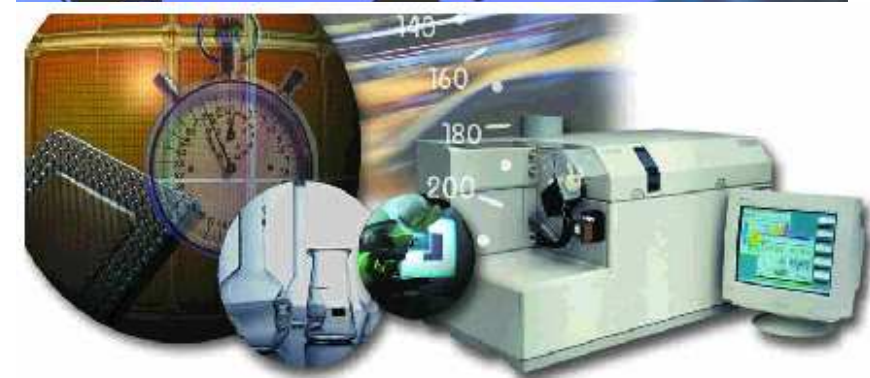


# СТАТИСТИКА И МЕТРОЛОГИЯ В ХИМИЯТА

Компютърна  
ХИМИЯ II курс

редовно обучение

**2008/09**



# ИМАТЕ ПРАВО!

- Да подлагате на съмнение всичко което ви кажа!
- Да използвате моите думи срещу мен!
- Да не приемате моите философски разсъждения, на които често ще бъдете подлагани.
- Да правите предложения за подобряване ефективността на образователния процес, като накрая на курса ще можете да оцените слабостите му в нарочна анонимна анкета


# НЯМАТЕ ПРАВО!

- Да сваляте отговорността за резултатите от себе си.
- Да прилагате получените знания със зла умисъл и във вреда на другите!



**Веселин Кметов** доц. д-р

E-mail [kmetov@uni-plovdiv.bg](mailto:kmetov@uni-plovdiv.bg)

 (032) 261 337

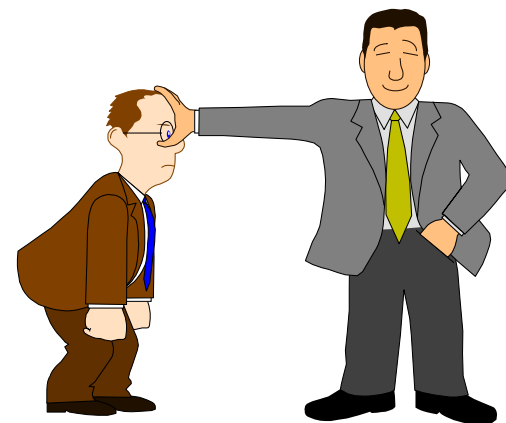
 <http://web.uni-plovdiv.bg/kmetov>



## ПРИНЦИПИ

- Придобиването на знания и компетентности е нашата цел. Преподавателят е тук за да помага на студентите в усилията им за изпълнението на тази ЦЕЛ.  
*“Учи мама, за да не работиш”*
- Студентите имат не по-нисък потенциал от мен. Превъзхождам ги само по възраст, житейски опит и времето отделено за усвояване на материала.
- Достойнството и добродетелността са над гениалността и целесъобразността

**Въпроси - по всяко време!**



## За курса

- Лекциите са доброволни (ЗАЩО и КАК ГО ПРАВИМ ),  
упражненията са задължителни (МОЖЕМ да ГО ПРАВИМ)  
**носете си електронни калкулатори, полезни са флаш памети**
- Самостоятелната подготовка е основния ключ към успеха !!! 2/0/2 60 + 25
- **Оценка = 0.3\*колоквиум + 0.7\* семестриален изпит + бонус**  
**(колоквиум – 2 АПРИЛ)**
- Студенти предоставили разработки (алгоритми, есета, методични материали)  
подпомагащи курса, получават бонус до 1 единица към крайната оценка.
- Колоквиума и изпита включват решаване на задачи с право на самостоятелно ползване  
на записки, пищови, помощна литература и изчислителна техника
- Критериите за оценяване са общи за всички и не подлежат на договаряне!

Prof. Antonio Canals – ERASMUS lecturer

“Univariate Methodological Calibration on Instrumental Analysis” – 9-10 Април

<http://argon.uni-plovdiv.bg/>

<http://forum.argon.acad.bg/>

<http://www.kosnos.com/>



ТУК МОЖЕТЕ ДА СВАЛИТЕ УЧЕБНИ ПРОГРАМИ И ЛЕКЦИОННИ МАТЕРИАЛИ

**ВНИМАНИЕ** - помагалата са НЕДОСТАТЪЧНИ за самоподготовка  
ПОСЕЩАВАЙТЕ ЛЕКЦИИТЕ И УПРАЖНЕНИЯТА!!!

Вашите критични забележки и коментари  
ще бъдат посрещнати с признателност от лектора!

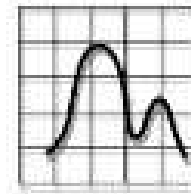
E-mail: [kmetov@uni-plovdiv.bg](mailto:kmetov@uni-plovdiv.bg)

## ОКС - БАКАЛАВЪР

<b>ИНСТРУМЕНТАЛНИ МЕТОДИ ЗА АНАЛИЗ I част</b>	<u>Химия</u> II курс редовно	<a href="#">програма</a> <a href="#">план упражнения</a> <b>ЛЕКЦИИ СВИТЪК ААС</b> <a href="#">задачи ИМА</a> <a href="#">оценки изпит</a>
	<u>Химия</u> III курс задочно	<a href="#">програма - лекции&amp;упраж.</a> <b>ЛЕКЦИИ СВИТЪК ААСz</b> <a href="#">оценки изпит</a>
<b>МЕТРОЛОГИЯ И СТАТИСТИКА В ХИМИЯТА</b>	<u>Химия</u> II курс редовно	<a href="#">програма</a> <a href="#">план упражнения</a> <b>ЛЕКЦИИ СВИТЪК МСХ</b> <a href="#">задачи за упражнения</a> <a href="#">неопределеност</a> <a href="#">изчисления с EXCEL</a> <a href="#">оценки изпит</a>
	<u>Химия</u> III курс задочно	<a href="#">програма</a> <a href="#">план упражнения</a> <b>ЛЕКЦИИ СВИТЪК МСХz</b> <a href="#">задачи за упражнения</a> <a href="#">неопределеност</a> <a href="#">изчисления с EXCEL</a> <a href="#">оценки изпит</a>
<b>СТАТИСТИКА И МЕТРОЛОГИЯ В ХИМИЯТА</b>	<u>Компютърна Химия</u> II курс редовно	<a href="#">програма</a> <a href="#">план упражнения</a> <b>ЛЕКЦИИ СВИТЪК СМХ</b> <a href="#">задачи за упражнения</a> <a href="#">неопределеност</a> <a href="#">изчисления с EXCEL</a> <a href="#">оценки изпит</a>
<b>КОМПЮТЪРНА КВАПИМЕТРИЯ</b> СВОБОДНО ИЗБИРАЕМ	<u>Компютърна Химия</u> IV курс редовно	<a href="#">програма</a> <a href="#">план упражнения</a> <a href="#">изчисления с EXCEL</a> <a href="#">оценки изпит</a>
<b>ИНСТРУМЕНТАЛЕН АНАЛИЗ</b>	<u>Компютърна Химия</u> III курс редовно	<a href="#">програма</a> <a href="#">план упражнения</a> <a href="#">ЗАДАЧИ за упражнения</a> <a href="#">оценки колоквиум</a>



# WHERE TO START?



ELECTRONIC TEXTBOOK  
**StatSoft**

# Statistics and Chemometrics for Analytical Chemistry

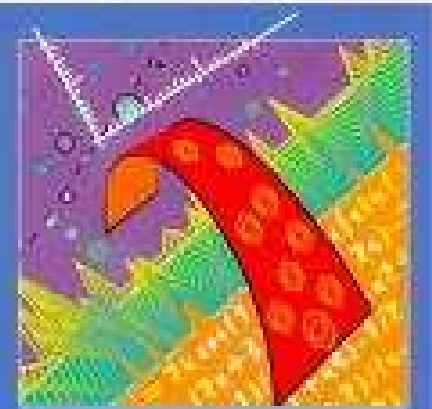
James N Miller & Jane C Miller  
Fifth edition

# Statistics and Chemometrics for Analytical Chemistry

Matthias Otto

# Chemometrics

Statistics and Computer Application in Analytical Chemistry



# STATISTICS in SPECTROSCOPY

Second Edition

Howard Mark  
Jerry Markman, Jr

# ЛИТЕРАТУРА

Свитък лекционен курс – разпечатка и pdf копие <http://web.uni-plovdiv.bg/kmetov>

**Л. Футеков, П. Пенчев, Теория на експеримента, Издателство на ПУ, Пловдив 1992/1998**

**Д. Дончев, М. Дилчева, В. Кинова Практическо ръководство по статистика, Автоспектър, П-в 2002**

К. Дьорфел. Статистика в аналитичната химия, Наука и изкуство 1987

J.C. Miller and J.N. Miller, Statistics and Chemometrics for Analytical Chemistry . Ellis Horwood Ltd. New York 2000

Kellner R., J-M Mermet, M. Otto, H.M. Widmer - "Analytical Chemistry" - An International Undergraduate Textbook, based on the DAC Curriculum, WILEY-VCH, 1998

БДС 17397:2005 Речник на основни и общи термини по метрология Издателство "Стандартизация" (1998)

VIM-International Vocabulary of Basic and General Terms in Metrology, ISO, 2005

БДС EN ISO 17025 Общи изисквания към лабораториите за изпитване и калибриране

Guide to the expression of Uncertainty in measurement ISO, Geneva 1993; ISBN 92-67-10188-9  
Eurachem/CITAC Guide: Quantifying Uncertainty in Analytical Measurement, 2000

Христо Радев, Васил Богев, Неопределеност на резултата от измерването, Софттрайд, София, 2001



# Направете предложение за солеността на Черно море – запишете го в тетрадката си



- Каква величина изписахте ?
- Какво разбирате под “соленост” ?
- Колко е NaCl?
- Каква е мерната единица ?
- Колко цифри използвахте и какво ви дават те ?
- Каква е достоверността на вашето твърдение ?
- Ако сте ползвали справочни данни колко точни са те?

В сравнение с всички океани и морета, които имат почти еднакво съдържание на сол (35‰), солеността на Черно море е два пъти по-ниска (18‰).

Таблични данни за океанска вода:  
23,6 g/l NaCl, 0,64 g/l KCl, 4,53 g/l MgCl<sub>2</sub>·6H<sub>2</sub>O,  
5,94 g/l MgSO<sub>4</sub>·7H<sub>2</sub>O и 1,3 g/l CaCl<sub>2</sub>·2H<sub>2</sub>O

# ЗАЩО УЧИМ СТАТИСТИКА И МЕТРОЛОГИЯ ?

- Опишете разликата между:

интуиция –

знание –

компетентност -

ПРОФЕСИОНАЛИЗЪМ –



- МОЖЕ ЛИ ДА ПРЕДСКАЗВАМЕ БЪДЕЩЕТО ?
- Как да се договаряме, да оценяваме риска и гарантираме доверие ?
- Как да оценяваме качеството? – КВАЛИМЕТРИЯ спец. курс
- Как да представяме и интерпретираме данни от химични измервания съгласно новите евро-изисквания ?

# Днес ще разгледаме:

## 1. Какво изучава МЕТРОЛОГИЯТА

- Евро-програми и стандартизация
- Мерни единици

## 2. Метрологична ПРОСЛЕДИМОСТ на измерванията

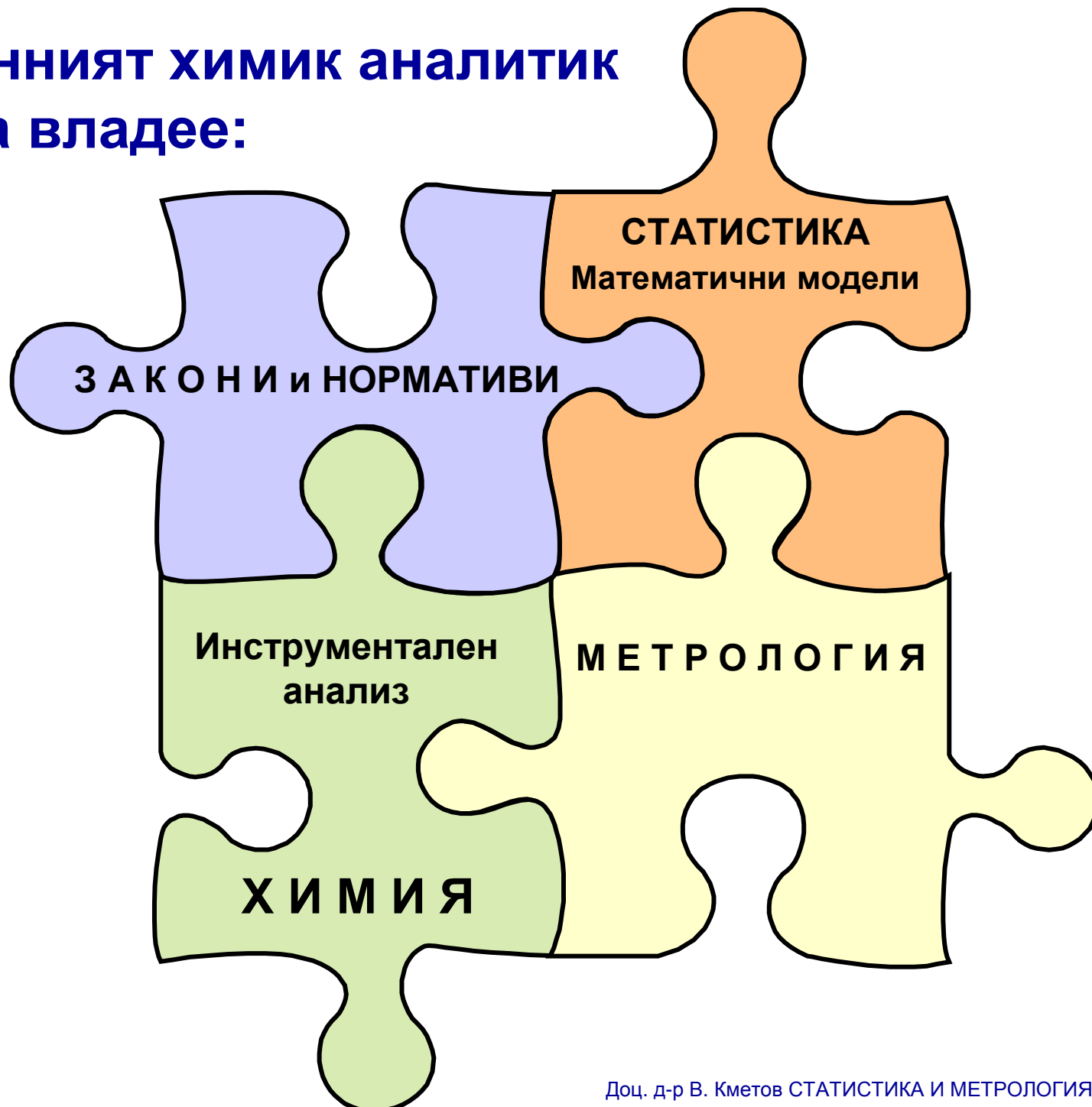
- Йерархия на съподчиненост на еталоните
- Кой я осигурява ?

## 3. Що е СТАТИСТИКА ?

## 4. Елементи на теория на вероятностите.

- Комбинаторика - някои основни понятия:  
*пермутации, вариации, комбинации*

# Съвременният химик аналитик трябва да владее:







EBI

Директорати

Външни  
Отношения

Проучвания  
Наука  
Технология

Обед  
Центр  
изслед

IRMM  
Гейл

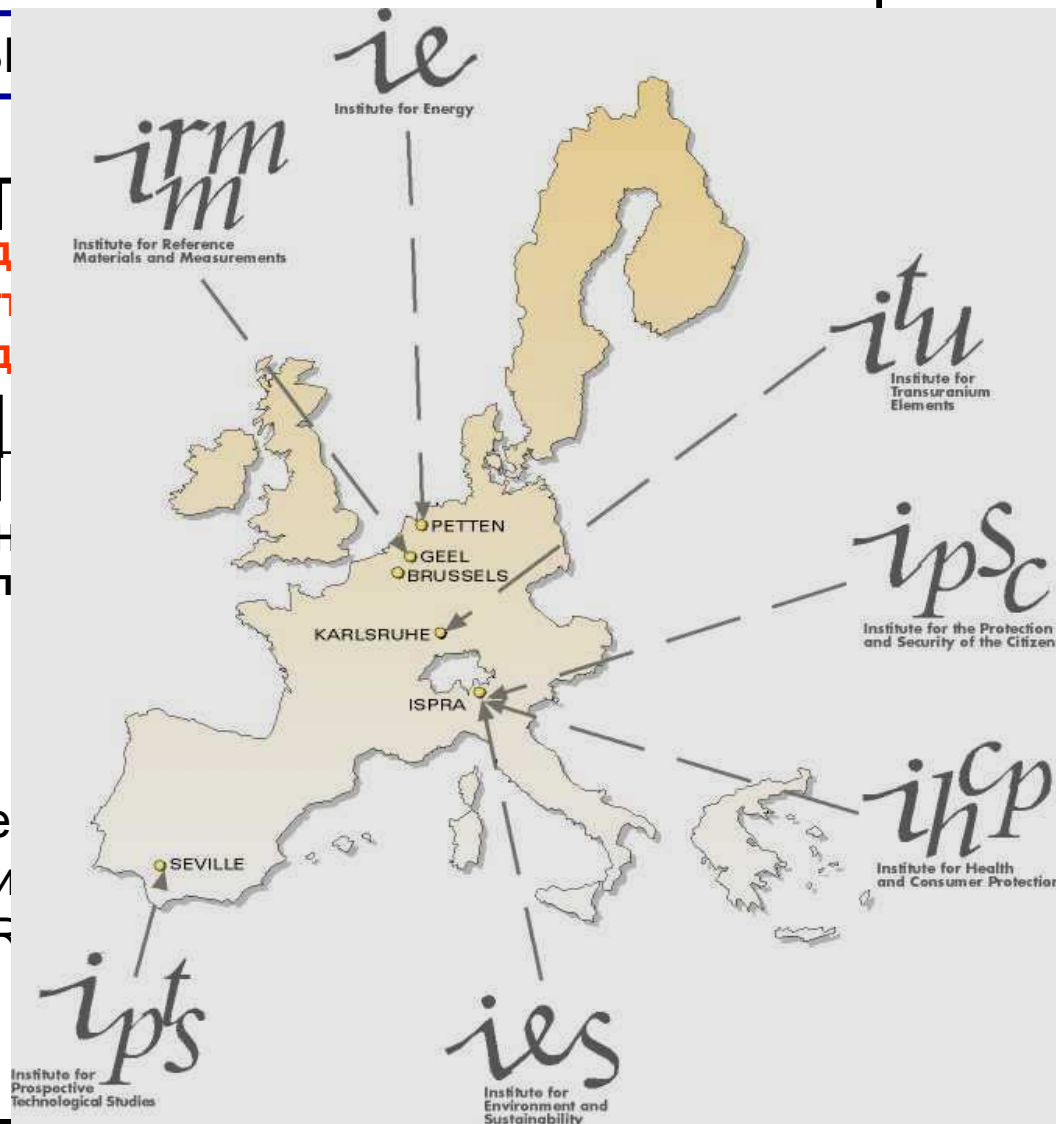
ITU  
Карлсруе

IPSC-И  
Исп

Неутрони  
Данни

Изотопни Измервания  
Метрология в Химията

Рефере  
Матери  
(BCR





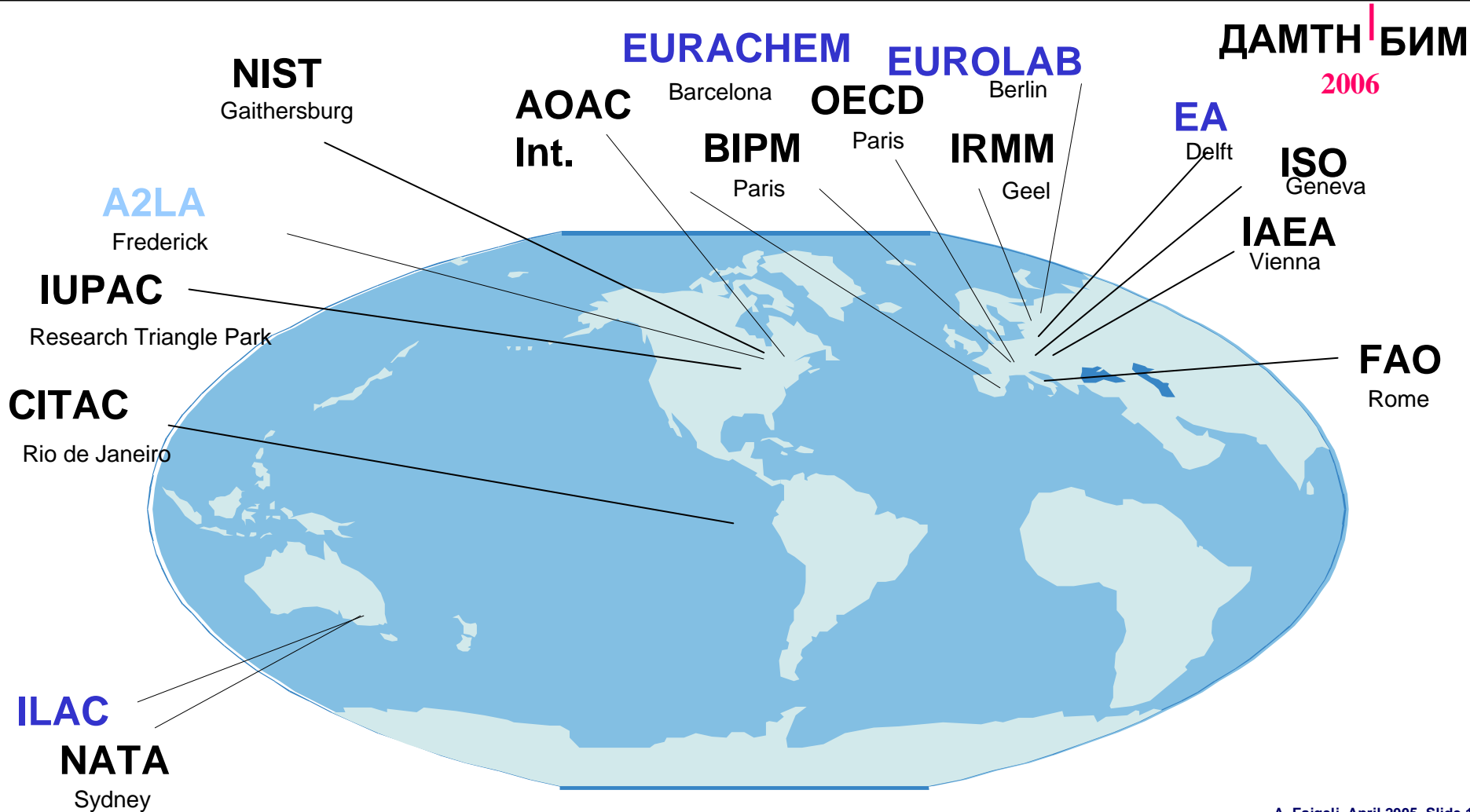
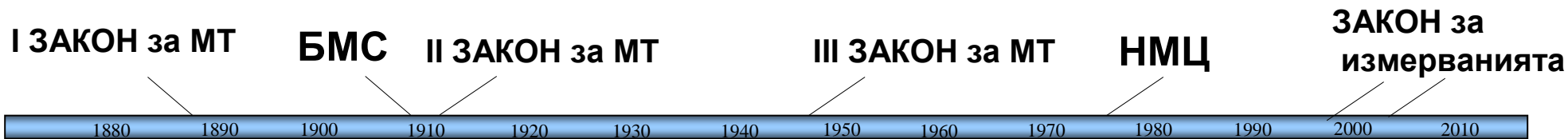
# TrainMiC

Training in Metrology in Chemistry

TrainMiC is organised and conducted by JRC-IRMM as a common international information platform to be used as a tool for the education in metrology in chemistry.



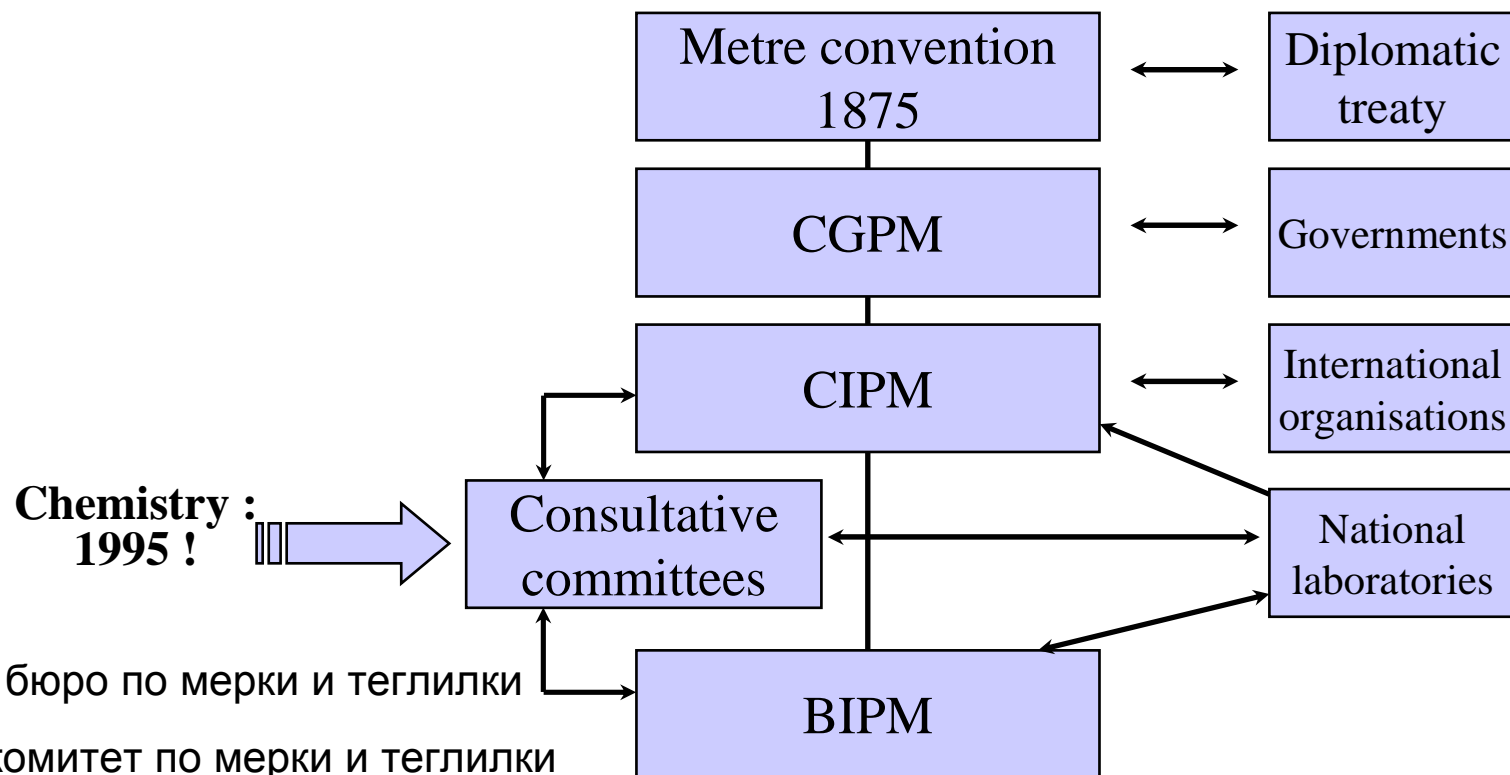
# Организации в областта на стандартизацията и метрологията



A. Fajgelj, April 2005, Slide 16 of 37



# Международна скала на измерванията



**BIPM** – Международно бюро по мерки и теглилки

**CIPM**- Международен комитет по мерки и теглилки

**CGPM** – Генерална конференция по мерки и теглилки

**CC** – Консултативни комитети

**БИМ** – Български Метрологичен институт  
**ДАМТН** – Държавна агенция за метрологичен и технически надзор

**НЦМ** – Национален център по Метрология

Всяко измерване е съпроводено с не съвсем известни отклонения, така че значимостта, която може да се придаде на измерването, трябва да отчита съответстващата **неопределеност**. Ето защо ние трябва да изразим с **ПРЕЦИЗНОСТ** самата тази **НЕПРЕЦИЗНОСТ**.

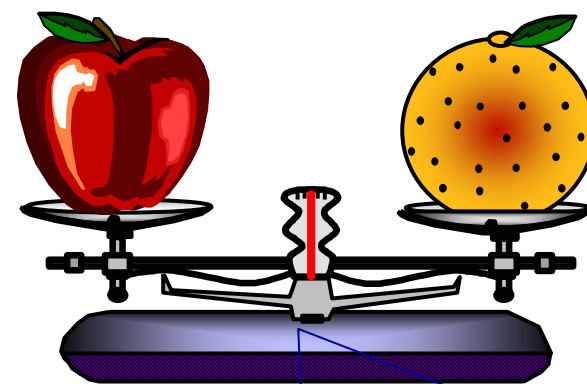
Prerre Giacomo - директор **VIM**

**ИЗМЕРВАНЕТО Е СРАВНЕНИЕ !**

**АНАЛИЗЪТ СРАВНЯВА  
ВЗАИМОДЕЙСТВИЯ**



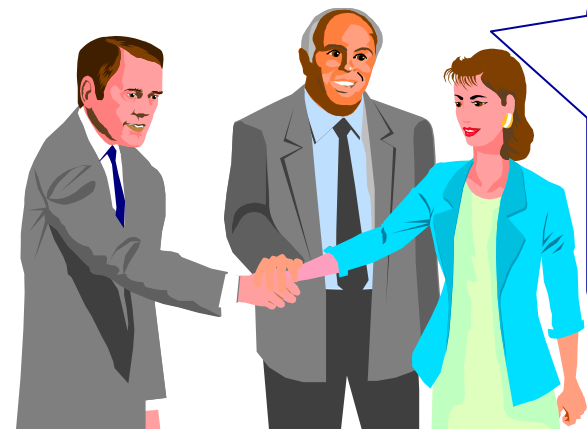
Какво прави аналитикът ?  
Какво продава ?  
Какъв е резултатът ?



Еднакво ли е теглото на ябълката и  
портокала ?  
**ПРИБЛИЗИТЕЛНО ДА!**  
**ИЗМЕРВАНЕТО СЪДЪРЖА**  
**НЕОПРЕДЕЛЕНОСТ (uncertainty)**

Резултатът от (химичното) измерване е просто  
**ДЕКЛАРАЦИЯ** за изолирано изчисление,  
**докато не се докаже противното !**

# КАК да се ГАРАНТИРА достоверността на резултата от измерването ?



ЗАКОН за акредитацията, извършвана от  
Изпълнителна Агенция Българска служба  
за акредитация **ИА БСА**  
в сила от 14.01.2006 г

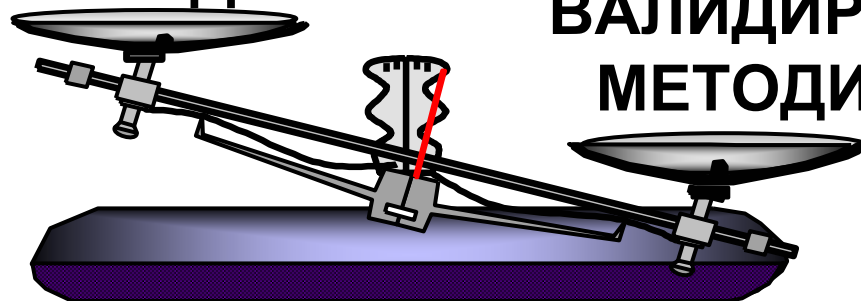
(2) Акредитация по смисъла на този закон  
е признаването от ИА БСА на  
компетентността на физически или  
юридически лица да извършват  
дейностите по чл. 1, ал. 1.

Документирани на  
цялостната процедура по  
един **прозрачен, ясен и**  
**стандартизиран начин**

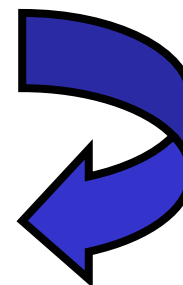
ISO/EC 17025:2005

**СТАНДАРТИЗИРАНИ**

**МЕТОДИ**



**ВАЛИДИРАНИ**  
**МЕТОДИКИ**



**Д О В Е Р И Е**



# КАКВО СЕ РАЗБИРА ПОД ПОНЯТИЕТО КАЧЕСТВО?

**Качество** - е съвкупността на характеристики на дадена единица, отнасяща се до нейната пригодност да изпълни указани или очаквани **изисквания**.

ISO8402 1994 Quality Measurement and Quality Assurance - Vocabulary

## ГАРАНТИРАНЕ КАЧЕСТВОТО НА РЕЗУЛТАТА ОТ ИЗМЕРВАНЕ

**Проследимост**  
(моя резултат е сравним - общ еталон)

**Бюджет неопределеността**  
(Колко добре познавам резултата )

**Валидирани методи**  
(метод, подходящ за целта )  
(fit –for-porous)

## ISO/EC 17025:(1999/2005)

General requirements for the competence of calibration and testing laboratories

## БДС ЕН ISO/EC 17025:2006

Общи изисквания към лабораториите за изпитване и калибриране

### Управленчески изисквания

- Обучение на персонала /компетентност/
- Контрол на документите
- Контрол на записите
- Съответствие на техническите средства
- Отговорности



#### ЗАКОН ЗА ИЗМЕРВАНИЯТА

- Чл. 1. Този закон урежда обществените отношения, свързани с осигуряване на проследимост, точност и достоверност на измерванията.

### Технически изисквания

- Процедури на **валидиране**
- Използване на ССМ
- Бюджет на **неопределеността**
- **Проследимост** на резултатите
- Междулабораторни сравнения

Валидиране е **потвърждение** чрез изследване и представяне на обективни доказателства, че конкретните изисквания за дадено специфично **предназначение** са изпълнени  
**(fit for purpose)**

# Метрология в Химията

*НОВА метрологична инфраструктура  
за химичните измервания*



**ОБЩ ЕЗИК - VIM**



**ОБЩ ПОДХОД ЗА ИЗРАЗЯВАНЕ НА  
КАЧЕСТВОТО НА ИЗМЕРВАНИЯТА - GUM**



**СРАВНИМОСТ НА РЕЗУЛТАТИТЕ И  
ВЗАИМНО ПРИЗНАВАНЕ - MRA**

## **VIM International Vocabulary of Basic and General Terms in Metrology ISO-1993 (2005)**

**БДС 17397:1998 Речник на основни и общи термини по метрология** Издателство "Стандартизация" (Проект 2005)

### В ТЕРМИНОЛОГИЯТА

- Нови термини - (проследимост; неопределеност; En....)
- Стари термини - (стандартно отклонение; корелация .....
- Видоизменени термини - (фактор на покритие; възпроизводимост....)

Споразумение за  
взаимно признаване

измерен веднъж,  
приет навсякъде

**the Mutual Recognition Arrangement**



**Mutual recognition**  
of national measurement standards  
and of calibration and measurement certificates  
issued by national metrology institutes

Paris, 14 October 1999

Comité international des poids et mesures

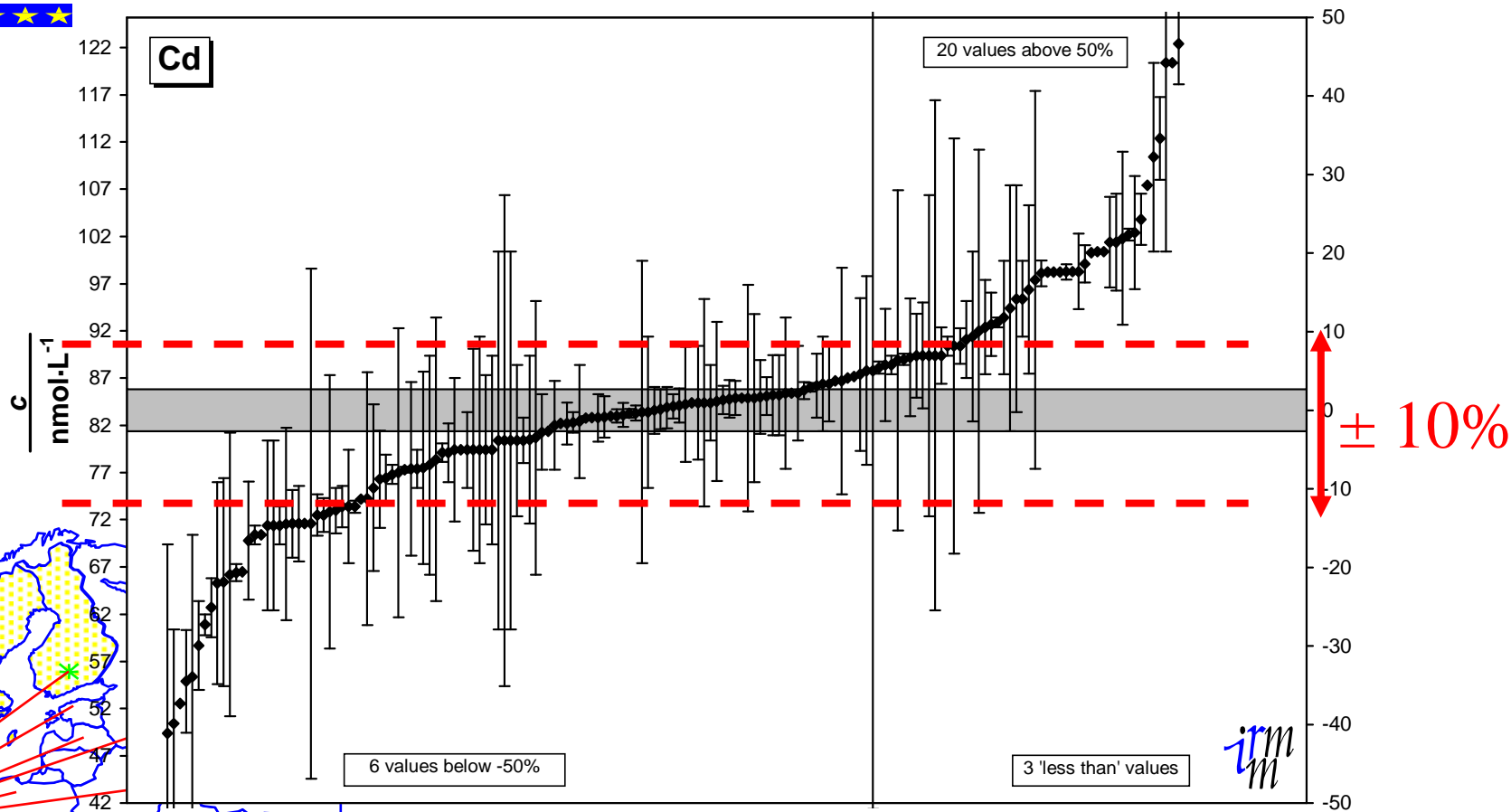
Bureau  
international  
des poids  
et mesures

Organisation  
intergouvernementale  
de la Convention  
du Mètre



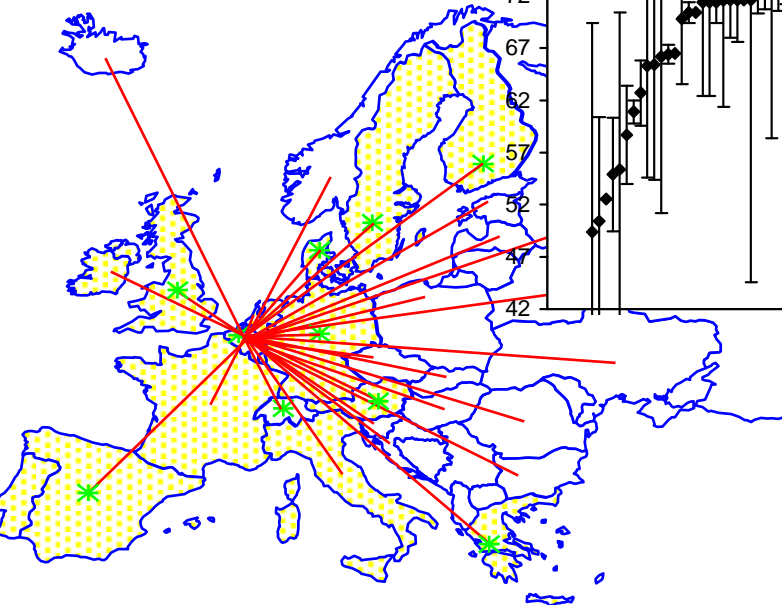
# IMEP- 9 : Trace elements in Water

Certified range ( $\pm U=2u_c$ ): 81.0 - 85.4 nmol·L<sup>-1</sup>



Results from all laboratories.

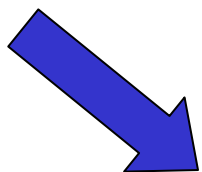
**ИЗИСКВАНО КАЧЕСТВО!**



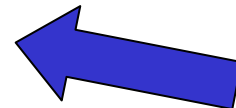
\* регионален съ-организатор

Валидирани на методи

*( метод, подходящ за целта )*

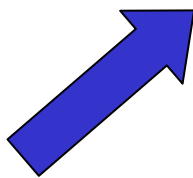


**Гарантиране  
качеството на  
резултата от  
измерване**



**Проследимост**

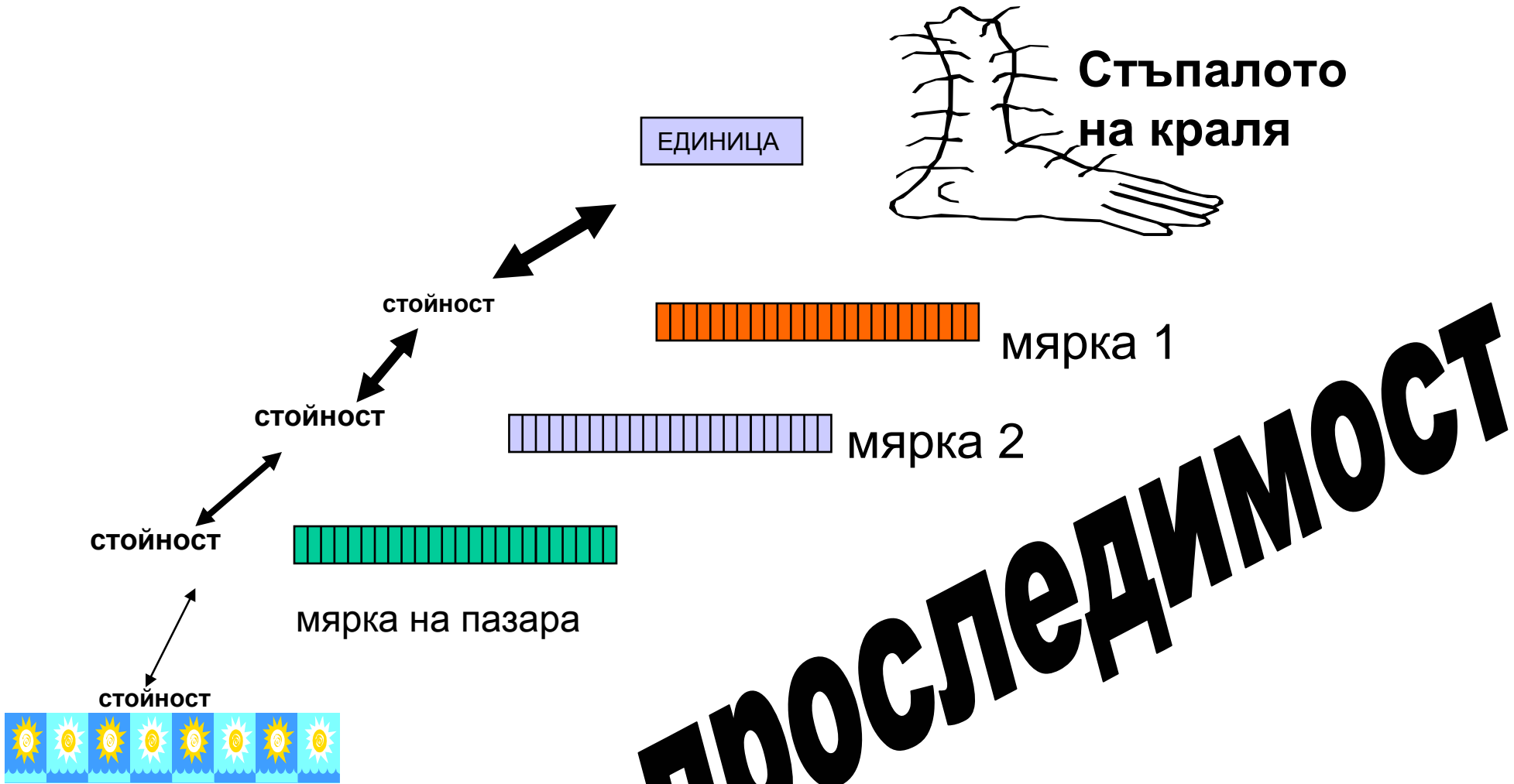
*(моя резултат е сравним -  
общ еталон)*



**Бюджет на неопределеността**

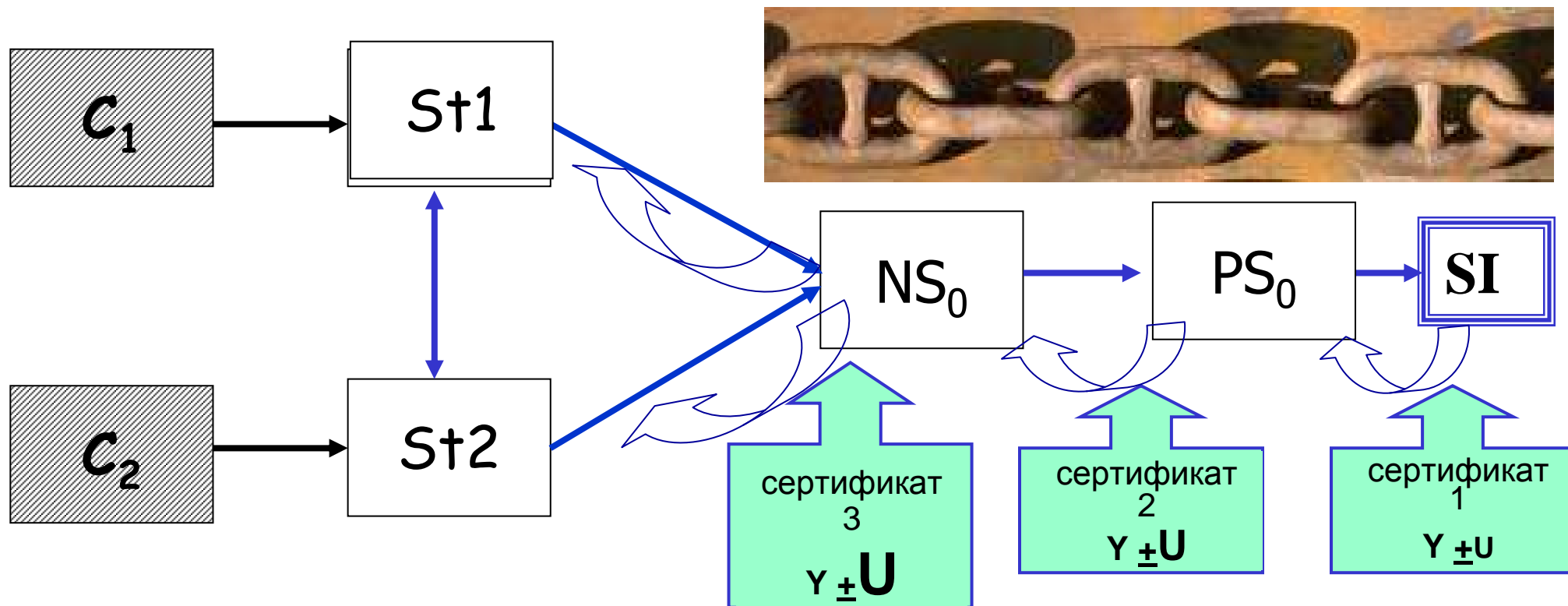
*(Колко добре познавам резултата )*

# ПРОСЛЕДИМОСТ при измерване на дължина



**ПРОСЛЕДИМОСТ**

## Асоциативност на сравнимостта – избиране на сравнителен материал



$C_1$  и  $C_2$  сравними ли са?

Да, чрез обща връзка  
йерархия на съподчиненост

# Метрологична ПРОСЛЕДИМОСТ (*Traceability*) ОПРЕДЕЛЕНИЕ

Проследимостта е свойство на резултат от измерване свързан с обявена метрологична основа (референт) чрез непрекъсната верига от калибрирания на система за измервания или сравнения като всяко има принос към обявената неопределеност на измерванията.

1

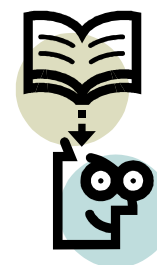
2

3

4

[VIM, проект 2005]

БЪДЕТЕ ГОТОВИ, ДА  
ПРЕДСТАВИТЕ ОБЕКТИВНИ  
ДОКАЗАТЕЛСТВА ЗА  
ПРОСЛЕДИМОСТТА НА  
ВАШИТЕ ИЗМЕРВАНИЯ



# Йерархия при калибриране

Метрологична проследимост ↑

## Средства

- Първичен СМ
- Международен СМ
- Национален СМ
- Сертифициран СМ
- Трансферен СМ
- Пътуващ СМ
- Работен СМ

неопределеност

## Осигуряващи услугата

- ВІРМ
- Нац. Метрол.Институти
- Акред. Калибриращи Лаб.
- Компании (in-house)
  - център за калибриране
  - тестова лаборатория

↓ Неопределеност

[ILAC-G2:1994]

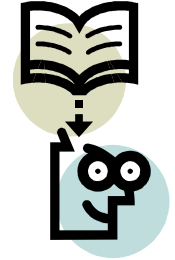
*Traceability of Measurements*

**сравнявай своите резултати със стандартни образци**  
**СВЪРЗАНИ С ВЪНШНИЯ СВЯТ**

# Дефиниции (VIM) БДС 17397 :

## Величина – (*quantity*)

Свойство на явление, тяло или вещество, на което може да се препише големина



## Единица; измерителна единица - *unit (of measurement)*

скаларна величина, определена и приета със спогодба, с която се **сравняват** други величини от същия вид, за да се изразят техните големина

## Международна система за единици SI - *International System of Units, SI*

Кохерентна система от единици, приета от Генералната конференция по мерки и теглилки (CGPM)

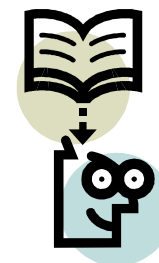
## SI основни единици

Величина	Наименование	Означение	ЕТАЛОН представляващ единицата
Дължина	Метър	m	
Маса	Килограм	kg	
Време	Секунда	s	
Електрически ток	Ампер	A	
Термодинамична темп.	Келвин	K	
Количество вещество	Мол	mol	
Интензитет на светлината	Кандела	cd	

## Стойност (на величина) - *value (of a quantity)*

Големината на дадена величина, изразявана като произведение на измерителната единица и число

**Дефиниции:**



# **СТАТИСТИКА**

**наука за събиране и изучаване  
на числови факти или данни и  
интерпретирането им  
с помощта на математически средства  
в тясна връзка  
с теорията на математическите вероятности**

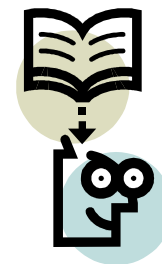


Статистиката е като бикините – показва много, и то привлекателни неща, но почти винаги най-важното остава скрито





# Комбинаторика (1)



ВАРИАНТ - ѿ

**множество**  $M = \{a_1, a_2, a_3, \dots, a_n\}$  с  $n$  елемента

$n=3$   $M \{Na Mg Zn\}$

**извадка** – подмножество от кои да е  $k$  елемента на  $M$   
*наредени, ненаредени, с повторение и без повторение*

1. Na Mg Zn
2. Na Zn Mg
3. Mg Zn Na
4. Mg Na Zn
5. Zn Na Mg
6. Zn Mg Na

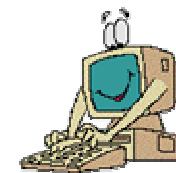
**пермутация** – наредена извадка без повторение от  $n$  елемента с обем  $n$

**$M$  е без еднакви елементи**  
**елементи**

$$P_n = n! = 1 * 2 * \dots * (n-1) * n$$

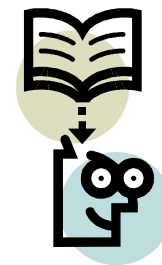
**$M$  е с  $j$  еднакви**

$$P_n^{r,s,t} = \frac{n!}{r! * s! * t!}$$





# Комбинаторика (2)



множество  $M = \{a_1, a_2, a_3, \dots, a_n\}$  с  $n$  елемента

**ВАРИАНТ -  $\omega$**

$n=3; k=2$   $M \{Na Mg Zn\}$

## Вариация

Всички **наредени** извадки от  $n$  елемента с обем  $k$  ( $k \leq n$ ) образуват множеството на **вариациите** от  $n$  елемента  $k$ -ти клас

Без повторение:

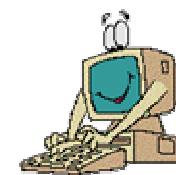
$$V_n^k \binom{n}{k} = \frac{n!}{(n-k)!}$$

1. Na Mg
2. Na Zn
3. Mg Zn
4. Mg Na
5. Zn Na
6. Zn Mg

С повторение:

$$\tilde{V}_n^k \binom{n}{k} = n^k$$

1. Na Mg
2. Na Zn
3. Mg Zn
4. Mg Na
5. Zn Na
6. Zn Mg
7. Na Na
8. Zn Zn
9. Mg Mg





# Комбинаторика (3)

$n=3; k=2$  M {Na Mg Zn}

## ВАРИАНТ - $\infty$

### КОМБИНАЦИЯ

Всички **НЕНаредени** извадки от  $n$  елемента с обем  $k$  ( $k \leq n$ ) образуват множеството на **комбинациите** от  $n$  елемента  $k$ -ти клас

Без повторение:

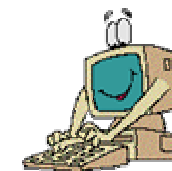
$$C_n^k \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

1. Na Mg и Mg Na
2. Na Zn и Zn Na
3. Mg Zn и Zn Mg

С повторение:

$$\tilde{C}_n^k \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}$$

1. Na Mg и Mg Na
2. Na Zn и Zn Na
3. Mg Zn и Zn Mg
4. Na Na
5. Mg Mg
6. Zn Zn



# Днес разгледахме:

## 1. Какво изучава МЕТРОЛОГИЯТА

- Евро-програми и стандартизация
- Мерни единици - **домашно!!**

## 2. Метрологична ПРОСЛЕДИМОСТ на измерванията

- Йерархия на съподчиненост на еталоните
- Кой я осигурява ?

## 3. Що е СТАТИСТИКА ?

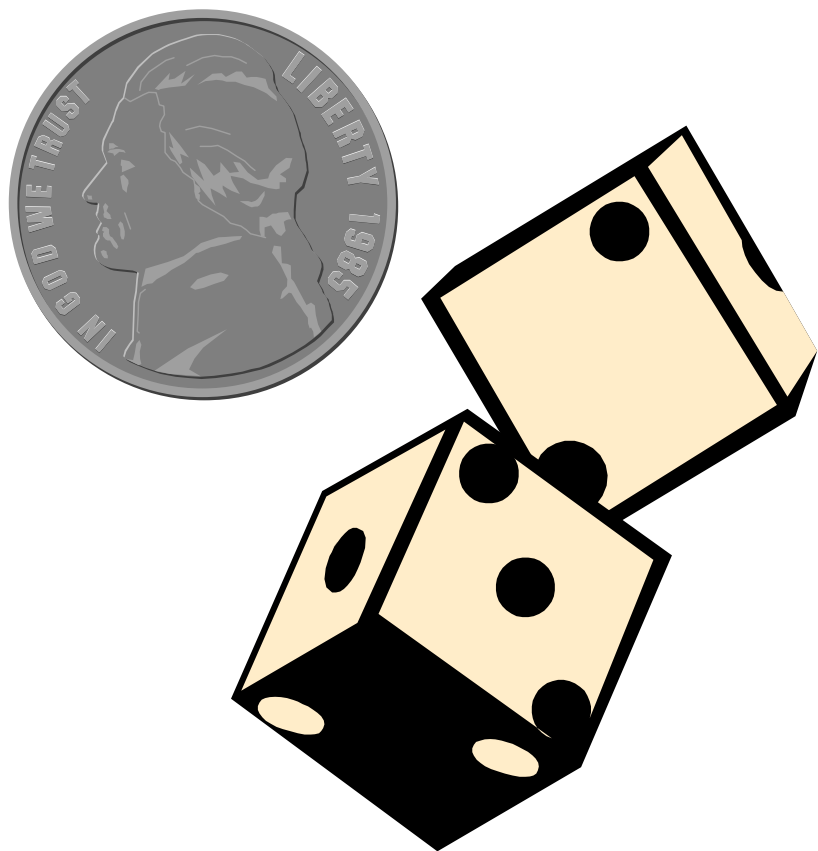
## 4. Елементи на теория на вероятностите.

- Комбинаторика - някои основни понятия:  
*пермутации, вариации, комбинации*

# Задачи за самостоятелна работа

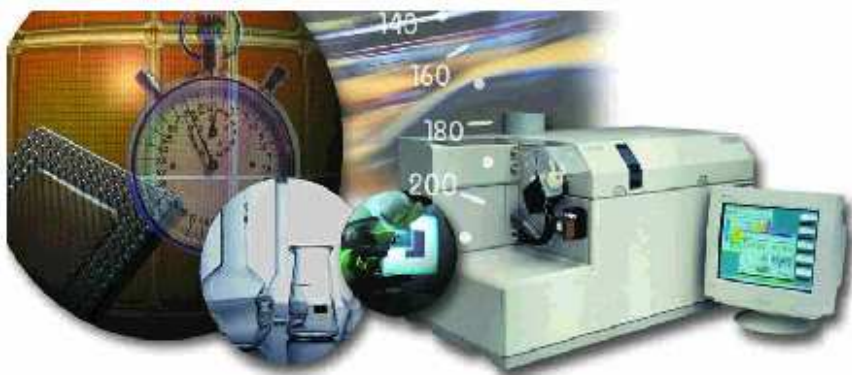


- Намерете следните сайтове и попълнете кой за какво е:
  - <http://www.smb-bg.org> .....
  - <http://www.nab-bas.bg> .....
  - <http://www.iso.org> .....
  - <http://www.ilac.org> .....
  - <http://www.bimp.fr> .....
  - <http://www.sp.se/metrology/eng/terminology.htm> .....
  - <http://www.statsoft.com> .....
  - <http://q-systems.uni.cc/index.php> .....
  - <http://www.irmm.jrc.be/imep/trainmic2> .....
  - <http://www.bds-bg.org> .....



# КЛАСИЧЕСКА ВЕРОЯТНОСТ

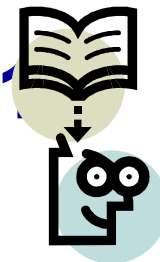
# СЛУЧАЙНИ ВЕЛИЧИНИ



# Днес ще разгледаме:

1. Събития и действия с тях
  - Пространство на елементарните събития
2. Класическа вероятност
  - Теореме за събиране и умножение на вероятности
  - Формула за пълна вероятност. Формула на Бейс
3. Случайни величини (СВ)
  - прекъснати и непрекъснати СВ
  - относителни честоти на поява
  - плътност и функция на разпределение на СВ

# Събития и действия с тях (1)



ОПИТ  $\Leftrightarrow$  ИЗМЕРВАНЕ  $\Leftrightarrow$  СРАВНЕНИЕ

РЕЗУЛТАТ  $\Leftrightarrow$  ЕЛЕМЕНТАРЕН ИЗХОД  $\Leftrightarrow$  СТОЙНОСТ  $\Leftrightarrow$  ЗНАЧЕНИЕ НА СВ  $\Leftrightarrow$  ВАРИАНТ -  $\omega$

**Събитие** – един или група от изходи (варианти, значения, резултата)

Съвкупността от всички **възможни** елементарни изходи за даден опит се нарича **ПРОСТРАНСТВО** на елементарни събития  $\Omega$   
ГЕНЕРАЛНА СЪВКУПНОСТ ??

**СЪБИТИЕ**  $\Leftrightarrow$  Подмножество на елементарни изходи  $A \in \Omega$

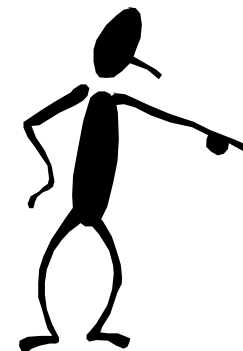


Дайте примери за  $\Omega$  и  $\omega$

- Хвърляне на монета
- Изотопи на водорода
- Изомери на динитробензена
- атомен радиус
- продукция на цимент
- .....

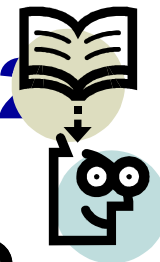
Дайте примерни желани събития

- .....
- .....
- .....
- .....
- .....

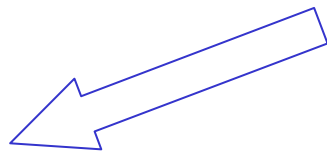




# Събития и действия с тях (2)

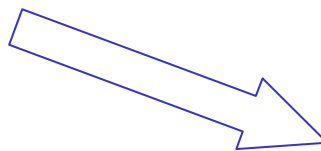


Съвкупността от всички възможни елементарни изходи (варианти) за даден опит се нарича **пространство на елементарни събития**  $\Omega$



**ДИСКРЕТНИ** – ако съдържат краен брой или изброимо много елементарни събития

- Хвърляне на монета
- Изотопи на водорода
- Изомери на динитробензена
- .....



**НЕПРЕКЪСНАТИ** – ако съдържат неизброимо много елементарни събития

- .....
- атомен радиус
- продукция на цимент годишно
- .....

# Събития и действия с тях (3)



Всяко подмножество на  $\Omega$  се нарича **събитие** или **случайно събитие**

Нека  $A, B, C \dots$  са събития

Елементарното събитие  $\omega_1$  е **благоприятно** за  $A$ , ако  $\omega_1 \in A$

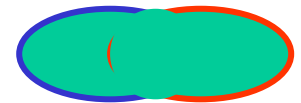
Събитие, което никога не настъпва се нарича **невъзможно  $\emptyset$**   
то не съдържа елементарни събития

Събитие, което винаги настъпва е **сигурно** събитие  
самото  $\Omega$  разглеждано като подсъвкупност на себе си е сигурно събитие,  
то настъпва при всяко едно провеждане на опита



## Връзки между събития

- $A \subset B$   $A$  **влече**  $B$  ако при съдване на  $A$  следва съдване на  $B$
- ако  $A \subset B$  и  $B \subset A$  то събитията са **еквивалентни  $A=B$**
- **обединение** на  $A \cup B$  е събитие, което се съдва при съдването на поне едно от  $A$  и  $B$
- **сечение** на  $A \cap B$  е събитие което се съдва при съдването на двете събития  $A$  и  $B$

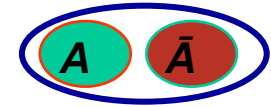


# Събития и действия с тях (4)



## Връзки между събития

• **допълнение** ( $\bar{A}$ ) на събитието  $A$  е събитие, което се сбъдва ако не се сбъдва  $A$

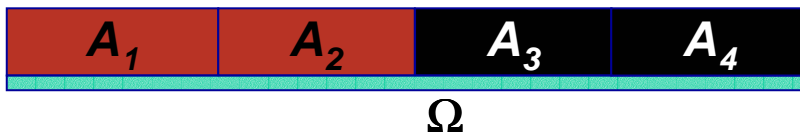


•  $A$  и  $\bar{A}$  се наричат **противоположни** ако за тях е в сила  $A \cup \bar{A} = \Omega$  и  $A \cap \bar{A} = \emptyset$

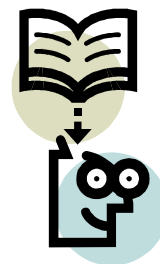
• Две събития се **несъвместими** ако  $A \cap B = \emptyset$  т.е. не могат да се сбъднат едновременно

• **Пълна група** несъвместими събития имаме, ако събитията  $A_i$  ( $i=1,2,\dots,n$ ) удовлетворяват условията:

$$\bigcup_{i=1}^n A_i = \Omega \quad \text{и} \quad A_i \cap A_j = \emptyset, i \neq j, j = 1, 2, \dots, n$$



# Какво е вероятност ?



**ВЕРОЯТНОСТ** - математически израз на възможността да се случи определено **събитие**.

Ако събитието ще настъпи с абсолютна сигурност - вероятността е **1** **приведено в проценти 100%**,

ако е сигурно, че няма да настъпи вероятността е **0** **приведено в проценти (0% )**

**МОДЕЛ ЗА ОПИСАНИЕ  
НА ВЪЗМОЖНОСТ  
ЧРЕЗ ЧИСЛО**

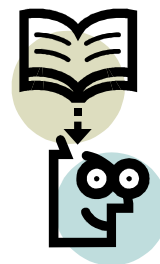


Дайте пример за 200 % вероятност.

Може ли да има отрицателна вероятност ?



# Класическа вероятност (1)



Нека  $\Omega$  е основно пространство, което :

1. Има краен **брой възможни** елементарни събития  $k_{\Omega}$
2. Всички елементарни събития са равно възможни

Ако  $A$  е случайно събитие с  $k_A$  **брой благоприятни** елементарни събития, вероятността на  $A$  може да се намери по формулата:

$$P(A) = \frac{k_A}{k_{\Omega}}$$

## СВОЙСТВА

1. За всяко събитие  $A$  е в сила  $0 \leq P(A) \leq 1$ , като  
 $P(A) = 0 \Leftrightarrow A = \emptyset$ ,  
 $P(A) = 1 \Leftrightarrow A = \Omega$ ,

$A$  и  $B$  са несъвместими

2. Теорема за събиране на вероятности:

$$P(A \cup B) = \begin{cases} P(A) + P(B) \\ P(A) + P(B) - P(A \cap B) \end{cases}$$

$A$  и  $B$  са произволни

3. За всяко събитие  $A$  е в сила  $P(\bar{A}) = 1 - P(A)$



# РЕШЕНИЕ ЗАДАЧА

Схема за производство на амоняк включва два независимо работещи агрегата с вероятност за отказ  $P(A) = 0,05$  и  $P(B) = 0,08$ .

• Да се намери вероятността  $P(C)$  да спре процеса на производство на амоняк при положение, че поне единия от агрегатите откаже:

Цехът ще спре ако спре агрегата А **или** агрегата Б (който и да е)

$\Rightarrow C = A \cup B$  тъй като А и В са **съвместими** (може и да спрат едновременно)

$$\Rightarrow P(C) = P(A) + P(B) - P(A) \cdot P(B) = 0,05 + 0,08 - 0,05 \cdot 0,08 = 0,126$$

• Да се намери вероятността  $P(C)$  да спре процеса на производство на амоняк при положение, че и двата от агрегатите откажат:

$\Rightarrow C = A \cap B$

$$\Rightarrow P(C) = P(A) \cdot P(B) = 0,05 \cdot 0,08 = 0,004$$

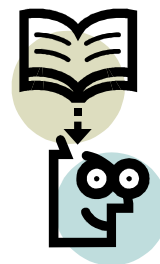
• Да се намери вероятността  $P(C)$  да спре процеса на производство на амоняк при положение, че само един от двата е спрял (ако спрат и двата процеса продължава) :

$\Rightarrow C = A \cap (1-B) + (1-A) \cap B$

$$\Rightarrow P(C) = \{P(A) \cdot [1 - P(B)]\} + \{[1 - P(A)] \cdot P(B)\} = 0,05 \cdot (1 - 0,08) + (1 - 0,05) \cdot 0,08 = 0,122$$



## Класическа вероятност(2)



### Условна вероятност

Нека  $A$  и  $B$  са произволни събития от  $\Omega$  и  $P(B) > 0$

Вероятността да настъпи  $A$  при условие, че се е сбъднало  $B$  е **условна вероятност** и се изчислява като:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

### Теорема за умножение на вероятности

Ако  $P(A) > 0$  и  $P(B) > 0$ , то вероятността на събитието  $A \cap B$  е:

$$P(A \cap B) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B)$$

При повече от две събития  $A_1, A_2, A_3, \dots, A_n$ , формулата приема вида:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cap A_2) \cdot \dots \cdot P(A_n/A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

Събитията  $A$  и  $B$  се наричат **независими**, ако  $P(A \cap B) = P(A) \cdot P(B)$

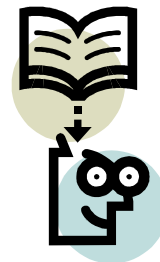
Тогава  $P(A/B) = P(A)$

В склад за химикали с 25 опаковки, 5 са се окислили.  
Ако се вземат две опаковки каква е вероятността и двете да са годни.  
I годна -  $P(A)$  II годна  $P(B)$   $P(C)$  и двете годни  $P(C) = P(A \cap B) =$   
 $P(A) \cdot P(B/A) = 20/25 \cdot 19/24$





# Класическа вероятност (3)



## Формула за пълна вероятност

Целта е да се намери вероятността на събитие **A**, която зависи от събъждането на краен брой предположения (хипотези) – от които винаги се събъдва само едно, т.е.

$$\bigcup_{i=1}^n H_i = \Omega \text{ и } H_i \cap H_j = \emptyset, i \neq j$$

$$\sum_{i=1}^n P(H_i) = 1$$

Следователно  $H_1, \dots, H_n$  образуват пълна група **несъвместими събития**:

Ако са известни  $P(H_i)$   $i = 1 \dots n$ , (наричани **априорни**) и е възможно намирането на условните вероятности на събитието **A** при различните условия (хипотези), то  $P(A)$  се пресмята по формулата за пълна вероятност:

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A / H_i)$$

Вероятностите на  $H_i$ , пресметнати при условие че след експеримента се е събдало събитието **A** се наричат **апостериорни** и се пресмятат по **формулата на Бейс**:

$$P(H_i / A) = \frac{P(H_i) \cdot P(A / H_i)}{\sum_i P(H_i) \cdot P(A / H_i)}$$

# ЗАДАЧА

- В лабораторията има **15** прахообразни реактива, от които **5** са се овлажнили и са негодни. Избират се по случаен начин **3** реактива

Определете вероятността:

A) и трите са качествени

$$P(A) = \frac{C_{10}^3}{C_{15}^3}$$

B) един от извадените е некачествен

$$P(B) = \frac{C_{10}^2 \cdot C_5^1}{C_{15}^3}$$

C) два са некачествени

$$P(C) = \frac{C_{10}^1 \cdot C_5^2}{C_{15}^3}$$

D) нито един не е качествен

$$P(D) = \frac{C_5^3}{C_{15}^3}$$

Имате изпит с конспект 15 въпроса от който пред комисията ще развиете 3. Не сте успели да научите 5 въпроса. Каква е вероятността .....

# Днес ще разгледаме:

## 1. Събития и действия с тях

- Пространство на елементарните събития

## 2. Класическа вероятност

- Теореме за събиране и умножение на вероятности
- Формула за пълна вероятност. Формула на Бейс

## 3. Случайни величини (СВ)

- прекъснати и непрекъснати СВ,
- относителни честоти на поява
- плътност и функция на разпределение на СВ

# Непрекъснати СЛУЧАЙНИ ВЕЛИЧИНИ

Дайте примери за непрекъснати СВ при лабораторните измервания :



.....

.....

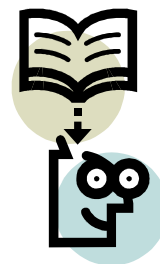
.....

ИСТИНСКА

ВЕЛИЧИНАТА НЯМА СТОЙНОСТ !!

Въпросът с каква вероятност непрекъснатата СВ заема конкретна стойност е **ЛИШЕН ОТ СМИСЪЛ**  
Предварително е ясно, че тази вероятност е нула (клони към нула)

# Дефиниции (1):



## Величина – (*quantity*)

Свойство на явление, тяло или вещество, на което може да се препише големина

### SI основни единици

#### Величина

#### представящ единицата

#### Наименование

#### Означение

#### ЕТАЛОН

Дължина

Метър

m

Маса

Време

Електрически ток

Термодинамична темп.

Количество вещество

Интензитет на светлината

Ампер

Келвин

Мол

Кандела

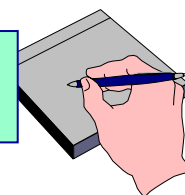
A

K

mol

cd

<http://www.ex.ac.uk/cimt/dictunit/dictunit.htm>



## Стойност (на величина) - *value (of a quantity)*

Големината на дадена величина, изразявана като произведение на измерителната единица и число

## Истинска стойност (на величина) - *true value (of a quantity)*

Стойност, която следва определянето на дадена конкретна величина – тя би се получила чрез **идеално** (съвършено)- измерване

**понятието истинска стойност да се избягва**

## Действителна стойност (на величина) - *conventional true value (of a quantity)*

Приписана стойност на конкретната величина и приета, понякога със спогодба, която има **неопределеност**, подходяща за дадената цел

## Основни единици

Това са основните единици, от които се получават всички останали. Техните [измерения](#) са независими едно от друго.

