Introduction to Data Analysis

- Analysis of Experimental Errors
- How to Report and Use Experimental Errors
- Statistical Analysis of Data
  - Simple statistics of data
  - Plotting and displaying the data
- Summary

Errors and Uncertainties

experimental error " mistake
experimental error " blunder

experimental error = inevitable uncertainty of the measurement

The measured value alone **is not enough**. We also need the experimental error.

Testing the Theories and Models

Experimental data should be consistent with you theory (model) and inconsistent with alternative ones to prove them wrong.

Example: Bending of the light near the Sun.

1. simplest classical theory 0’’
2. careful classical analysis 0.9’’
3. Einstein’s general relativity 1.8’’

Solar eclipse needed to check: Dyson, Eddington, Davidson, year 1919, a=2’’,
95% confidence, 1.7’’ < α < 2.3’’.
Consistent with 1.8’’ and inconsistent with 0.9’’!!

Types of Experimental Errors

Random errors

- rotating table + stopwatch
- voltage measurements

i) revealed by repeating the measurements

ii) may be estimated statistically

Systematic errors

- measuring glass
- car performance
- calibration errors

in general are invisible

detected by comparison with results of alternative method or equipment

How to Report & Use the Errors

Experimental result:

measured value of $x = x_{best} \pm \delta x$

the best estimate for x

uncertainty, error, margin of error

$$x_{best} - \delta x < x < x_{best} + \delta x$$

Usually 95% confidence is suggested:

95% sure x inside the limits

5% chance x is outside the limits

Some Basic Rules

Experimental errors should be always rounded to one significant digit.

$$g = 9.82 \pm 0.02385 \qquad \text{wrong}$$

$$g = 9.82 \pm 0.02 \qquad \text{correct}$$

Thus error calculations become simple estimates.

Exception: if the leading significant digit of the error is 1, keep one digit more.

Everest is 8848 ± 1.5 m high

Some Basic Rules

The last significant figure in the answer should be the same order of magnitude as the uncertainty.

$$92.8 \pm 0.3$$

$$92 \pm 3$$

$$90 \pm 30$$

During the calculation retain one more digit than is finally justified to decrease the error.

Error in a Sum or a Difference

Two (or more) independently measured variables:

$x = x_{best} \pm dx$

$$\delta(x \pm y) = \delta x + \delta y$$

$y = y_{best} \pm dy$

$$\delta(x \pm y \pm z \pm \ldots.) = \delta x + \delta y + \delta z + \ldots.$$

If x, y, z,…. are large, but $(x \pm y \pm z \pm \ldots.)$ is small → trouble! Overall error is large.

Calculating Relative Errors

relative error (fractional uncertainty) $= \dfrac{\delta x}{x}$

$$x = 50 \pm 1 \text{ cm} \qquad \frac{\delta x}{x} = 0.02$$

$$x = 100000 \pm 1 \text{ cm} \qquad \frac{\delta x}{x} = 0.00001$$

Relative errors of product and ratio of two variables.

$$x = x_{best} \pm \delta x \qquad \delta x << x$$

$$y = y_{best} \pm \delta y \qquad \delta y << y$$

$$\frac{\delta(xy)}{xy} = ?$$

$$xy = (x_{best} \pm \delta x)(y_{best} \pm \delta y) \approx x_{best} y_{best} \pm (x_{best}\delta y + y_{best}\delta x)$$

$$\delta(xy) = x_{best}\delta y + y_{best}\delta x$$

$$\frac{\delta(xy)}{xy} = \frac{\delta x}{y_{best}} + \frac{\delta y}{x_{best}}$$

For product xy the relative error is the sum of relative errors of x and y.

Relative errors of product and ratio of two variables.

$$\frac{\delta\left(x/y\right)}{x/y} = ? \qquad x/y = \frac{x_{best} \pm \delta x}{y_{best} \pm \delta y} \approx \frac{x_{best}}{y_{best}} \frac{1 \pm \delta x / x_{best}}{1 \pm \delta y / y_{best}}$$

$$\max = \frac{x_{best}}{y_{best}} \frac{1 + \delta x / x_{best}}{1 - \delta y / y_{best}} \approx \frac{x_{best}}{y_{best}} \left(1 + \frac{\delta x}{x_{best}} + \frac{\delta y}{y_{best}} \right)$$

$$\min = \frac{x_{best}}{y_{best}} \frac{1 - \delta x / x_{best}}{1 + \delta y / y_{best}} \approx \frac{x_{best}}{y_{best}} \left(1 - \frac{\delta x}{x_{best}} - \frac{\delta y}{y_{best}} \right)$$

$$\frac{\delta(x/y)}{x/y} = \frac{\delta x}{y_{best}} + \frac{\delta y}{x_{best}}$$

2 Simple Rules.

When the measured quantities are added or subtracted the errors add.

