

The role of biostatistics and informatics in the every day medical practice

The **purpose** of medical science:

- Prevention of diseases,
- Healing of the sick

Diagnostics: **scientific** methodology of recognition of diseases.

Therapy

Auxiliary sciences: e.g. anatomy, physiology, physics, chemistry, biology; *and*

Biostatistics and informatics

How can we approach this rather abstract subject?

As an example, let us take one by one the course of forming diagnosis:

Examination: first impressions, collection of **data**, anamnesis, physical examinations, additional examinations.

Thinking: evaluation of **data**, grading the symptoms, relationship among the symptoms, recognition of syndromes.

Answers for further questions (one or more diseases)

Thus **data** must be **collected, processed, conclusions have to be drawn** and sometimes **decisions** should be made.

General experience:

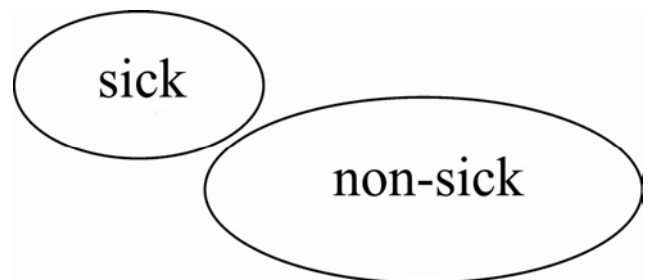
many of us make inappropriate conclusions easily, and make decisions based on them. (whisky)

Statistics is a field of science that is able to do all these.

It helps to arrange one's thoughts critically and keeps alive one's skepticism, the base of every intellectual activity.

The most important fundamental concepts and the connected problems

Who is sick and who is healthy?



1. Set: collection of distinct objects, considered as an object in its own right. They are characterized uniquely. Things belonging to the set are the **elements** of the set.

In general: **variable**

In which set the given element can be found?

Systematization, classification, separation (Greek)



What is the similarity between this toy and the diagnosis?

3 very different **bodies**

Characteristics:

„case/variable”	shape	color	size
1	sphere	yellow	4,3 cm
2	tetrahedron	blue	4,5 cm
3	cube	red	3,8 cm

3 very different **holes**

Characteristics:

„case/variable”	shape	color	size
1	circle	yellow	4,3 cm
2	triangle	blue	4,5 cm
3	square	red	3,8 cm

„Can not” miss it.

There is **one-to-one** correspondence.

Same (exists only theoretically)

("We both step and do not step in the same rivers." Heraclitus)

Instead, more or less **similar**

The **absence of uniqueness** can cause the problems.

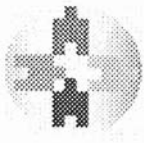
We are not able to take into account all of the circumstances.

Measure of similarity: confidence

The **essence** of the problem:

There are no two **identically** sick people.

There are diseases which can produce **very similar** symptoms.



Honvédelmi Minisztérium Állami Egészségügyi Központ
1134 Budapest, Róbert Károly krt. 44. Tel.: 06-1-465-1800 Fax: 06-1-340-3129

Főigazgató:

Működési engedély száma:

Központi Laboratóriumi Diagnosztikai Osztály

Osztályvezető főorvos:

Laboratóriumi eredmények

Megnevezés	Érték	M.e.	Megjegyzés	Eltérés	Referencia értékek	
Klinikai kémia						
Glukóz	3,5	mmol/l			3,1	5,6
Karbamid	6,1	mmol/l			1,7	8,3
Kreatinin meghat.	75	μmol/l			44	80

Zuglói Egészségügyi Szolgálat
1148 Budapest, Őrs Vezér tér 23.
Telefon: 469-4600

LABORATÓRIUMI LELET

Szakorvosi Rendelőintézet
Laboratórium

Labor vezető:

Páciens neve:

TAJ szám:

Született:

Anyja neve:

Nem:

Napi sorszám: **749**

Beut. egység: **340092019** Azon.: 012101003

Lelet kelte:

Kért vizsgálatok:	Eredmény: mértékegység	Referencia érték:
VÉRKÉP XT WBC	10,71 10 ³ /u	4,0 - 13,0
RBC vvt szám	4,22 10 ⁶ /u	3,9 - 5,6
KARBAMID	+ 9,6 mmol/l	1,7 - 8,3
KREATININ	+ 113,0 μmol/l	50,0 - 110,0