Име	Означение	Мярка за	Определение
<a href="#">метър</a>	<b>m</b>	Дължина	Единицата за дължина е дължината на пътя, изминат от <a href="#">светлината</a> във <a href="#">вакуум</a> за интервал от време $\frac{1}{299\,792\,458}$ от секундата.
<a href="#">килограм</a>	<b>kg</b>	Маса	Единицата за маса е равна на масата на международния <a href="#">прототип</a> на килограма ( <a href="#">цилиндър</a> , изработен от <a href="#">сплав</a> на <a href="#">платина</a> и <a href="#">иридий</a> ), съхраняван в Международното бюро по мерки и теглилки (BIPM) в <a href="#">Париж</a> .
<a href="#">секунда</a>	<b>s</b>	Време	Единицата за време е продължителността на 9 192 631 770 <a href="#">периода</a> на <a href="#">лъчението</a> , съответстващо на прехода между двете свръхфини нива на основното състояние на атома на <a href="#">Цезий-133</a> .
<a href="#">ампер</a>	<b>A</b>	Електрически ток	Единицата за електрически ток е постоянен електрически ток, който при протичане по два успоредни праволинейни <a href="#">проводника</a> с безкрайна дължина и незначително кръгово напречно сечение, поставени на разстояние 1 метър един от друг във вакуум, създава между тези два проводника взаимодействие със сила $2 \cdot 10^{-7}$ нютона на всеки метър от тяхната дължина.
<a href="#">келвин</a>	<b>K</b>	Термодинамична температура	Единицата за термодинамична <a href="#">температура</a> представлява 1/273,16 част от термодинамичната температура на <a href="#">тройната точка</a> на <a href="#">водата</a> .
<a href="#">мол</a>	<b>mol</b>	Количество вещество	Единицата за количество вещество е количеството вещество на система, съдържаща толкова структурни единици (елементи), колкото <a href="#">атома</a> се съдържат в 0,012 килограма <a href="#">въглерод</a> 12 (структурните единици са $6,023 \times 10^{23}$ или <a href="#">число на Авогадро</a> ).
<a href="#">кандела</a>	<b>cd</b>	Интензитет на светлина	Единицата за интензитет на светлина е интензитетът на светлината в дадена посока на източник, излъчващ монохроматично лъчение с <a href="#">честота</a> $540 \cdot 10^{12}$ херца и интензитет в тази посока 1/683 вата на стерадиан.

## Дефиниции (2):



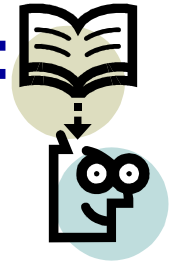
**СЛУЧАЙНА ВЕЛИЧИНА - *variable* (СВ )** - стойностите и се различават една от друга независимо от постоянството на експерименталните условия  
количествена мярка за характеристика на наблюдаван процес или явление и под влияние на случайни фактори приема различни стойности (значения)

**Неслучайна величина (константа)** - величина с точно определена стойност която може да се предскаже преди експеримента и резултатите от измерванията са равни една на друга.

**Дискретна величина** - възможните значения на една СВ са краен или изброимо безкраен брой. (има стъпка)

**Непрекъснатата величина** - ако съдържа неизброимо много елементарни събития (значения)

## Дефиниции (3):



**Честота на поява** - броя опити при които СВ заема дадено значение  
 $k_A$  брой благоприятни елементарни събития

**Относителна честота** на поява на значение на СВ се дефинира като **отношението "M/N"** между честотата на поява на дадено значение **M** към общия брой значения **N**  
 $k_\Omega$  брой възможни елементарни събития

При увеличаване на броя на измерванията относителните честоти се стабилизират и клонят във вероятностен смисъл към определени числа наречени вероятности.

*При увеличаване на измерванията  $\Rightarrow N$  достатъчно голямо  
 $\Rightarrow M$  клони към  $k_A$  а  $N$  клони към  $k_\Omega$*

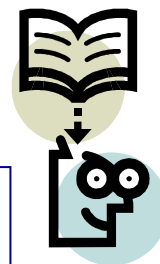
благоприятни

$$\frac{M_A}{N} \cong P(A) = \frac{k_A}{k_\Omega}$$

всевъзможни



# Дискретни СВ



Съответствието между стойностите на дискретна СВ  $X_i$  и вероятностите с които тя заема тези стойности се нарича **закон за разпределение на дискретна случайна величина**

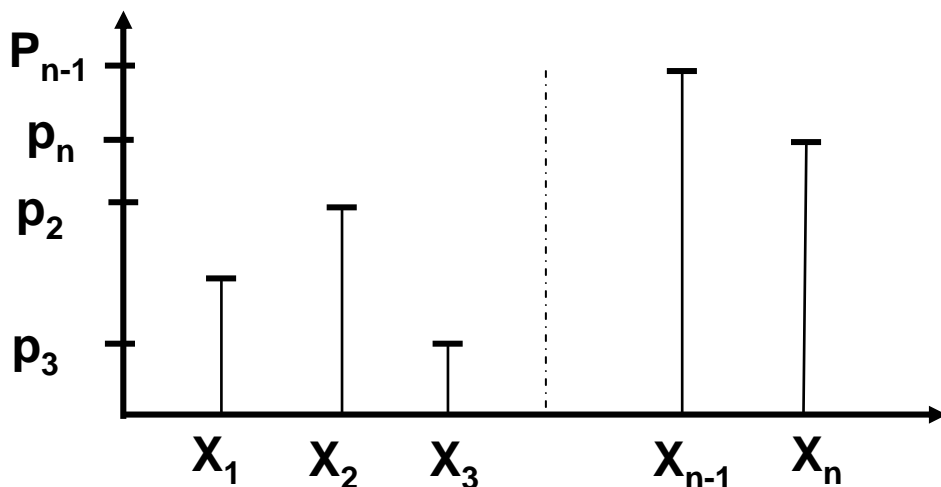
Дискретна СВ е **зададена** ако се знае законът и на разпределение

- Законът може да се зададе най-лесно таблично при което е в сила

$\xi$	$x_1$	$x_2$	...	$x_n$
$P(\xi)$	$p_1$	$p_2$	...	$p_n$

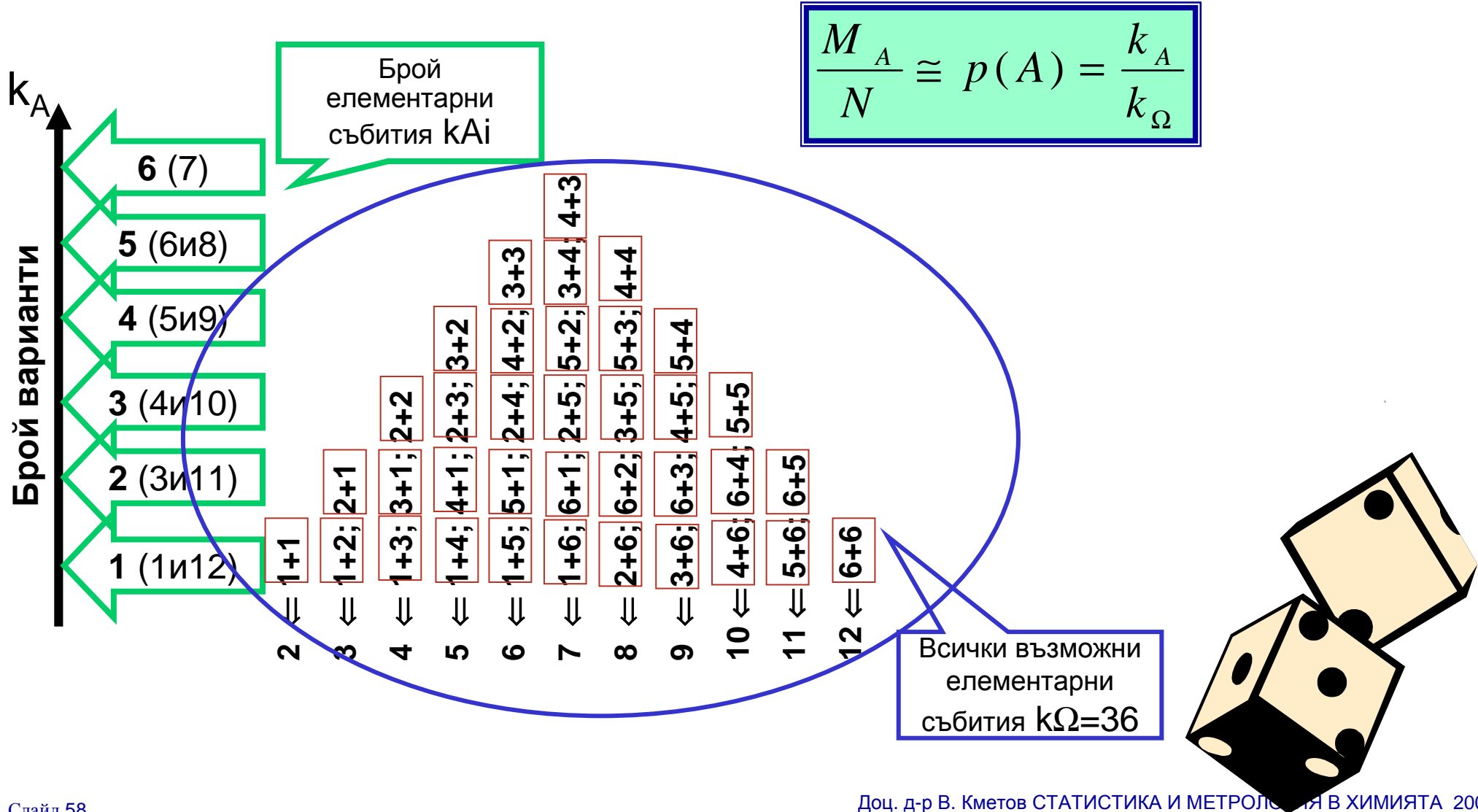
$$\sum_{i=1}^n p_i = 1$$

- Законът може да се зададе графично



Представете графично закона на разпределение на точки от хвърлено зарче

# Изчисляване вероятността да се падне определена сума точки при хвърляне на 2 зарчета



# Изчисляване на вероятността за да се падне определена сума от две хвърлени зарчета

Вероятностен ред ЗАКОН НА РАЗПРЕДЕЛЕНИЕ на случайната величина

сл. величина	2	3	4	5	6	7	8	9	10	11	12
вероятност	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Функцията на разпределение на случайната величина "сума от точките на две хвърлени зарчета"

сл. величина	2	3	4	5	6	7	8	9	10	11	12
вероятност	1/36	3/36	6/36	10/36	15/36	21/36	26/36	30/36	33/36	35/36	36/36

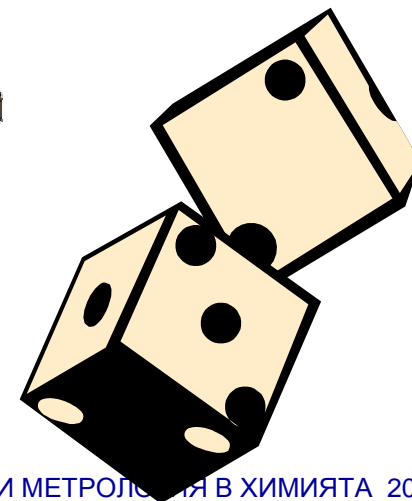
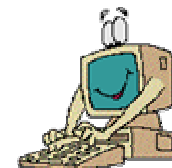
## ФУНКЦИЯ на РАЗПРЕДЕЛЕНИЕ

кривата която показва натрупаната честота на поява на значенията на СВ по-малки или равни на дясната граница да се паднат четен брой точки



Изчислете вероятността да се паднат по-големи и равни на 10

.....



Съдържание на Zn mg/L  
в питейна вода



# ЧЕСТОТИ НА ПОЯВА НА непрекъснатата СВ

Как да определим  
честоти на поява ?

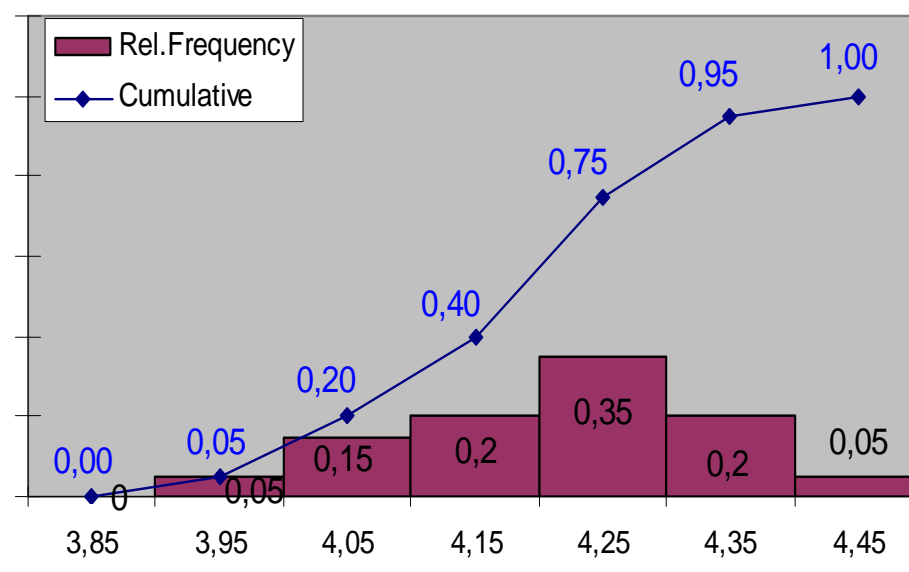
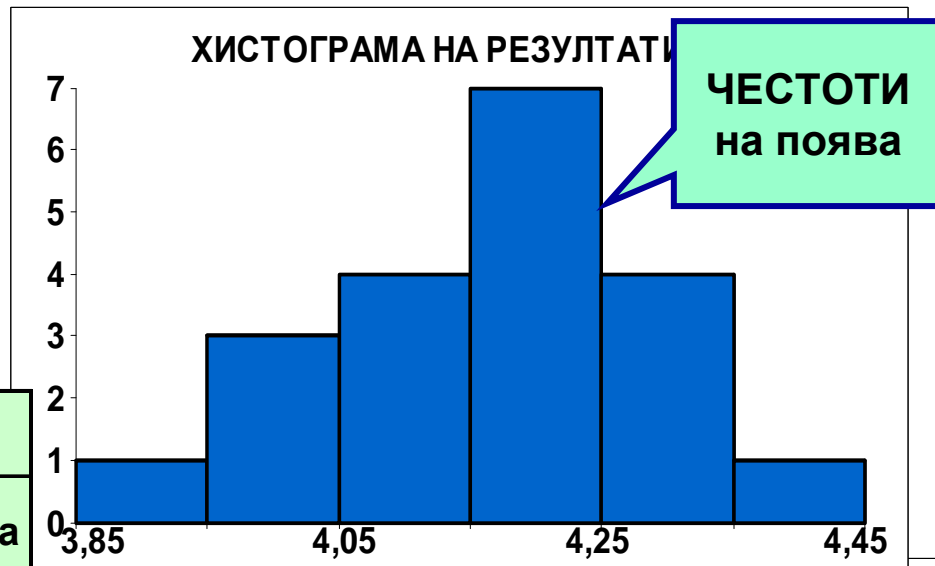
относителните

Как да построим  
функцията на  
разпределение ?

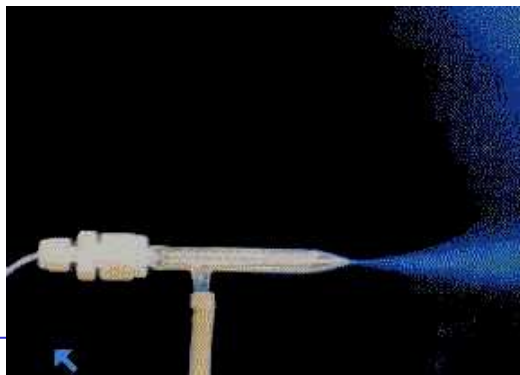
функцията на

## Съдържание на цинк в питейна вода

№ на измерването	Съдържание на Zn mg/l	№ на измерването	Съдържание на Zn mg/l
1	4,23	11	4,11
2	3,97	12	4,05
3	4,18	13	4,27
4	4,29	14	4,43
5	4,00	15	4,31
6	4,17	16	4,15
7	4,12	17	4,24
8	4,08	18	4,18
9	4,20	19	4,25
10	3,88	20	4,35



# Размер на капчици

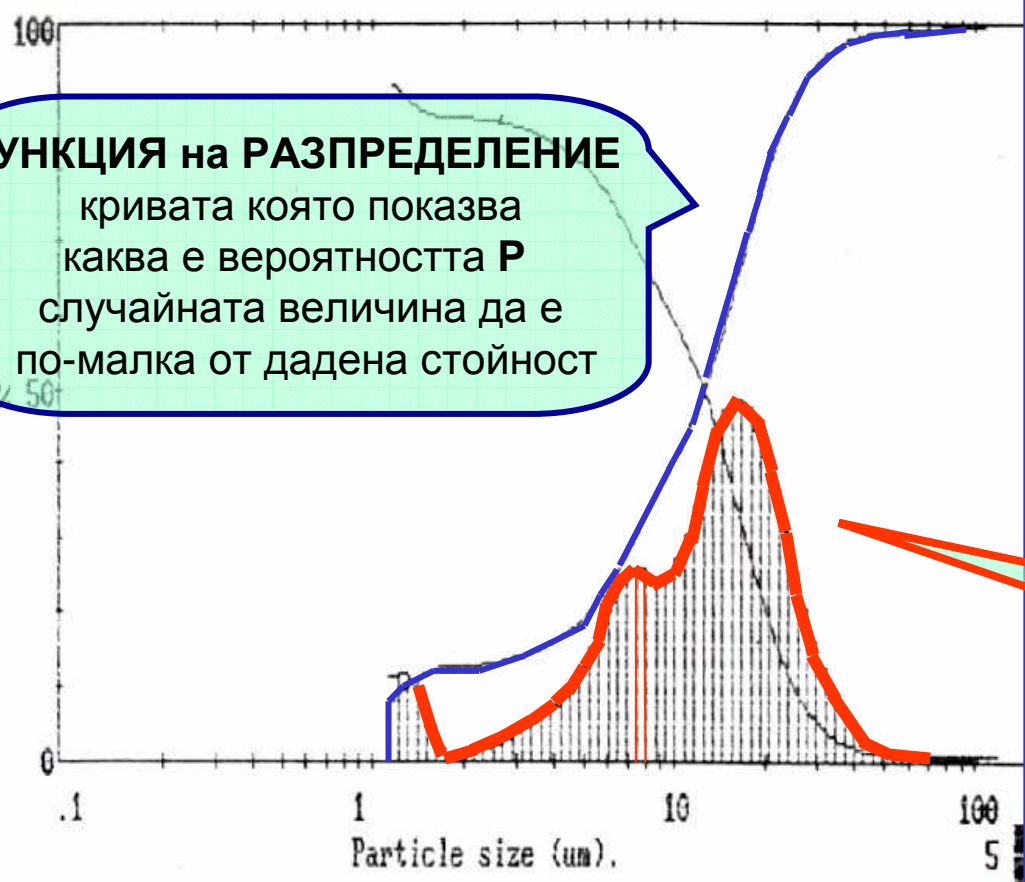


MALVERN

Upper in Lower Under				Upper in Lower Under				Upper in Lower Under			
36.3	2.4	31.4	92.7	6.18	3.4	5.33	20.1				
31.4	5.3	27.1	87.3	5.33	2.3	4.60	17.8				
27.1	7.0	23.3	80.4	4.60	2.0	3.97	15.8				
23.3	6.9	20.2	73.5	3.97	1.8	3.42	14.0				
20.2	7.5	17.3	66.0	3.42	1.4	2.95	12.6				
17.3	7.4	15.0	58.6	2.95	0.8	2.55	11.8				
15.0	7.1	12.9	51.5	2.55	0.5	2.19	11.3				
12.9	5.7	11.1	45.8	2.19	0.2	1.90	11.0				
11.1	5.2	9.63	40.6	1.90	0.3	1.64	10.8				
9.63	4.6	8.31	35.1	1.64	1.3	1.41	9.5				
8.31	4.2	7.16	28.8	1.41	2.0	1.22	7.5				
7.16	5.3	6.18	23.5	1.22	7.5	0.32	0.0				

118 0.0 102 100  
 102 0.0 99.2 100  
 18 Jun 1996 11:19 a

**ФУНКЦИЯ на РАЗПРЕДЕЛЕНИЕ**  
 кривата която показва  
 каква е вероятността  $P$   
 случайната величина да е  
 по-малка от дадена стойност



**ПЛЪТНОСТ на РАЗПРЕДЕЛЕНИЕ  $p(X)$**   
 кривата получена от относителните честоти  
 на поява на стойностите на СВ

# Честоти на поява на резултати от спектрометрично измерване

Trial	Absorbance	Trial	Absorbance
1	0.488	26	0.483
2	0.480	27	0.482
3	0.486	28	0.491
4	0.473	29	0.481
5	0.475	30	0.469

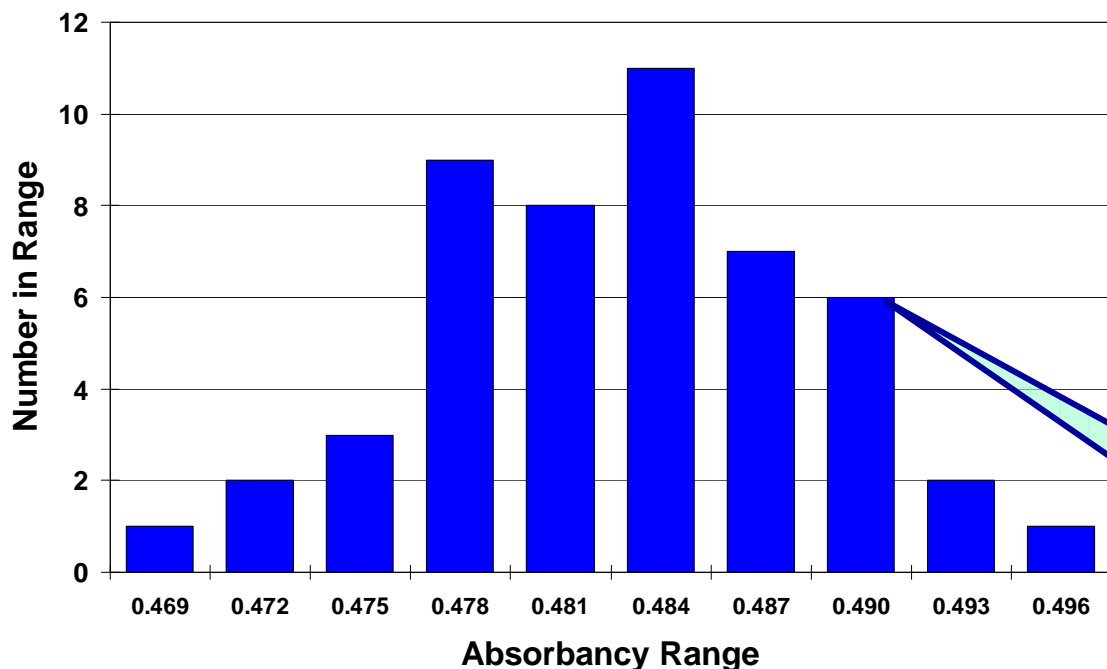


Bin = 0,03

Absorbance Range	Number in Range	Frequency
0.469	1	2.0%
0.472	2	4.0%
0.475	3	6.0%
0.478	9	18.0%
0.481	8	16.0%
0.484	11	22.0%
0.487	7	14.0%
0.490	6	12.0%
0.493	2	4.0%
0.496	1	2.0%
	50	100.0%

Стъпка на хистограмата от-до (dx)

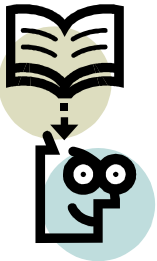
Frequency Distribution



ОТНОСИТЕЛНИ ЧЕСТОТИ на поява

ЧЕСТОТИ на поява

# Плътност на разпределение и функция на разпределение



(cumulative probability distribution function- CDF)  
**ФУНКЦИЯ на РАЗПРЕДЕЛЕНИЕ** – кривата показваща каква е вероятността **P**, случайната величина да е по-малка от дадена стойност

С малко “**x**” бележим елементарния изход

• за дискретна СВ

$$F(X_i) = \sum_k p(x_k); k \leq i$$

• за Непрекъснатата СВ

$$F(X) = \int_{-\infty}^x p(x) dX$$

За непрекъснатата СВ **F(X)** е нарастваща, непрекъснатата и диференцируема функция. Първата и производна е плътността на разпределение

(probability density function- pdf)  
**ПЛЪТНОСТ на РАЗПРЕДЕЛЕНИЕ** **p(X)**

кривата получена от относителните честоти на поява на стойностите на СВ

Ако СВ е реализуема то вероятността тя да се намира в интервала  $(-\infty, +\infty)$  е **1**.

Този интервал е всъщност съдържа **Ω** на СВ:

дискретна

Непрекъснатата

$$P(X_{\min}; X_{\max}) = \sum_i p(x_i) = 1$$

$$P(-\infty; +\infty) = \int_{-\infty}^{+\infty} C \cdot p(x) dX = 1$$

## За непрекъснатата СВ



- Функцията на разпределение на непрекъснатата СВ е *примитивна функция* на плътността на разпределение:

$$F'(X) = p(x)$$

- Функцията на разпределение на непрекъснатата СВ притежава следните свойства

$$F(X) \geq 0 \text{ за всяко } X$$

$F(X)$  е нарастваща функция

$$\lim_{n \rightarrow -\infty} F(X) = 0 \text{ и } \lim_{n \rightarrow +\infty} F(X) = 1$$

- Вероятността непрекъснатата СВ да заема интервалите  $(a, b]$ ,  $[a, b]$  и  $[a, b)$  е една и съща, защото вероятността за заемане на конкретно елементарно значение е  $0$ ,
- Дефинирането на непрекъснатата СВ чрез функция на разпределение или плътност на разпределение е еквивалентно
- Горната граница  $X_a$  за която се изчислява  $F(X_a)$  се нарича *квантил* на разпределението



# Днес разгледахме:

## 1. Събития и действия с тях

- Пространство на елементарните събития

## 2. Класическа вероятност

- Теореме за събиране и умножение на вероятности
- Формула за пълна вероятност. Формула на Бейс

## 3. Случайни величини (СВ)

- прекъснати и непрекъснати СВ,
- относителни честоти на поява
- плътност и функция на разпределение на СВ

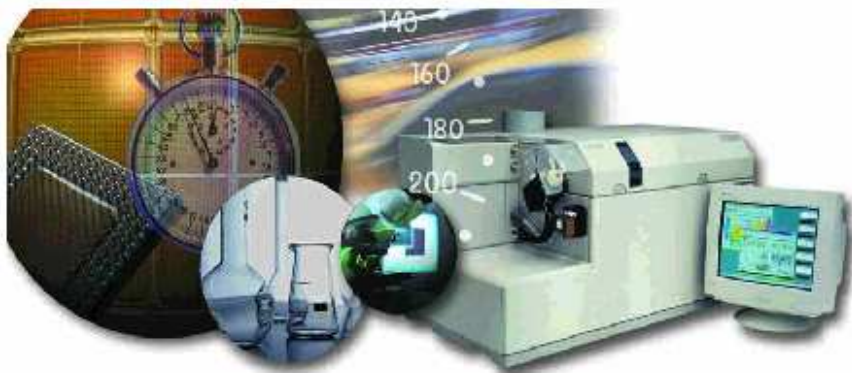
# Задачи за самостоятелна работа:

- Изчислете емпиричните относителни честоти на поява на резултата точки от 3 зарчета и постройте емпиричната хистограма на функцията на разпределение
- Изчислете вероятността да се паднат четни и не четни суми от точки
- Ако на изпит теглите два въпроса от конспект съдържащ 30 въпроса – каква е вероятността да отговорите и на двата въпроса ако сте научили само 20 въпроса от конспекта.
- Разгледайте биномните вероятности по схемата на Бернули

# Лекция №3



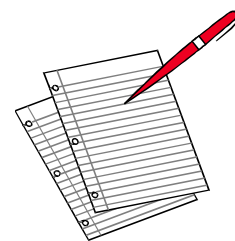
# СЛУЧАЙНА ВЕЛИЧИНА РАЗПРЕДЕЛЕНИЯ



# Днес ще разгледаме:

1. Полигони на случайните величини
2. Числови характеристики на разпределение на СВ:  
мода, медиана, размах
  - **Математическо очакване и дисперсия на СВ.**
  - Свойства на числовите характеристики
3. Разпределения на случайни величини
  - Равномерно (правоъгълно)
  - **Нормално разпределение и стандартно разпределение**
  - Параметри на нормалното разпределение,
  - Интеграл на Лаплас и намирането му от таблици

# Да си припомним



- Какво разбираме под СЛУЧАЙНА ВЕЛИЧИНА?
- Какви видове СВ разгледахме?
- Кога една случайна величина е зададена ?
- Какво разбираме под ВЕРОЯТНОСТ ?
- Как изчисляваме вероятността ?
- Разграничете плътност и функция на разпределение на СВ?
- Как числово да характеризираме една СВ ?

# Характеризиране на СВ

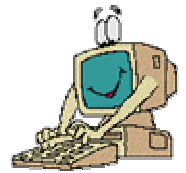
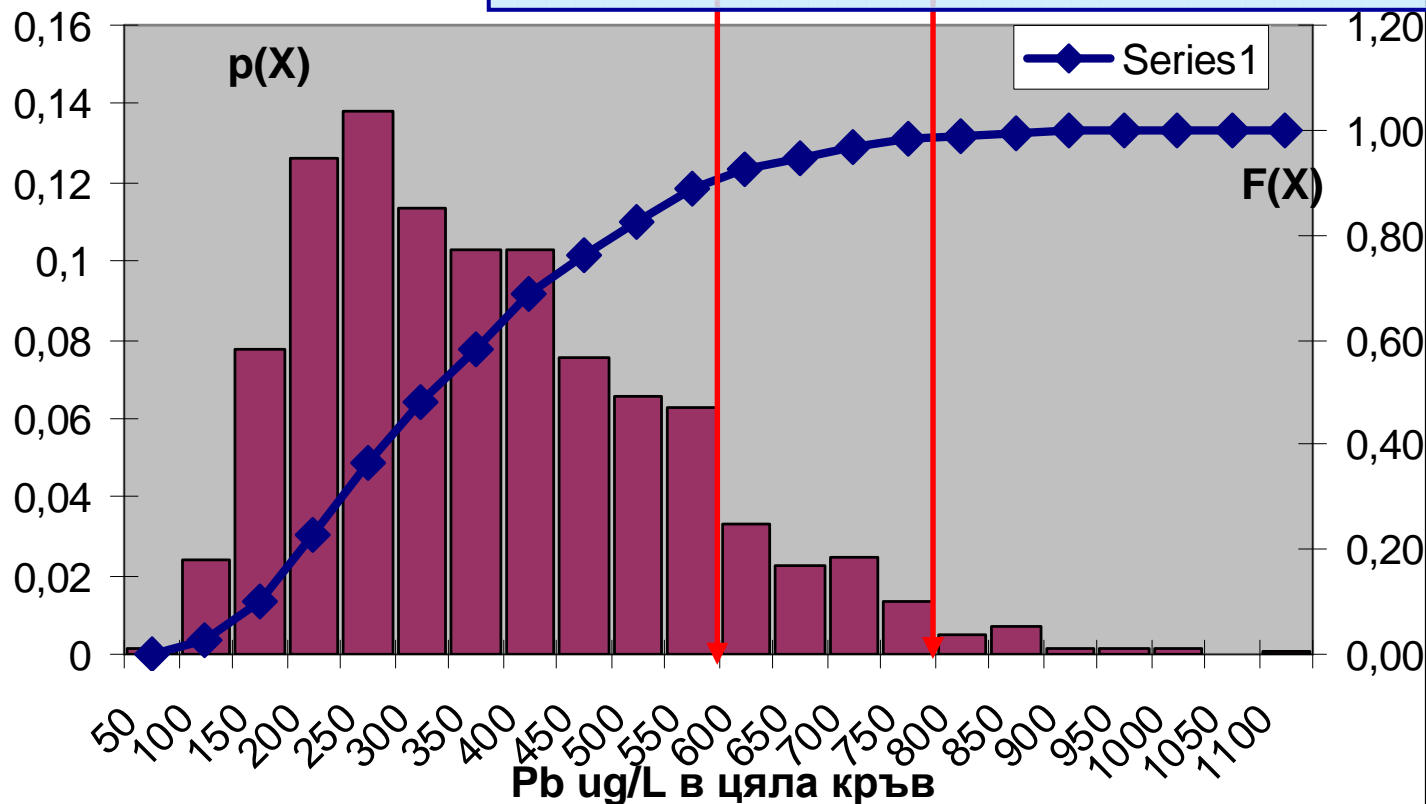
**ЗАДАЧА:** На профилактика с ЕДТА подлежат работници от КЦМ АД със съдържание на олово в кръвта от 600 до 800  $\mu\text{g/L}$ .

Опишете СВ олово в кръв на база изследване на 1763 работника (1993 г.)

Каква е вероятността работник в подобно металургично предприятие да се нуждае от третиране с ЕДТА?

**Решение:**

$$P(600;800) = \int_{600}^{800} p(X) dX = F(800) - F(600) = 0,099$$

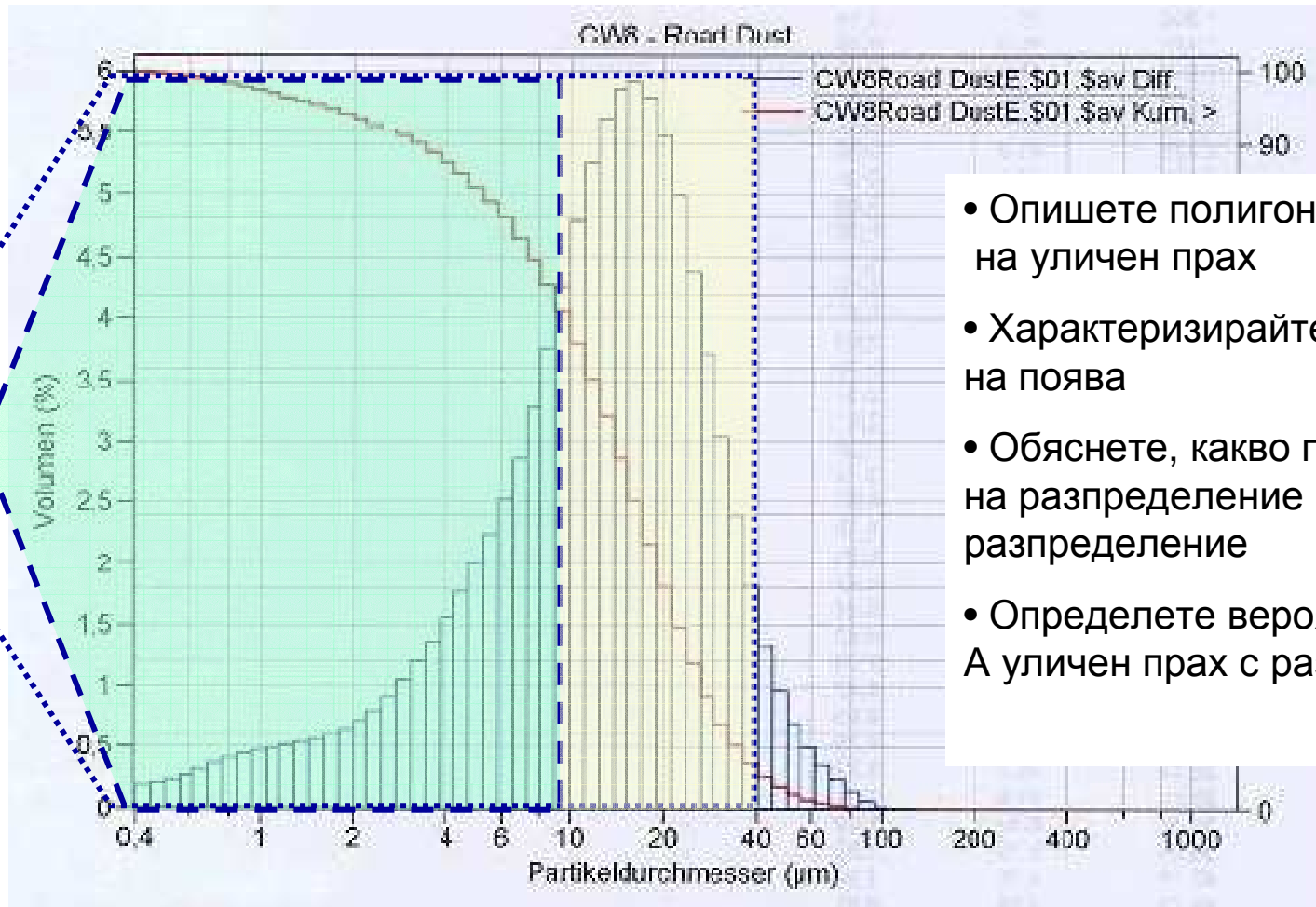


$M(X) =$

$D(X) =$

$Mo(X) =$

# ОЦЕНКА НА СЪБИТИЕ като множество от елементарни изходи

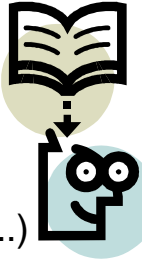


- Опишете полигона на СВ – диаметър на уличен прах
- Характеризирайте относителните честоти на поява
- Обяснете, какво представлява функцията на разпределение и плътността на разпределение
- Определете вероятността на събитието А уличен прах с размер от 10 до 40 µm

$$P(10 \div 40) = P(-\infty \div 40) - P(-\infty \div 10)$$

$$P(10 \div 40) = F(40) - F(10)$$

# ЧИСЛОВИ ХАРАКТЕРИСТИКИ НА СЛУЧАЙНИТЕ ВЕЛИЧИНИ



- **Мода** ( $M_o$ ) – най-срещаното значение СВ, това което тя заема с най-голяма относителна честота  $k_x$  и  $p(X) \max$ ;



Excel.Ink

**MODE**(number1,number2,...)

- **Медиана** ( $M_d$ ) – значението, което разделя подреденото по големини множеството от значения на две равни части



Excel.Ink

**MEDIAN**(number1,number2,...)

- **Математическо очакване – mathematical estimation** ( $M$ ) или ( $E$ )

за **дискретна СВ** - сумата от произведенията на значенията и съответните им вероятности

за **непрекъснатата СВ** – пълния интеграл от произведенията на значенията и съответните им вероятности

$$M(X) = \sum_{k=1}^{\infty} p(X_k) \cdot X_k$$

$$M(X) = \int_{-\infty}^{+\infty} X \cdot p(X) dX$$

$$D(X) = M[X - M(X)]^2$$

- **Дисперсия на СВ – variance** ( $D$ ) или ( $V$ )

Математическото очакване на квадратите на разликите между значенията на СВ и нейното математическо очакване

за **дискретна СВ** – сума .....

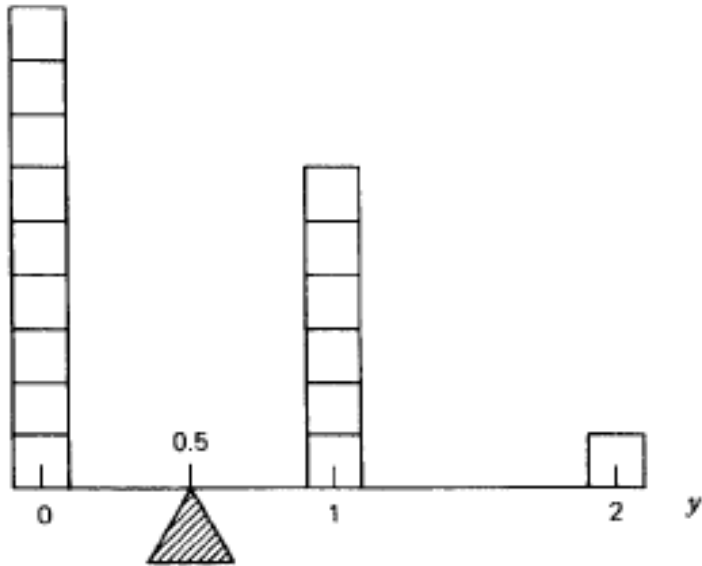
за **непрекъснатата СВ** – пълния интеграл от ...

$$D(X) = \sum_{k=1}^{\infty} p(X_k) \cdot [X_k - M(X)]^2$$

$$D(X) = \int_{-\infty}^{+\infty} p(X) \cdot [X - M(X)]^2 dX$$

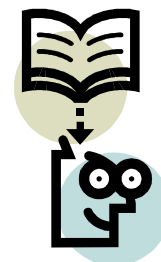


# Математическо очакване



Изчислете  
МАТЕМАТИЧЕСКОТО  
ОЧАКВАНЕ,  
МЕДИАНАТА и  
МОДАТА ?

# ХАРАКТЕРИСТИКИ на математическото очакване и на дисперсията



- Физическия смисъл на Математическото очакване е, че то представлява центъра на тежестта на множеството от значения които СВ заема
- Математическото очакване може да е произволно число. То дори може да не е елемент на множеството от значения на случайната величина
- Модата, медианата, математическото очакване са числови характеристики на СВ и имат размерността (мерната единица) на СВ
- Дисперсията винаги е положително число и размерността е квадрата на мерната единица на СВ
- Физическия смисъл на дисперсията е, че тя се явява мярка за разсейването на стойностите които СВ заема около математическото очакване.
- По-удобно е използването корен втори от дисперсията, чийто физичен смисъл е – центъра на тежестта на полу-интервала на разбягване на стойностите са случайната величина около математическото очакване (размерността е на СВ) :

$$SD = \sqrt{D(X)}$$

# СВОЙСТВА на матем. очакване и на дисперсията



1. Математическото очакване на неслучайна величина -константа е равно на константата

$$M(C) = C$$

2. Неслучайната величина константа, може да се изнася от знака на матем. очакване

$$M(CX) = C.M(X)$$

3. Матем. очакване на сума от случайни величини е равно на сумата от матем. им очаквания:

$$M(X+Y+..Z) = M(X) + M(Y) + .. M(Z)$$

4. Матем. очакване на произведение на **НЕЗАВИСИМИ** случайни величини е равно на произведението на матем. им очаквания:

$$M(X.Y...Z) = M(X).M(Y)...M(Z)$$

5. Дисперсията на неслучайна величина -константа е равна на **НУЛА**

$$D(C) = 0$$

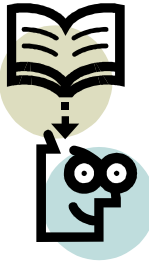
6. Неслучайната величина константа, може да се изнася извън знака на дисперсията, като при това **изнесената константа се повдига на квадрат**

$$D(C.X) = C^2 . D(X)$$

7. Дисперсията на сума от **НЕЗАВИСИМИ** случайни величини е равно на сумата от техните дисперсии

$$D(X+Y+..Z) = D(X) + D(Y) + .. D(Z)$$

# СВОЙСТВА на матем. очакване и на дисперсията



- На база свойства от 1-7, всяка една случайна величина може да се нормира по формулата:

$$Z = \frac{[X - M(X)]}{\sqrt{D(X)}}$$

В учебника “Теория на Експеримента”  
на Футеков и Пенчев  
вместо Z е записано Y

Това се нарича – ЗЕТ ТРАНСФОРМАЦИЯ

След нея получената Z величина е **стандартна**  
със следните характеристики:

$$M(Z) = 0 \quad \text{и} \quad D(Z) = 1;$$

# Днес ще разгледаме:

1. Полигони на случайните величини
2. Числови характеристики на разпределение на СВ:  
мода, медиана, размах
  - **Математическо очакване и дисперсия на СВ.**
  - Свойства на числовите характеристики
3. Разпределения на случайни величини
  - **Равномерно (правоъгълно)**
  - **Нормално разпределение и стандартно разпределение**
  - **Параметри на нормалното разпределение,**
  - **Интеграл на Лаплас и намирането му от таблици**

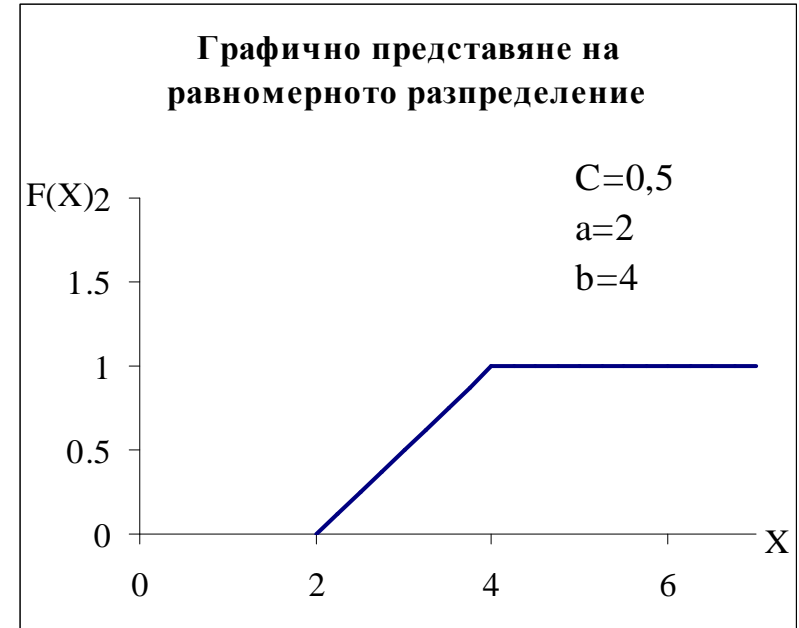
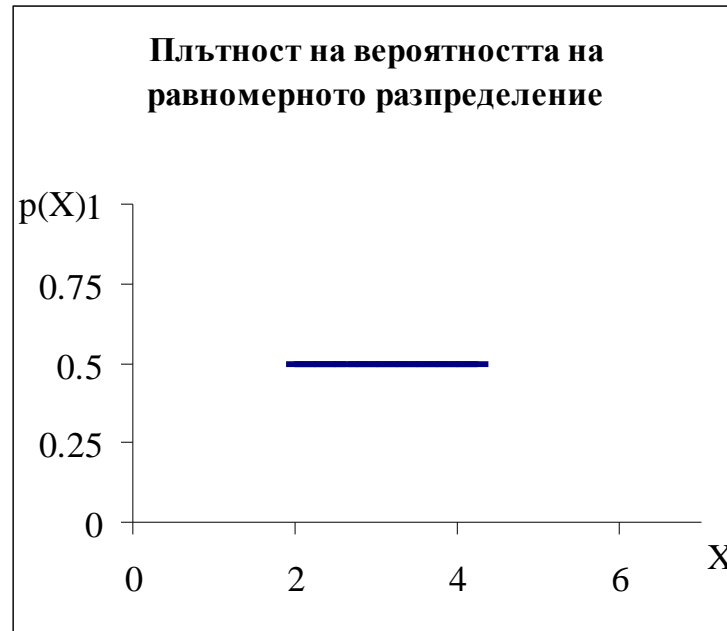
# Равномерно (правоъгълно) разпределение

$$p(X) = \frac{1}{b-a}$$

$$M(X) = \frac{b+a}{2}$$

$$D(X) = \frac{(b-a)^2}{12}$$

$$u_s = S = \frac{b-a}{2\sqrt{3}}$$



## Особености:

- Всички елементарни значения са еднакво вероятни -  $k\omega_i = \text{const}$ ,  $p(\omega_i) = \text{const}$
- Ограничено е в ясни граници на  $\Omega$

Плътност

$$p(X) = \begin{cases} C : a \leq X \leq b \\ 0 : X < a, X > b \end{cases}$$

Функция на разпределение

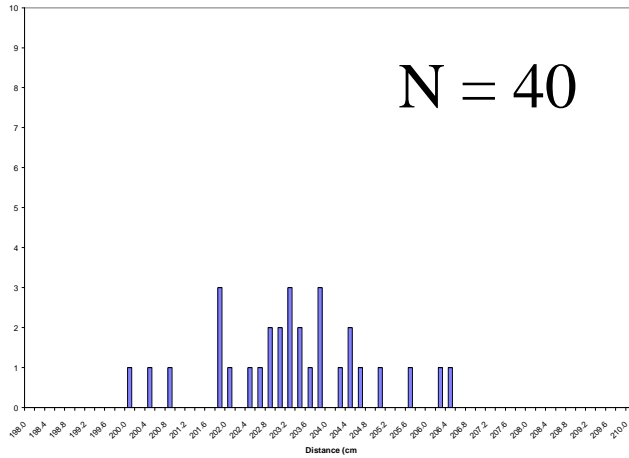
$$\int_{-\infty}^{+\infty} p(X) dx = \int_{-\infty}^a 0 dx + \int_a^b C dx + \int_b^{+\infty} 0 dx = \int_a^b C dx$$

# НОРМАЛНО РАЗПРЕДЕЛЕНИЕ

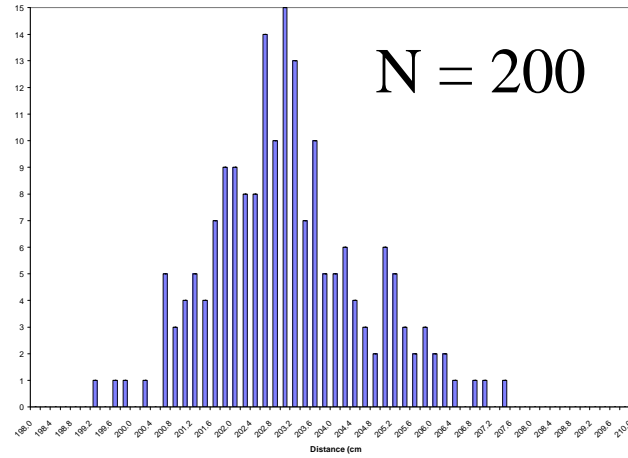
Дистанция на куршум изстрелян от детски пистолет,

$N$  – брой изстрели

Distribution of distance measurements for proj

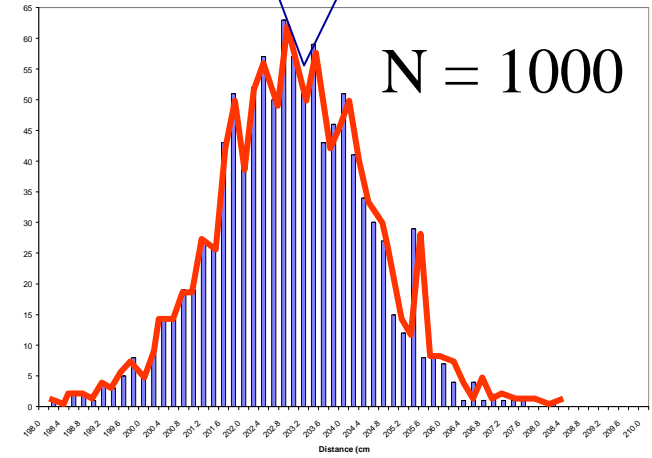


Distribution of distance measurements for proj

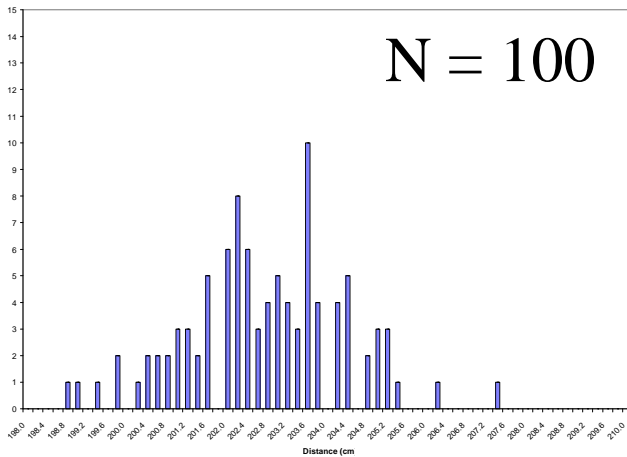


Полигон на СВ

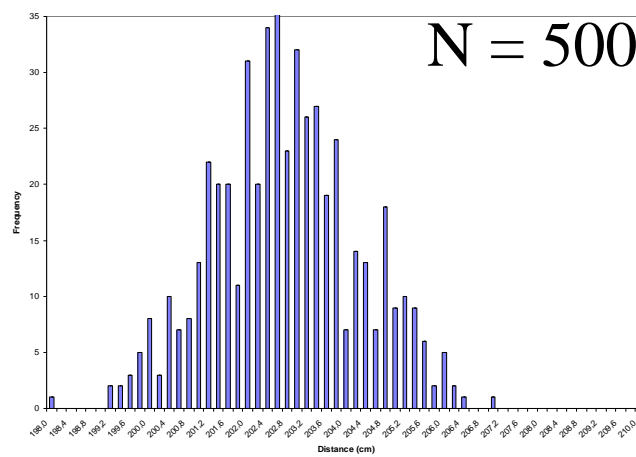
Distribution of distance measurements for proj



Distribution of distance measurements for proj

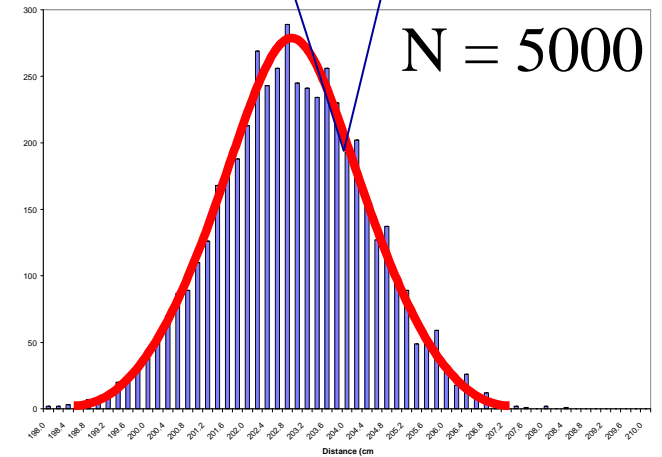


Distribution of distance measurements for projectile

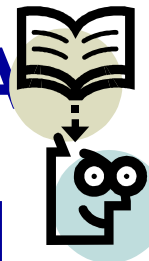


Гаусова камбана

Distribution of distance measurements for proj



# ЦЕНТРАЛНА ГРАНИЧНА ТЕОРЕМА



сумарната **СЛУЧАЙНА ВЕЛИЧИНА** на голям брой произволно разпределени, случайни величини, които са **независими** и стойностите им са малки в сравнение с тази на сумата - **КЛОНИ КЪМ НОРМАЛНО РАЗПРЕДЕЛЕНИЕ**.

ТОВА Е ТЕОРЕТИЧЕН МОДЕЛ,  
към който отнасяме поведението на  
повечето реално измервани СВ



ДЖОНА САН ПИЕТРО СКАЛЪ - 1999 - 1998

$$F(X) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}} dx$$

$$y = p(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2}$$

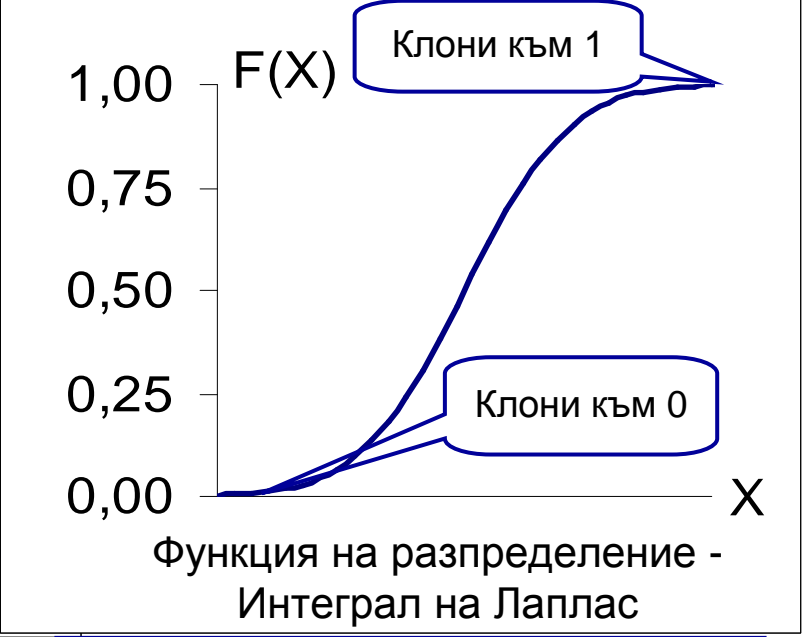
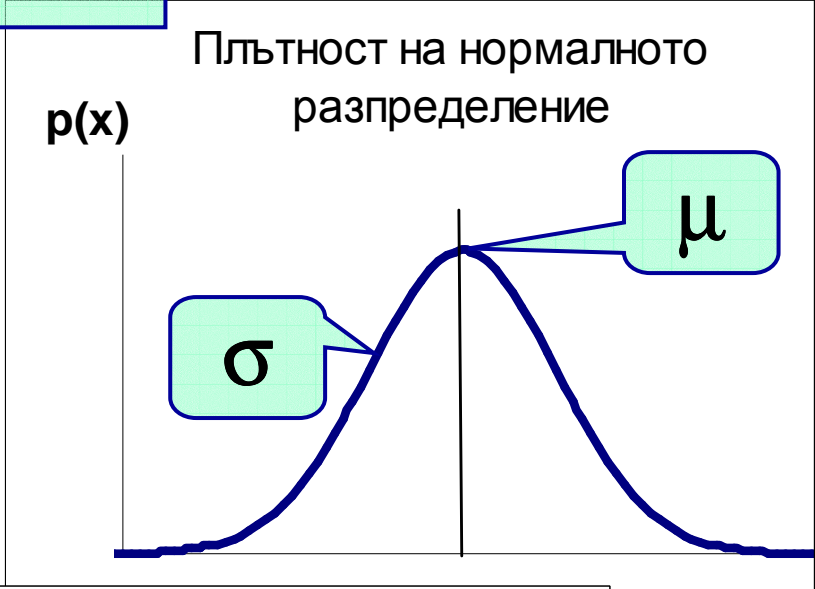


# Нормално (Гаусово) разпределение $N(\mu; \sigma^2)$

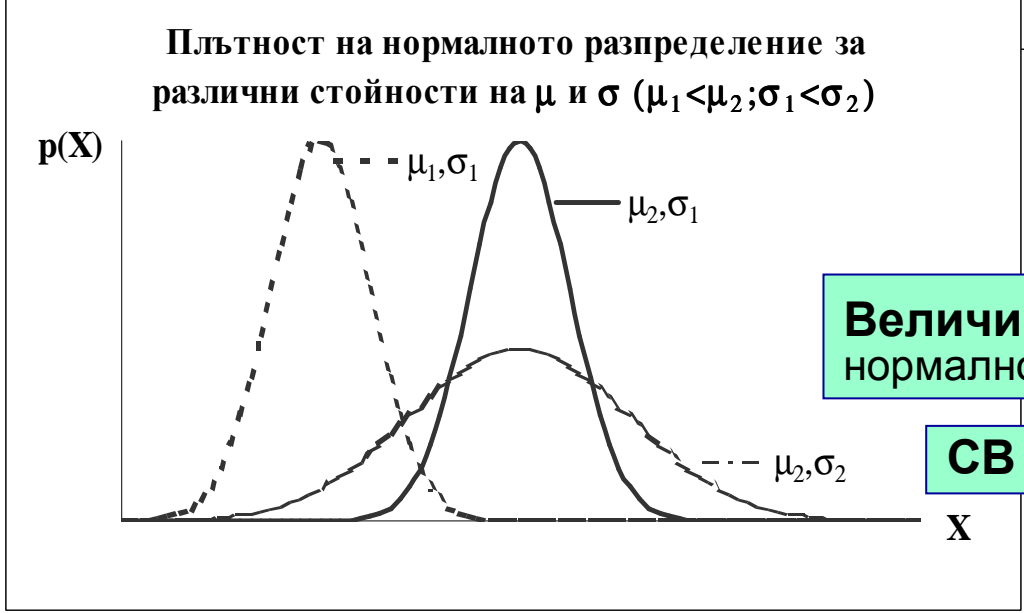
$$p(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}}$$

$$M(X) = \mu$$

$$D(X) = \sigma^2$$



$$F(X) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}} dx$$



Величините  $\mu$  и  $\sigma$  се наричат **ПАРАМЕТРИ** на нормалното разпределение и напълно го характеризират

**СВ** заема стойности в отворен интервал

# СТАНДАРТНО разпределение $N(\mu=0;\sigma=1)$

Нормално разпределение, което е нормирано спрямо  $\sigma$  посредством подмяна на променливата  $X$  със зет трансформация  $Z = (X - \mu)/\sigma$

Параметрите на стандартното разпределение са  $\mu = 0$  и  $\sigma = 1$

Стойности на интеграла на Лаплас за стандартното разпределение са дадени в ПРИЛОЖЕНИЕ 1 или в EXCEL



Excel.Ink

**NORMDIST(x,mean,standard\_dev,cumulative)**



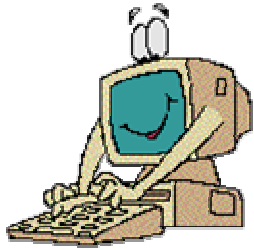
Как може да преведем всяко едно нормално разпределение към стандартно ?

Може ли да определяме вероятността за дадено събитие на произволна нормално разпределена величина посредством Приложение 1?

Подзадача - стойности по-малки от гранична стойност

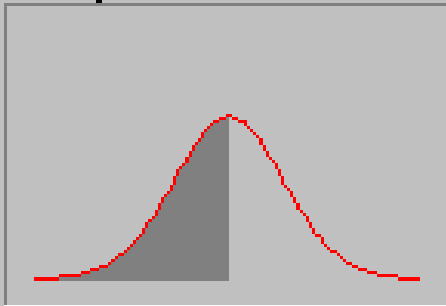
Подзадача - стойности във симетричен интервал около  $\mu$

# СТАНДАРТНО разпределение $N(\mu=0; \sigma=1)$

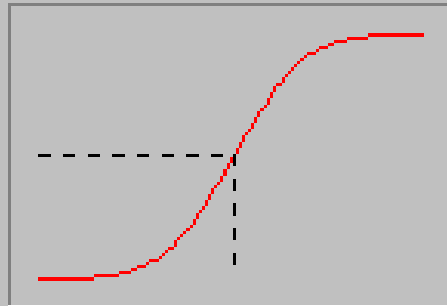


$$F(X) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}} dx$$

Density Function:



Distribution Function:

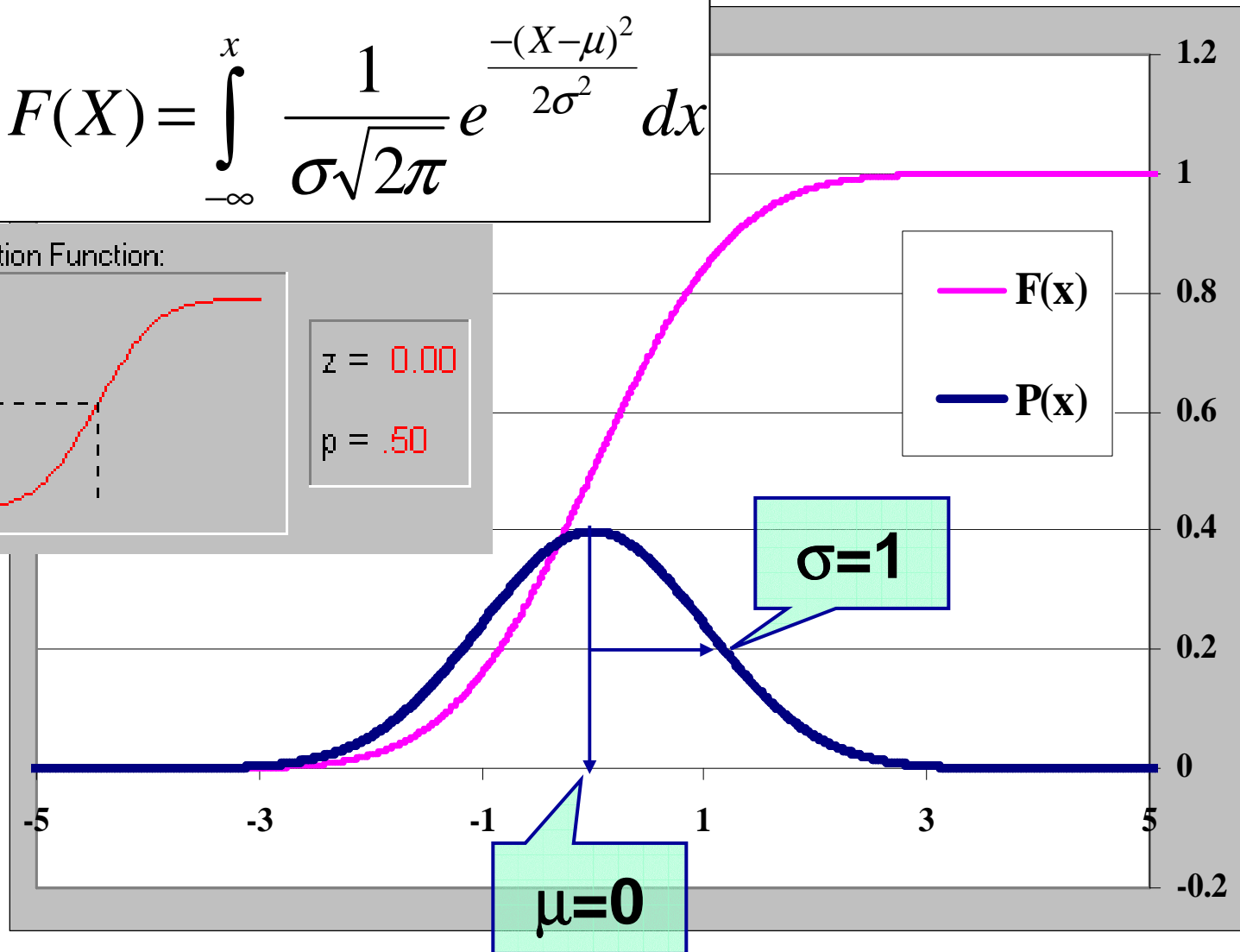


$z = 0.00$

$p = .50$

$$M(X) = \mu = 0$$

$$D(X) = \sigma^2 = 1$$

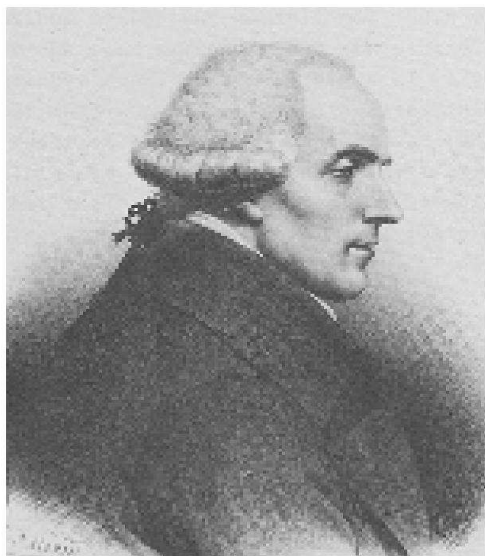


# Приложение №1 Интеграл на Лаплас (в граници $\sigma$ )

КВАНТИЛИ  
интегриране  $z$   
 $F(-\infty; z)$

	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8829
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

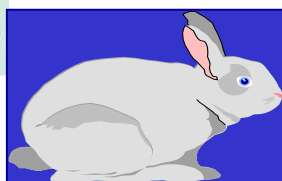
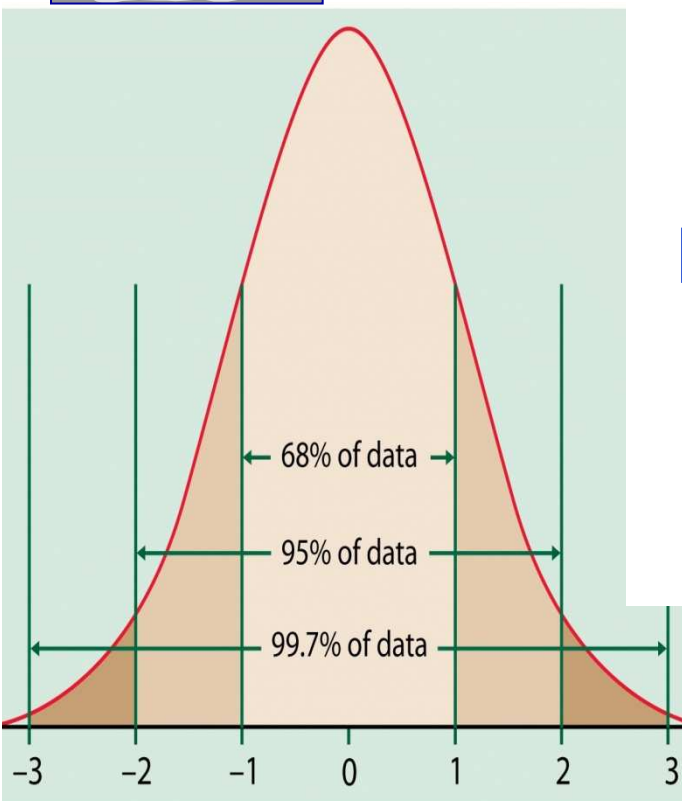
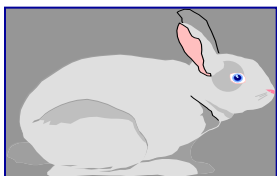
Площ  
 $F(-\infty; z)$



# Днес разгледахме:

1. Полигони на случайните величини
2. Числови характеристики на разпределение на СВ:  
мода, медиана, размах
  - **Математическо очакване и дисперсия на СВ.**
  - Свойства на числовите характеристики
3. Разпределения на случайни величини
  - Равномерно (правоъгълно)
  - **Нормално разпределение и стандартно разпределение**
  - Параметри на нормалното разпределение,
  - Интеграл на Лаплас и намирането му от таблици

# Z-трансформация на нормално разпределени величини за превеждане към стандартно разпределение



$$Z = \frac{X - \mu}{\sigma}$$

# Днес ще разгледаме:

1. Статистическа сигурност и ниво на значимост при оценяване на стойностите на случайна величина.
2. Важни връзки при изчисляване на вероятността стандартно разпределена СВ да заема стойности в зададен интервал.
3. Правило  $3\sigma$  - за нормално разпределена величина.
4. Изчисляване вероятността за настъпване на събитие на нормално разпределена величина
  - Преминаване към стандартно разпределение
  - Z трансформация
  - Едностранно и двустранно ограничени интервали

# КОНТРОЛНИ ВЪПРОСИ



Дефинирайте понятието разпределение на СВ.

Характеризирайте равномерното разпределение

Как се описва нормално разпределена СВ ?

Какво представлява стандартно разпределената СВ?

Как се работи с таблици представящи квантилите на стойностите на интеграла на Лаплас ?



