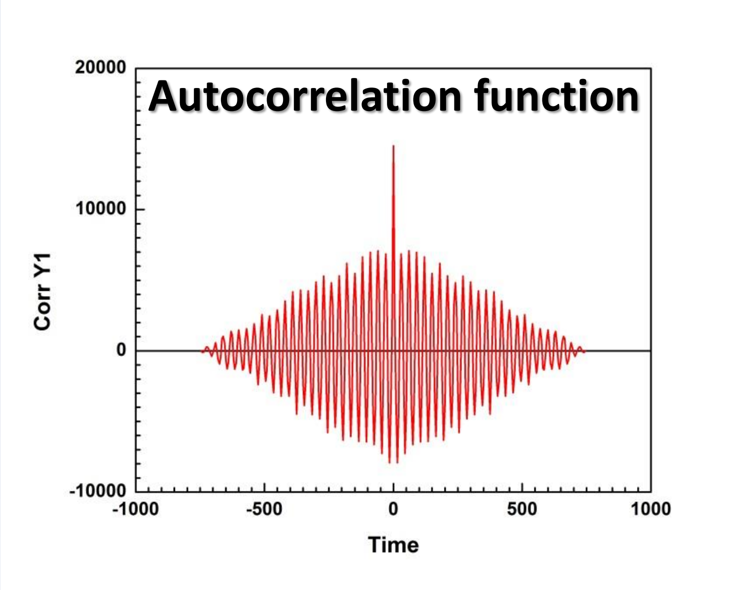




Histogram does not follow the normal distribution and there is frequency of 0.333 is present in spectrum

Fitting. Analysis of the residuals. Non "ideal" case

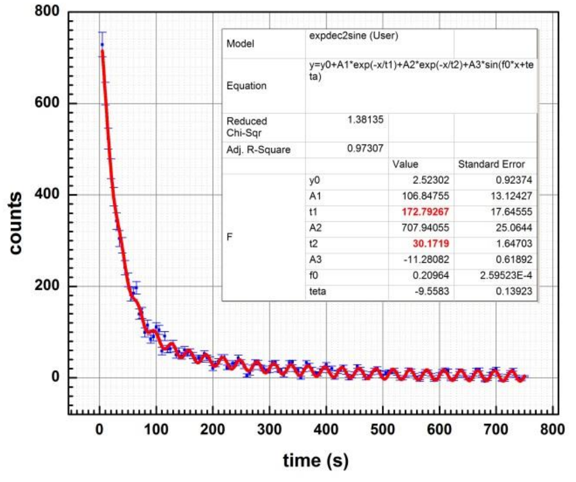


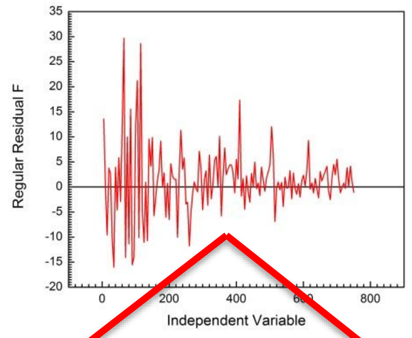


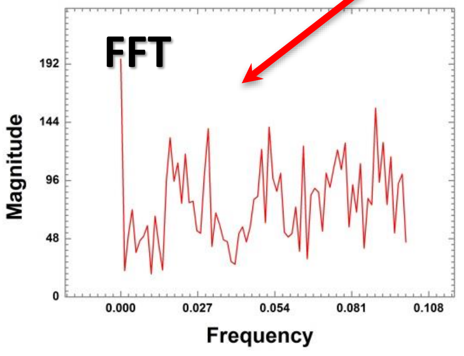
Conclusion: fitting function should be modified by adding an additional term:

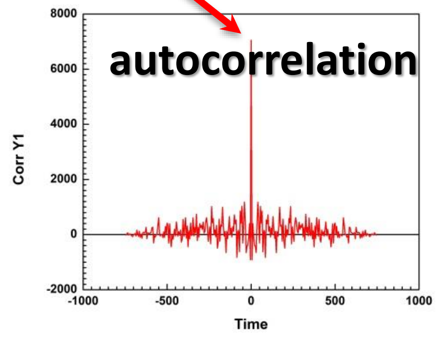
$$y(t) = y_0 + A_1 \exp\left(\frac{-t}{t_1}\right) + A_2 \exp\left(\frac{-t}{t_2}\right) + A_3 \sin(\omega t + \theta)$$

Fitting. Analysis of the residuals. Non “ideal” case









	Clear experiment	Data + noise	Modified fitting
t_1(s)	177.76	145.89	172.79
t_2(s)	30.32	27.94	30.17

Error Analysis. Millikan oil drop experiment.

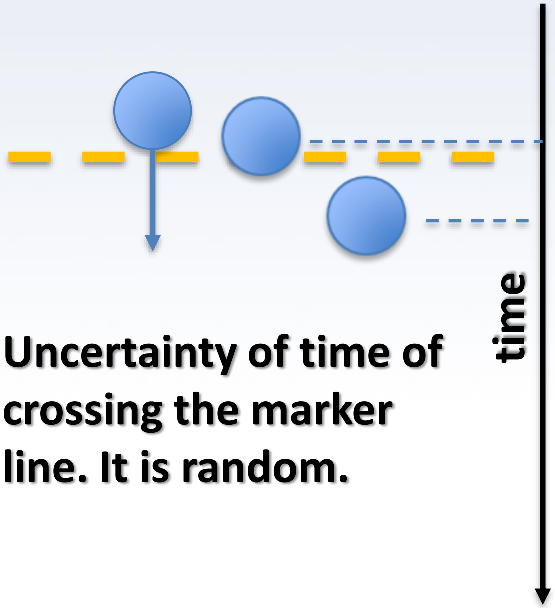
In general we could expect both components of errors

$$Q_{meas} = Q_{true} + e_s + e_r$$

e_s **- systematic error comes from uncertainties of plates separation distance, applied DC voltage, ambient temperature etc.**

$V = V_{DC} \pm \Delta V$, $d = d_0 \pm \Delta d$ …

e_r **- random errors are related to uncertainty of the knowledge of the actual t_g and t_{rise}.**



Uncertainty of time of crossing the marker line. It is random.

illinois.edu

Systematic component. Error propagation. Millikan oil drop experiment.