When the measured quantities are multiplied or divided, the relative errors add.

Propagation of Errors

We have upper bounds on errors for sum/difference and product/quotient of 2 measurables. Can we do any better?

If errors are independent and random:
the errors are *added in quadrature*.

$$q = x + y$$

$$\delta q = \sqrt{(\delta x)^2 + (\delta y)^2} \le \delta x + \delta y$$

$$q = x + ... + z - (u + ... + w)$$

$$\delta q = \sqrt{(\delta x)^2 + ... + (\delta z)^2 + (\delta u)^2 + ... + (\delta w)^2} \le$$

$$\le \delta x + ... + \delta z + \delta u + ... + \delta w$$

Propagation of Errors

$$l_1 = 5.3 \pm 0.2cm$$

$$l_2 = 7.2 \pm 0.2cm$$

$$l = l_1 + l_2$$

$$\delta l = \sqrt{\left(\delta l_1\right)^2 + \left(\delta l_2\right)^2} = \sqrt{\left(0.2\right)^2 + \left(0.2\right)^2} \approx 3mm$$

$$= \delta l_1 + \delta l_2 = 2mm + 2mm = 4mm$$

For 2 measurables there is no great difference, but for n measurables the difference is $1/\sqrt{n}$.

Propagation of Errors

Relative error of product/quotient:

$$q = \frac{x \times \ldots \times z}{u \times \ldots \times w}$$

$$\frac{\delta q}{|q|} = \sqrt{\left(\frac{\delta x}{x}\right)^2 + \ldots + \left(\frac{\delta z}{z}\right)^2 + \left(\frac{\delta u}{u}\right)^2 + \ldots + \left(\frac{\delta w}{w}\right)^2} \leq$$

$$\leq \frac{\delta x}{x} + \ldots + \frac{\delta z}{z} + \frac{\delta u}{u} + \ldots + \frac{\delta w}{w}$$

If the relative errors for n measurables are the same, we gain $1/\sqrt{n}$ in relative error.

Propagation of Errors

D.C. electric motor, V - voltage, I - current.

$$efficiency,\quad e = \frac{work\,done\,by\,motor}{energy\,delivered\,to\,motor} = \frac{mgh}{VIt}$$

$$relative\,error\;\;for\quad m, h, V, I \;=1\%$$

$$relative\,error\;\;for\quad t \;=5\%$$

$$\frac{\delta q}{|q|} = \sqrt{\left(\frac{\delta m}{m}\right)^2 + \left(\frac{\delta h}{h}\right)^2 + \left(\frac{\delta V}{V}\right)^2 + \left(\frac{\delta I}{I}\right)^2 + \left(\frac{\delta t}{t}\right)^2} =$$

$$\sqrt{(1\%)^2 + (1\%)^2 + (1\%)^2 + (1\%)^2 + (5\%)^2} = \sqrt{29}\% \approx 5\%$$

$$\frac{\delta m}{m} + \frac{\delta h}{h} + \frac{\delta V}{V} + \frac{\delta I}{I} + \frac{\delta t}{t} = 1\% + 1\% + 1\% + 1\% + 5\% = 9\%$$

Propagation of Errors

If many measurables, the error is dominated by the one from the “worse” measurable:

$$\frac{\delta e}{e} \approx \frac{\delta t}{t}$$

What if measure x, but need an error for f(x)?

- Suggest that the error is small
- Do Taylor expansion of f about x_{best}

$$f(x) = f(x_{best}) \pm \left| \frac{df(x)}{dx} \right| \delta x$$

Propagation of Errors

Relative error of a power:

$$q = x^n$$

$$\frac{\delta q}{|q|} = |n| \frac{\delta x}{|x|}$$

General case - function of several variables:

$$q(x, \ldots, z)$$

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x}\delta x\right)^2 + \ldots + \left(\frac{\partial q}{\partial z}\delta z\right)^2}$$

$$\delta q \leq \left|\frac{\partial q}{\partial x}\right|\delta x + \ldots + \left|\frac{\partial q}{\partial z}\right|\delta z$$

Propagation of Errors

Example: measuring g with a simple pendulum, L-length, T-oscillation period

$T=2\pi (L/g)^{1/2}$ → $g=4\pi^2 L/T^2$

$$\frac{\delta(T^2)}{T^2}=2\frac{\delta T}{T} \qquad \frac{\delta g}{g}=\sqrt{\left(\frac{\delta L}{L}\right)^2+\left(2\frac{\delta T}{T}\right)^2}$$