Excel.Ink

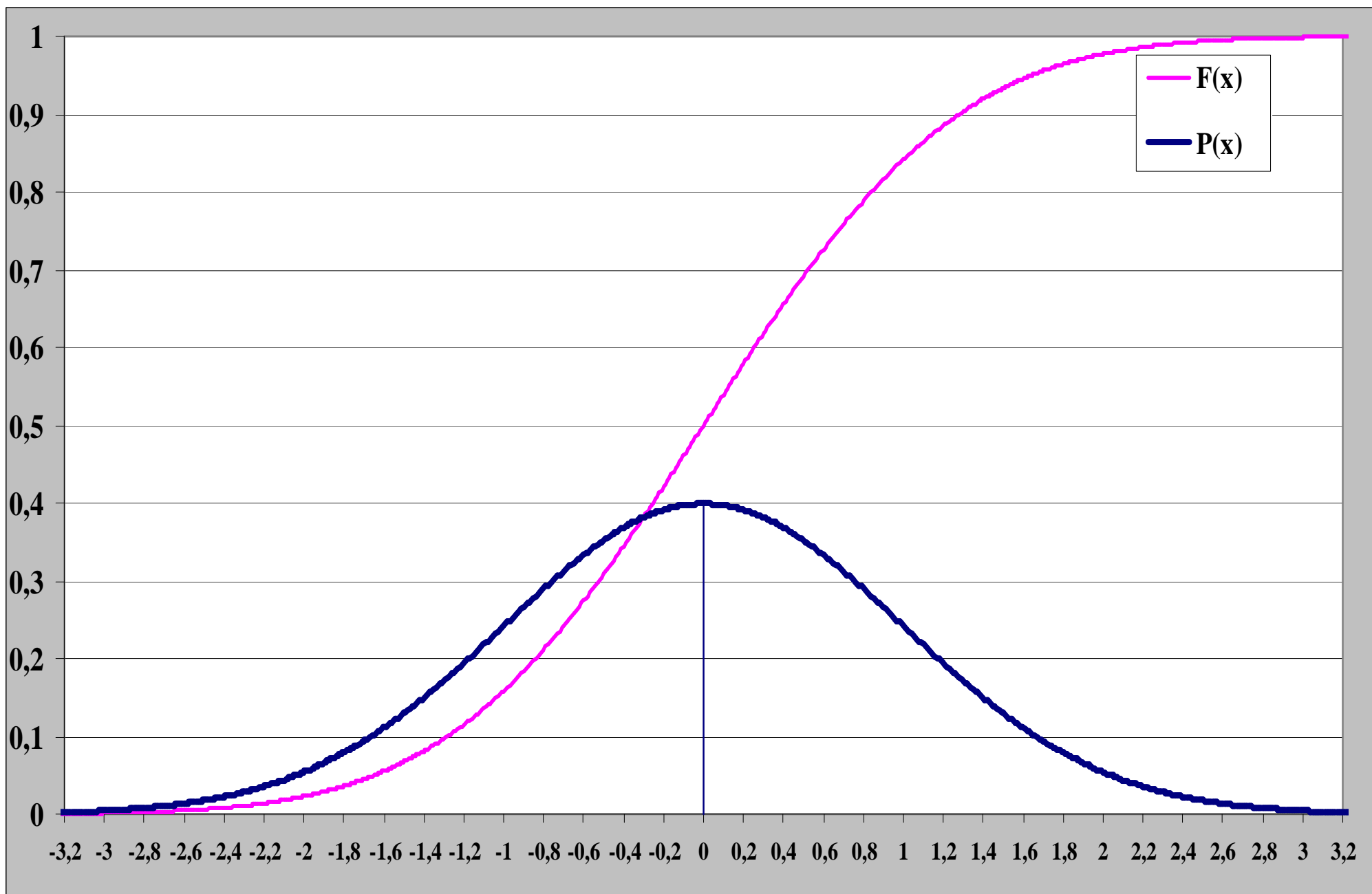
**NORMDIST(x,mean,standard\_dev,cumulative)**



## ПРИЛОЖЕНИЕ 1

Стойности на интеграла на Лаплас в граници от $-\infty$ до X										
X	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,00	0,50000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,10	0,53983	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,20	0,57926	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,30	0,61791	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,40	0,65542	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,50	0,69146	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,60	0,72575	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,70	0,75804	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,80	0,78815	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,90	0,81594	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,00	0,84134	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,10	0,86433	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,20	0,88493	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,30	0,90320	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,40	0,91924	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,50	0,93319	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441

1,60	0,94520	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,70	0,95543	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,80	0,96407	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,90	0,97128	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,00	0,97725	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,10	0,98213	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,20	0,98610	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,30	0,98928	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,40	0,99180	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,50	0,99379	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,60	0,99534	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,70	0,99653	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,80	0,99744	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,90	0,99813	0,9982	0,9983	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,00	0,99865	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990



# Важни връзки !

P – статистическа сигурност

$$P = 1 - \alpha$$

$\alpha$  – ниво на значимост

a – квантил

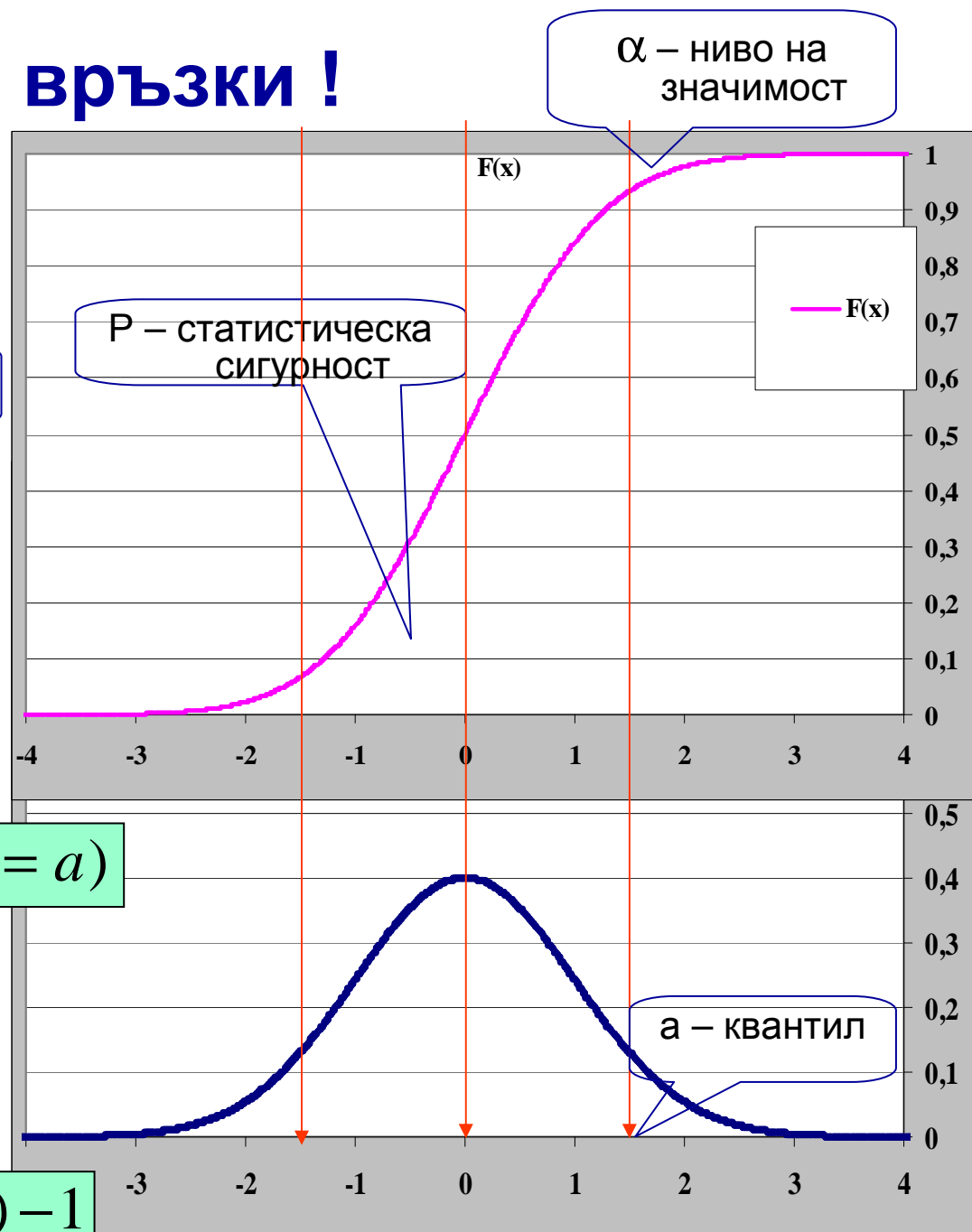
$$P(X \leq a) = F(X = a)$$

$$P(X > a) = 1 - F(X = a)$$

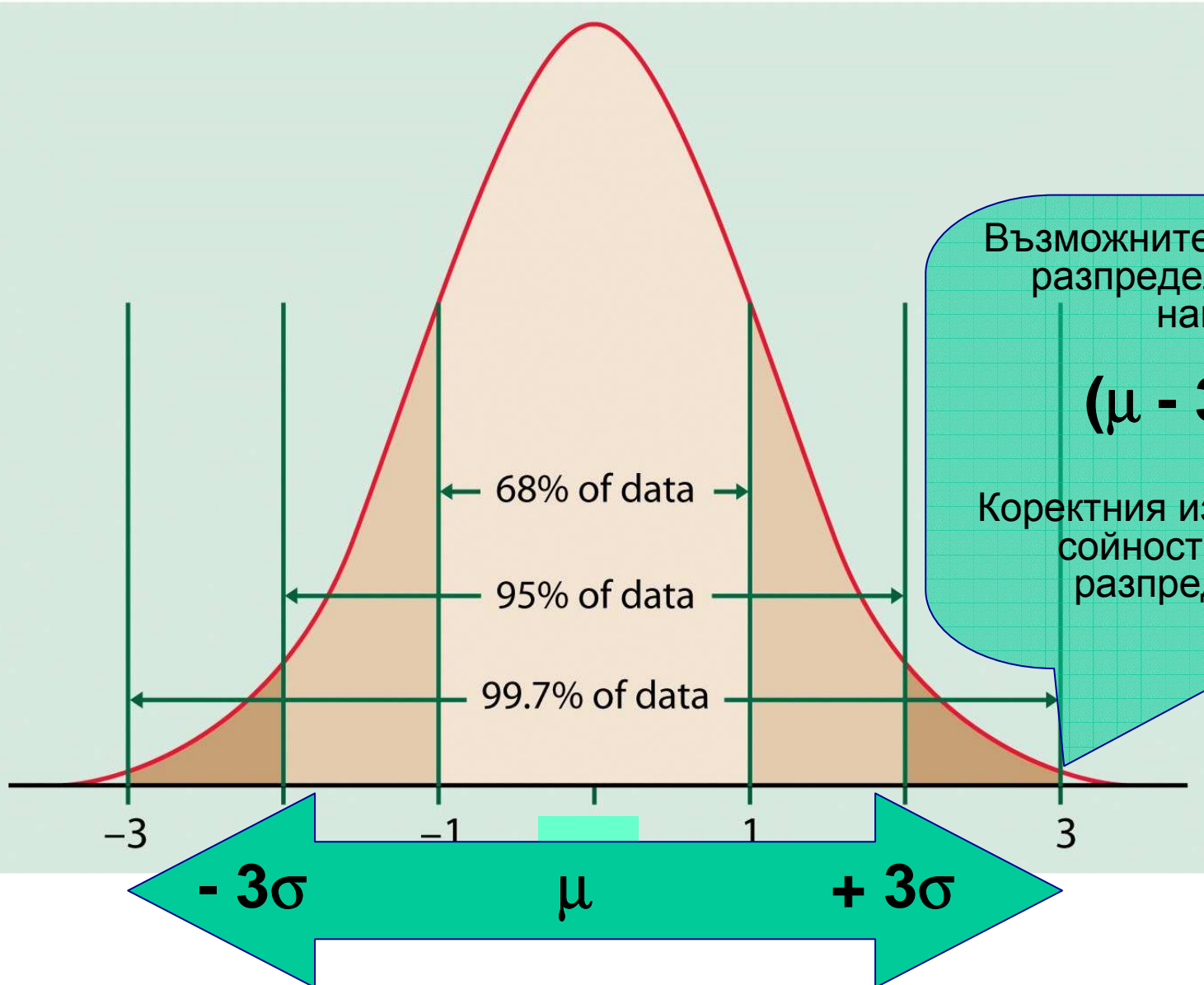
$$P(a < X < b) = F(X = b) - F(X = a)$$

$$P(-\infty; -a) = 1 - F(X = a)$$

$$P(\mu - a < X < \mu + a) = 2F(X = a) - 1$$



# ПРАВИЛО ТРИТЕ СИГМА



Възможните стойности на нормално разпределена СВ практически не напускат интервала

$$(\mu - 3\sigma < x < \mu + 3\sigma)$$

Коректния израз е 99,73% от всички стойности на **ВСЯКА** нормално разпределена величина са в този интервал

# СТАНДАРТНО разпределение $N(\mu=0;\sigma=1)$

Нормално разпределение което е нормирано спрямо  $\sigma$  посредством подмяна на променливата  $X$  със зет трансформация  $Z = (X - \mu)/\sigma$

Параметрите на стандартното разпределение са  $\mu = 0$  и  $\sigma = 1$

Стойности на интеграла на Лаплас за стандартното разпределение са дадени в ТЕ ПРИЛОЖЕНИЕ 1 на или в EXCEL



Excel.Ink

**NORMDIST(x,mean,standard\_dev,cumulative)**



Как може да преведем всяко едно нормално разпределение към стандартно ?

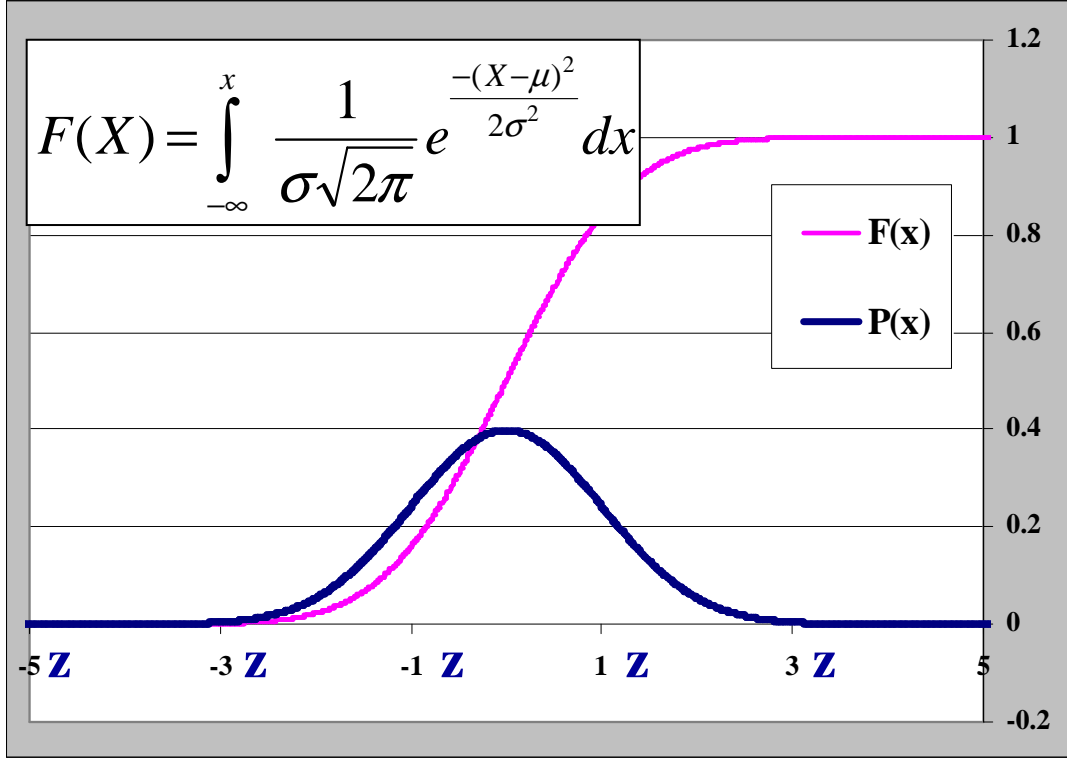
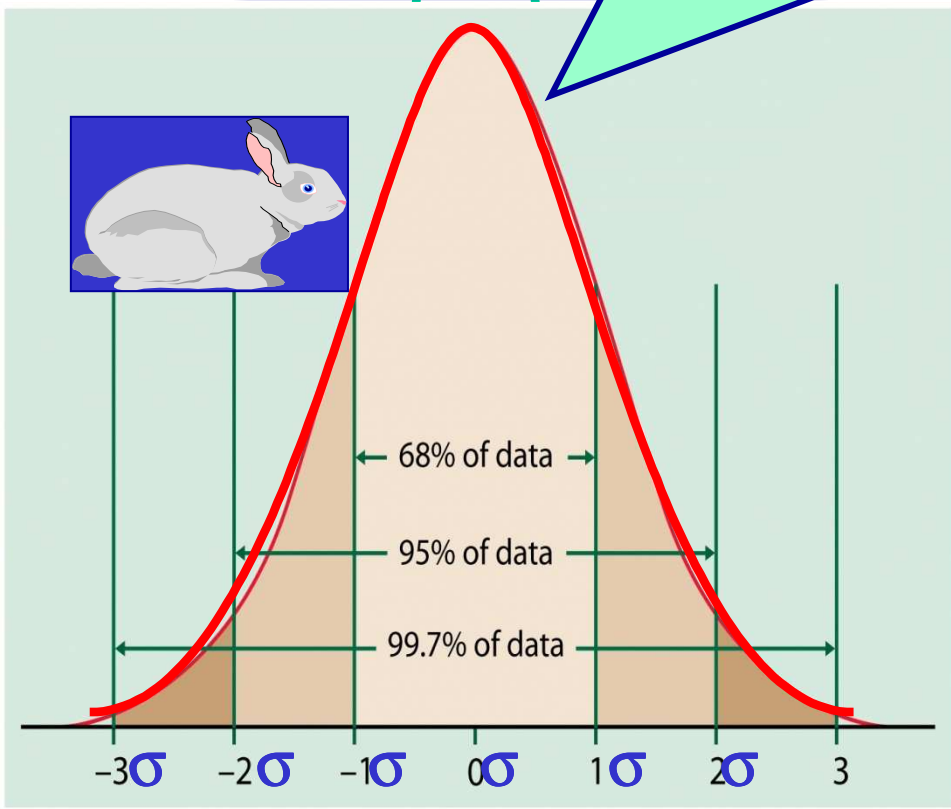
Може ли да определяме вероятността за дадено събитие на произволна нормално разпределена величина посредством Приложение 1?

Подзадача - стойности по-малки от гранична стойност

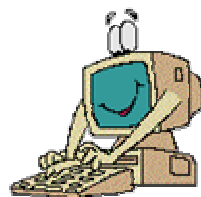
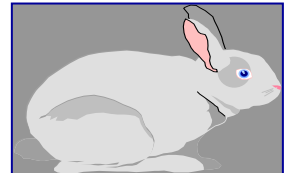
Подзадача - стойности във симетричен интервал около  $\mu$

# Z - трансформация на нормално разпределена величина до стандартно разпределена - $N(\mu=0; \sigma_2= 1)$

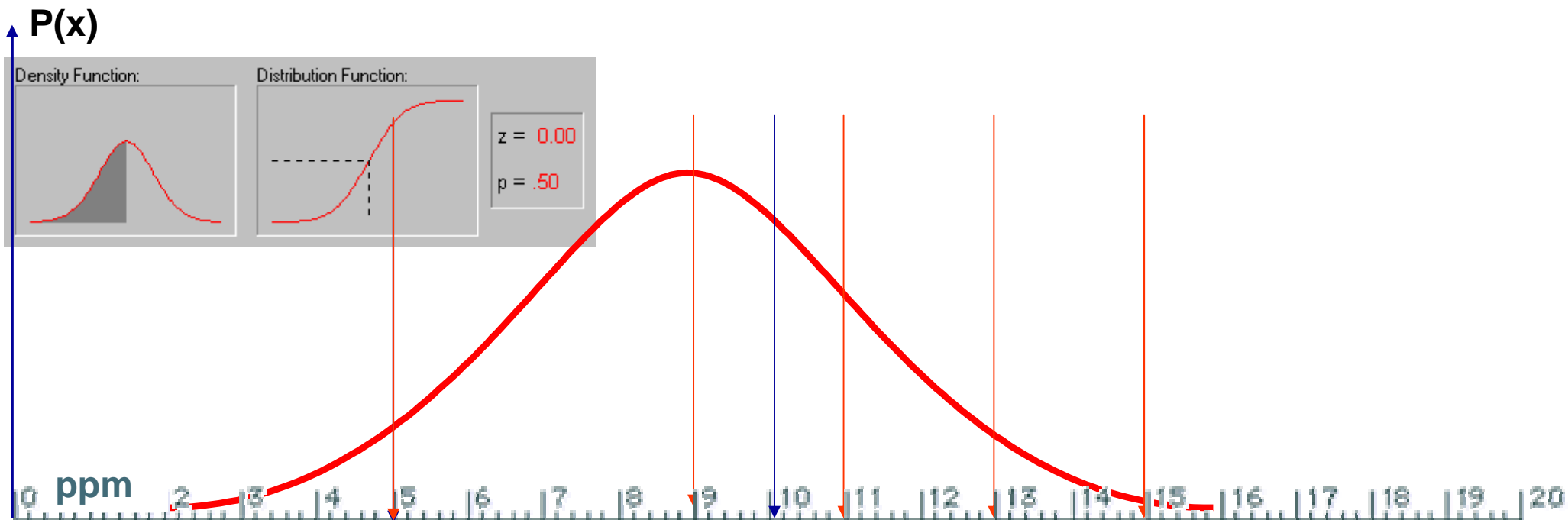
Всяка нормално разпределена СВ заема камбана с еднотипен полигон спрямо  $\mu$  и  $\sigma$



$$Z = \frac{X - \mu}{\sigma}$$



# Съдържание на Au в полиметална руда



## Задача:

Съдържание на злато в руда ppm

Съдържанието на Au (ppm) в полиметална руда от Родопите е случайна нормално разпределена величина  $N(9;2^2)$

## Изчислете:

Симетричния интервал в който попадат 95% от значенията на СВ

Симетричния интервал в който попадат 99% от значенията на СВ

Каква е вероятността рудата да съдържа злато над 10 mg/g ?

Каква е вероятността златото да е по-малко от 5 mg/g ?

# Задачи за самостоятелна работа:

Разгледайте в EXCEL разпределенията

- биномно разпределение



**BINOMDIST(number\_s, trials, probability\_s, cumulative)**

Excel.Ink

- Хипергеометрично разпределение



**HYPGEOMDIST(sample\_s, number\_sample, population\_s, number\_population)**

Excel.Ink

- Поасоново разпределение



**POISSON(x, mean, cumulative)**

Excel.Ink



# Днес разгледахме:

1. Статистическа сигурност и ниво на значимост при оценяване на значенията на случайна величина.
2. Важни връзки при изчисляване на вероятността стандартно разпределена СВ да заема стойности в зададен интервал.
3. Правило  $3\sigma$  - за нормално разпределена величина.
4. Изчисляване вероятността за настъпване на събитие на нормално разпределена величина
  - Преминаване към стандартно разпределение
  - Z трансформация
  - Едностранно и двустранно ограничени интервали

## Лекции на Prof. Dr. Antonio Canals Hernandez

March 11, 2008 at 15:23 · Filed under Събития



По програмата за академичен обмен Еразъм  
в Химически факултет на Пловдивски университет "Паисий  
Хилендарски"

гостува

**Prof. Dr. Antonio Canals Hernandez**  
**University of Alicante**  
**SPAIN**

който ще изнесе интензивен курс  
***Univariate Calibration on Instrumental Analysis***

Лекциите са от 9:00 до 12:30 и от 14:00 до 17:00 на 20 март 2008  
год. (четвъртък)  
в 4 аудитория, Ректорат.

Поканват се студенти, преподаватели и гости!

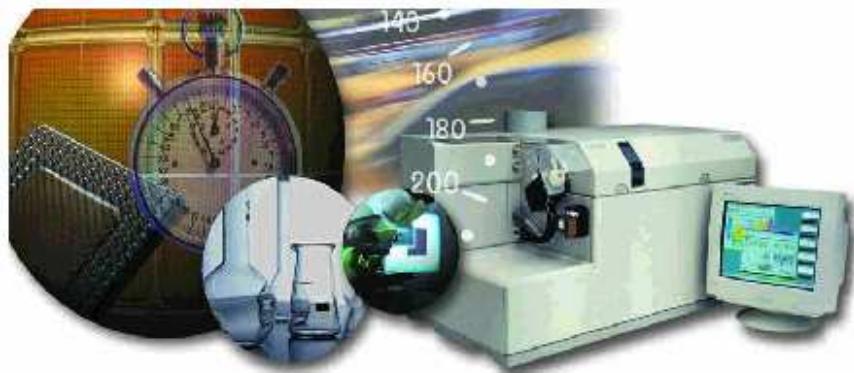
**English version**  
За факултета >

- 3. Състав
- 4. Обучение
  - А: Бакалавър
  - Б: Магистър
  - В: Доктор
  - Г: СДК
- 5. Документи
- 6. Факултетен съвет
- 7. Научна конференция
- 8. Научни трудове по химия
- 1. История
- 2. Катедри

Категории >

- За страницата
- Мрежата
- Проекти

# СТАТИСТИЧЕСКИ ОЦЕНКИ НА ИЗВАДКИ С ОГРАНИЧЕН ОБЕМ



# Днес ще разгледаме:

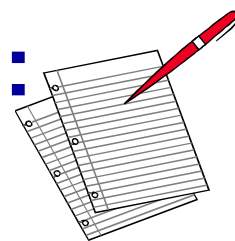
## 1. Статистически оценки.

- Средно-аритметична стойност и средно квадратично отклонение, относително стандартно отклонение
- Начини за изчисляване на статистическите оценки от извадки с ограничен обем
- Неизместеност и състоятелност на статистическите оценки.

## 2. Разпределения – характеризиращи статистическите оценки

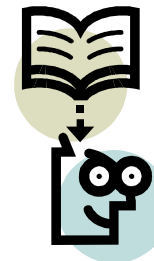
- t-разпределение;
- $\chi^2$ -разпределение;
- F-разпределение

## Да си припомним:



- Кои са числовите характеристики на една случайна величина?
- Кое разпределение наричаме равномерно ?
- Какво е характерно за нормалното разпределение ?
- Какво е стандартно разпределение ?
- Как изчисляваме вероятността стандартно разпределена величина да заема стойности в определени интервали ?
- Как изчисляваме вероятността на нормално разпределена СВ да заема стойности в определени интервали?
- Какво е статистическа сигурност и какво ниво на значимост за интервалите на значенията на една СВ ?

# Статистическа извадка (1)



## Генерална съвкупност (general population)

множество от всички елементи на СВ

аналогът е  $\Omega$  пространството -  
пълният обем на възможните  
елементарни изходи  
това е извадка с неограничен обем

**извадка (sample)** – подмножество  
на генералната съвкупност

Извадката е набор от  
експериментални стойности и  
позволява чрез ограничен брой  
измервания да бъдат направени  
изводи относно поведението на ГС

Извадките удовлетворяват  
следните три условия:

1. дефинирани са върху същото пространство на което е определена и ГС
2. имат същото разпределение като ГС
3. съставните им са НЕЗАВИСИМИ

**Представителна (статистическа) извадка** - такава извадка,  
чиито параметри представят параметрите на генералната съвкупност

**ПРЕДСТАВИТЕЛНА ПРОБА**

**Сборна проба** -> Лабораторна проба -> Тестова проба - реплика

При вземането на проби и подготовката на лабораторни проби инспекторите вземат необходимите предпазни мерки с цел недопускане на промени в пробите, които могат да окажат влияние върху съдържанието на замърсителите, върху аналитичното определяне или представителността на общата проба.

Чл. 17. (1) **Точкови проби** се вземат от различни места, разпределени в цялата партида или подпартида. Всяко отклонение от това правило се посочва в протокола за вземане на проби.

(2) Когато е практически невъзможно да се изпълни изискването по ал. 1, пробите се вземат от различни места от достъпната част на партидата, което обстоятелство се отбелязва в протокола за вземане на проби.

Чл. 18. **Общата проба** се формира чрез обединяване и смесване на точковите проби. Общата проба се хомогенизира в лабораторията.

Чл. 19. **Двойни проби** за целите на държавния контрол и контролни експертизи или за анализ по молба на физически и юридически лица се вземат по реда на Наредба № 22 за условията и реда за вземане на проби от храни. Те се образуват от хомогенизираната обща проба, когато това е практически възможно.

Чл. 20. (1) **След смесване общата проба се разделя на равни лабораторни проби.**

(2) Количеството на всяка лабораторна проба трябва да бъде достатъчно за извършването на най-малко два анализа.

Чл. 21. (1) Всяка проба се поставя в чиста опаковка, която да осигурява защита от замърсяване и повреди при транспортиране, както и от загуби на анализирания материал, вследствие на абсорбцията от вътрешните стени на опаковката. По време на транспортиране или съхранение се вземат предпазни мерки за недопускане на промени в състава на пробата.

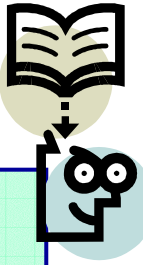
(2) Опаковката по ал. 1 трябва да бъде изработена от инертни в химично отношение, с антикорозионни свойства материали, на основата на които са разработени специални средства за дезинфекция и други химични вещества, използвани за обработка на храните.

Чл. 22. Опаковането, затварянето и маркирането на пробите се извършват по изискванията на Наредба № 22 за условията и реда за вземане на проби от храни.

Чл. 23. (1) Държавните инспектори съставят протокол за взетите проби храни по реда на чл. 4, ал. 2 от Наредба № 22 за условията и реда за вземане на проби от храни.

ЦЕЛТА Е след като се проведе изпитване на характеристиките на **ПРОБАТА(те)** да се **ПРЕПИШАТ** тези характеристики на **ГЕНЕРАЛНАТА СЪВОКУПНОСТ - ОБЕКТА**

# Статистическа извадка (2)



**СТАТИСТИКА** – се нарича всяка функция  $\varphi = \phi (X_1, X_2, \dots, X_n)$  която зависи само от елементите на извадката

**СТАТИСТИЧЕСКИ ОЦЕНКИ** – изчислени параметри на разпределение на СВ на база извадки с ограничен обем

Статистическите оценки са изчислими на база експериментални наблюдения, а истинските параметри са най-често непознаваеми, поради невъзможността ГС да се реализира експериментално

Самите статистически оценки са също случайни величини – те са **състоятелни** и **неизместени** оценки на параметрите на СВ

- **Състоятелна оценка** - която при нарастване на обема на извадката клони по вероятност към оценяваната стойност
- **Неизместена оценка** - статистическа оценка чието математическо очакване съвпада с оценяваната величина



# СТАТИСТИЧЕСКИ ОЦЕНКИ

ПАРАМЕТРИ на  
ГЕНЕРАЛНАТА  
СЪВОКУПНОСТ

$M(X)$   
Математическо  
очакване

ПАРАМЕТРИ на  
Нормално  
разпределение

$\mu$

Статистически  
оценки на  
параметрите

$\bar{X}$

Средно  
аритметично

$$\bar{X} = \frac{1}{N} \sum_{i=1}^n (x_i)$$

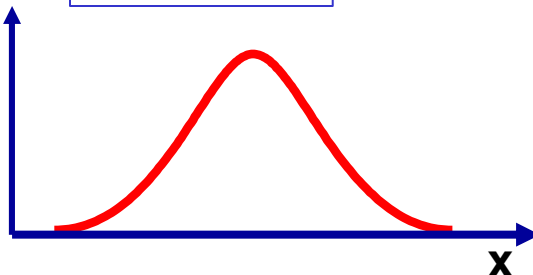
$D(X)$   
Дисперсия

$\sigma^2$

$S^2$

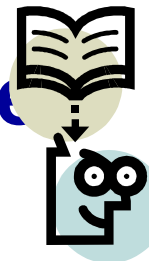
$$S = \sqrt{\frac{\sum_i (x_i - \bar{X})^2}{(N-1)}}$$

$P(x)$



Стандартно отклонение  
(средно квадратично  
отклонение)

# Алгоритми на изчисляване



- ИЗЧИСЛЯВАНЕ НА СРЕДНО АРИТМЕТИЧНО (arithmetic mean)

$$\bar{X} = \frac{(x_1 + x_2 + \dots + x_N)}{N} = \frac{\sum_{k=1}^n x_k}{N}$$

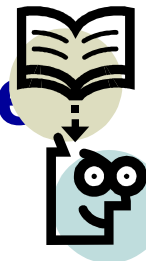
$$\bar{X} = \frac{(n_1 x_1 + n_2 x_2 + \dots + n_m x_m)}{N} = \sum_{k=1}^M \frac{n_k}{N} x_k$$

$$\bar{X} = p_1 x_1 + p_2 x_2 + \dots + p_M x_M = M(X)$$

$$M(X) = \int_{-\infty}^{+\infty} X \cdot p(X) dX$$

$$M(X) = \sum_{k=1}^{\infty} p(X_k) \cdot X_k$$

# Алгоритми на изчисляване



- ИЗЧИСЛЯВАНЕ НА СРЕДНО КВАДРАТИЧНО ОТКЛОНЕНИЕ (стандартно отклонение) (*standard deviation*)

$$S = \sqrt{\frac{(\bar{X} - x_1)^2 + (\bar{X} - x_2)^2 + \dots + (\bar{X} - x_N)^2}{N - 1}}$$

$$S = \sqrt{\frac{\left(\sum_{i=1}^N x_i^2\right) - \frac{\left(\sum_{i=1}^N x_i\right)^2}{N}}{N - 1}}$$

Използват се степените на свобода

(**degrees of freedom,  $N - 1$** ), вместо  $N$ .

за да не изместим оценката тъй като една стойност е дефинирана – т.е. имаме една степен свобода по-малко. При голям брой измервания ефективната разлика между  $S$  изчислено с  $N$  и  $N - 1$  се заличава

$$D(X) = \sum_{k=1}^{\infty} p(X_k) [X_k - M(X)]^2$$

$$D(X) = \int_{-\infty}^{+\infty} p(X) [X - M(X)]^2 dX$$

## Средно квадратично отклонение на резултати

Повторяемост на метод за определяне на селен в храни 9 измервания на лаб проба осреднена от 20 пакета ориз.

$$S = \sqrt{\frac{(\bar{X} - x_1)^2 + (\bar{X} - x_2)^2 + \dots + (\bar{X} - x_N)^2}{N-1}}$$

$$S = \sqrt{\frac{\left(\sum_{i=1}^N x_i^2\right) - \frac{\left(\sum_{i=1}^N x_i\right)^2}{N}}{N-1}}$$

Проба	Se (mg/g) ( $x_i$ )	$x_i^2$
1	0.07	0.0049
2	0.07	0.0049
3	0.08	0.0064
4	0.07	0.0049
5	0.07	0.0049
6	0.08	0.0064
7	0.08	0.0064
8	0.09	0.0081
9	0.08	0.0064
Sum $x_i$	=0.69	Sum $x_i^2$ =0.0533

Средно =  $Sx_i/N = 0.077 \text{ mg/g}$

$(Sx_i)^2/N = 0.4761/9 = 0.0529$

Стандартно отклонение:  $s = \sqrt{\frac{0.0533 - 0.0529}{9-1}} = 0.00707106 = 0.007 \text{ mg/g}$

Коефициент на вариация = 9.2% **ИЗМЕРВАНА ВЕЛИЧИНА =  $77 \pm 7 \mu\text{g/g}$**

## Таблица за изчисляване на статистическите оценки :

mg/g	Номер на измерването	Стойност на измерването	X-Xcp	Suma(X-Xcp)2	Дисперсия	Standard deviation
	1	0,07	-0,006667	0,00004444	0,0000500	<b>0,00707</b>
	2	0,07	-0,006667	0,00004444		
	3	0,08	0,0033333	0,00001111		
	4	0,07	-0,006667	0,00004444		
	5	0,07	-0,006667	0,00004444		
	6	0,08	0,0033333	0,00001111		
	7	0,08	0,0033333	0,00001111		
	8	0,09	0,0133333	0,00017778		
	9	0,08	0,0033333	0,00001111		
	X_средно	0,0766667		0,00040000		



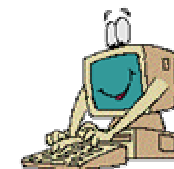
**AVERAGE(D31:D39)**

Excel.Ink

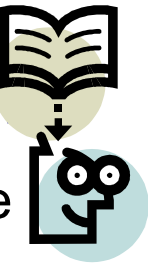


**STDEV(D3:D39)**

Excel.Ink



# ДЕФИНИЦИИ



**ПРЕЦИЗНОСТ (2.35 nVIM)** – степен на съвпадение между стойностите на величината, получени чрез **повтаряне на измервания** на величина при определени условия

**характеризира близостта на резултатите един спрямо друг**

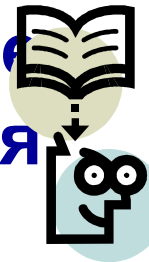
**(Precision: The closeness of data to each other)**

**ЕЛЕМЕНТИТЕ ХАРАКТЕРИЗИРАЩИ ПРЕЦИЗНОСТТА НА РЕЗУЛТАТИТЕ СА:**

- **ПОВТОРЯЕМОСТ** (repeatability) – VIM 5.27 – свойство на **система** за измерване да дава близки подобни **показания** за повтарящи се измервания на същата **величина при повтарящи се условия** (в условия на повторяемост) - количествена характеристика  **$S_r$**

- **ВЪЗПРОИЗВОДИМОСТ** (reproducibility) – условия на възпроизводимост (2.40 nVIM) условие на измерване от съвкупност от условия включващи **различни места, оператори, и системи за измерване** – **ВЪЗПРОИЗВОДИМОСТ НА ИЗМЕРВАНЕ** – прецизност в условия на възпроизводимост - количествена характеристика  **$S_R$**

# Относително стандартно отклонение коэффициент на вариация



## НЕЗАБРАВЯЙТЕ !!!

Стандартното отклонение  
(средноквадратичното отклонение)  
има размерност на **измерваната** величина

$$\bar{X} \rightarrow \text{units} \leftarrow S$$

ОТНОСИТЕЛНОТО СТАНДАРТНО  
ОКЛОНЕНИЕ е безразмерна величина  
(*relative standard deviation*),  
(*coefficient of variation*)

$$RSD \equiv CV = \frac{S(\text{units})}{\bar{X}(\text{units})} = \frac{S}{\bar{X}}$$

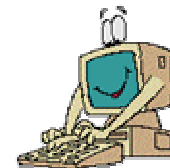
Означава се като RSD или CV

Много често относителното стандартно  
отклонение се дава в проценти RSD%

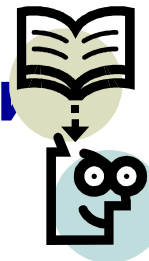
$$RSD \% = \frac{S}{\bar{X}} \cdot 100$$

**ПОЯТИЕТО ПРОЦЕНТНА ГРЕШКА  
ДА НЕ СЕ ИЗПОЛЗВА !!!**

показва каква част (може и в  
проценти) е разсейването на  
результатите спрямо  
големината на измерваната  
величина



# Стандартно отклонение на средни стойности

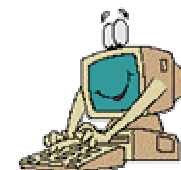


- Оценка на математическото очакване (резултата) може да се направи и чрез средна стойност на средни стойности.
- Средната стойност на средни стойности от  $m$  серии съвпада със средната стойност на обединените резултати от всички серии по  $n$  измервания
- Стандартното отклонение на средните стойности получени от  $N$  измервания е с **корен втори от  $N$  пъти по-малко** от стандартното отклонение на самата СВ

$$\bar{X}_{aver} = \bar{X}_1 + \bar{X}_2 + \dots + \bar{X}_m$$

$$\bar{X}_{aver} = \bar{X}_N \quad \text{за} \quad N = n * m$$

$$S_{\bar{X}} = \frac{S_X}{\sqrt{N}}$$



**Това е статистическо а не абсолютно равенство !!!**

$$D(\bar{X}) = \left(\frac{1}{N}\right)^2 \sum_{k=1}^M D(x_k) = \frac{1}{N} D(X)$$



# Днес ще разгледаме:

## 1. Статистически оценки.

- Средно-аритметична стойност и средно квадратично отклонение, относително стандартно отклонение
- Начини за изчисляване на статистическите оценки от извадки с ограничен обем
- Неизместеност и състоятелност на статистическите оценки.

## 2. Разпределения – характеризиращи статистическите оценки

- t-разпределение;
- $\chi^2$ -разпределение;
- F-разпределение

# Разпределение на Стюдънт (t-разпределение)

William Sealey Gosset  
1908

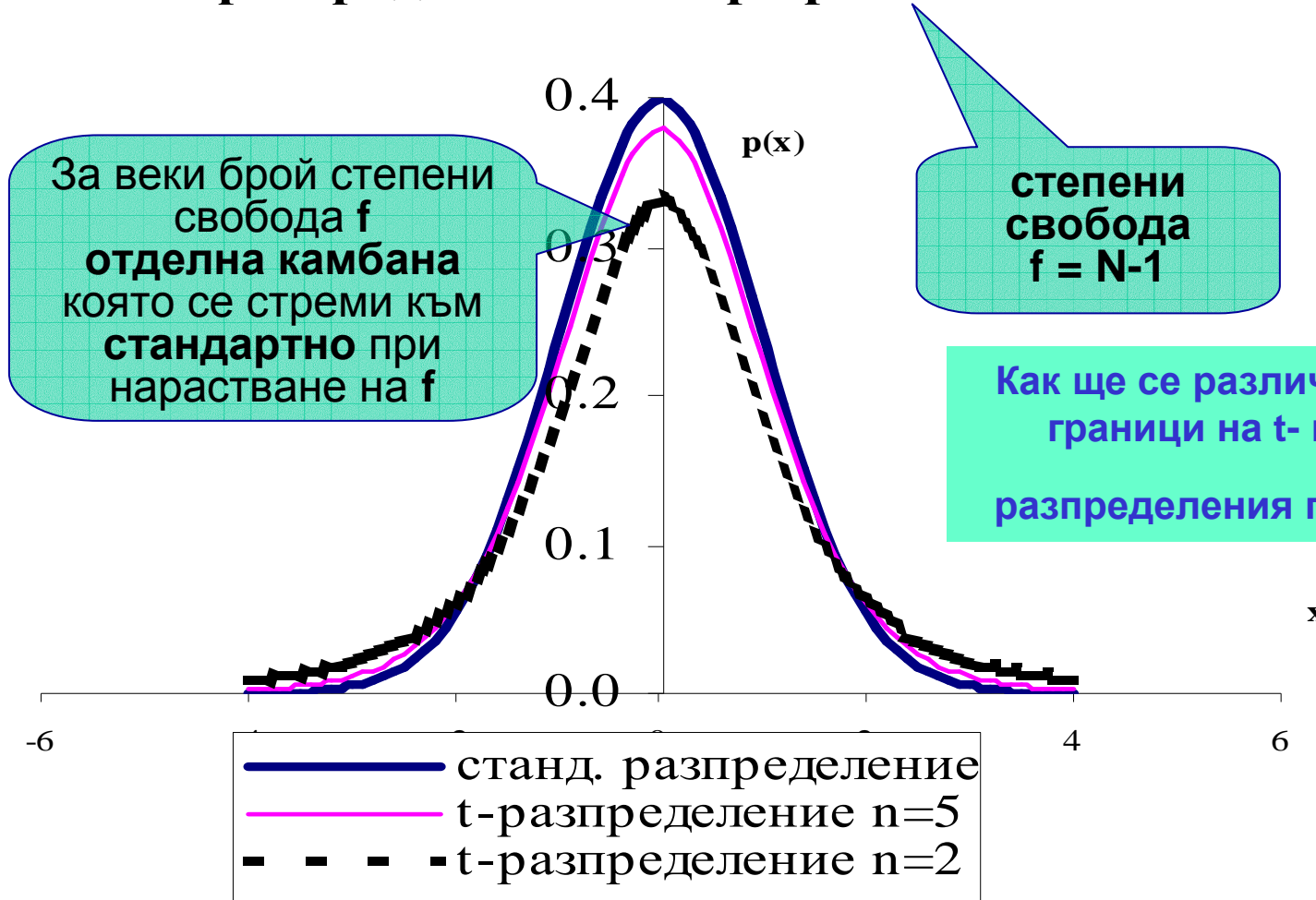


$$p(X) = C \cdot \left(1 + \frac{X^2}{N}\right)^{-\frac{(N+1)}{2}}$$

$$C = \frac{\gamma[(N+1)/2]}{[\sqrt{N\pi} \cdot \gamma(N/2)]}$$

# Разпределение на Стюдънт (t-разпределение)

Плътност на стандартното ( $\mu=0; \sigma=1$ ) и t-разпределението при различни  $n$



$$\frac{(\bar{X} - \mu) * \sqrt{N}}{S}$$

# ЗАПОМНИ ТОЗИ ОБОБЩЕН МОДЕЛ !!



за ВСИЧКИ СВ значенията им могат да заемат стойности около  $M(X)$  с вероятност  $P \leq 1$  (100 %)

за **нормално разпределени** СВ значенията им се разполагат в области около  $\mu$  като относителните им честоти на поява образуват камбана (Гаусова крива) с инфлексна точка  $\sigma$

за **ВСИЧКИ** нормално разпределени СВ значенията им се разполагат с **ЕДНАКВА ВЕРОЯТНОСТ** в съответни интервали измерени в  $\sigma$  спрямо  $\mu$  например:

$$\mu \pm 1 \sigma = 0,6827$$

$$\mu \pm 2 \sigma = 0,9545$$

$$\mu \pm 3 \sigma = 0,9973$$

ТОВА Е В СИЛА БЕЗ  
ЗНАЧЕНИЕ ОТ  $\mu$  и  $\sigma$   
т.е.  
за **ВСЯКО**  $\mu$  и  $\sigma$

**Z**

$$Z = \frac{X - \mu}{\sigma}$$

$$t = \frac{X - \bar{X}}{S / \sqrt{N}}$$

# Таблица на t-разпределение

Table 4-2 Values of Student's *t*

Двустранна постановка	75	95	95,7	99,0	99,5	99,75	99,995
Degrees of freedom	50	90	95	98	99	99.5	99.9
1	1.000	6.314	12.706	31.821	63.657	127.32	636.619
2	0.816	2.920	4.303	6.965	9.925	14.089	31.598
3	0.765	2.353	3.182	4.541	5.841	7.453	12.924
4	0.741	2.132	2.776	3.747	4.604	5.598	8.610
5	0.727	2.015	2.571	3.365	4.032	4.773	6.869
6	0.718	1.943	2.447	3.143	3.707	4.317	5.959
7	0.711	1.895	2.365	2.998	3.500	4.029	5.408
8	0.706	1.860	2.306	2.896	3.355	3.832	5.041
9	0.703	1.833	2.262	2.821	3.250	3.690	4.781
10	0.700	1.812	2.228	2.764	3.169	3.581	4.587
15	0.691	1.753	2.131	2.602	2.947	3.252	4.073
20	0.687	1.725	2.086	2.528	2.845	3.153	3.850
25	0.684	1.708	2.060	2.485	2.787	3.078	3.725
30	0.683	1.697	2.042	2.457	2.750	3.030	3.646
40					2.704	2.971	3.551
60					2.660	2.915	3.460
120					2.617	2.860	3.373
∞					2.576	2.807	3.291

Едностранны  
постановка

Статистическа  
сигурност

t - квантили

ВИЖ "ТЕ"!  
ПРИЛОЕНИЕ 3

ОБЯСНИ  
едностранна и  
двустранна  
постановка на  
задачата

Двустранна  
постановка

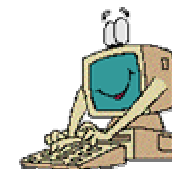
степени  
свобода

Тук това е  
 $\alpha = 1 - P$



Excel.lnk

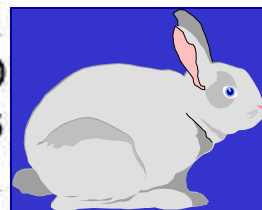
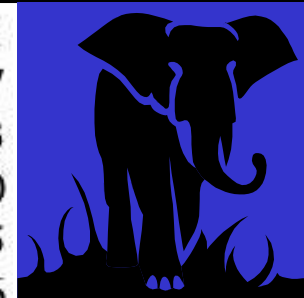
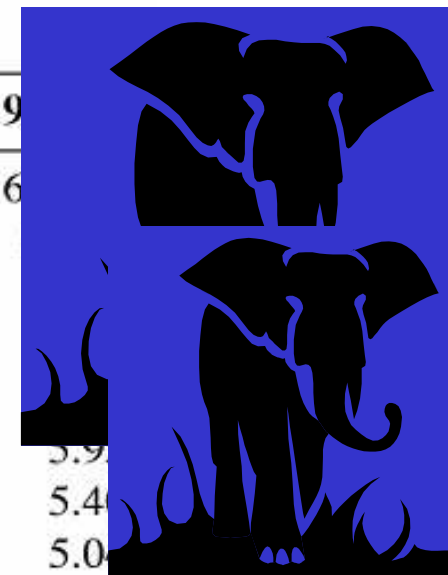
TINV(probability,degrees\_freedom)



# Таблица на t-разпределение

Table 4-2 Values of Student's *t*

Degrees of freedom	Confidence level (%)						
	50	90	95	98	99	99.5	99.9
1	1.000	6.314	12.706	31.821	63.657	127.32	636.619
2	0.816	2.920	4.303	6.965	9.925	14.089	31.821
3	0.765	2.353	3.182	4.541	5.841	7.453	12.941
4	0.741	2.132	2.776	3.747	4.604	5.598	11.716
5	0.727	2.015	2.571	3.365	4.032	4.773	10.998
6	0.718	1.943	2.447	3.143	3.707	4.317	10.591
7	0.711	1.895	2.365	2.998	3.500	4.029	10.291
8	0.706	1.860	2.306	2.896	3.355	3.832	10.028
9	0.703	1.833	2.262	2.821	3.250	3.690	9.797
10	0.700	1.812	2.228	2.764	3.169	3.581	9.597
15	0.691	1.753	2.131	2.602	2.947	3.252	9.148
20	0.687	1.725	2.086	2.528	2.845	3.153	8.962
25	0.684	1.708	2.060	2.485	2.787	3.078	8.845
30	0.683	1.697	2.042	2.457	2.750	3.030	8.787
40	0.681	1.684	2.021	2.423	2.704	2.971	8.716
60	0.679	1.671	2.000	2.390	2.660	2.915	8.660
120	0.677	1.658	1.980	2.358	2.617	2.860	8.616
∞	0.674	1.645	1.960	2.326	2.576	2.807	8.581

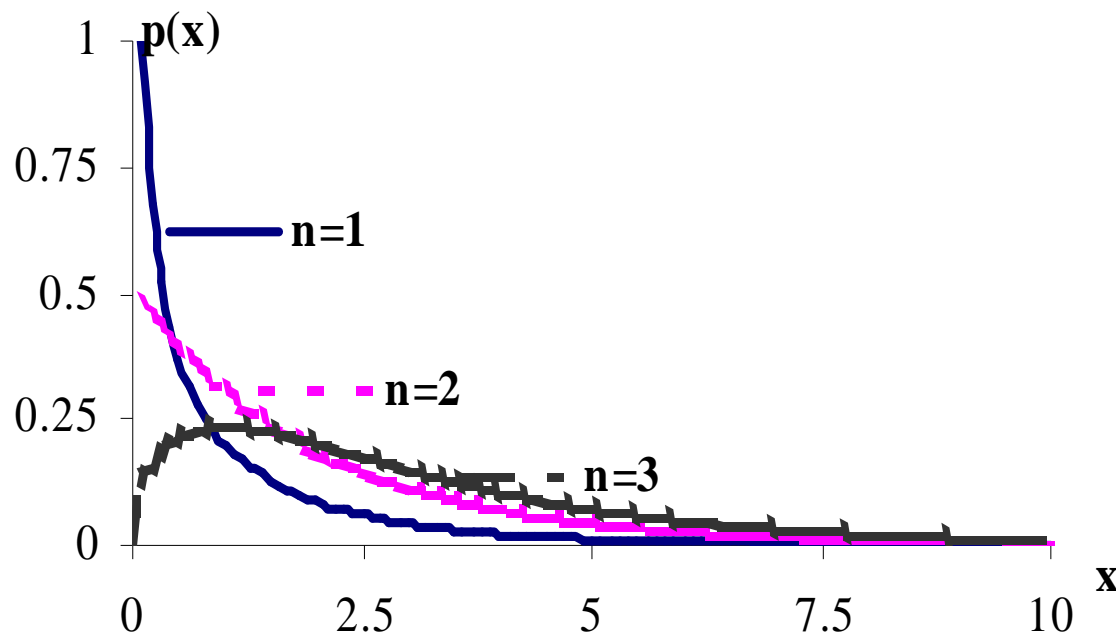


# Хи-квадрат разпределение

Стандартното отклонение е също случайна величина, като величината  $\chi^2$  е **ХИ** разпределена СВ

$$\chi^2 = S^2 \cdot (N - 1) / \sigma^2$$

Плътност на  $\chi^2$  - разпределението при няколко степени на свобода



$$p(X) = \begin{cases} C \cdot e^{-\frac{x}{2}} \cdot X^{\frac{N}{2}-1} & \\ 0; & X \leq 0 \end{cases}$$

**ВИЖ “ТЕ” !  
ПРИЛОЕНИЕ 2**

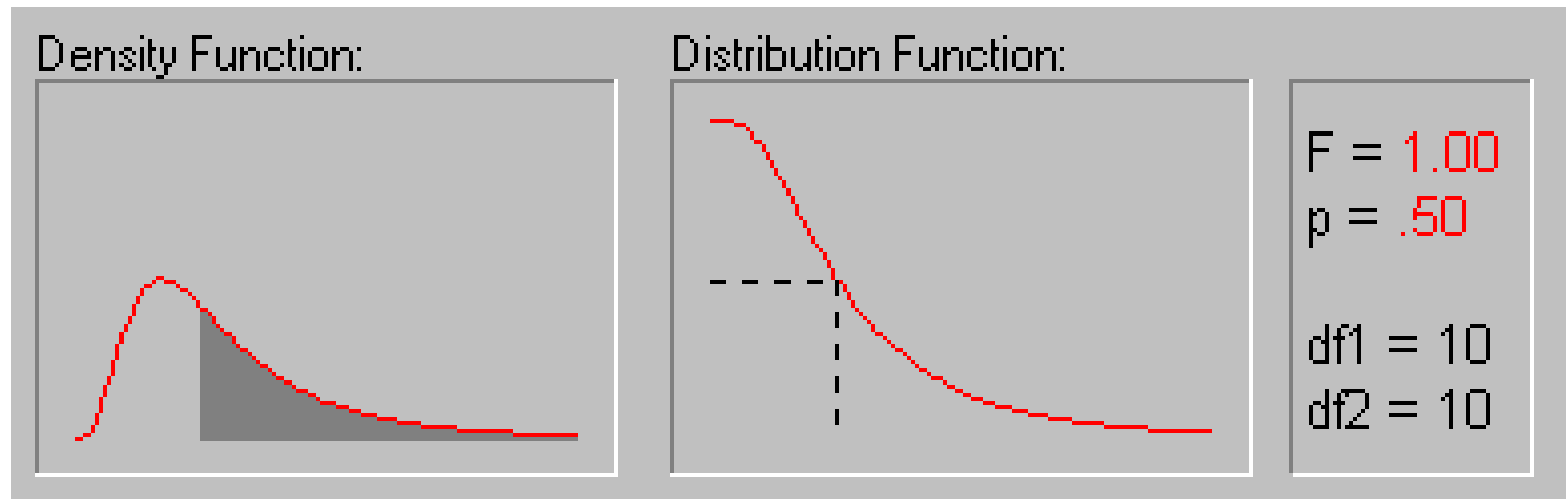
**ОБЯСНИ**  
Степени на  
свобода  $f=N-1$

# F – разпределение на Фишер

Това е разпределение  $F_{f_1, f_2}$  - със степени на свобода  $f_1 = N_1 - 1$  и  $f_2 = N_2 - 1$

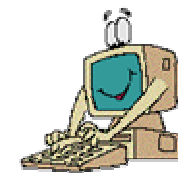
$$F = \frac{S_1^2}{S_2^2}$$

Задължително условие е:  
 $S_2 < S_1$



Excel.Ink

`FINV(probability,degrees_freedom1,degrees_freedom2)`



**ВИЖ “ТЕ” !  
ПРИЛОЖЕНИЕ 4**



# Днес разгледахме:

## 1. Статистически оценки.

- Средно-аритметична стойност и средно квадратично отклонение, относително стандартно отклонение
- Начини за изчисляване на статистическите оценки от извадки с ограничен обем
- Неизместеност и състоятелност на статистическите оценки.

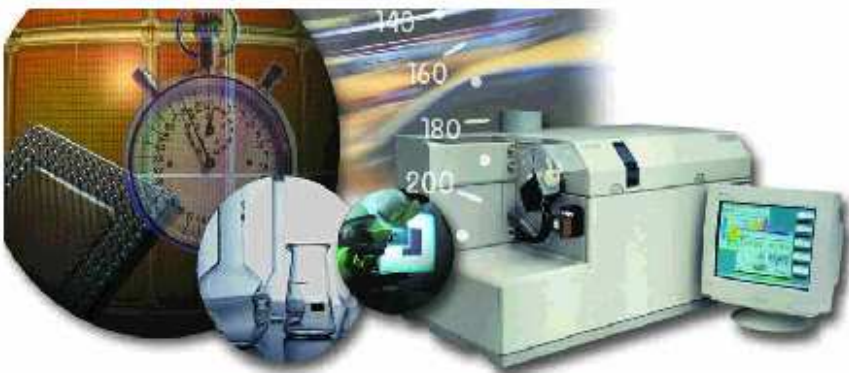
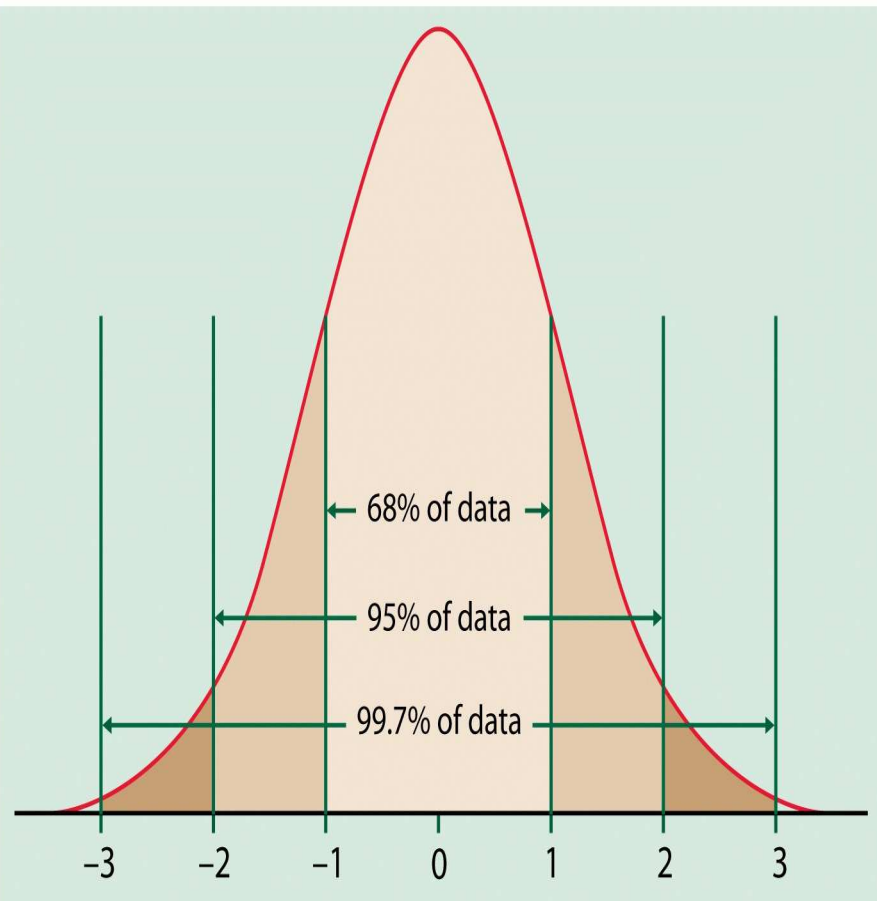
## 2. Разпределения – характеризиращи статистическите оценки

- t-разпределение;
- $\chi^2$ -разпределение;
- F-разпределение

# ИНТЕРВАЛНИ ОЦЕНКИ

значащи цифри

ДОВЕРИТЕЛЕН  
ИНТЕРВАЛ



# ВНИМАНИЕ – КОЛОКВИУМ

**На 9 АПРИЛ 2009  
(четвъртък)**



Education and Culture

**Socrates**

## ПО ПРОГРАМАТА ЗА АКАДЕМИЧЕН ОБМЕН **ЕРАЗЪМ**

В ХИМИЧЕСКИ ФАКУЛТЕТ при  
ПУ "Паисий ХИЛЕНДАРСКИ" гостува:

**Prof. Dr. Antonio Canals  
Hernández**

**UNIVERSITY of ALICANTE -  
SPAIN**



КОИТО ЩЕ ИЗНЕСЕ ИНТЕНЗИВЕН КУРС

### **Univariate Calibration on Instrumental Analysis**

ЛЕКЦИИТЕ СА ОТ 9:00 до 12:30 и от 14:00 до 17:00 НА  
**10.04.2009 (петък) В**

**5 Аудитория, Ректорат**

ПОКАНВАТ СЕ СТУДЕНТИ, ПРЕПОДАВАТЕЛИ, ГОСТИ

Лекциите са на АНГЛИЙСКИ ЕЗИК

# Днес ще разгледаме:

1. **Значещи цифри – правилно представяне на експериментални данни – правило  $\frac{1}{4} S$**
2. **ИНТЕРВАЛНИ ОЦЕНКИ - статистическа сигурност и ниво на значимост**
  - **ДОВЕРИТЕЛЕН ИНТЕРВАЛ НА СВ при ограничен брой измервания**
  - **Методи за изчисляване на доверителния интервал**
2. **СТАТИСТИЧЕСКИ ХИПОТЕЗИ**
  - **Основни понятия и същност на проверката на СХ.**
  - **Хипотеза за равенство на  $\mu$  на дадена стойност**

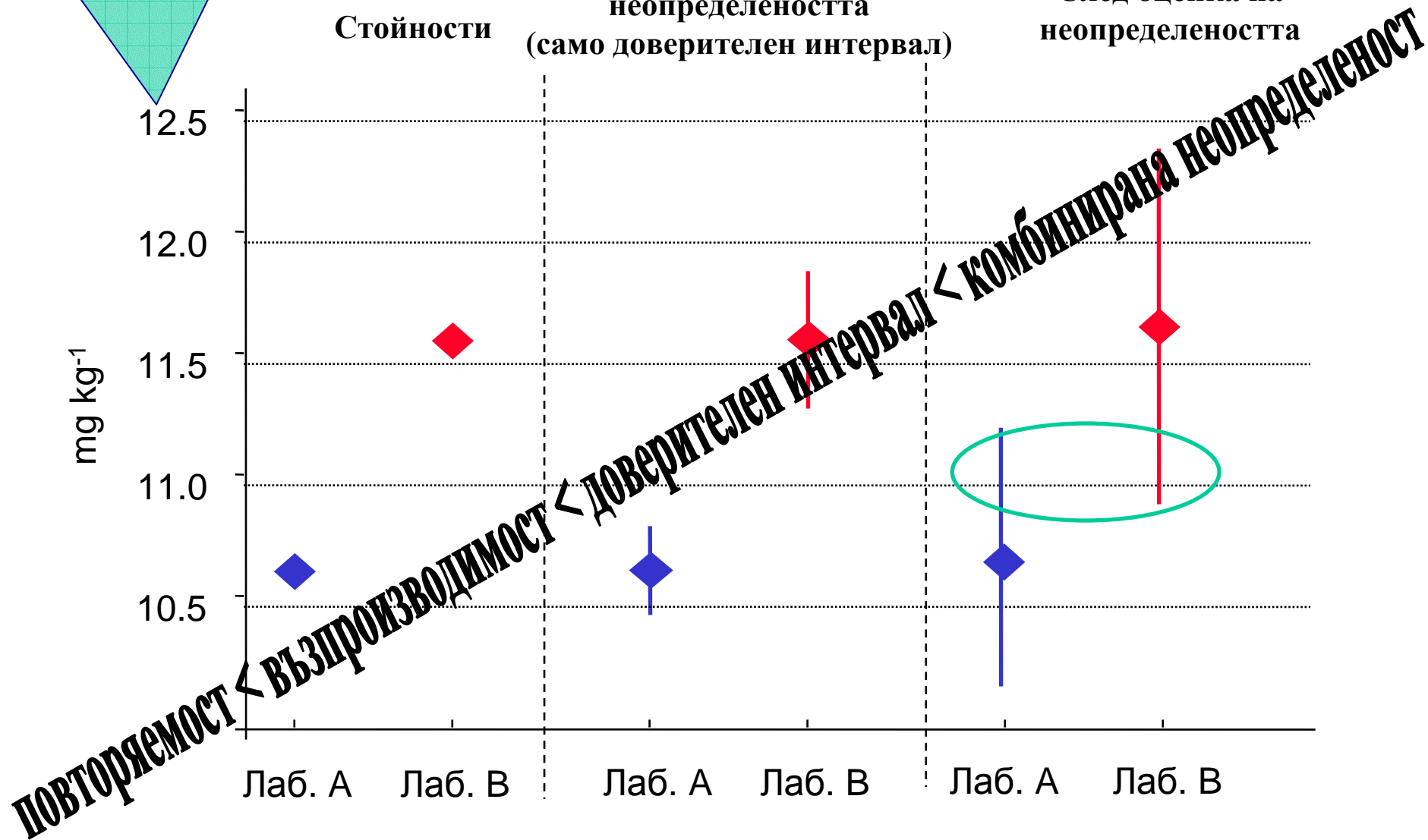
# ИНТЕРВАЛНИ ОЦЕНКИ

ЧИСЛАТА ПРЕДСТАВЯТ ГОЛЕМИНИ

Стойности

Без оценка на  
неопределеността  
(само доверителен интервал)

След оценка на  
неопределеността



**В химичните измервания стойностите на случайната величина могат да се разполагат в даден интервал, като заемат коя да е реална числена стойност в него**

**ВЕЛИЧИНАТА НЯМА СТОЙНОСТ !!!!!!!!**

**Приписана  
стойност !**



Колко е нивото на холестерола в кръвта ?

Има ли DDT в почвата ?

Токсични ли са пластелините за деца ?

Има ли злато в полиметалната руда ?

Замърсява ли Идеал Стандарт водата в Севлиево?

Колко е радиоактивността на отпадък от АЕЦ Козлодуй?

Вашите резултати отговарят ли на РТ?

Кой лаборант работи по-добре в лабораторията?

**! ТОЧКОВА оценка** - число представящо стойността на величината - обикновено състоятелна и неизместена оценка.

**! ИНТЕРВАЛНА оценка** - представя интервала от значения на СВ в който се очаква тя да се прояви с определена (изисквана) вероятност !

- **Пример:** Равновесната константа, изразяваща дисоциацията на оцетна киселина във воден разтвор при 25°C има стойност  $K=1,75 \cdot 10^{-5}$
- $pK = -\log K = 4.76$

ЧИСЛАТА ПРЕДСТАВЯТ  
ГОЛЕМИНИ

$$K_a = \frac{[H_3O^+].[CH_3COO^-]}{[CH_3COOH]}$$

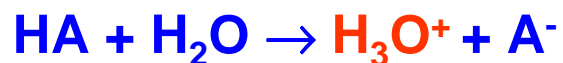


**TABLE 15.2 Ionization Constants of Some Weak Acids and Weak Bases in Water at 25 °C**

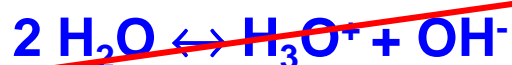
Ionization Equilibrium		Ionization Constant, $K$	p <i>K</i>
<b><i>Inorganic Acids</i></b>		$K_a =$	
Chlorous acid	$\text{HClO}_2 + \text{H}_2\text{O} \rightleftharpoons \text{H}_3\text{O}^+ + \text{ClO}_2^-$	$1.1 \times 10^{-2}$	1.96
Nitrous acid	$\text{HNO}_2 + \text{H}_2\text{O} \rightleftharpoons \text{H}_3\text{O}^+ + \text{NO}_2^-$	$7.2 \times 10^{-4}$	3.14
Hydrofluoric acid	$\text{HF} + \text{H}_2\text{O} \rightleftharpoons \text{H}_3\text{O}^+ + \text{F}^-$	$6.6 \times 10^{-4}$	3.18
Hypochlorous acid	$\text{HOCl} + \text{H}_2\text{O} \rightleftharpoons \text{H}_3\text{O}^+ + \text{OCl}^-$	$2.9 \times 10^{-8}$	7.54
Hypobromous acid	$\text{HOBr} + \text{H}_2\text{O} \rightleftharpoons \text{H}_3\text{O}^+ + \text{OBr}^-$	$2.5 \times 10^{-9}$	8.60
Hydrocyanic acid	$\text{HCN} + \text{H}_2\text{O} \rightleftharpoons \text{H}_3\text{O}^+ + \text{CN}^-$	$6.2 \times 10^{-10}$	9.21
<b><i>Carboxylic Acids</i></b>		$K_a =$	
Chloroacetic acid	$\text{CH}_2\text{ClCOOH} + \text{H}_2\text{O} \rightleftharpoons \text{H}_3\text{O}^+ + \text{CH}_2\text{ClCOO}^-$	$1.4 \times 10^{-3}$	2.85
Formic acid	$\text{HCOOH} + \text{H}_2\text{O} \rightleftharpoons \text{H}_3\text{O}^+ + \text{HCOO}^-$	$1.8 \times 10^{-4}$	3.74
Benzoic acid	$\text{C}_6\text{H}_5\text{COOH} + \text{H}_2\text{O} \rightleftharpoons \text{H}_3\text{O}^+ + \text{C}_6\text{H}_5\text{COO}^-$	$6.3 \times 10^{-5}$	4.20
Acetic acid	$\text{CH}_3\text{COOH} + \text{H}_2\text{O} \rightleftharpoons \text{H}_3\text{O}^+ + \text{CH}_3\text{COO}^-$	$1.8 \times 10^{-5}$	4.74
<b><i>Inorganic Bases</i></b>		$K_b =$	
Ammonia	$\text{NH}_3 + \text{H}_2\text{O} \rightleftharpoons \text{NH}_4^+ + \text{OH}^-$	$1.8 \times 10^{-5}$	4.74
Hydrazine	$\text{H}_2\text{NNH}_2 + \text{H}_2\text{O} \rightleftharpoons \text{H}_2\text{NNH}_3^+ + \text{OH}^-$	$8.5 \times 10^{-7}$	6.07
Hydroxylamine	$\text{HONH}_2 + \text{H}_2\text{O} \rightleftharpoons \text{HONH}_3^+ + \text{OH}^-$	$9.1 \times 10^{-9}$	8.04
<b><i>Amines</i></b>		$K_b =$	
Dimethylamine	$(\text{CH}_3)_2\text{NH} + \text{H}_2\text{O} \rightleftharpoons (\text{CH}_3)_2\text{NH}_2^+ + \text{OH}^-$	$5.9 \times 10^{-4}$	3.23
Ethylamine	$\text{CH}_3\text{CH}_2\text{NH}_2 + \text{H}_2\text{O} \rightleftharpoons \text{CH}_3\text{CH}_2\text{NH}_3^+ + \text{OH}^-$	$4.3 \times 10^{-4}$	3.37
Methylamine	$\text{CH}_3\text{NH}_2 + \text{H}_2\text{O} \rightleftharpoons \text{CH}_3\text{NH}_3^+ + \text{OH}^-$	$4.2 \times 10^{-4}$	3.38
Pyridine	$\text{C}_5\text{H}_5\text{N} + \text{H}_2\text{O} \rightleftharpoons \text{C}_5\text{H}_5\text{NH}^+ + \text{OH}^-$	$1.5 \times 10^{-9}$	8.82
Aniline	$\text{C}_6\text{H}_5\text{NH}_2 + \text{H}_2\text{O} \rightleftharpoons \text{C}_6\text{H}_5\text{NH}_3^+ + \text{OH}^-$	$7.4 \times 10^{-10}$	9.13

### 3.1.1 pH в разтвори на силни киселини

- Ако концентрацията на киселината е  $C_{HA} > 10^{-6}$  (mol/l)



1.  $C_{HA} = [A^-] = [H_3O^+] > 10^{-6}$



2.  $[H_3O^+].[OH^-]=K_w=10^{-14}$

3.  $[H_3O^+]$  от автопротолиза на вода може да се пренебрегне

следователно  $[H_3O^+] = C_{HA}$

$$pH = - \lg C_{HA}$$

Съкратено  
уравнение

- За силно разредени киселини, когато концентрацията е  $C_{HA} < 10^{-6}$  (mol/l) се решава пълното квадратно уравнение

$$[H_3O^+] = \frac{C_{HA} + \sqrt{C_{HA}^2 + 4K_w}}{2}$$

## Изчисляване на рН в разтвори на слаби киселини - примери

6) опростяване: HF е средно силна киселина и е изпълнено условието:

$$\checkmark C_{HA} \cdot K_a = 0,163 \cdot 10^{-4} = 6,3 \cdot 10^{-5} \gg 10^{-14}$$

$$K_a = \frac{[H_3O^+]^2}{C_{HA} - [H_3O^+]} \quad [H_3O^+] = \frac{-K_a + \sqrt{K_a^2 + 4K_a \cdot C_{HA}}}{2}$$

$$[H_3O^+] = \frac{-6,3 \cdot 10^{-4} + \sqrt{6,3^2 \cdot 10^{-8} + 4 \cdot 6,3 \cdot 10^{-5}}}{2}$$

$$[H_3O^+] = \frac{-6,3 \cdot 10^{-4} + \sqrt{\cancel{3,97 \cdot 10^{-7}} + 2,52 \cdot 10^{-4}}}{2}$$

$$\begin{aligned} [H_3O^+] &= \frac{-6,3 \cdot 10^{-4} + 1,59 \cdot 10^{-2}}{2} = \\ &= \frac{0,0159 - 0,00063}{2} = \frac{0,0153}{2} = 0,00765 \end{aligned}$$

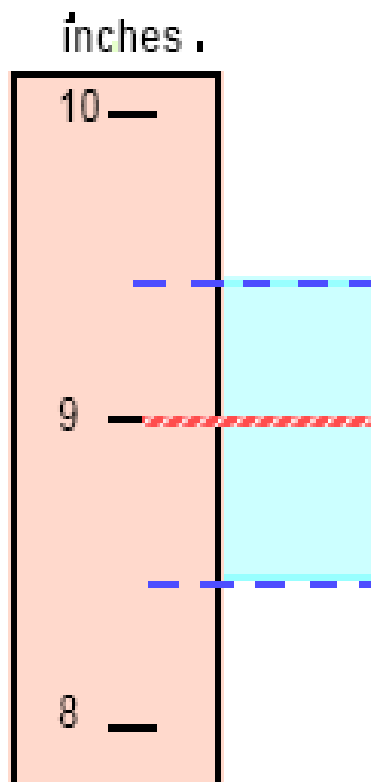
# Дискретизация при числови данни



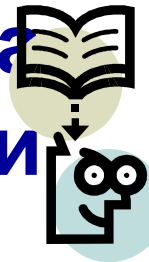
Какво означава размер 9 инча ?