$X_{meas} = X_{true} + e_s + e_r$ **Systematic error**

$$Q = F \bullet S \bullet T = \left(\frac{1}{f_c^{3/2}}\right)\frac{9\pi d}{V}\sqrt{\frac{2\eta^3 x^3}{g\rho}}\frac{1}{\sqrt{t_g}}\left(\frac{1}{t_g}+\frac{1}{t_{rise}}\right)$$

$$F = \frac{1}{f_c^{3/2}} \approx 1-\left(\frac{t_g}{\tau_g}\right)^{\frac{1}{2}}$$

$$S = \frac{9\pi d}{V}\sqrt{\frac{2\eta^3 x^3}{g\rho}}$$

$$T = \frac{1}{\sqrt{t_g}}\left(\frac{1}{t_g}+\frac{1}{t_{rise}}\right)$$

$$\Delta Q = \sqrt{\left(\frac{dQ}{dF}\right)^2(\Delta F)^2+\left(\frac{dQ}{dS}\right)^2(\Delta S)^2+\left(\frac{dQ}{dT}\right)^2(\Delta T)^2} \approx \sqrt{\left(\frac{dQ}{dS}\right)^2(\Delta S)^2+\left(\frac{dQ}{dT}\right)^2(\Delta T)^2}$$

$$= \sqrt{(FT)^2(\Delta S)^2+(FS)^2(\Delta T)^2} = Q\sqrt{\left(\frac{\Delta S}{S}\right)^2+\left(\frac{\Delta T}{T}\right)^2}$$

Systematic component. Error propagation. Millikan oil drop experiment.

Systematic error

$$X_{meas} = X_{true} + e_s + e_r$$

$$\Delta Q \approx Q\sqrt{\left(\frac{\Delta S}{S}\right)^2 + \left(\frac{\Delta T}{T}\right)^2}$$

$$\frac{\Delta S}{S} = \sqrt{\left(\frac{\Delta d}{d}\right)^2 + \left(\frac{\Delta V}{V}\right)^2 + \left(\frac{3}{2}\frac{\Delta x}{x}\right)^2 + \left(\frac{3}{2}\frac{\Delta \eta}{\eta}\right)^2 + \left(\frac{1}{2}\frac{\Delta \rho}{\rho}\right)^2 + \left(\frac{1}{2}\frac{\Delta g}{g}\right)^2} \approx \sqrt{\left(\frac{\Delta d}{d}\right)^2 + \left(\frac{3}{2}\frac{\Delta x}{x}\right)^2}$$

$$\Delta T = \sqrt{\left(\frac{3/2}{t_g^{5/2}} + \frac{1/2}{t_g^{3/2}}\frac{1}{t_{rise}}\right)^2 \Delta t_g^{\,2} + \left(\frac{1}{t_g^{1/2}}\frac{1}{t_{rise}^{\,2}}\right)^2 \Delta t_{rise}^{\,2}}$$

Appendix #1. Analyzing of the statistical data.

Step 1. Origin Project For Raw Data :

\\engr-file-03\PHYINST\APL Courses\PHYCS401\Students\2. Millikan Raw Data

Data Analysis for Millikan Oil Drop Experi...	2/22/2008 9:36 AM	Adobe Acrobat D...	59 KB
Millikan_raw data.opj	10/5/2017 4:50 PM	OPJ File	15 KB
Millikan_raw data1.opj	9/25/2018 1:38 PM	OPJ File	14 KB
Millikan1_calc.opj	9/26/2018 1:36 PM	OPJ File	95 KB
Millikan1_no_calc.opj	9/25/2017 2:03 PM	OPJ File	66 KB
T measurement.opj	9/26/2018 1:36 PM	OPJ File	316 KB

All these projects with raw data should be stored in:

\\engr-file-03\PHYINST\APL Courses\PHYCS401\Students\2. Millikan Raw Data

Here should only the files with raw data but not other files which you using for calculations. All other files you can save in your personal folder

Appendix #1. Analyzing of the statistical data.

The “raw data” area is common place and plese remove from there the files which have no relation to the experimental results!

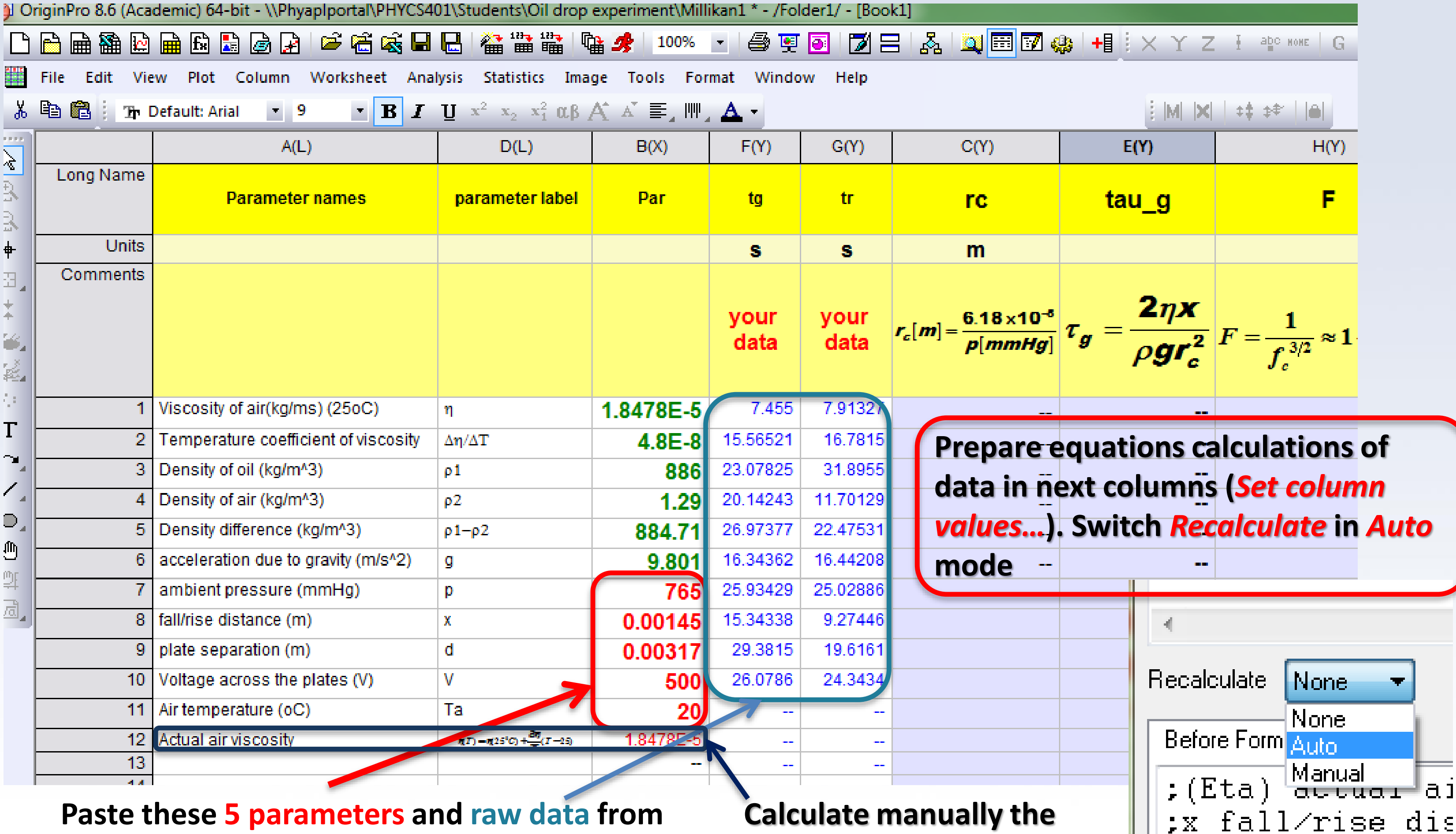
Name	Date modified	Type	Size
important	9/25/2018 3:52 PM	File folder	
Origin Folder.lnk	9/25/2018 1:28 PM	File folder	
Origin Samples Folder.lnk	9/25/2018 1:28 PM	File folder	
A1A2B1B2.bmp	11/9/2016 9:52 AM	BMP File	2 KB
AppsTabs.xml	9/25/2018 3:52 PM	XML Document	1 KB
Backup.opj	10/4/2017 5:09 PM	OPJ File	317 KB
Custom.ogs	4/19/2015 4:16 PM	OGS File	1 KB
excel.otw	11/9/2016 9:52 AM	Origin Worksheet ...	21 KB
Filter.ini	11/9/2016 9:52 AM	Configuration sett...	1 KB
MRFiles.ini	9/25/2018 1:29 PM	Configuration sett...	1 KB
NLSF.ini	9/25/2018 1:28 PM	Configuration sett...	10 KB
NLSFOide.ini	11/9/2016 9:52 AM	Configuration sett...	7 KB
NLSFwiz_OPS.bmp	11/9/2016 9:52 AM	BMP File	3 KB
NLSFwizard.ini	11/9/2016 9:52 AM	Configuration sett...	1 KB
Olbtedit.ini	11/9/2016 9:52 AM	Configuration sett...	7 KB
Origin.ini	9/25/2018 3:53 PM	Configuration sett...	19 KB
Origin94.INI	9/25/2018 1:28 PM	Configuration sett...	1 KB
OUbtn.ini	9/25/2018 1:28 PM	Configuration sett...	1 KB
Oubtn2.ini	9/25/2018 1:28 PM	Configuration sett...	1 KB
OubtnA1A2B1B2.ini	9/25/2018 1:28 PM	Configuration sett...	1 KB
Template.ini	11/9/2016 9:52 AM	Configuration sett...	2 KB
Userdef.bmp	11/9/2016 9:52 AM	BMP File	2 KB
userdef2.bmp	11/9/2016 9:52 AM	BMP File	3 KB