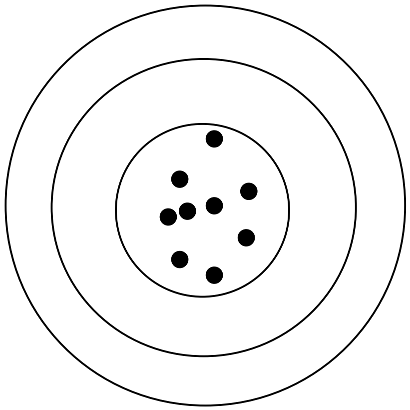
L = 92.95±0.1 cm, T = 1.936±0.004 sec.

$$g_{best}=\frac{4\pi^2\times(92.95\,cm)}{(1.936\sec)^2}=979\,cm/\sec^2$$

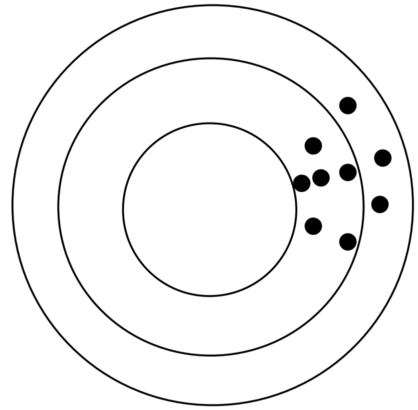
δL/L=0.1% δT/T=0.2% $\frac{\delta g}{g}=\sqrt{(0.1)^2+(2\times 0.2)^2}\,\%=0.4\%$

δg = 0.004 x 979 cm/sec^2 = 4 cm/sec2

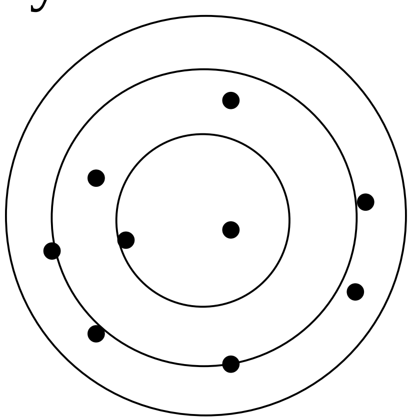
Statistical Analysis of Random Errors.



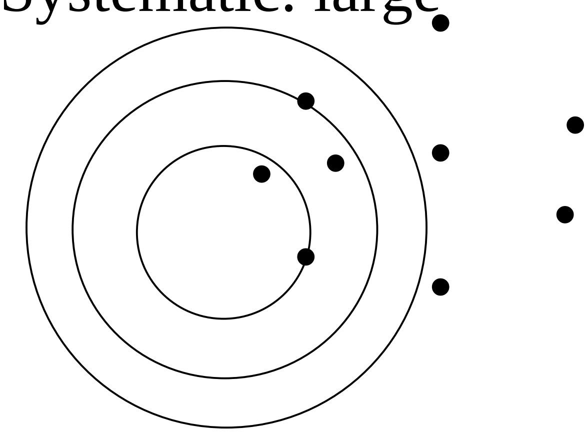
Random: small
Systematic: small



Random: small
Systematic: large

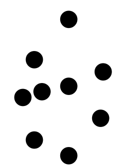


Random: large
Systematic: small

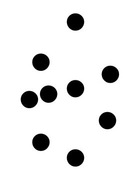


Random: large
Systematic: large

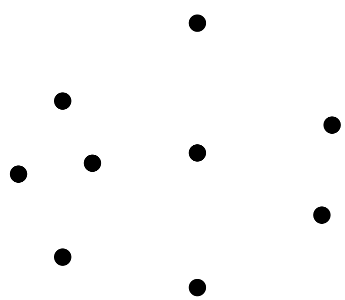
Statistical Analysis of Random Errors.



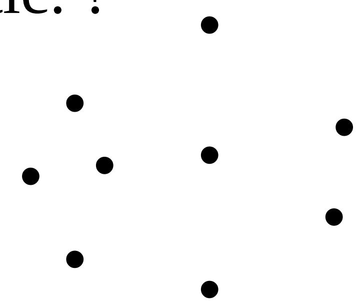
Random: small
Systematic: ?



Random: small
Systematic: ?



Random: large
Systematic: ?



Random: large
Systematic: ?

Statistical Analysis of Random Errors.