Как ще изразите този размер в см ?

**ВСЯКО ЦИФРОВО ПРЕДСТАВЯНЕ НА ГОЛЕМИНАТА НА  
ВЕЛИЧИНА Е ДИСКРЕТИЗИРАНЕ НА ТАЗИ ВЕЛИЧИНА**



# Значещи цифри - представяне на резултати



ВСЯКО ЦИФРОВО ПРЕДСТАВЯНЕ НА ГОЛЕМИНАТА НА ВЕЛИЧИНА Е ДИСКРЕТИЗИРАНЕ НА ТАЗИ ВЕЛИЧИНА

Значещи цифри - доказуемост на резултата - последната цифра от резултата е **НЕСИГУРНА**, а предпоследната и останалите - **СИГУРНИ**

**⊗** **Правило (1/4)S** - стандартното отклонение се разделя на четири и се закръглява до една цифра, която се приема за последна (несигурна)

№	X	S	1/4 S	закръглено	резултат	
1	18,3465	0,025	0,006	0,01	18,35 ± 0,02	ЕДИНИЦИ
2	18,3465	0,53	0,13	0,1	18,3 ± 0,5	ЕДИНИЦИ
3	18,3465	1,17	0,29	0,1	18,3 ± 1,2	ЕДИНИЦИ
4	18,3465	3,56	0,9	1	18 ± 4	ЕДИНИЦИ

# Значещи цифри - представяне на резултати



☒ Правило - неопределеността на математическия израз не може да надминава неопределеността на най-неопределения (неточен) член в него.

☒ Неопределеността на сума (разлика) от случайни величини е равна на сумата на абсолютните им стойности на абсолютните им неопределености |,

☒ Относителната неопределеност на произведение (частно) от случайни величини е равна на сумата на абсолютните им стойности на относителните им неопределености |,

# Значещи цифри - представяне на резултати

- ☒ При сумиране и разлика на закръглени числа всички числа се представят с толкова цифри след десетичната запетая, колкото има числото с най-малко значещи цифри

$$1,67367 + 33,4 + 0,21 = 1,7 + 33,4 + 0,2 = 35,3$$

- ☒ Ако се изчислява произведение или частно – то резултатът е число със значещи цифри от определени от този множител с **най-голяма относителна неопределеност**.

$$\text{но } 1,67367 \cdot 33,4 \cdot 0,21 = 11,74 = 12$$

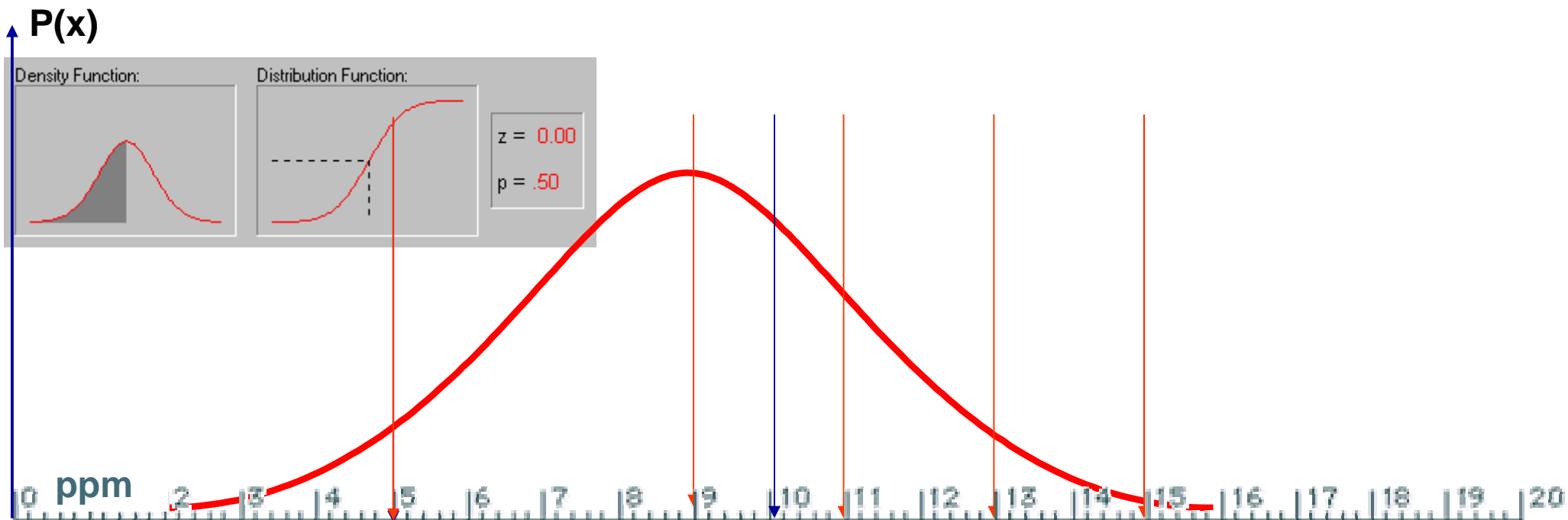
Резултат от калкулатор	Закръглена стойност	Пояснение
$\frac{21.023}{1.14} = 18.4422807$	18.4	Числителят е представен с неопределеност - 1 част от 21 000, но знаменателя е известен до 1 част от 114 (0,9%); закръгления резултат е с неопределеност 1 част от 184, или 0,5%.
$\frac{5.030 \times 10^{-9} \times (1.19 \times 10^6)}{3.1 \times 10^{-9}} = 1.930871 \text{ E6}$	1.9E6	Факторът “3.1” е с най-голяма неопределеност и е представен като 1 част от 31 (3%). Резултатът - 1.9 е закръглен до 1 част от 19 (5%), което е съизмеримо с неопределеността на фактора “3.1”.
<p>Определена книга има дебелина от 117mm; намерете ширината на 24 еквивалентни книги:</p> $24 \text{ books} \times \frac{117 \text{ mm}}{1 \text{ book}} = 2808 \text{ mm}$	2810 mm	“24” и “1” са константи и единствения фактор, имащ значение при закръглянето е дебелината на всяка книга, представена с 3 значещи цифри. Закръгля се до 3 значещи цифри.
$\begin{array}{r} .01 \\ + .0007 \\ \hline .4 \end{array} = 10.4107$	10.4	При събиране или изваждане, погледнете члена с най-малък брой числа след десетичната запетая и закръглете до тази значеща цифра.
<p>Превърнете “9 инча” в сантиметри:</p> $9 \text{ in} \times \frac{2.54 \text{ cm}}{1 \text{ in}} = 22.06 \text{ cm}$	23 cm	(Виж по-горе)



# Днес ще разгледаме:

1. Значещи цифри – правилно представяне на експериментални данни – правило  $\frac{1}{4} S$
2. **ИНТЕРВАЛНИ ОЦЕНКИ - статистическа сигурност и ниво на значимост**
  - **ДОВЕРИТЕЛЕН ИНТЕРВАЛ НА СВ** при ограничен брой измервания
  - Методи за изчисляване на доверителния интервал
2. **СТАТИСТИЧЕСКИ ХИПОТЕЗИ**
  - Основни понятия и същност на проверката на СХ.
  - Хипотеза за равенство на  $\mu$  на дадена стойност

# Съдържание на Au в полиметална руда



## Задача:

Съдържание на злато в руда ppm

Съдържанието на Au (ppm) в полиметална руда от Родопите е случайна нормално разпределена величина  $N(9;2^2)$

## Изчислете:

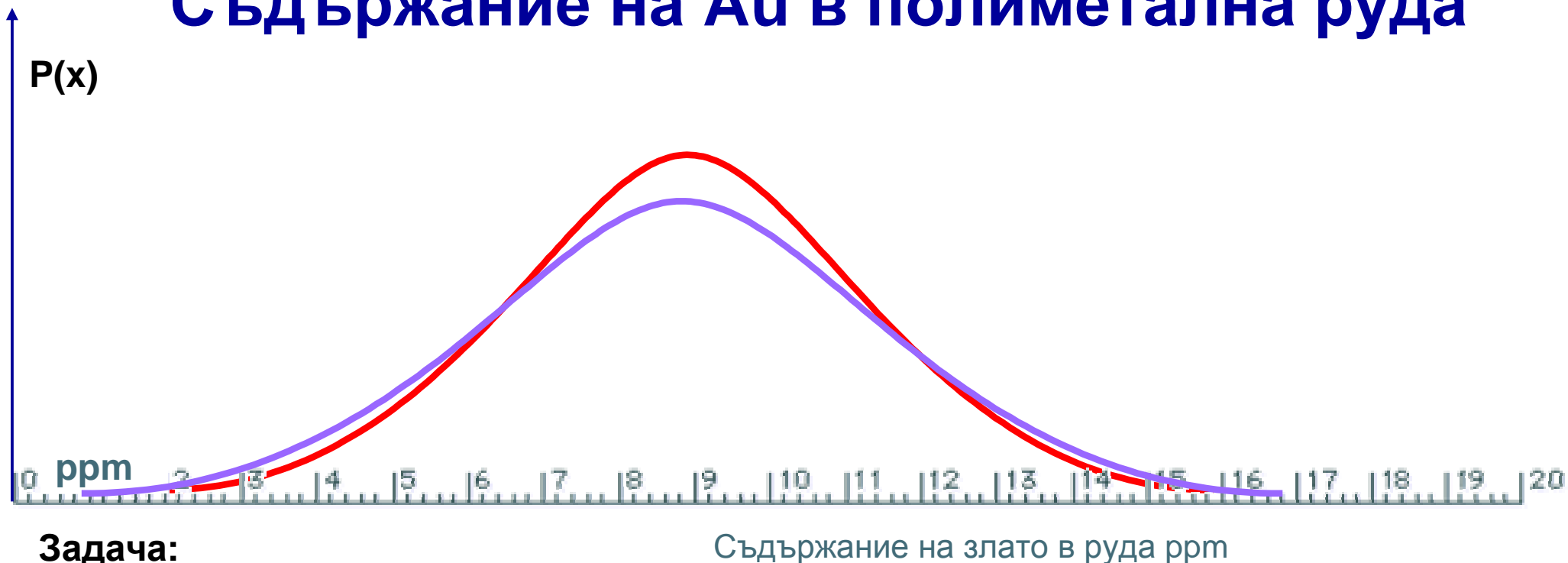
Симетричния интервал в който попадат 95% от значенията на СВ

Симетричния интервал в който попадат 99% от значенията на СВ

Каква е вероятността рудата да съдържа злато над 10 mg/g ?

Каква е вероятността златото да е по-малко от 5 mg/g ?

# Съдържание на Au в полиметална руда



## Задача:

Съдържанието на Au (ppm) в полиметална руда от Родопите е случайна нормално разпределена величина със статистически оценки от 5 измервания –  $\bar{X} = 9$  ppm и  $S = 2$  ppm

## Изчислете:

Симетричния интервал в който попадат 95% от значенията на СВ

Симетричния интервал в който попадат 99% от значенията на СВ

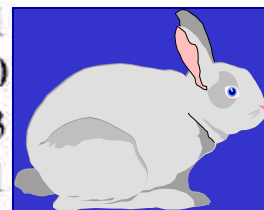
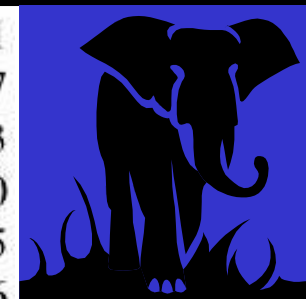
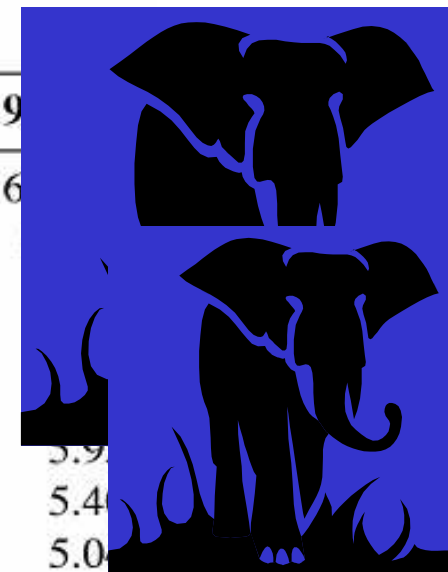
Каква е вероятността рудата да съдържа злато над 10 mg/g ?

Каква е вероятността златото да е по-малко от 5 mg/g ?

# Таблица на t-разпределение

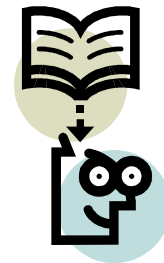
Table 4-2 Values of Student's *t*

Degrees of freedom	Confidence level (%)						
	50	90	95	98	99	99.5	99.9
1	1.000	6.314	12.706	31.821	63.657	127.32	631.0
2	0.816	2.920	4.303	6.965	9.925	14.089	31.821
3	0.765	2.353	3.182	4.541	5.841	7.453	10.136
4	0.741	2.132	2.776	3.747	4.604	5.598	9.000
5	0.727	2.015	2.571	3.365	4.032	4.773	8.591
6	0.718	1.943	2.447	3.143	3.707	4.317	8.163
7	0.711	1.895	2.365	2.998	3.500	4.029	7.794
8	0.706	1.860	2.306	2.896	3.355	3.832	7.501
9	0.703	1.833	2.262	2.821	3.250	3.690	7.259
10	0.700	1.812	2.228	2.764	3.169	3.581	7.063
15	0.691	1.753	2.131	2.602	2.947	3.252	6.578
20	0.687	1.725	2.086	2.528	2.845	3.153	6.375
25	0.684	1.708	2.060	2.485	2.787	3.078	6.256
30	0.683	1.697	2.042	2.457	2.750	3.030	6.191
40	0.681	1.684	2.021	2.423	2.704	2.971	6.131
60	0.679	1.671	2.000	2.390	2.660	2.915	6.080
120	0.677	1.658	1.980	2.358	2.617	2.860	6.040
∞	0.674	1.645	1.960	2.326	2.576	2.807	6.000



# ДОВЕРИТЕЛЕН ИНТЕРВАЛ

## Confidential interval (CI)



ГОРНА И ДОЛНА ГРАНИЦА МЕЖДУ КОЯТО СЕ ПРИЕМА ЧЕ СЕ НАМИРА ИСТИНСКАТА СТОЙНОСТ НА ИЗМЕРВАНАТА ВЕЛИЧИНА СЪС СЪОТВЕТНА СТАТИСТИЧЕСКА СИГУРНОСТ

Low\_стойност < ИСТИНСКА СТОЙНОСТ < High\_стойност

6.5 ppm < истинска стойност < 11.5 ppm

при  $S = 2$  ppm и статистическа сигурност  $P=0,95$

За симетричен около  $\mu$  (или средната стойност) = 9 интервал доверителния интервал може да се задава чрез  $\pm$  полу интервала

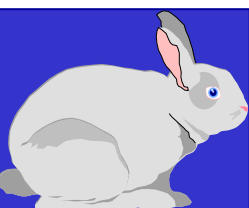
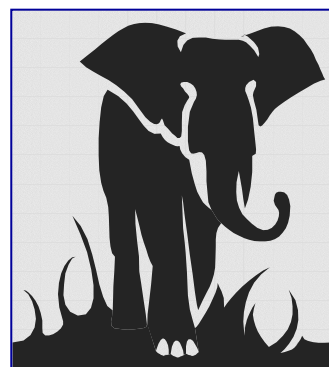
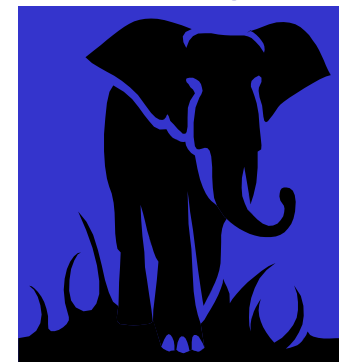
истинска стойност  $\in 9,0 \pm 2,5$  ppm

при  $S = 2$  ppm и статистическа сигурност  $P=0,95$

# АНАЛОГИЯ СЪС Z-трансформация при t-разпределена величина



- Припомнете си Z-трансформацията !!
  - Как ще procedираме когато не знаем  $\sigma$  ?
  - Как ще procedираме когато не знаем  $\mu$  ?

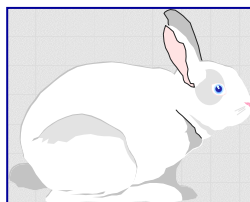


$$Z = \frac{X - \mu}{\sigma}$$

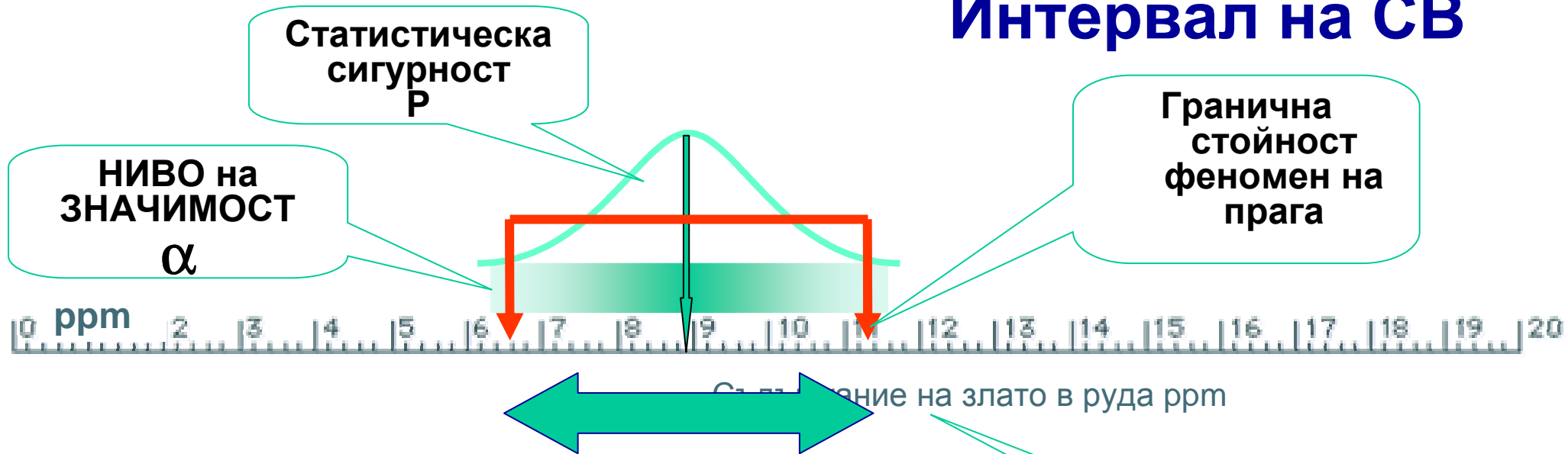


$$Z_t = \frac{(X - \bar{X}) * \sqrt{N}}{S}$$

t – квантил  
при  $f = N-1$   
и  $\alpha = ??$



# Интервал на СВ



?

$$(\mu - Z\sigma < x < \mu + Z\sigma)$$

$$(\mu - 3\sigma < x < \mu + 3\sigma)$$

$$(\bar{X} - 3\sigma/\sqrt{N} < x < \bar{X} + 3\sigma/\sqrt{N})$$

$$(\bar{X} - t^*S/\sqrt{N} < x < \bar{X} + t^*S/\sqrt{N})$$

$$(\bar{X} - U < x < \bar{X} + U)$$

# ДОВЕРИТЕЛЕН ИНТЕРВАЛ (confidential interval) интервална оценка за истинската стойност

- Изчислява се средно аритметичната стойност  $X_{\text{ср}}$
- Изчислява се стандартното отклонение  $S$  от измерванията
- Изчисляват се степените на свобода  $f = (N-1)$
- Избира се  $\alpha$  и се изчислява  $P = 1 - \alpha$  (най-често  $\alpha = 0.05$  т.е.  $P=95\%$ )
- Намира се стойността на  $t(f, \alpha)$  като интегрална граница при двустранна постановка, за съответното  $\alpha$  и  $f$
- Изчислява се ДИ съгласно формулата

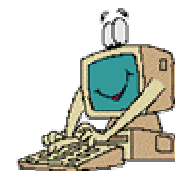
истинската стойност се намира с вероятност  $P = (1-\alpha)$  в интервала:

$$\bar{X} - \frac{t(f, \alpha).S}{\sqrt{N}} < \mu < \bar{X} + \frac{t(f, \alpha).S}{\sqrt{N}}$$

ВИЕ  
ИЗБИРАТЕ  
 $\alpha = ??$

$$\mu \in \bar{X} \pm t_{(f, \alpha)} S_{\bar{X}} = \bar{X} \pm \frac{t_{(f, \alpha)} \times S}{\sqrt{N}}$$

☒ Доверителният интервал намалява при -  
намаляване на  $S$  и увеличаване на  $\alpha$  и  $N$





# Таблица на t-разпределение

Table 4-2 Values of Student's *t*

Confidence level (%)

Degrees of freedom	50	90	95	98	99	99.5	99.9
1	1.000	6.314	12.706	31.821	63.657	127.32	636.619
2	0.816	2.920	4.303	6.965	9.925	14.089	31.598
3	0.765	2.353	3.182	4.541	5.841	7.453	12.924
4	0.741	2.132	2.776	3.747	4.604	5.598	8.610
5	0.727	2.015	2.571	3.365	4.032	4.773	6.869
6	0.718	1.943	2.447	3.143	3.707	4.317	5.959
7	0.711	1.895	2.365	2.998	3.500	4.029	5.408
8	0.706	1.860	2.306	2.896	3.355	3.832	5.041
9	0.703	1.833	2.262	2.821	3.250	3.690	4.781
10	0.700	1.812	2.228	2.764	3.169	3.581	4.587
15	0.691	1.753	2.131	2.602	2.947	3.252	4.073
20	0.687	1.725	2.086	2.528	2.845	3.153	3.850
25	0.684	1.708	2.060	2.485	2.787	3.078	3.725
30	0.683	1.697	2.042	2.457	2.750	3.030	3.646
40	0.681	1.684	2.021	2.423	2.704	2.971	3.551
60	0.679	1.671	2.000	2.390	2.660	2.915	3.460
120	0.677	1.658	1.980	2.358	2.617	2.860	3.373
∞	0.674	1.645	1.960	2.326	2.576	2.807	3.291

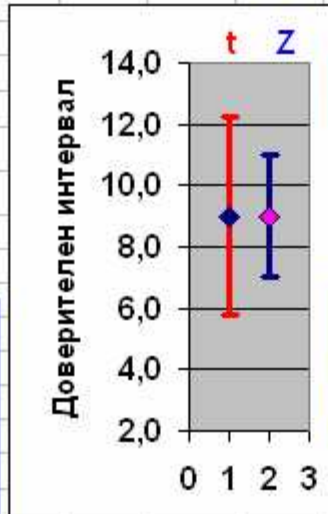
# Изчисляване на доверителен интервал

!!ВНИМАНИЕ - функцията CONFIDENCE - работи със стандартно разпределение (СИГМА) и НЕ МОЖЕ ДА СЕ ИЗПОЛЗВА !!!  
 Може да се използва Discriptive Statistics за извадка измервания (от Tools - Data Analysis)

$$\mu \in \bar{X} \pm \frac{t(f, \alpha) \cdot S}{\sqrt{N}}$$

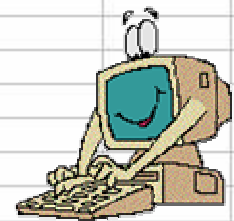
ДИ намалява при намаляване на S  
и увеличаване на  $\alpha$  и N

14	n=	4
8	f=(n-1)=	3
9	SQRT(n)	2
10	Alfa=	0,05
11	P=	95%
12	X-aver=	9,0
13	SD=	2,0
14	t(f,alfa)=	3,182449
15	CI=	3,2181
16	за СИГМА	1,9819
17	Low=	5,7819
18	High=	12,2181

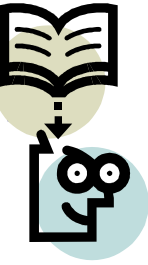


1	7,1
2	9,9
3	8,2
4	7,7
5	12,1
AVER	9,0
SD	2,0
RSD	22,5%

Column 1	
Mean	9
Standard Error	0,9044335
Median	8,2
Mode	#N/A
Standard Deviation	2,0223748
Sample Variance	4,09
Kurtosis	0,1678433
Skewness	1,0554296
Range	5
Minimum	7,1
Maximum	12,1
Sum	45
Count	5
Largest(1)	12,1
Smallest(1)	7,1
Confidence Level(95,0%)	2,5111152



- ДО Тука



~~✗~~ ДОВЕРИТЕЛНА ВЕРОЯТНОСТ (probability)  $P$  ( $1-\alpha$ )  
**СТАТИСТИЧЕСКА СИГУРНОСТ**

И

~~✗~~ НИВО НА ЗНАЧИМОСТ (significance level)  $\alpha = 1 - P$  или  
размер на критерия (грешка от I род) (ОПАШЛЕТО – както казва Киро)

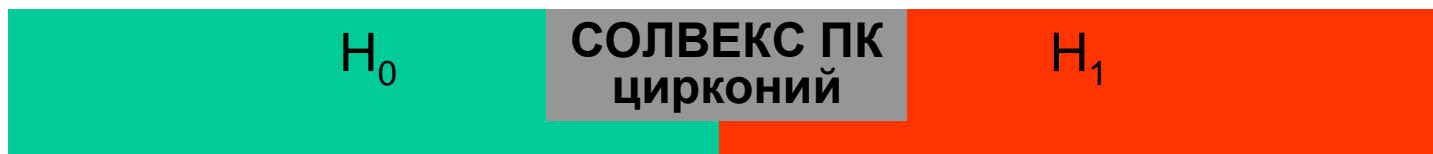
**ГРЕШКА ОТ ПЪРВИ РОД (Type I error)** - отхвърляне на ВЯРНАТА и приемане на ГРЕШНАТА - когато  $H_0$  е вярна, но стойностите на критерия попаднат в критичните области  
(бракува се годна продукция)

**ГРЕШКА ОТ ВТОРИ РОД (Type II error)**- приема се  $H_0$  при вярна алтернативната хипотеза. (брак минава за годен)

**Статистическите оценки** имат свое поведение и свои функции на разпределение,  
може да се прецени с определена вероятност какви стойности тези оценки заемат.

**ТЕ СЕ ЯВЯВАТ КРИТЕРИИ ЗА ПРОВЕРКА на статистически хипотези.**

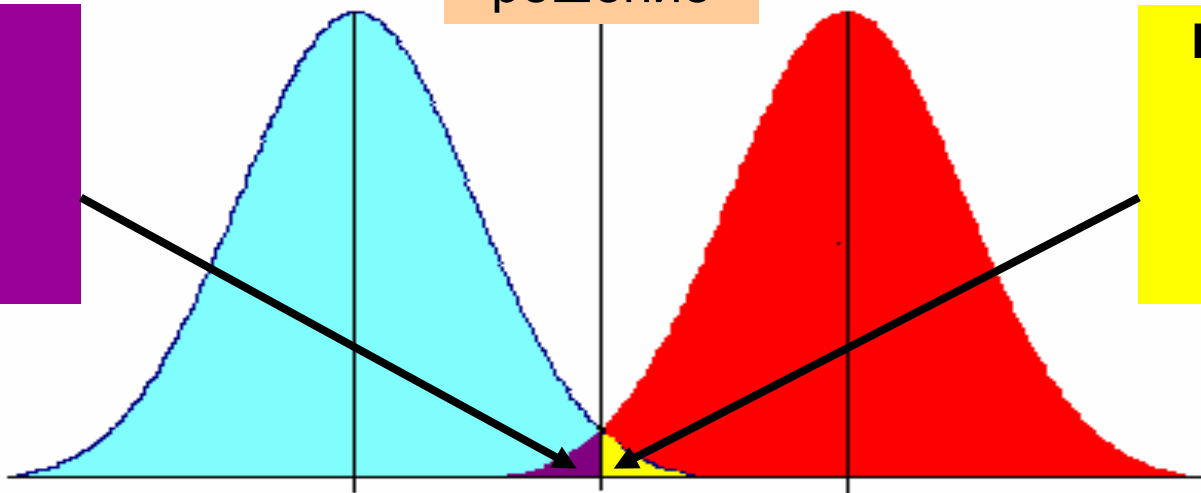
# ХИПОТЕЗИ - грешка от първи и от втори род



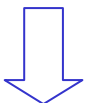
Граница на решение

FP област грешка от II род  $\beta$

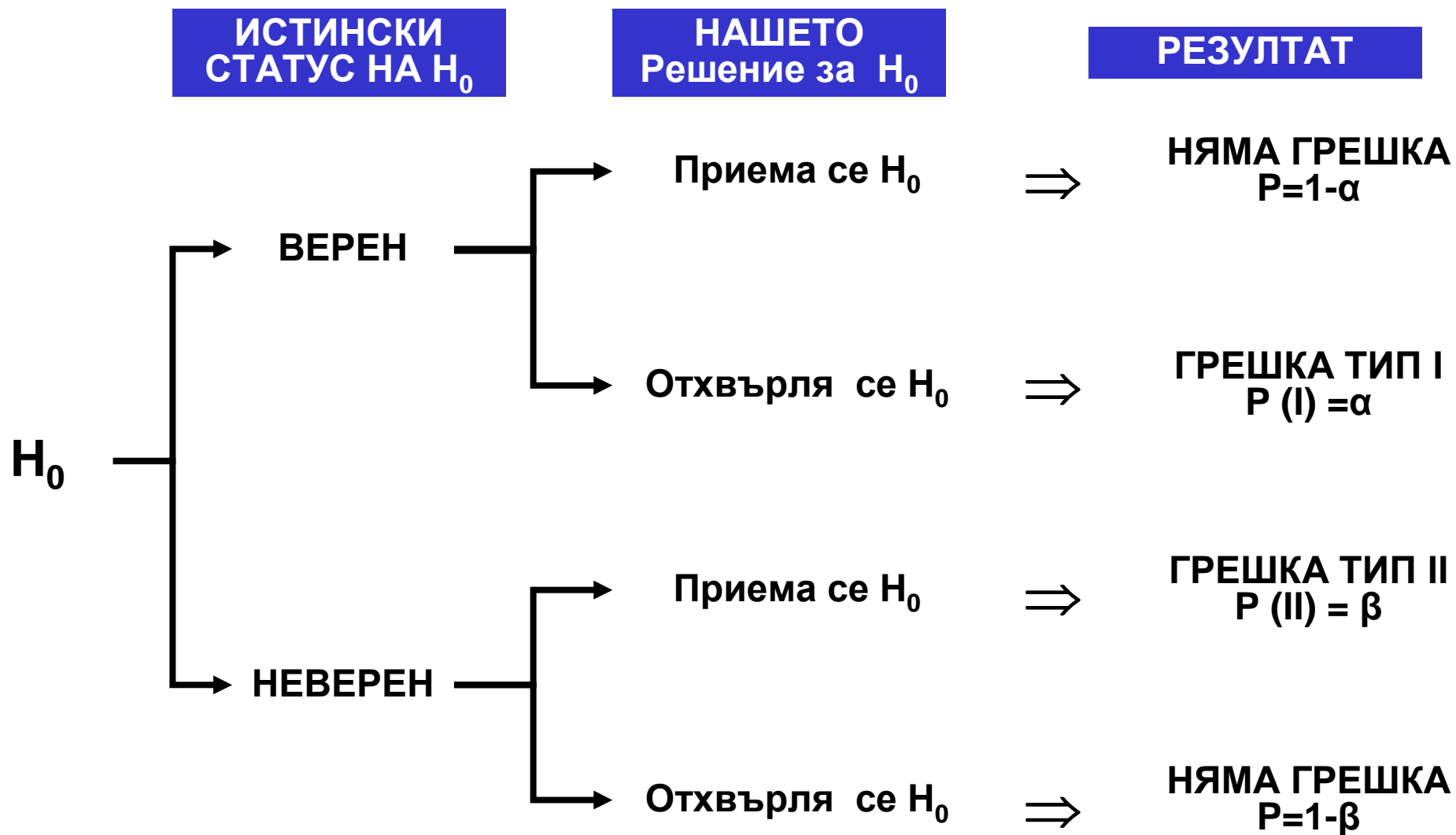
FN област грешка от I род  $\alpha$



приемаме



	ВЯРНА Е $H_0$	$H_0$ не е вярна
$H_0$ $X \in N_0$ - позитив	TP вярно приемане $1-\alpha$	FN грешно отриц. $\alpha$
$H_1$ $X \in N_1$ - негатив	FP грешно приемане $\beta$	TN вярно отхвърляне $1-\beta$





🔗 **ХИПОТЕЗА (hypothesis)** - научно обосновано предположение обясняващо наблюдавани явления.

🔗 **СТАТИСТИЧЕСКА ХИПОТЕЗА** - твърдение за параметър(и) на разпределение на СВ (параметрични) или вида на това разпределение (непараметрични).

**Проста** - изказва твърдение за всички параметри на разпределението

**Сложна** - твърдения при които някой параметрите е произволен

Насочени ( $H_0 \mu < \mu_0$ ) ( $H_0 \mu > \mu_0$ ) → едностранен критерий

Ненасочени ( $H_0 \mu = \mu_0$ ) ( $H_0 \mu \neq \mu_0$ ) → двустранен критерий

🔗 **НУЛЕВА ХИПОТЕЗА (null hypothesis)** - изказаната хипотеза  $H_0$ .

🔗 **АЛТЕРНАТИВНА ХИПОТЕЗА (alternative)**

- несъвместимата с нулевата (обратната) хипотеза  $H_1$ .

Нулевата хипотеза може да има голям брой алтернативни хипотези

ФИЛОСОФИЯ НА ВЗЕМАНЕ НА РЕШЕНИЯ -  
вероятностен смисъл - оценка на СЪБИТИЕ

статистическа сигурност  **$P = 0.99$**

ниво на значимост -  **$\alpha = 1 - P$**



# КРИТЕРИИ ЗА ОЦЕНКА

ИЗПОЛЗВА СЕ ЛОГИКАТА ЗА ХАРАКТЕРИЗИРАНЕ НА СЪБИТИЯ И ТЯХНАТА ВЕРОЯТНОСТ ДА СЕ РЕАЛИЗИРАТ

Ако вероятността средната стойност да попада в интервал съдържащ дадена стойност  $A = \mu_0$  е малка – например 0,01 то **със статистическа сигурност  $P$  99% може да се предположи**, че СВ има значения различни от  $A$

и обратно,

ако вероятността да настъпи събитието -  $X$  заема стойности (елементарни изходи) с вероятност 0,95 между които е и  $A$  то **може да се приеме, с 95% вероятност**, че СВ съвпада с  $A$



# АЛГОРИТЪМ ЗА ПРОВЕРКА НА ХИПОТЕЗИ (1)

❶ **ДАДЕНО ПО УСЛОВИЕ** - известни ли са параметрите на СВ(и)  $\mu$  и  $\sigma$  или на техните статистически оценки: **средно аритметично**  $\bar{X}$ , **стандартно отклонение**  $S$ , както и брой измервания -  $N$  и съответните **степени на свобода**  $f$  ?

❷ **ТЪРСИ СЕ !** Дефиниране на нулевата хипотеза  $H_0$  и на алтернативната(и)  $H_1$  (насочени или ненасочени)

❸ **ИЗБОР НА СТАТИСТИКА - С КОЯ ТАБЛИЦА ЩЕ РАБОТИМ !**

Преценка кой критерий следва да бъде използван и какво е неговото разпределение !!! до колко познаваме СВ знаем ли  $\sigma$  ( за две - еднакви ли са или различни)

❹ **ИЗЧИСЛЕНИЕ** изчисляване на критерия  $X_{кр}$  чрез подходящата статистика на база данните от извадката

# АЛГОРИТЪМ ЗА ПРОВЕРКА НА ХИПОТЕЗИ (2)

⑤ **ПРЕЦЕНКА НА РИСКА** избор на ниво на значимост  $\alpha = 1 - P$  (размер на критерия, грешка от първи род) и съответната доверителна вероятност  $P = 1 - \alpha$  = статистическа сигурност

⑥ **ЛОГИКА** Намиране на  $X_{\text{tab}}$  като граница на интеграла от разпределението на критерия при избраното  $\alpha$ - внимание преценка за лява, дясна или двустранна граница ( $P=1- \alpha/2$ )

⑦ **ОЦЕНКА** Сравняване на  $X_{\text{кр}}$  и  $X_{\text{таб}}$

⑧ **РЕШЕНИЕ** Приемане или отхвърляне на хипотезата със съответната статистическа сигурност  $P$

# ПРОВЕРКА НА ХИПОТЕЗА ЗА РАВЕНСТВО НА МАТЕМАТИЧЕСКОТО ОЧАКВАНЕ НА СВ НА ДАДЕНА СТОЙНОСТ

Предполага се, че СВ и съответно нейните наблюдения са нормално разпределени

Хипотезите, които могат да се формулират са:

- 1) Нулева хипотеза –  $H_0 : \mu = \mu_0$  и алтернативна  $H_1 : \mu \neq \mu_0$  **ненасочена**
- 2) Нулева хипотеза –  $H_0 : \mu \leq \mu_0$  и алтернативна  $H_1 : \mu > \mu_0$  **насочена**
- 3) Нулева хипотеза –  $H_0 : \mu \geq \mu_0$  и алтернативна  $H_1 : \mu < \mu_0$  **насочена**

където  $\mu$  е математическото очакване, а  $\mu_0$  е стойността, с която то се сравнява.

Ако са известни  $\mu$  и  $\sigma^2$  - се работи с интеграла на Лаплас използвайки

$$Z \text{ трансформация } Z = (X - \mu_0) / \sigma$$

който се явява критерий за проверка на хипотезите чрез изчислението:

**Ами ако  $\mu$  не е известно ?**

$$Z_{кр} = \frac{(\mu - \mu_0)}{\sigma}$$



# ПРОВЕРКА НА ХИПОТЕЗАТА ЗА РАВЕНСТВО НА МАТЕМАТИЧЕСКОТО ОЧАКВАНЕ НА СЛУЧАЙНА ВЕЛИЧИНА НА ДАДЕНА СТОЙНОСТ

Ако  $\mu$  **НЕ** е известно – извършват се  $N$  наблюдения над случайната величина:  $X_1, X_2, \dots, X_N$  и се изчислява  $X_{\text{ср}}$

При **ИЗВЕСТНА** дисперсия критерият за проверка на хипотезата е следния:

$$Z_{kp} = \frac{(\bar{X} - \mu_0) \cdot \sqrt{N}}{\sigma}$$

където  $\sigma^2$  е дисперсията на случайната величина.

## ХИПОТЕЗА ЗА РАВЕНСТВО НА МАТЕМАТИЧЕСКОТО ОЧАКВАНЕ НА СВ НА ДАДЕНА СТОЙНОСТ

В зависимост от формулираната нулева хипотеза се процедира по различен начин

ако - Нулева хипотеза –  $H_0: \mu = \mu_0$  и алтернативна  $H_1: \mu \neq \mu_0$

Изчислява се средната стойност от наблюденията по формулата :

$$\bar{X} = \frac{\sum_{i=1}^N x_i}{N}$$

Пресмята се величината на критерия  $Z_{кр}$

$$Z_{кр} = \frac{(\bar{X} - \mu_0) \cdot \sqrt{N}}{\sigma}$$

**Избира** се ниво на значимост на нулевата хипотеза -  $\alpha$ .

От приложение 1 се намира стойността  $X_{табл}$ , за която интеграла на Лаплас в границите от  $-Z_{табл}$  до  $+Z_{табл}$  има решение  $1-\alpha$ . Понеже повечето таблици са за интеграл на Лаплас от  $-\infty$  до  $X$  е удобно  $Z_{табл}$  да се търси като решение на уравнението

$$F(X) = 1 - \frac{\alpha}{2}$$

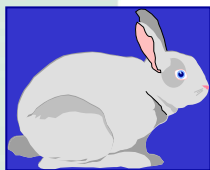
където  $F(Z)$  е интеграла на Лаплас от  $-\infty$  до  $X=Z$ .

При  $|Z_{кр}| \leq |Z_{табл}|$ , хипотезата **СЕ ПРИЕМА**

Ако  $|Z_{кр}| > |Z_{табл}|$ , хипотезата **НЕ СЕ ПРИЕМА**

**Ами ако  $\sigma^2$  не ни е  
известно ?**

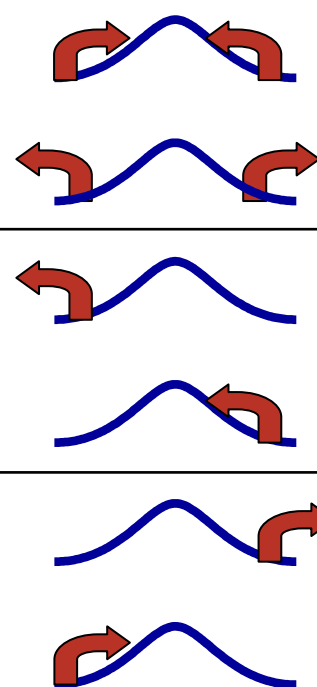
# ПРОВЕРКА НА ХИПОТЕЗИ



Ненасочени ( $H_0 \mu = \mu_0$ ) ( $H_0 \mu \neq \mu_0$ ) → **ДВУСТРАННА** постановка  
 Насочени ( $H_0 \mu < \mu_0$ ) ( $H_0 \mu > \mu_0$ ) → **ЕДНОСТРАННА** постановка

$$Z_{кр} = \frac{(\bar{X} - \mu_0) \cdot \sqrt{N}}{\sigma}$$

ХИПОТЕЗИ	критерий	$H_0$ СЕ ПРИЕМА ако
$H_0 \mu = \mu_0$	<u>двустранен</u>	$ Z_{кр}  \leq Z_{(1-1/2\alpha)}$
$H_0 \mu \neq \mu_0$	<u>двустранен</u>	$ Z_{кр}  > Z_{(1-1/2\alpha)}$
$H_0 \mu < \mu_0$	<u>едностранен</u>	$Z_{кр} \leq -Z_{(1-\alpha)}$
$H_0 \mu \leq \mu_0$	<u>едностранен</u>	$Z_{кр} < Z_{(1-\alpha)}$
$H_0 \mu > \mu_0$	<u>едностранен</u>	$Z_{кр} \geq Z_{(1-\alpha)}$
$H_0 \mu \geq \mu_0$	<u>едностранен</u>	$Z_{кр} > -Z_{(1-\alpha)}$



# Днес разгледахме:

- 1. Значещи цифри – правилно представяне на експериментални данни – правило  $\frac{1}{4} S$**
- 2. ИНТЕРВАЛНИ ОЦЕНКИ - статистическа сигурност и ниво на значимост**
  - **ДОВЕРИТЕЛЕН ИНТЕРВАЛ НА СВ** при ограничен брой измервания
  - Методи за изчисляване на доверителния интервал
- 2. СТАТИСТИЧЕСКИ ХИПОТЕЗИ**
  - Основни понятия и същност на проверката на СХ.
  - Хипотеза за равенство на  $\mu$  на дадена стойност

# ВНИМАНИЕ – КОЛОКВИУМ

**На 9 АПРИЛ 2009  
(четвъртък)**

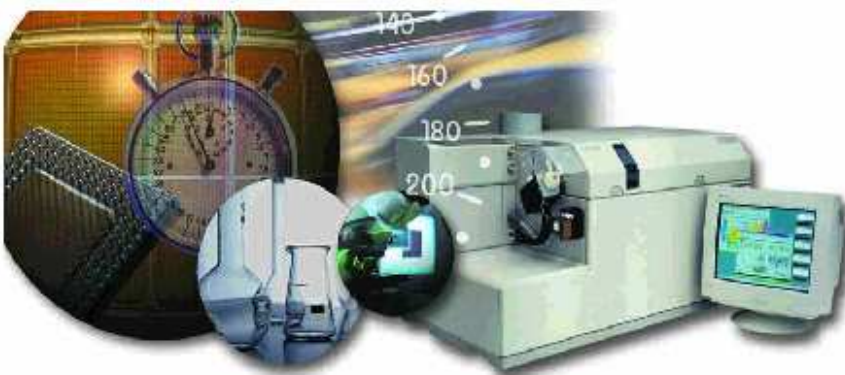
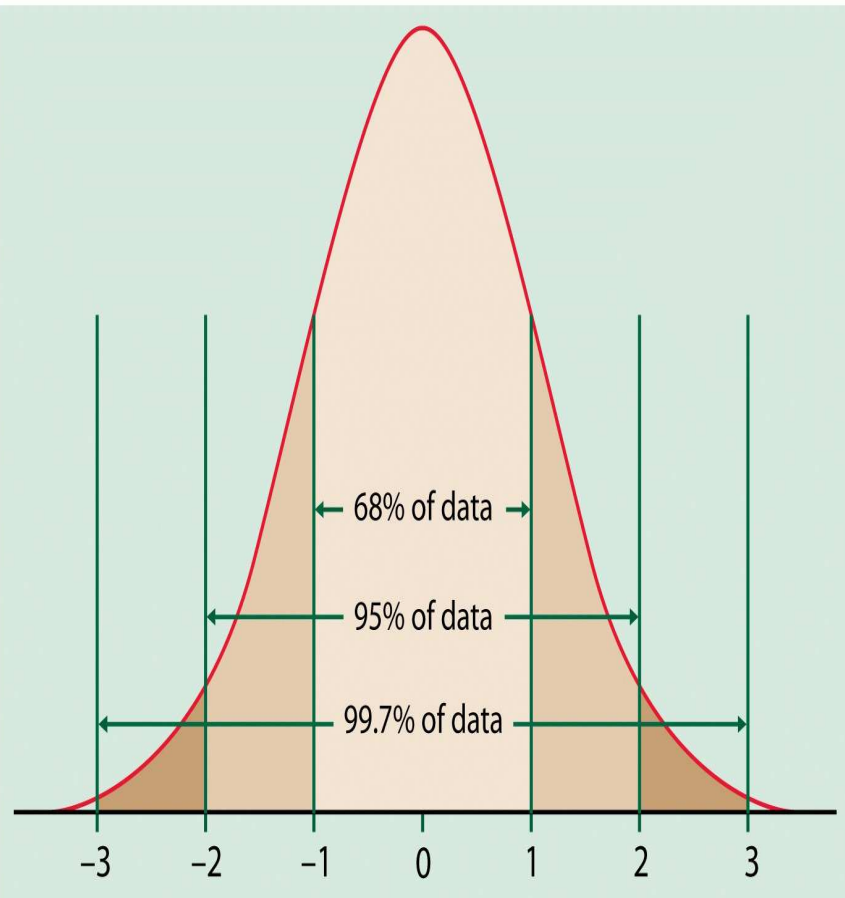


лекция № 8

# Статистика на набор от измервания

# СТАТИСТИЧЕСКИ ХИПОТЕЗИ

продължение



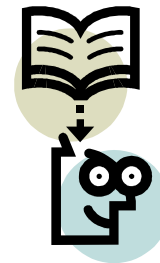
# Днес ще разгледаме:

## 1. СТАТИСТИЧЕСКИ ХИПОТЕЗИ

- Основни понятия и същност на проверката на **СХ**.
- **СХ** за равенство на  $X$  на дадена стойност при известно и **неизвестно**  $\sigma^2$ .
- **СХ** за равенство на две математически очаквания – при известни дисперсии
- **СХ** за равенство на две дисперсии - F - критерий
- Хипотези за равенство на две математически очаквания
  - при **известни** дисперсии
  - при **неизвестни но равни** дисперсии
  - при **неизвестни и неравни** дисперсии.

## 2. ОЦЕНКА НА СТАТИСТИЧЕСКИ ХИПОТЕЗИ ЧРЕЗ EXCEL

# ВАЖНО ДА СИ ПРИПОМНИМ !!



- ☒ Величината  $z = \frac{(X - \mu)}{\sigma}$  е стандартно разпределена величина **N(0, 1)**
- ☒ Величината  $\bar{X}$  е нормално разпределена СВ - **N( $\mu$ ,  $\sigma^2/N$ )**
- ☒ Величината  $\frac{(N-1) \cdot S^2}{\sigma^2}$  е  $\chi^2$ - разпределена със степени на свобода **f = N-1**
- ☒ Величината  $(\bar{X} - \mu)$  е нормално разпределена СВ - **N(0,  $\sigma^2/N$ )**, а  $\frac{(\bar{X} - \mu)}{\sigma/\sqrt{N}}$  е СТ.Р - **N(0, 1)**
- ☒ Величината  $\frac{(\bar{X} - \mu) \cdot \sqrt{N}}{S}$  е t- разпределена СВ със степени на свобода **f = N-1**
- ☒ Величината  $S_1^2 / S_2^2$  е F-разпределена случайна величина с **f<sub>1</sub>=N-1** и **f<sub>2</sub>= M-1**

$$P = 1 - \alpha$$

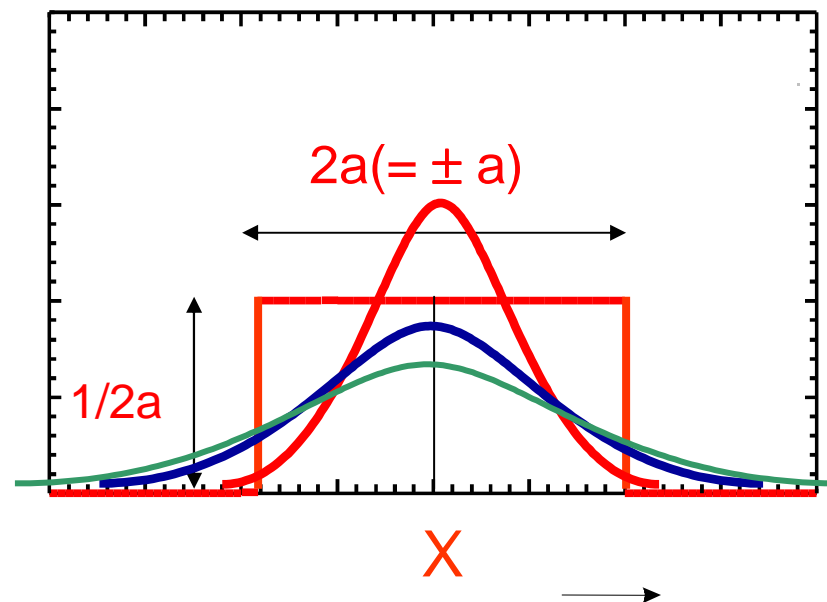
$$P(X \leq a) = F(z_a)$$

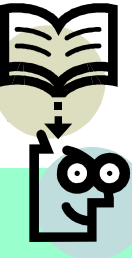
$$P(X > a) = 1 - F(z_a)$$

$$P(a < X < b) = F(z_b) - F(z_a)$$

$$F(-a) = 1 - F(z_a)$$

$$P(\mu - a < X < \mu + a) = 2F(z_a) - 1$$





**ХИПОТЕЗА (hypothesis)** - научно обосновано предположение обясняващо наблюдавани явления.

**СТАТИСТИЧЕСКА ХИПОТЕЗА** - твърдение за параметър(и) на разпределение на СВ (параметрични) или вида на това разпределение (непараметрични).

**Проста** - изказва твърдение за всички параметри на разпределението

**Сложна** - твърдения при които някои параметрите е произволен

Насочени ( $H_0 \mu < \mu_0$ ) ( $H_0 \mu > \mu_0$ ) → едностранен критерий

Ненасочени ( $H_0 \mu = \mu_0$ ) ( $H_0 \mu \neq \mu_0$ ) → двустранен критерий

**НУЛЕВА ХИПОТЕЗА (null hypothesis)** - изказаната хипотеза  $H_0$ .

**АЛТЕРНАТИВНА ХИПОТЕЗА (alternative)**

- несъвместимата с нулевата (обратната) хипотеза  $H_1$ .

Нулевата хипотеза може да има голям брой алтернативни хипотез

ФИЛОСОФИЯ НА ВЗЕМАНЕ НА РЕШЕНИЯ -  
вероятностен смисъл - оценка на СЪБИТИЕ

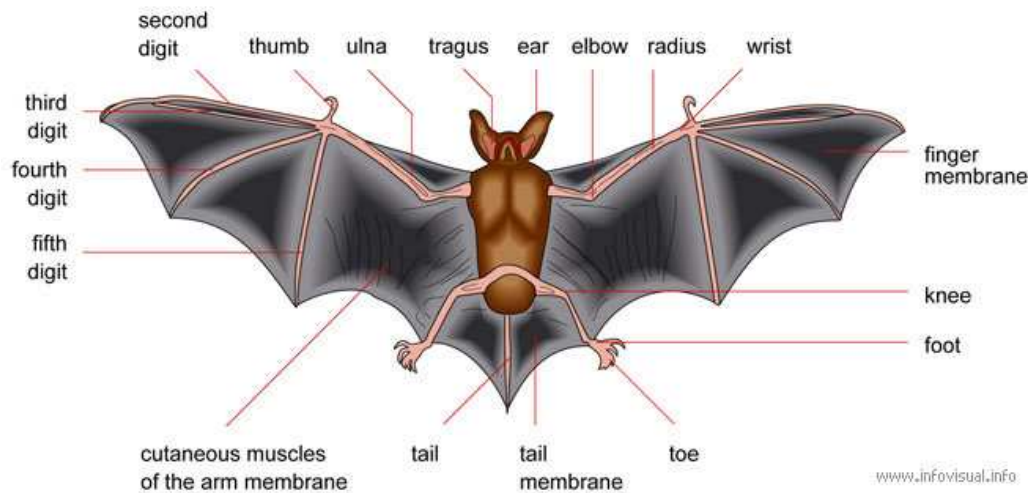
статистическа сигурност  $P = 0.99$

ниво на значимост  $\alpha = 1 - P$

Прогностични  
методи в  
астрологията



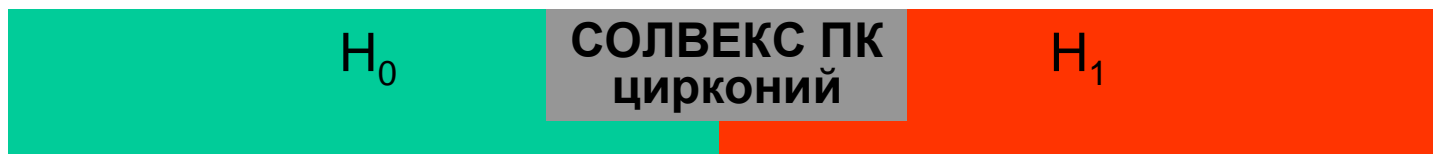
### MORPHOLOGY OF A BAT



www.infovisual.info



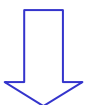
# ХИПОТЕЗИ - грешка от първи и от втори род



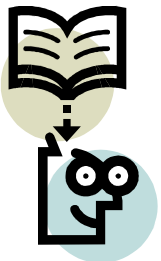
FP област грешка от II род  $\beta$

FN област грешка от I род  $\alpha$

приемаме



	ВЯРНА Е $H_0$	$H_0$ не е вярна
$H_0$ $X \in N_0$ - позитив	TP вярно приемане $1-\alpha$	FN грешно отриц. $\alpha$
$H_1$ $X \in N_1$ - негатив	FP грешно приемане $\beta$	TN вярно отхвърляне $1-\beta$



# АЛГОРИТЪМ ЗА ПРОВЕРКА НА ХИПОТЕЗИ (1)

**❶ ДАДЕНО ПО УСЛОВИЕ** - известни ли са параметрите на СВ(и)  $\mu$  и  $\sigma$  или на техните статистически оценки: **средно аритметично**  $\bar{X}$ , **стандартно отклонение**  $S$ , както и брой измервания -  $N$  и съответните **степени на свобода**  $f$  ?

**❷ ТЪРСИ СЕ !** Дефиниране на **нулевата хипотеза**  $H_0$  и на **алтернативната(и)**  $H_1$  (насочени или ненасочени)

**❸ ИЗБОР НА СТАТИСТИКА - С КОЯ ТАБЛИЦА ЩЕ РАБОТИМ !**

Преценка кой критерий следва да бъде използван и какво е неговото разпределение !!! до колко познаваме СВ знаем ли  $\sigma$  ( за две - еднакви ли са или различни)

**❹ ИЗЧИСЛЕНИЕ** изчисляване на критерия  $X_{кр}$  чрез подходящата статистика на база данните от извадката



# АЛГОРИТЪМ ЗА ПРОВЕРКА НА ХИПОТЕЗИ (2)

⑤ **ПРЕЦЕНКА НА РИСКА** избор на ниво на значимост  $\alpha = 1 - P$  (размер на критерия, грешка от първи род) и съответната доверителна вероятност  $P = 1 - \alpha$  = статистическа сигурност

⑥ **ЛОГИКА** Намиране на  $X_{tab}$  като граница на интеграла (квантил) от разпределението на критерия при избраното  $\alpha$ - внимание преценка за лява, дясна или двустранна граница (критична стойност) ограничаваща съответно площ  $P=1- \alpha/2$

⑦ **ОЦЕНКА** Сравняване на  $X_{кр}$  и  $X_{таб}$

⑧ **РЕШЕНИЕ** Приемане или отхвърляне на хипотезата със съответната статистическа сигурност  $P$



# АЛГОРИТЪМ ЗА ПРОВЕРКА НА ХИПОТЕЗИ

- ➊ **ДАДЕНО ПО УСЛОВИЕ** - известни ли са параметрите на СВ(и)  $\mu$  и  $\sigma$  или на техните статистически оценки: средно аритметично  $\bar{X}$ , стандартно отклонение  $S$ , както и брой измервания -  $N$  и съответните степени на свобода  $f$  ?
- ➋ **ТЪРСИ СЕ !** Дефиниране на нулевата хипотеза  $H_0$  и на алтернативната(и)  $H_1$  (насочени или ненасочени)
- ➌ **С КОЯ ТАБЛИЦА ЩЕ РАБОТИМ !** Преценка кой критерий следва да бъде използван и какво е неговото разпределение !!! до колко познаваме СВ знаем ли  $\sigma$  ( за две - еднакви ли са или различни)
- ➍ **ИЗЧИСЛЕНИЕ** изчисляване на критерия  $X_{кр}$
- ➎ **ПРЕЦЕНКА НА РИСКА** избор на ниво на значимост  $\alpha = 1 - P$  (размер на критерия, грешка от първи род) и съответната доверителна вероятност  $P = 1 - \alpha$  = статистическа сигурност
- ➏ **ЛОГИКА** Намиране на  $X_{таб}$  като граница на интеграла от разпределението на критерия при избраното  $\alpha$ - внимание преценка за лява, дясна или двустранна граница ( $P=1- \alpha/2$ )
- ➐ **ОЦЕНКА** Сравняване на  $X_{кр}$  и  $X_{таб}$
- ➑ **РЕШЕНИЕ** Приемане или отхвърляне на хипотезата със съответната статистическа сигурност  $P$

## ПРОВЕРКА НА ХИПОТЕЗАТА ЗА РАВЕНСТВО НА МАТЕМАТИЧЕСКОТО ОЧАКВАНЕ НА СЛУЧАЙНА ВЕЛИЧИНА НА ДАДЕНА СТОЙНОСТ

Ако - Нулева хипотеза е  $H_0 : \mu = \mu_0$  и алтернативна  $H_1 : \mu \neq \mu_0$

(ако  $\mu$  е неизвестно, но знаем  $\sigma$ )

Изчислява се средната стойност от наблюденията по формулата :

$$\bar{X} = \frac{\sum_{i=1}^N x_i}{N}$$

Пресмята се критерия  $Z_{кр}$  - 
$$Z_{кр} = \frac{(\bar{X} - \mu_0) \cdot \sqrt{N}}{\sigma}$$

Избира се ниво на значимост на нулевата хипотеза -  $\alpha$ .

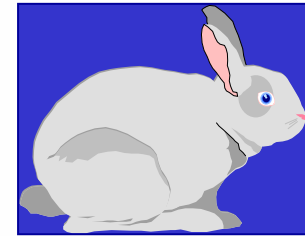
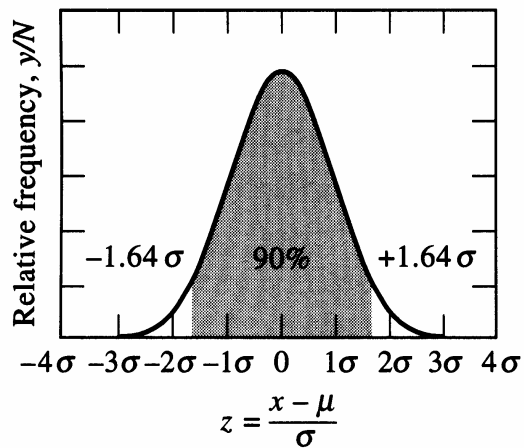
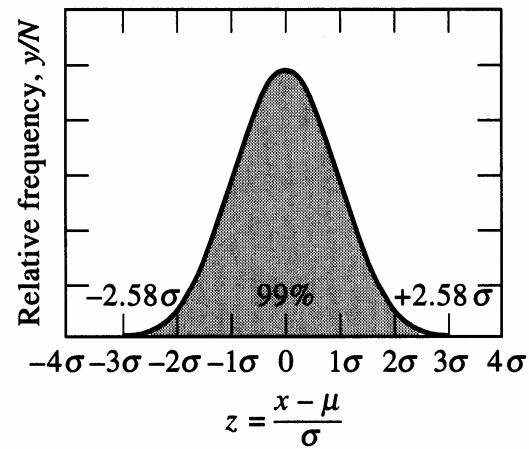
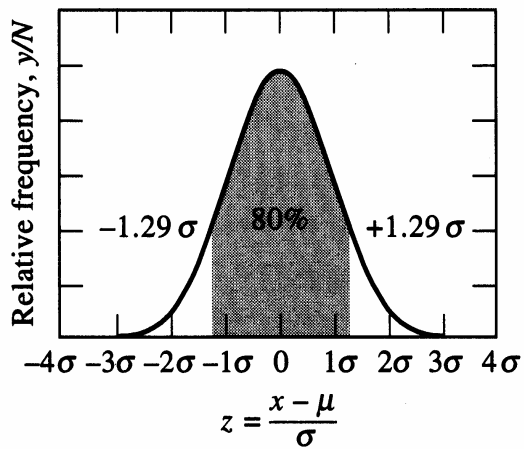
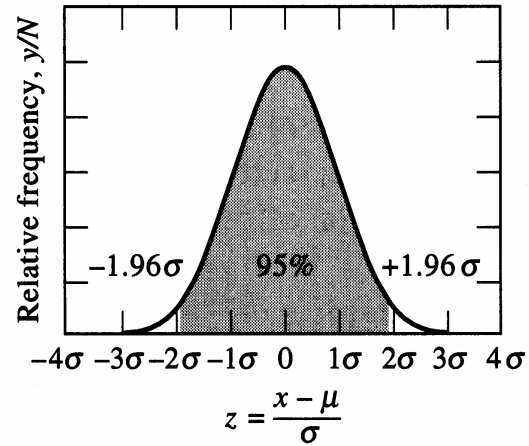
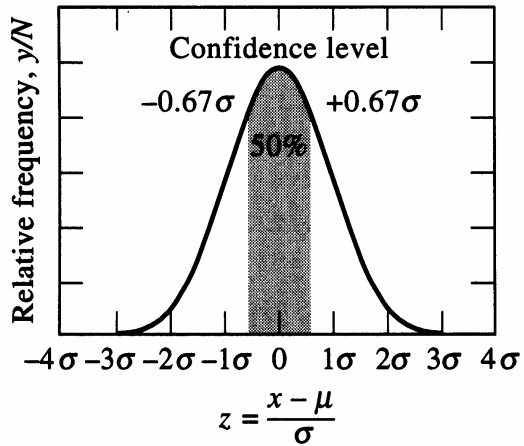
От приложение 1 се намира стойността  $Z_{табл}$ , за която интеграла на Лаплас в границите от  $-Z_{табл}$  до  $+Z_{табл}$  има решение  $1 - \alpha$ . Понеже повечето таблици са за интеграл на Лаплас от  $-\infty$  до  $X$  е удобно  $Z_{табл}$  да се търси като решение на уравнението

$$F(X) = 1 - \frac{\alpha}{2}$$

където  $F(Z)$  е интеграла на Лаплас от  $-\infty$  до  $X=Z$ .

При  $|Z_{кр}| \leq Z_{табл}$ , хипотезата **СЕ ПРИЕМА**

Ако  $|Z_{кр}| > Z_{табл}$ , хипотезата **НЕ СЕ ПРИЕМА**



$$Z_{kp} = \frac{(\bar{X} - \mu_0) \cdot \sqrt{N}}{\sigma}$$

**Figure a1-5** Confidence levels for various values of  $z$ . The shaded areas under the curves represent the probability that a given measurement lies within  $\pm z$  of the mean.

# ПРОВЕРКА НА ХИПОТЕЗАТА ЗА РАВЕНСТВО НА МАТЕМАТИЧЕСКОТО ОЧАКВАНЕ НА СЛУЧАЙНА ВЕЛИЧИНА НА ДАДЕНА СТОЙНОСТ

Ако - Нулева хипотеза –  $H_0 : \mu = \mu_0$  и алтернативна  $H_1 : \mu \neq \mu_0$   
(ако  $\mu$  е неизвестно)

Изчислява се средната стойност от наблюденията по формулата :  $\bar{X} = \frac{\sum_{i=1}^N x_i}{N}$

Пресмята се критерия  $Z_{кр}$  -  ~~$Z_{кр} = \frac{(\bar{X} - \mu_0) \cdot \sqrt{N}}{\sigma}$~~

$$t_{кр} = \frac{(\bar{X} - \mu_0) \cdot \sqrt{N}}{S}$$

Избира се ниво на значимост на нулевата хипотеза -  $\alpha$ .

От приложение 1 се намира стойността  $Z_{табл}$ , за която интеграла на Лаплас в границите от  $-Z_{табл}$  до  $+Z_{табл}$  има решение  $1 - \alpha$ . Понеже повечето таблици са за интеграл на Лаплас от  $-\infty$  до  $Z$  е удобно  $Z_{табл}$  да се търси като решение на уравнението

$$F(X) = 1 - \frac{\alpha}{2}$$

където  $F(Z)$  е интеграла на Лаплас от  $-\infty$  до  $X=Z$ .

При  $|t_{кр}| \leq t_{табл}(\alpha, f)$ , хипотезата **СЕ ПРИЕМА**

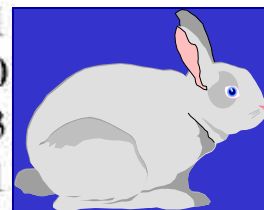
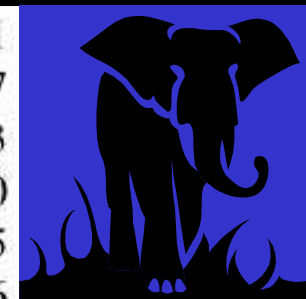
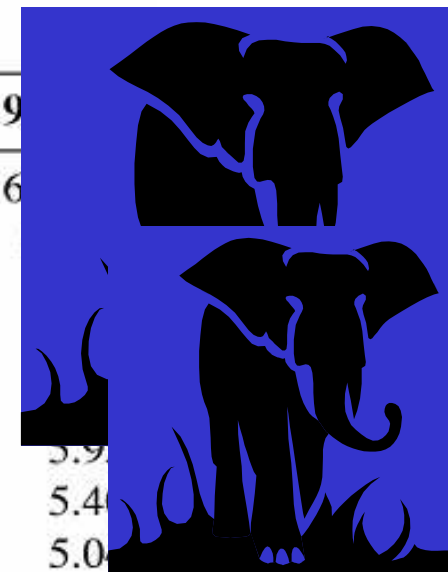
Ако  $|t_{кр}| > t_{табл}(\alpha, f)$ , хипотезата **НЕ СЕ ПРИЕМА**

**Ами ако и  $\sigma$  е  
неизвестно ?**

# Таблица на t-разпределение

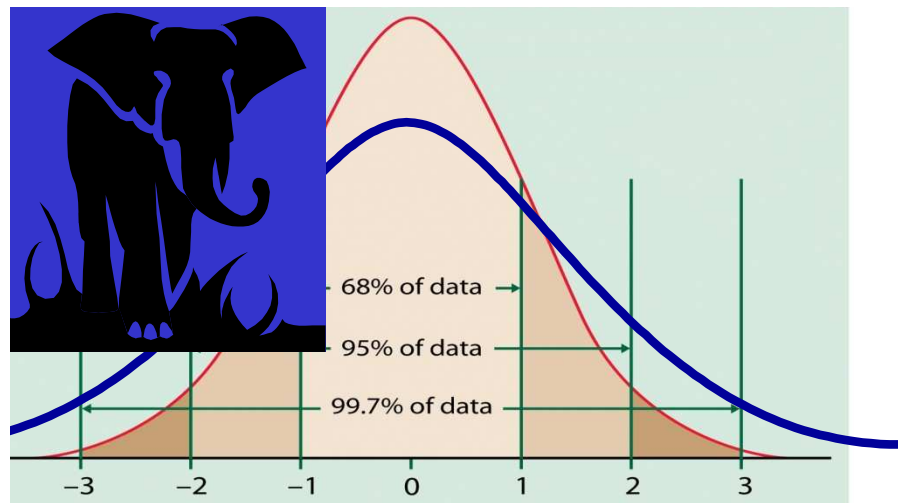
Table 4-2 Values of Student's *t*

Degrees of freedom	Confidence level (%)						
	50	90	95	98	99	99.5	99.9
1	1.000	6.314	12.706	31.821	63.657	127.32	631.97
2	0.816	2.920	4.303	6.965	9.925	14.089	31.821
3	0.765	2.353	3.182	4.541	5.841	7.453	12.941
4	0.741	2.132	2.776	3.747	4.604	5.598	11.709
5	0.727	2.015	2.571	3.365	4.032	4.773	10.811
6	0.718	1.943	2.447	3.143	3.707	4.317	9.915
7	0.711	1.895	2.365	2.998	3.500	4.029	9.147
8	0.706	1.860	2.306	2.896	3.355	3.832	8.451
9	0.703	1.833	2.262	2.821	3.250	3.690	7.877
10	0.700	1.812	2.228	2.764	3.169	3.581	7.367
15	0.691	1.753	2.131	2.602	2.947	3.252	6.314
20	0.687	1.725	2.086	2.528	2.845	3.153	5.478
25	0.684	1.708	2.060	2.485	2.787	3.078	4.949
30	0.683	1.697	2.042	2.457	2.750	3.030	4.571
40	0.681	1.684	2.021	2.423	2.704	2.971	4.191
60	0.679	1.671	2.000	2.390	2.660	2.915	3.851
120	0.677	1.658	1.980	2.358	2.617	2.860	3.519
∞	0.674	1.645	1.960	2.326	2.576	2.807	3.291

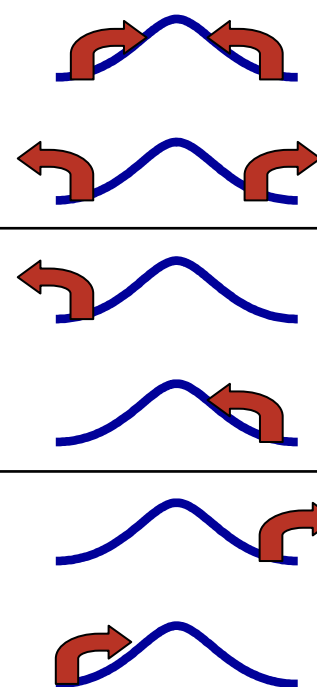


# ПРОВЕРКА НА ХИПОТЕЗИ

$$t_{кр} = \frac{(\bar{X} - \mu_0) \cdot \sqrt{N}}{S}$$



ХИПОТЕЗИ	критерий	$H_0$ СЕ ПРИЕМА ако
$H_0 \mu = \mu_0$	<u>дву</u> странен	$ t_{кр}  \leq t_{(1-\frac{1}{2}\alpha)}$
$H_0 \mu \neq \mu_0$	<u>дву</u> странен	$ t_{кр}  > t_{(1-\frac{1}{2}\alpha)}$
$H_0 \mu < \mu_0$	<u>едно</u> странен	$t_{кр} \leq -t_{(1-\alpha)}$
$H_0 \mu \leq \mu_0$	<u>едно</u> странен	$t_{кр} < t_{(1-\frac{1}{2}\alpha)}$
$H_0 \mu > \mu_0$	<u>едно</u> странен	$t_{кр} \geq t_{(1-\frac{1}{2}\alpha)}$
$H_0 \mu \geq \mu_0$	<u>едно</u> странен	$t_{кр} > -t_{(1-\frac{1}{2}\alpha)}$



# СХ за равенство на две математически очаквания при ИЗВЕСТНИ дисперсии

Нека  $X_1, X_2, X_3, \dots, X_{N_1}$  и  $Y_1, Y_2, Y_3, \dots, Y_{N_2}$  са две извадки, където  $N_1$  и  $N_2$  са съответно техните обеми

Нека и двете случайни величини са разпределени нормално.