Type: File folder
Date modified: 9/25/2018 1:28 PM

Appendix #1. Analyzing of the statistical data.

Step 2. Working on personal Origin project

Make a copy of the Millikan1 project to your personal folder and open it



	A(L)	D(L)	B(X)	F(Y)	G(Y)	C(Y)	E(Y)	H(Y)
Long Name	Parameter names	parameter label	Par	tg	tr	rc	tau_g	F
Units				s	s	m		
Comments				your data	your data	$r_c[m]=\frac{6.18\times10^{-8}}{p[mmHg]}$	$\tau_g=\frac{2\eta x}{\rho g r_c^2}$	$F=\frac{1}{f_c^{3/2}}\approx 1$
1	Viscosity of air(kg/ms) (25oC)	η	1.8478E-5	7.455	7.91327			
2	Temperature coefficient of viscosity	Δη/ΔT	4.8E-8	15.56521	16.7815			
3	Density of oil (kg/m^3)	ρ1	886	23.07825	31.8955			
4	Density of air (kg/m^3)	ρ2	1.29	20.14243	11.70129			
5	Density difference (kg/m^3)	ρ1−ρ2	884.71	26.97377	22.47531			
6	acceleration due to gravity (m/s^2)	g	9.801	16.34362	16.44208			
7	ambient pressure (mmHg)	p	765	25.93429	25.02886			
8	fall/rise distance (m)	x	0.00145	15.34338	9.27446			
9	plate separation (m)	d	0.00317	29.3815	19.6161			
10	Voltage across the plates (V)	V	500	26.0786	24.3434			
11	Air temperature (oC)	Ta	20	--	--			
12	Actual air viscosity	[illegible]	1.8478E-5	--	--			
13			--	--	--			

Paste these 5 parameters and raw data from Section L1-L4.opj projects

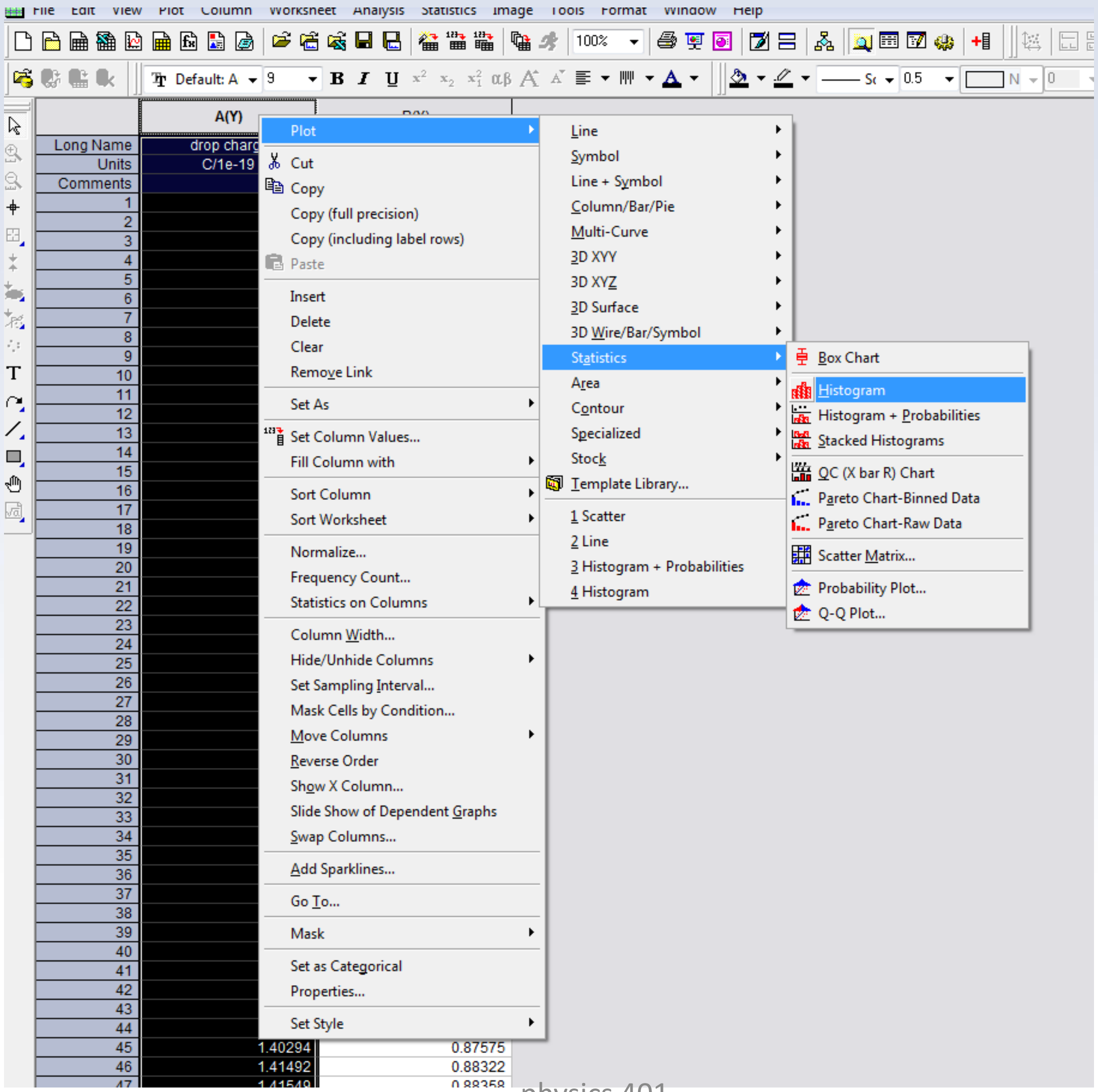
Calculate manually the actual air viscosity