1st case: we knew the value of the measured quantity.
a) not realistic, b) not very interesting
2nd case - realistic, no chance of determining the systematic error.

Concentrate on random errors and repeat the measurements several times.

$x_1, x_2, x_3, x_4, \ldots\ldots$

Statistical Analysis of Random Errors

• measure x several times, get a set of data:
$x_1, x_2, \ldots, x_N$
• try to estimate the variability (dispersion) of the measurable and it's best value.

Example: manufacturer of metal parts
Complaints: non-uniformity of melting temperatures.

Analysis: Pick a representative batch of parts and measure melting temperatures.
Get a set of data: $t_1, t_2, \ldots, t_N$ for all N parts from the sample.

Statistical Analysis of Random Errors

Calculate the average (mean) t: $t_{best} = \bar{t} = \frac{1}{N}\sum_{i=1}^{N} t_i$
(our estimate for the true value of t)

Sort the data: $t_1 \le t_2 \le \ldots \le t_N$

t_j	310	311	312	313	314	315
n_j	1	3	8	6	2	0
n_j/N	1/20	3/20	8/20	6/20	2/20	0

n_j / N – probability of $t_j \le t \le t_{j+1}$

Statistical Analysis of Random Errors

$$\sum_{i=1}^{N} \frac{t_i}{N} = \sum_{j} \frac{t_j n_j}{N} = 1 - \text{normalized probability}$$

Plot a histogram.

Increase N → get a smoother histogram.
If the errors are random and small, we get a Gaussian bell-shaped curve at N → infinity.

But we can get more information on random errors before that...

Statistical Analysis of Random Errors

Example: 5 measurements of some value.

71, 72, 72, 73, 71

$$x_{best} = \bar{x} = \frac{71+72+72+73+71}{5} = 71.8 = \frac{\sum_{i=1}^{N} x_j}{N}$$

Trial number	Measured value	Deviation d_i	Deviation squared
1	71	-0.8	0.64
2	72	0.2	0.04
3	72	0.2	0.04
4	73	1.2	1.22
5	71	-0.8	0.64

$d_i = x_i - \bar{x}$ $\quad \sum d_i = 0$ $\quad \sum d_i^{\,2} = 2.8$

Statistical Analysis of Random Errors

Linear deviations are no good to characterize the statistics of random errors: have zero average,

Let's go for squares.

Definition: Standard deviation σ_x:

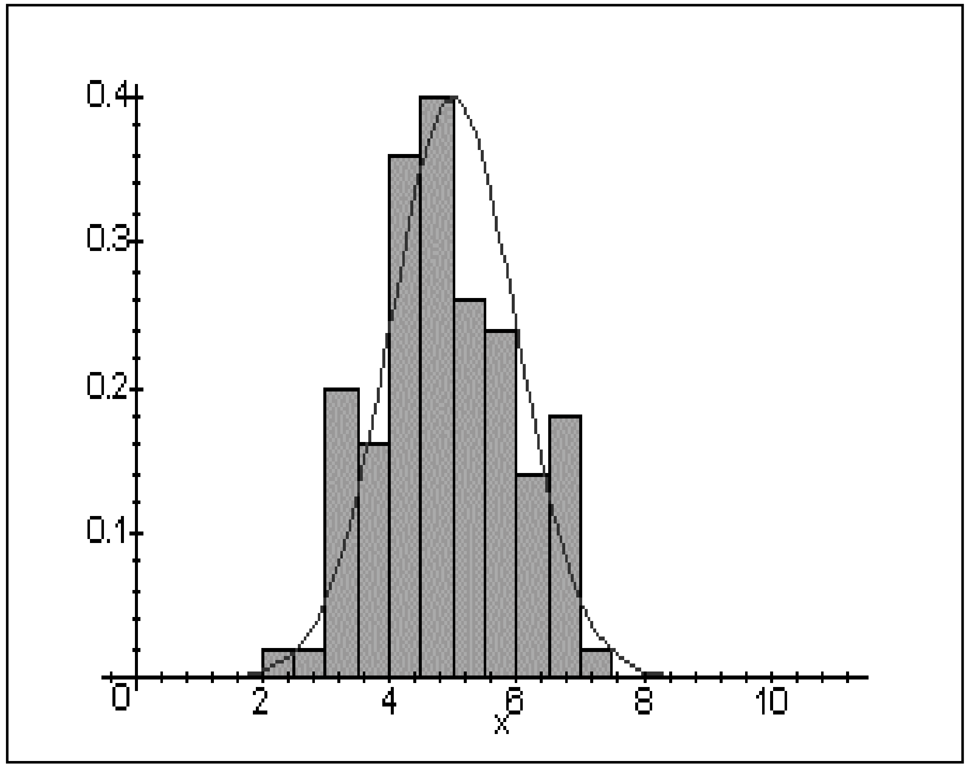
$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(d_i\right)^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \overline{x}\right)^2}$$

Root mean square (RMS) deviation of the measurements $x_1, x_2, \ldots, x_N$

Statistical Analysis of Random Errors

Standard deviation - measure of the accuracy of a single measurement or of the width of the data distribution (histogram).

