Tutorial Statistics Limits Part 1

M. Herbst

influenced by many other unknowing contributors,
mentioned where possible

today	**tomorrow**
Statistics/ Probability	*Hypothesis Testing*
Frequentist/ Bayesian	*Error Classification*
Probability Density Function	*Size/ Power of Test*
Confidence Level/ p-Value	*Test Statistics/ Chisquare Dist.*
Confidence Intervals	*NP Lemma/ Wilks' theorem*
Exercises	*Likelihood Function*
	Systematics
	POI, Nuissance Parameters
	Profile Likelihood Ratio

Coverage/ Flip-Flopping/ Asymptotic Limit/ Look-Elsewhere

current ATLAS discussion: Power Constraint Limits

What is Statistics?

“The mathematics of the collection, organization, and interpretation of numerical data, ...”

www.thefreedictionary.com

first part inspired by Roger Wolf (CMS, Uni Hamburg)

What is Statistics?

"there are three kinds of lies:

lies, damned lies and statistics"

Benjamin Disraeli

"the only statistics you can trust

are those you falsified yourself"

Winston Churchill

Long History of Statistics

Oldest Census in Egypt/ China (2600 b.c.)

number of people, wealth, weapons suitability

Demographic Studies in Great Britain (19th century)

mass collection of demographic characteristics

"Descriptive Statistics"

you have to be good at counting

Long History of Statistics

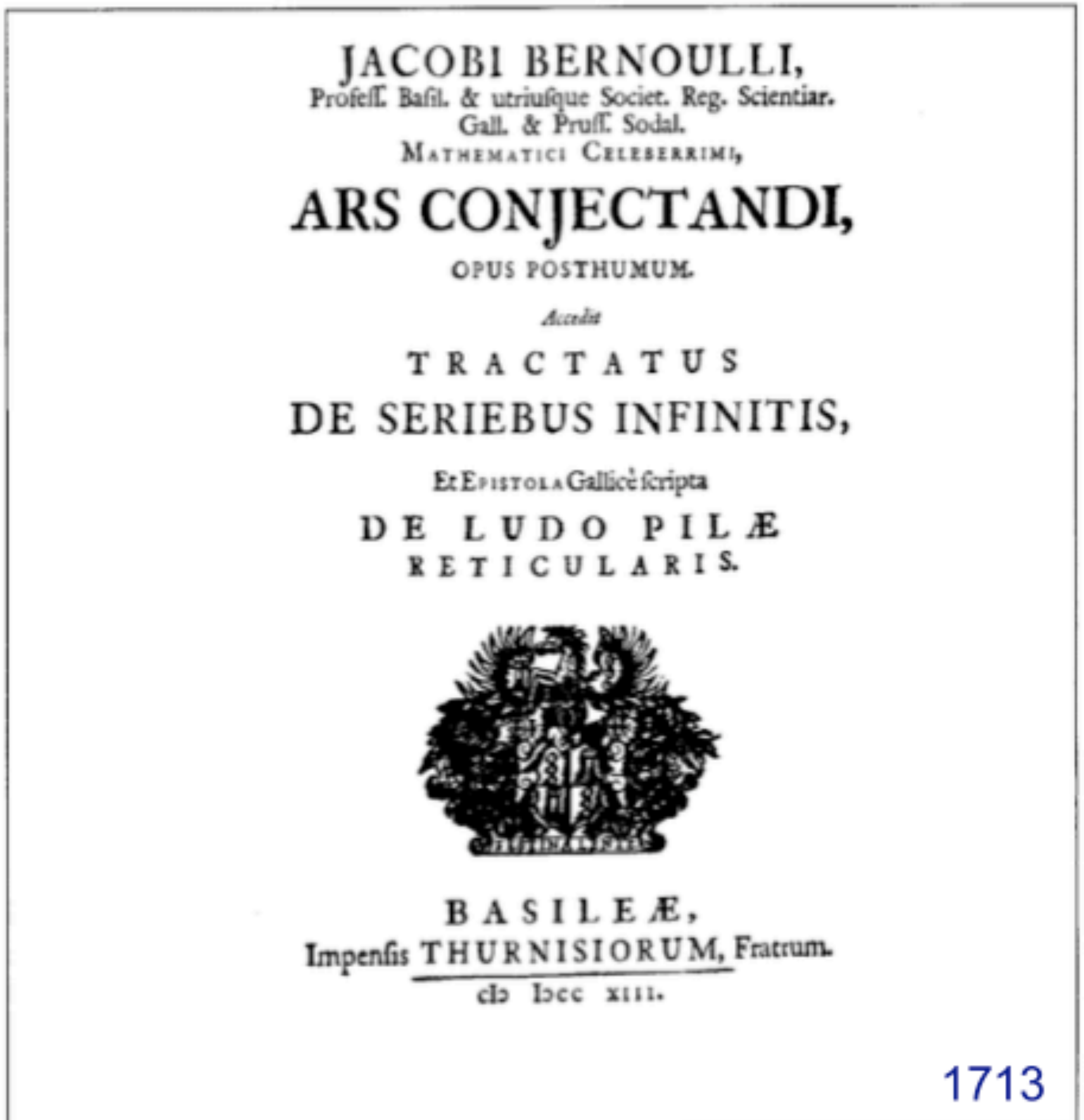
Analytical Statistics (19th/20th century)