Ако с  $\mu_1$  и  $\mu_2$  се означат съответните математически очаквания то могат да се формулират следните хипотези:

1) Нулева хипотеза  $H_0: \mu_1 = \mu_2$   
и алтернативна  $H_a: \mu_1 \neq \mu_2$

2) Нулева хипотеза  $H_0: \mu_1 \leq \mu_2$   
и алтернативна  $H_1: \mu_1 > \mu_2$

3) Нулева хипотеза  $H_0: \mu_1 \geq \mu_2$   
и алтернативна  $H_1: \mu_1 < \mu_2$

$H_0$  се приема със статистическа сигурност  $P$  ако:

$|X_{кр}| \leq X_{таб}$  за който  $F(X_{таб}) = 1 - \frac{1}{2} \alpha$

$X_{кр} \leq X_{таб}$  за който  $F(X_{таб}) = 1 - \alpha$

$X_{кр} \geq -X_{таб}$  за който  $F(X_{таб}) = 1 - \alpha$

**АКО  $\sigma_1^2$  и  $\sigma_2^2$  са ИЗВЕСТНИ !!**

Избира се  $\alpha$  и изчислява  $P = 1 - \alpha$  от таблицата на ЛАПЛАС се намира квантила  $X_{таб}$  за който  $F(X_{таб}) = P$

$$X_{кр} = \frac{(\bar{X} - \bar{Y})}{\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}}$$

# Хипотеза за равенството на две дисперсии

Ако две дисперсии са оценени чрез техните статистически оценки  $S_1^2$  и  $S_2^2$  получени от извадки с обем съответно  $N_1$  и  $N_2$  то може да се повери хипотезата за равенство на двете дисперсии:

$H_0$  се приема със статистическа сигурност  $P$   
ако:

Нулева хипотеза  $H_0: \sigma_1 = \sigma_2$   
и алтернативна  $H_a: \sigma_1 \neq \sigma_2$

$$1/F(f_2, f_1, \alpha) < F_{kr} < F(f_1, f_2, \alpha)$$

Изчислява се статистиката :

Избира се  $\alpha$  и изчислява  $P=1-\alpha$  от таблица за разпределение на ФИШЕР се намира квантила

$$F(f_1, f_2, \alpha)$$

където  $f_1 = N_1 - 1$  и  $f_2 = N_2 - 1$

$$F_{kr} = \frac{S_1^2}{S_2^2}$$



# СХ за равенство на две математически очаквания при НЕизвестни но РАВНИ дисперсии

Нека  $X_1, X_2, X_3, \dots, X_{N_1}$  и  $Y_1, Y_2, Y_3, \dots, Y_{N_2}$  са две извадки, където  $N_1$  и  $N_2$  са ..... да се формулират следните хипотези:

$H_0$  се приема със статистическа сигурност  $P$   
ако:

1) Нулева хипотеза  $H_0: \mu_1 = \mu_2$   
и алтернативна  $H_1: \mu_1 \neq \mu_2$

$|t_{кр}| \leq t_{таб}$  за който  $t(f, \alpha) = 1 - \frac{1}{2} \alpha$

2) Нулева хипотеза  $H_0: \mu_1 \leq \mu_2$   
и алтернативна  $H_1: \mu_1 > \mu_2$

$t_{кр} \leq t_{таб}$  за който  $t(f, \alpha) = 1 - \alpha$

3) Нулева хипотеза  $H_0: \mu_1 \geq \mu_2$   
и алтернативна  $H_1: \mu_1 < \mu_2$

$t_{кр} \geq -t_{таб}$  за който  $t(f, \alpha) = 1 - \alpha$

АКО  $\sigma^2_1 = \sigma^2_2$  са  
НЕИЗВЕСТНИ – но РАВНИ !!

Избира се  $\alpha$  и изчислява  $P=1-\alpha$   
от t-таблица се намира квантила

$t_{(\alpha, f)}$

$$t_{кр} = \frac{(\bar{X} - \bar{Y}) \cdot \sqrt{N_1 + N_2 - 2}}{\sqrt{\left(\frac{1}{N_1} + \frac{1}{N_2}\right) \cdot [(N_1 - 1) \cdot S_1^2 + (N_2 - 1) \cdot S_2^2]}}$$

$f = N_1 + N_2 - 2$

# СХ за равенство на две математически очаквания при НЕизвестни и РАЗЛИЧНИ дисперсии

Нека  $X_1, X_2, X_3, \dots, X_{N_1}$  и  $Y_1, Y_2, Y_3, \dots, Y_{N_2}$  са две извадки, където  $N_1$  и  $N_2$  са ..... да се формулират следните хипотези:

$H_0$  се приема със статистическа сигурност  $P$  ако:

1) Нулева хипотеза  $H_0: \mu_1 = \mu_2$   
и алтернативна  $H_a: \mu_1 \neq \mu_2$

$|t_{кр}| \leq t_{таб}$  за който  $t(f, \alpha) = 1 - \frac{1}{2} \alpha$

2) Нулева хипотеза  $H_0: \mu_1 \leq \mu_2$   
и алтернативна  $H_1: \mu_1 > \mu_2$

$t_{кр} \leq t_{таб}$  за който  $t(f, \alpha) = 1 - \alpha$

3) Нулева хипотеза  $H_0: \mu_1 \geq \mu_2$   
и алтернативна  $H_1: \mu_1 < \mu_2$

$t_{кр} \geq -t_{таб}$  за който  $t(f, \alpha) = 1 - \alpha$

АКО  $\sigma^2_1 \neq \sigma^2_2$  са  
НЕИЗВЕСТНИ – и РАЗЛИЧНИ !!

Избира се  $\alpha$  и изчислява  $P=1-\alpha$   
от t-таблица се намира квантил

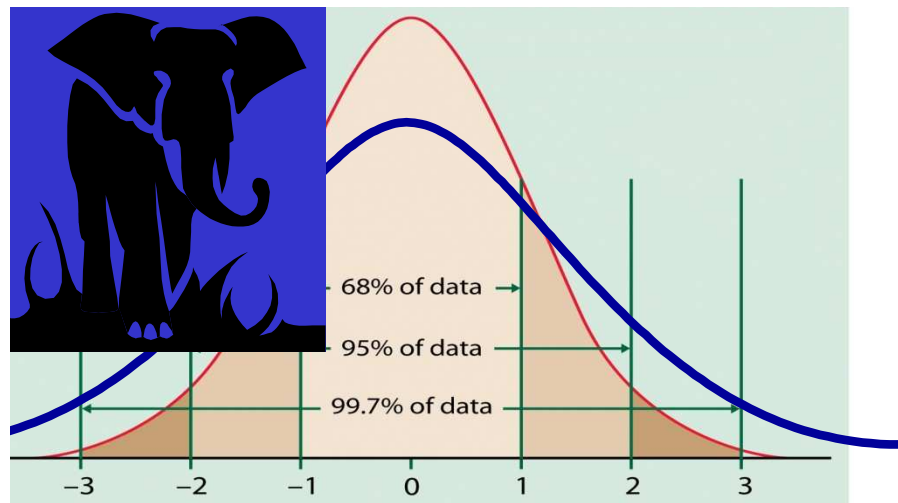
$t(\alpha, f)$

$$f = \frac{\left( \frac{S_1^2}{N_1} + \frac{S_2^2}{N_2} \right)^2}{\frac{\left( \frac{S_1^2}{N_1} \right)^2}{N_1 + 1} + \frac{\left( \frac{S_2^2}{N_2} \right)^2}{N_2 + 1}} - 2$$

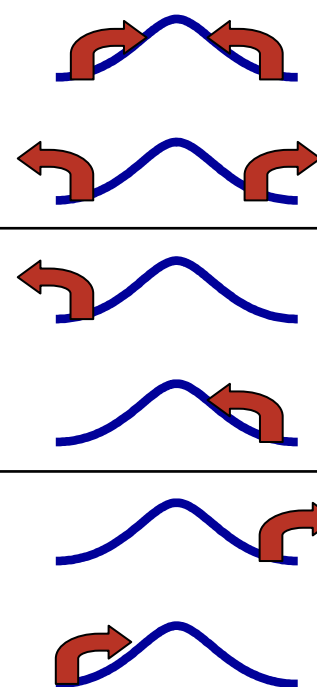
$$t_{кр} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}}$$

# ПРОВЕРКА НА ХИПОТЕЗИ

$$t_{кр} = \frac{(\bar{X} - \bar{Y})}{S / \sqrt{N}}$$



ХИПОТЕЗИ	критерий	$H_0$ СЕ ПРИЕМА ако
$H_0 \mu_1 = \mu_2$	<u>дву</u> странен	$ t_{кр}  \leq t_{(1-1/2\alpha)}$
$H_0 \mu_1 \neq \mu_2$	<u>дву</u> странен	$ t_{кр}  > t_{(1-1/2\alpha)}$
$H_0 \mu_1 < \mu_2$	<u>едно</u> странен	$t_{кр} \leq -t_{(1-\alpha)}$
$H_0 \mu_1 \leq \mu_2$	<u>едно</u> странен	$t_{кр} < t_{(1-1/2\alpha)}$
$H_0 \mu_1 > \mu_2$	<u>едно</u> странен	$t_{кр} \geq t_{(1-1/2\alpha)}$
$H_0 \mu_1 \geq \mu_2$	<u>едно</u> странен	$t_{кр} > -t_{(1-1/2\alpha)}$



ЗАДАЧА	ХИПОТЕЗИ	ПРОВЕРКА	РЕЗУЛТАТ	ЗАКЛЮЧЕНИЕ
1. <i>Хипотеза за равенството на математическото очакване на дадена стойност при неизвестна дисперсия.</i>	Нулева $H_0: \mu = \mu_0$	$t_{kr} = \frac{(\bar{X} - \mu_0)}{S} \sqrt{N}$	$ t_{kr}  \leq  t(f, \alpha) $	$H_0$ се приема
	Алтернативна $H_1: \mu \neq \mu_0$		$ t_{kr}  >  t(f, \alpha) $	$H_0$ се отхвърля
2. <i>Хипотеза за равенство на две математически очаквания при неизвестни, но <u>равни</u> дисперсии.</i>	Нулева $H_0: \mu_1 = \mu_2$	$t_{kr} = \frac{(\bar{X} - \bar{Y}) \cdot \sqrt{N_1 + N_2 - 2}}{\sqrt{\left(\frac{1}{N_1} + \frac{1}{N_2}\right) \cdot [(N_1 - 1) \cdot S_1^2 + (N_2 - 1) \cdot S_2^2]}}$	$ t_{kr}  \leq t_{tab}$	$H_0$ се приема
	Алтернативна $H_1: \mu_1 \neq \mu_2$		$f = N_1 + N_2 - 2$	
3. <i>Хипотеза за равенство на две математически очаквания при неизвестни и <u>различни</u> дисперсии.</i>	Нулева $H_0: \mu_1 = \mu_2$	$t_{kr} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}}$	$ t_{kr}  \leq t_{tab}$	$H_0$ се приема
	Алтернативна $H_1: \mu_1 \neq \mu_2$		$f = \frac{\left(\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}\right)^2}{\frac{\left(\frac{S_1^2}{N_1}\right)^2}{N_1 + 1} + \frac{\left(\frac{S_2^2}{N_2}\right)^2}{N_2 + 1}} - 2$	
4. <i>Хипотеза за равенство на две дисперсии.</i>	Нулева $H_0: \sigma_1^2 = \sigma_2^2$	$F_{kr} = \frac{S_1^2}{S_2^2}$	$\frac{1}{F(f_2, f_1, \alpha/2)} < F_{kr} < F(f_1, f_2, \alpha/2)$	$H_0$ се приема

# Днес ще разгледаме:

## 1. СТАТИСТИЧЕСКИ ХИПОТЕЗИ

- Основни понятия и същност на проверката на СХ.
- СХ за равенство на  $X$  на дадена стойност при неизвестно  $\sigma^2$ .
- СХ за равенство на две математически очаквания – при известни дисперсии
- СХ за равенство на две дисперсии - F - критерий
- Хипотези за равенство на две математически очаквания
  - при известни дисперсии
  - при неизвестни но равни дисперсии
  - при неизвестни и неравни дисперсии.

## 2. ОЦЕНКА НА СТАТИСТИЧЕСКИ ХИПОТЕЗИ ЧРЕЗ EXCEL

## ПРОВЕРКА НА ХИПОТЕЗИ чрез t критерий

Хипотеза за равенство на  $\mu$  - на дадена стойност  $\mu_0$  при вариращи  $n, \bar{X}_{av}, S, H_0:\mu, n \alpha$ .  
**ВНИМАНИЕ**- Правилно отнесете към едностранна лява или едностранна дясна или двустранна постановка (критерий) на задачата да отнесете статистиката.  
 Бележки: 1) само стойности за  $|t| < 8$  ще се изобразят на графиката  
 2) Използвайте  $n > 2$

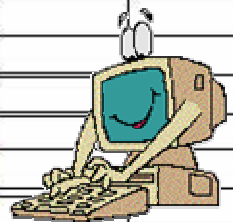
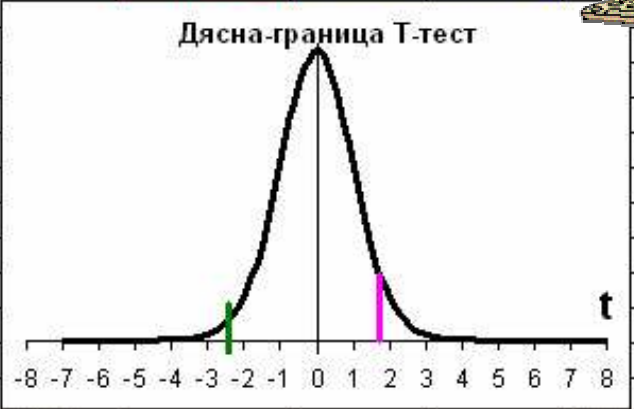
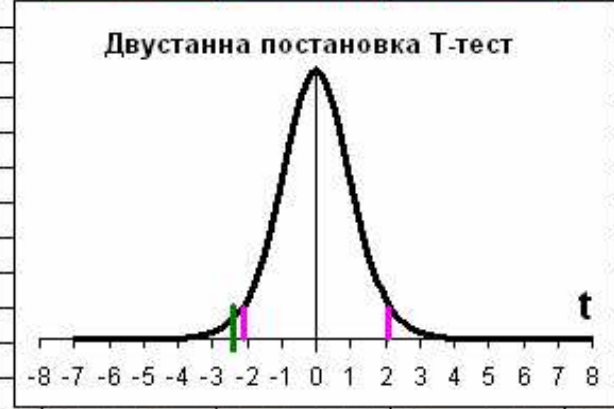
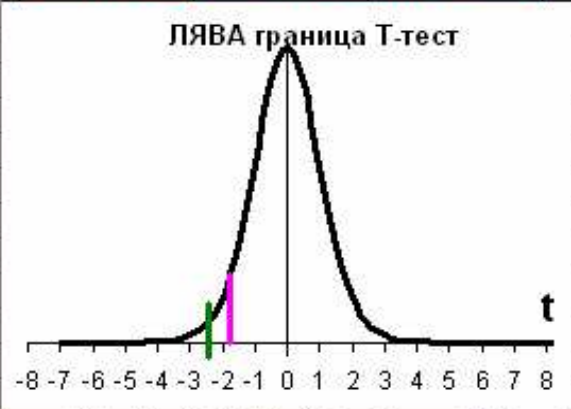
$$t_{kp} = \frac{(\bar{X} - \mu_0) \cdot \sqrt{N}}{S}$$

n	$\bar{X}$	S	P	$H_0:\mu$	$\alpha$	df	t-критерий
17	-2.5	6	95%	1	0,05	16	-2,4051

ЛЯВА ГРАНИЦА T-тест	
$H_0:\mu \geq 1$	1
$H_a:\mu < 1$	1
$-t_{\alpha, n-1}$	-1.7459
p-стойност	0,0143
$H_0:\mu \geq \mu_0$	ОТХВЪРЛЯ СЕ

ДВУСТРАННА постановка T-тест	
$H_0:\mu = 1$	1
$H_a:\mu \neq 1$	1
$-t_{\alpha/2, n-1}$	-2,1199
$+t_{\alpha/2, n-1}$	2,1199
p-стойност	0,0286
$H_0:\mu = \mu_0$	ОТХВЪРЛЯ СЕ

ДЯСНА ГРАНИЦА T-тест	
$H_0:\mu \leq 1$	1
$H_a:\mu > 1$	1
$+t_{\alpha, n-1}$	1.7459
p-стойност	0,9857
$H_0:\mu \leq \mu_0$	НЕ СЕ ОТХВЪРЛЯ



# ПРИЛОЖЕНИЯ В EXCEL

The image shows a screenshot of the Microsoft Excel application window. The title bar reads "Microsoft Excel - VK'Statistics'EDU". The menu bar includes "File", "Edit", "View", "Insert", "Format", "Tools", "Data", "Window", "Help", and "Adobe PDF". The "Tools" menu is open, displaying a list of options: Spelling..., Error Checking..., Speech, Share Workbook..., Track Changes, Compare and Merge Workbooks..., Protection, Euro Conversion..., Online Collaboration, Solver..., Goal Seek..., Scenarios..., Formula Auditing, Tools on the Web..., Macro, Add-Ins..., AutoCorrect Options..., Customize..., Options..., Conditional Sum..., Lookup..., and Data Analysis... (highlighted). The "Data Analysis" dialog box is open in the foreground, titled "Data Analysis". It contains a list of "Analysis Tools": Histogram, Moving Average, Random Number Generation, Rank and Percentile, Regression, Sampling, t-Test: Paired Two Sample for Means, t-Test: Two-Sample Assuming Equal Variances (highlighted), t-Test: Two-Sample Assuming Unequal Variances, and z-Test: Two Sample for Means. On the right side of the dialog box are buttons for "OK", "Cancel", and "Help". The background shows a blank Excel spreadsheet with columns labeled A through Q and rows numbered 1 through 31. The status bar at the bottom includes a "Draw" button and "AutoShapes" options.

# Проверка хипотезата за равенство на две дисперсии

## F-Test Two-Sample for Variances

	<i>Variable 1</i>	<i>Variable 2</i>
Mean	1,455	1,199
Variance	<u>0,027</u>	0,019
Observations	10	10
df	9	9
F	1,388	
P(F<=f) one-tail	0,316	
F Critical one-tail	3,179	

**F < FCritical one-tail**  
Двете дисперсии  
СЪВПАДАТ

ДИСПЕРСИИТЕ СА СТАТИСТИЧЕСКИ НЕОТЛИЧИМИ



# Проверка хипотезата за равенство на две математически очаквания с неизвестни но равни дисперсии

## t-Test: Two-Sample Assuming Equal Variances

Да се внимава за алфа - не се отпечатва тук е 0,01

	<i>Variable 1</i>	<i>Variable 2</i>
Mean	1,455	1,199
Variance	0,027	0,019
Observations	10	10
Pooled Variance	0,023	
Hypothesized Mean Difference	0,000	
df	18	
t Stat	3,790	
P(T<=t) one-tail	0,001	
t Critical one-tail	2,552	
P(T<=t) two-tail	0,001	
t Critical two-tail	2,878	

**tStat > t Critical two-tail**

**Двете математически очаквания НЕСЪВПАДАТ**

**M1 различно от M2**

# Проверка хипотезата за равенство на две математически очаквания с неизвестни но равни дисперсии

## t-Test: Two-Sample Assuming Unequal Variances

	<i>Variable 1</i>	<i>Variable 2</i>
Mean	84,42727273	87,2909091
Variance	3,996181818	0,42490909
Observations	11	11
Hypothesized Mean Difference	0	
df	12	
t Stat	-4,516993739	
P(T<=t) one-tail	0,000352741	
<i>t Critical one-tail</i>	1,782286745	
P(T<=t) two-tail	0,000705481	
t Critical two-tail	2,178812792	

**tStat > t Critical two-tail**

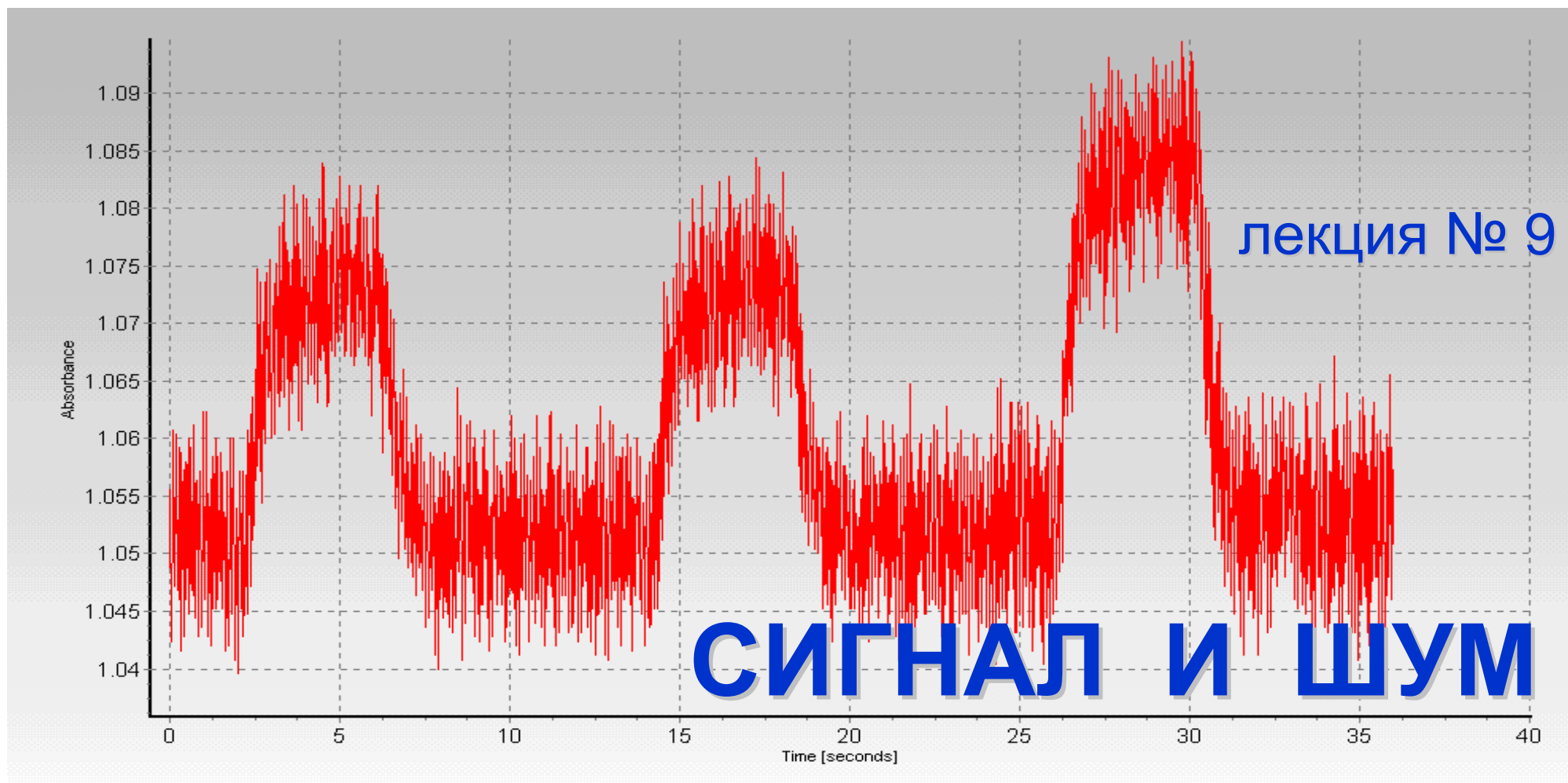
**Двете математически очаквания  
НЕСЪВПАДАТ**

# Днес разгледахме:

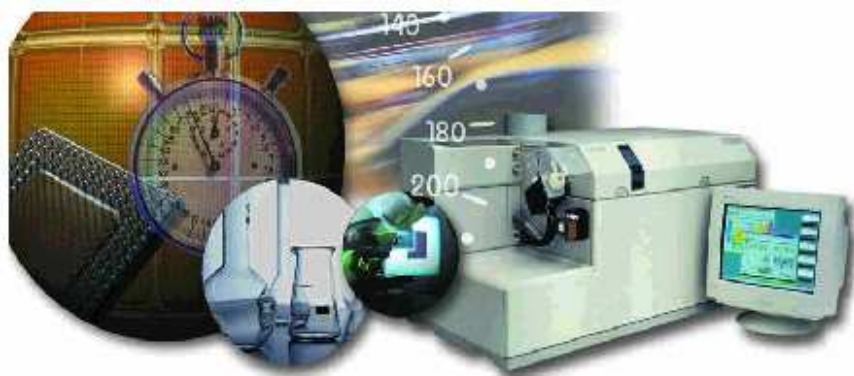
## 1. СТАТИСТИЧЕСКИ ХИПОТЕЗИ

- Основни понятия и същност на проверката на СХ.
- СХ за равенство на  $X$  на дадена стойност при неизвестно  $\sigma^2$ .
- СХ за равенство на две математически очаквания – при известни дисперсии
- СХ за равенство на две дисперсии - F - критерий
- Хипотези за равенство на две математически очаквания
  - при **известни** дисперсии
  - при неизвестни но **равни** дисперсии
  - при неизвестни и **неравни** дисперсии.

## 2. ОЦЕНКА НА СТАТИСТИЧЕСКИ ХИПОТЕЗИ ЧРЕЗ EXCEL



**граница на  
откриване  
LOD**



# Днес ще разгледаме:

1. Определение за сигнал и шум.
2. Откриване на сигнала. Дефиниране на :
  - граница на откриване
  - граница на гарантирано откриване
  - граница на определяне
3. Оценка на границата на откриване с  $t$ -разпределението.
4. Отношение сигнал/шум и методи за неговото повишаване

Signal - Wikipedia, the free encyclopedia - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites

Address http://en.wikipedia.org/wiki/Signal

Google signal

Sign in / create account

article discussion edit this page history

Your continued donations keep Wikipedia running!

# Signal

From Wikipedia, the free encyclopedia

Signal, signals or signalling may refer to:

**Scientific concepts** [edit]

- Signal (electrical engineering), a varying quantity that can carry information
- Signal processing, the field of techniques used to extract information from signals
- Signal (computing), an event, message, or data structure transmitted between computational processes
- Cell signalling (biology), the system of communication that governs basic cellular activities and coordinates cell actions
- Signals (biology), electrochemical activity in an organism

**Communications** [edit]

- Signals** used in various kinds of transport:
  - Road traffic signal
  - Railway signal
  - Semaphore
  - Beacon
  - International maritime signal flags, using the International Code of Signals
- Signals (military), a historical name for the military communications
- Signalling theory in biology, how organisms signal their state to others
- Distress signal
- Smoke signal (See also Category:Early telecommunications)
- Signalling (telecommunications), a part of some communication protocols

Wiktionary Look up *signal, signaling* in Wiktionary, the free dictionary.

**Contents** [hide]

- Scientific concepts
- Communications
- Names
- Other

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

interaction

- About Wikipedia
- Community portal
- Recent changes
- Contact us
- Make a donation
- Help

search

Go Search

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version

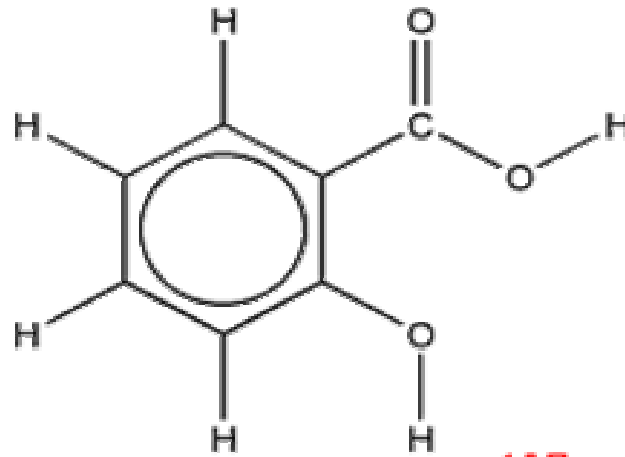
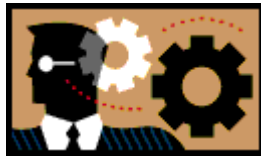
Start

2 Win... 5 Mic... 3 Int... SA Dict... Micros... 2 Mic... EN 08:35

# КАКВО РАЗБИРАТЕ ПОД ПОНЯТИЕТО СИГНАЛ ?

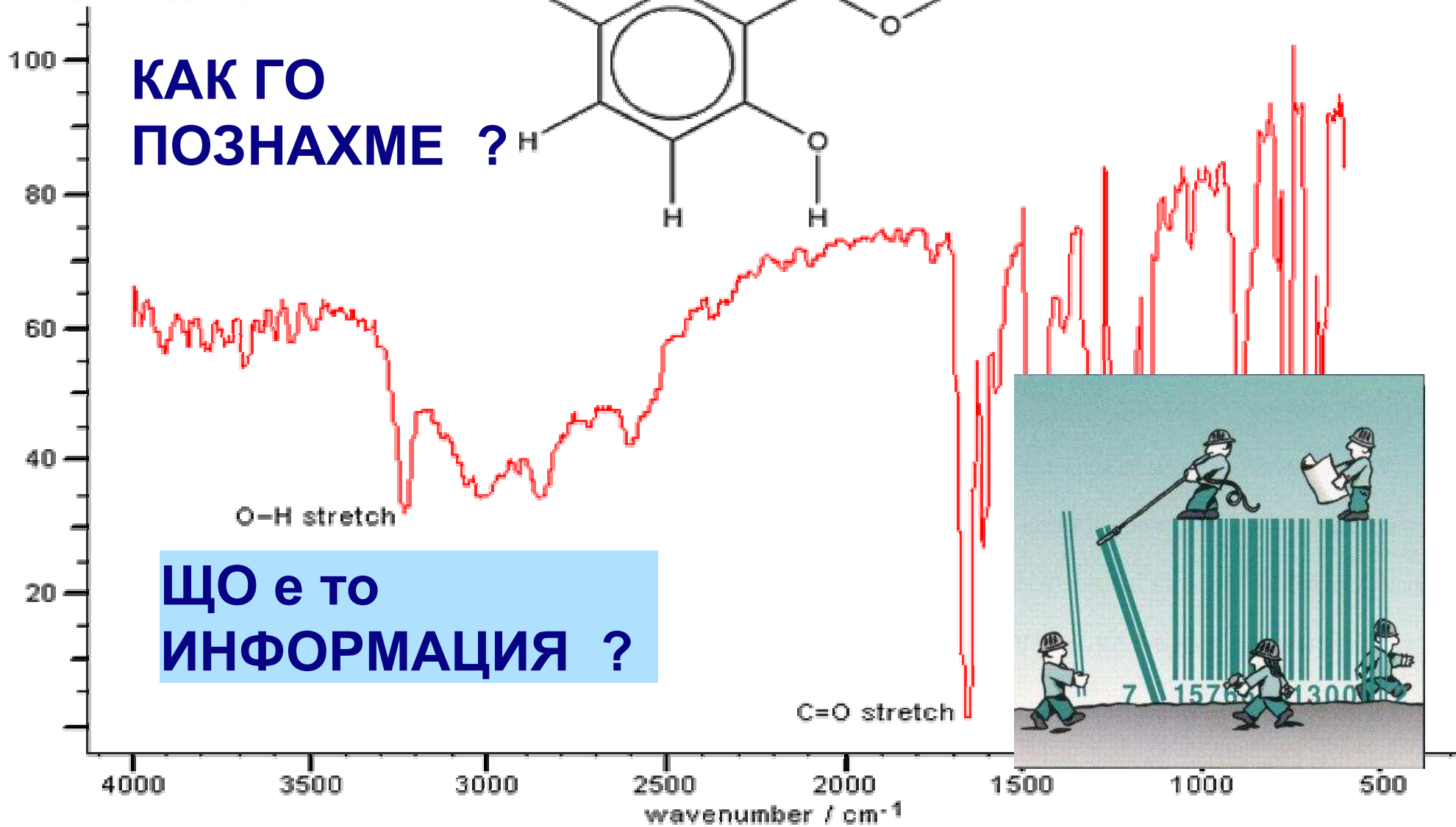
information

# КАКВО Е ТОВА?



transmittance / %

КАК ГО  
ПОЗНАХМЕ ?



# КОНЦЕПЦИЯ на анализа



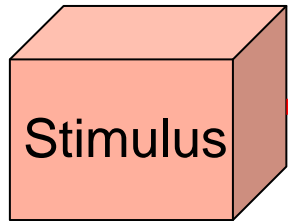
- ❖ Какво е веществото?
  - Идентификация - **Качествен анализ**
- ❖ В какво количество - пропорции?
  - **Количествен анализ** - абсолютен и относителен





# КОНЦЕПЦИЯ на анализа

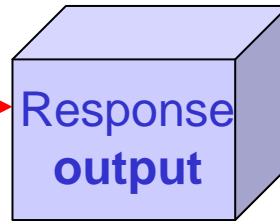
**Измервана величина (Measurand) –**  
quantity intended to be measured  
VIM 2.3 (2.6)



Енергиен  
източник



Обект на  
анализ



Аналитична  
информация

**СИГНАЛ**  
показание на системата за  
измерване  
(indication of measuring system)  
quantity value provided as **output** of  
measuring system  
VIM 4.1 (3.2)

Показанието се дава чрез положението на  
дисплея за аналогови изходящи сигнали,  
числото, показано за цифрови изходящи  
сигнали, от кодовия тип на кодовите сигнали,  
от номиналната или обявена стойност за  
материалните мерки

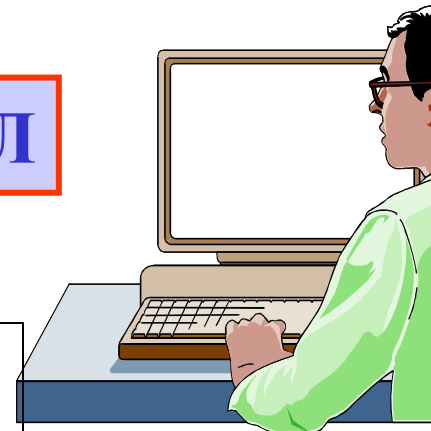
**ИЗМЕРВАНА  
ВЕЛИЧИНА**

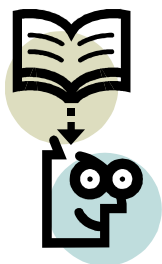
MEASURAND  
≠  
ANALYTE

↑  
Случайна  
величина

**СИГНАЛ**

↑  
Случайна  
величина





## ОСНОВНИ ПОНЯТИЯ

В химичния анализ по-пълното понятие е **ИЗПИТВАНЕ** което представя комплексността при оценката на П-Р-П-С

'Измерване': определяне стойността на величина

'Измервана величина': величина, която се измерва

'Определяем компонент' **АНАЛИТ**: веществото, химичната форма, която се измерва

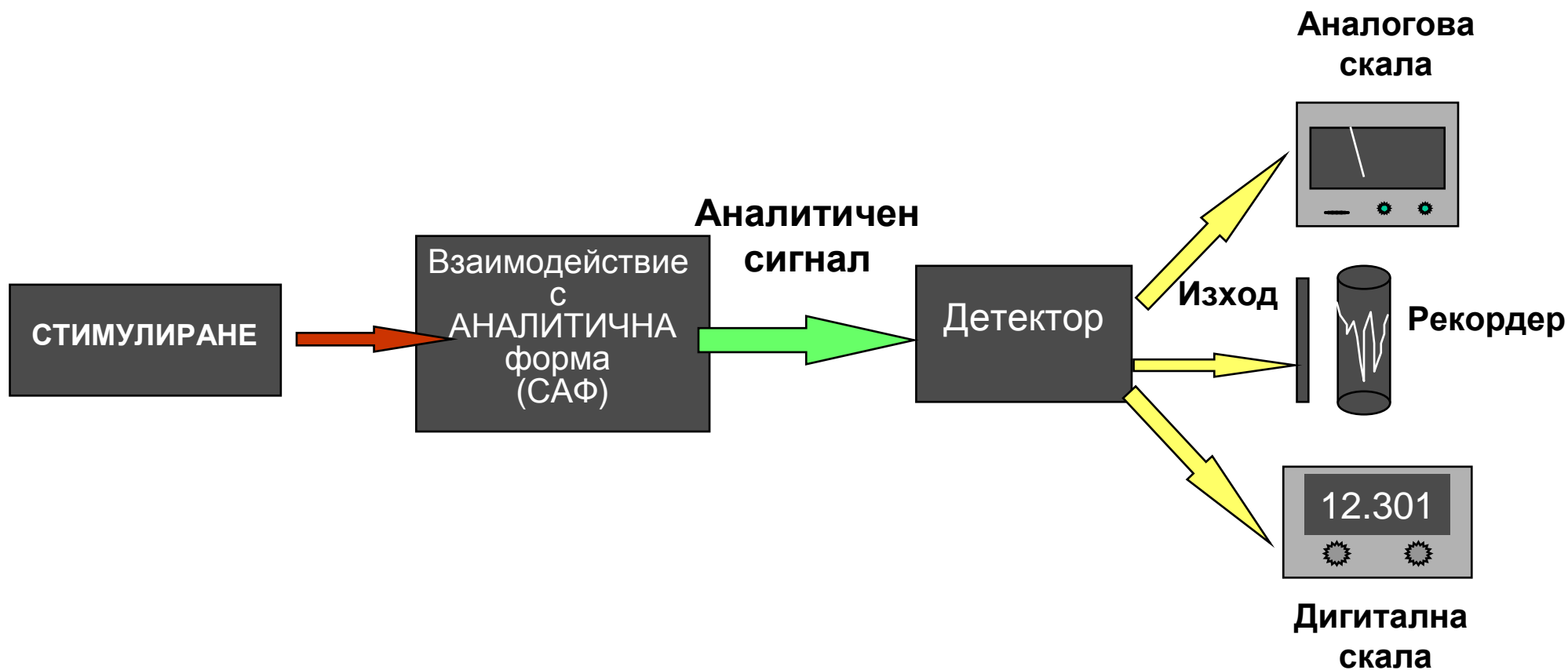
✂ **ИЗМЕРИТЕЛЕН СИГНАЛ (measurement signal)** - величина която представлява измерваната величина и е функционално свързана с нея. БДС 17397 (2.8)

[VIM, 1993]

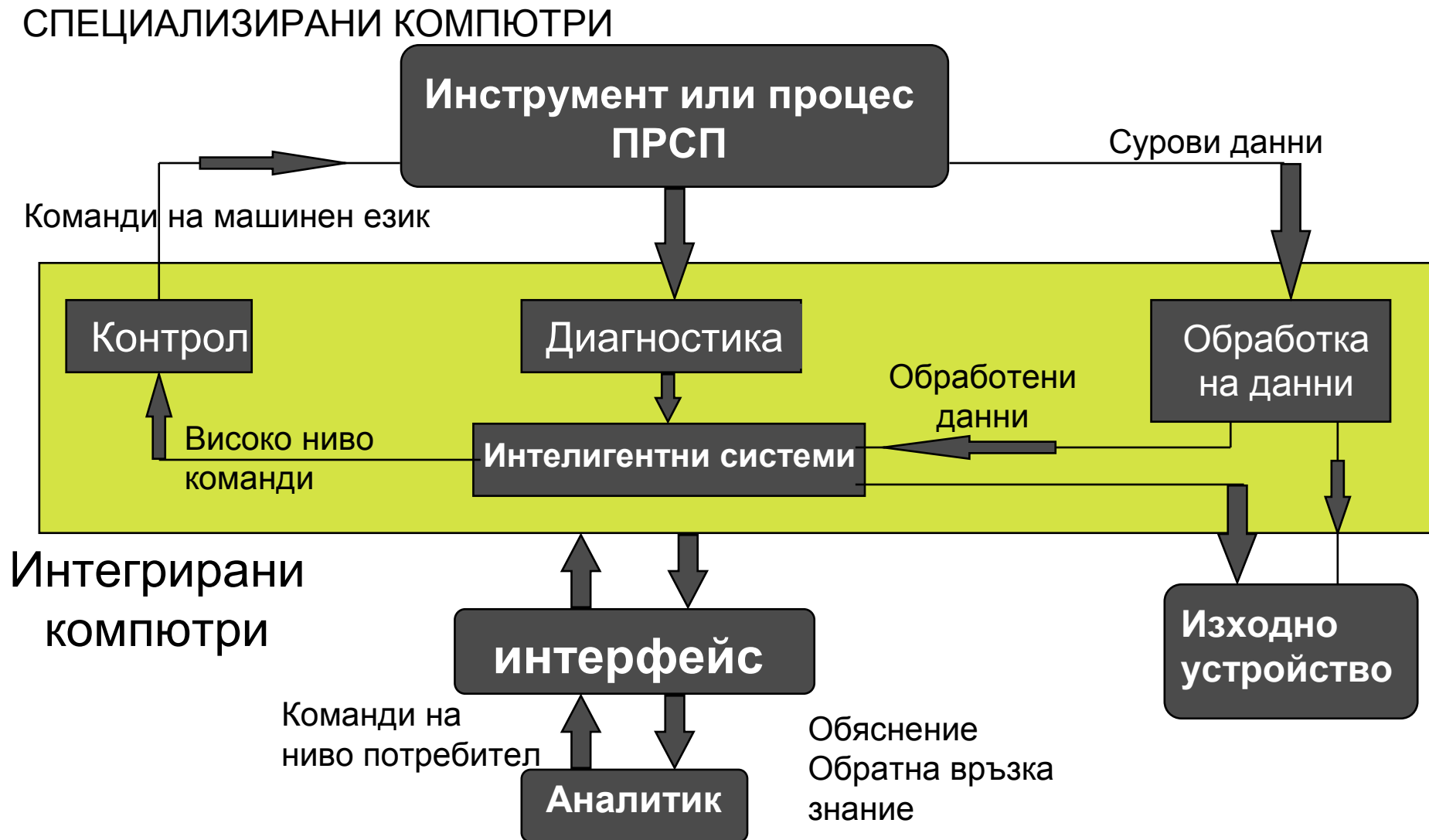
✂ СИГНАЛЪТ е откликът /отговора/ на прибора на определено въздействие.

# ИНСТРУМЕНТИ НА АНАЛИЗА

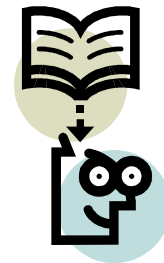
Компоненти на типичен инструмент



# ПЕТА генерация аналитични инструменти



# ДЕФИНИЦИИ



✎ Всяка една от регистриращите се електрични величини е **сигнал**, а тяхното разсейване около средната стойност **шум**.

✎ **Сигнал** е изменението на полезната величина причинно-следствено свързано с анализа, а **шум** е изменението от случайни фактори.

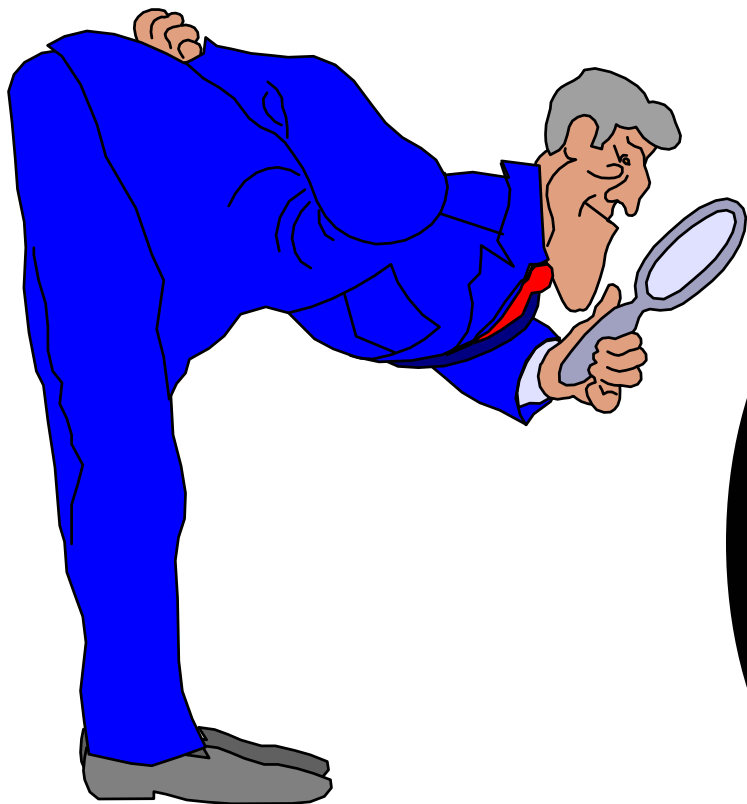
Шумът е част от сигнала не носеща полезна информация

(8<sup>-б</sup> - клас)



# ОТКРИВАНЕ НА СИГНАЛА

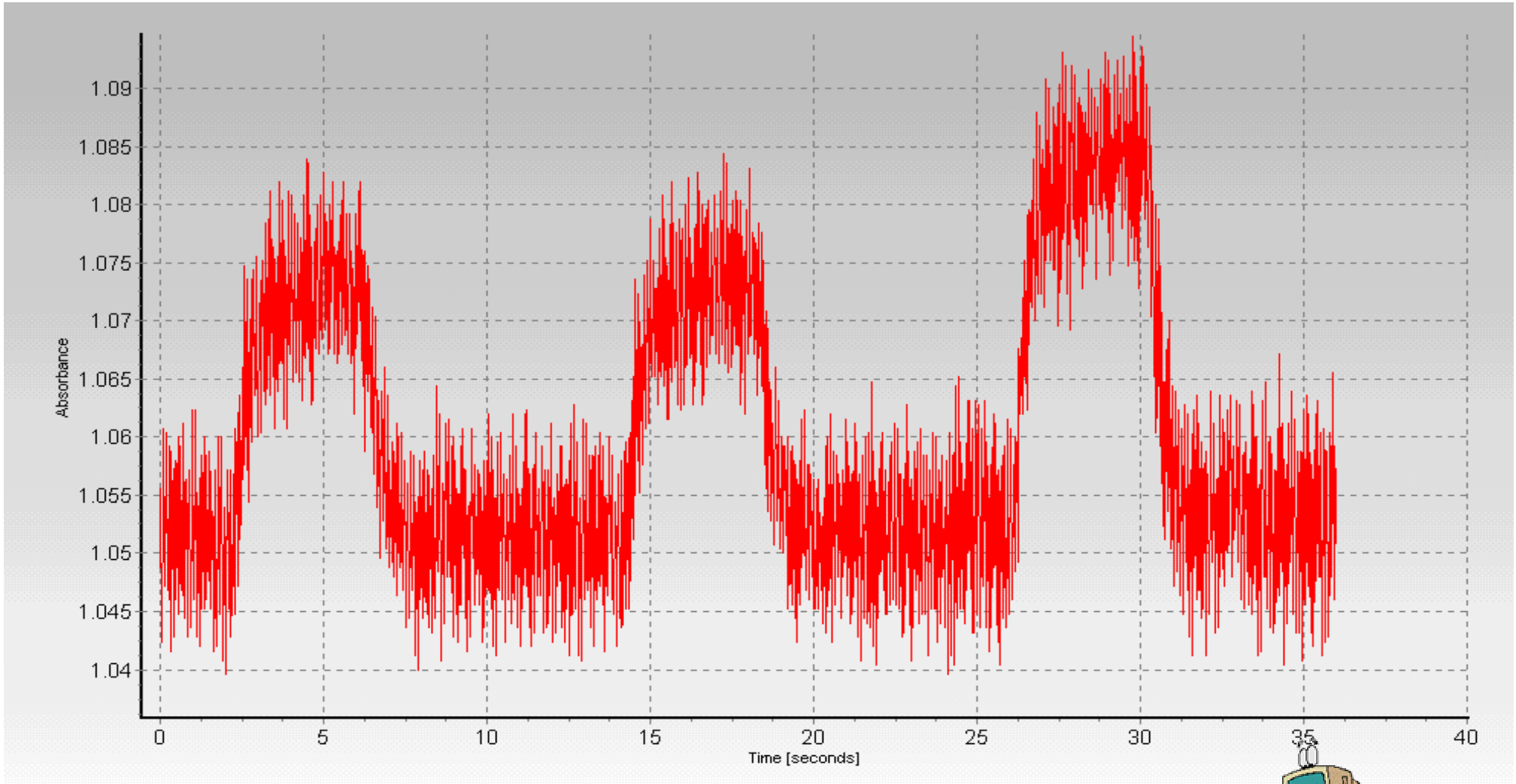
Виждате ли точката ?



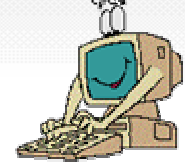
а



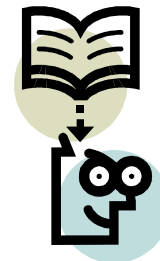
# ЗАПИС НА СИГНАЛ - ASDI Spec



**FAAS определяне на Mn в мъхове**



# ДЕФИНИЦИИ



✎ **ГРАНИЦА НА ОТКРИВАНЕ** - (Limit of detection - LOD)

Най-малкото количество аналит, което дава сигнал, статистически различим от нулевия

$$LOD_{sig} = \bar{X}_0 + 3S_0$$

✎ **ГРАНИЦА НА ГАРАНТИРАНО ОТКРИВАНЕ** – Такова количество аналит което дава сигнал с математическо очакване отстоящо на  $6\sigma_0$

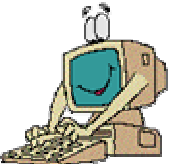
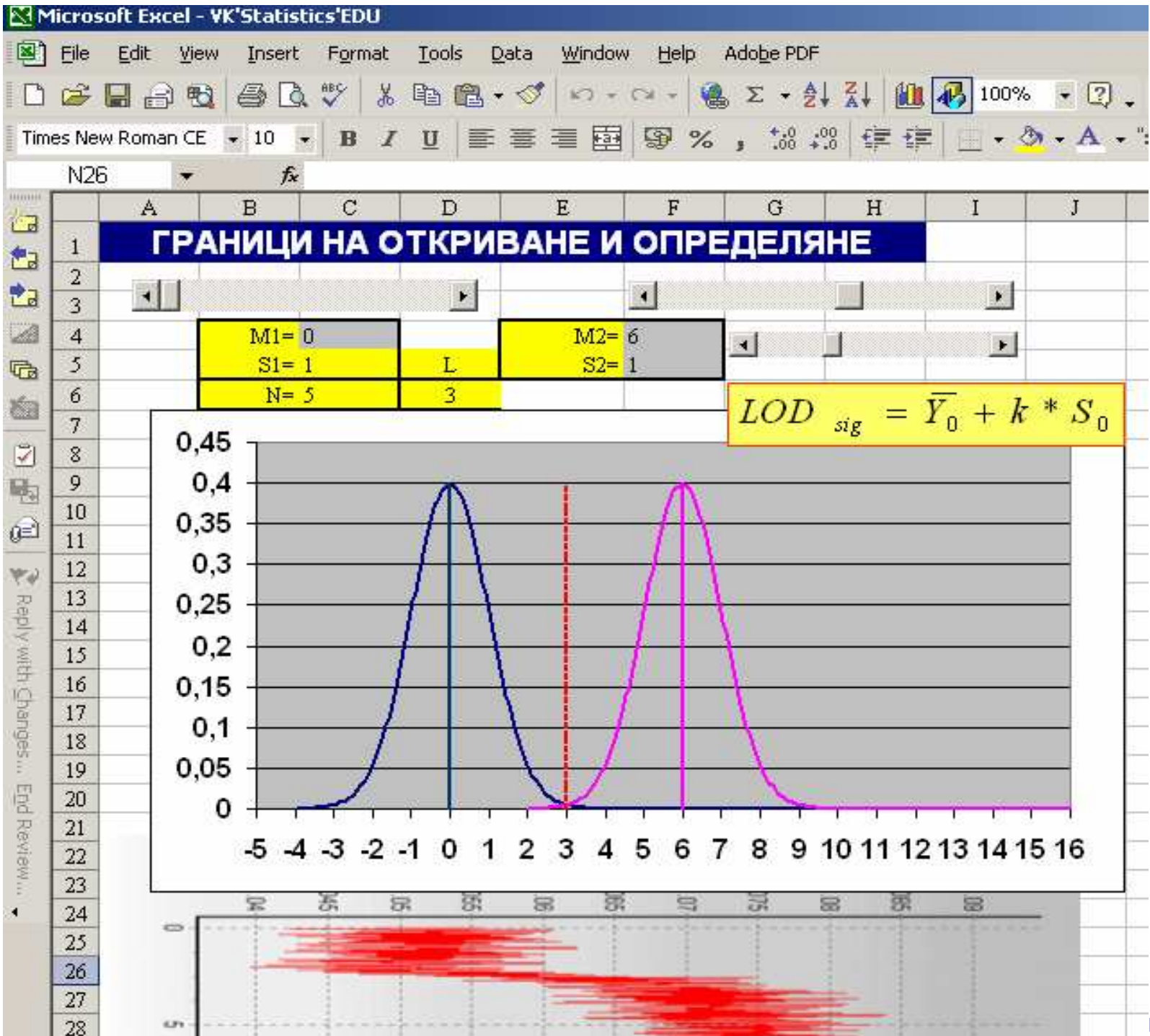
✎ **ГРАНИЦА НА ОПРЕДЕЛЯНЕ** - (Limit of quantitation -LOQ)

Такова количество аналит което дава сигнал с математическо очакване отстоящо на  $10\sigma_0$

$$LOD(Q) = \frac{k \cdot S_0}{m}$$

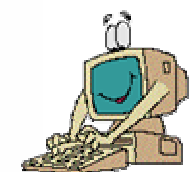
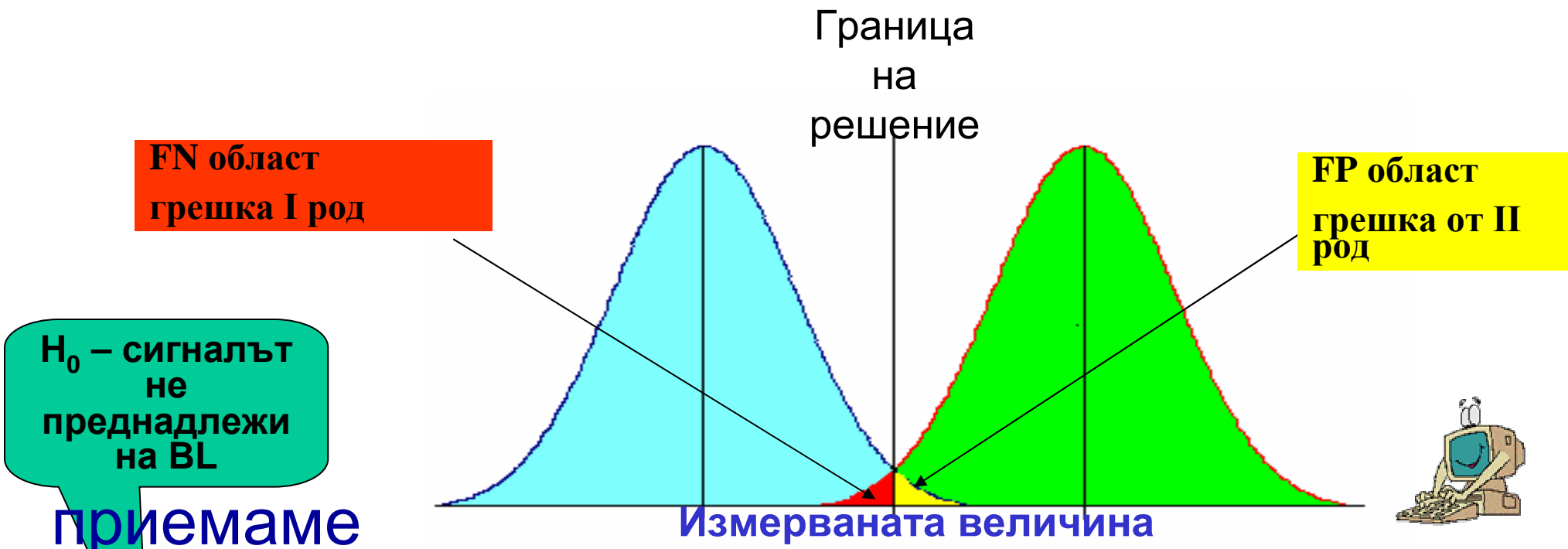
- $K = 3$  - граница на ОТКРИВАНЕ
- $K = 6$  - граница на ГАРАНТИРАНО ОТКРИВАНЕ
- $K = 10$  - граница на ОПРЕДЕЛЯНЕ





# МОДЕЛ НА ГРАНИЦА НА ОТКРИВАНЕ

Отсъствие на определяемия компонент	Присъствие на определяемия компонент
-------------------------------------	--------------------------------------



	НЕ принадлежи	ПРИНАДЛЕЖИ
$H_0$ Sig $\neq$ BL	TN вярно отрицание	FN грешно отриц.
$H_1$ Sig = BL	FP грешно приемане	TP вярно приемане

# Важни забележки!

☒ Модел на LOD за единично измерване и за многократно измерване - работа със средни стойности - МЕТОД НА ПОСЛЕДОВАТЕЛНОТО ОТКРИВАНЕ

☒ Да се казва - **НЕ МОЖЕ ДА СЕ ДОКАЖЕ** присъствие на **аналит а не че няма анализ**



съдържанието на анализа е  
ПОД ГРАНИЦИТЕ НА  
ОПРЕДЕЛЯНЕ на метода

$$C_{\text{analyt}} < LOQ$$

# ТИПОВЕ ГРАНИЦИ НА ОТКРИВАНЕ



- ☒ **ФУНДАМЕНТАЛНА (fundamental LOD)** - при чисти условия
- ☒ **ИНСТРУМЕНТАЛНА (instrumental LOD)** - отчита шума на инструмента и влиянието на матрицата при анализ, но без процедурата за пробоподготовка
- ☒ **МЕТОДИЧНА (methodological LOD)** - отчита абсолютно цялостната аналитична процедура

# Оценка на границата на откриване с t-разпределението ?

На практика на аналитика са известни само оценките на  $\sigma_0$  и  $\mu$  а именно  $X_{cp}$  и  $S_0$ .

Следователно оценката на ГО се прави чрез разглеждане на хипотезата за равенство на две математически очаквания ползвайки Разпределението на Стюdent и статистика базирана на

t- критерии – (метод на последователното откриване)

$$t_{kr} = \frac{(\bar{X} - \bar{X}_0) \cdot \sqrt{N_A + N_0 - 2}}{\sqrt{\left(\frac{1}{N_A} + \frac{1}{N_0}\right) \cdot [(N_A - 1) \cdot S_A^2 + (N_0 - 1) \cdot S_0^2]}}$$

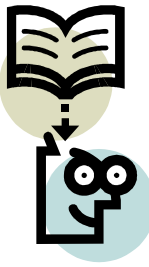
$$f = N_A + N_0 - 2$$

# Оценка на границата на откриване с t-разпределението ?

- Следователно ГРАНИЦАТА НА ОТКРИВАНЕ Е  
НАЙ-МАЛКАТА стойност на  $\bar{X}$  за която все още е в сила  $t_{kp} \geq t(f, \alpha)$
- Или


$$LOD = \bar{X}_0 + \frac{t(f, \alpha) \cdot \sqrt{\left(\frac{1}{N_A} + \frac{1}{N_0}\right) \cdot [(N_A - 1) \cdot S_A^2 + (N_0 - 1) \cdot S_0^2]}}{\sqrt{N_A + N_0 - 2}}$$

# Съотношение сигнал / шум (*signal-to-noise ratio S/N*)

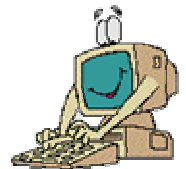


Съотношението СИГНАЛ/ШУМ – е основна характеристика на измервателната система (техническото средство)

ЦЕЛТА Е ДА СЕ ПОСТИГНЕ **ВИСОКО** съотношение СИГНАЛ/ШУМ

 Нормалното усилване на сигнала – не води до подобряване на съотношението, защото пропорционално расте и шума !!!!!

Съотношението СИГНАЛ/ШУМ е реципрочно на относителното стандартно отклонение



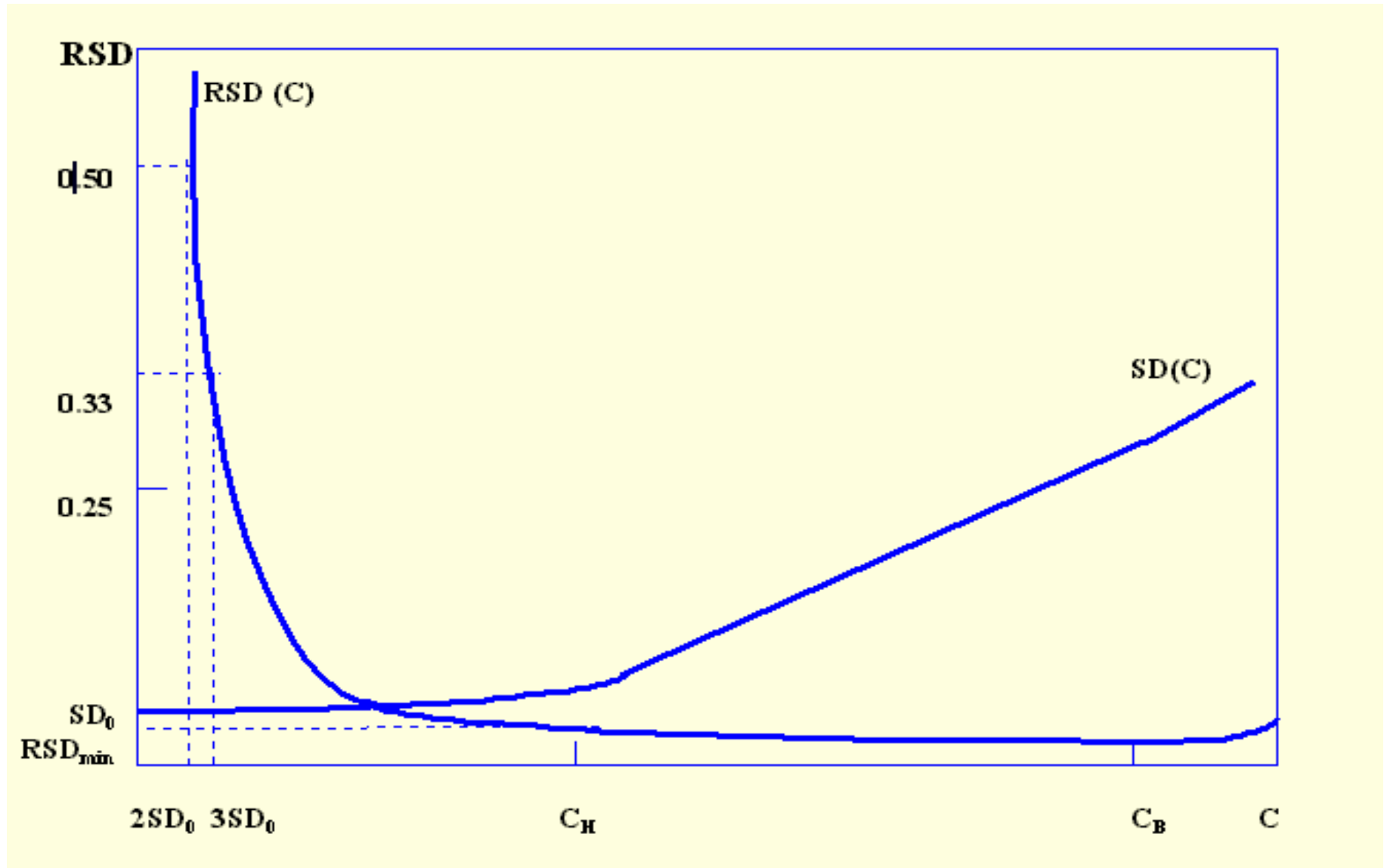
$$RSD\% = \frac{S}{\bar{X}} * 100$$

Шум N

Сигнал S

$$RSD\% = \frac{Nois}{3Nois * t(f, \alpha)} * 100 \approx 33\%$$

# СКЕДАСТИЧЕСКИ КРИВИ





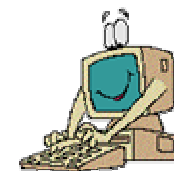
# ПОДОБРЯВАНЕ НА СЪОТНОШЕНИЕТО СИГНАЛ/ШУМ

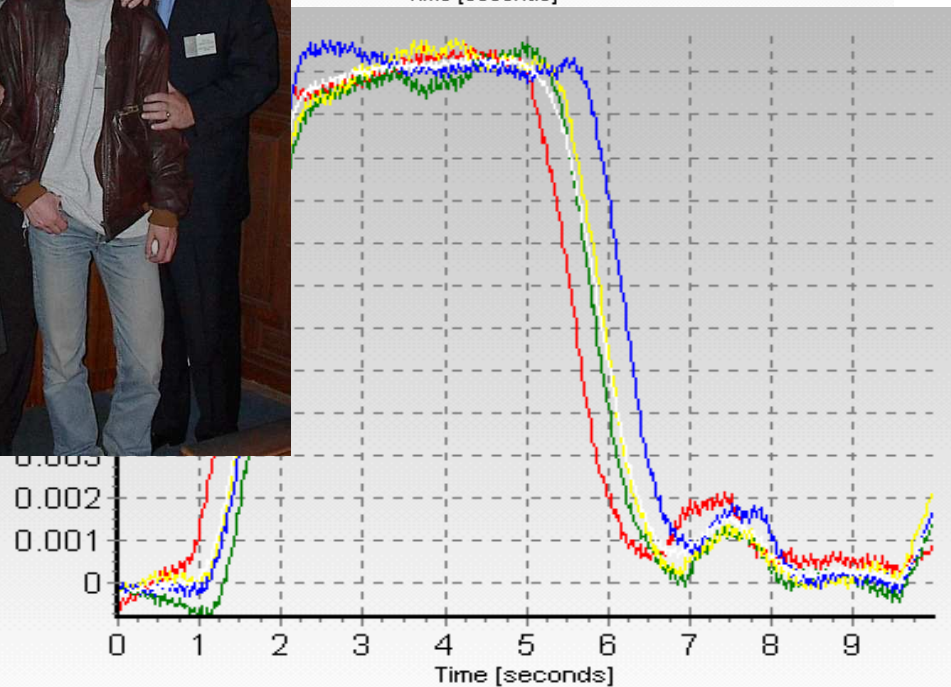
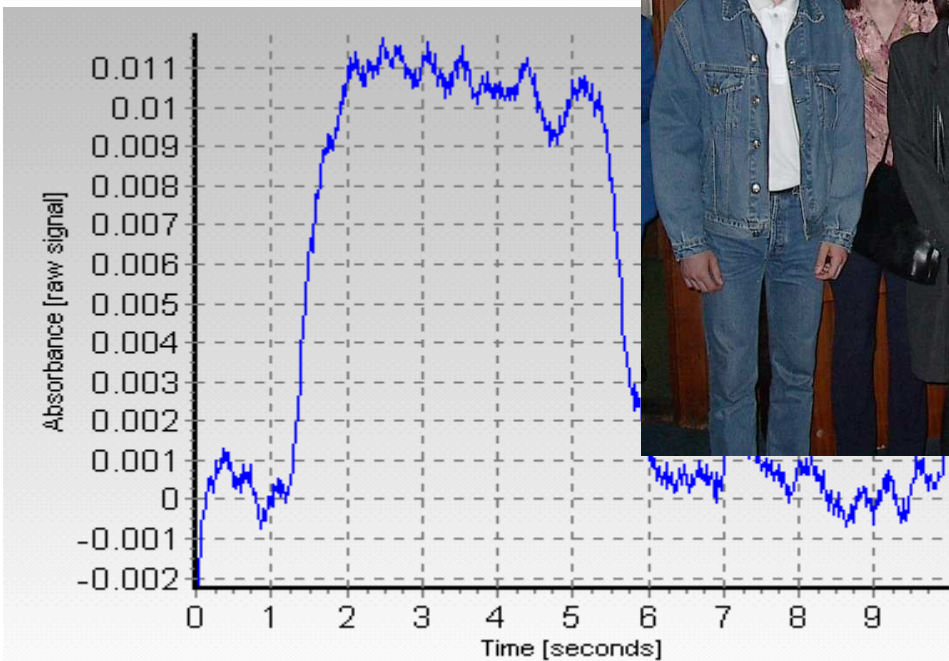
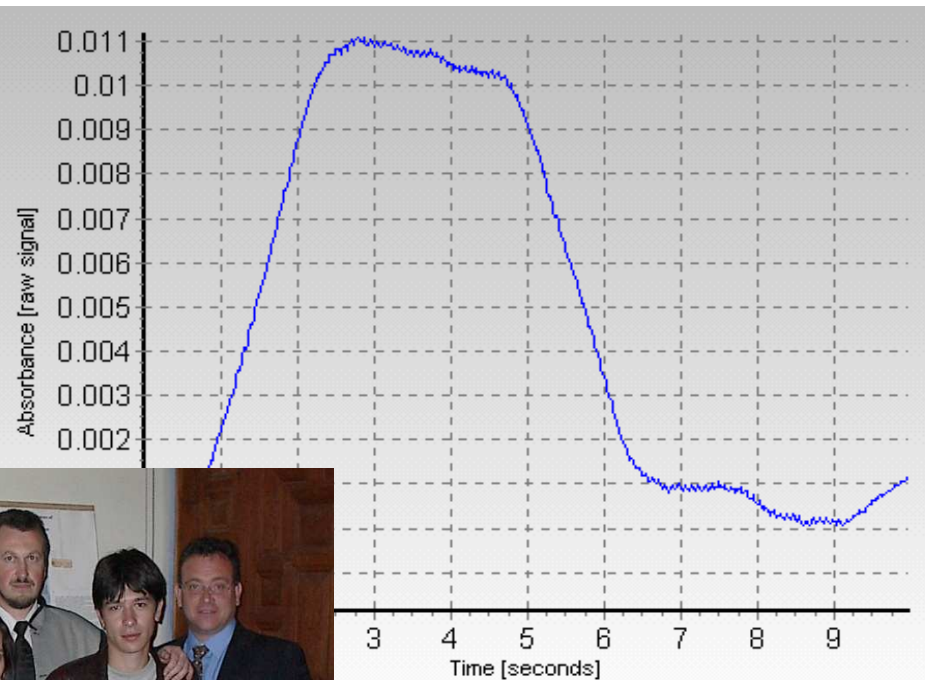
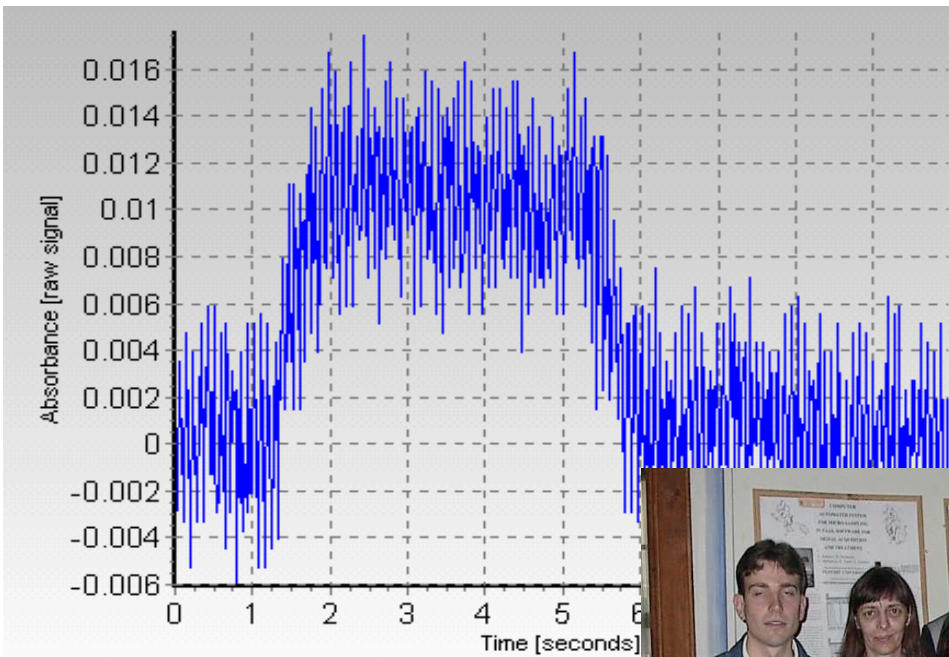
## НАЧИНИ :

- Оптимизация
- Усредняване
- Изглаждане (smoothing)

(в движещ се прозорец, Савитски-Голай и др.)