Appendix #1. Analyzing of the statistical data.

Millikan oil drop experiment

Step 3. Histogram graph

First use the data from the column with drop charges and plot the histogram



illinois.edu

Appendix. Analyzing of the statistical data.

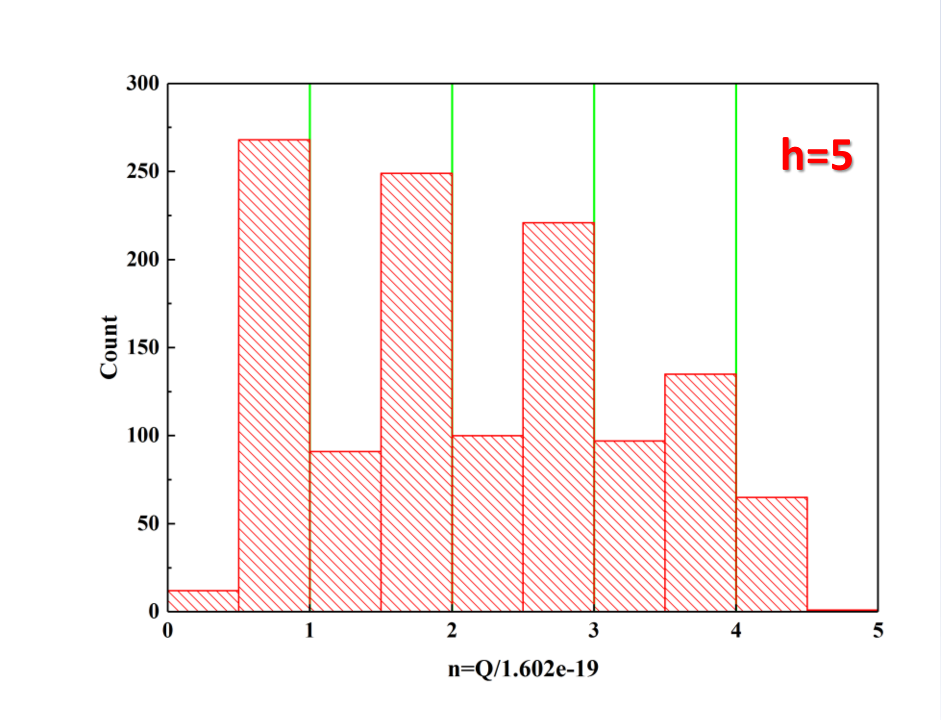
Millikan oil drop experiment

Step 4. Histogram. Bin size

Origin will automatically but not optimally adjust the bin size h. In tis page figure **h=0.5**. There are several theoretical approaches how to find the optimal bin size.

$$h = \frac{3.5\sigma}{n^{1/3}}$$

σ Is the sample standard deviation and **n** is total number of observation. For presented in Fig.1 results good value of **h ~0.1**

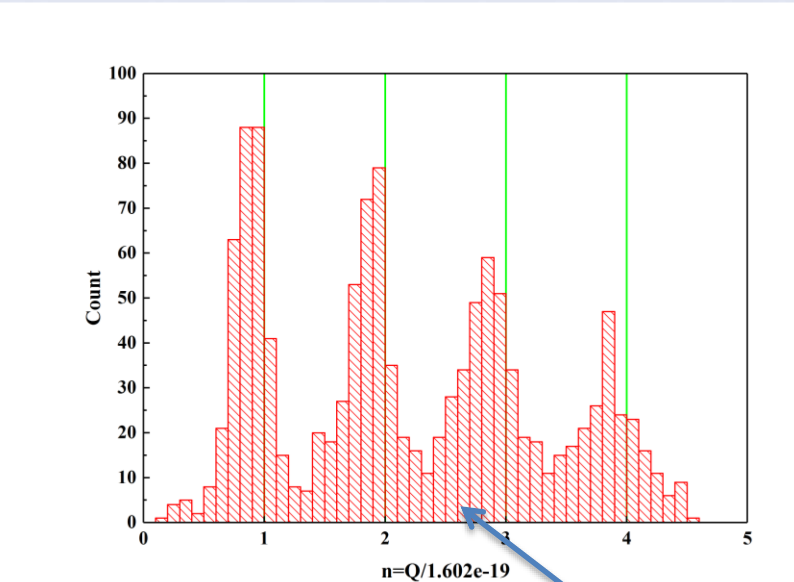


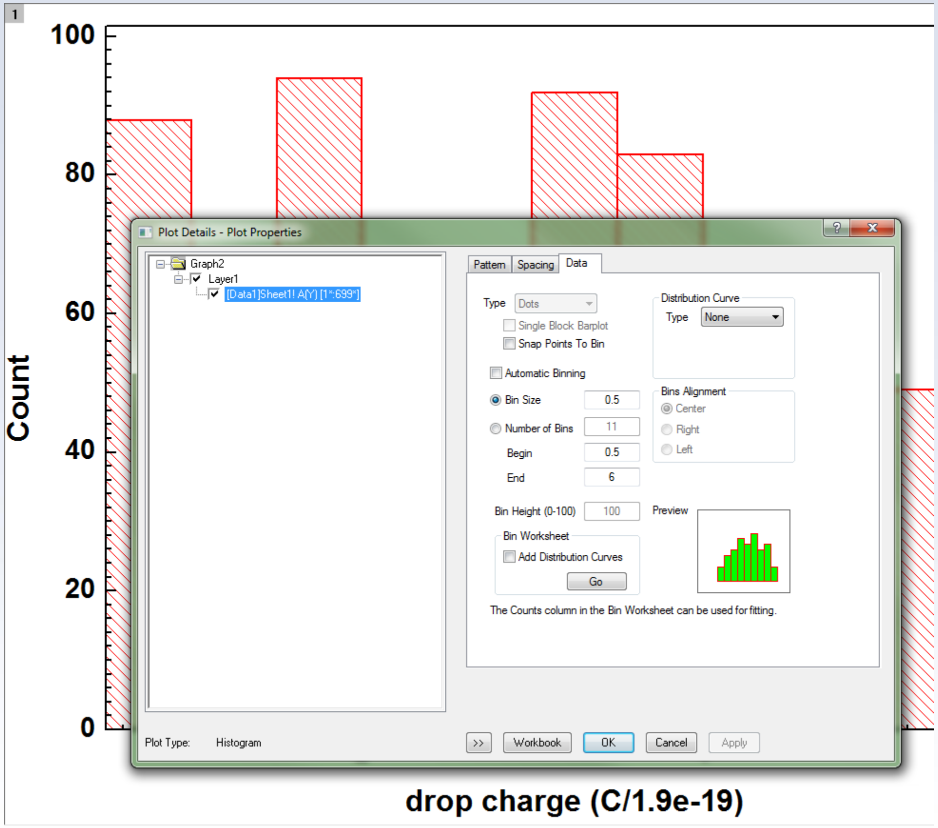
Appendix #1. Analyzing of the statistical data.