introduction of concepts of probability calculus

(random) test sample replaces mass collection

subset allows drawing conclusions on full (data-)set

The Art of Estimating


JACOBI BERNOULLI,
Profess. Basil. & utriusque Societ. Reg. Scientiar.
Gall. & Pruss. Sodal.
MATHEMATICI CELEBERRIMI,
ARS CONJECTANDI,
OPUS POSTHUMUM.
Accedit
TRACTATUS
DE SERIEBUS INFINITIS,
Et EPISTOLA Gallicè scripta
DE LUDO PILÆ
RETICULARIS.

BASILEÆ,
Impensis THURNISIORUM, Fratrum.
cIↄ Iↄcc XIII.

1713

Stochastic Theory

= ars conjectandi

(Jacob Bernoulli)

Probability Calculus

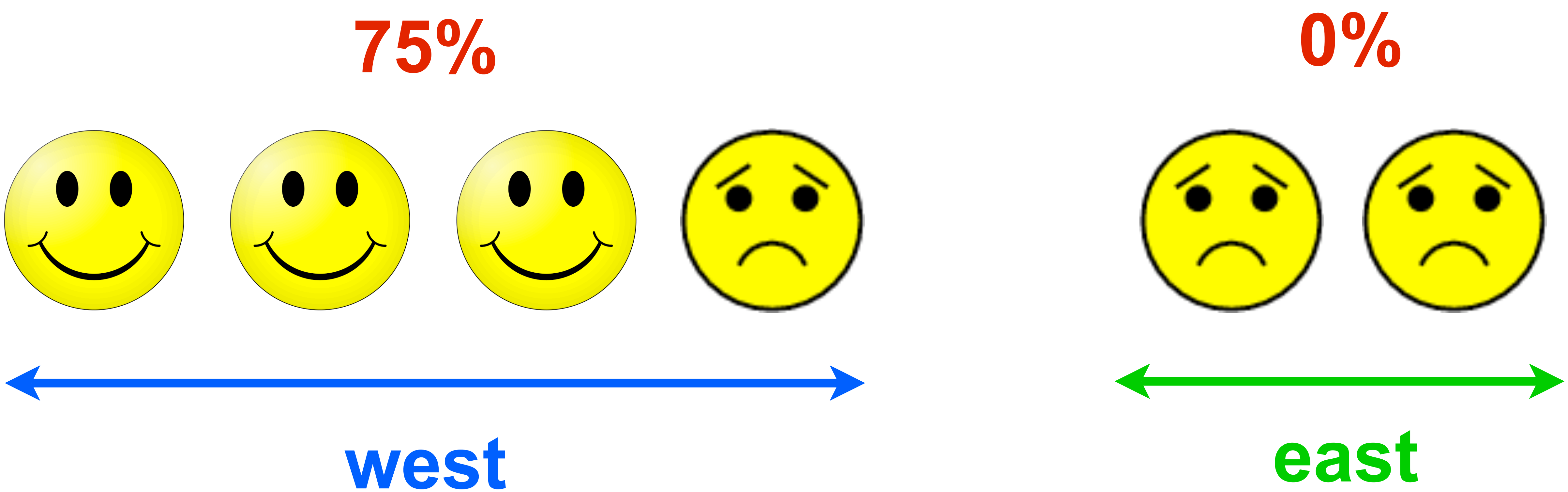
+

Analytical Statistics

Example

A market research company makes a survey for a new product.