- Ансамбловото сумиране



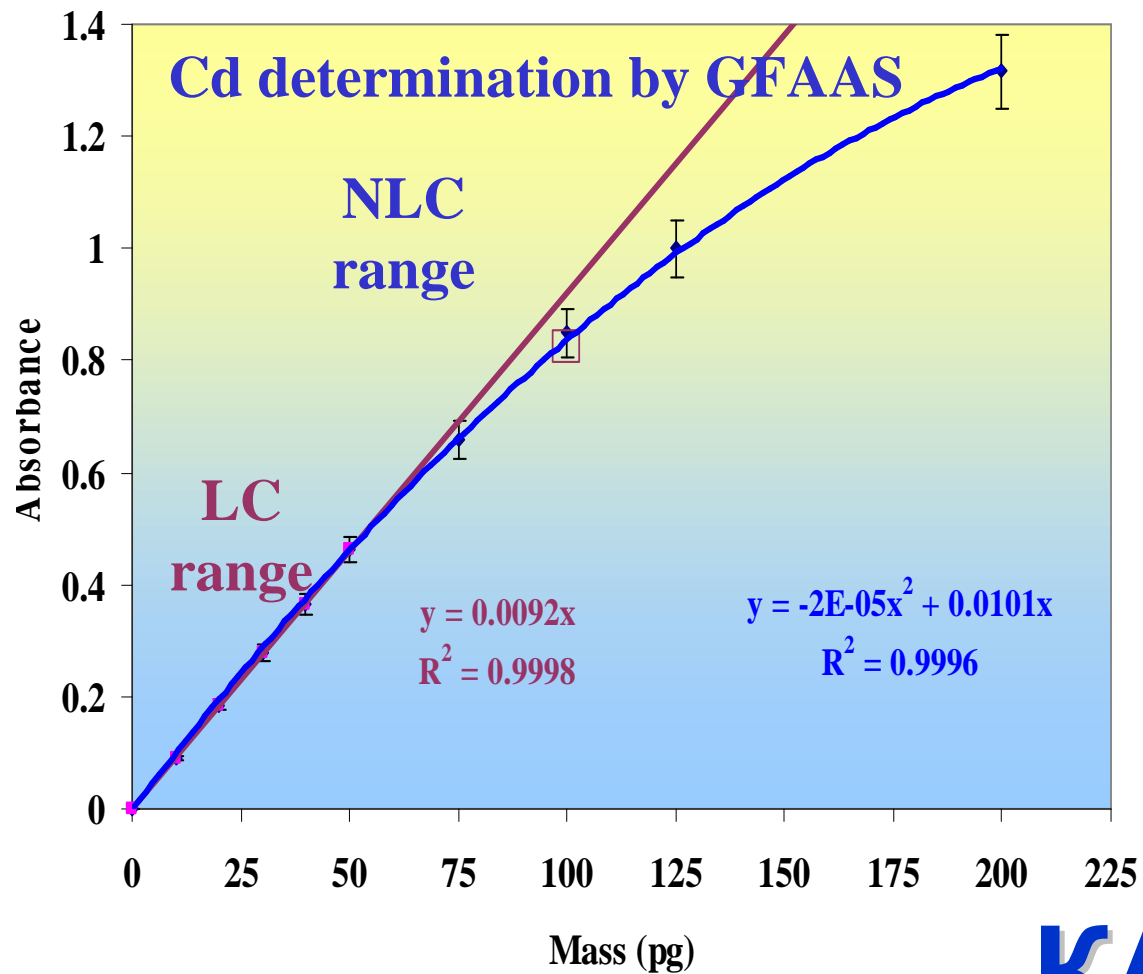


# Днес разгледахме: СИГНАЛ И ШУМ

1. Определение за сигнал и шум. Основни понятия.
2. Откриване на сигнала. Граница на откриване и определяне.

$$LOD_{sig} = \bar{X}_0 + 3S_0$$

3. Оценка на границата на откриване с t-разпределение.
4. Отношение сигнал/шум и методи за неговото повишаване



лекция № 10

# КАЛИБРИРАНЕ ПРИ ХИМИЧНИТЕ ИЗМЕРВАНИЯ



# Днес ще разгледаме:

## КАЛИБРИРАНЕ ПРИ ХИМИЧНИТЕ ИЗМЕРВАНИЯ

### 1. Корелация и регресия.

Зависимост и независимост на две случайни величини

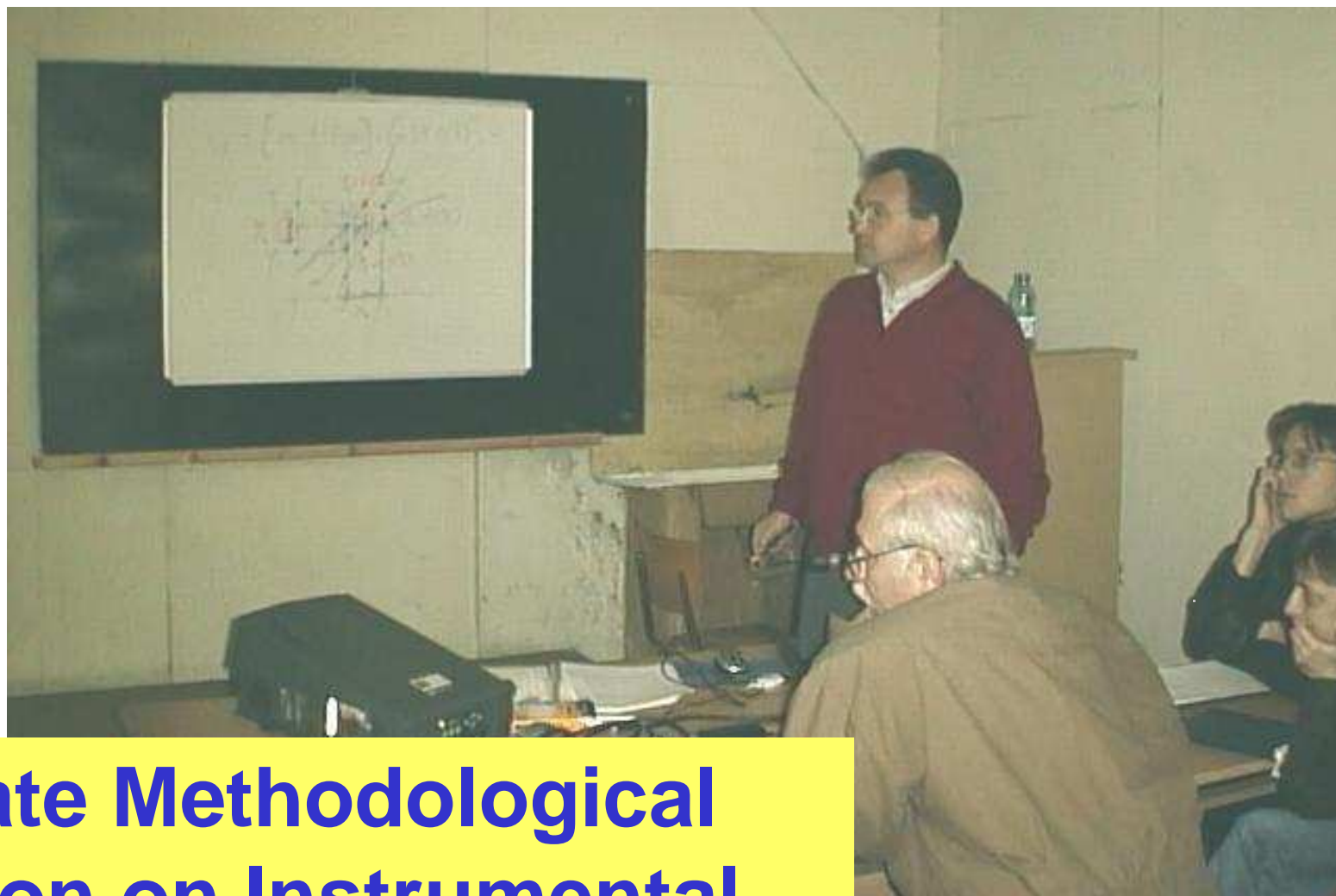
- коефициент на корелация

**2. Метод на най-малките квадрати** - намиране на уравнението на регресионна права

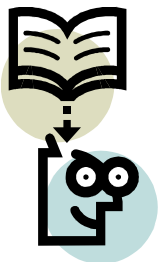
### 3. Калибрация.

- калибрационни стандарти - ССМ и еталони, проследимост
- калибрационна права - приложение

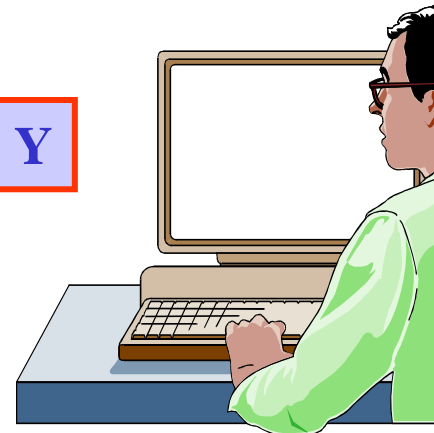
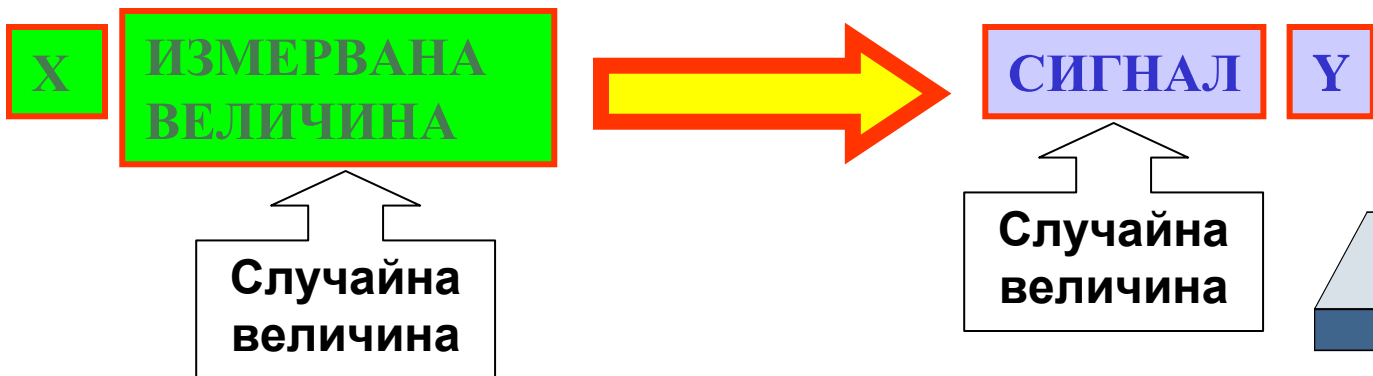
# Prof. Dr. Antonio Canals Hernández



## Univariate Methodological Calibration on Instrumental Analysis

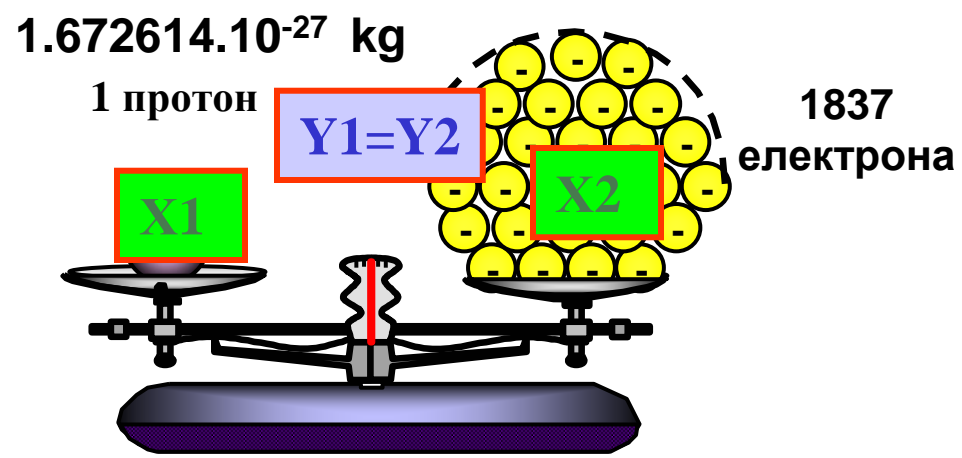


# АНАЛИЗЪТ Е СРАВНЕНИЕ !!!!



Всяка стойност се разкрива при определено **въздействие** върху околния свят.

Стойността на величината се проявява при връзката на тази величина с други величини или при **сравнението** на въздействието и с това на еталони или стандарти.



**Стохастическа връзка** - функцията на разпределение на  $Y$  зависи от тази на  $X$   
**частен случай**

**КОРЕЛАЦИОННА връзка**- математическото очакване на  $Y$  зависи от това на  $X$



# Корелация и регресия

**Correlation = Degree of association between two variables X and Y**

- При корелация

– заданието е да се установи степента на асоциираност (зависимост) между две величини. Оценява се чрез коефициента на корелация.

**Calibration graph = Regression**

$$y = f(x)$$

- При регресия

– заданието е да се установи стойността на измерваната величина чрез регресионните параметри (и техните стандартни отклонения) описващи функционалната връзка между две величини.



# КОЕФИЦИЕНТ НА КОРЕЛАЦИЯ

Корелационната зависимост между две величини се оценява с коефициента на корелация означаван като  $K$ ,  $R$ ,  $R^2$  или  $r$

ЗАЕМА СТОЙНОСТИ от  $-1$  ;  $0$  ;  $+1$

Ако  $X$  и  $Y$  са независими - то  $R = 0$   
(необходимо но НЕдостатъчно условие за независимост)

ОБРАТНОТО ТВЪРДЕНИЕ **НЕ Е**  
ЗАДЪЛЖИТЕЛНО ВЯРНО -  
ако  $R = 0$  то **МОЖЕ**  $X$  и  $Y$  да са независими

## КОВАРИАЦИЯ COV

$$R_{XY} = \frac{M[(X - M(X))(Y - M(Y))]}{\sqrt{M(X - M(X))^2 \cdot M(Y - M(Y))^2}}$$

## СТАТИСТИЧЕСКА ОЦЕНКА НА КОЕФИЦИЕНТА НА КОРЕЛАЦИЯ

$$R_{XY} = \frac{\sum_{K=1}^M (X_K - \bar{X})(Y_K - \bar{Y})}{\sqrt{\sum_{K=1}^M (X_K - \bar{X})^2 \cdot \sum_{K=1}^M (Y_K - \bar{Y})^2}}$$

Величината:

$$t_{kp} = \frac{R_{XY} \sqrt{(M-2)}}{\sqrt{(1-R_{XY}^2)}}$$

е  $t$ -разпределена с  $f = M-2$  степени на свобода. Избира се ниво на значимост  $\alpha$  и се намира съответстващата интегрална граница  $t(f, \alpha)$  при двустранна постановка на въпроса. Ако  $|t_{kp}| > t(f, \alpha)$ , то коефициентът на корелация е статистически отличим от нула, при  $|t_{kp}| \leq t(f, \alpha)$  той е не отличим от нула

# КОЕФИЦИЕНТ НА КОРЕЛАЦИЯ

## Correlation coefficient

$$SS_{xx} = \sum (x_i - \bar{X})^2 = \sum x_i^2 - \sum x_i^2 / N$$

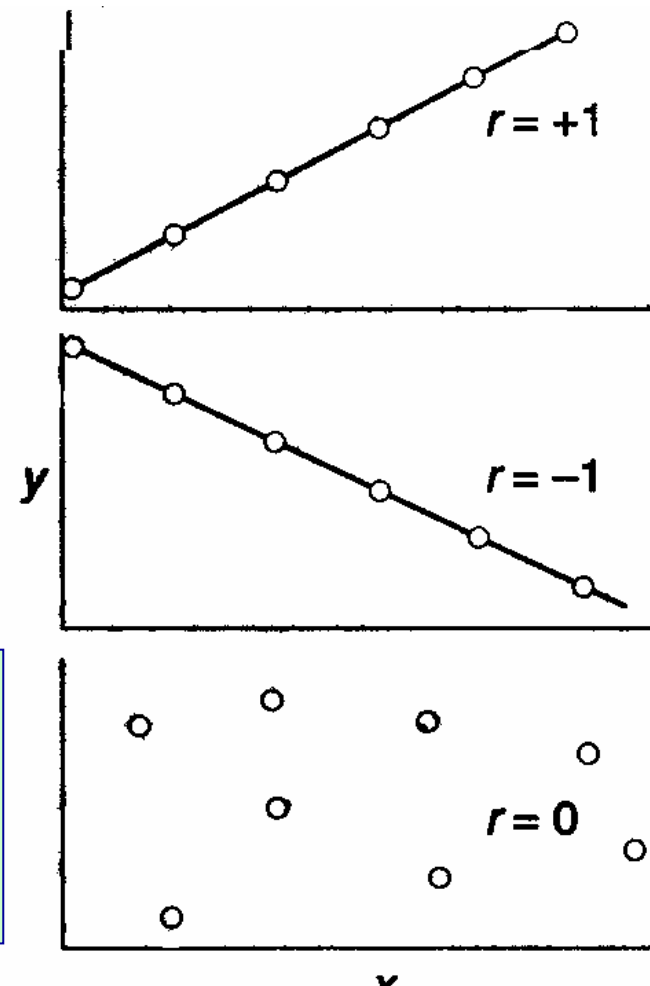
$$SS_{yy} = \sum (y_i - \bar{Y})^2 = \sum y_i^2 - \sum y_i^2 / N$$

$$SS_{xy} = \sum (x_i - \bar{X}) \cdot (y_i - \bar{Y}) = \sum x_i y_i - \sum x_i \sum y_i / N$$

$$r = S_{xy} / (S_{xx} S_{yy})^{1/2}$$

$$R = \frac{SS_{xy}}{\sqrt{SS_{xx} \cdot SS_{yy}}}$$

$$r = \sum (x_i - x_m)(y_i - y_m) / [\sum (x_i - x_m)^2 \sum (y_i - y_m)^2]^{1/2}$$



# КОЕФИЦИЕНТ НА КОРЕЛАЦИЯ

ВИНАГИ калибрационната крива трябва да се изчертае за да се прецени вида на зависимостта между  $X$  и  $Y$

- Бележки:
  - Дори зле изглеждаща корелация т.е. при значителни отклонения по  $y$ -направление, може да има  $r$  стойност близка до  $\pm 1$ .
  - Степените на свобода винаги трябва да бъдат оказани. Например линейна регресия само с две точки винаги има  $\rightarrow r = \pm 1$  (Перфектна корелация!).

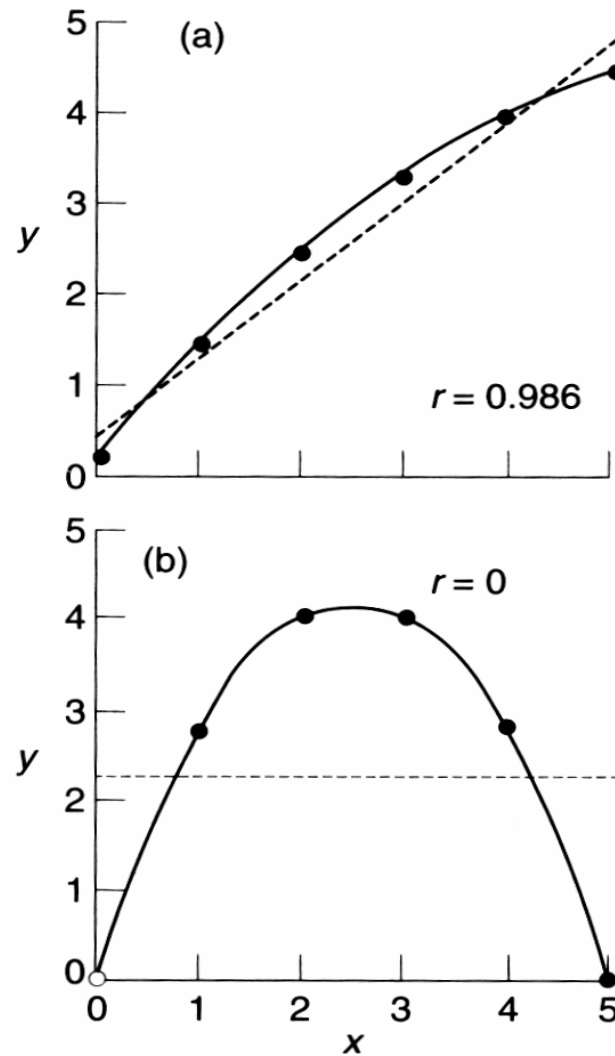


Figure 5.4 Misinterpretation of the correlation coefficient,  $r$ .

# Аналитичните методи биват:

- **АБСОЛЮТНИ МЕТОДИ**

- Моделът на връзката сигнал измервана величина е известен параметрите му са постоянни и следователно измерваната величина може да бъде определена без сравнителни измервания. (електрогравиметрия, био-тестове, радиоактивен разпад) .

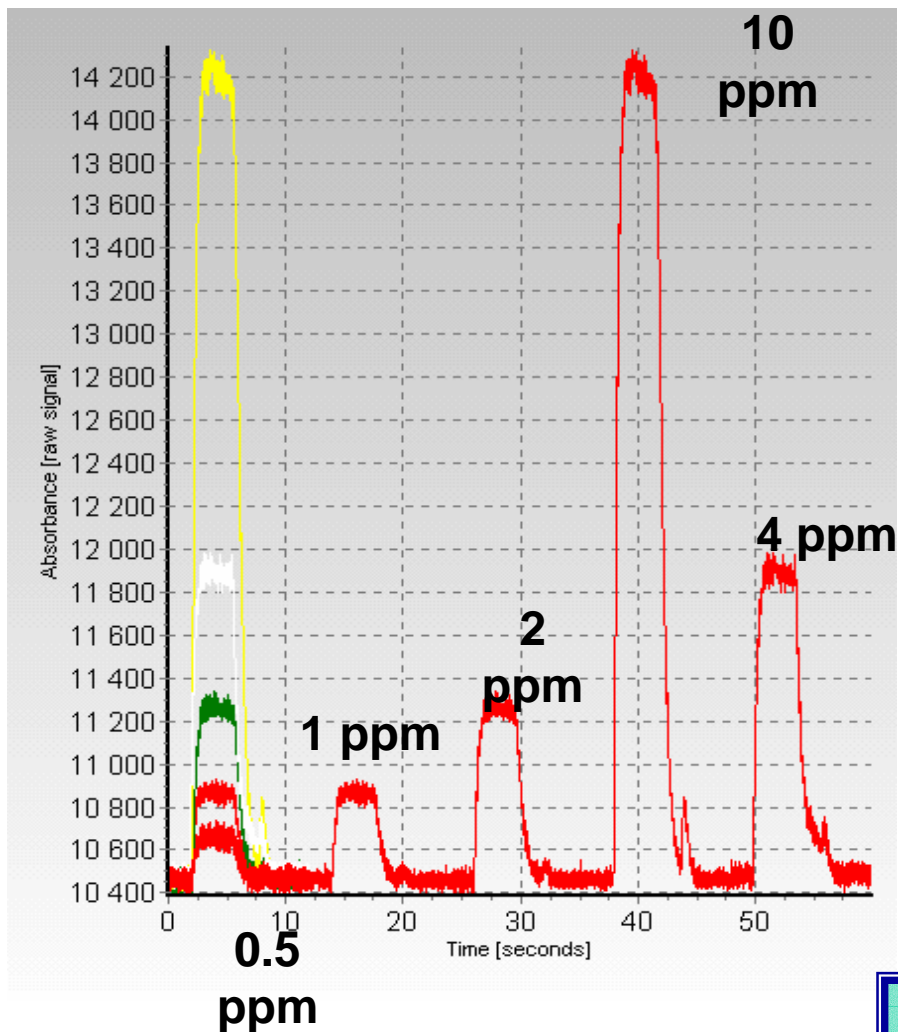
- **ОТНОСИТЕЛНИ МЕТОДИ**

- Моделът на връзката е неизвестен, или известен, но неговите параметри са неизвестни и трябва да бъдат експериментално установени. СРАВНЕНИЕ С ЕТАЛОНИ- КАЛИБРИРАНЕ

- **ВАЖНО** да се избере подходящ модел  
най-често се стремим към линеен  $Y = a_1X + a_0$   
но може и произволен полином :

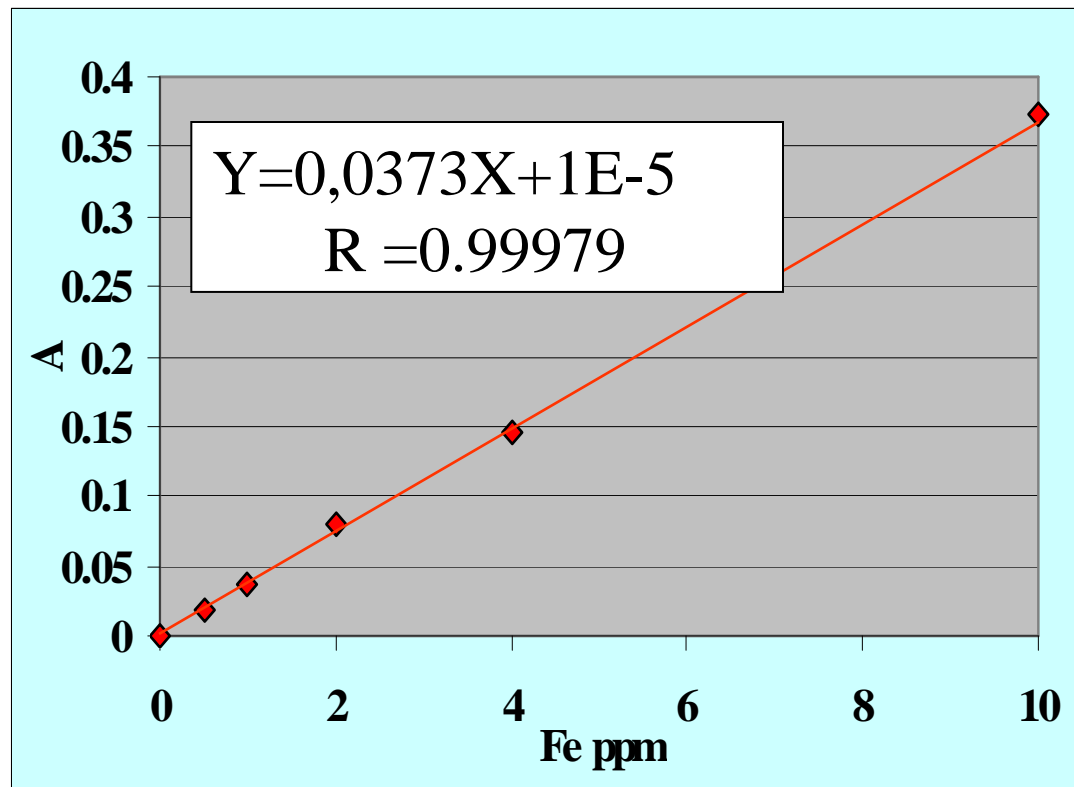
$$Y = a_1X + a_2X^2 + a_3X^3 + a_0$$

# Как да построим зависимостта на X и Y ?



**Fe - био-плазма**

Графично - на милиметрова хартия ?



**ЧРЕЗ МЕТОДА НА НАЙ - МАЛКИТЕ  
КВАДРАТИ**

# МЕТОД НА НАЙ-МАЛКИТЕ КВАДРАТИ

отклонения  
residuals

$$Sum = \sum_{K=1}^M (Y_K - \hat{Y}_K)^2$$

Нека са дадени извадки с еднакъв обем на случайните величини X и Y (K=1...M). Необходимо е да се намерят стойностите на коефициентите (оценки на истинските стойности  $a_1$  и  $a_0$ ), за които сумата:

$$Sum = \sum_{K=1}^M (Y_K - b_1 X_K - b_0)^2$$

има минимум. Първите частни производни на тази сума по  $b_1$  и  $b_0$  се приравняват на 0 (необходимо условие за минимум)

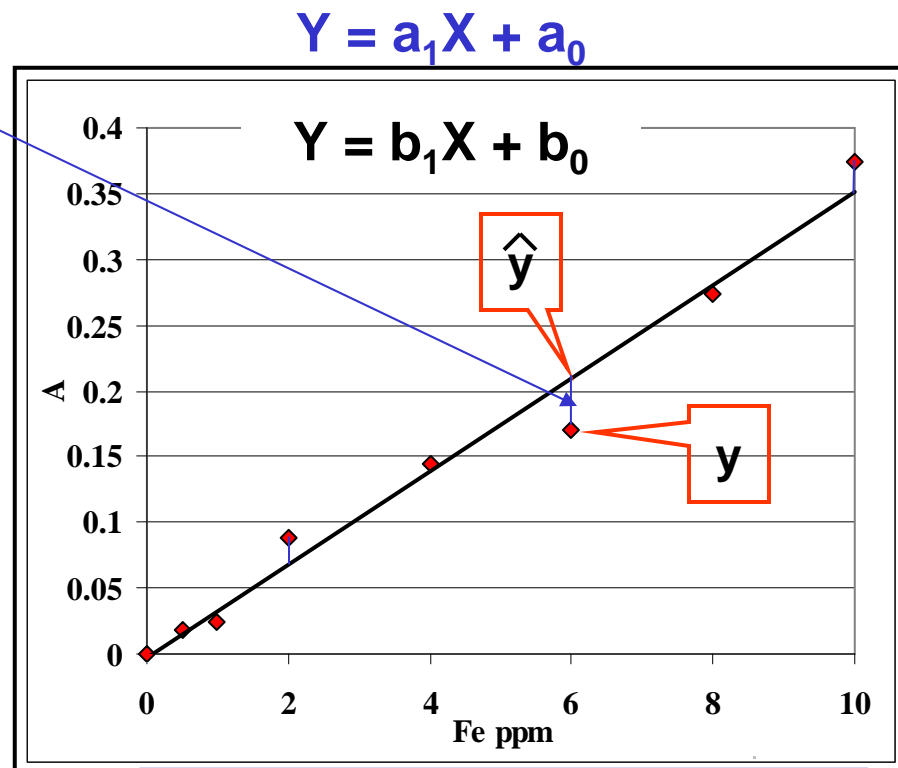
$$\frac{\delta Sum}{\delta b_0} = 2 \sum_K (Y_K - b_1 X_K - b_0)(-1) = 0$$

$$\frac{\delta Sum}{\delta b_1} = 2 \sum_K (Y_K - b_1 X_K - b_0)(-X_K) = 0$$

получава се система от две уравнения с две неизвестни

$$\begin{cases} b_1 \sum X_K + b_0 \cdot M = \sum Y_K \\ b_1 \sum X_K^2 + b_0 \cdot \sum X_K = \sum X_K \cdot Y_K \end{cases}$$

която се решава с формулите на Крамер и коефициентите (параметрите на правата) са:



$$b_1 = \frac{\sum X_K \sum Y_K - M \cdot \sum X_K Y_K}{(\sum X_K)^2 - M \cdot \sum X_K^2}$$

$$b_0 = \frac{(\sum X_K Y_K) \sum X_K - \sum X_K^2 \sum Y_K}{(\sum X_K)^2 - M \cdot \sum X_K^2}$$

$$b_1 = \frac{SS_{xy}}{SS_{xx}}$$

$$b_0 = \bar{Y} - b_1 \bar{X}$$

# Оценка на точките при регресия

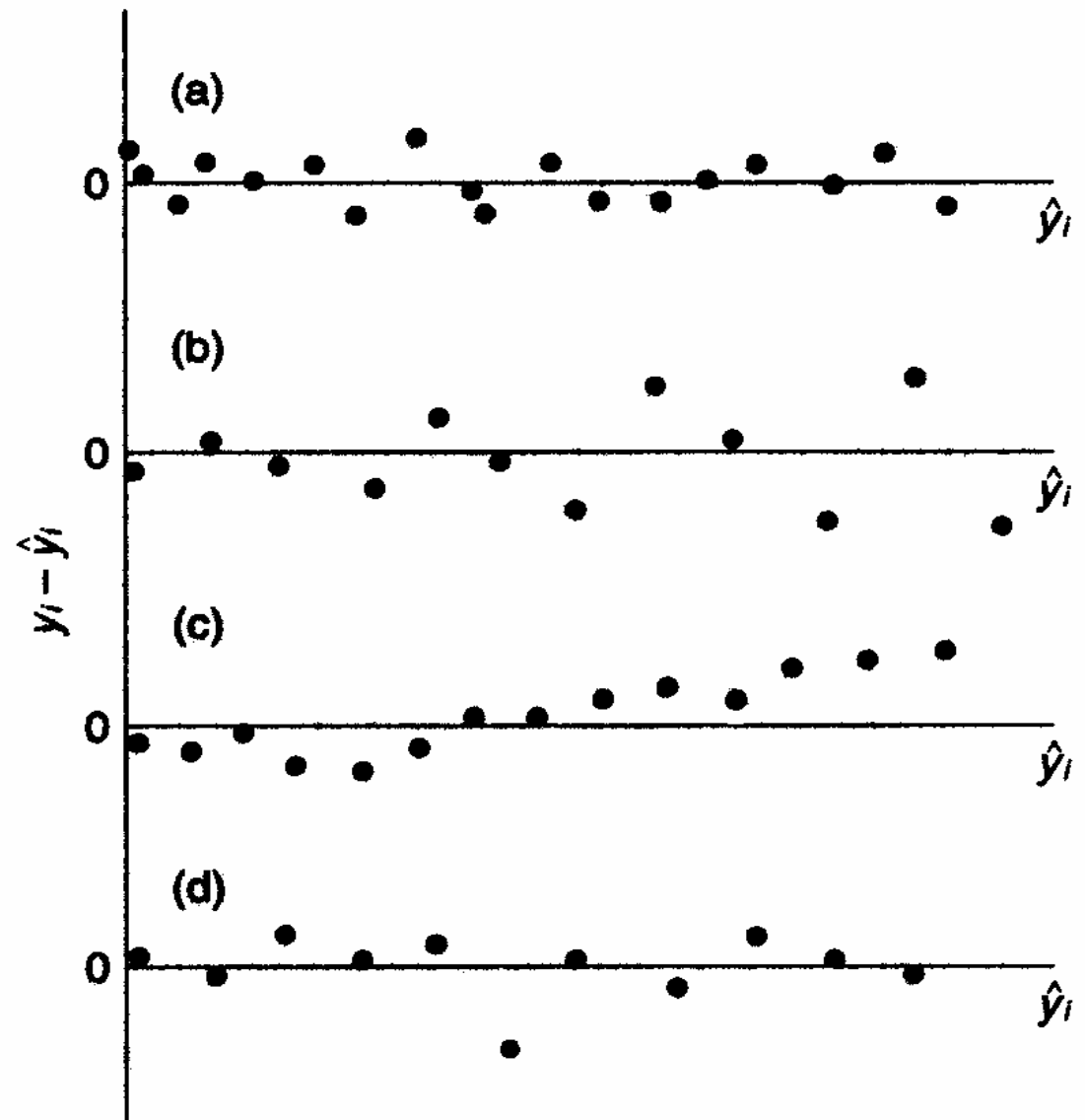
## Диагностика на остатъците (residuas)

a) задоволително разпределение на остатъците

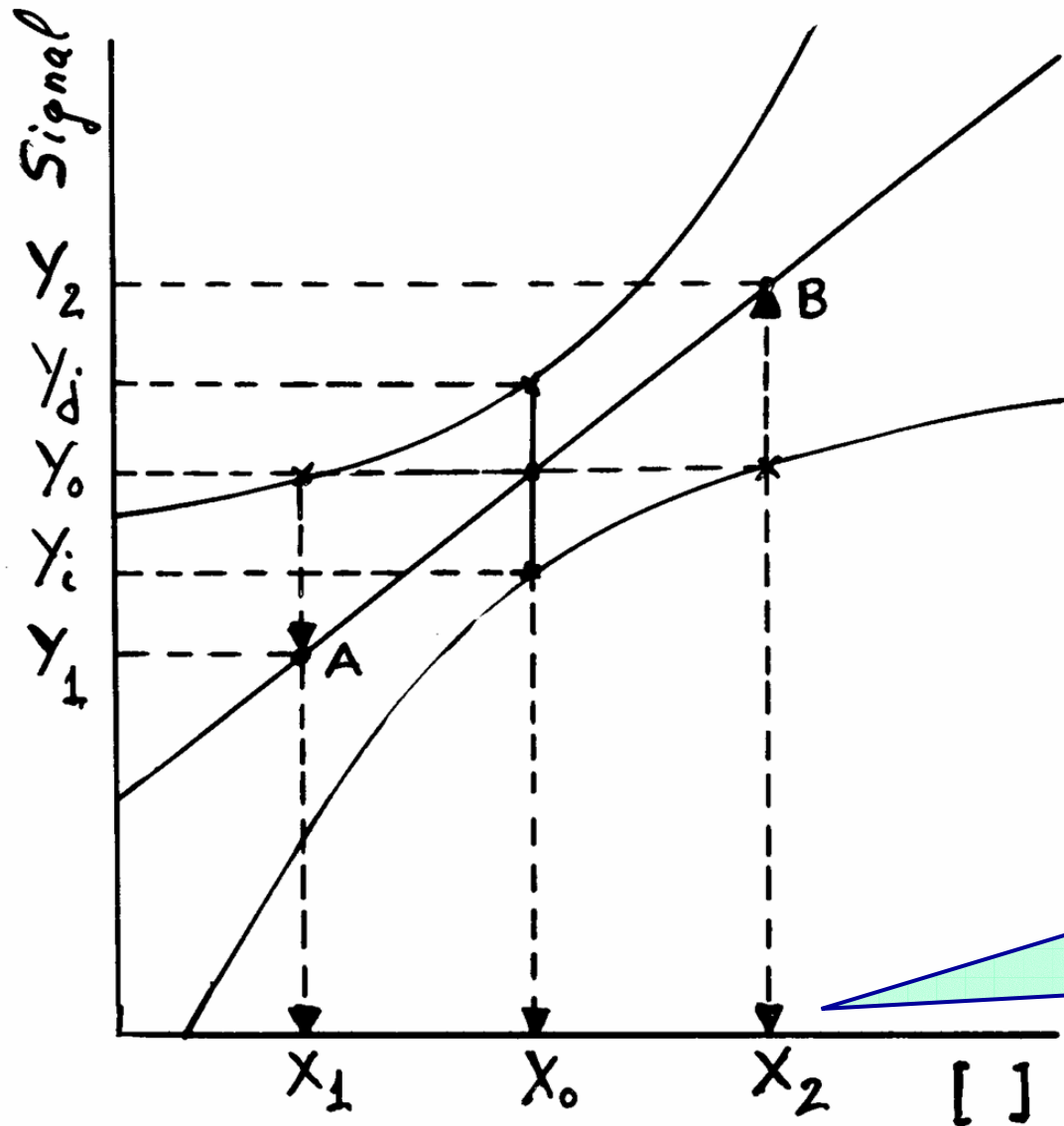
b) Остатъците нарастват с нарастване на  $Y$

c) има тенденции на нарастване и на намаляване - неправилен избран модел на кривата

d) задоволително разпределение с един беглец (outlier)



# Standard deviation of predicted values



**ВНИМАНИЕ !**  
**ПРАВА и**  
**ОБРАТНА задача**

**МНК**  
Приема, че  
неопределеността  
по X е нула



# СТАТИСТИЧЕСКИ ОЦЕНКИ НА ПАРАМЕТРИТЕ

ОЦЕНКА НА ДИСПЕРСИЯТА НА  
ОСТАТЪЦИТЕ (residuals) по Y

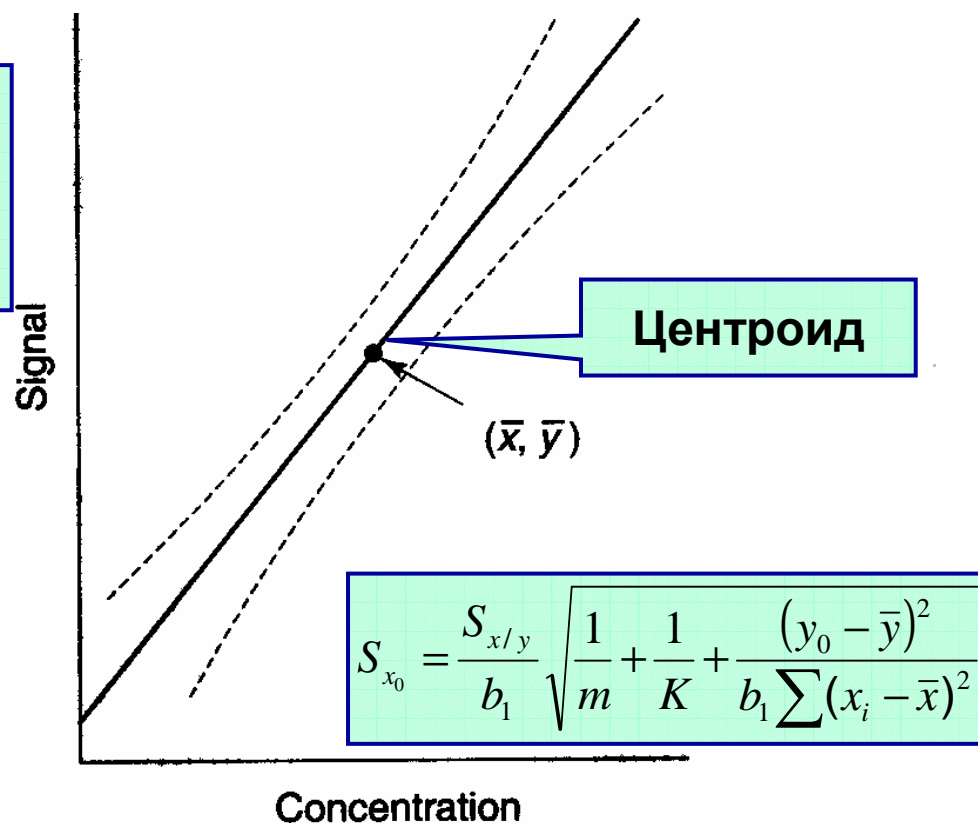
$$S_{Y/X}^2 = \frac{\sum_{K=1}^M (Y_K - \bar{Y})^2 - b_1^2 \sum (X_K - \bar{X})^2}{M - 2}$$

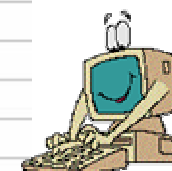
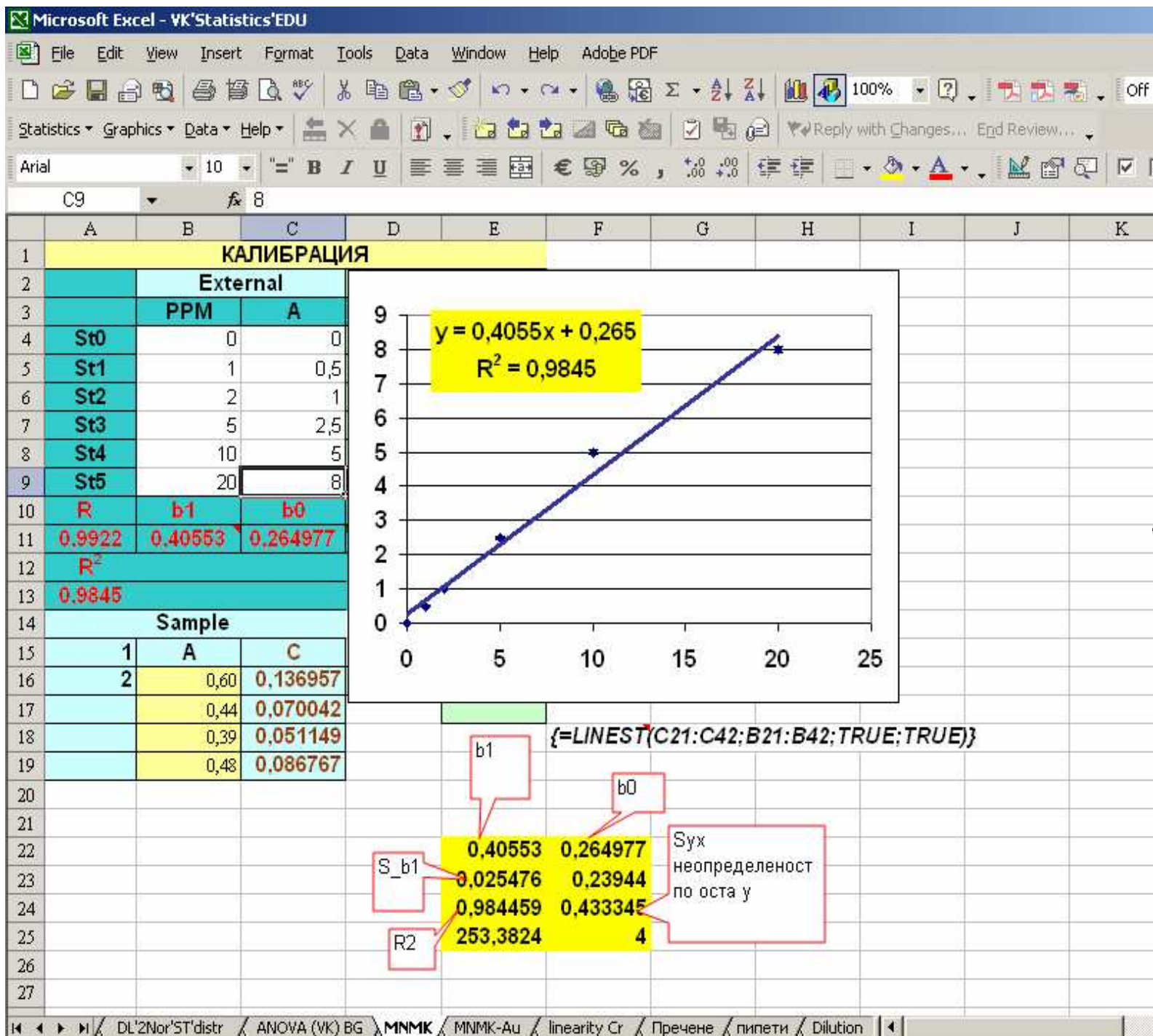
ОЦЕНКИ НА ДИСПЕРСИИТЕ НА ДВАТА  
ПАРАМЕТЪРА  $b_1$  и  $b_0$

$$S_{b_1}^2 = \frac{S_{Y/X}^2}{\sum (X_K - \bar{X})^2}$$

$$S_{b_0}^2 = \frac{S_{Y/X}^2 \cdot \sum X_K^2}{M \cdot \sum (X_K - \bar{X})^2}$$

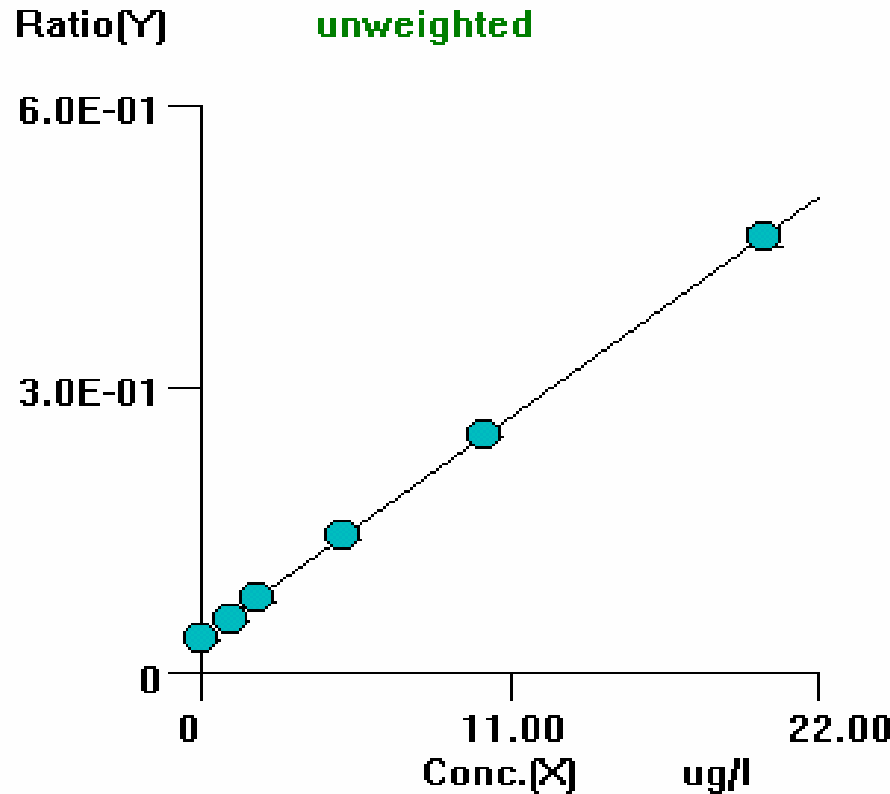
Параметрите  $a$  и  $a_0$  са случайни  
величини със съответните статистически  
оценки  $b$  и  $b_0$  и съответно оценки на  
дисперсиите  $Sb_1^2$  и  $Sb_0^2$





# КАЛИБРИРАНЕ

Mass: 75 As  
 Element:   
 IS: 103  
 Units: ug/l



Lv.	Conc.	Ratio	RSD %
1	0.00	3.61E-02	P 6.92
2	1.00	5.56E-02	P 3.24
3	2.00	7.76E-02	P 2.73
4	5.00	1.45E-01	P 3.64
5	10.00	2.50E-01	P 4.45E-01
6	20.00	4.57E-01	P 2.12
* 7	—	—	—
* 8	—	—	—
* 9	—	—	—
* 10	—	—	—
* 11	—	—	—
* 12	—	—	—
* 13	—	—	—
* 14	—	—	—
* 15	—	—	—
* 16	—	—	—
* 17	—	—	—
* 18	—	—	—
* 19	—	—	—
* 20	—	—	—

$$C_x = \frac{Y_{C_x} - b_0}{b_1}$$

Curve Fit:

r = 0.9999  
 Y = 2.112E-002\*X + 3.645E-002  
 X = 4.734E+001\*Y - 1.726E+000

Min Conc:

# Днес ще разгледаме:

## КАЛИБРИРАНЕ ПРИ ХИМИЧНИТЕ ИЗМЕРВАНИЯ

### 1. Корелация и регресия.

Зависимост и независимост на две случайни величини

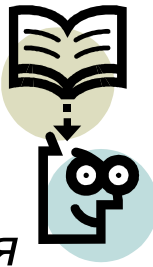
- коефициент на корелация

2. Метод на най-малките квадрати - намиране на уравнението на регресионна права

### 3. Калибрация.

- калибрационни стандарти - ССМ и еталони, проследимост
- калибрационна права - приложение

# КАЛИБРИРАНЕ



Калибриране:

Съвкупност от операции, които при определени условия установяват, зависимост между стойностите, отчетени от средството за измерване и съответните известни стойности на измерваната величина (стойности реализирани от еталоните или стандартите).

Трябва да бъде изпълнено с помощта на сравнителни материали с доказана **проследимост** и достатъчно малка **неопределеност**.

**Проследимостта е свойство** на резултата от измерване или стойността на еталон, които могат да бъдат свързани с установени еталони (обикновено национални или международни), чрез една непрекъсната верига от сравнения, които имат обявени неопределености.

[VIM, 6.10]

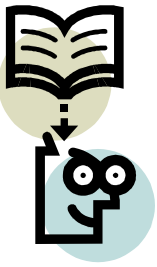
## Вие вече сте чували за това...

### ✓ Еталони

- Първичен и вторичен еталон
- Международен (измервателен) еталон
- Национален еталон
- Еталон за калибриране
- Измервателен еталон

### ✓ Сравнителен материал (ССМ, ЕСМ (еталонен СМ), ...)

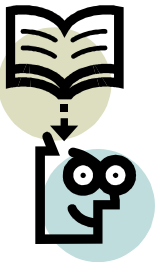
- Лабораторен СМ
  - Вътрешен, “лабораторен” СМ
  - Матричен СМ
- + Калибрационен разтвор  
+ Контролен образец



## Дефиниция за сравнителен материал (СМ)

*В съответствие с VIM*

Материал или вещество, на което едно или повече характеристики са достатъчно **ХОМОГЕННИ** и добре установени, за да се използват за калибриране на апарат за измерване, за оценяване на метод за измерване или за приписване на стойности на материали



## Дефиниция за сертифициран сравнителен материал (ССМ)

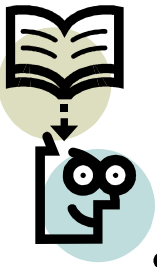
*В съответствие с VIM,*

Сравнителен материал, придружен със **сертификат (свидетелство)**, на който една или повече стойности на характеристиките му са сертифицирани по процедура, която установява проследимост към точното реализиране на единицата, в която се изразяват стойностите на характеристиките, като всяка сертифицирана стойност се придружава от неопределеност с обявена доверителна вероятност.



## Типове ССМ в съответствие с тяхното използване

- Чисти вещества за калибриране (напр. разтвор на Pb за приготвяне на калибрационен разтвор за AAS)
- Чисти вещества за напасване на матрици (напр. Си с висока чистота за приготвяне на Zn/Si серии за калибриране на искров емисионен анализ)
- матрични ССМ  
(напр. холестерол в серум. )

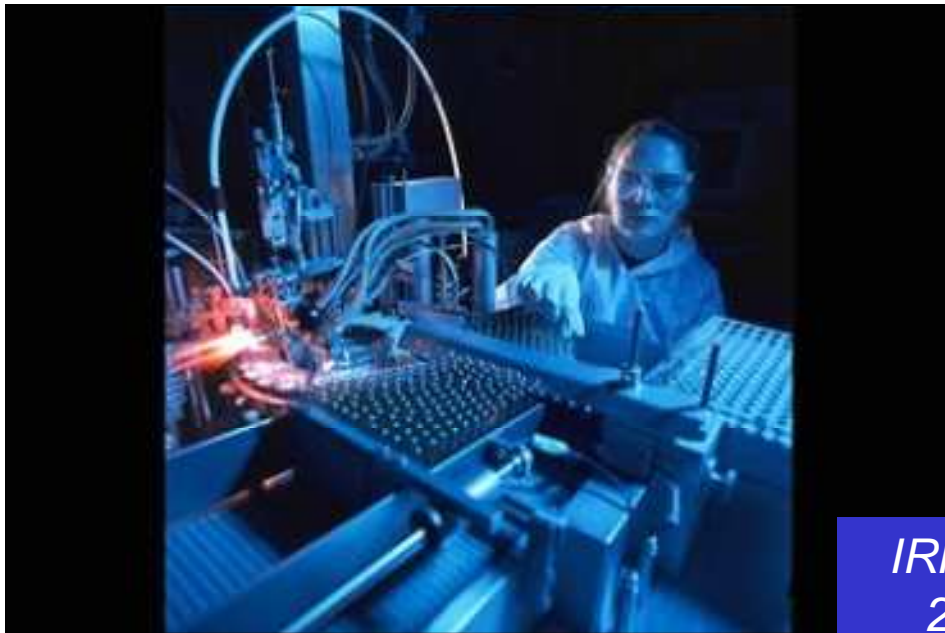


## Как работим/манипулираме със ССМ ?

- Следваме “инструкцията за използване”, дадена от доставчика/производителя
- **Съобразяване с предписания минимум образец за използване**
- Съобразяване с предписаната температурата за съхранение (-20, +4, +18 °C ?)
- Съобразяване с предписаната влажност на пробата, както и поемането на влага от пробата ( т.нар. биологични дейности)
- Избягване на замърсяванията
- Прилагане на предписания метод съгласно протокола



IRMM-RM  
2003 ©



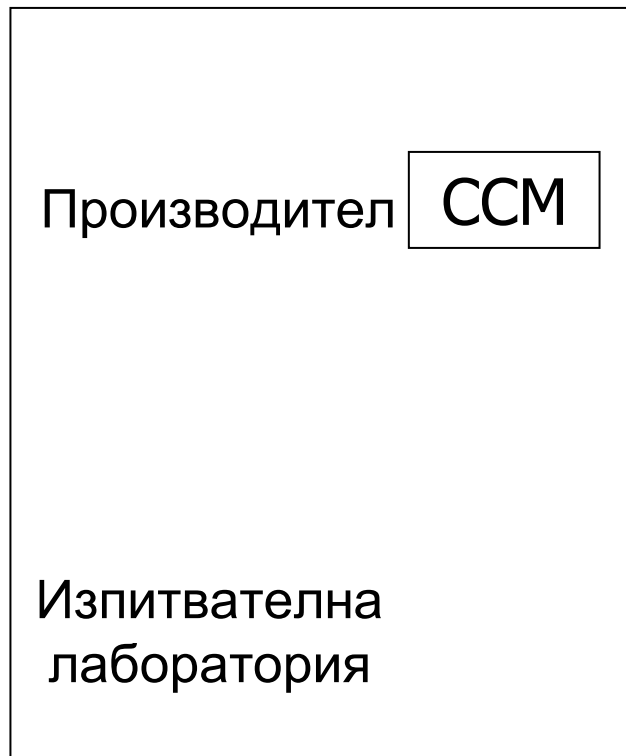
*IRMM-RM*  
2003 ©

# ССМ - сертифицирани сравнителни материали

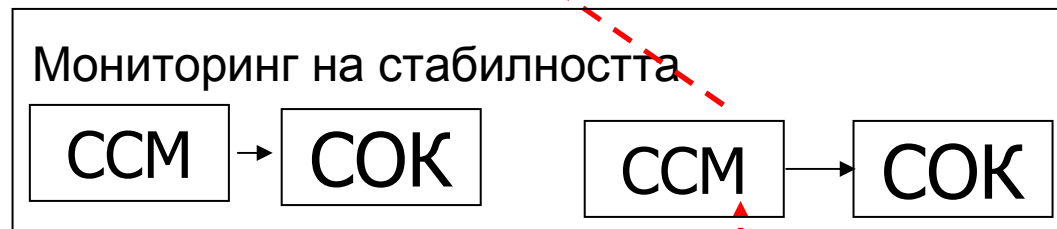
## Времева линия на проследимостта



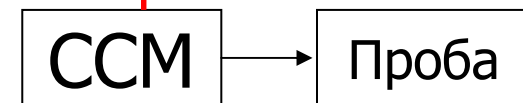
*Проследим до SI чрез  
Международно признати възможности за измерване - например NIST*



ССМ Производство и Сертификация



*Стойност и разпространение на  
неопределеността*



*Проследимост*



# inorganic ventures / iv labs

195 lehigh avenue, suite 4, lakewood, nj 08701 usa  
phone: 800-669-6799 • 732-901-1900 • fax: 732-901-1903  
e-mail: ivsales@ivstandards.com • website: www.ivstandards.com

## certificate of analysis

**1.0 Inorganic Ventures / IV Labs is an ISO Guide 34-2000 Certified Reference Material (CRM) Manufacturer: Certificate #883-02.** The certificate is designed and the certified value(s) and uncertainty(ies) are determined in accordance with ISO Guide 31-2000 (Reference Materials - Contents of certificates and label(s)), ISO Guide 34-2000 "Quality System Guidelines for the Production of Reference Materials," and ISO Guide 35-1989 "Certification of Reference Materials - General and Statistical Principles."

**2.0 DESCRIPTION OF CRM** Custom-Grade 10000 µg/mL Lead in 0.35% (abs) HNO<sub>3</sub>

Catalog Number: CGPB10-1 and CGPB10-5  
Lot Number: T-PB02143  
Starting Material: Pb(NO<sub>3</sub>)<sub>2</sub>  
Starting Material Purity (%): 99.99997  
Starting Material Lot No 22150  
Matrix: 0.35% (abs) HNO<sub>3</sub>

### 3.0 CERTIFIED VALUES AND UNCERTAINTIES

**Certified Concentration:** 10,009 ± 22 µg/mL  
**Certified Density:** 1.014 g/mL (measured at 22° C)

The Certified Value is based upon the most precise method used to analyze this CRM. The following equations are used in the calculation of the certified value and the uncertainty:

$$\text{Certified Value } (\bar{x}) = \frac{\sum x_i}{n}$$

$$\text{Uncertainty } (\pm) = \frac{2t(\sum s_i)^2}{(n)^2}$$

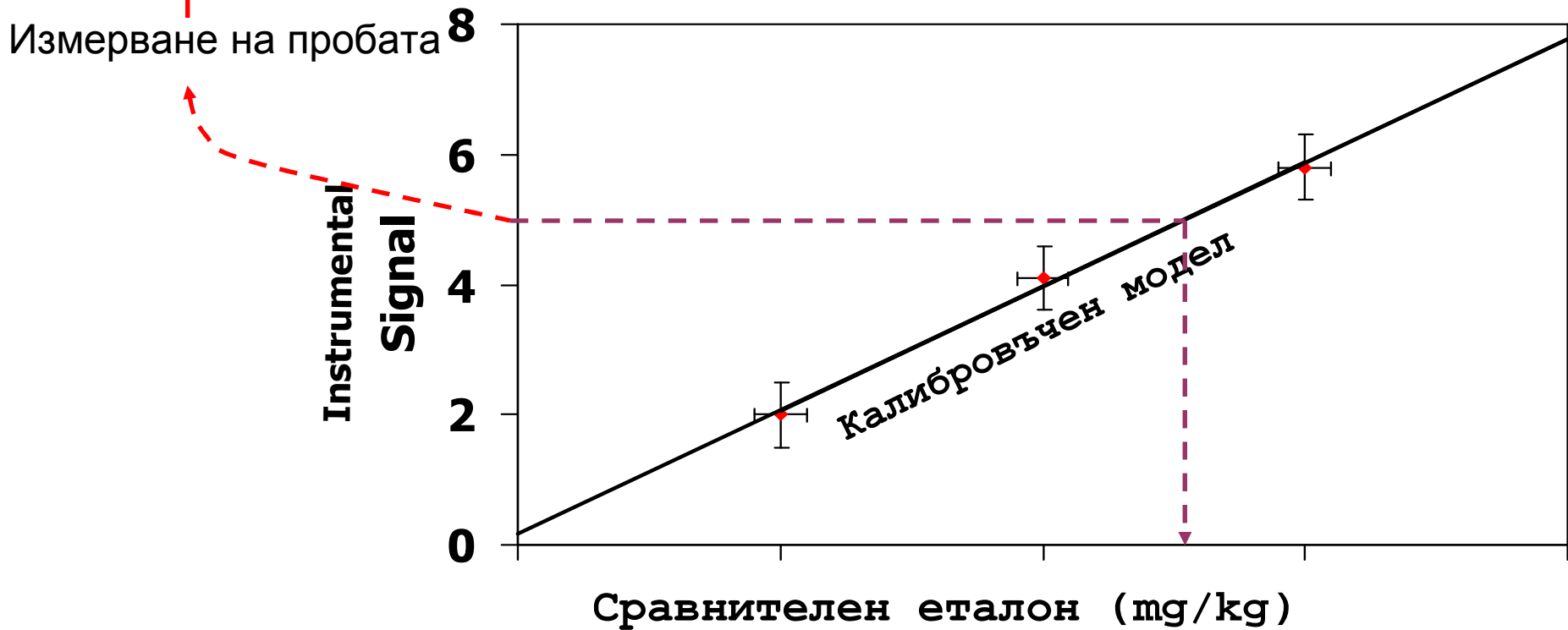
( $\bar{x}$ ) = mean  
 $x_i$  = individual results  
 $n$  = number of measurements  
 $\sum s_i$  = The summation of all significant estimated errors.  
(Most common are the errors from instrumental measurement, weighing, dilution to volume, and the fixed error reported on the NIST SRM certificate of analysis.)

The independent samples t-test was used to determine if there is agreement between the above assay methods at the 95% confidence interval. Both methods were compared and showed agreement within the stated uncertainties. This agreement is a confirmation of the accuracy of this CRM.

### 4.0 TRACEABILITY TO NIST AND VALUES OBTAINED BY INDEPENDENT METHODS

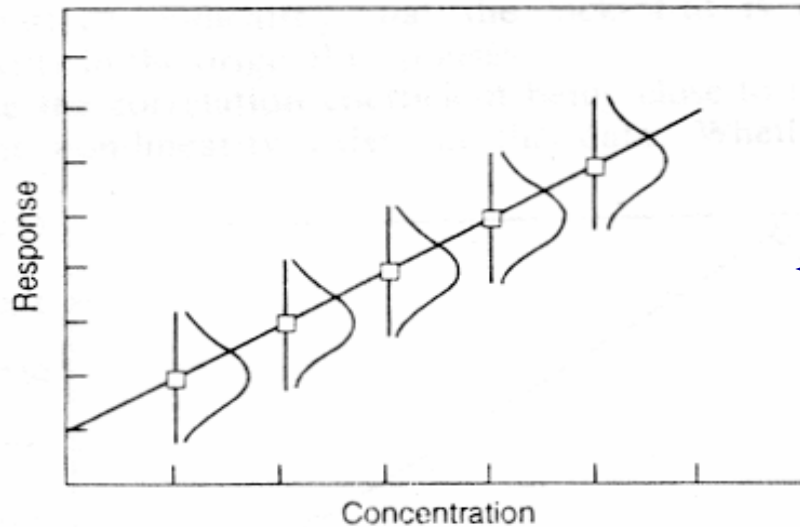
Резултат проследим чрез стойността  
на сравнителния стандарт

Проследим до? Чрез?



Проследим до? Чрез?

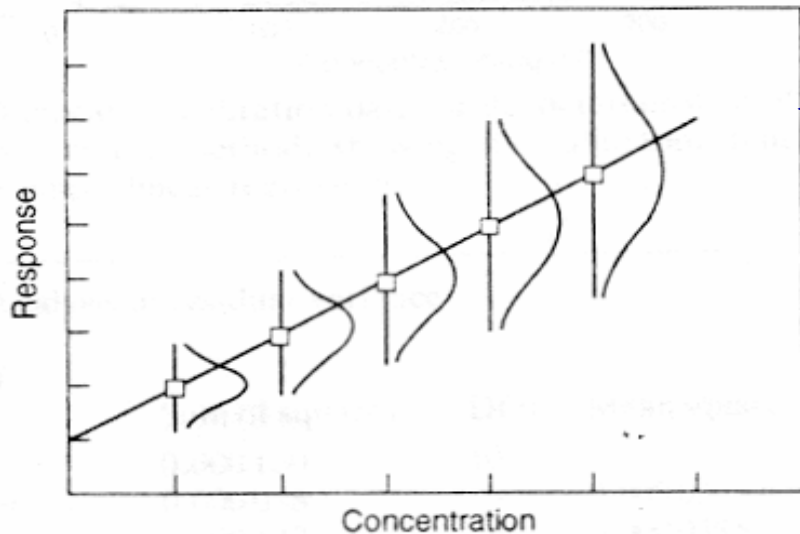
# Претеглен метод



Хомоскедастичен модел

МНМК  
unweighted

Fig. 1 Model for linear regression (homoscedastic case), showing distributions of observations around the true calibration function.



Хетероскедастичен модел

Претеглена регресия  
МНМК  
weighted

Fig. 2 Model for linear regression (heteroscedastic case), showing distributions of observations around the true calibration function.

$$b_1 = \frac{\sum X_K W_k \sum Y_K W_k - \sum W_k \sum X_K Y_K W_k}{(\sum X_K W_k)^2 - \sum W_k \sum X_K^2 W_k}$$

$$b_0 = \frac{(\sum X_K Y_K W_k) \sum X_K W_k - \sum X_K^2 W_k \sum Y_K W_k}{(\sum X_K W_k)^2 - \sum W_k \sum X_K^2 W_k}$$



**Днес разгледахме:**

## **КАЛИБРИРАНЕ ПРИ ХИМИЧНИТЕ ИЗМЕРВАНИЯ**

### **1. Корелация и регресия.**

Зависимост и независимост на две случайни величини

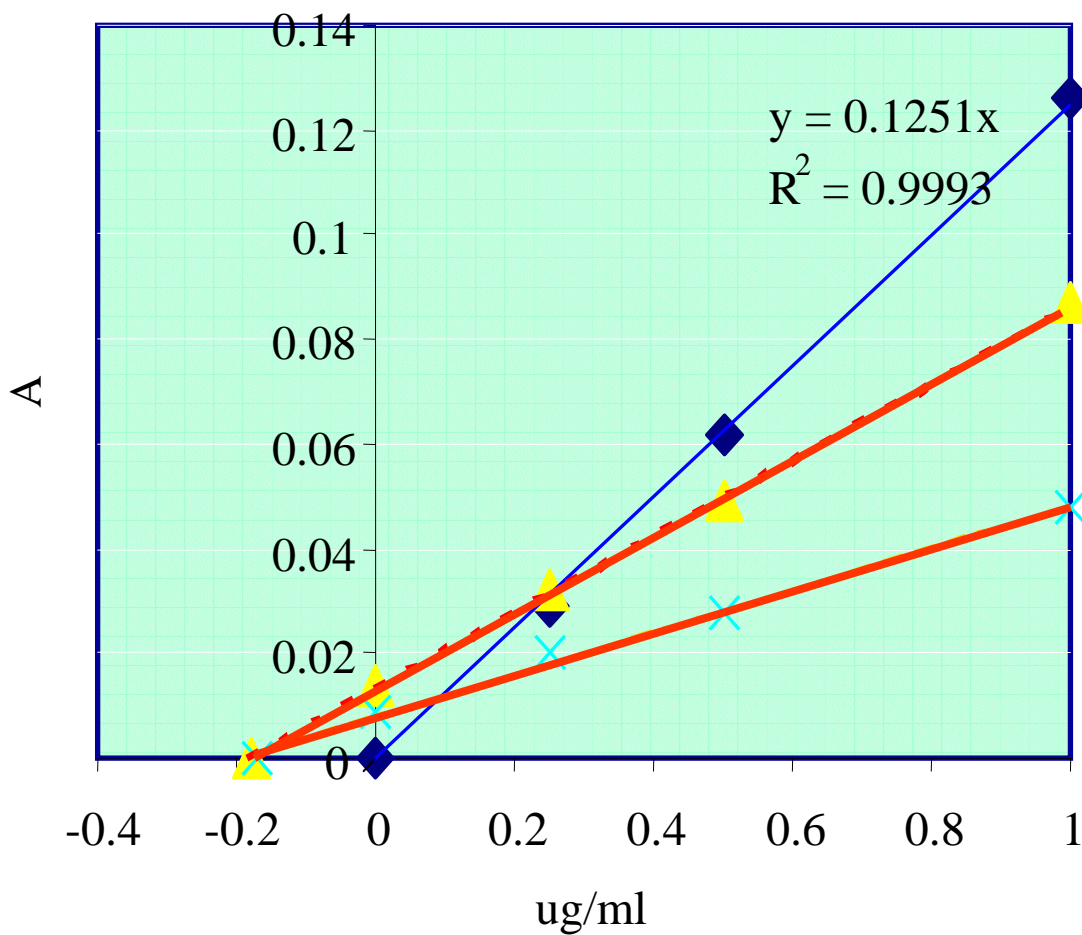
- коефициент на корелация

**2. Метод на най-малките квадрати** - намиране на уравнението на регресионна права

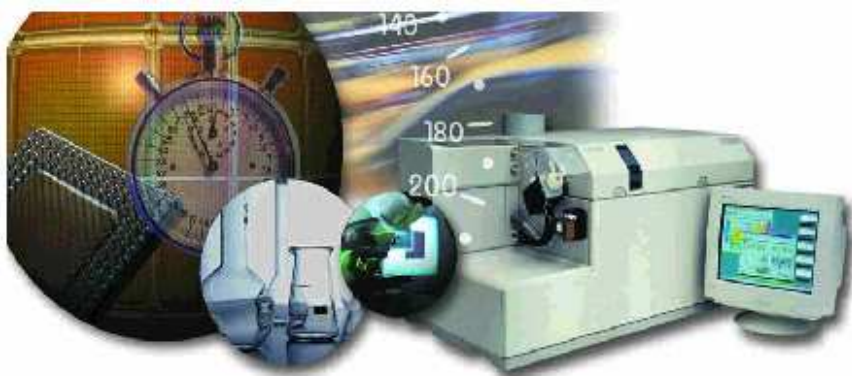
### **3. Калибрация.**

- калибрационни стандарти - ССМ и еталони, проследимост
- калибрационна права - приложение

лекция № 11



# КАЛИБРИРАНЕ ПРИ ХИМИЧНИТЕ ИЗМЕРВАНИЯ II част МАТРИЧЕН ЕФЕКТ



# Днес ще разгледаме:

## КАЛИБРИРАНЕ ПРИ ХИМИЧНИТЕ ИЗМЕРВАНИЯ ПРОДЪЛЖЕНИЕ

### 1. Регресионна права – продължение

- Преминаване от сигнали към стойности на измервана величина
- чувствителност; работна област; линеен динамичен диапазон
- изчисляване на границите на откриване и определяне в домейн на измервана величина.

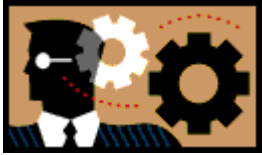
### 2. Матричен ефект

- мултипликативно и адитивно пречене
- дрейф на чувствителността

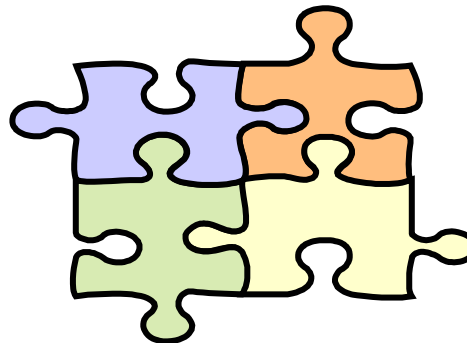
### 3. Методи за намаляване и коригиране на матричния ефект

- Метод на стандартната добавка
- Метод на вътрешния стандарт

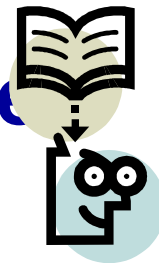
# ПРИПОМНЕТЕ СИ !



- Какво е относителен метод ?
- Какво е калибриране ?
- Какво ви е нужно за да извършите калибриране?
- Как ще проведете процедурата на калибриране и какво ще получите след нейното провеждане?