Millikan oil drop experiment

Step 4. Histogram. Bin size

To change the bin size click on graph and unplug the "*Automatic Binning*" option





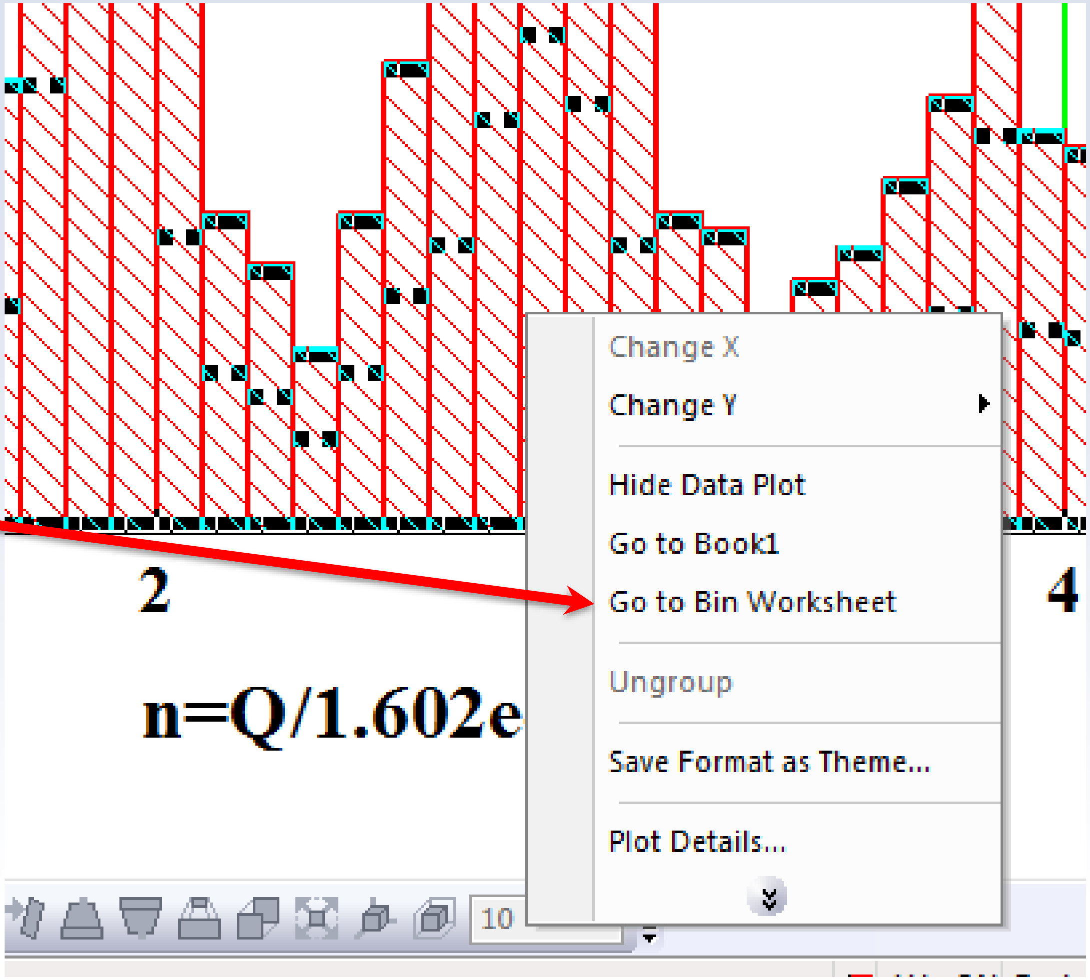
Bin size in this histogram is 0.1

Appendix #1. Analyzing of the statistical data.