"Are you satisfied with the new product?"



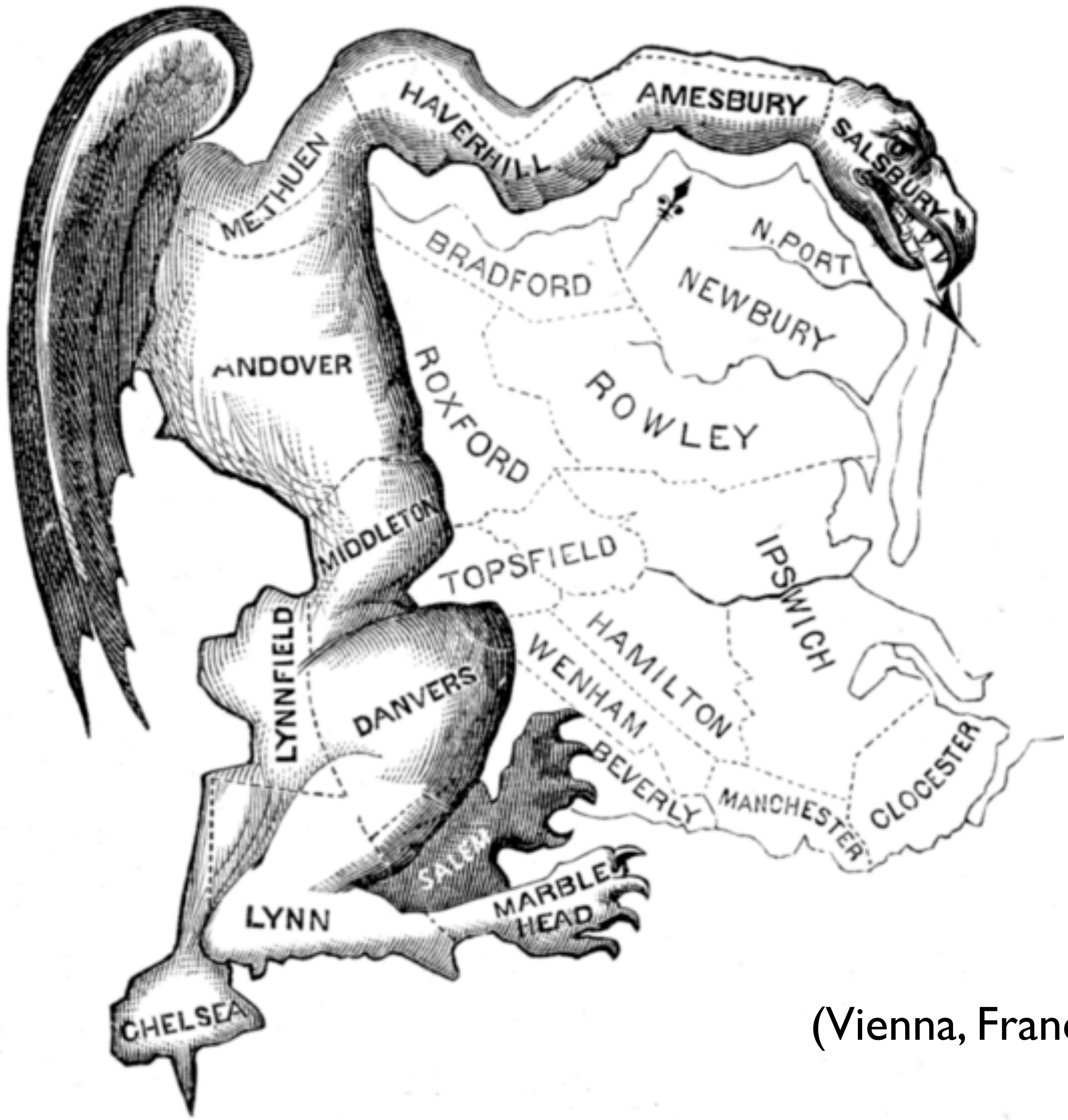
Example

A clever survey sales-man moves the border a bit.

