Modern Physics

Why do we do experiments?
Introduction to data analysis

Laszlo Mihaly, 2022 Spring

Textbook:
L. Lyons, “A Practical Guide to Data Analysis for Physical Science Students”

Partially based on slides by Prof. Joanna Kiryluk, Prof. Giacinto Piacquadio and graduate students Darin Mihalik & Jonathan Pachter

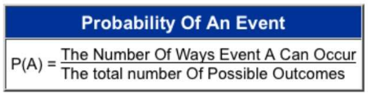
---

Experiment, Outcome, Event, Probability

- ❑ An **experiment** is a situation involving chance or probability that leads to results called outcomes.
- ❑ The **outcomes** are the possible results of a repeated experiments.
- ❑ An **event** is one possible outcome.
- ❑ The **probability** is the measure of how likely an event is.

- ❑ The **experiment** is throwing a dice.
- ❑ The **outcomes** are the top face showing 1 or 2 or .. Or 6 dots
- ❑ An **event** is when you get 3.
- ❑ For a fair dice, the **probability** of getting a 3 is ???

In order to measure probabilities, mathematicians have devised the following formula for finding the probability of an event.



Probability Of An Event

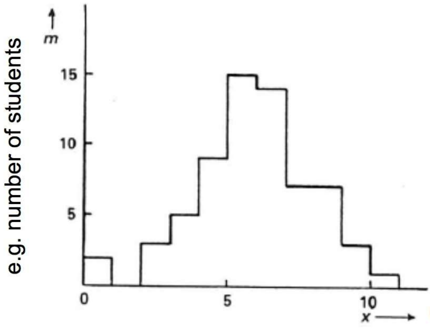
$$P(A) = \frac{\text{The Number Of Ways Event A Can Occur}}{\text{The total number Of Possible Outcomes}}$$

$$0 \leq P(A) \leq 1$$

- The probability of an event is the measure of the chance that the event will occur as a result of an experiment.

Discrete data/results

If data comes in discrete number → **Histogram**



e.g. number of students

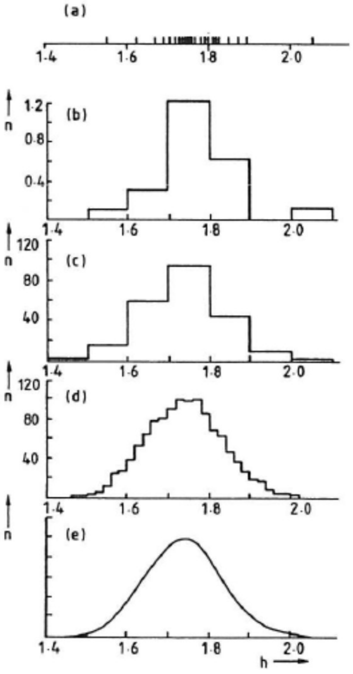
e.g. quiz scores

x in this histogram represents a range of experimental result between x and x+ Δ, where Δ is the bin size. x increases in steps of Δ.

y is the number of events (experimental results falling in that range

---

Data continuous: Measure the height of 30 year old men



(a) each experiment represented by a bar → difficult to visualize distribution

Binning: count how many events fall within a certain range. In (b) and (c) the bin size is 0.2m, but (c) has much more events. The more the events (data), the finer it can be binned.

In order to make the height of the histogram basically independent on the binning, choose Y scale properly.

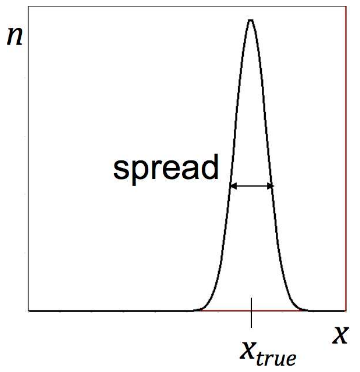
Number of events in each bin
→ Use number of events in each bin, divided by bin length (~dN/dh)

$N_{exp} \to \infty$

(e) In the limit of an infinite number of experiments, a continuous probability distribution f(h) is obtained

Continuous probability distribution (pdf)

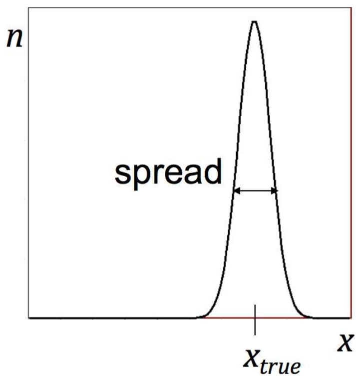
- Can be a good approximation already when the number of performed experiments is large!
- Let's see how such distribution typically looks like:



- The experimental measurements are typically spread around the true value x_{true}, that we'd like to measure.

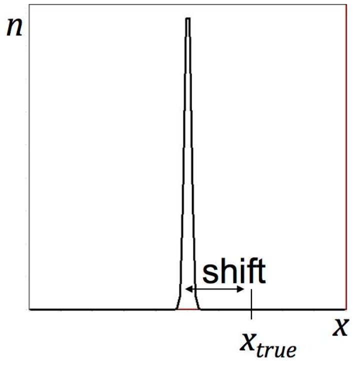
Type of uncertainties: Random

Continuous distribution (infinite number of measurements)



- Statistical uncertainties: arise from the inherent statistical nature of the phenomena being observed, for example, nuclear decay experiments and/or limited instrumental precision, for example the fifth digit of the voltmeter fluctuates randomly.)
- A series of repeated measurements results in parameters "x" randomly distributed around the true value we want to measure "x_{true}"
- May be handled by the theory of statistics

Type of uncertainties: Systematic

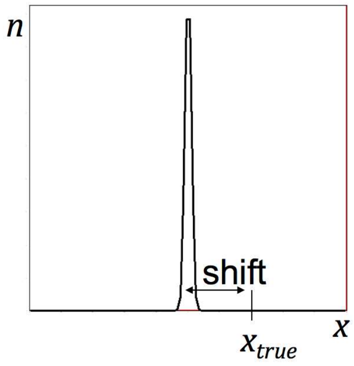


- Comes from a possible **bias** of the experimental result from the true value we want to measure
- E.g. a series of repeated experiments results in measurements that are **systematically** shifted in the same direction by the same amount from the true value
- Can't be cured by accumulating more data
- Possible sources of this uncertainty are typically difficult to identify.

How to avoid/reduce them?

(1) Ensure apparatus is properly calibrated and zeroed
(2) No simple rule for eliminating systematic errors: good theory knowledge + common sense + experience!

Type of uncertainties: Mistakes



Similar to systematic uncertainties in nature

It somewhat differs from the systematic uncertainties since you don't expect it, and thus typically don't associate an error to it

Example 1:
Writing 2.34 kHz instead of 2.43 kHz in your lab book. If not immediately corrected, it will affect the correctness of the result.

Other examples:
Misreading scales, confusion of units, a physics effect you forgot to consider, etc.

A good experimentalist avoids such mistakes by careful cross-checks: e.g. understand step-by-step if results are in line with expectations, use multiple methods to verify them and their systematic errors, etc.