# Типове методологично калибриране



- ОТНОСИТЕЛНИ МЕТОДИ

- Инди́ректно или въ́ншно калибриране

- Пробата с неизвестното количество аналит и стандартите с известно количество аналит са в различни съдове или са отделни препарати.

- Метод на вътрешния стандарт

- преди измерване в пробата се добавя фиксирано количество различен но подобен на анализа компонент (вътрешен стандарт).

- Метод на стандартната добавка

- Аналита сам по себе си действа като вътрешен стандарт.

## ВАЖНИ СТЪПКИ



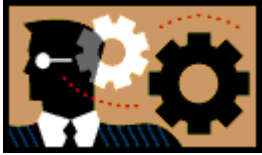
- Избира се вида на модела (например линеен).
- Прави се план на експеримента.
- Извършват се измерванията на еталоните.
- Изчисляват се оценките на параметрите на модела.
- За валидиране на модела (analysis of variance and/or analysis of residuals).
- Изчисляват се доверителните граници на параметрите на регресията
- Измерва се сигнала на изследваната проба и от регресионното уравнение се изчислява обратната задача - за определяне концентрацията (съдържанието) на търсения аналит.

# Критични аспекти при методологично калибриране

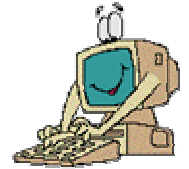
Трябва да се има в предвид:

- Калибрационните стандарти да покриват целия диапазон от концентрации за съответните измервания. Неизвестните концентрации да се определят чрез **интерполация** а не чрез **екстраполация**.
- Изключително важно е да се отчете стойността на празната проба в калибрационната крива.
- Приема се че неопределеността е само по посока на **у**-стойностите.
- Обикновено регресионните методи приемат - нормално Гаусово разпределение на СВ стойности по **у**.

# ВЪПРОС ?



- Приготвили сте стандартни разтвори
- Измерили сте техните сигнали
- Изчислили сте регресионните параметри и сте построили регресионната права
- Измервате сигнала на неизвестна проба -

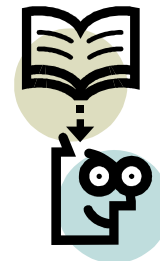


## КАК ЩЕ ОПРЕДЕЛИТЕ СЪДЪРЖАНИЕТО НА АНАЛИТА В НЕИЗВЕСТНАТА ПРОБА ?

**St - Na 20 ppm → 40 сигнал**  
**Вода ПУ Na X ppm → 20 сигнал**  
**X=?**

$$C_{sample} = \frac{Y_{C_x} - b_0^{St}}{b_1^{St}}$$





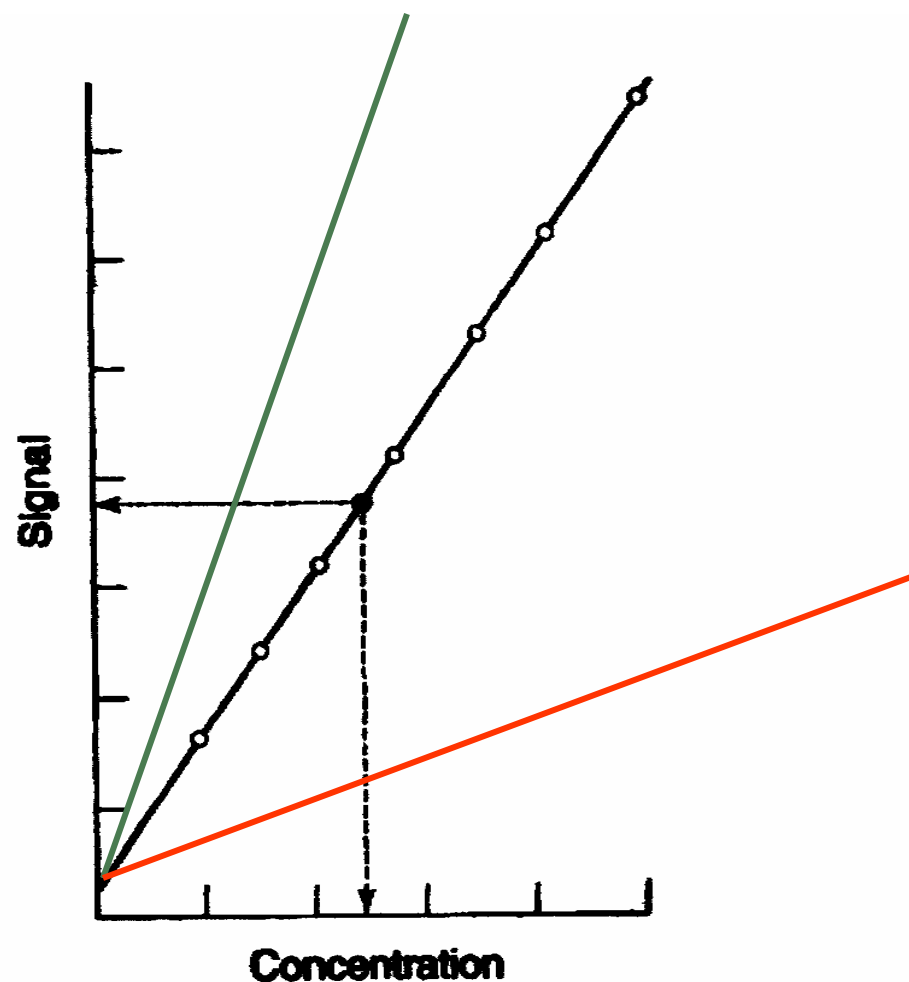
- Дефиниция:

Промяната в сигнала на измервателния прибор, разделена на съответната промяна на измерваното свойство

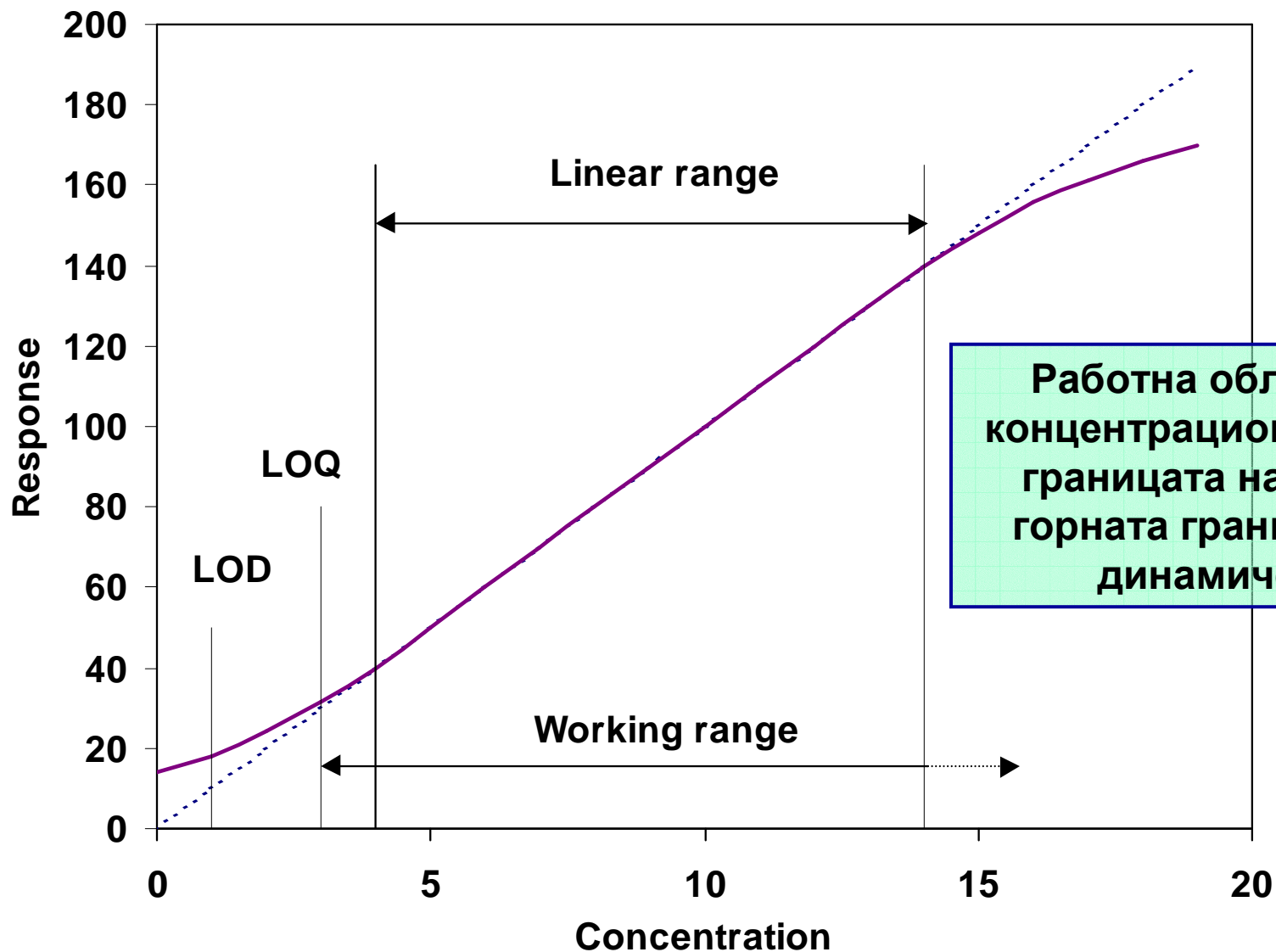
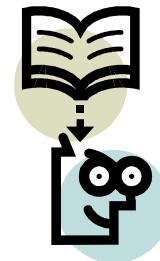
(VIM 1993)

- *Какво означава това:*  
Градиентът (наклонът) на калибрационната права

$$b_1 = m = \frac{d\text{Sig}}{dC} = \text{tg } \alpha$$

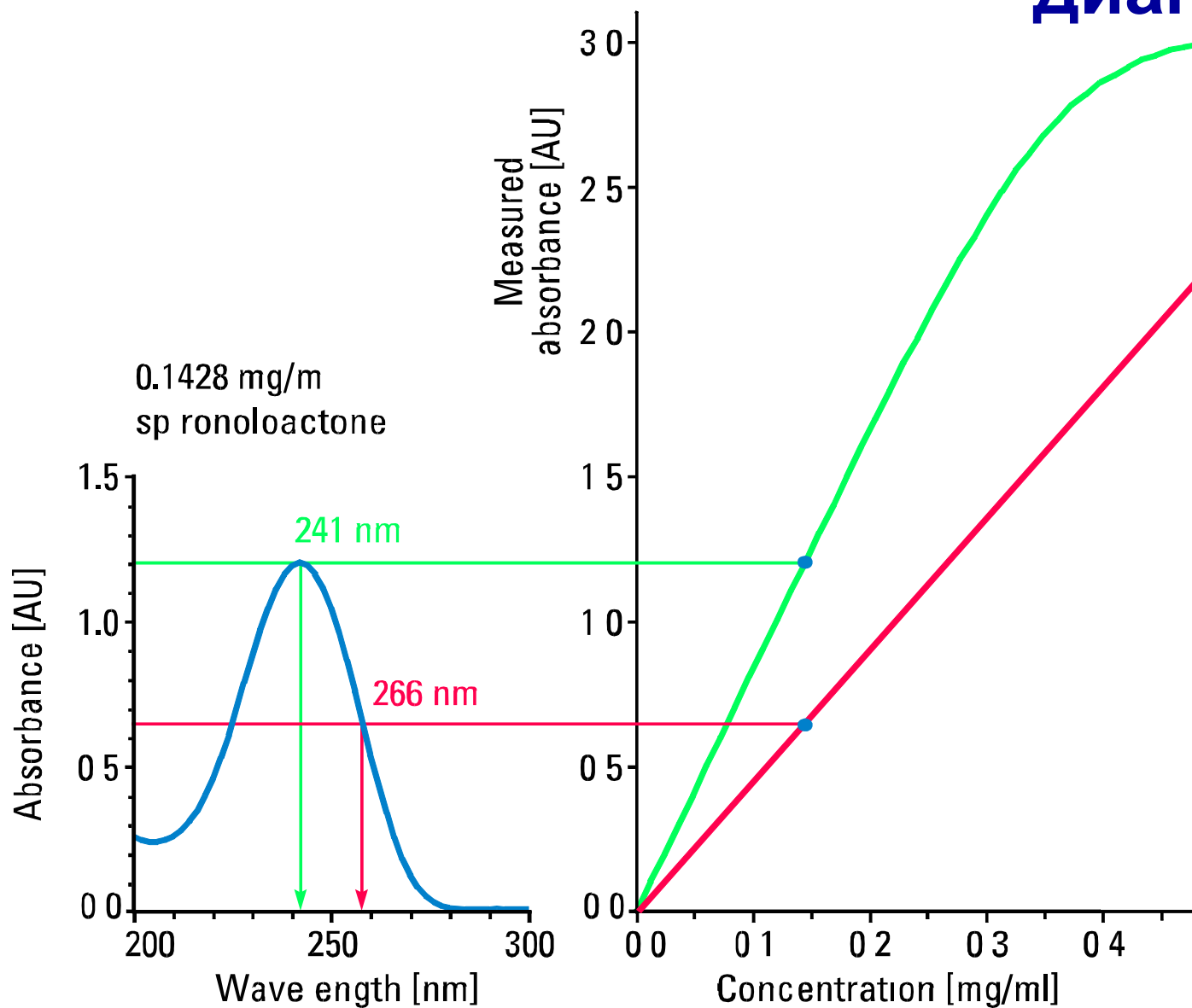


# Линейност Работна област



Работна област на метода -  
концентрационния интервал от  
границата на определяне до  
горната граница на линейния  
динамичен интервал

# Разширяване на динамичния линейен диапазон



Граници на...  
Откриване (LOD)  
Определяне (LOQ)

‘Празна проба’

- от измервателния прибор
- от пробоподготовката  
(напр. замърсяване при разлагането, пречистването)

Калибрационно уравнение:

$$\text{Сигнал} = b_1 * c + b_0$$

$Y_{bl}$  = сигнал на празната проба;

$S_{bl}$  = стандартно отклонение на празната проба в домейн на сигнала

$$Y_{LOD} = Y_{bl} + 3 S_{bl} \rightarrow LOD = (Y_{LOD} - b_0)/b_1$$

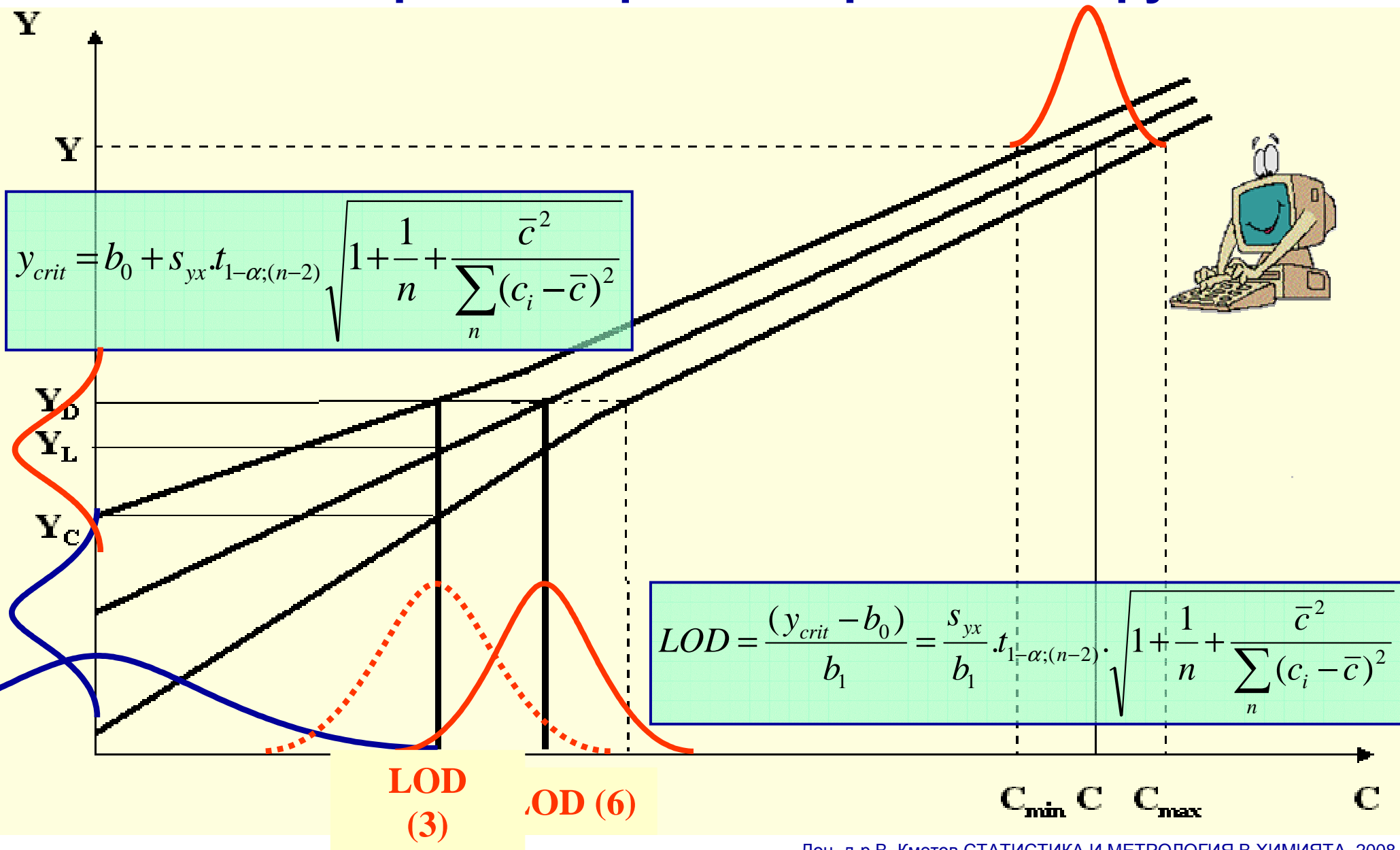
$$Y_{LOQ} = Y_{bl} + 10 S_{bl} \rightarrow LOQ = (Y_{LOQ} - b_0)/b_1$$

При нулиране по  
празна проба и  $b_0=0$

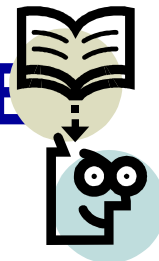
$$LOD (Q) = \frac{k \cdot S_0}{m}$$

- $K = 3$  - граница на ОТКРИВАНЕ
- $K = 6$  - граница на ГАРАНТИРАНО ОТКРИВАНЕ
- $K = 10$  - граница на ОПРЕДЕЛЯНЕ

# Граници на откриване и граница на гарантирано откриване при калибрационна функция



# ТИПОВЕ ГРАНИЦИ НА ОТКРИВАНЕ



- ☒ **ФУНДАМЕНТАЛНА (fundamental LOD)** - при чисти условия
- ☒ **ИНСТРУМЕНТАЛНА (instrumental LOD)** - отчита шума на инструмента и влиянието на матрицата при анализ, но без процедурата за пробоподготовка
- ☒ **МЕТОДИЧНА (methodological LOD)** - отчита абсолютно цялостната аналитична процедура

$$LOD_{метод} = DF \cdot LOD_{инструм}$$

мерни единици на измервана величина – изходна проба

В мерни единици на концентрация на разтвор

# Фактор на разреждане

- Нека изходната проба има съдържание на анализит  $C_{изх.}$ .
- Ако дозираме аликвотна част от изходната проба с обем  $V_{из}$  в колба с обем  $V_{кр.}$  и долеем с разтворител ( $H_2O$ ) то изходната проба ще се разрежи с фактор

$$DF = \frac{V_{кр.}}{V_{из}}$$

Резултантната концентрация при разреждане на разтвори се изчислява съгласно формулите:

$$C_{изх.} = DF \cdot C_{кр.}$$

$$DF = \frac{C_{изх.}}{C_{кр.}}$$

$$C_{изх.} \cdot V_{из} = C_{кр.} \cdot V_{кр.}$$

$$C_{изх.} = \frac{C_{кр.} \cdot V_{кр.}}{V_{из}}$$

= DF

# Днес ще разгледаме:

## КАЛИБРИРАНЕ ПРИ ХИМИЧНИТЕ ИЗМЕРВАНИЯ ПРОДЪЛЖЕНИЕ

### 1. Регресионна права – продължение

- Преминаване от сигнали към стойности на измервана величина
- чувствителност; работна област; линеен динамичен диапазон
- изчисляване на границите на откриване и определяне в домейн на измервана величина.

### 2. Матричен ефект

- мултипликативно и адитивно пречене
- дрейф на чувствителността

### 3. Методи за намаляване и коригиране на матричния ефект

- Метод на стандартната добавка
- Метод на вътрешния стандарт



## Основни понятия

В химичния анализ по-пълното понятие е **ИЗПИТВАНЕ** което представя комплексността при оценката на ПРОВОКАЦИОННО-РЕСПОНСНАТА-ПОВЕДЕНЧЕСКА СХЕМА

‘Измерване’: определяне стойността на величина

‘Измервана величина’: величина, която се измерва

‘Определяем компонент’ **АНАЛИТ**: веществото, химичната форма, която се измерва

‘Модел’: уравнението, което използвате, за да се изчисли крайният резултат (*вие винаги използвате такова!*).

Този модел е (*приблизително*) описание на реалността

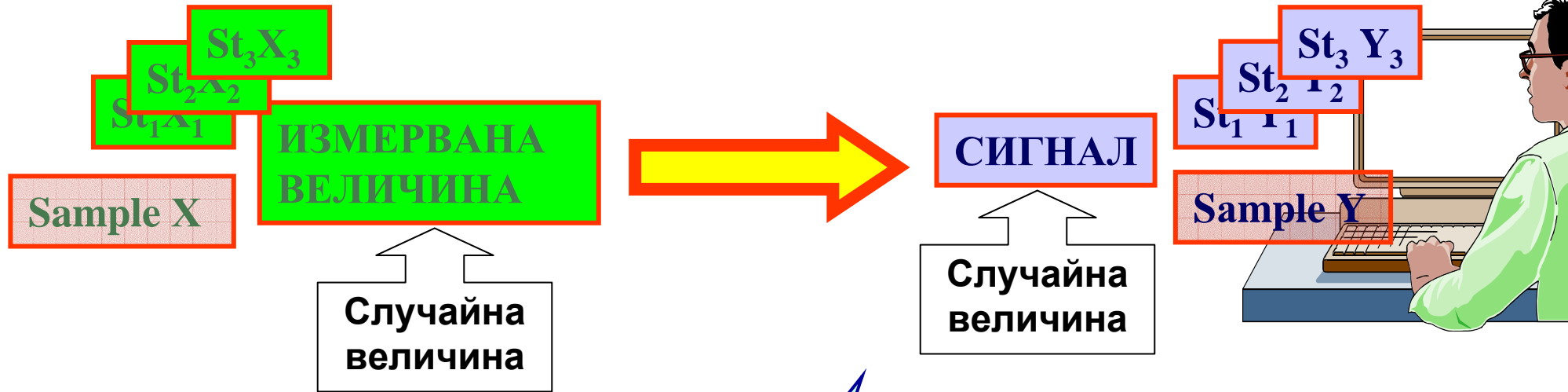
[VIM, 1993]

Измерваната величина е характеристика  
на **аналита** в **КОНКРЕТНО ОБКРЪЖЕНИЕ**  
- **МАТРИЦА**

## Общ език

Величина	Опр.комп	Изм.величина	Единици	Обявен еталон
Концентрация	DDT	$C_{\text{DDT}}$ в почва	ng/kg	SI
Количество/ съдържание	Pb	w(Pb) в отпадна вода	ng/l	SI
Брой	E.Coli	брой E.Coli/повърх.	m <sup>-2</sup>	SI
АКТИВНОСТ	Амилаза	A(Амилаза)	Катал	SI
pH	H <sup>+</sup> йони	C (H <sup>+</sup> ) в питейна вода	pH един.	pH скала
Твърдост			Mohs твърдост	Mohs скала
Октаново число			октаново число	Скала на ОКТ. ЧИСЛО

# АНАЛИЗЪТ Е СРАВНЕНИЕ !!!!



Всяка стойност се разкрива при определено **въздействие** върху околния свят.

Стойността на величината се проявява при връзката на тази величина с други величини или при **сравнението** на въздействието и с това на еталони или стандарти.

**ПРПС**  
 ПРОВОКАЦИОННО-  
 РЕСПОНСНА  
 ПОВЕДЕНЧЕСКА СХЕМА

$$Y = f(X)$$

**ПРЕЧЕНЕ – когато :**

$$f_{St} \neq f'_{проба}$$

# ВЛИЯНИЕ НА МАТРИЦАТА



**ПРЕЧЕНЕ** - промяна, или симулиране на въздействие на измерваната величина не причинено от анализа.

$$Y = f_{st}(X) \neq f_{matrix}(X)$$

$$Y = b_{1st} X + b_{0st} \neq b_{1matrix} X + b_{0matrix}$$

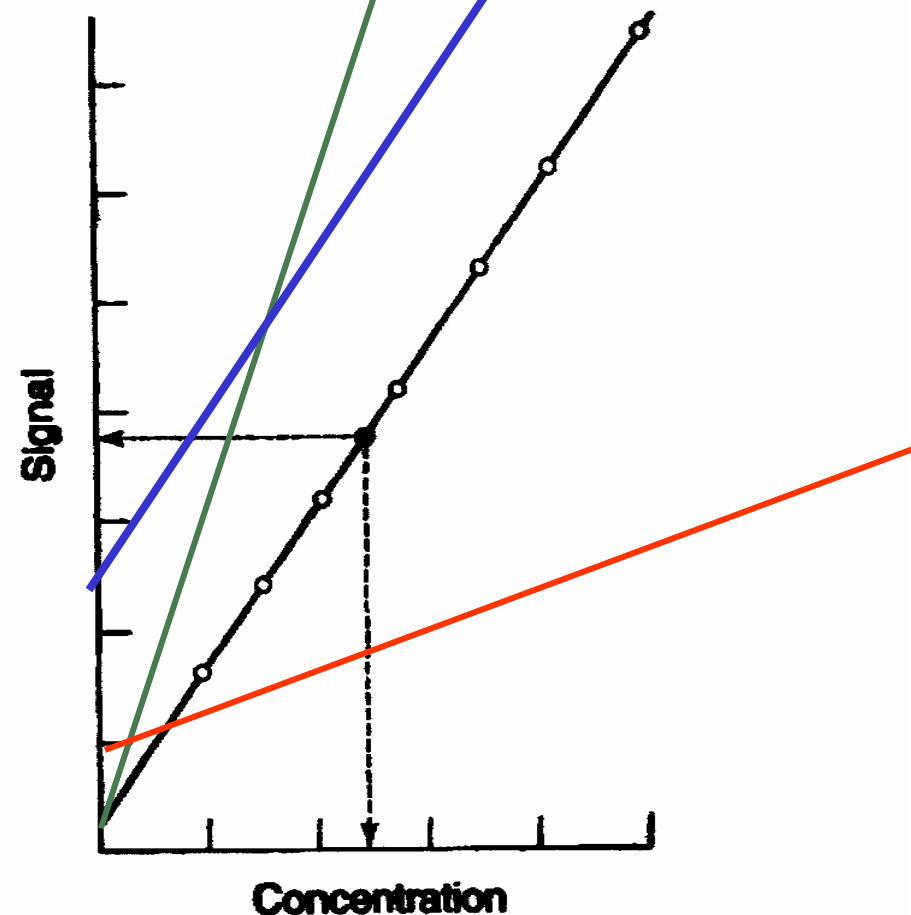
**МАТРИЧЕН ЕФЕКТ** - значимо въздействие на обкръжението върху поведението на анализа.

**Адитивно пречене** - матрицата добавя постоянен сигнал към този на анализа.

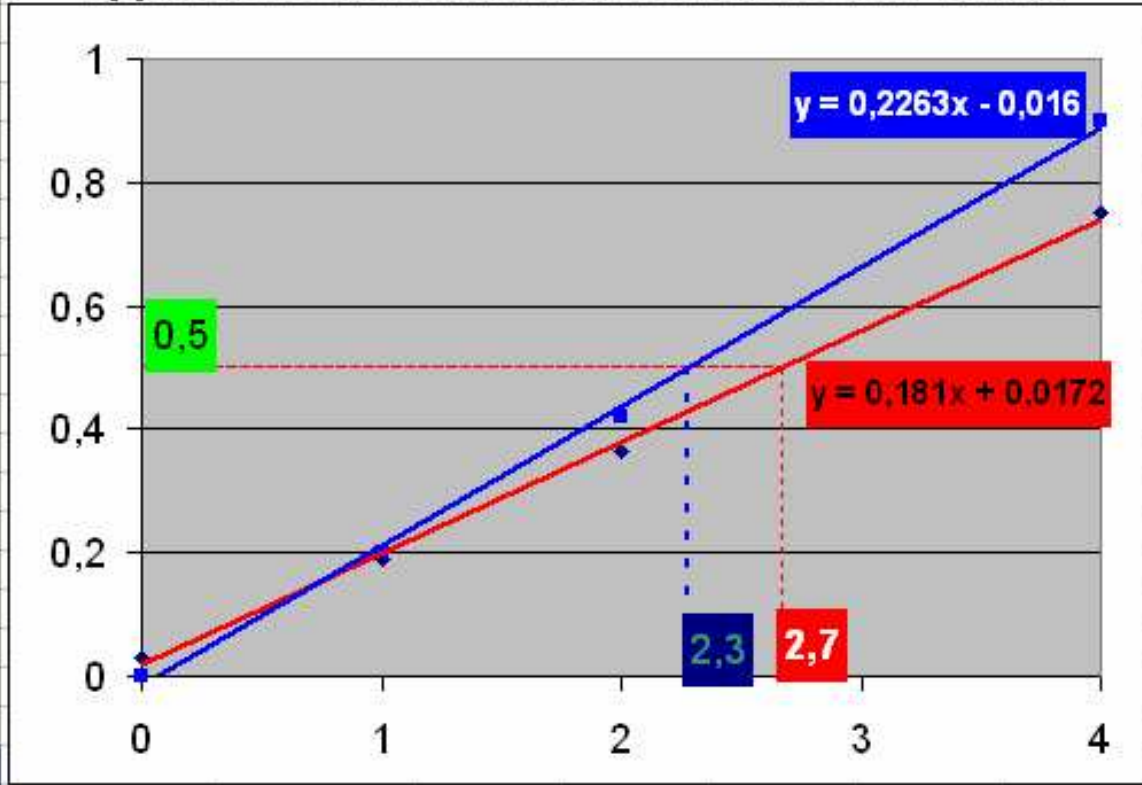
$$b_0' \neq b_0$$

**Мультипликативно** - матрицата изменя величината на сигнала на анализа изменяйки чувствителността.

$$b_1' \neq b_1$$

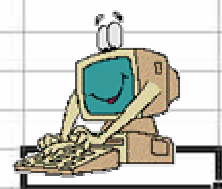


	A	B	C	D	E	F	G	H	I	J	K	L	
1	<b>ВИДОВЕ ПРЕЧЕНЕ</b>												
2		Multiplicative		Y St	X	Y matrix							
3		8		0	0	0,03			50				
4		0,181028571		0,2	1	0,19		Y samp	X	Y1	X	Y2	
5				0,42	2	0,366	0,1	0,5	0,01	0,5	2,666982	0	
6		Additive		0,9	4	0,75			2,666982	0,5	2,666982	0,5	
7		3		0,226286	b1 и b1m	0,181029		Y samp	X	Y1	X	Y2	
8		0,0172		-0,016	b0 и b0m	0,0172		0,5	0,01	0,5	2,280303	0	
9	<b>АДИТИВНО И МУЛТИПЛИКАТИВНО ПРЕЧЕНЕ</b>								2,280303	0,5	2,280303	0,5	



$$C_{sample} = \frac{Y_{C_x} - b_0^{St}}{b_1^{St}}$$

$$C_{sample} = \frac{Y_{C_x} - b_0^{matrix}}{b_1^{matrix}}$$



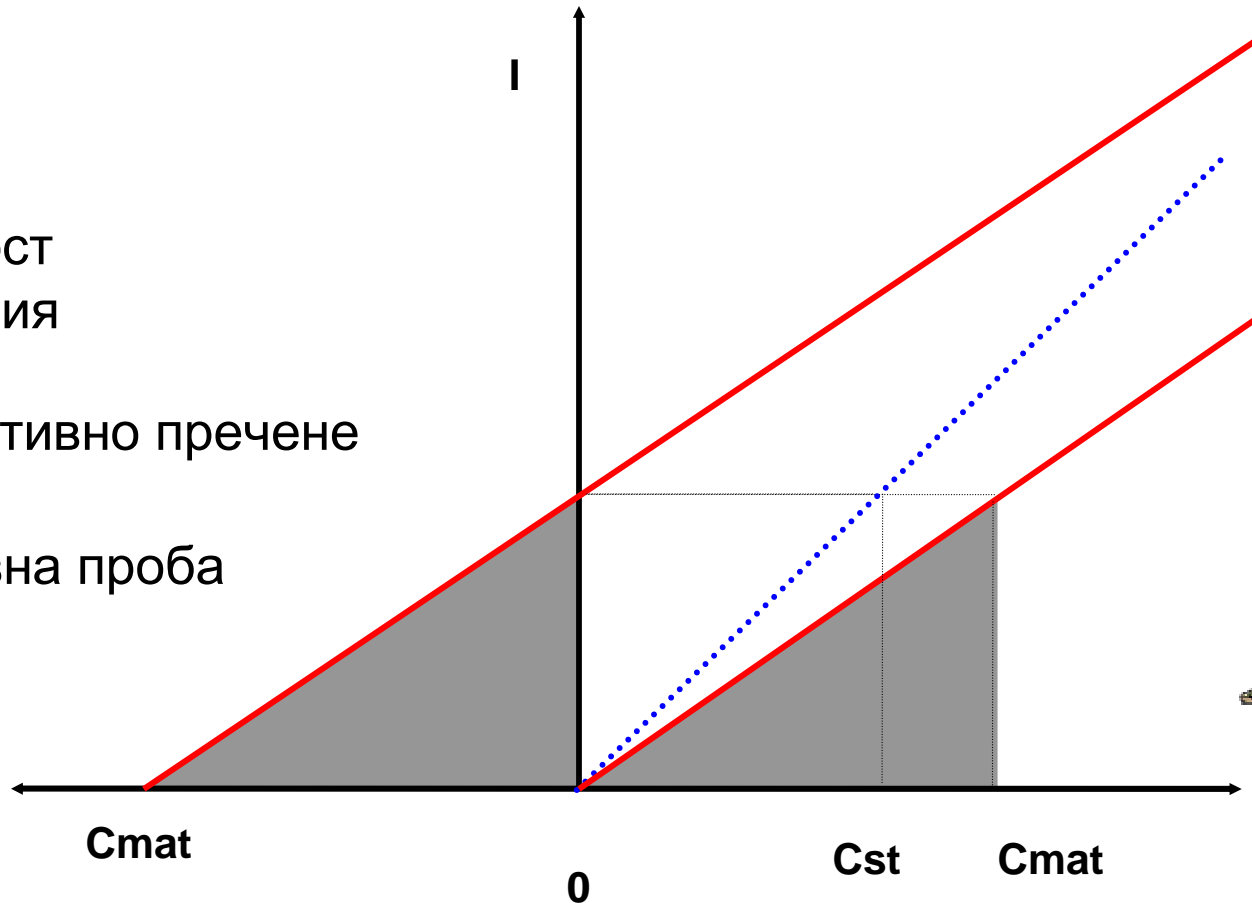
# Метод на стандартната добавка

## ИЗИСКВАНИЯ

линейна зависимост  
сигнал/концентрация

- отсъствие на адитивно пречене
- нулиране по празна проба

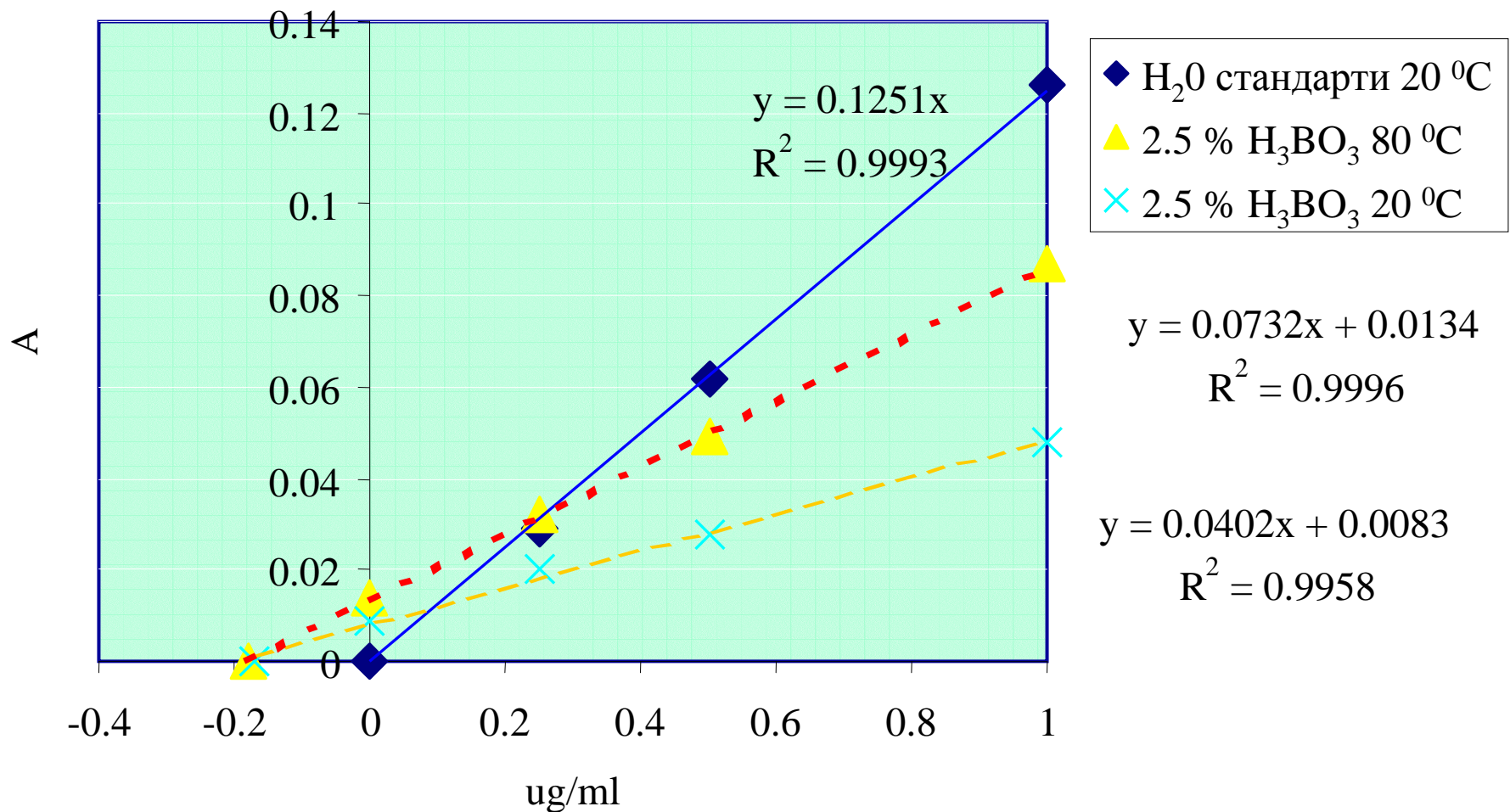
$$C_x = \frac{b_0^{\text{проба}}}{b_1^{\text{матрица}}}$$



Методът на стандартната добавка отчита мултипликативно пречене и е неприложим при наличие на адитивно пречене

# Метод на стандартната добавка

## Определяне на Си в борна киселина



# Метод на вътрешния стандарт

- В разтворите на празна проба, проби и на стандарти се внася едно едно и също количество (концентрация) от подобен на анализа компонент – наречен вътрешен стандарт (internal standard)

$$C_{IS} = C_{IS}^{Bl} = C_{IS}^{samp} = C_{IS}^{St1} = C_{IS}^{St2} \dots$$

- Измерва се резултатен сигнал (ratio)  $R_{sig}$  представляващ отношението на сигналите – на анализа  $Sig(A)$  и вътрешния стандарт  $Sig(IS)$ :

$$R_{sig} = \frac{Sig_A}{Sig_{IS}}$$

- Регресионната зависимост се построява на база
  - $X$  - концентрация на анализа ;
  - $Y$  – резултатния сигнал  $R_{sig}$

$$R_{sig} = b_1'' C_A + b_0''$$

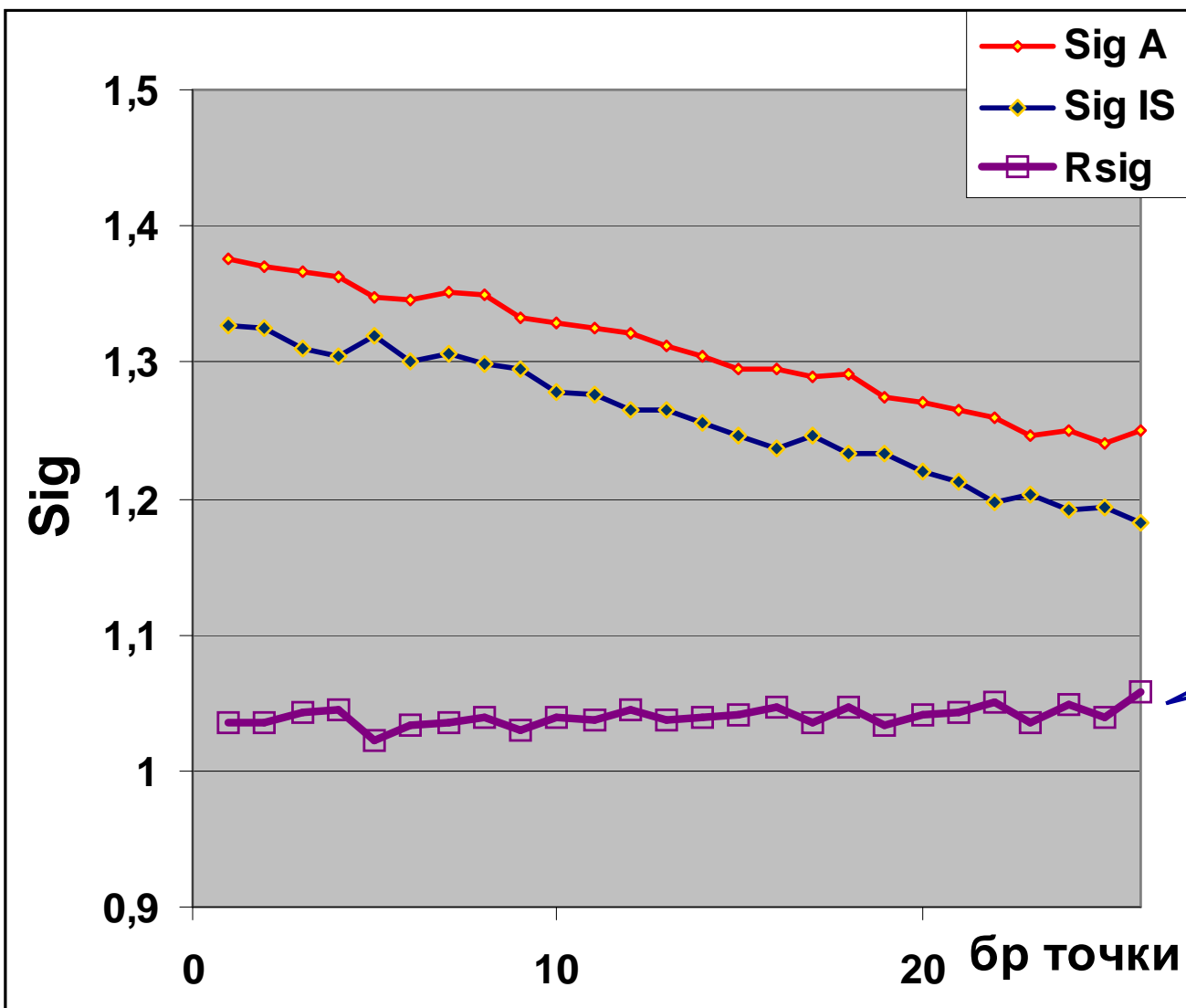


# Метод на вътрешния стандарт II

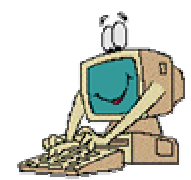
- Ако е изпълнено условието – МАТРИЦАТА действа еднакво на ПРПС на анализа и на вътрешния стандарт то  $R_{sig}$  е независим от матрицата параметър при което следва, че изчислените по ММК параметри  $b_1''$  и  $b_0''$  са също независими от матричния ефект.
- СЛЕДОВАТЕЛНО -

$$R_{sig} = b_{1St}'' X + b_{0St}'' = b_{1matrix}'' X + b_{0matrix}''$$

# Елиминиране на дрейф на чувствителността чрез метода на вътрешния стандарт



*Ако дрейфът на чувствителността е еднакъв за анализа и за вътрешния стандарт, то **Rsig** остава постоянен с времето*



# Днес разгледахме:

## КАЛИБРИРАНЕ ПРИ ХИМИЧНИТЕ ИЗМЕРВАНИЯ ПРОДЪЛЖЕНИЕ

### 1. Регресионна права – продължение

- Преминаване от сигнали към стойности на измервана величина
- чувствителност; работна област; линеен динамичен диапазон
- изчисляване на границите на откриване и определяне в домейн на измервана величина.

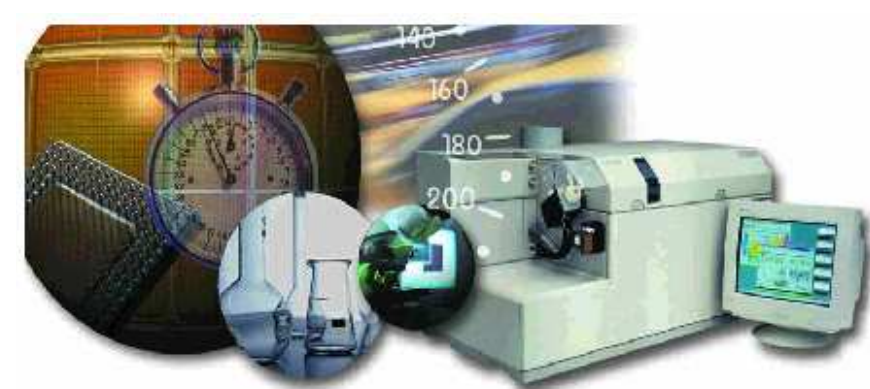
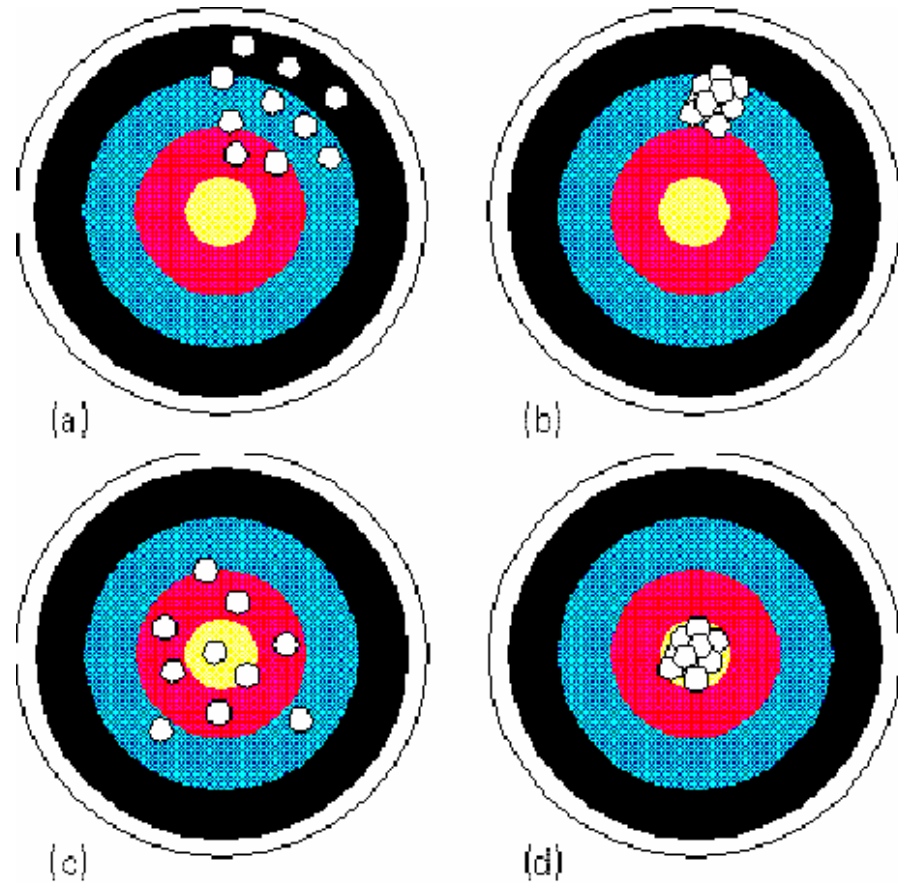
### 2. Матричен ефект

- мултипликативно и адитивно пречене
- дрейф на чувствителността

### 3. Методи за намаляване и коригиране на матричния ефект

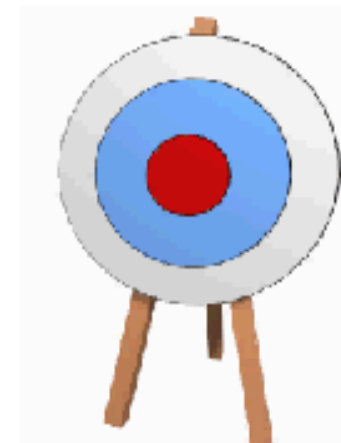
- Метод на стандартната добавка
- Метод на вътрешния стандарт

# Видове грешки в химичния експеримент ОЦЕНЯВАНЕ на ТОЧНОСТТА



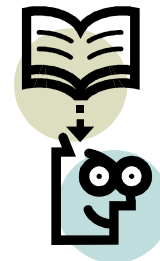
# Днес ще разгледаме:


## ГРЕШКИ В ХИМИЧНИЯ ЕКСПЕРИМЕНТ



1. Точност и прецизност на резултатите от измерване
2. Видове грешки.
  - Случайна и систематична грешка.
  - Оценка на систематичната грешка. Постоянна и променлива систематична грешка. Причини.
3. Оценка на случайната грешка - НЕОПРЕДЕЛЕНОСТ .
  - Прецизност - възпроизводимост - повторяемост.
  - Относителни и абсолютни грешки.
4. Оценка на точността
  - Сравнение със ССМ
  - Участие в междулабораторни сравнения и тестове за пригодност

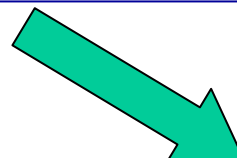
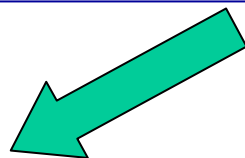
# ТОЧНОСТ И ПРЕЦИЗНОСТ





 **ТОЧНОСТ** на измерване (ACCURACY) - степен на съвпадение на резултат от измерване с *истинската стойност* на измерваната величина. **(НЕ Е прецизност)** точността е качествено понятие.

≠

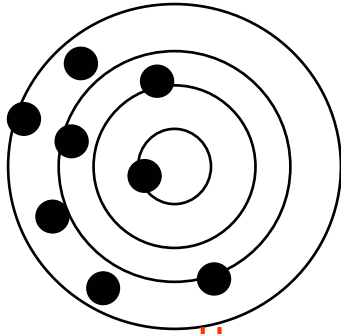
 **ПРЕЦИЗНОСТ (precision)** - характеристика на разсейването на резултатите



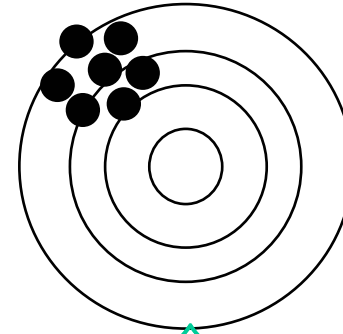
 **ПОВТОРЯЕМОСТ repeatability** - степен на съвпадение на резултатите от последователни измервания на една и съща СВ извършени при ЕДНИ И СЪЩИ условия на измерване

 **ВЪЗПРОИЗВОДИМОСТ reproducibility** - степен на съвпадение на резултатите от последователни измервания на една и съща СВ извършени при ПРОМЕНЕНИ условия на измерване

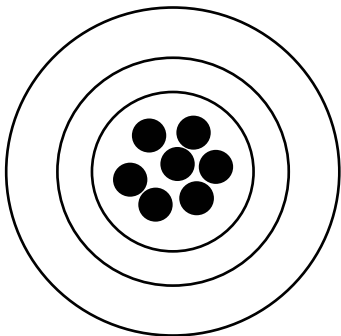
# ТОЧНОСТ И ПРЕЦИЗНОСТ (2)



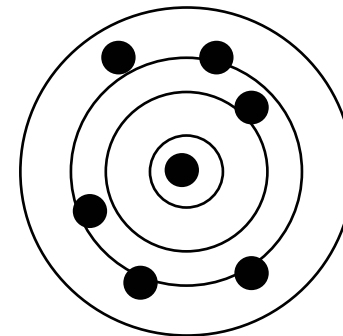
↓ точност , ↓ прецизност



↓ точност , ↑ прецизност



↑ точност , ↑ прецизност



↑ точност , ↓ прецизност



**ДАЙТЕ ПРИМЕР ЗА ПОВТОРЯЕМОСТ И ВЪЗПРОИЗВОДИМОСТ НА ПОПАДЕНИЯТА НА СТРЕЛЕЦА**

# НАРЕДБА № 12 от 21.05.2002 г. за норми за максимално допустими количества от тежки метали като замърсители в храни

Изда на 21.05.2002 г.

кн. Приложение № 5 към чл. 20, ал. 1

ДОПЪ Изисквания към прецизността (изразена чрез RSDr и RSDR, в %) на методите за определяне съдържанието на олово, кадмий и живак в храни

§ 1. Таблица с данни за прецизността

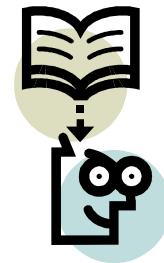
1.	2. Повторяемост			3. Възпроизводимост		
	4. Концентрационен интервал, mg/kg	5. препоръчителна стойност на RSDr, (%)	6. максимално разрешена стойност на RSDr, (%)	7. препоръчителна стойност на RSDR, (%)	8. максимално разрешена стойност на RSDR, (%)	9. брой измервания
10.	0,004 - 0,01	" 24	37	" 37	56	на една
11.	0,01 - 0,1	" 21	32	" 32	48	ля на
12.	0,1 - 1,0	" 11	16	" 16	24	, кратък


- "Sr" е стандартно отклонение, изчислено от резултатите, получени в условия на повторяемост.
- "RSDr" е относително стандартно отклонение, изчислено от резултатите, получено в условия на повторяемост  $[(Sr/x) \times 100]$ , където x е средната стойност на измерванията, изчислена от резултатите, получени в условия на повторяемост.
- "Възпроизводимост" (R) е степента на съвпадение на резултатите от измерванията на една и съща величина със съответната вероятност (обикновено 95%), извършени при променени условия на измерване (различни оператори, в различни лаборатории, но идентична проба и стандартизиран метод за анализ);  $R = 2,8 \times Sr$ .
- "RSDR" е относително стандартно отклонение, изчислено от резултатите, получени в условия на възпроизводимост  $[(Sr/x) \times 100]$ .

$$r = 2\sqrt{2}S_r = 2 * 1,41 * S_r = 2,8S_r$$




# ДЕФИНИЦИИ - БДС 17397 (речник по метрология)



 **ГРЕШКА** в химичния експеримент - разлика между резултата от измерване и истинската (действителната) стойност на измерваната величина.

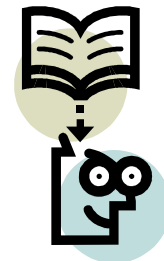
Промени в резултатите причинени от съвкупност от фактори, които повече или по-малко изменят търсения от химика резултат.

Някои действат едностранно, други двустранно.

 **Относителна** грешка - отношението на грешката и истинската стойност на измерената величина

$$X_{er} = \frac{|\bar{X} - \mu|}{\mu}$$

## (2) ДЕФИНИЦИИ - БДС 17397 (речник по метрология)



### ГРЕШКИТЕ БИВАТ

А) ✎ **СИСТЕМАТИЧНА ГРЕШКА (systematic error)** - разликата между **средноаритметичната стойност** която би се получила от безкраен брой измервания на една и съща СВ извършени при повтарящи се условия и **истинската стойност** на величината  $|X_{cp} - \mu|$

Систематичната грешка е равна на разликата между грешката и случайната грешка. Получава се при преобладаващо действие на един или няколко фактора.

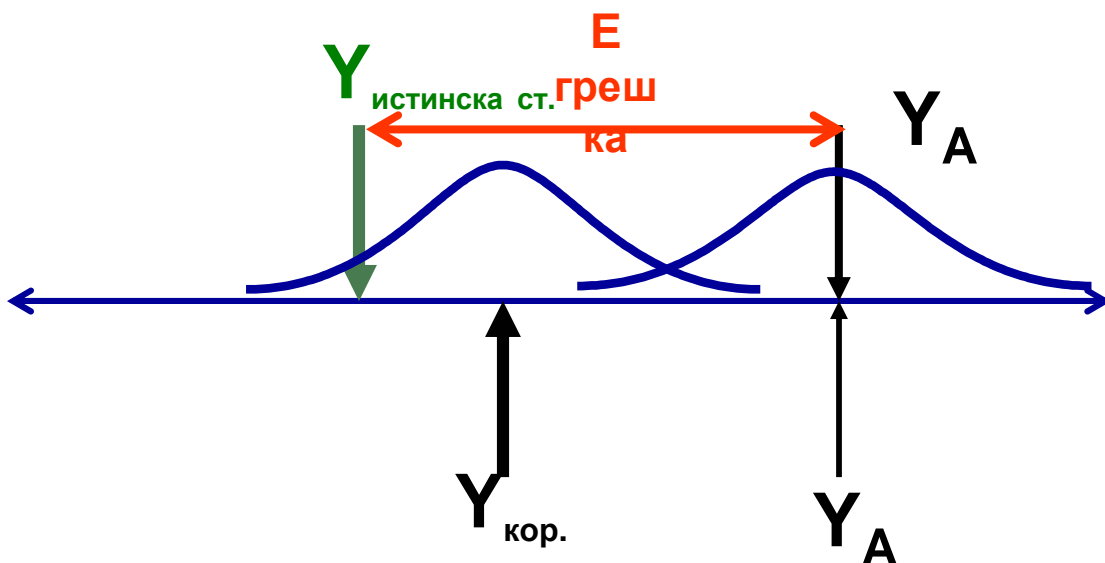
В) ✎ **СЛУЧАЙНА ГРЕШКА (random error)**- разликата между **резултата от измерване** и **средно-аритметичната стойност**, която би се получила от безкраен брой измервания на една и съща измервана величина извършена при повтарящи се условия  $|X - X_{cp}|$ .

Определя се от разсейването на отделните резултати около средната стойност. Оценява се чрез **S** и има пряка връзка с **НЕОПРЕДЕЛЕНОСТТА** на резултатите.

# Представяне на резултати и експериментална грешка

- Резултатите от измерванията се представят след математическа обработка на “суровите данни” по общоприет и разбираем начин, за да могат да бъдат ползвани и коментирани.
- Всички измервания включват (експериментална грешка) неопределеност
- Грешката бива
  - **Систематична (Systematic)** - еднопосочна определима и отстранима грешка
  - **Случайна (Random)** - разнопосочна не отстранима грешка
  - **GUM- неопределеност**

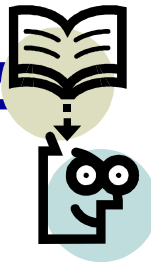
# НОВО в GUM - Разлика между грешка и неопределеност



**Грешка** - разлика между резултата от измерване и истинската стойност.

Тя не може да се определи количествено ако не се знае истинската стойност !!!!

# ЦЕНТРАЛНА ГРАНИЧНА ТЕОРЕМА



сумарната **СВ** на голям брой произволно разпределени, случайни величини които са **независими** и стойностите им са малки в сравнение с тази на сумата, - **КЛОНИ КЪМ НОРМАЛНО РАЗПРЕДЕЛЕНИЕ.**

# СИСТЕМАТИЧНА Грешка

- Възниква от провали в апаратурата и експерименталния дизайн
- Има определена стойност и познаваема причина
- Проявява се възпроизводимо при повторни измервания
- Може да се коригира относително лесно
- Примери
  - Аналита присъства във внесените реагенти (Празна проба)
  - Груба грешка в приготвените стандарти
  - Неправилно калибриран инструмент

**ВИЕ НОСИТЕ ОТГОВОРНОСТ ЗА НЕЯ !!!!**

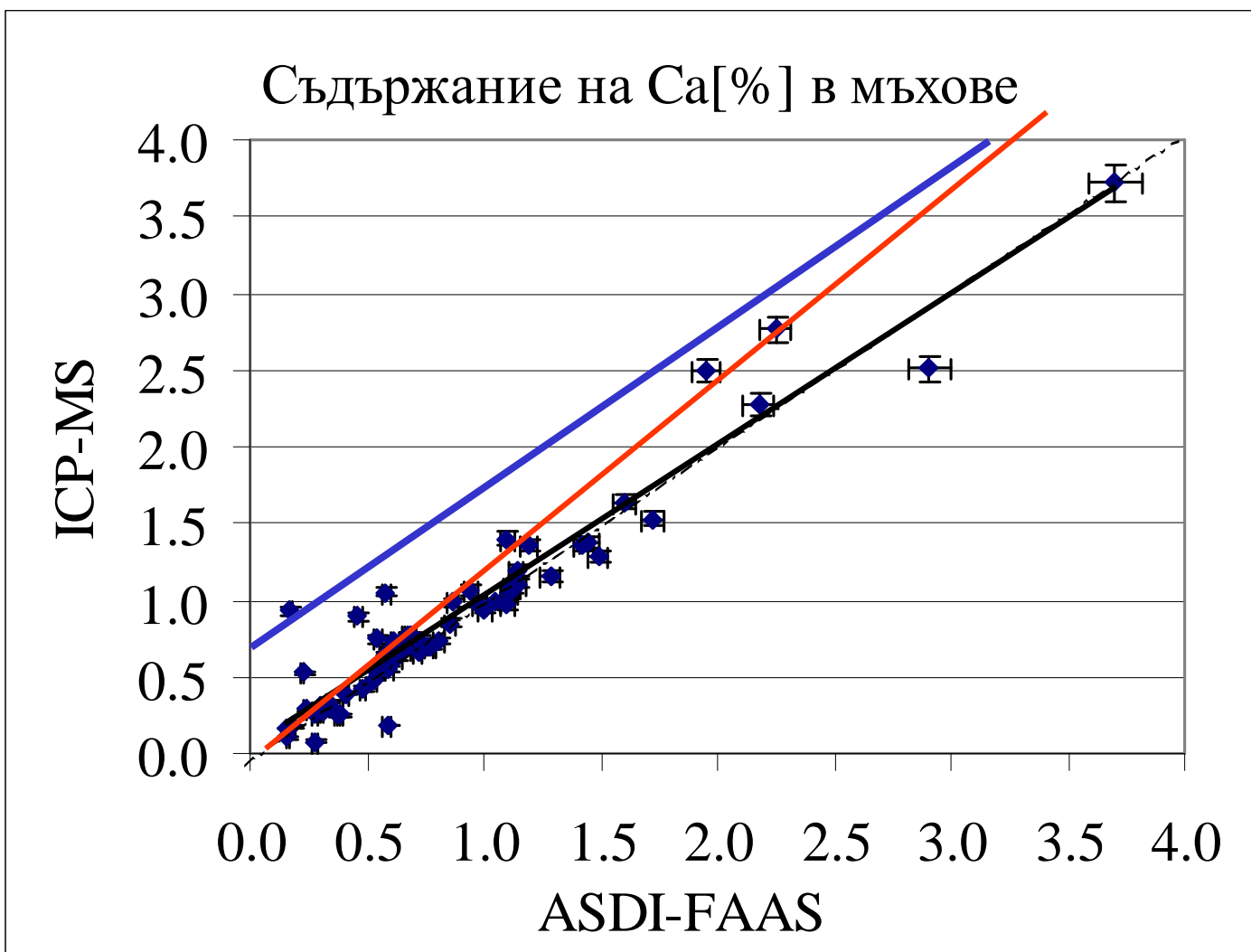


# Оценяване на систематична грешка

- Проверка на изчисленията чрез елементарен модел - логичен ли е резултата ?
- Анализирание на проба с известен състав
  - Сертифициран Сравнителен Материал CCM.  
Standard Reference material SRM
  - Коригиране на калибрационната линия (функция)
- Анализ на празни “blank” проби
  - Проверка дали инструмента ще даде - нулев резултат
- Сравнение с резултати от измервания на същата проба чрез други инструменти или от други лаборатории (ILC; PT)
  - Валидиране на процедурата на измерване

**ПОМОЩ ОТ ВЪН**

# ПОСТОЯННА И ПРОМЕНЛИВА ГРЕШКА



**ПОСТОЯННА** не зависи от  
количеството аналит  
**и**  
**ПРОМЕНЛИВА** зависи  
(обикновено пропорционално)  
от количеството аналит  
**СИСТЕМАТИЧНА ГРЕШКА..**

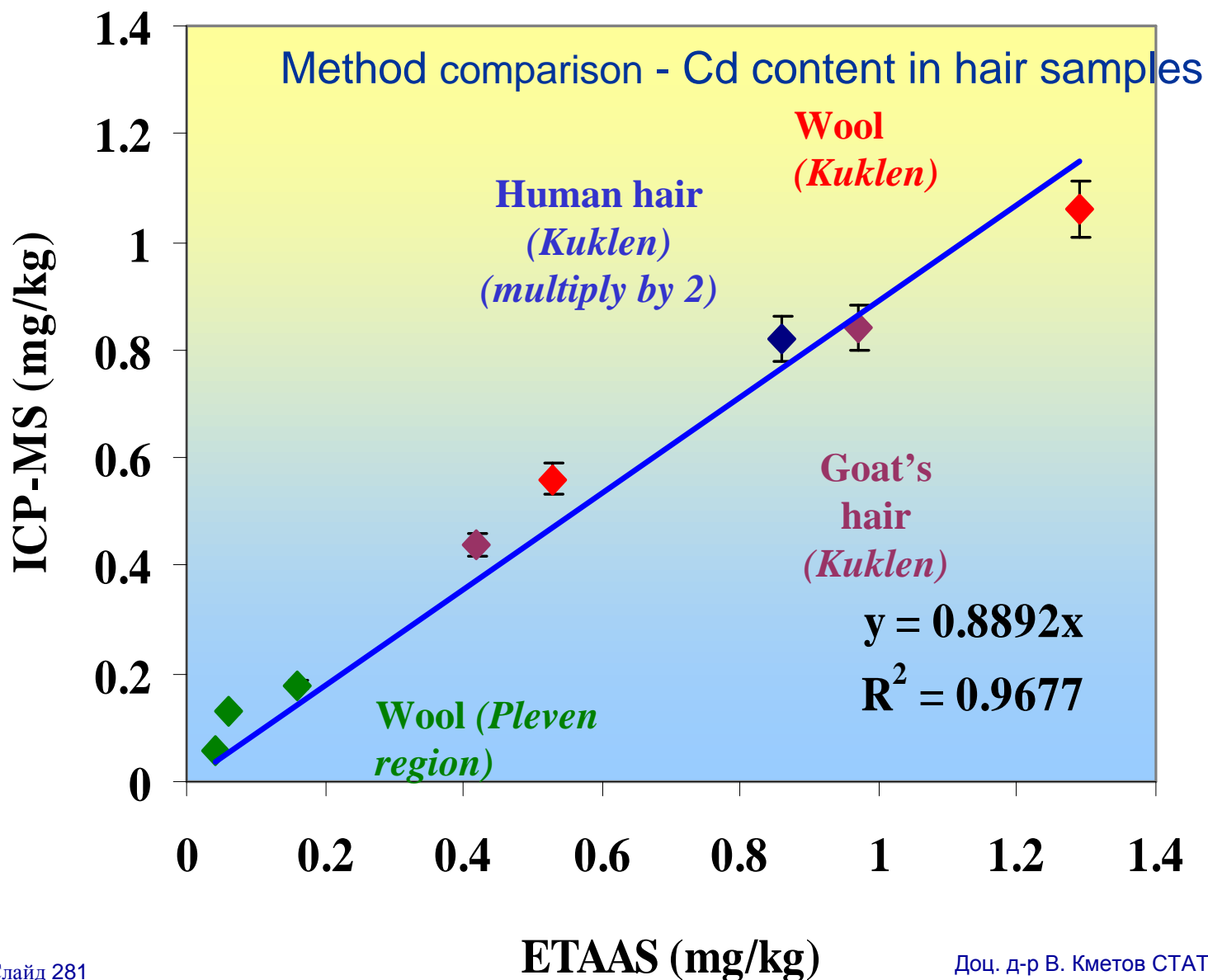


Каква ще грешката  
при невярно  
установена  
концентрация на  
титранта ?

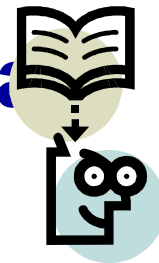


# DETERMINATION OF Pb AND Cd IN WOOL AND GOAT'S HAIR FROM THE ANIMAL POPULATION IN THE VICINITY OF NON-FERROUS METALS PLANT (PLOVDIV) BY ETAAS AND ICP-MS

V. KMETOV, A. SANCHEZ\*, A. CANALS\*, D. GEORGIEVA, V. STEFANOVA *Department of Analytical Chemistry, University of Plovdiv, Bulgaria. \*Department of Analytical Chemistry, University of Alicante, Spain.*



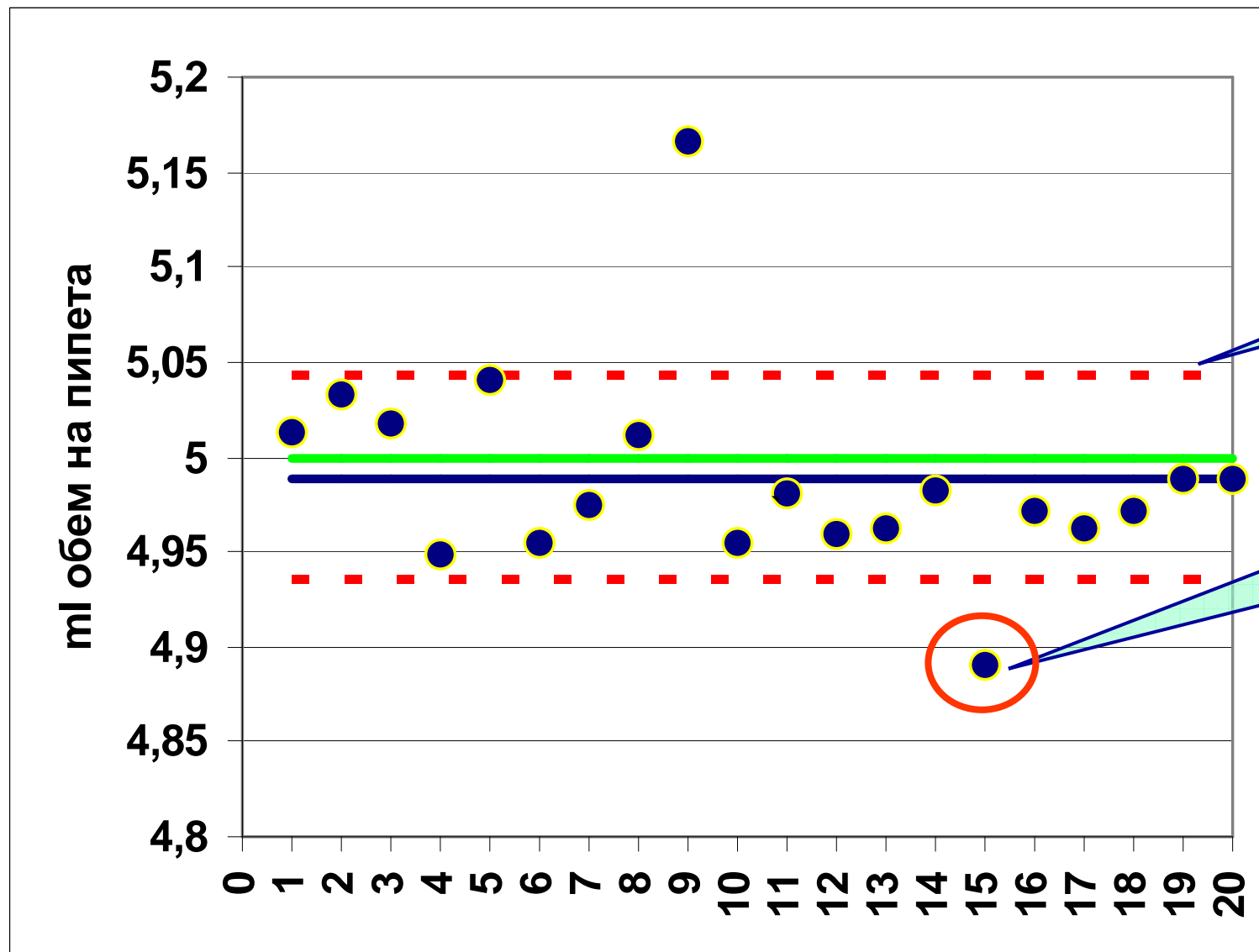
# Случайна Грешка



- Възниква от най-често не познаваеми и неконтролируеми източници
- Примери
  - Вариации при калибрирането и отчитането на сигнала от измервателната апаратура
  - Инструменталния шум
- Винаги присъства и **не може** да бъде коригирана.
- Подлежи на статистическа оценка и трябва да бъде коректно обявена.
- ЗАВИСИ ОТ БРОЯ НА ИЗМЕРВАНИЯТА!