Semmelweis Egyetem ÁOK Központi Laboratórium
1083 Budapest, Korányi Sándor u. 2/a.
Intézetvezető:
Tel: 06 1 2100 278/1522,1457

LABORATÓRIUMI EREDMÉNYKÖZLŐ LAP

Név :

Születési idő :

TAJ/azonosító :

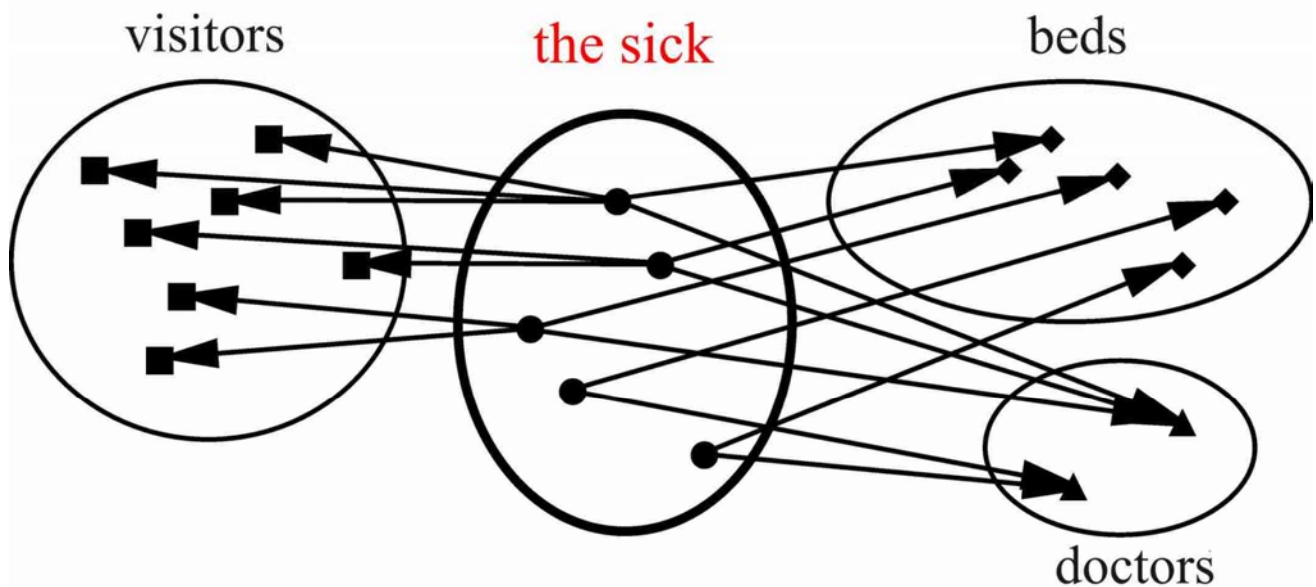
Nem:

Rendelés sorszáma: 6037990

Vizsgálat	Eredmény	M.Egység	Ref.tart
VVt süllyedés	2	mm/h	1-20
Karbamid	6,7	mmol/l	2,5-8
Kreatinin	108	μmol/l	62-106

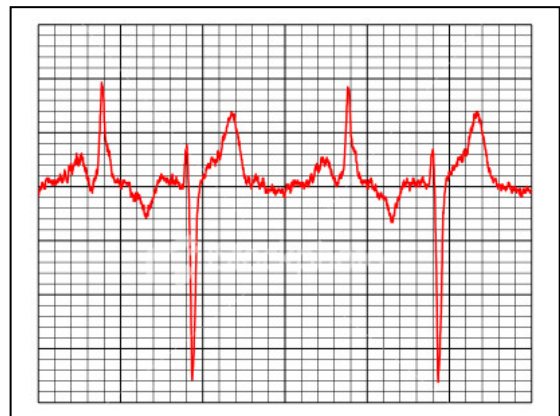
Here the set is a range of values. The question is whether the measured value is inside or outside the range.

2. Function (mapping), but there are exceptions



3. Data: facts for the cognition, characterization of somebody or something;
qualitative and quantitative characteristics of the surrounding world.