Step 4. Find the bin Worksheet

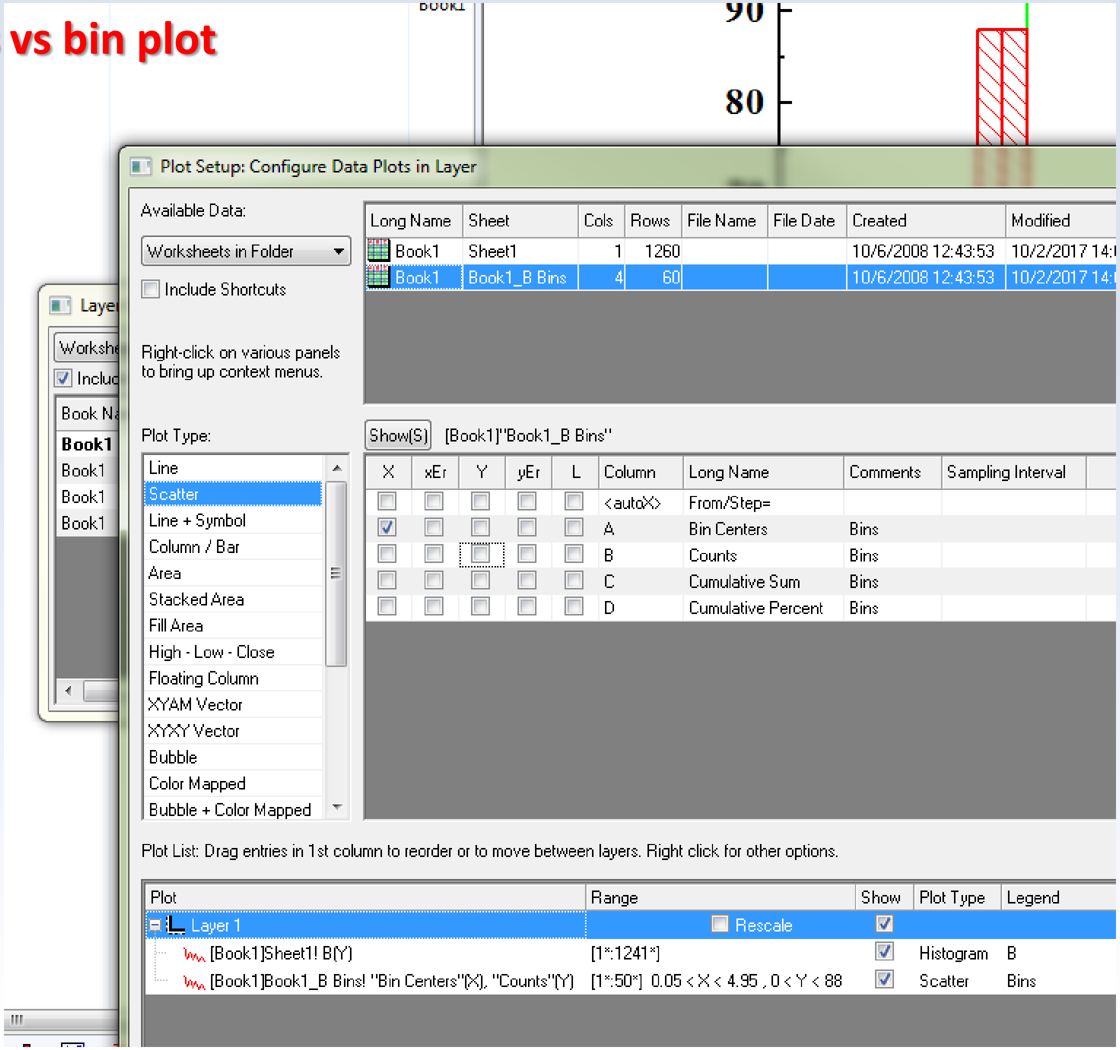
Millikan oil drop experiment

Right click on histogram and choose "*Go to Bin Worksheet*".



Appendix #1. Analyzing of the statistical data.

Step 5. Add Counts vs bin plot

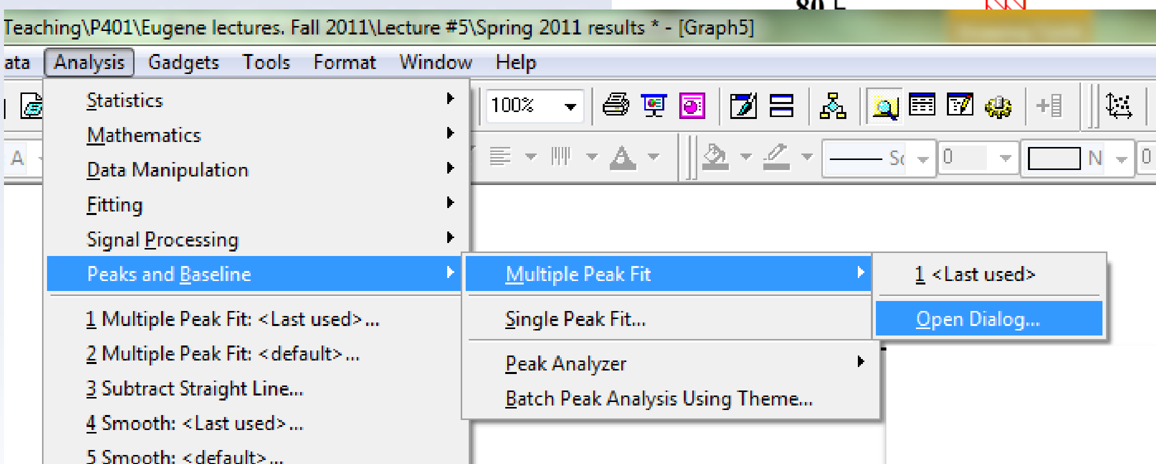


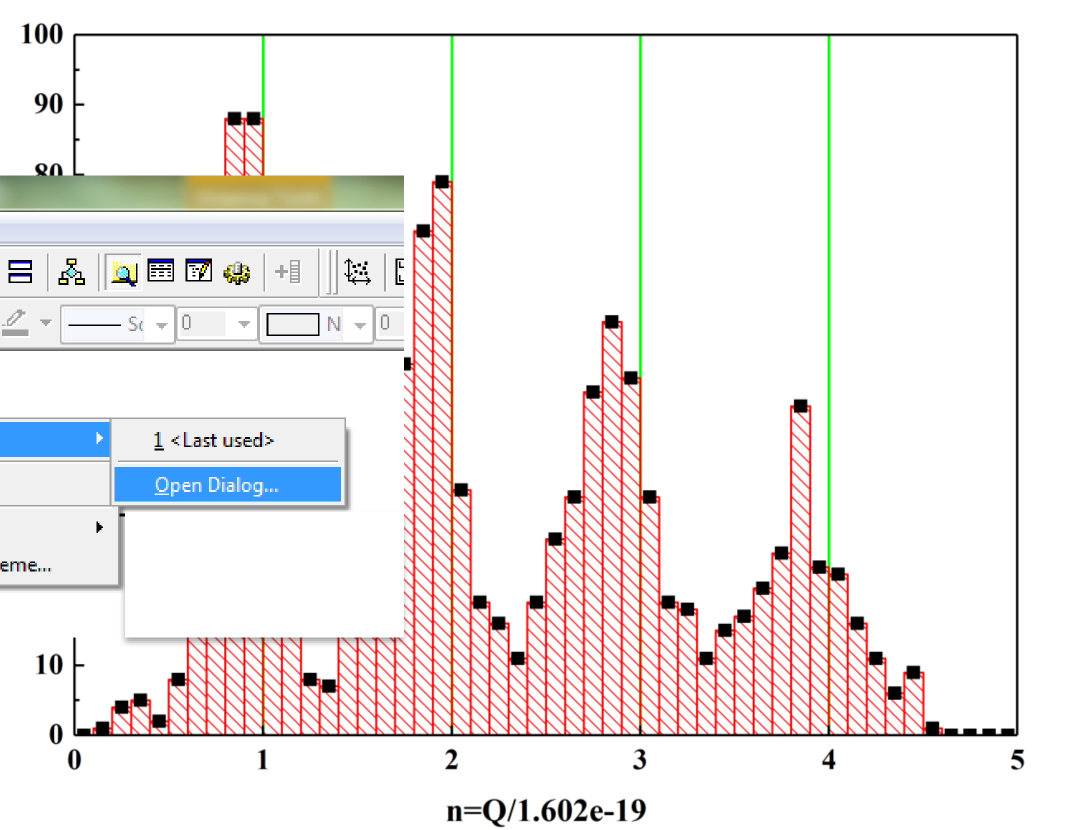
Appendix #1. Analyzing of the statistical data.