25% more satisfied customers in east and west!

In the “Real” World: The Gerrymander

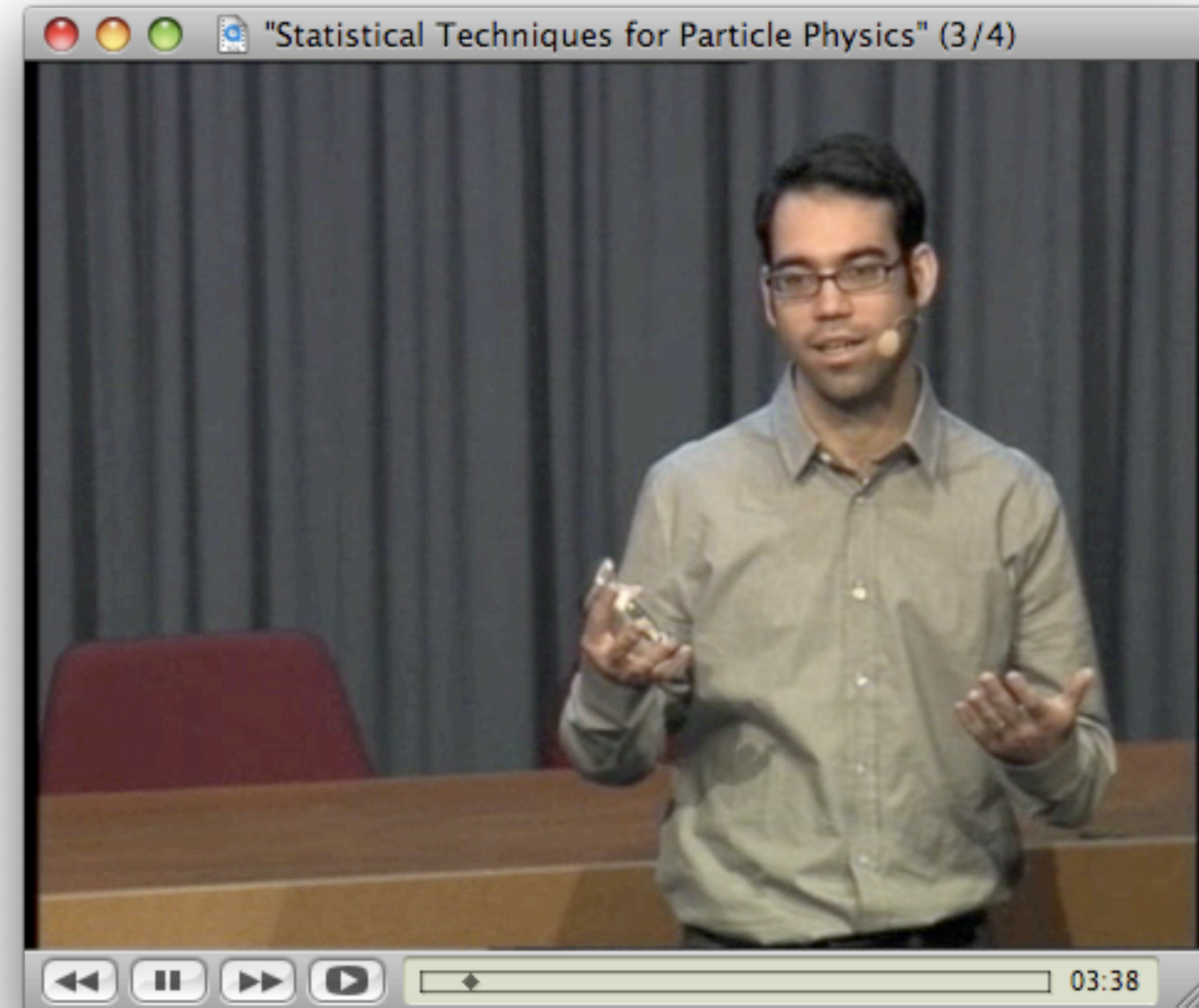
http://de.wikipedia.org/wiki/Gerrymandering