4. Signals: transmitter units of **data** (suitable for description of **data**)
changes in space or time, (or in both) e.g. change of light, sound, any sensation or a measurable quantity



Types of data and their characteristics

Categorical (qualitative)

Nominal e.g. blood group: A, B, AB, 0;

gender: woman, man

Relations: $A = A$ and $A \neq B$

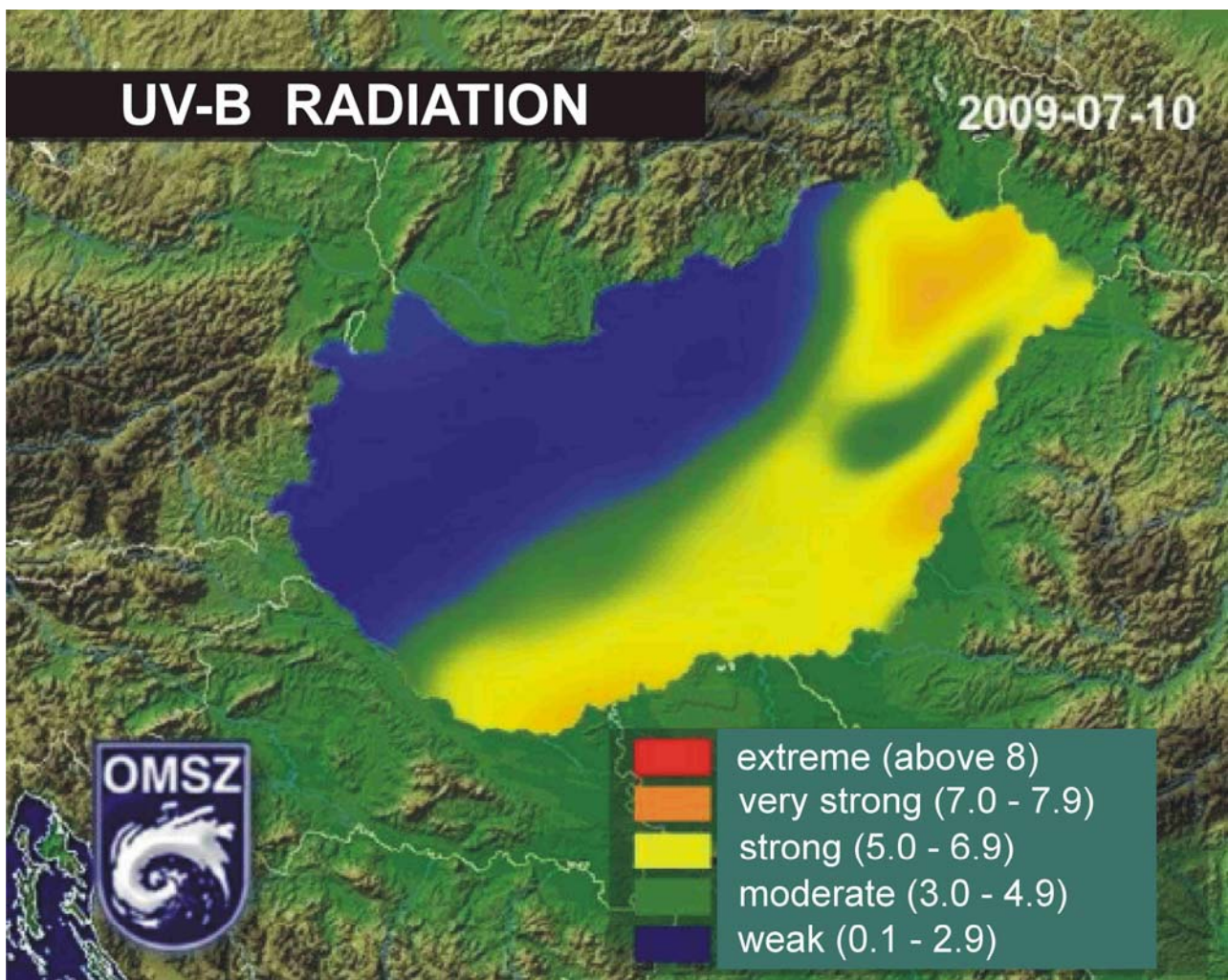
Ordinal e.g. grades: fail („1”), pass („2”), fair („3”), good („4”), excellent („5”);

Grading of a disease:

weak, moderate, serious

Relations:

fail < pass < fair < good < excellent



Numerical (quantitative)

Discrete e.g. number of teeth

Continuous e.g. blood pressure (p), body temperature (T)