Mistakes

Media and Press Relations

OPERA experiment reports anomaly in flight time of neutrinos from CERN to Gran Sasso

23 Sep 2011

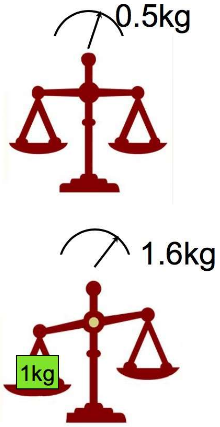
Mistakes can happen even to senior scientists

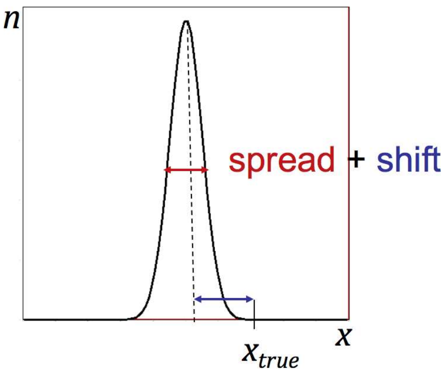
However, no result is accepted in the scientific community before further careful cross-checks (especially if it violates a cornerstone of physics as the Special Theory of Relativity)
At the end, the original authors of the study found out their mistake (a loose cable!). Nevertheless, such mistakes can cost a lot in terms of career! Learn how to avoid them!

Embattled neutrino project leaders step down

No-confidence vote follows confirmation of faults in experiment's cable and clock.

Most realistic situation:
random and *systematic* uncertainties





x can have a meaning of any measured quantity (e.g. box weight, acceleration due to gravity, etc.)

Characteristics of a distribution

Sample mean $\bar{x} = \sum x_i / N$
(central value of distribution)

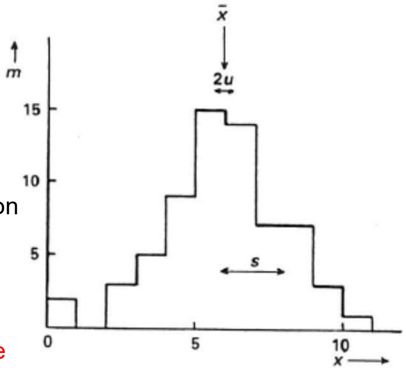
Sample variance $s^2 = \sum (x_i - \bar{x})^2 / (N-1)$
(width of distribution). Cannot work for N=1

A single measurement out of this distribution has an uncertainty of *s*.

Result of an individual measurement: $\bar{x} \pm s$

Result of many measurements $\bar{x} \pm u$, where

$$u^2 = \frac{\sum (x_i - \bar{x})^2}{N(N-1)} = \frac{s^2}{N}$$



One entry (x) in this histogram means one measurement

More on this later

Sloppy wording, but common: Uncertainty = error

How to present final experimental results → proper rounding

Incorrect: (1.89999679 ± 0.00346) [m]

How to write it correctly?

1. Look at the uncertainty: 0.00346 and then round it to 2 most significant digits. If the 3rd digit is ≥ 5 then the 2nd significant number must be increased by 1, i.e. 0.00346 ~ 0.0035.
2. Round the measurement itself such that the number of decimal digits is the same as for the (rounded) uncertainty

If the uncertainty is 0.0035, then it does not make sense to keep as many numbers in the 1.89999679 as possible. Numbers in purple are not significant.

Correct: (1.9000 ± 0.0035) [m]
1.9000(35) [m]
(19000 ± 35) x 10^{-4} [m]
19000(35) x 10^{-4} [m]

How to present final experimental results → proper rounding

Important:
In Lab Reports some points will be subtracted if rounding is not done properly!

Exercises:

A. (1.9 +/- 0.189) [m]
B. (1.89999679) +/- 0.189 [m]
C. (1.90 +/- 0.19) [m]
D. (1.9 +/- 0.2) [m]

E. (23.24555 +/- 2.234) [m]
F. (23.2 +/- 2.2) [m]
G. (23 +/- 2) [m]

H. (0.00012378 +/- 0.00000568) [m]
I. (0.0001238 +/- 0.0000057) [m]
J. (0.000124 +/- 0.000006) [m]
K. (1.24 +/- 0.06)x10^{-4} [m]
L. 1.24(6) x10^{-4} [m]

Which are correct and which are incorrect?

Probability interpretations

- “It is possible for an exp. physicist to spend a lifetime analyzing data without realizing that there are two different fundamental approaches to statistics” L. Lyons

1. Relative frequency (frequentism)
   A and B are outcomes of a repeatable experiment

$$\boxed{P(A) = \lim_{N\to\infty} \frac{times\ outcome\ is\ A}{N}}$$

Most common in Experimental Physics (this course!)

   e.g. particle scattering, radioactive decay

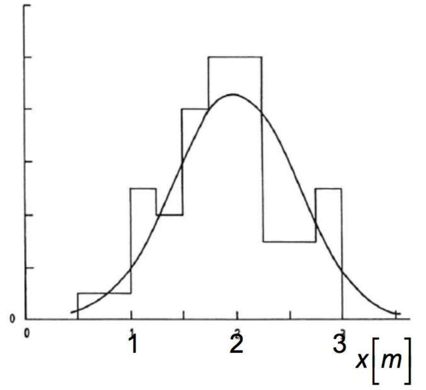
2. Subjective probability (bayesian)
   A, B are hypothesis (statements that are true or false)

$$\boxed{P(A) = \text{degree of belief that } A \text{ is true}}$$

   answers more directly the question we are interested in, but additional dependence on “prior belief” (e.g. 30% chance of rain tomorrow)

Probability distribution functions, expectation values and moments

Multiple measurements: distribution



Continuous line is a known function, so called Probability Density Function (PDF)

For $N \to \infty$ the histogram approaches the PDF

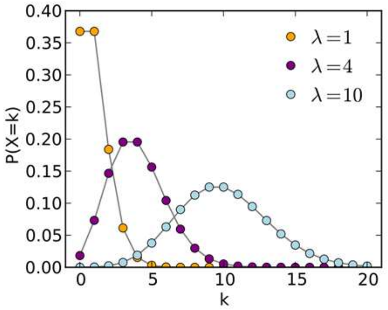
Many PDFs exist , but a large number of problem in physics are described by a small number of theoretical distributions

Binomial, Poisson, Gaussian PDFs - most common in experimental physics. See Appendix 3 and 4 of the textbook (L. Lyons)

Poisson distribution

Typical for counting experiments, for example nuclear decay rate measurement. For example, we record the number of clicks of a GM counter for 60 seconds and repeat that measurement many times. The probability distribution will be

$$P(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$



Here λ is the expected number of counts, and *k* is the actual measured number. The average of the measured *k* values, $\bar{k} = \frac{\sum k_i}{N}$ is our best guess for the value of λ. The variance is $s^2 = \bar{k}^2$. Accordingly, the result of a measurement should be reported as

$$\bar{k} \pm \sqrt{\bar{k}}$$

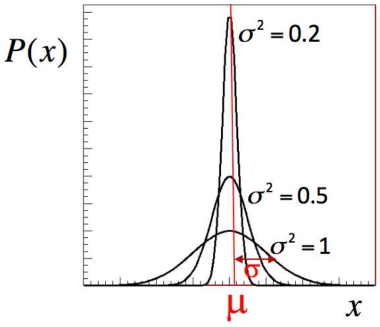
Measuring for longer time means larger $\bar{k}$. For $\bar{k} \to \infty$ the distribution approaches a Gaussian centered around λ and the variance is $\sigma = \sqrt{\lambda}$.