**Можете да я оценявате статистически !**

# Преценка на грубите грешки - контролни карти



КАК ДА  
ОПРЕДЕЛИМ  
ИНТЕРВАЛА?

Това беглец  
ли е ?

# Проверка за бегълци

Тест на Диксон (Dixon's test) – Q test

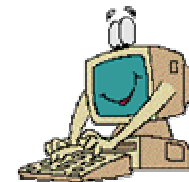
$$Q = \frac{|suspected - nearest|}{max - min}$$

Ако  $Q > Q_{tab} \Rightarrow$  **БЕГЛЕЦ**

Тест на Груб (Grubbs' test) – G test

$$G = \frac{suspected - \bar{X}}{S}$$

Ако  $G > G_{tab} \Rightarrow$  **БЕГЛЕЦ**



## Дефиниции

### Междулабораторни сравнения

(Inter-Laboratory Comparison - ILC)

*‘Организация, провеждане и оценка на изпитването на едни и същи или подобни обекти от две или повече лаборатории при определени условия’*

(Лабораторни) **Тестове за компетентност** (Proficiency Testing - PT )

*‘Определяне на компетентността на лабораторията посредством междулабораторни сравнителни изпитания’*

*[ISO/IEC Guide 43:1997]*

## Цели на МЛС

- Чрез МЛС се демонстрира компетентност и се установява степен на еквивалентност между резултатите на участващите лаборатории
- МЛС се използват за приписване на сертифицирани стойности на СМ
- Чрез МЛС се стандартизира/валидира метод (определяне на повторяемост, възпроизводимост, ...)
- МЛС като обучение и упражнение за подобряване на уменията

## Кой организира ILCs/PT?

- CCQM ([www.bipm.fr](http://www.bipm.fr))
- IMEP при IRMM ([www.imep.ws](http://www.imep.ws))
  - външни сравнителни стойности, свързани са с международните измервателни възможности
  - подкрепа от ЕА (Европейско сътрудничество в акредитацията)
  - по въпроси, отнесени към оп ЕУ директиви, пресечни граници на сектори & географски региони
- FAPAS ([www.fapas.com](http://www.fapas.com))
- AFSSA ([www.afssa.fr](http://www.afssa.fr))
- ЕА ([www.european-accreditation.org](http://www.european-accreditation.org))
- Национални референтни лаборатории (NRLs)
- Други → провери [www.eptis.bam.de](http://www.eptis.bam.de)  
(Европейска информационна система за РТ схеми)

## Концентрация на Pb във вино (IMEP-16)



EC Directive 2001/22

EC Regulation 2676/90

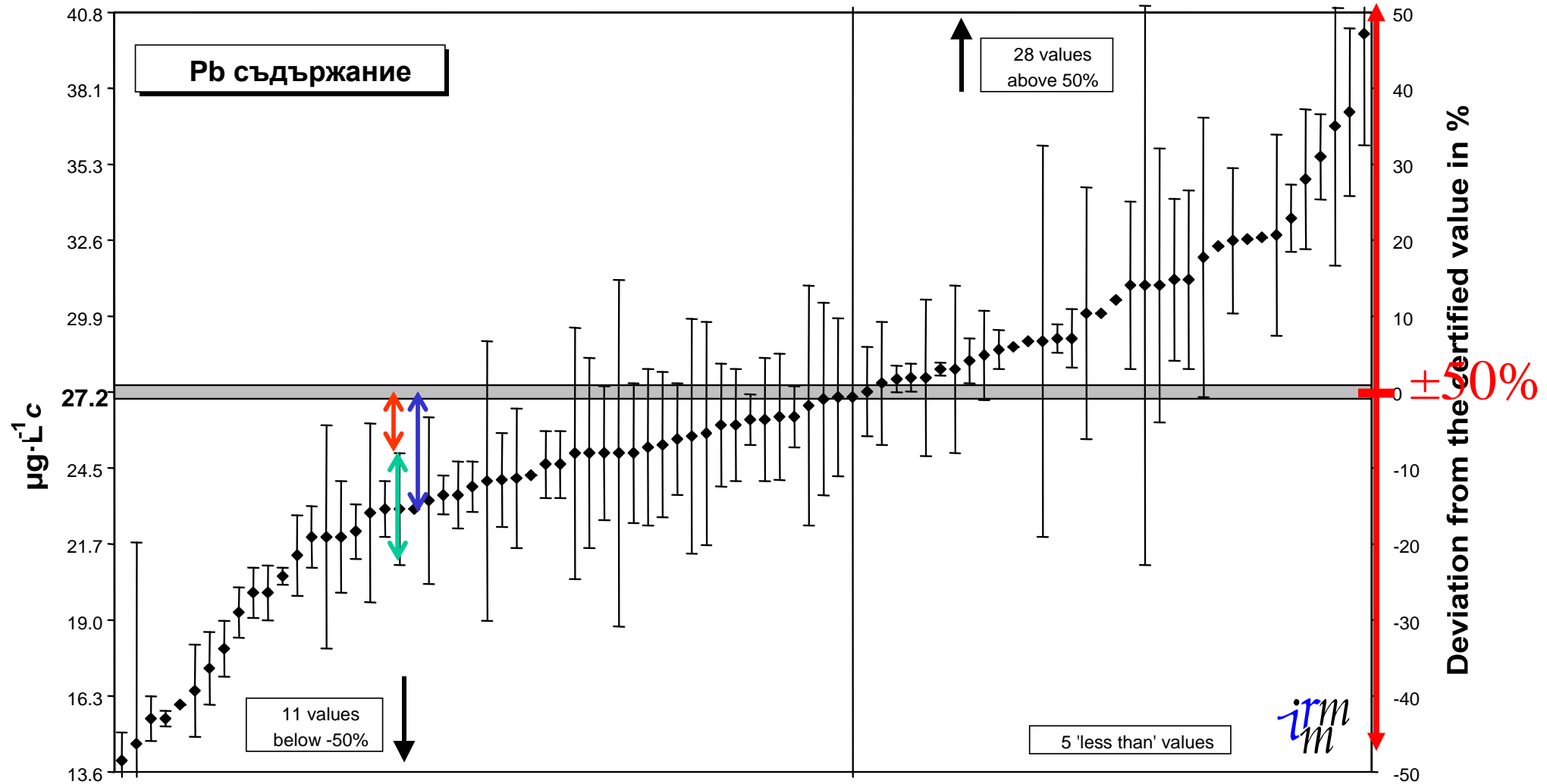
EC Regulation 466/2001

⇒ *пределно допустима стойност  
от 0.2 mg Pb /Kg*



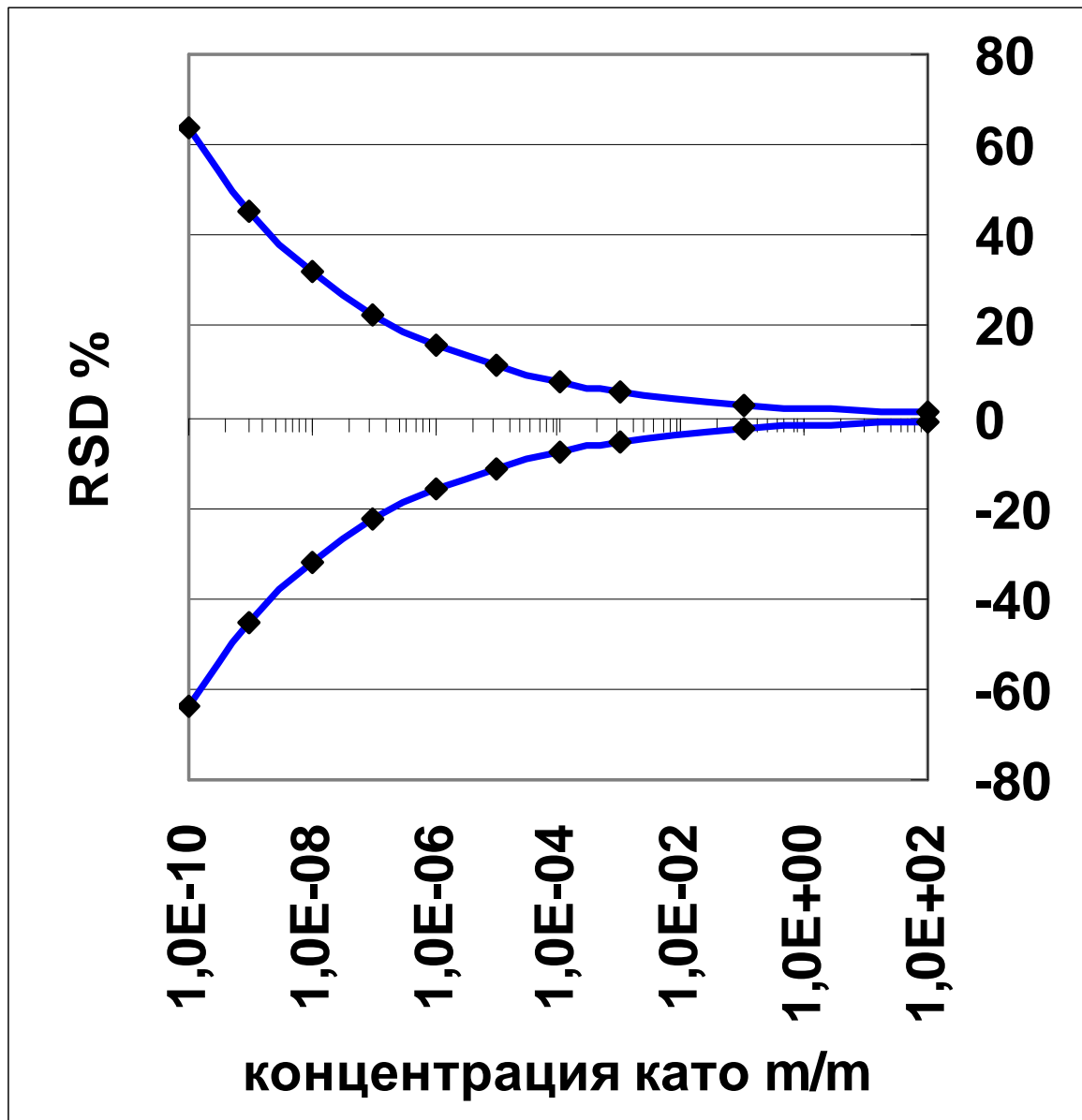
# IMEP- 16: Pb във вино

Сертифицирана стойност :  $27.18 \pm 0.25 \mu\text{g}\cdot\text{L}^{-1}$  [ $U=k\cdot u_c$  ( $k=2$ )]

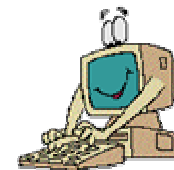


Резултати от всички участници

# Трумпет на Хорвиц



$$RSD = \pm 2^{(1-0,5 \log C)}$$



## ISO Дефиниция за неопределеност

‘a parameter associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the *measurand*’

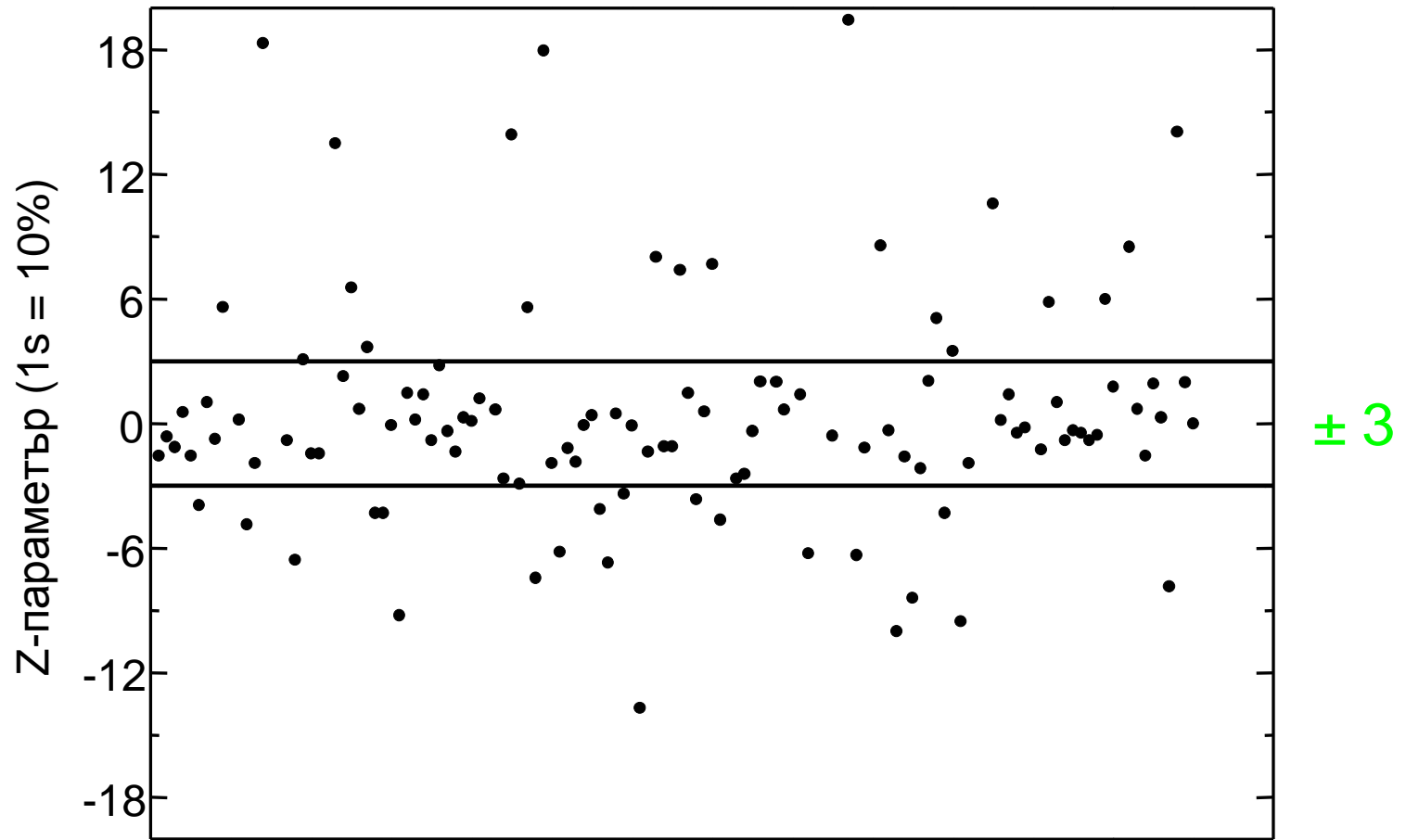
✎ “параметър, асоцииран към резултата от измерване, характеризиращ дисперсията на стойностите, които могат да бъдат основателно приписани на измерваната величина”

☒ **Резултат = стойност ± неопределеност**

$$(22.7 \pm 4.8) \text{ mg/kg}$$

стойността е между 17.9 и 27.5 mg/kg  
(т.е. в област, интервал)

# Оценка на данни Z-параметри



Сортирани по възходящ номер на лабораториите

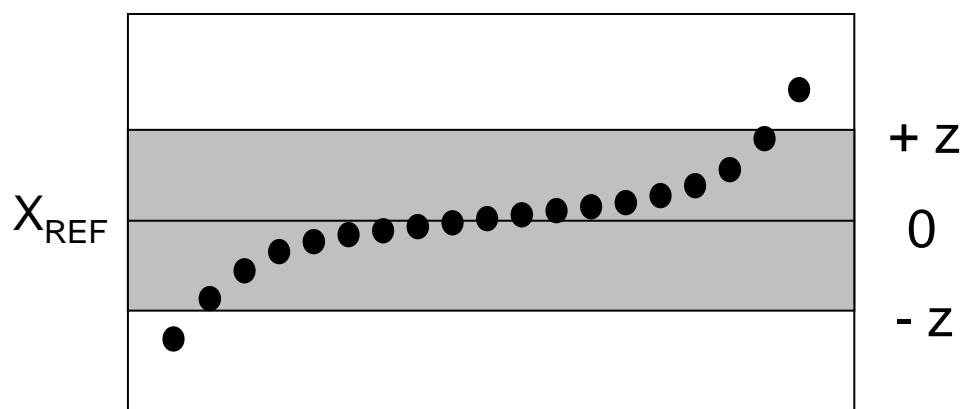
*1s = 10%*

## Индикатори за изпълнението

– Процентна грешка;  $\%E = (x_{\text{lab}} - x_{\text{ref}}) / x_{\text{ref}}$

– **Z стойност;**  $Z = (x_{\text{lab}} - x_{\text{ref}}) / s$

–  $E_n$  -стойност;  $E_n = (x_{\text{lab}} - x_{\text{ref}}) / (u_{\text{lab}}^2 + u_{\text{ref}}^2)^{1/2}$



Оценка на изпълнението,

$|Z| \leq 1$  много добро

$|Z| \leq 2$  добро

$|Z| \leq 3$  задоволително

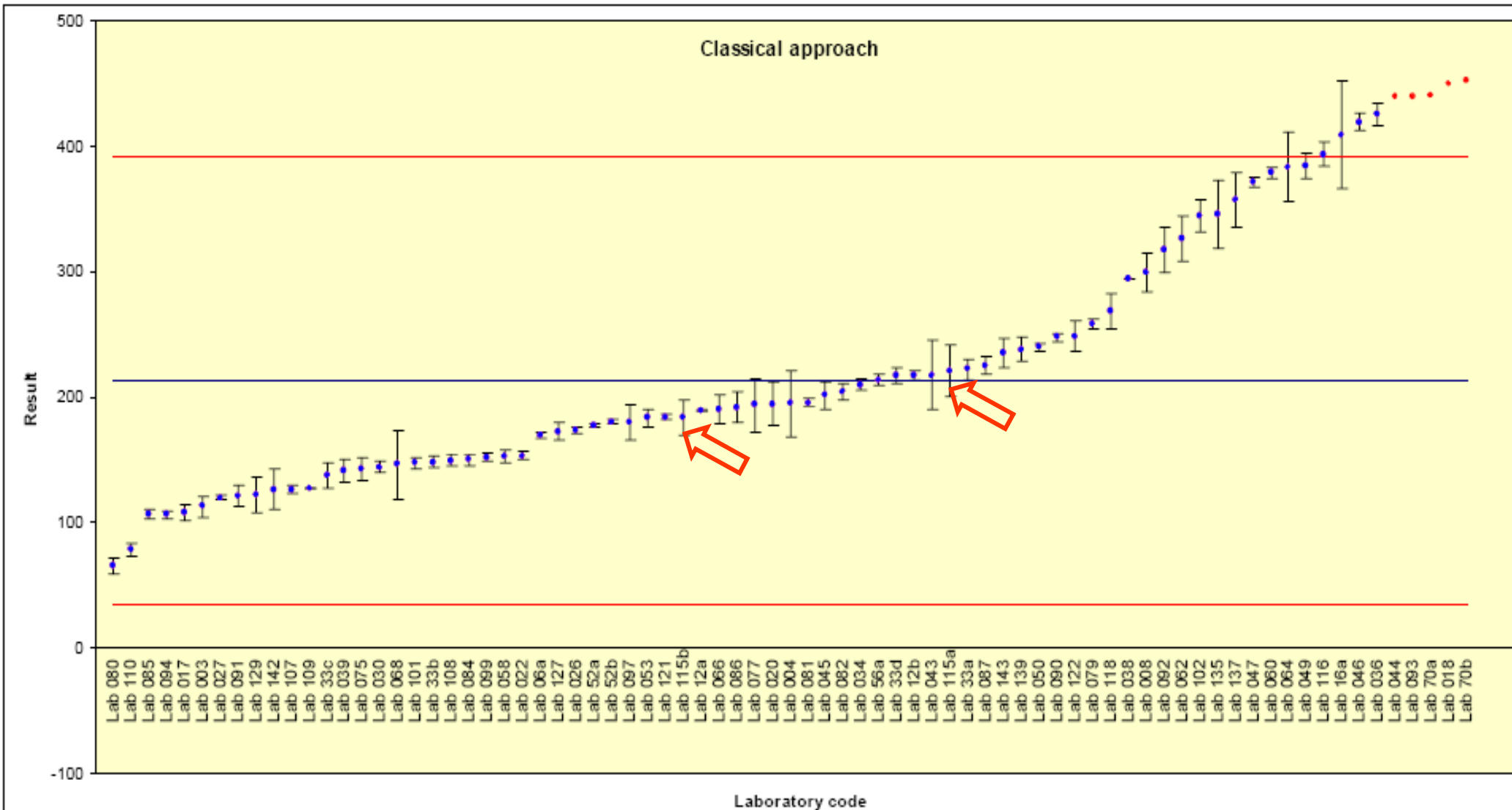
$|Z| > 3$  незадоволително

**превантивни мерки**

# Improving the infrastructure for metrology in chemistry in the candidate New Member States (newly associated states)

## Qua-NAS

QuaNAS: Cr in urban dust [mg/kg] - all data



Analyte:	As
Matrix:	Urban dust
Unit:	mg/kg

**Results by Laboratory**

Label	Comment1	n	Average	St.Dev.	RSD	Z-Score	*/**
Lab 143	HGAAS	3	6.31	0.29	4.6%	-2.0	*
Lab 020	GFAAS	3	6.67	0.13	1.9%	-1.9	
Lab 084	HGAAS	6	7.42	0.50	6.7%	-1.6	
Lab 080	HGAAS	10	8.27	0.47	5.7%	-1.2	
Lab 127	ETAAS	3	8.31	1.03	12.4%	-1.2	
Lab 068	HGAAS	3	8.55	0.17	2.0%	-1.1	
Lab 135	ICPAES	5	8.74	0.96	11.0%	-1.1	
Lab 097	ICPAES	6	8.98	0.45	5.0%	-1.0	
Lab 036	HGAAS	3	9.12	0.31	3.4%	-0.9	
Lab 142	AASTHGA	6	9.41	0.55	5.8%	-0.8	
Lab 004	GFAAS	10	9.47	0.73	7.7%	-0.8	
Lab 099	HGAAS	6	9.49	0.34	3.6%	-0.7	
Lab 045	ICPOES	6	9.55	0.11	1.2%	-0.7	
Lab 008	ICPOES	9	9.82	0.45	4.6%	-0.6	
Lab 039	ICPOES	6	10.13	1.31	12.9%	-0.5	
Lab 101	ICPOES	3	10.13	0.23	2.3%	-0.5	
Lab 038	ICPMS	1	10.30		#VALUE!	-0.4	
Lab 139	HGAAS	10	10.31	0.55	5.3%	-0.4	
Lab 060	AAS	3	10.57	0.08	0.8%	-0.3	
Lab 022	ICPOES	3	10.70	0.60	5.6%	-0.3	
Lab 092	ICPMS	5	10.74	0.37	3.4%	-0.2	
Lab 064	ICPMS	4	10.83	0.92	8.5%	-0.2	
Lab 090	ICPMS	10	11.01	0.71	6.5%	-0.1	
Lab 041	HGAAS	3	11.19	1.06	9.5%	-0.1	
Lab 026	GFAAS	3	11.33	0.58	5.1%	0.0	
Lab 046	HGAAS	10	11.86	0.25	2.1%	0.2	
Lab 115b	ETAAS	3	11.97	0.97	8.1%	0.3	
Lab 108	ETAAS	10	11.97	1.40	11.7%	0.3	
Lab 12a	ICPOES	3	12.07	0.55	4.6%	0.3	
Lab 12b	ICPOES	3	12.13	0.80	6.6%	0.3	
Lab 33a	ICPMS	4	12.33	0.25	2.0%	0.4	
Lab 030	HGAAS	3	12.43	0.15	1.2%	0.5	
Lab 003	ETAAS	6	12.46	0.89	7.1%	0.5	
Lab 075	ETAAS	3	12.70	0.87	6.8%	0.6	
Lab 70b	k0-INAA	3	12.78	0.06	0.5%	0.6	
Lab 56a	ICPAES	3	12.83	0.75	5.9%	0.6	
Lab 047	HGAAS	3	13.47	0.45	3.3%	0.9	
Lab 145	ISPAES	3	13.96	0.13	0.9%	1.1	
Lab 115a	ICPMS	3	14.13	1.00	7.1%	1.1	
Lab 044	ETAAS	3	14.20	0.22	1.6%	1.2	

## МЛС при валидиране на метода

- Задача: определяне на повторяемостта “ $s_r$ ” и възпроизводимостта “ $s_R$ ” на процедурата за измерване
- Оценка чрез ANOVA (дисперсионен анализ **на променливите**)
- Проверка за груби грешки (преди осредняване/включване)
  - Тест на Cochran за променливи груби грешки,
  - Тест на Grubbs за средни груби грешки

*[ISO5725-2]*

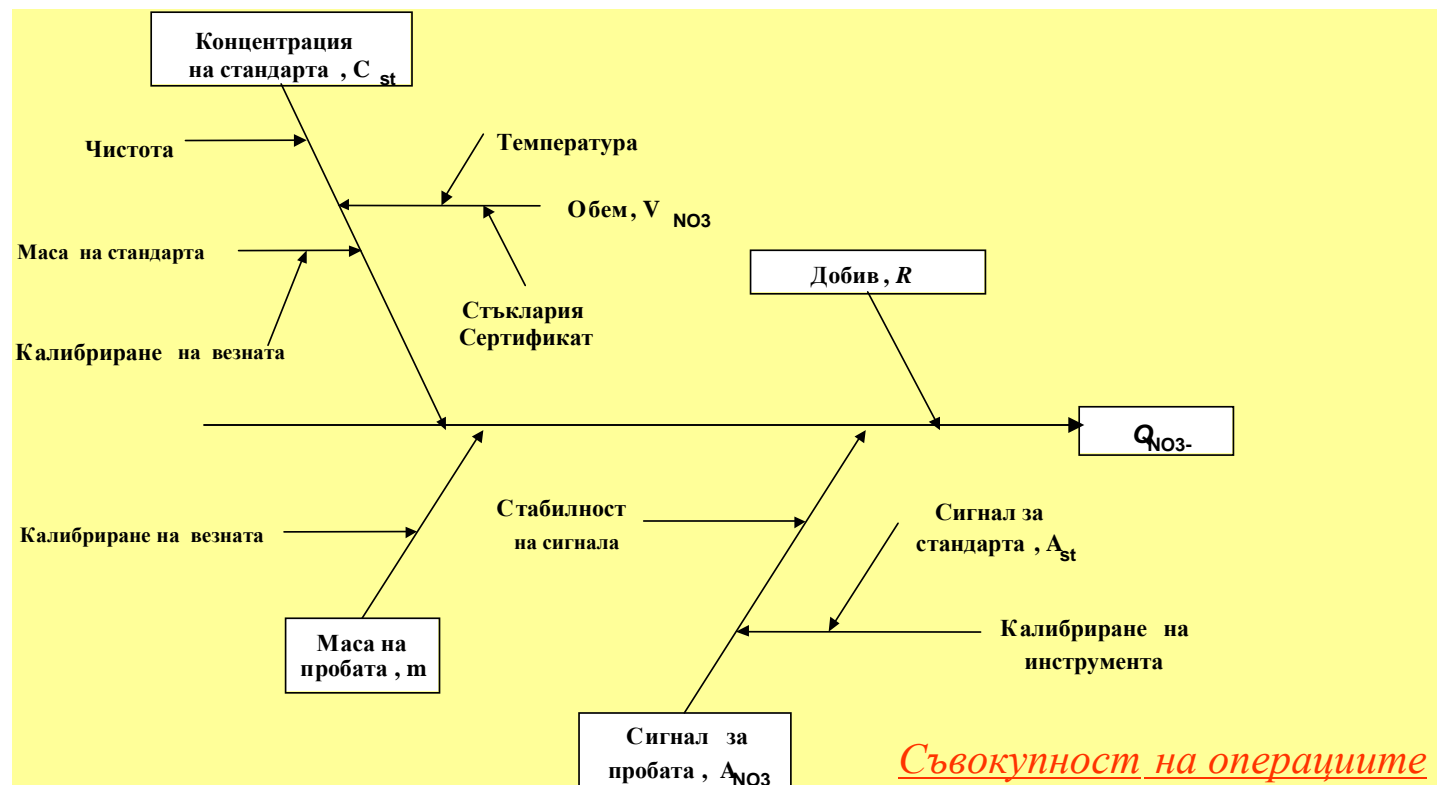


# Днес разгледахме:

## ГРЕШКИ В ХИМИЧНИЯ ЕКСПЕРИМЕНТ

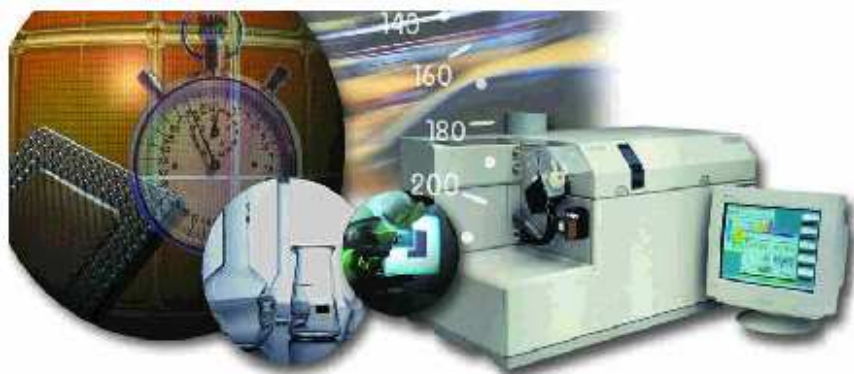
1. Точност и прецизност на резултатите от измерване
2. Видове грешки.
  - Случайна и систематична грешка.
  - Оценка на систематичната грешка. Постоянна и променлива систематична грешка. Причини.
3. Оценка на случайната грешка - НЕОПРЕДЕЛЕНОСТ .
  - Прецизност - възпроизводимост - повторяемост. Стандартно отклонение.
  - Относителни и абсолютни грешки.
4. Оценка на точността
  - Сравнение със ССМ
  - Участие в междулабораторни сравнения и тестове за пригодност

## лекция № 13



# НЕОПРЕДЕЛЕНОСТ на резултатите от измерване

# БЮДЖЕТ на неопределеност



# Днес ще разгледаме:

## НЕОПРЕДЕЛЕНОСТ на резултатите при измерване

### 1. Що е то неопределеност ?

- Кое налага нейното приложение ?
- Кое е новото?
- Разлика между случайна грешка и неопределеност

### 2. Видове неопределеност

- ТИП **A** и ТИП **B**
- Средноквадратична - Комбинирана и Разширена

### 3. GUM процедура за изграждане бюджета на неопределеността

- 10 стъпки
- Моделно уравнение
- Закон за разпространение на грешката

### 4. Примери

## За какво ни е нужна неопределеността?

- Изисква се от ISO 17025 - акредитация.
- Неопределеността на резултата показва метрологичното КАЧЕСТВО на измерванията (*това не е измерване с най-малката достижима неопределеност*).
- Тя подобрява знанието за дадена измервателна процедура.
- Вътре в лабораторията → документира по прозрачен начин измервателната процедура.
- За крайния потребител → дава резултата с приемлива доверителност.
- Позволява да се сравняват резултати.
- Идентифицира главните източници на неопределеност - разкрива способности за подобряване на процедурата
- Показва съответствие с граничните норми (законови или договорени) и спомага за установяване на критерии за приемане (допуск).

⇒ **Вашата най-добра защита при дискусии!**

- ☒ Повтаряне на измерването 2, 10 или 100 пъти не Ви дава необходимата цялостна информация, за да считате че резултатите Ви са надеждни!

***GUM (1995) :***  
***ISO Guide on the expression of***  
***Uncertainty in Measurements***

***(Ръководство за изразяване на неопределеността  
при измерванията)***

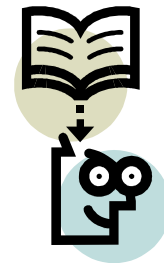
**EURACHEM/CITAC**

<http://www.eurachem.ul.pt/guides/QUAM2000-1.pdf>



**Guide Quantifying Uncertainty in Analytical  
Measurement, 2000**

# ISO Дефиниция за неопределеност uncertainty



‘a parameter associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the *measurand*’

“параметър, асоцииран към резултата от измерване, характеризиращ дисперсията на стойностите, които могат да бъдат основателно (разумно) приписани на измерваната величина”

**Резултат = стойност ± неопределеност**

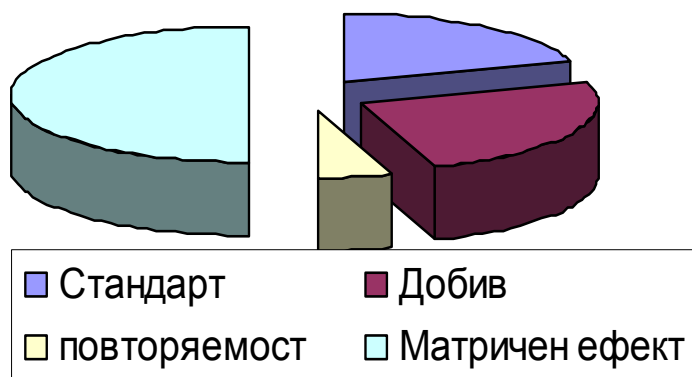
**(5.15 ± 0.36) mg/kg**

стойността е между 5.51 и 4.79 mg/kg  
(т.е. в област, интервал)

полу-интервал

НОВА КОНЦЕПЦИЯ за  
**неопределеността** :  
включва оценка на всички  
елементи на несигурност -  
вътрешни и външни за  
лабораторията

несигурно твърдение



Открий и Определи

**Заявявам**

Концентрация на Fe във вино  
 $92.3 \pm 4.6 \mu\text{mol/L}$   
 $5.15 \pm 0.26 \text{ mg/L}$

истинската стойност се намира с  
вероятност  $P = (1-\alpha)$  в интервала

$$\mu \in \bar{X} \pm \frac{t_{(f, \alpha)}}{\sqrt{N}} \cdot S$$

**Доказвам**

**Демонстрирам**

# ОБЩА НЕОПРЕДЕЛЕННОСТ

Всеки **процес** предполага **продукт**

Всеки **продукт** притежава **характеристики** – качества

Всяка **количествено определима** характеристика може да бъде **измерена**

**ОБЩАТА НЕОПРЕДЕЛЕННОСТ:**

**Дисперсията на характеристиките на продукта**

**+**

**Дисперсията на измервателния процес**

---

**= ОБЩА ДИСПЕРСИЯ -> НЕОПРЕДЕЛЕННОСТ**



## Различават ли се резултатите?

*Няма резултати без неопределеност !*

$$R1 = 5.8 \text{ mg/L}$$

$$R2 = 7.1 \text{ mg/L}$$

- Традиционният подход: доверителен интервал

$$R1 = (5.8 \pm 0.2) \text{ mg/L}$$

$$R2 = (7.1 \pm 0.3) \text{ mg/L}$$

- GUM подхода: разпространение на неопределеността (комбинирана неопределеност) взема под внимание приноса на всички компоненти

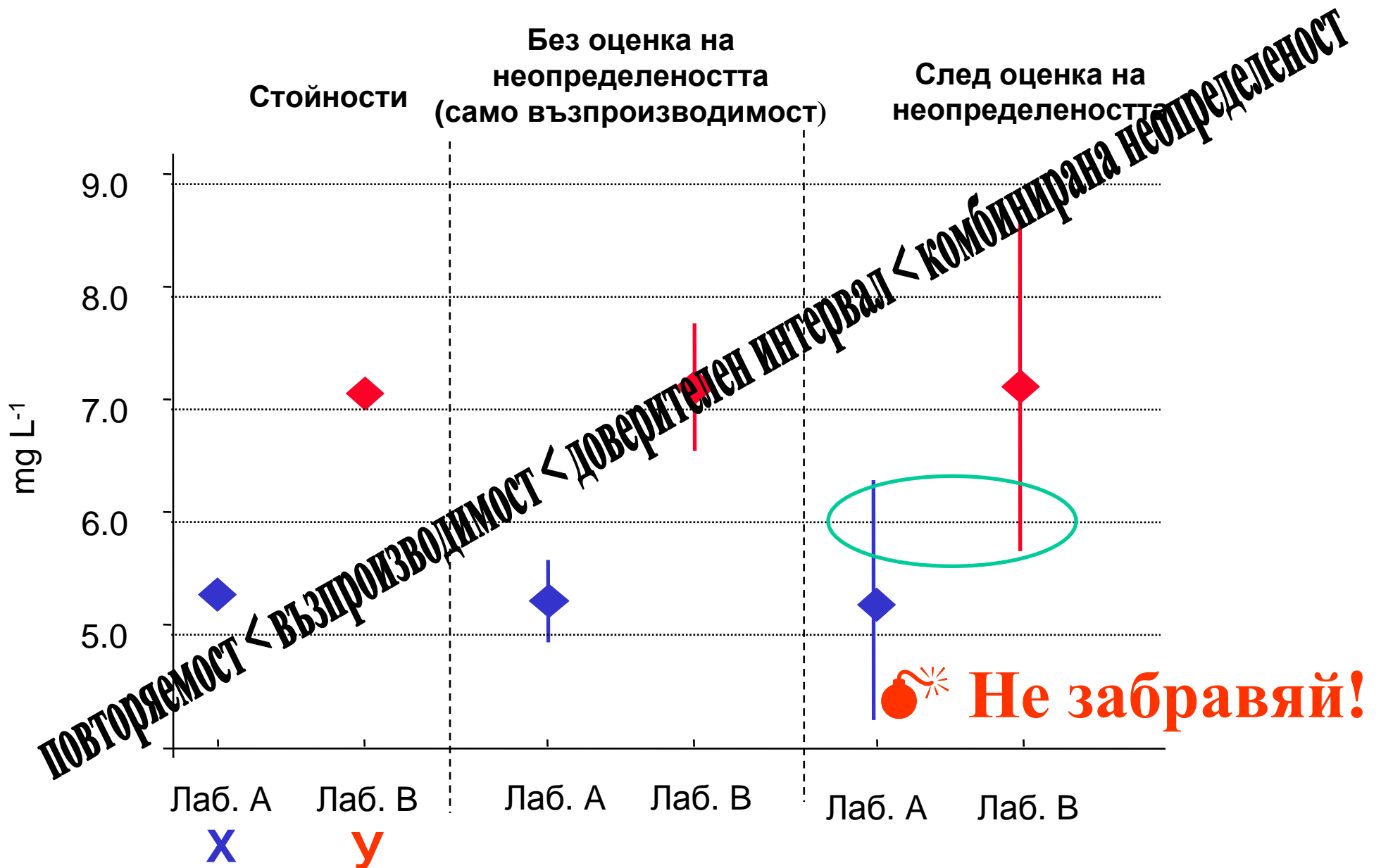
$$R1 = (5.8 \pm 0.7) \text{ mg/L}$$

$$R2 = (7.1 \pm 0.7) \text{ mg/L}$$

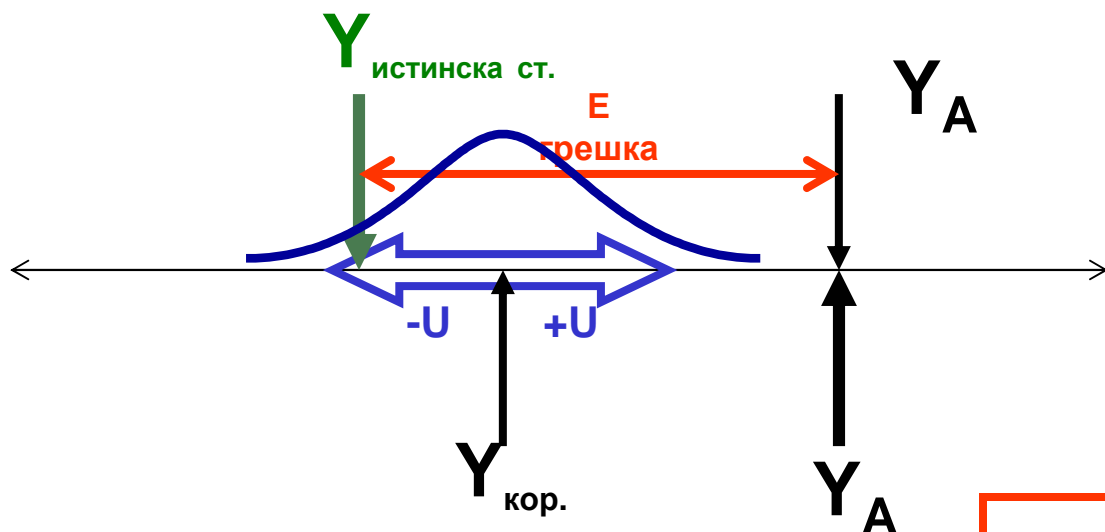
Статистическите тестове не се изискват от GUM (почти) .../...

Визуално сравнение → припокриване на интервали ДА/НЕ ?

## Различават ли се резултатите?



# НОВО в GUM - Разлика между грешка и неопределеност



- GUM е указател за прозрачно, просто и стандартизирано документиране на процедурата на измерване.
- Използвайте оценки на неопределеността, такива като **тип A** (измерени в лабораторията) и **тип B** (други).

**ИЗБЯГВАЙТЕ** понятията  
случайна и систематична  
грешка !

Грешка - разлика между резултата от измерване и истинската стойност.

Тя не може да се определи количествено !!!!

- Прилагайте данни за комбинирана и разширена неопределеност.

# Видове неопределеност

## UNCERTAINTY



✓ Средноквадратична  
неопределеност  
 $u_s$  (А или В)

Неопределеност на резултат от измерване, изразена като средноквадратично отклонение (стандартно отклонение)

БИВА тип А и тип В

✓ КОМБИНИРАНА  
средноквадратична  
неопределеност  
 $u_c$

Средноквадратична неопределеност на резултат от измерване, получен от стойности на известен брой други величини.

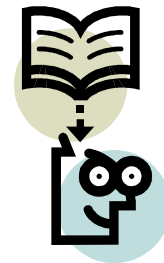
Квадратен корен от сума от членовете които са дисперсиите и ковариациите на тези други величини, претеглени в зависимост от това как влияят на резултата

✓ РАЗШИРЕНА  
неопределеност  
 $U$  (  $K=2$  )

Величина (полуширина) на интервал около резултата от измерване за който може да се очаква, че обхваща голяма част от разпределението на стойностите които биха могли да се припишат на измерваната величина.

Получава се като комбинираната средноквадратична неопределеност се умножи по фактор на покриване  $K$  (най-често  $K=2$ )

## Неопределеност “Тип”



Оценяване на неопределеността **ТИП А** :

статистически анализ на серии от наблюдения.

**Тип А** стандартна неопределеност се измерва от повторими експерименти и нейната количествена мярка е *стандартното отклонение* на измерените резултати.

Оценяване на неопределеността **ТИП В** :

чрез други способи, а не статистически анализ  
(*предишни експерименти, литературни данни, информация от производителя*)

[GUM, 1993]

## Кога ще оценявате неопределеността на резултатите от измерванията ?

- Когато се въвежда нова процедура във Вашата лаборатория
- Когато се промени критичен фактор в процедурата (инструмент, оператор, ...)
- По време и съвместно с процеса на валидиране

➔ **За всеки отделен получен резултат  
НЕ е нужна индивидуална оценка!**

## Разбиране на измерването !

“...The evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of measurement...”

[GUM § 3.4.8]

“...оценяването на неопределеността не е рутинна практическа задача, нито чисто математическа; тя зависи от детайлното познаване на природата на измерваната величина и измерването ...”

## Какво Ви е необходимо да знаете ?

### *Някои основни положения в статистиката*

- средно аритметично на набор от данни;
- стандартно отклонение;
- закон за разпространение на неопределеността (грешката);
- разпределение (нормално, правоъгълно, триъгълно ...)

## Как да прилагаме GUM неопределеността ?

*Документирайте входящите данни, които използвате при измерването !!*

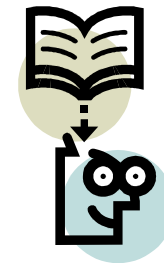
“...Пакетът от информация може да включва:

- ✓ данни от предишни измервания;
- ✓ данни за валидиране;
- ✓ експертни знания или опит за поведението и/или свойствата на съответни материали или инструменти;
- ✓ спецификации на производителя;
- ✓ данни от калибрирания или от други сертификати;
- ✓ неопределеност, приписана на данни, взети от справочници”

*[GUM § 4.3.1]*



## 10-те стъпки на GUM последователността



- 1 - Дефиниране на измерваната величина
- 2 - Описание на Моделното уравнение (*за процедурата на измерване*)
- 3 - Идентифициране на възможните източници на неопределеност
- 4 - Оценяване на всички входящи величини  $Y = f(X_1, X_2, X_3, \dots, X_i)$
- 5 - Оценяване на стандартната неопределеност  $u_i(X_i)$  (1S)  
за всяка от входящите величини  $Y = f(X_1, X_2, X_3, \dots, X_i) - S_1, S_2, S_3, \dots, S_i$
- 6 - Изчисляване на стойността на измерваната величина (използвайки моделното уравнение)
- 7 - Изчисляване на комбинираната стандартна неопределеност на резултата  $u_c(Y)$
- 8 - Изчисляване на разширената стандартна неопределеност  $U(k=2)$  (с избрано  $k$ , 2 или 3)
- 9 - Анализ на преносния индекс в неопределеността (*МИСЛИ !!*)
- 10- Документиране на всички стъпки в Отчет.

## Стъпка 1 – Измервана величина

## Протокол на експеримента

**Измервана величина** = конкретната величина, която е обект на измерване

**Пример:** Определяне на нитрати в растителен материал чрез йонна хроматография **съдържание на  $\text{NO}_3^-$  в (mg/g)**

АНАЛИТ



## Стъпка 2 - Уравнение - модел

Моделът на измервателната процедура е функционалната връзка между входните X величини и изходната Y величина (резултат):

$$Y = f(X_1, X_2, \dots, X_n)$$

Изходната величина Y зависи от входните величини  $X_1, X_2, \dots, X_n$ :

МОДЕЛЪТ на измерване представлява уравнението, което използвате, за да изчислите Вашия резултат !

**Вие вече го имате !**

Входни величини ( $X_i$ ) могат да бъдат величини, чиито стойност и неопределеност са директно определени в конкретното измерване (**Тип А, статистически анализ на серия от наблюдения**) или привнесени към измерването от външни източници (**Тип В, предходни експерименти, литературни данни, информация от производителя**)

## Моделно уравнение

$$Q_{\text{NO}_3^-} = C_{\text{st}} \frac{A_{\text{NO}_3^-} \cdot V_{\text{NO}_3^-}}{A_{\text{st}} \cdot m} \times f_{\text{di}} \times \frac{1}{R}$$

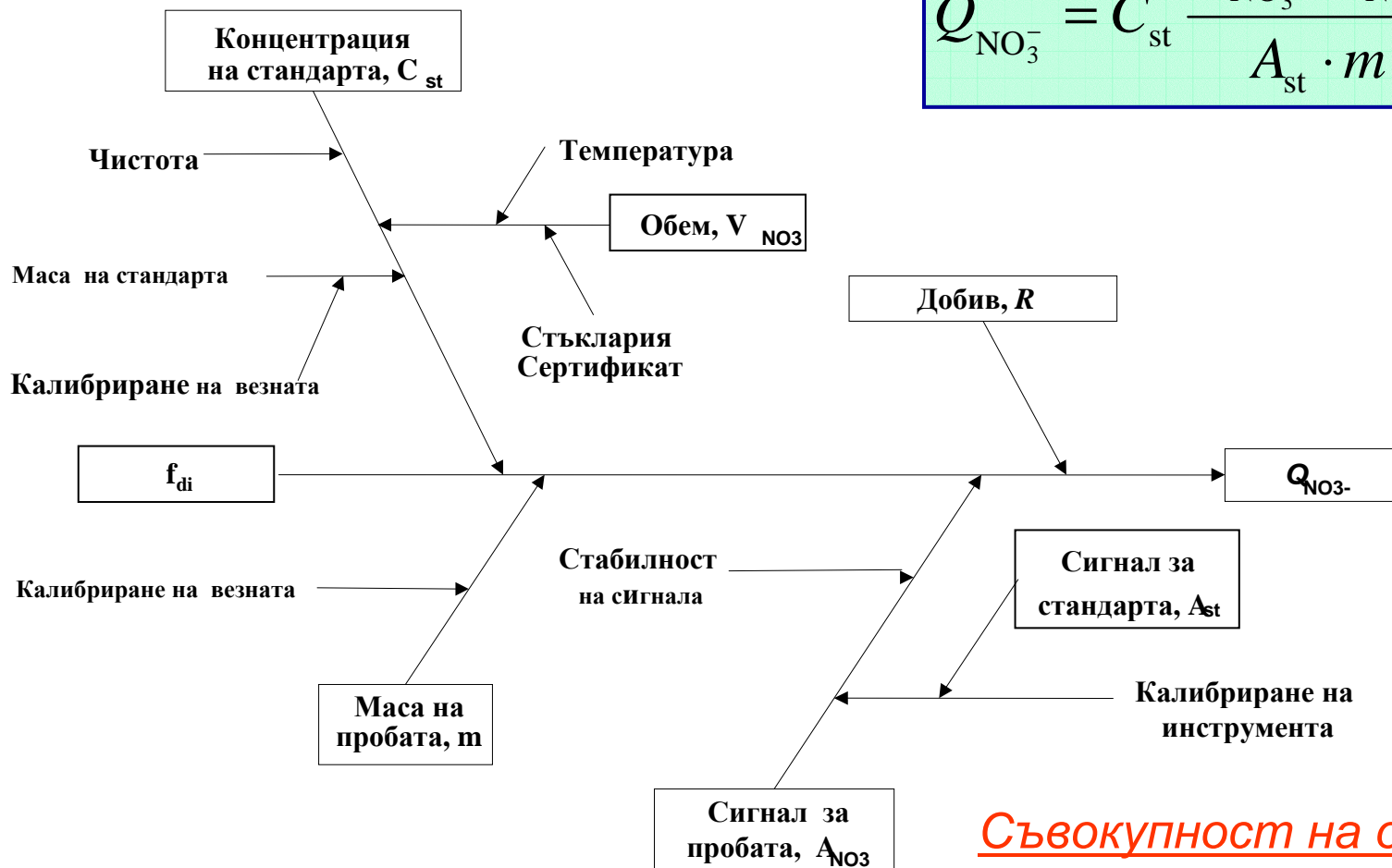
- $Q_{\text{NO}_3^-}$  нитратно съдържание в пробата (mg/g) -  $Y$
- $C_{\text{st}}$  нитратно съдържание в стандартния разтвор (mg/l) –  $X_1$  тип B
- $A_{\text{NO}_3^-}$  интензитет на сигнала (РА) на пробния разтвор  $X_2$  тип A
- $A_{\text{st}}$  интензитет на сигнала (РА) на стандартния разтвор  $X_3$  тип A
- $V_{\text{NO}_3^-}$  обем на разтвора с проба (l)  $X_4$  тип B
- $m$  масата на пробата (g)  $X_5$  тип B
- $f_{\text{di}}$  фактор на разреждане (безразмерна величина);  $X_6$  тип A
- $R$  аналитичен добив (вж. пробоподготовка)  $X_7$  тип B или A

## Стъпка 3 - Възможни източници на неопределеност

- ☑ аналитичен добив от комплексни матрици
- ☑ условия на съхранение
- ☑ чистота на реагентите
- ☑ приета стехиометрия
- ☑ пробоподаване
- ☑ инструментални параметри на измерване
- ☑ инструментален отклик
- ☑ отклонения на инструмента
- ☑ разделителна способност
- ☑ неопределеност на стандартите и на CRM
- ☑ вариации на повторните наблюдения

## Стъпка 3 - възможни източници на неопределеност

$$Q_{\text{NO}_3^-} = C_{\text{st}} \frac{A_{\text{NO}_3^-} \cdot V_{\text{NO}_3^-}}{A_{\text{st}} \cdot m} \times f_{\text{di}} \times \frac{1}{R}$$



Съвокупност на операциите  
Входни и влияещи величини

## Стъпка 4 - неопределеност на входните величини (оценка на типове A & B)

- повторни наблюдения (A)
- валидиращи експерименти (A и/или B)
- спецификации на производителя (B)
- калибрационни сертификати (B)
- резултати от междулабораторни сравнения за валидиране на метод (B)
- от натрупан опит и/или литература (B)

## Стъпка 5 - изразяване на стандартна неопределеност

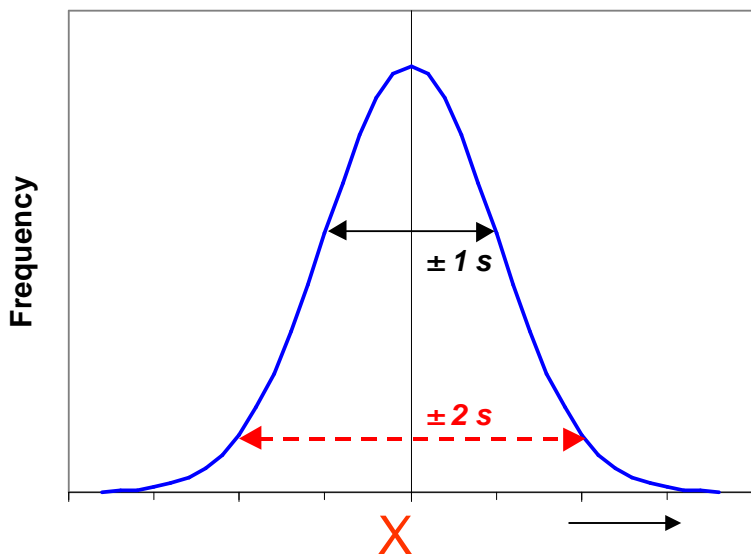
Преди комбиниране всички допринасящи към неопределеността компоненти трябва да бъдат изразени/превърнати в размерност на стандартна неопределеност.

**Когато наличната размерност е :**

- стандартно отклонение : не се променя
- доверителен интервал : превърни в S
- обявен интервал : превърни в S
- разширена неопределеност: превърни в S



# Нормално разпределение



За набор от  $n$  стойности на  $x_i$

Средна стойност

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n (x_i)$$

Стандартно отклонение =  $u_s$  =

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

Дисперсия на средната стойност

$$V(x) = \sigma^2 \cong S^2$$

Относително стандартно отклонение  
RSD (абс. стойност или %)  
= коефициент на вариация CV

$$RSD = \frac{S}{\bar{x}}$$

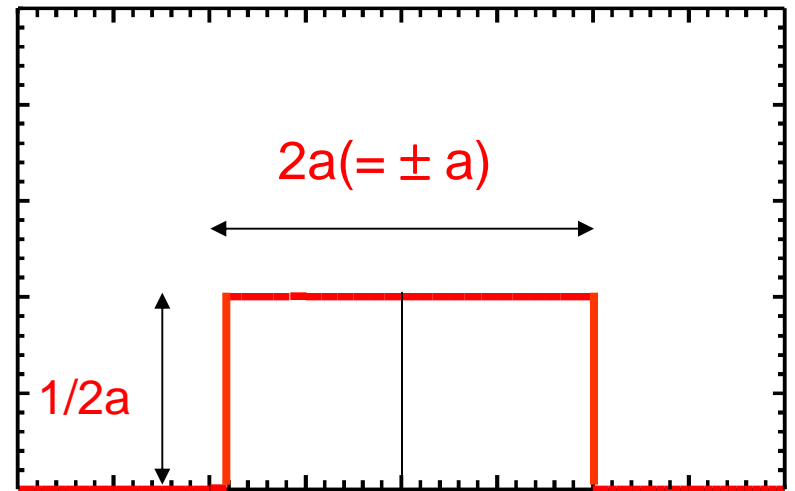
## Правоъгълно (равномерно) разпределение

Стойността  $x$  е в границите: ↑

$$a_- \dots a_+$$

Очакваната стойност на  $y$  е:

$$y = x \pm a$$



Приемано стандартно отклонение:

$$S = \frac{a}{\sqrt{3}}$$



Допуска се, че с равна вероятност стойността може да е разположена където и да е в тези граници

Концентрация на стандарт за калибриране

Чистотата на образец

## Пример за Правоъгълно разпределение

“Вероятно стойността е някъде в този обхват ( $\pm a$ );”

Удостоверения (сертификати) или други определения относно границите в които би могла да се намира стойността, без посочване нивото на надеждност (или степените свобода).

### Примери:

Концентрацията на стандарт за калибриране е декларирана  $1000 \pm 2 \text{ mg/L}$ .

Приемайки равномерно разпределение на стойността на концентрацията на стандарта – допускът се превръща в стандартно отклонение:

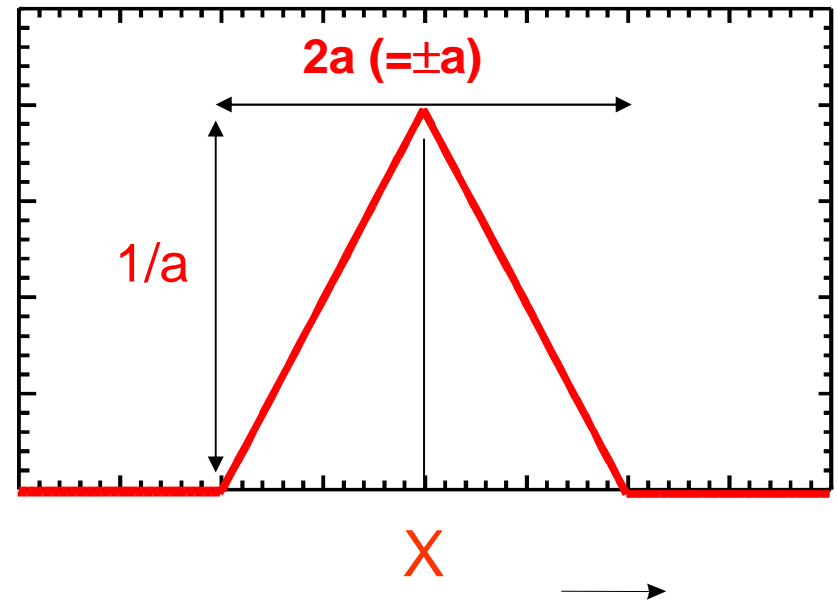
$$S = u(x) = a / \sqrt{3} = 2 / \sqrt{3} = 1.16 \text{ mg / L}$$

Чистотата на образец Cd указана в сертификат е  $99.99 \pm 0.01 \%$  Приемайки модела за правоъгълното разпределение неопределеността на стойността на стандарта е:

$$S = u(x) = a / \sqrt{3} = 0.01 / \sqrt{3} = 0.0058 \%$$

# Триъгълно разпределение

Разпределение използвано в случаите когато е направено допускането, че стойностите намиращи се в средата на доверителния интервал са по-вероятни от тези намиращи се в краищата му



$$y = x \pm a$$

Приемано стандартно отклонение:

Напр. обем за мерителна колба

$$S = \frac{a}{\sqrt{6}}$$

**Ако се съмнявате, използвайте равномерно разпределение**

## Пример за Триъгълно разпределение

*Близките до  $x$  стойности са по-вероятни от тези разположени в границите на интервала*

ПРИМЕР - мерителна колба

Производителят дава обем за мерителна колба от

$100 \pm 0.1$  ml при  $T = 20^\circ\text{C}$ .

Номиналната стойност е най-вероятна!

Приемайки модела на триъгълно разпределение, неопределеността на тази стойност е:

$$S = u(x) = a/\sqrt{6} = 0.1/\sqrt{6} = 0.04 \text{ ml}$$

**Ако се съмнявате, използвайте равномерно (ПРАВОЪГЪЛНО) разпределение**

## Стъпка 6 - изчисляване на стойността на измерваната величина

Използвай моделното уравнение за изчисляване на стойността на изходната величина  $Y$  ( $Q_{\text{NO}_3^-}$ )

$$Q_{\text{NO}_3^-} = C_{\text{st}} \frac{A_{\text{NO}_3^-} \cdot V_{\text{NO}_3^-}}{A_{\text{st}} \cdot m} \times f_{\text{di}} \times \frac{1}{R}$$

Пример

$$Q_{\text{NO}_3^-} = 0.801 \times \frac{0.0131 \times 0.1000}{0.0232 \times 1.142} \times 10 \times \frac{1}{0.78}$$

$$Q_{\text{NO}_3^-} = 0.508 \text{ mg / g}$$

# Закон за “Разпределение на неопределеността” *без корелация*

$$Y = f(X_1, X_2, \dots, X_n)$$

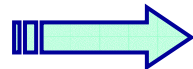
$$u_c(Y) = \sqrt{\sum \left( \frac{\partial f}{\partial X_i} \right)^2 \cdot u_{X_i}^2}$$

$$Y = (X_1 + X_2)$$
$$Y = (X_1 - X_2)$$



$$u_c(Y) = \sqrt{u_{X_1}^2 + u_{X_2}^2}$$

$$Y = (X_1 * X_2)$$
$$Y = (X_1 / X_2)$$



$$\frac{u_{C(Y)}}{Y} = \sqrt{\left( \frac{u_{X_1}}{X_1} \right)^2 + \left( \frac{u_{X_2}}{X_2} \right)^2}$$

## Стъпка 7 - Комбинирана стандартна неопределеност

Когато входните величини не корелират помежду си, **комбинираната стандартна неопределеност** се оценява като квадратен корен от обединените дисперсии, съгласно уравнението:

**Закон за разпространение на неопределеността,**

$$u_c(Y) = \sqrt{\sum \left( \frac{\partial f}{\partial X_i} \right)^2 \cdot (u_{X_i})^2}$$

където

$u_c(Y)$  = комбинирана стандартна неопределеност

$u_{X_i}$  = стандартна неопределеност на всяка от входните величини

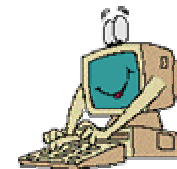
**Може да се изчисли с електронна таблица  
или чрез специализиран софтуер!**



## Комбинирана стандартна неопределеност

$$u_{c,RSu}(Q_{NO_3^-}) = \sqrt{RSu(C_{st})^2 + RSu(A_{NO_3^-})^2 + RSu(A_{st})^2 + RSu(V_{NO_3^-})^2 + RSu(m)^2 + RSu(f_{di})^2 + RSu(R)^2}$$

където  $RSu(X_i) = u(X_i)/X_i$  (относителни стандартни неопределености)



$$u_{c,RSu}(Q_{NO_3^-}) = \sqrt{\left(\frac{0.00058}{0.801}\right)^2 + \left(\frac{0.0003}{0.0131}\right)^2 + \left(\frac{0.0006}{0.0232}\right)^2 + \left(\frac{0.0003}{0.1000}\right)^2 + \left(\frac{0.00058}{1.1420}\right)^2 + \left(\frac{0.023}{10.000}\right)^2 + \left(\frac{0.04}{0.78}\right)^2}$$

$$u_c(Q_{NO_3^-}) = u_{c,RSu}(Q_{NO_3^-}) \times Q_{NO_3^-} = 0.032 \text{ mg/g}$$

## Стъпка 8 - Разширена неопределеност

Разширената неопределеност  $U$ , се получава чрез умножение на комбинираната стандартна неопределеност  $u_c(y)$  с фактор на покриване  $k$ :

$$U = k * u_c$$

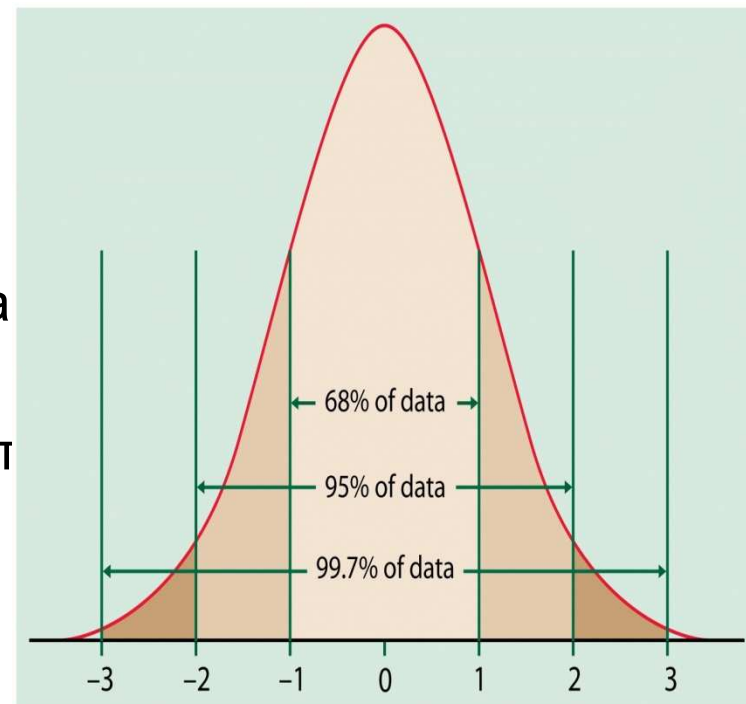
Резултатът се представя като: **Резултат =  $y \pm U$  ( $k = ??$ )**

*Например:*  $Q_{\text{NO}_3^-} = 0.51 \pm 0.06 \text{ mg/g} , k = 2$

- “ $y$ ” е най-добрата оценка за стойността, която може да се отдаде на измерваната величина,
- Интервалът  **$[y - U, y + U]$**  е областта, която се очаква, че обхваща голяма част от разпределението на стойности, присъщи на измерената величина **в рамките на разумното**.

## Стъпка 8 - Разширена неопределеност (2)

- Разширената неопределеност дава по-реалистичен обхват на възможни стойности.
- Обикновено факторът на покриване се избира като  **$k = 2$** , представлящ покриване на около 95% от вероятните значения, ако разпределението е нормално.

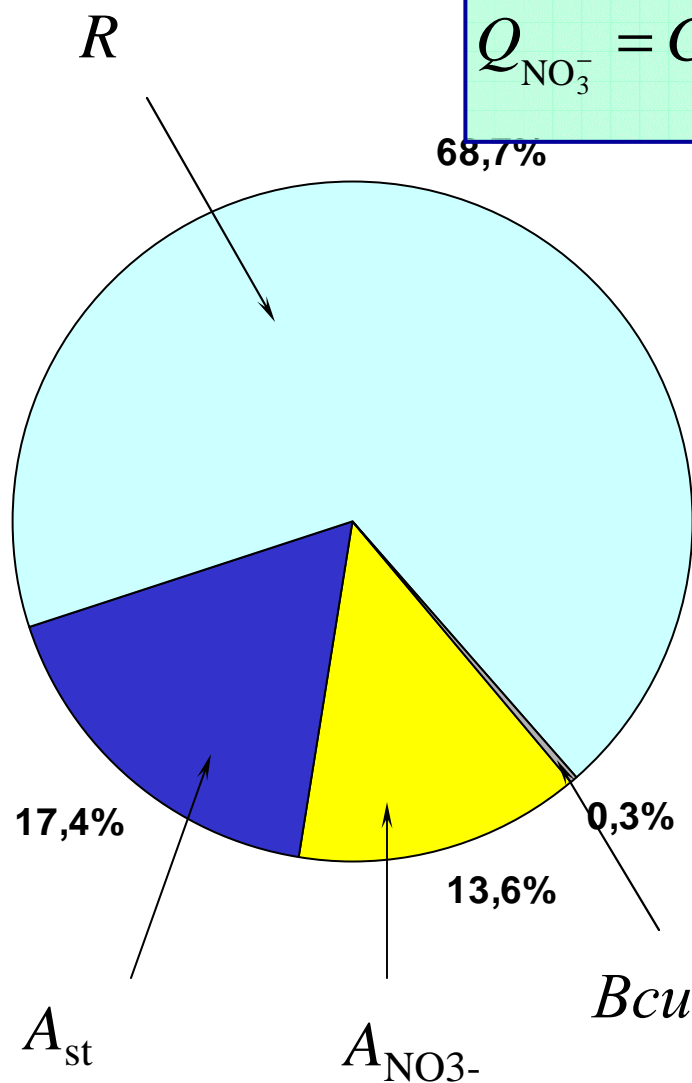


Стандартната неопределеност следва да бъде използвана  
вътре в лабораторията  
(за прилагане на разпространението на  
неопределеността)

Разширената неопределеност касае по-реалистичен обхват  
и следва да съпътства резултата за крайния потребител.

## Стъпка 9 - Приноси в неопределеността

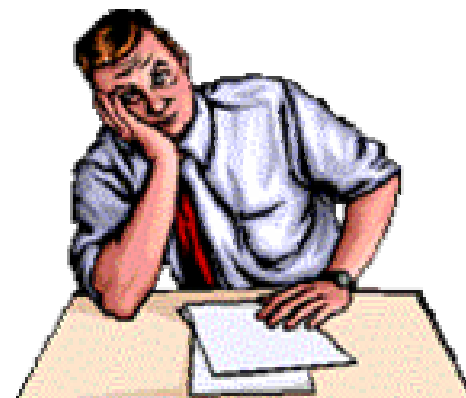
$$Q_{\text{NO}_3^-} = C_{\text{st}} \frac{A_{\text{NO}_3^-} \cdot V_{\text{NO}_3^-}}{A_{\text{st}} \cdot m} \times f_{\text{di}} \times \frac{1}{R}$$



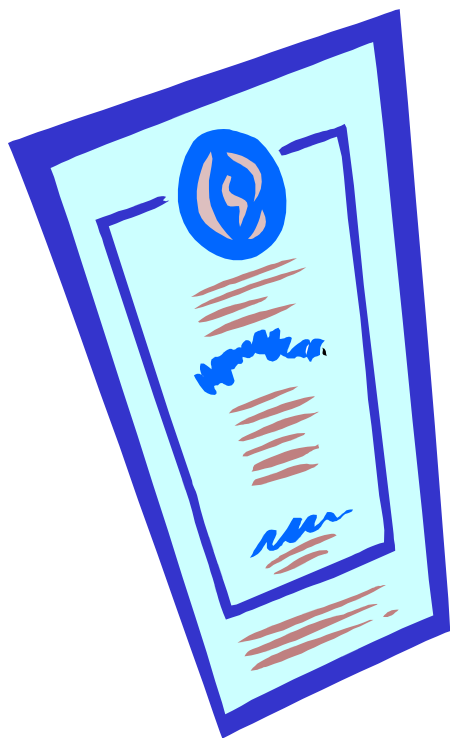
Главни компоненти:

- Тип В? ☹
- Тип А? ☺
- Повторения ?
- Повече работа ?
- Контролни карти ?

**МИСЛИ!**



## Стъпка 10 - отчет на резултатите



$$Q_{NO_3^-} = (0.51 \pm 0.06) \text{ mg/g } (*)$$

(\*)

*Отчетената неопределеност е разширената неопределеност, изчислена при фактор на покриване = 2, даващ ниво на сигурност от около 95%*

## Резюме: относно неопределеността

- **Неопределеност** оценена съгласно **GUM**, е обща концепция за оценка на резултатите при измерване за **всички сектори** на измерванията в частност и за химията!
- Тя позволява на аналитика да комбинира предварително придобити знания с такива от наблюденията по логичен и добре дефиниран начин;
- Тя позволява другите (например оценителите) да разберат какво & как са извършени нещата
- Тя не изисква да се измерва с най-малката достижима неопределеност, а с тази която е най-реалистична.
- Концепцията е възприета и призната от международните институции, като националните метрологични институти и BIPM; IUPAC, OIML и акредитационните комитети като EA (BCA) и ILAC; CEN внедрява тези концепции
- **Тя се изисква от ISO 17025 за акредитация**

# Днес разгледахме:

## НЕОПРЕДЕЛЕНОСТ на резултатите при измерване

### 1. Що е то неопределеност ?

- Кое налага нейното приложение ?
- Кое е новото?
- Разлика между случайна грешка и неопределеност

### 2. Видове неопределеност

- **ТИП А** и **ТИП В**
- Средноквадратична - Комбинирана и Разширена

### 3. GUM процедура за изграждане бюджета на неопределеността

- 10 стъпки
- Моделно уравнение
- Закон за разпространение на грешката

### 4. Примери

An iceberg floating in the ocean. The tip of the iceberg is above the water surface, and the much larger base is submerged. Two green arrows point to the tip: the top one is labeled 'ОТЛИЧЕН 6' and the bottom one is labeled 'Среден 3'. The tip is labeled 'Резултат на изпита'. The submerged part contains the text 'САМОСТОЯТЕЛНА РАБОТА', 'Вили + Пламен + Киро .....', 'КОЛОКВИУМ', and 'ЛЕКЦИИ Весо'.

**ОТЛИЧЕН 6**

**Среден 3**

**Резултат на  
изпита**

**САМОСТОЯТЕЛНА  
РАБОТА**

**Вили + Пламен + Киро .....**

**КОЛОКВИУМ**

**ЛЕКЦИИ Весо**