Possible operations: *difference*
ratio

Scales and transformations (examples)

$$x(\text{Celsius}) = (x + 273,16)^\circ (\text{Kelvin})$$

$$x(\text{Kelvin}) = (x - 273,16)^\circ (\text{Celsius})$$

$$x(\text{inch}) = 2,54x(\text{cm})$$

$$x(\text{cm}) = x/2,54(\text{inch})$$

$$x^\circ (\text{Celsius}) = (1,8x + 32)^\circ (\text{Fahrenheit})$$

$$x^\circ (\text{Fahrenheit}) = \frac{(x - 32)^\circ}{1,8} (\text{Celsius})$$

Linear transformations

Variable can be

multiplied (or divided) **by a number**, and
we may **add** (or subtract) **a number** to (from) it

Richter scale

logarithmic ($\approx 32\times$)

The role of „change” in theory and practice

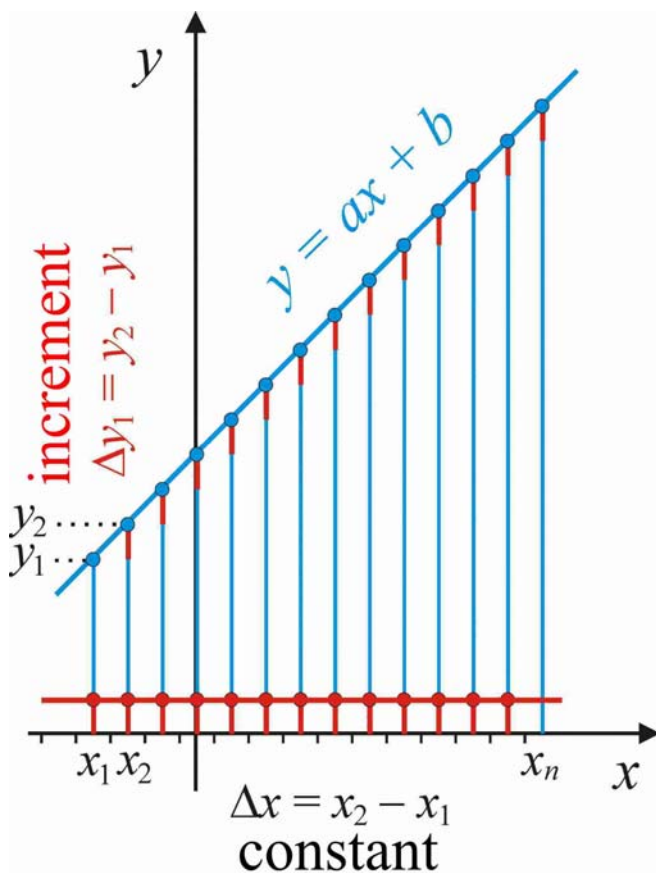
The “most important” feature of a function is the **change**.

How does it change?

Increases or decrease; quickly or slowly

How large is the **increment**?

The simplest function is the linear one: $y = ax + b$
(in most cases we prefer it)

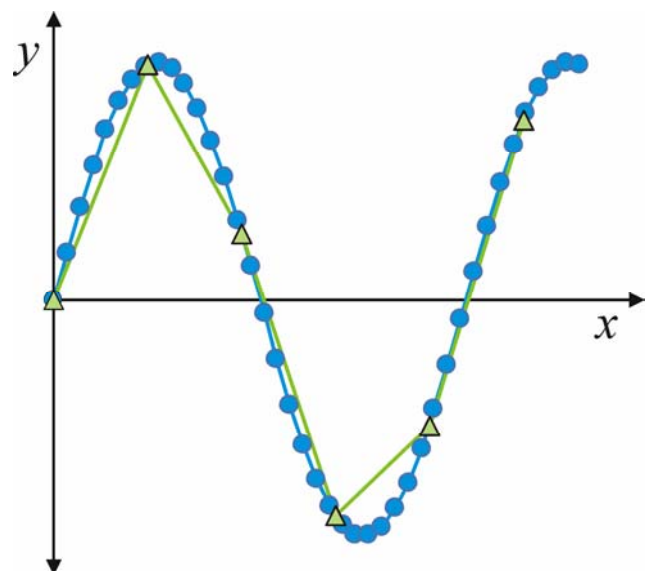
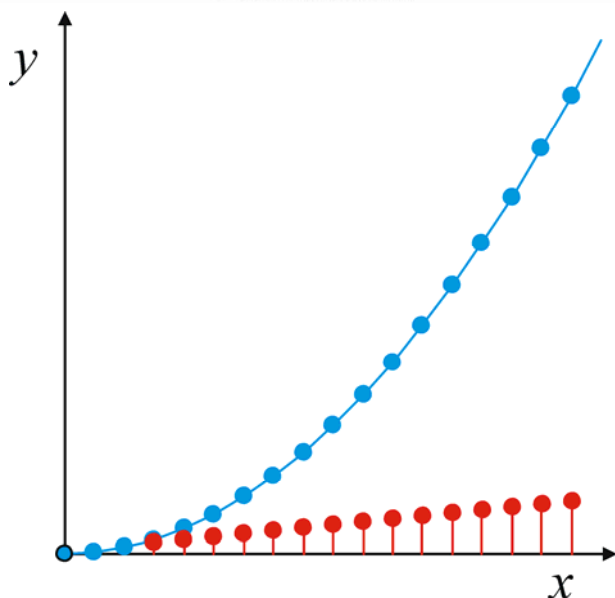