For the rest of the discussion we will focus on the Gaussian PDF.

The Gaussian Distribution

The Gaussian (also called "normal") PDF also plays a central role in all of statistics, and thus in science. Even in cases where its application is not strictly correct, the Gaussian often provides a good approximation to the true PDF. It is defined as:

$$P(x) = y = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Described by two parameters: *μ, σ*
For large *N*, $\bar{x} \to \mu$ and $s \to \sigma$

Expectation value (mean): E[*x*] = *μ*
Variance: V[*x*] = σ^2
Standard deviation ("error"): *σ*
Relative error: *σ/ μ*

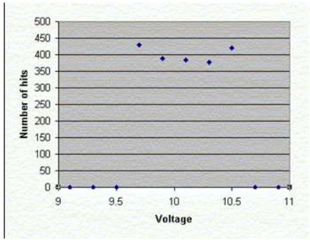
Why is Gaussian so important?

Central limit theorem: Consider any pdf. Repeat measurements *n* times. Take average. For *n* → infty the average will follow Gaussian distribution.

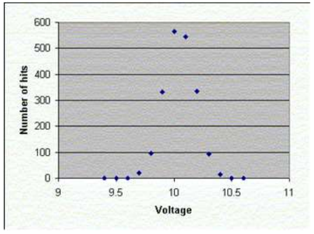
Example: throwing a single dice, then x={1,…,6} and P(x)=1/6 The pdf is uniform.

Throw *n* dice and calculate the average score. Repeat that *N* times and make the histogram. In the limit of *n, N* goes to infinity the average will follow a Gaussian distribution.

Works for any pdf. For example, a voltmeter reading randomly with this distribution:



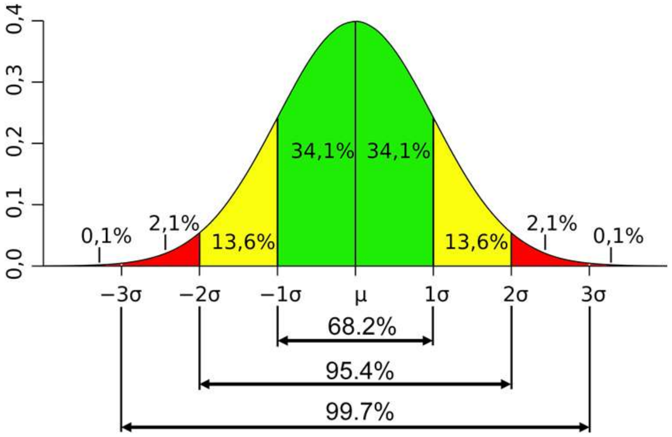
Average of 5 measurements:



http://solidstate.physics.sunysb.edu/teaching/2020_fall/phy251/labs/statmeth/random.xls
http://solidstate.physics.sunysb.edu/teaching/2020_fall/phy251/labs/statmeth/random5.xls

Characteristics of Probability Functions

The area under the Gaussian curve between integral intervals of σ is an important practical quantity



Expectation Values, Distribution Moments

These can be defined mathematically without specifying P(x)

Expectation value of x: $E[x] = \bar{x} = \int x\ P(x)dx$

More general: The *r*-th moment of *x* around *x₀* is:

$$E[(x - x_0)] = \int (x - x_0)^r\ P(x)dx$$

1st moment, if $x_0 = 0$

$$\boxed{Mean = \int x\ P(x)dx}$$

2nd moment

$$\boxed{Variance = \int x^2\ P(x)dx}$$

$$\boxed{Standard\ deviation = \sqrt{Variance}}$$

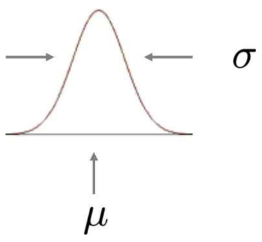
Expectation Values, Distribution Moments

For a Gaussian distribution $P(x) = y = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$$\boxed{Mean = \mu}$$

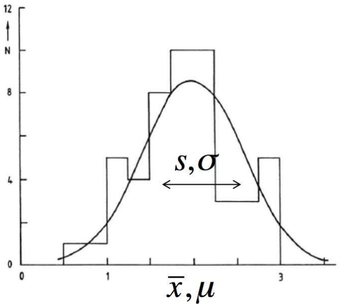
$$\boxed{Variance = \sigma^2}$$

$$\boxed{Standard\ deviation = \sigma}$$



Characteristics of the distribution

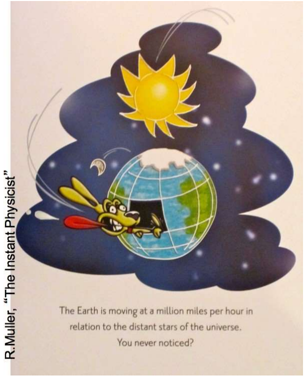
- Sample mean x
  → provides an estimate of the true value μ
- Sample variance s^2
  → estimate of the variance σ^2



One entry (x) in this histogram means one measurement

Notice:
s = uncertainty on a single measurement
u = uncertainty on the mean

Example



R.Muller, "The Instant Physicist"

- In an experiment consisting of 10 independent measurements, we measured the speed of Earth v_E in its revolution around the Sun and got the following results:

1. v_E= 29.7 [km/s]
2. v_E= 29.9 [km/s]
3. v_E= 29.9 [km/s]
4. v_E= **29.9** [km/s]
5. v_E= 29.8 [km/s]
6. v_E= 30.0 [km/s]
7. v_E= **29.7** [km/s]
8. v_E= 29.9 [km/s]
9. v_E= 29.8 [km/s]
10. v_E= 30.0 [km/s]

Questions:
What is the best estimate (and its uncertainty) for v_E ?
What is a single measurement uncertainty on v_E?

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i =$$
$$= \frac{1}{10}(29.7 + 29.9 + 29.9 + 29.9 + 29.8 + 30.0 + 29.7 + 29.9 + 29.8 + 30.0)\ [\mathrm{km/s}]$$
$$= 29.853394[\mathrm{km/s}]$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_u - \overline{x})^2$$
$$= \frac{1}{9}[(29.7-\overline{x})^2 + (29.9-\overline{x})^2 + (29.9-\overline{x})^2 + (29.9-\overline{x})^2 + (29.8-\overline{x})^2$$
$$+ (30.0-\overline{x})^2 + (29.7-\overline{x})^2 + (29.9-\overline{x})^2 + (29.8-\overline{x})^2 + (30.0-\overline{x})^2][\mathrm{km}^2/\mathrm{s}^2]$$
$$= 0.009456[\mathrm{km}^2/\mathrm{s}^2]$$

$$u^2 = \frac{s^2}{n} = \frac{0.009456}{10}[\mathrm{km}^2/\mathrm{s}^2] = 0.0009456[\mathrm{km}^2/\mathrm{s}^2]$$

$$u = 0.030751[\mathrm{km/s}] \approx 0.03[\mathrm{km/s}]$$

Result:

$$\boxed{\overline{v}_E \pm \sigma_{v_E} = \overline{x} \pm u = (29.85 \pm 0.03)\ [\mathrm{km/s}]}$$

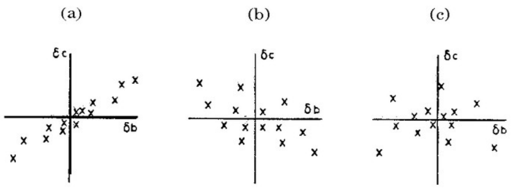
(1 significant digit)

Notice: a single measurement has an uncertainty s=√s² (not u!), i.e. each measurement of the previous page e.g. v = 29.7 ± 0.1

More than one variable – error propagation

Want to evaluate a quantity *c*, that depends on two variables, *a* and *b* in a very simple way: *a = b - c.* Assume a and b follow Gaussian PDF with σ_b and σ_c . What is the standard deviation (error) of *a*?