At N→infinity histogram turns into a bell-shaped Gaussian curve.



Statistical Analysis of Random Errors

Standard deviation in a single measurement:
σ_x characterizes the average error of the measurements x_1, x_2, …, x_N. At large N the distribution approaches Gaussian and $P(x_{true} - \sigma_x < x < x_{true} + \sigma_x) = 68\%$.

If you make a single measurement knowing σ_x for the method used, it's 68% probable that your x is distance σ_x or less form x_{true}.

You may be 68% confident that the single measurement is within σ_x from the correct answer.

Statistical Analysis of Random Errors

Standard deviation of the mean:

$x_1, x_2, \ldots, x_N$ suggest $x_{best} = \Sigma x_i / N$.

How good (accurate) is our knowledge of x_{best}?

It may be proven that:

$$\sigma_{\bar{x}} = \sigma_x \frac{1}{\sqrt{N}}$$

Standard deviation of the mean

Or: Average of N measurements gives a $1/N^{1/2}$ smaller error, than a single measurement.

$$x = x_{best} \pm \frac{\sigma_x}{\sqrt{N}}$$

Statistical Analysis of Random Errors

Example: Measure elasticity constants of springs.
86, 85, 84, 89, 85, 89, 87, 85, 82, 85

$$\overline{k} = 85.7\,N/m,\ \sigma_k\, 2.16\,N/m \approx 2\,N/m$$

If we make 10 measurements and get $\overline{k} = 85.7\,N/m$

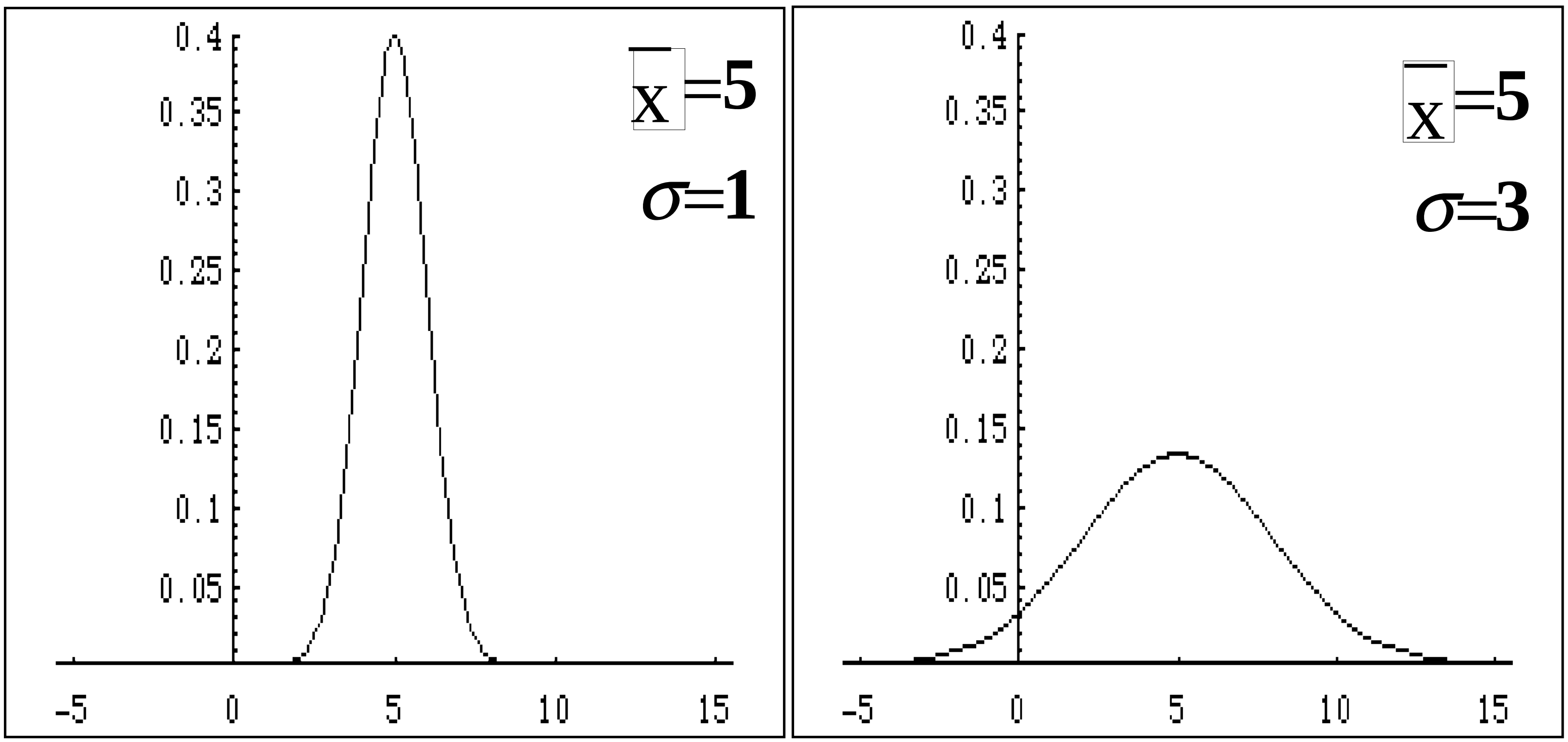
$$\delta\overline{k} \text{ will be } \frac{\sigma_k}{\sqrt{10}} \approx 0.7\,N/m$$