Millikan oil drop experiment

Step 5. Multipeak Gaussian fitting

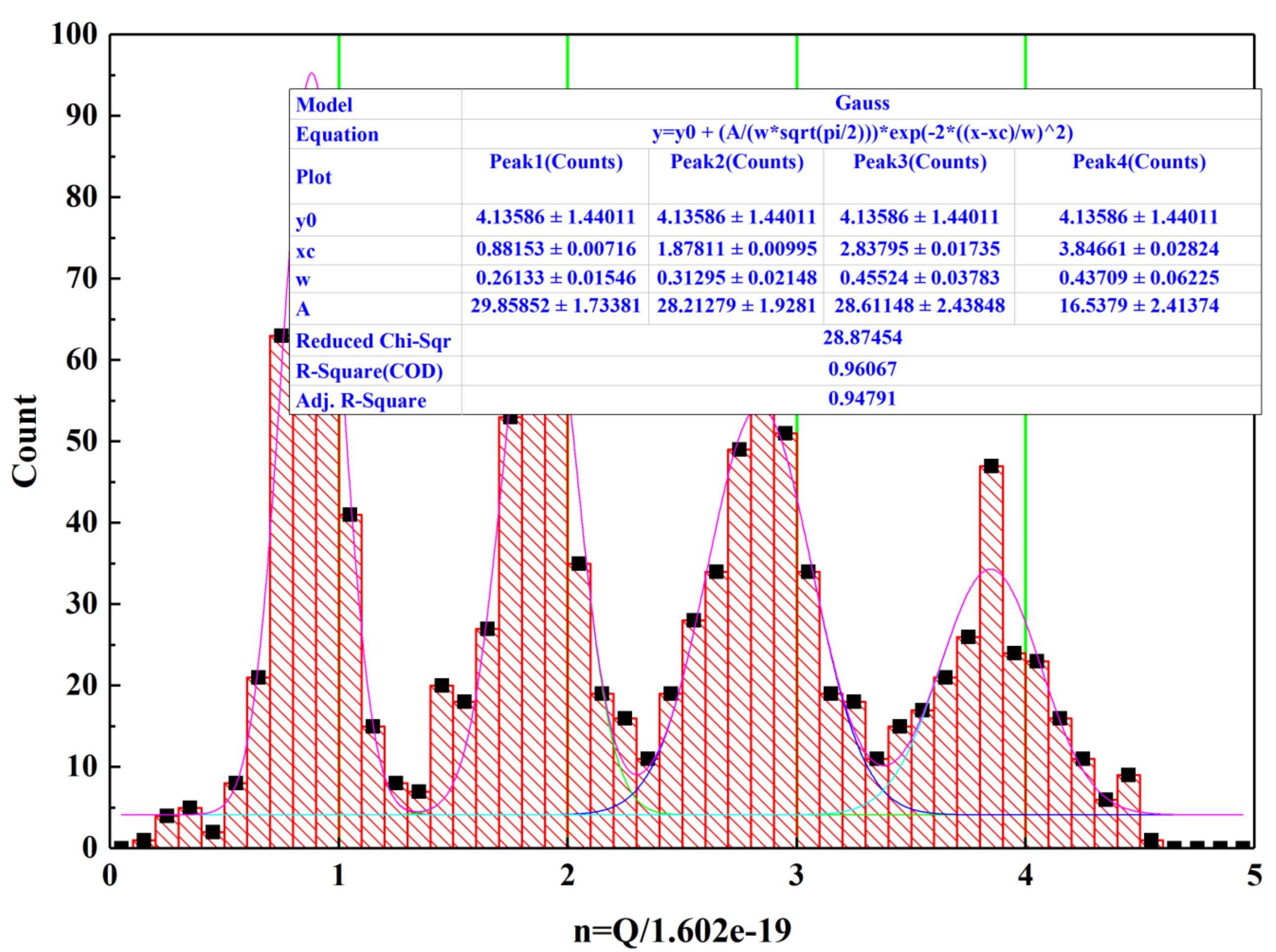
This plot can be used for peak fitting.





Appendix #1. Analyzing of the statistical data.