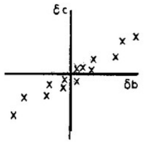
Two situations: The errors in *a* and *b* are correlated or uncorrelated. Plot the deviation from the average value for each measurement:



The calculation of the error of *a* is very different in the two cases.

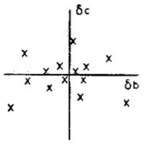
More than one variable – error propagation

Correlated errors



When b is off, c is also off in the same direction. Systematic error.
$\sigma_a = \sigma_b + \sigma_c$

Uncorrelated errors



When b is off c may be off in the same direction, or in the opposite direction. Truly random error.
$\sigma_a^2 = \sigma_b^2 + \sigma_c^2$

See book for proof.

Adding uncorrelated errors – general case

We calculate *f* that depends on measured parameters x_1, x_2 …. For any given measurement the deviation from the mean is δx_1 , δx_2 , ….

If all δx is zero except for δx_i , the deviation from the mean value of f is $\delta f = \frac{\partial f}{\partial x_i}\delta x_i$. The typical value of δx_i is σ_i, so the contribution to the error of *f* is $\sigma_f = \frac{\partial f}{\partial x_i}\sigma_i$

In the spirit of previous discussions, if the errors are uncorrelated, we obtain the error of *f* by

$$\sigma_f^2 = \sum \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_i^2$$

Works for any function, small errors only.

Combining different experiments

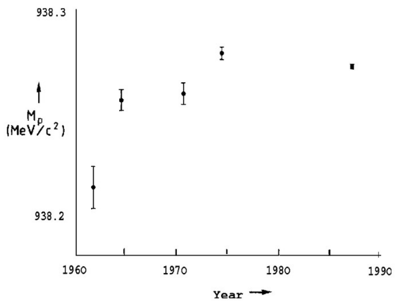
We do several measurements to determine the same quantity *a*. Each measurement has its own error and they may be different from each other. What is the best way to calculate the average? What is the error of the average?

$$\bar{a} = \frac{\sum a_i/\sigma_i^2}{\sum 1/\sigma_i^2} \qquad \frac{1}{\sigma^2} = \sum \frac{1}{\sigma_i^2}$$

Special case: all $\sigma_i = \sigma_0$ are equal, there are N measurements.

$$\sigma = \sigma_0/\sqrt{N}$$

Measured Proton mass



Least squares (χ^2) fit

Quantity *y* depends on *x*. For example, $y = ax + b$. We set the value of *x* to x_i (with no error) and measure the value y_i^{obs} and uncertainty σ_i, and repeat this several times. How can we determine the parameters *a* and *b*? What is their error?

First, for each x_i, calculate from the formula the corresponding y_i^{th}.

Calculate $\chi^2 = S = \sum \left(\frac{y_i^{th}(a,b) - y_i^{obs}}{\sigma_i} \right)^2$

Change the parameters *a* an *b* until this quantity reaches a minimum value.

Works for any function! Works for any number of parameters. We need (much) more measurements than the number of parameters we want to determine.

Least squares (χ^2) fit

Useful to:

- Fit (determine) parameters of a function
  - Determine the parameters (+ uncertainties) of a function that fits the data
    - Example: I measure the distance a car has moved at various different times, and I want to determine its velocity v = distance / time (the function assumes the car is traveling at constant speed)
- Hypothesis testing
  - Determine how compatible the data is with being described by a certain function
    - Example (following from above): I want to understand how compatible my data is with the hypothesis I made that the motion is happening at constant speed.

χ^2 >>1 means that there is systematic deviation from the data. For example you do a linear fit, but the actual dependence is quadratic.

Fitting straight lines is special (simple)

We do not need to go through a time-consuming minimization process. There is an algebraic solution. Take

$$y = A + Bx$$

To calculate the intercept A of the best fit line

$$A = \frac{1}{\Delta}(\sum \frac{x_i^2}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2})$$

where $\Delta = \sum \frac{1}{\sigma_i^2} \sum \frac{x_i^2}{\sigma_i^2} - (\sum \frac{x_i}{\sigma_i^2})^2$

and $B = \frac{1}{\Delta}(\sum \frac{1}{\sigma_i^2} \sum \frac{x_i y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2})$

The uncertainties are

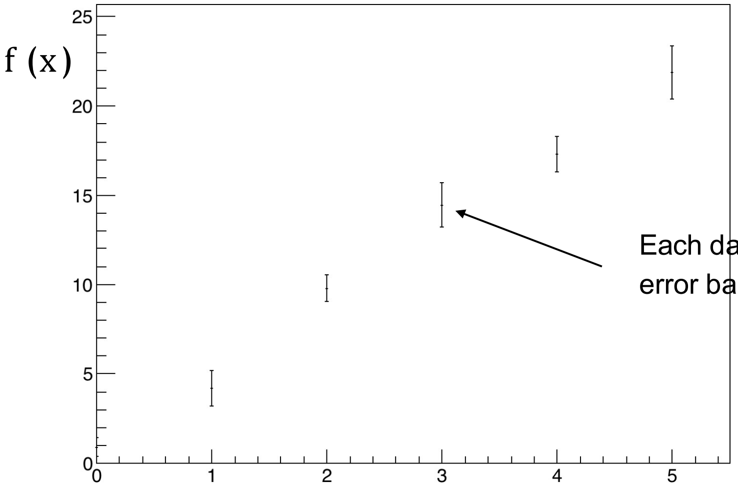
$$\sigma_A^2 = \frac{1}{\Delta} \sum (\frac{x_i^2}{\sigma_i^2}) \qquad \sigma_B^2 = \frac{1}{\Delta} \sum (\frac{1}{\sigma_i^2})$$

Implemented here

https://www.ic.sunysb.edu/class/phy141md/doku.php?id=phy131studio:labs:plottingtool

An example (II)

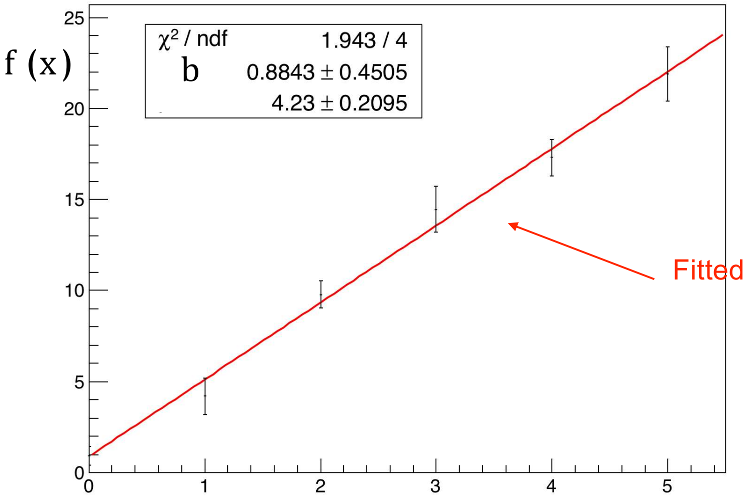
- First plot the data on a two-dimensional x y graph (y=f(x)):



Each data point $y_i=f_i$ has its error band σ_i.