$$\overline{k} = 85.7 \pm 0.7\,N/m$$

The more measurements we make or the larger is the sample, the more accurate is the measurement!!

Gaussian Distribution

If the errors of the measurement are random and small, we approach the Gaussian distribution at large N.



Gaussian Distribution

$$G_{\bar{x},\sigma} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{\left(x-\bar{x}\right)^2}{2\sigma^2} \right\}$$

$$P(x, x+dx) = G(x)dx \qquad \text{(probability density)}$$

$$P(a \le x \le b) = \int_a^b G(x)dx$$

$$\int_{-\infty}^{\infty} G(x)dx = 1 \qquad \text{(normalized)}$$

$$G(\bar{x}+dx) = G(\bar{x}-dx) \qquad \text{(symmetric with respect to } \bar{x}\text{)}$$

Gaussian Distribution

$$\int_{-\infty}^{\infty} x\, G(x)dx = \bar{x} = x_{best} \qquad \text{(first moment)}$$

$$\int_{-\infty}^{\infty} \left(x - \bar{x}\right)^2 G(x)dx = \sigma^2 \qquad \text{(second moment)}$$

remember "experimental" $\bar{x}$ and σ?

At $N \to \infty$ $\sum$ is replaced by $\int$.

$$\int_{\bar{x}-\sigma}^{\bar{x}+\sigma} x\, G(x)dx \approx 0.68 \ \ (68\%)$$

$$\int_{\bar{x}-2\sigma}^{\bar{x}+2\sigma} x\, G(x)dx \approx 0.95 \ \ (95\%)$$

Least Squares Fitting

So far: Single measurable x, though multiple measurements to get the statistics. Analysis of random errors.

Now: The statistics is known. Multiple measurables y_1, y_2, …, y_N , at different values of x: x_1, x_2, …, x_N; $y_i = f(x_i)$. We are figuring out f(x).

Two possible approaches:

Either : 1. Measure y_i,
2. Plot the data.
3. Guess the form of y = f(x) & make the fit.

Or: 1. Have an idea of y = f(x).
2. Set the experiment.
3. Plot the data to check the initial guess (model).

Least Squares Fitting

The lore:

1. Get many points.
2. Plot the data.
3. Play with the coordinates (lin-lin, log-lin, log-log).
4. Look for the simplest possible fits.

Least Squares Fitting

The simplest case - linear dependence:

$$y = \mathbf{a}\,x + \mathbf{b}.$$

$x_1, x_2, \ldots, x_N$ & $y_1, y_2, \ldots, y_N$

What are the best **a** and **b** to fit the data?

- Suggest Gaussian distribution for each y_i, know σ_y.
- Assume the fit curve is the "true" f(x), ax_i+b is the "true" value of y_i.
- Construct $P(y_1, y_2, \ldots, y_N; a,b)$ – probability of having the set of data $y_1, y_2, \ldots, y_N$.
- Maximize $P(y_1, y_2, \ldots, y_N; a,b)$ with respect to **a** & **b** and find the corresponding **a** and **b**.
- Compare **a** and **b** with theory (if we have any).