Step 5. Multipeak Gaussian Fitting

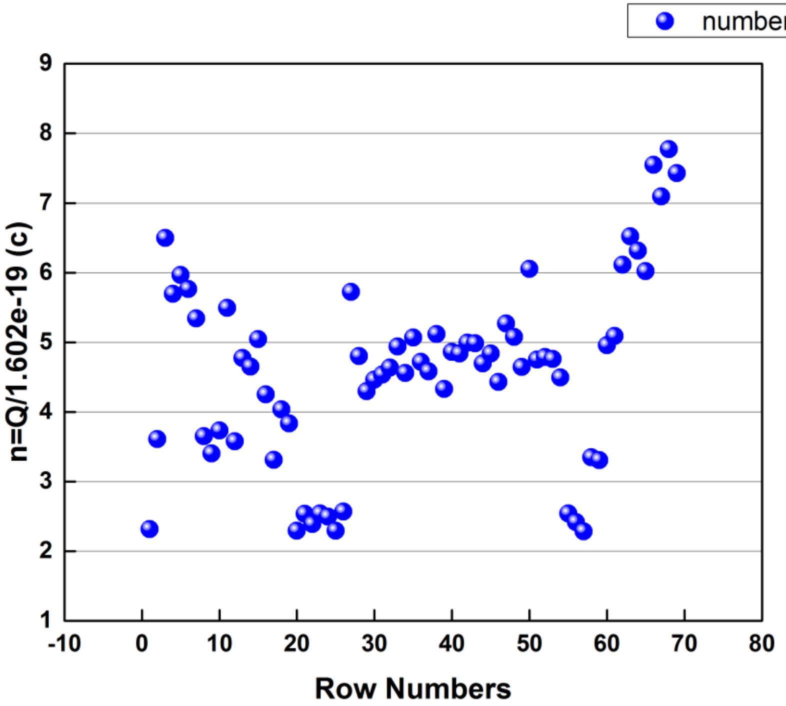


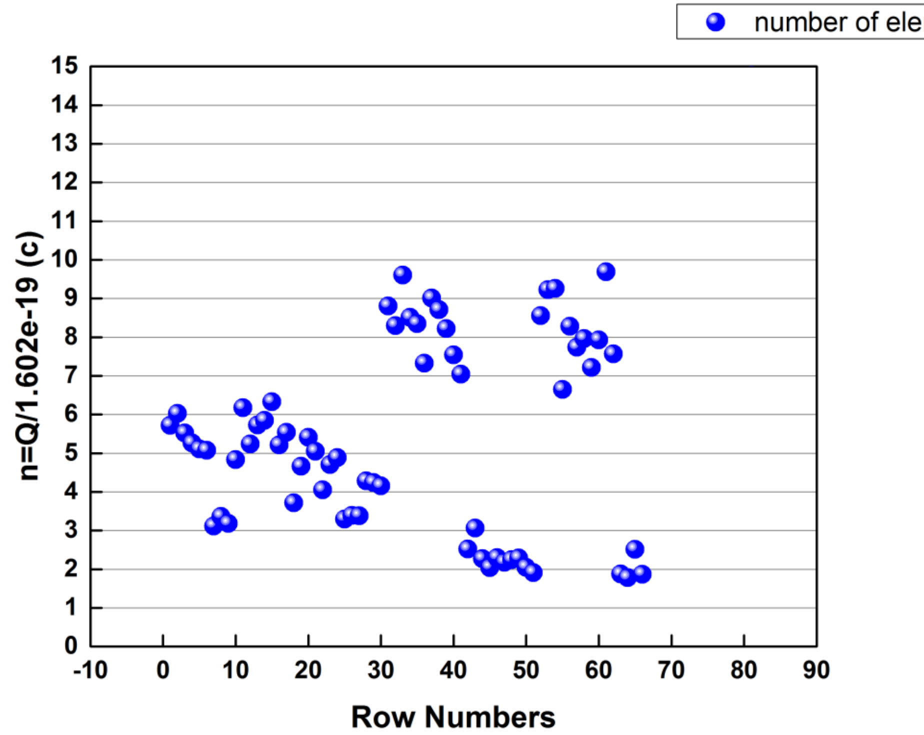
Model	Gauss			
Equation	y=y0 + (A/(w*sqrt(pi/2)))*exp(-2*((x-xc)/w)^2)			
Plot	Peak1(Counts)	Peak2(Counts)	Peak3(Counts)	Peak4(Counts)
y0	4.13586 ± 1.44011	4.13586 ± 1.44011	4.13586 ± 1.44011	4.13586 ± 1.44011
xc	0.88153 ± 0.00716	1.87811 ± 0.00995	2.83795 ± 0.01735	3.84661 ± 0.02824
w	0.26133 ± 0.01546	0.31295 ± 0.02148	0.45524 ± 0.03783	0.43709 ± 0.06225
A	29.85852 ± 1.73381	28.21279 ± 1.9281	28.61148 ± 2.43848	16.5379 ± 2.41374
Reduced Chi-Sqr	28.87454			
R-Square(COD)	0.96067			
Adj. R-Square	0.94791			

For 1st peak x_c~0.882±0.007

Appendix #1. oil Drop Data Issue.

Be careful with data selection obtained by different teams!





For more details how to create the histogram plot and do the analysis see "*Working with Histogram Graph. Millikan Oil Drop Experiment*"

Appendix #1. Analyzing of Oil Drop experiment Errors . How to Increase the Accuracy of the Experiment.

$$Q_{meas} = Q_{true} + e_s + e_r$$

e_s – systematic error: can be reduced by more accurate knowledge of parameters of the experiment like x, d, V, temperature etc. (usually it is limited to the exsisting measuring equipment)

e_r – random or statistical error: can be reduced by only by increasing of the number raw data point (no limits)

Appendix #2. Fitting. Main Idea.

(x_i, y_i) is an experimental data array. **x_i** is an independent variable and **y_i** - dependent

$f(x,\beta)$ is a model function and β is the vector of fitting (adjustable) parameters

The goal of the fitting procedure is to find the set of parameters which will generate the function ***f*** closest to the experimental points.

To reach this goal we will try to minimize the sum of squared deviation function (**S**):

$$S(\beta) = \sum_{i=1}^{m} \left[y_i - f(x_i, \beta) \right]^2$$

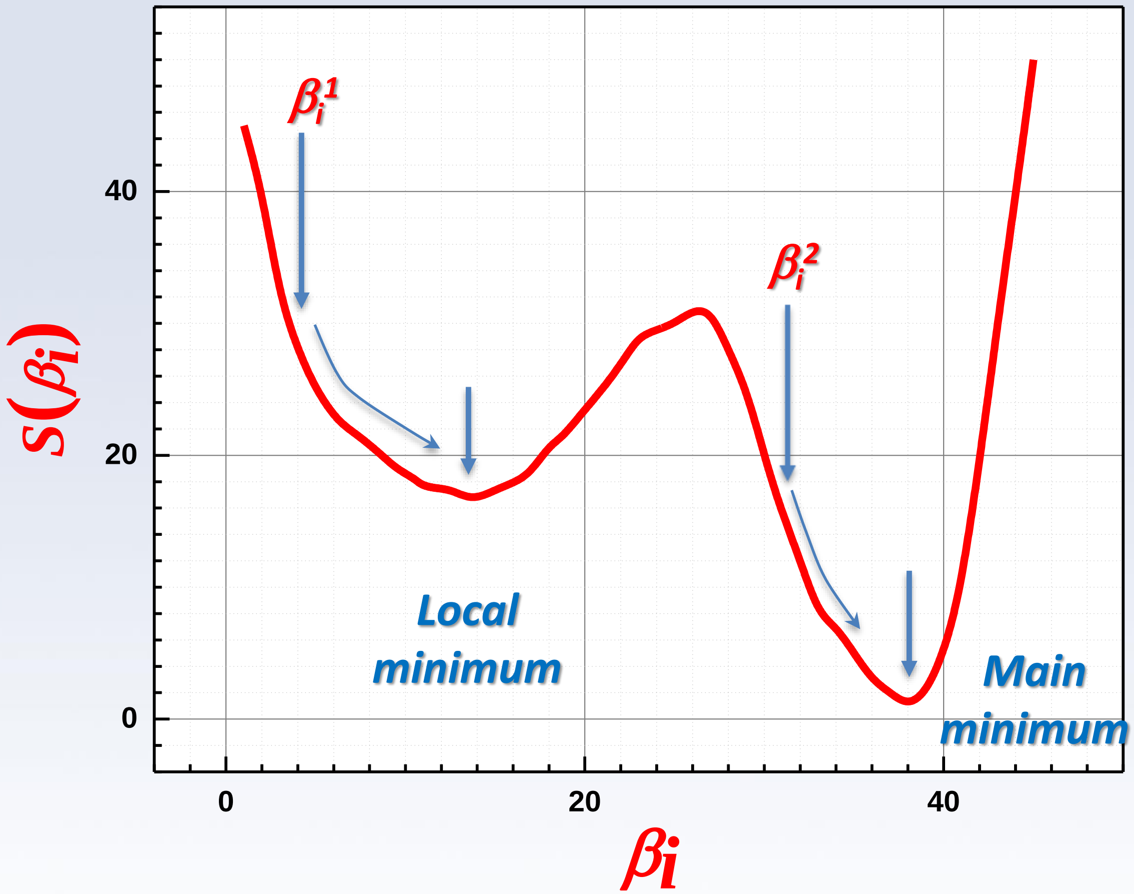
Appendix #2. Fitting. The Choice of Parameters.

The goal of fitting is not only to find the curve best matching the experimental data but also to find the corresponding parameters which in majority cases are the important physical parameters

There are several known mathematical algorithms for optimizing these parameters but in general the fitting procedure could have not only unique solution and the choice of initial parameters is very important issue

$$S(\beta) = \sum_{i=1}^{m} \left[y_i - f(x_i, \beta) \right]^2$$

Appendix #2. Fitting. The Choice of Parameters.



Let we have the ***S*** function dependent on parameter β_i as shown on this graph