An example (III)

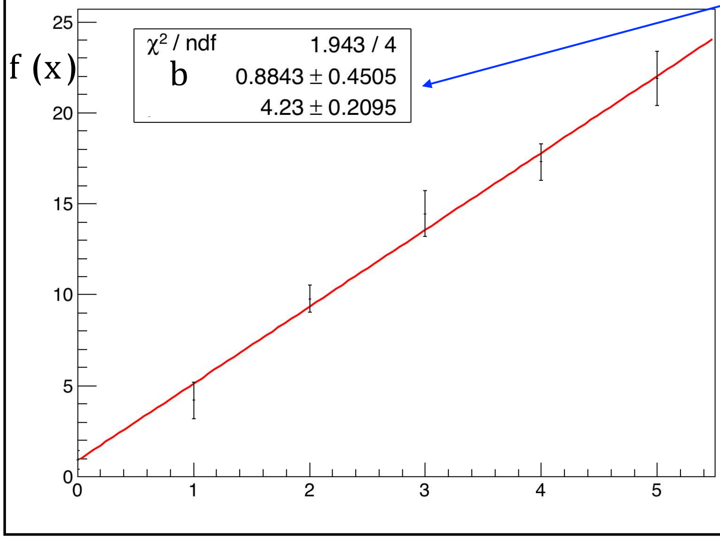
- Now fit a straight line to the data (y=ax+b), determine a and b:



Fitted line

An example (IV)

- Now fit a straight line to the data (y=ax+b), determine a and b:



Fitted parameters

Computer programs often do not round correctly. Please do in lab reports!

2 sign. digits:

a = 4.23 ± 0.21b

= 0.88 0.45

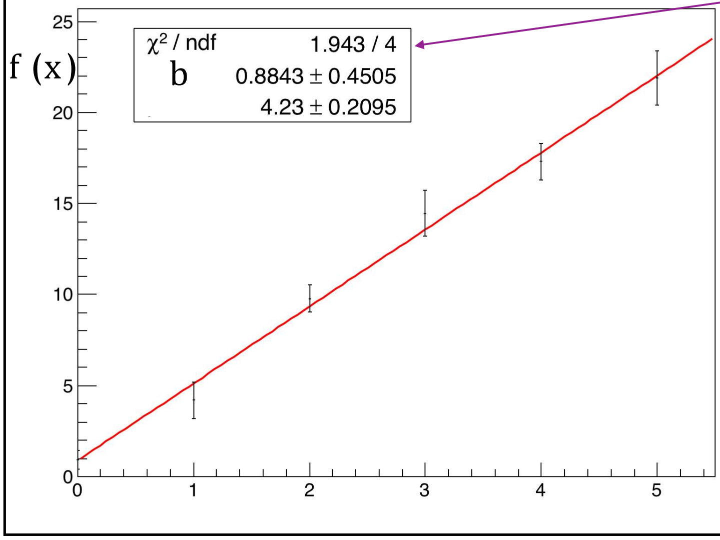
1 sign. digit:

a = 4.2 ± 0.2

b = 0.9 ± 0.5

An example (V)

- Now fit a straight line to the data (y=ax+b), determine a and b:



χ^2 and ndf value

- ndf (number of degrees of freedom) is the number of data points minus the number of fitted parameters
- χ^2 is the value of S at minimum

Another Example: Measure "g"

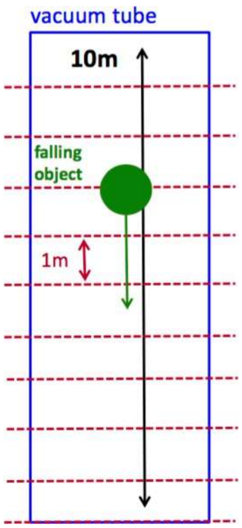
In this experiment we want to measure the acceleration due to gravity (or our hypothesis for the law governing the change of velocity per time)

$$h(t) = \frac{1}{2}gt^2$$

We therefore need to know the time it takes an object to travel a known distance under the influence of gravity.

Our experiment will consist of dropping an object from a specific height and recording the time from release until it hits the ground

Setup of the Experiment



- You will drop a massive object in a 10 meter long vacuum tube (neglect air resistance)
- You precisely know the position of the markers (no uncertainty in the position)
- You will measure the time from release to when the object passes each successive meter mark
- You measure time with a stopwatch and therefore, this measurement has uncertainty
- Each time measurement has the same uncertainty

$$\sigma_t = 0.05\ s$$

Recorded Data

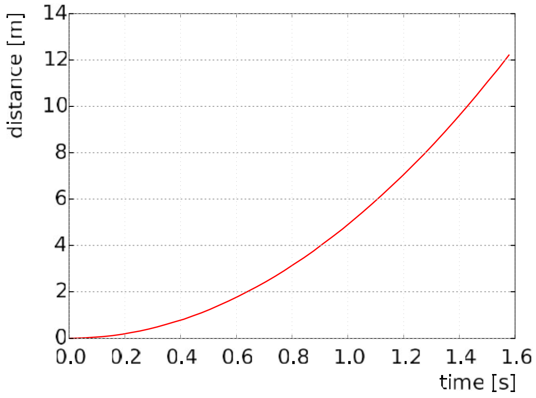
Distance [m]	Time [s]
0	0.0
1	0.43
2	0.6
3	0.73
4	0.91
5	1.05
6	1.1
7	1.15
8	1.26
9	1.26
10	1.47

distance is a quadratic function of time!

$$h(t) = \frac{1}{2}gt^2$$

↓

g = 9.81 m/s^2



This is a 1 parameter estimation problem

We have to calculate an estimate of g from our experimental data

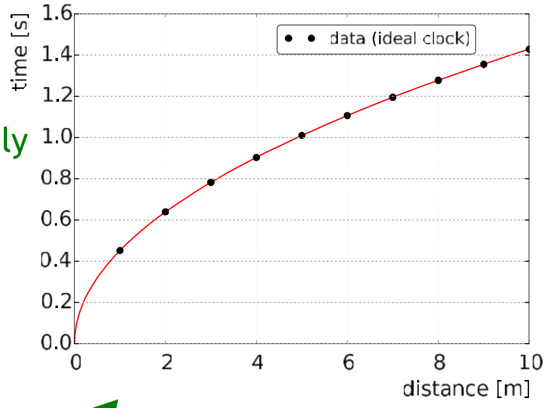
Remember:

In our experiment we actually measure time as function of distance! We have to invert

$$h(t) = \frac{1}{2}gt^2$$

To get

$$t(h) = \sqrt{\frac{2h}{g}}$$



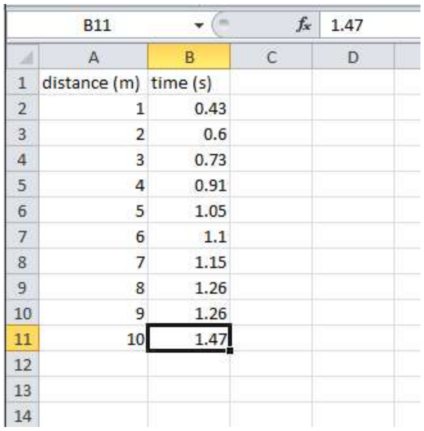
g=9.81 m/s^2

Assuming we know the value of g;
this is what we expect.