Least Squares Fitting

For a single measurement:

$$P(y_i) \sim \frac{1}{\sigma_y} \exp\left\{ -\frac{\left(y_i - (a\,x_i + b)\right)^2}{2\,\sigma_y{}^2} \right\}$$

measured value: y_i; "true" value: $a\,x_i + b$

For a set of measurements:

$$P(y_1, y_2, ..., y_N) \sim P(y_1)P(y_2)...P(y_N) \sim \frac{1}{\sigma_y{}^N} \exp\left\{ -\frac{\chi^2}{2} \right\}$$

$$\chi^2 = \sum_{i=1}^{N} \frac{\left(y_i - (a\,x_i + b)\right)^2}{\sigma_y{}^2}$$

$$\text{We look for: } \frac{\partial \chi^2}{\partial a} = 0 \;\&\; \frac{\partial \chi^2}{\partial b} = 0$$

Least Squares Fitting

We have a system of linear equations for a and b:

$$\begin{cases} \dfrac{\partial \chi^2}{\partial a} = -\dfrac{2}{\sigma_y{}^2} \displaystyle\sum_{i=1}^{N} x_i \left(y_i - (ax_i + b) \right) = 0 \\ \dfrac{\partial \chi^2}{\partial b} = -\dfrac{2}{\sigma_y{}^2} \displaystyle\sum_{i=1}^{N} \left(y_i - (ax_i + b) \right) = 0 \end{cases}$$

$$\begin{cases} a \sum x_i{}^2 + b \sum x_i = \sum x_i y_i \\ a \sum x_i + Nb = \sum y_i \end{cases}$$

Least Squares Fitting

The solutions are:

$$a = \frac{\sum x_i y_i - \sum x_i \sum y_i}{\Delta}$$

$$b = \frac{\sum x_i^{\,2} y_i - \sum x_i \sum x_i y_i}{\Delta}, \text{ where}$$

$$\Delta = \sum x_i^{\,2} - \left(\sum x_i\right)^2$$

A x + b - least squares fit to the data, or line of regression of y on x.

Linear Least Squares, General case

Our fitting function in general case is:

$$F(x) = a_1 f_1(x) + a_2 f_2(x) + \ldots + a_n f_n(x)$$

Note that the function itself does not have to be linear for the problem to be linear in the fitting parameters.

Let us find a compact way to formulate the least squares problem.

Linear Least Squares, General case

Thus we have: vectors x, y and a:

$$x = \begin{pmatrix} x_1 \\ \ldots \\ x_N \end{pmatrix} \begin{matrix} \text{points} \\ \text{where the data} \\ \text{was taken} \end{matrix}, \quad y = \begin{pmatrix} y_1 \\ \ldots \\ y_N \end{pmatrix} \text{the data},$$

$$a = \begin{pmatrix} a_1 \\ \ldots \\ a_n \end{pmatrix} \begin{matrix} \text{fitting} \\ \text{parameters} \end{matrix}$$

and functions $f_1(x), f_2(x), \ldots, f_n(x)$.

Linear Least Squares, General case

The problem now looks like:

$y_i = F(x_i) + e_i,$ where e_i is a residual: mismatch between the measured value and the one predicted by the fit.

Let's intorduce vector e:

$$e = \begin{pmatrix} e_1 \\ \dots \\ e_N \end{pmatrix}$$

Linear Least Squares, General case

Let us express the problem in matrix notation:

$$Z=\begin{bmatrix} f_1(x_1) & f_2(x_1) & \dots & f_n(x_1) \\ f_1(x_2) & f_2(x_2) & \dots & f_n(x_2) \\ \dots & \dots & \dots & \dots \\ f_1(x_N) & f_2(x_N) & \dots & f_n(x_N) \end{bmatrix}$$

Overall we have now:

$$y = Z \cdot a + e$$

Fitting problem in matrix notation.

Look for $\min\left(\sum_{i=1}^{N} e_i^{\,2}\right) = \min(e^T e)$

Linear Least Squares, General case

$$\text{Look for } \min\left(\sum_{i=1}^{N} e_i^{\,2}\right) = \min\left(\sum_{i=1}^{N}\left(y_i - \sum_{j=1}^{n} z_{ij} a_j\right)^2\right) =$$

$$\min\left(\left(y - z\cdot a\right)^T \cdot \left(y - z\cdot a\right)\right)$$

$$\frac{\partial\left(e^T e\right)}{\partial a_k} = 0 \quad for\ 1 \le k \le n$$

$$\left(\frac{\partial\left(y - z\cdot a\right)}{\partial a_k}\right)^T \cdot \left(y - z\cdot a\right) = 0$$