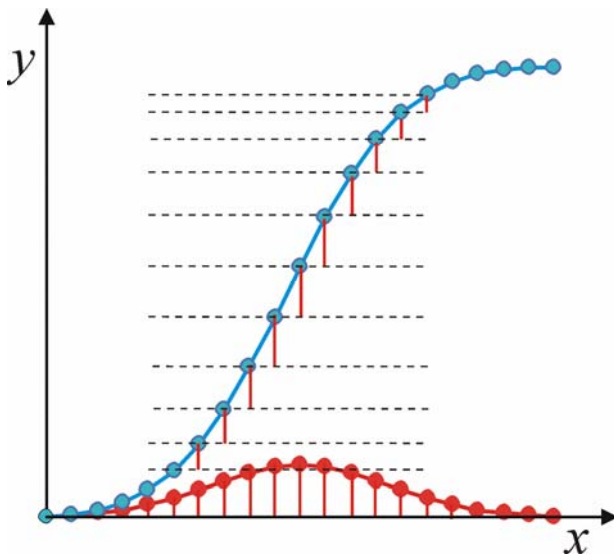
Differentiation:

make a difference
subtract

Difference quotient:
increment divided by the
step size $\frac{\Delta y}{\Delta x}$

Δx should be small





The **inverse operation** of the differentiation is also important.

Integration:
unite
sum

(calculation of area)

Origin of some important functions

1. Linear function

If Δy and Δx proportional, means $\Delta y = a\Delta x$, where a constant.
than $y = ax + y_0$

2. Exponential function

If $\Delta y = ay\Delta x$, where a constant, means
than $y = y_0 e^{ax}$

$$\frac{\Delta y}{y} = a\Delta x$$

3. Logarithmic function

$$y = a \ln x + k$$

$$\Delta y = a \frac{\Delta x}{x}$$

4. Powerfunction

$$y = kx^a$$

$$\frac{\Delta y}{y} = a \frac{\Delta x}{x}$$

Remarks: $e^{\ln 2} = 2$

$$e = 2^{\frac{1}{\ln 2}} = 2^{\frac{1}{0,693}} = 2,7182\dots$$

thus

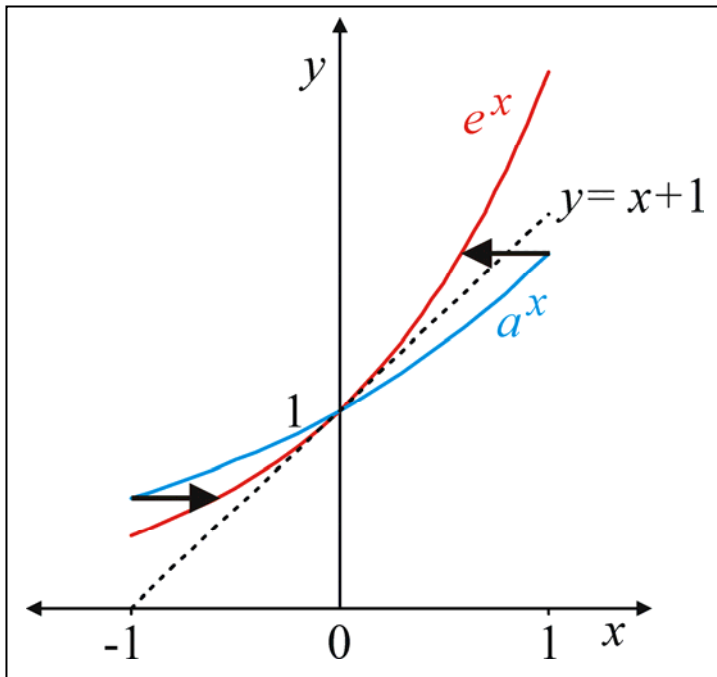
$$y = y_0 e^{ax} = y_0 2^{\frac{a}{\ln 2}x}$$