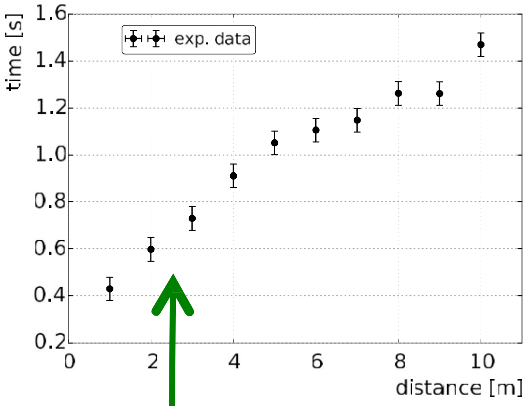
Let's turn on uncertainty

σ_t=0.05s

measured times t_i are now random variables



	A	B	C	D
1	distance (m)	time (s)		
2	1	0.43		
3	2	0.6		
4	3	0.73		
5	4	0.91		
6	5	1.05		
7	6	1.1		
8	7	1.15		
9	8	1.26		
10	9	1.26		
11	10	1.47		
12				
13				
14				



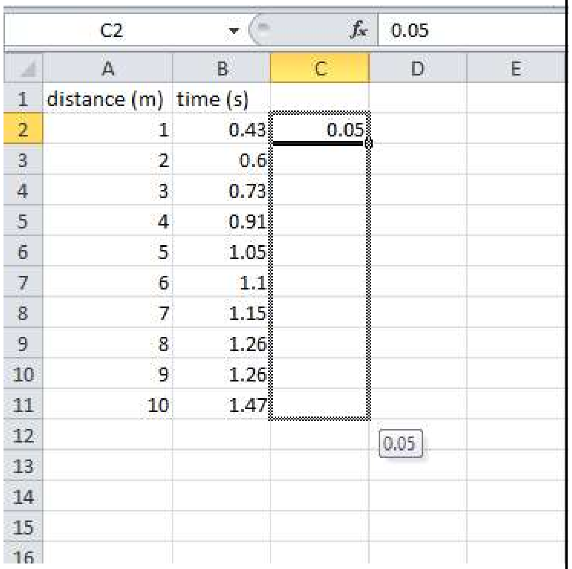
Actual measurement in experiment with uncertainty

We'll use the data set shown on the left, in Excel

Need column for uncertainty.

Two methods:
-Create whole column of identical values
-Refer to fixed box, with the one value 0.05

We'll do the former, for now



	A	B	C	D	E
1	distance (m)	time (s)			
2	1	0.43	0.05		
3	2	0.6			
4	3	0.73			
5	4	0.91			
6	5	1.05			
7	6	1.1			
8	7	1.15			
9	8	1.26			
10	9	1.26			
11	10	1.47			
12					
13					
14					
15					
16					

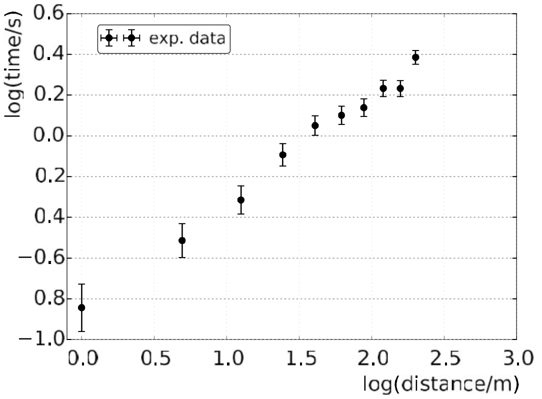
you know how to do a "straight line fit". Good!

But time is a non-linear function of distance.

$$t(h) = \sqrt{\frac{2h}{g}}$$

Make the problem linear using logarithms and algebra!

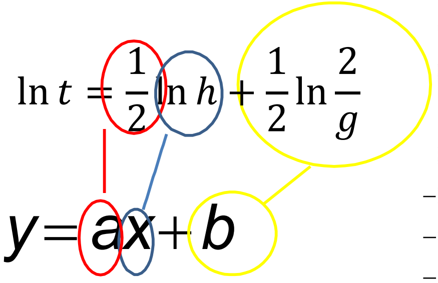
$$\ln t = \frac{1}{2}\ln h + \frac{1}{2}\ln\frac{2}{g}$$



$$\ln(ab) = \ln(a) + \ln(b)$$

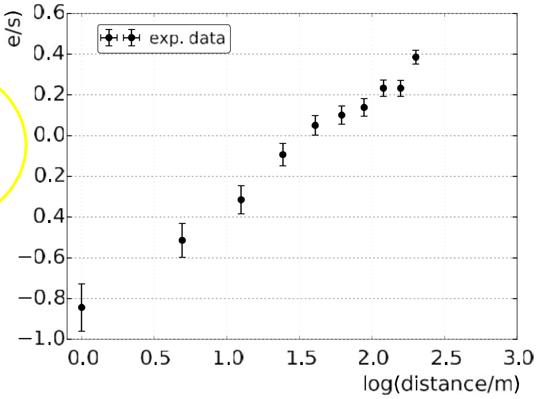
$$\ln(a^b) = b\ln(a)$$

Why can this be viewed as linear?



$$\ln t = \frac{1}{2}\ln h + \frac{1}{2}\ln\frac{2}{g}$$

$$y = ax + b$$



and identify

$$x = \ln(h)$$

$$y = \ln(t)$$

$$\longrightarrow$$

$$\boxed{\begin{gathered} a = \frac{1}{2} \\ b = \frac{1}{2}\ln\frac{2}{g} \\ g = 2\exp(-2b) \end{gathered}}$$

We can find g from the offset *b* of the straight line

$$\ln t = \frac{1}{2}\ln h + \frac{1}{2}\ln\frac{2}{g}$$

compare to

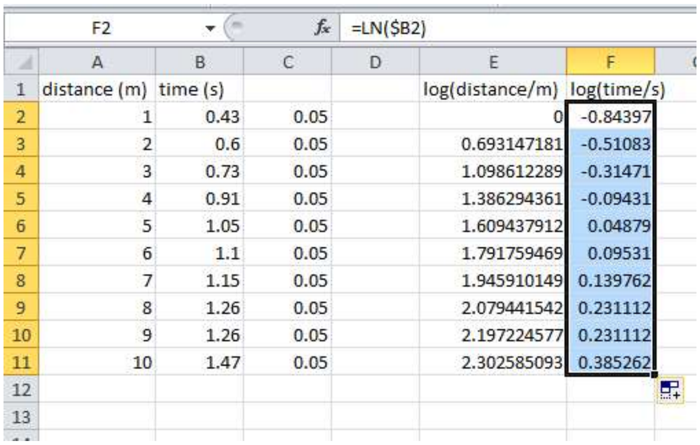
$$y = ax + b$$

Now compute:

$$x = \ln(s)$$

$$y = \ln(t)$$

$$o_s = 0 \rightarrow \sigma_x = 0$$



=LN($B2)