Linear Least Squares, General case

$$\left(-z \cdot \frac{\partial(a)}{\partial a_k}\right)^T \cdot (y - z \cdot a) = 0$$

$$\left(z \cdot (0\,0 \ldots 1 \ldots 0)\right)^T \cdot (y - z \cdot a) = 0$$

$$\left(z_{1k}\ z_{2k} \ldots z_{Nk}\right)^T \cdot (y - z \cdot a) = 0 \quad for\ 1 \le k \le n$$

Using Matlab colon notation:

$$\left(z_{:,k}\right)^T \cdot (z \cdot a) = \left(z_{:,k}\right)^T \cdot y$$

Or after putting all n equations together:

$$\boxed{z^T \cdot z \cdot a = z^T \cdot y}$$

Linear Least Squares, General case

In general case linear lest squares problem can be formulated as a set of linear equations.

Ways to solve:

1. Gaussian elimination.
2. To calculate the matrix inverse:

$$a = \left(z^T \cdot z\right)^{-1} \cdot z^T \cdot y$$

Suitable for Matlab, see homework 9.

Nonlinear Regression (Least Squares)

What if the fitting function is not linear in fitting parameters?

We get a nonlinear equation (system of equations).

Example:

$$f(x) = a_1(1 - e^{-a_2 x}) + e$$

$$y_i = f(x_i; a_1, a_1, ..., a_m) + e_i \text{ or just } y_i = f(x_i) + e_i$$

Again look for the minimum of $\sum_{i=1}^{N} e_i^2$ with respect to the fitting parameters.

Matlab Function FMINSEARCH.

Accepts as input parameters:

1. Name of the function (FUN) to be minimized
2. Vector with initial guess X0 for the fitting parameters

Returns: Vector X of fitting parameters providing the local minimum of FUN.

Function FUN accepts vector X and returns the scalar value dependent on X.

In our case (hw10) FUN should calculate $\sum_{i=1}^{N} e_i^2$ dependent on the fitting parameters b, m, A_1, A_2, …

Matlab Function FMINSEARCH.

Syntax: x = FMINSEARCH(FUN,X0) or
x = FMINSEARCH(FUN,X0,OPTIONS)
See OPTIMSET for the detail on OPTIONS.

x = FMINSEARCH(FUN,X0,OPTIONS,P1,P2,..)
in case you want to pass extra parameters to FMINSEARCH

If no options are set use OPTIONS = [] as a place holder.

Use "@" to specify the FUN:

```
x = fminsearch(@myfun,X0)
```

Gauss-Newton method for nonlinear regression

$y_i = f(x_i; a_1, a_1, ..., a_m) + e_i$ or just $y_i = f(x_i) + e_i$

Look for the minimum of $\sum_{i=1}^{N} e_i^2$ with respect to a_i.

1. Make an initial guess for a: a0.

2. Linearize the equations (use Taylor expansion about a0).

3. Solve for Δa - correction to a0 $\rightarrow$ a1=a0+Δa - improved a-s and our new initial guess.

4. Back to (1).

5. Repeat until $\left| a_{k,j+1} - a_{k,j} \right| < \varepsilon$ for any k.

Gauss-Newton method for nonlinear regression

Linearization by Taylor expansion:

$$y_i = f(x_i) + e_i \approx f(x_i, a0) + \sum_{j=1}^{n} \frac{\partial f(x_i, a0)}{\partial a_n} + e_i$$

$$y_i - f(x_i, a0) = \sum_{j=1}^{n} \frac{\partial f(x_i, a0)}{\partial a_n} + e_i \quad \text{for i = 1,2,...,N}$$

or in matrix form:

$$D = Z \cdot \Delta a + e, \quad \text{where}$$

$$D = \begin{pmatrix} y_1 - f(x_1, a0) \\ .... \\ y_N - f(x_N, a0) \end{pmatrix}, \quad \text{and} \quad Z = \begin{bmatrix} \frac{\partial f(x_1, a0)}{\partial a_1} & ... & \frac{\partial f(x_1, a0)}{\partial a_n} \\ ... & ... & ... \\ \frac{\partial f(x_N, a0)}{\partial a_1} & ... & \frac{\partial f(x_N, a0)}{\partial a_n} \end{bmatrix}$$

Gauss-Newton method for nonlinear regression

Linear regression: $y = Z \cdot a + e$

Now, nonlinear regression : $D = Z \cdot \Delta a + e.$

Old good linear equations with Δa in plce of a, D in place of y and Z with partial derivatives in place of Z with values of functions.

Solve it for Δa, use a1=a0+Δa as the new initial guess and repeat the procedure untill the convergence criteria are met.....