$$2. \ln y = \ln y_0 + a(x)$$

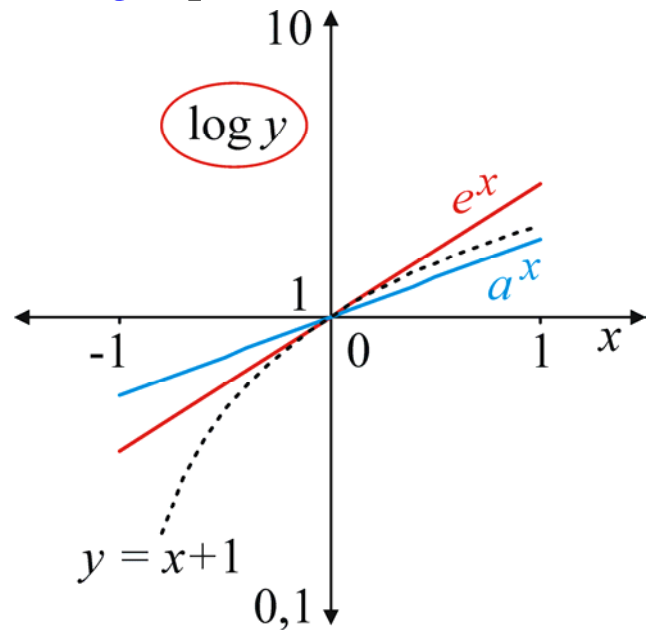
$$4. \ln y = \ln k + a(\ln x)$$

After this transformation we get a linear function in both cases.

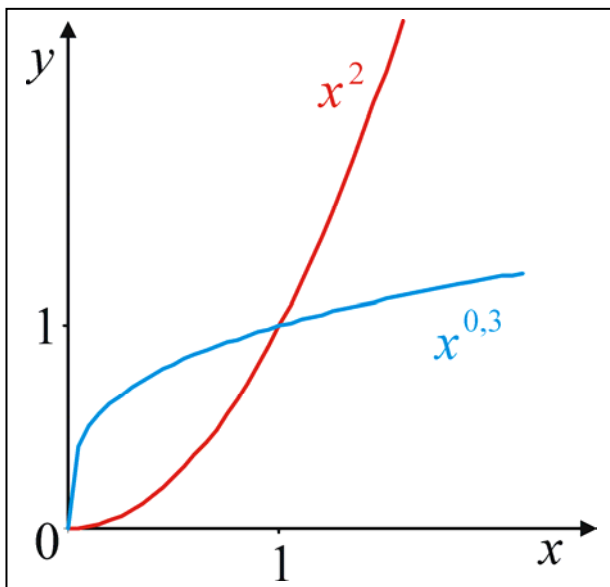
2.



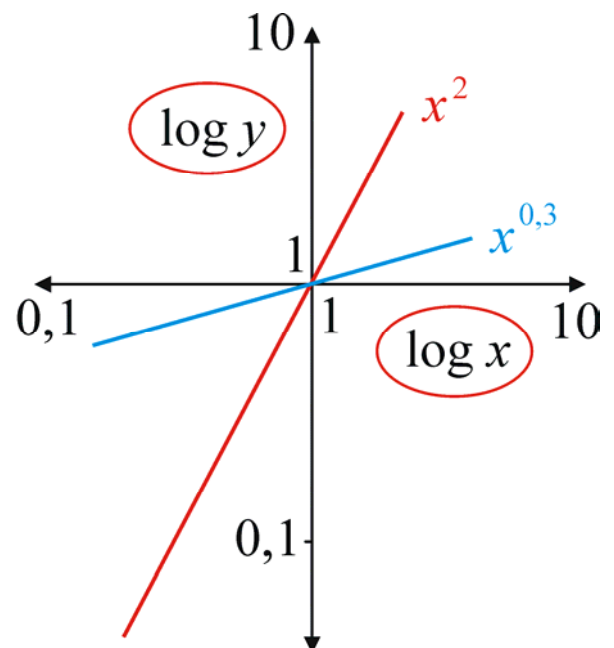
lin-log representation



4.



log-log representation



Deterministic part (determined by circumstances, which **could be taken into account**) and **stochastic** part (determined by circumstances, which **could not be taken into account**) of changes appear simultaneously.