	A	B	C	D	E	F
1	distance (m)	time (s)			log(distance/m)	log(time/s)
2	1	0.43	0.05		0	-0.84397
3	2	0.6	0.05		0.693147181	-0.51083
4	3	0.73	0.05		1.098612289	-0.31471
5	4	0.91	0.05		1.386294361	-0.09431
6	5	1.05	0.05		1.609437912	0.04879
7	6	1.1	0.05		1.791759469	0.09531
8	7	1.15	0.05		1.945910149	0.139762
9	8	1.26	0.05		2.079441542	0.231112
10	9	1.26	0.05		2.197224577	0.231112
11	10	1.47	0.05		2.302585093	0.385262
12						
13						

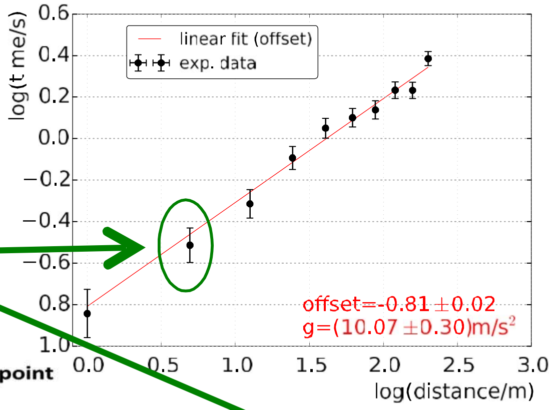
Need to find uncertainties on

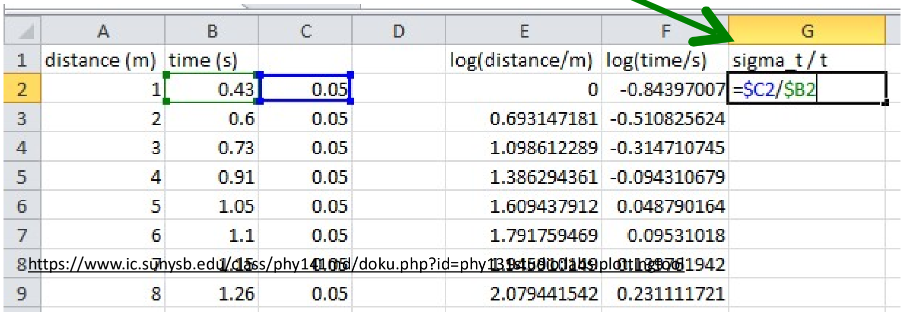
$$y_i = \ln(t_i)$$

from error propagation

$$\sigma_{y_i} = \left|\frac{\partial y_i}{\partial t_i}\right| \sigma_{t_i} = \frac{\sigma_{t_i}}{t_i}$$

Note: the subscript *i* refers to each data point





	A	B	C	D	E	F	G
1	distance (m)	time (s)			log(distance/m)	log(time/s)	sigma_t / t
2	1	0.43	0.05		0	-0.84397007	=$C2/$B2
3	2	0.6	0.05		0.693147181	-0.510825624	
4	3	0.73	0.05		1.098612289	-0.314710745	
5	4	0.91	0.05		1.386294361	-0.094310679	
6	5	1.05	0.05		1.609437912	0.048790164	
7	6	1.1	0.05		1.791759469	0.09531018	
8	7	1.15	0.05		1.945910149	0.139761942	
9	8	1.26	0.05		2.079441542	0.231111721	

https://www.ic.sunysb.edu/class/phy141md/doku.php?id=phy131studio:labs:plotting

Recall:

$$\ln t = \frac{1}{2}\ln h + \frac{1}{2}\ln\frac{2}{g}$$

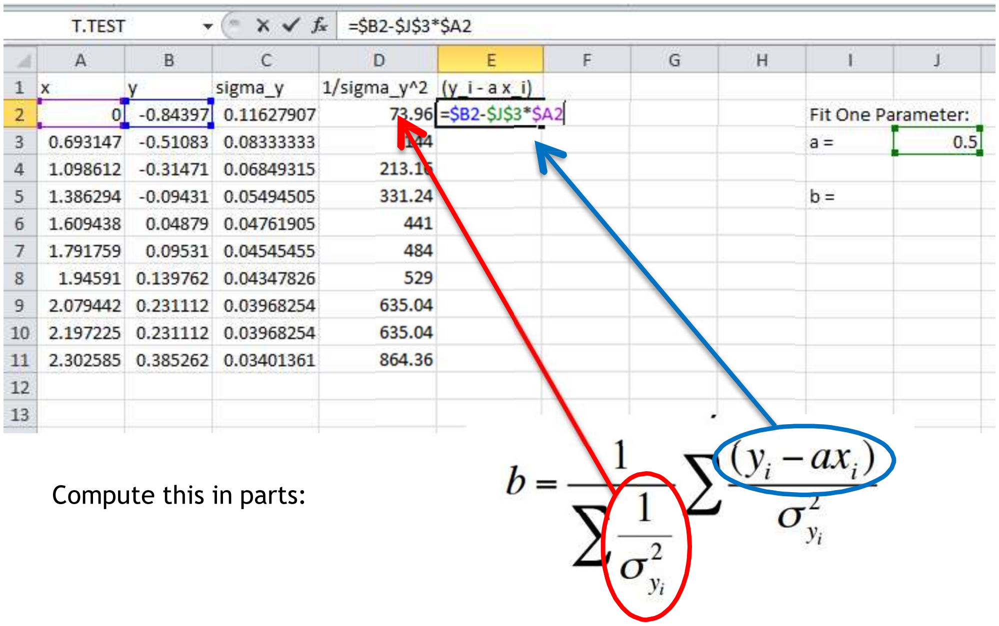
$$y = ax + b$$

Now we're just doing linear analysis on y, x, with a, and b

But *a* is fixed to be ½!

This is the case when we only fit one parameter b:

$$b = \frac{1}{\sum \frac{1}{\sigma_{y_i}^2}} \sum \frac{(y_i - ax_i)}{\sigma_{y_i}^2} \qquad \sigma_b = \sqrt{\frac{1}{\sum \frac{1}{\sigma_{y_i}^2}}}$$

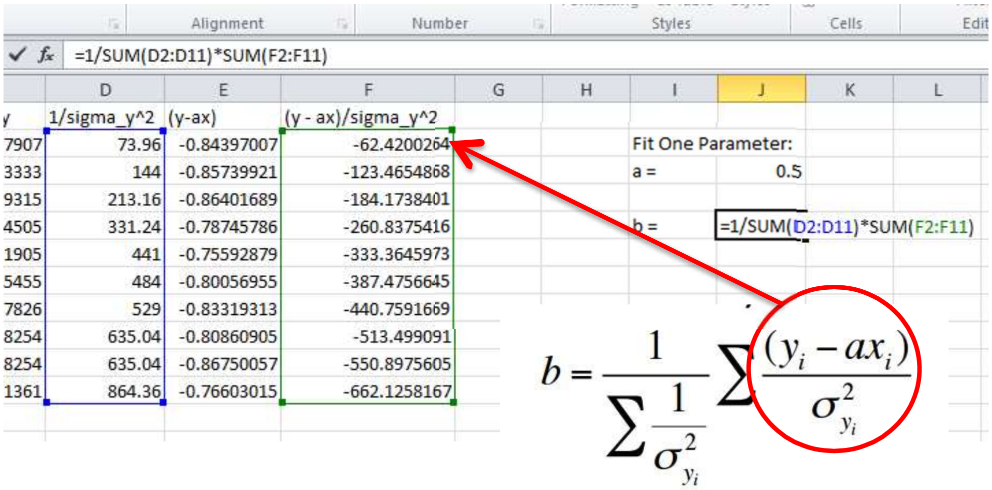


T.TEST | =$B2-$J$3*$A2

	A	B	C	D	E	F	G	H	I	J
1	x	y	sigma_y	1/sigma_y^2	(y_i - a x_i)					
2	0	-0.84397	0.11627907	73.96	=$B2-$J$3*$A2				Fit One Parameter:	
3	0.693147	-0.51083	0.08333333	144					a =	0.5
4	1.098612	-0.31471	0.06849315	213.16						
5	1.386294	-0.09431	0.05494505	331.24					b =	
6	1.609438	0.04879	0.04761905	441						
7	1.791759	0.09531	0.04545455	484						
8	1.94591	0.139762	0.04347826	529						
9	2.079442	0.231112	0.03968254	635.04						
10	2.197225	0.231112	0.03968254	635.04						
11	2.302585	0.385262	0.03401361	864.36						
12										
13										

Compute this in parts:

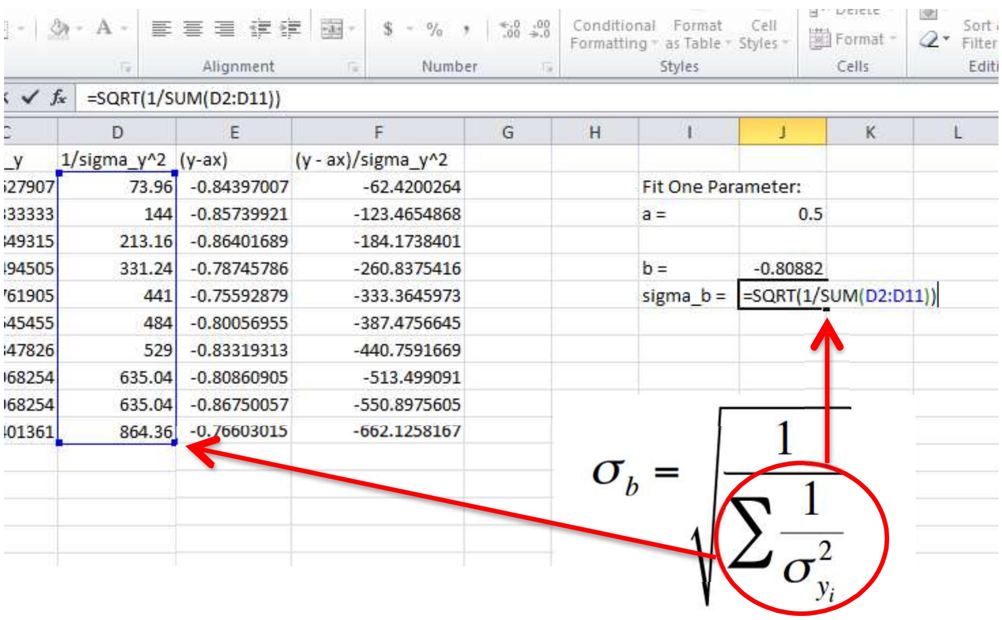
$$b = \frac{1}{\sum \frac{1}{\sigma_{y_i}^2}} \sum \frac{(y_i - ax_i)}{\sigma_{y_i}^2}$$



=1/SUM(D2:D11)*SUM(F2:F11)

	D	E	F	G	H	I	J	K	L
y	1/sigma_y^2	(y-ax)	(y - ax)/sigma_y^2						
7907	73.96	-0.84397007	-62.4200264			Fit One Parameter:			
3333	144	-0.85739921	-123.4654868			a =	0.5		
9315	213.16	-0.86401689	-184.1738401						
4505	331.24	-0.78745786	-260.8375416			b =	=1/SUM(D2:D11)*SUM(F2:F11)		
1905	441	-0.75592879	-333.3645973						
5455	484	-0.80056955	-387.4756645						
7826	529	-0.83319313	-440.7591669						
8254	635.04	-0.80860905	-513.499091						
8254	635.04	-0.86750057	-550.8975605						
1361	864.36	-0.76603015	-662.1258167						

$$b = \frac{1}{\sum \frac{1}{\sigma_{y_i}^2}} \sum \frac{(y_i - ax_i)}{\sigma_{y_i}^2}$$



=SQRT(1/SUM(D2:D11))

C	D	E	F	G	H	I	J
_y	1/sigma_y^2	(y-ax)	(y - ax)/sigma_y^2				
27907	73.96	-0.84397007	-62.4200264			Fit One Parameter:	
33333	144	-0.85739921	-123.4654868			a =	0.5
49315	213.16	-0.86401689	-184.1738401				
94505	331.24	-0.78745786	-260.8375416			b =	-0.80882
61905	441	-0.75592879	-333.3645973			sigma_b =	=SQRT(1/SUM(D2:D11))
45455	484	-0.80056955	-387.4756645				
47826	529	-0.83319313	-440.7591669				
68254	635.04	-0.80860905	-513.499091				
68254	635.04	-0.86750057	-550.8975605				
01361	864.36	-0.76603015	-662.1258167				

$$\sigma_b = \sqrt{\frac{1}{\sum \frac{1}{\sigma_{y_i}^2}}}$$

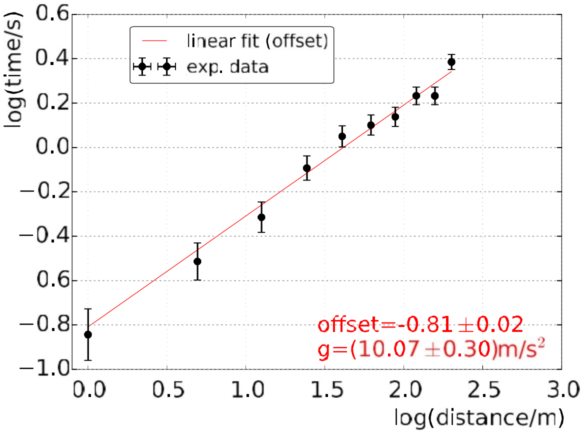
51

Q: Is this value consistent with 9.81m/s²

Calculate uncertainty on g ✐ easy! (error propagation)

$$\sigma_g = \left|\frac{\partial g}{\partial b}\right| \sigma_b$$

$$\sigma_g = 4\exp(-2b)\sigma_b$$



52