E.g. reproduction of a bacteria population

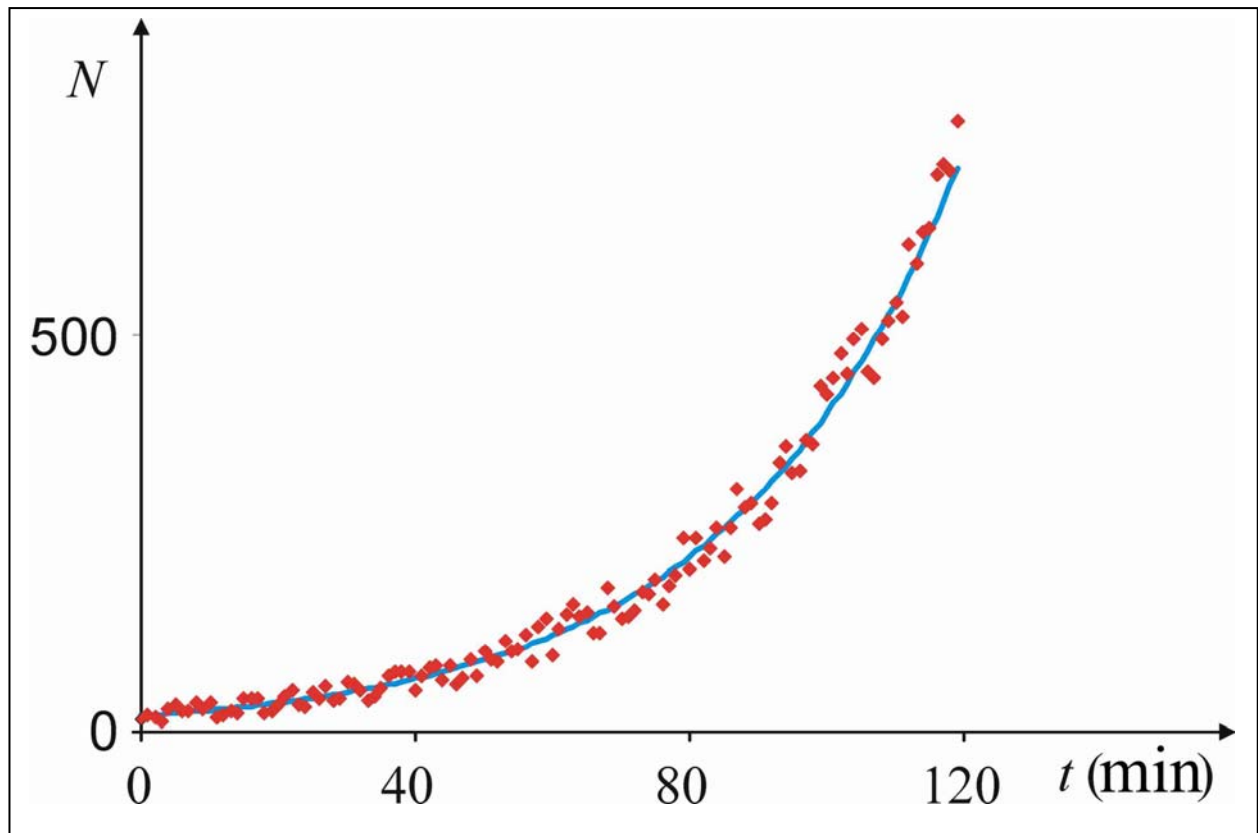
Theory

$\Delta N = aN\Delta t$ and, at time $t = t_0$ $N = N_0$
than

$$N(t) = N_0 e^{at} \quad \text{or} \quad N(t) = N_0 2^{\frac{t}{T}}$$

Practice (we should measure)

$$N_0 = 20$$



Doubling the number (N) 20, 40, 80, 160, 320, 640

Doubling time (T)
 ≈ 23 minute

There are **uncertainties** which come from the **measurement**, but they can also be caused by the properties of the **measured quantity**.