named after a governor of massachusetts in the early 19th century “Elbridge Gerry”

still practice, not only in US

(Vienna, France, Great Britain, North Ireland, Belgium)



second part follows lectures by Kyle Cranmer (ATLAS, NYU)

http://cdsweb.cern.ch/search?cc=Video+Lectures&ln=en&jrec=1&p=statistics

What does Probability mean?

Kolmogorov Axioms (1933):

For every event E of event space Ω a probability p(E) can be attributed.

1. *Probabilites are non-negative:* $p(E) \geq 0$
2. *Probability for the "certain" event:* $p(\Omega) = 1$
3. *If events are disjunct, probability for one OR the other is sum of probabilities.*

$$P(A \cup B) = P(A) + P(B) \quad \text{if } A \cap B = \emptyset$$

Frequentist vs. Bayesian

Likelihood Methods

Frequentist

probability defined as limit of long-term frequency

- *flip a coin 50-50*
- *roll a dice 1/6*
- *Monte Carlo methods*

P (Data | Theory)

conditional prob. data given theory

P (Theory | Data)

instructive example

$$P(A|B) \neq P(B|A)$$

P (pregnant | female) ≈ 3%

P (female | pregnant) >> 3%

Bayesian: Bayes' Theorem

cond. prob. of A given B — *prior Prob.*

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

posterior/ conditional probability of A given B

prior/ marginal Prob. (normalization)

P(A) is unknown! Subjective priors!
update your knowledge.

P (Data | Theory) **P (Theory | Data)**
using a prior!

Frequentist vs. Bayesian

Frequentist always restrict to statements:

P (Data | Theory)

deductive reasoning

Bayesian can address:

P (Theory | Data) $\propto$ P (Data | Theory) P (Theory)

inductive reasoning

needs prior on theory (subjective/ empirical (objective) priors)

How Likelihood Methods fit in

Frequentist always restrict to statements:

P (Data | Theory)

deductive reasoning

Bayesian can address:

P (Theory | Data) $\propto$ P (Data | Theory) P (Theory)

inductive reasoning

needs prior on Theory (subjective/ empirical (objective) priors)

Likelihood Methods **e.g. MINUIT/ MINOS**

approximately frequentist methods

enjoy nice properties of Bayesian without need of priors

A philosophical question!

Frequentist vs. Bayesian
reason for many heavy "philosophical" discussions
different, strong opinions exist under experts/ experiments

anonymous quotes:

"frequentist for discovery, bayesian for limits"
"bayesians tend to be aggressive and optimistic"
"frequentist statisticians are more cautious and defensive"

both are legitimate, scientific approaches!

language is important!
frequentists determine confidence intervals!
⇔ *interval covers true value 68% (95%) of the time*

bayesians infer credible intervals!
⇔ *posterior has prob. that true value inside (prior assumption)*

Jokes

A Bayesian is one who, vaguely expecting a horse, and catching a glimpse of a donkey, strongly believes he has seen a mule.

Some Basic Ingredients

Probability Density Functions:

Gaussian (Normal), Log-Normal, Poisson, Binomial, ...

$$\text{Mean value}: \quad \overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\text{Variance}: \quad V(x) = \frac{1}{N}\sum_{i=1}^{n} (x_i - \bar{x})^2 = \overline{x^2} - \bar{x}^2$$

$$\text{Standard deviation}: \quad \sigma = \sqrt{V(x)} = \sqrt{\overline{x^2} - \bar{x}^2}$$

Cumulative Density Functions:

Confidence Level, p-value

Probability Density Functions (PDFs)

$$P(x \in [x, x+dx]) = f(x)dx$$

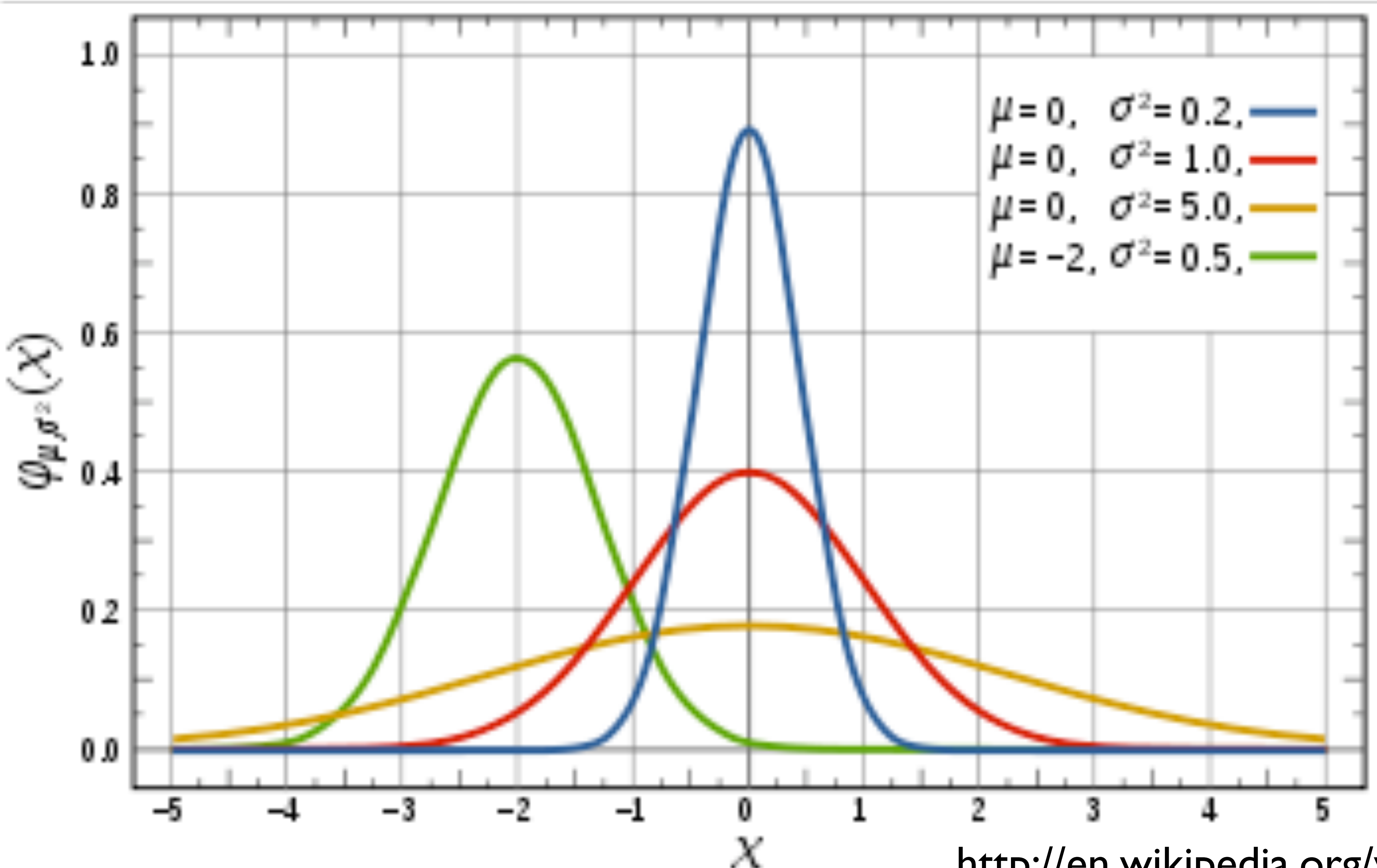
f(x) not a probability!

but $$\int_{-\infty}^{\infty} f(x)dx = 1$$

obey 2. axiom from Kolmogorov

The Gaussian PDF

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$



normal distrib.

mean = μ

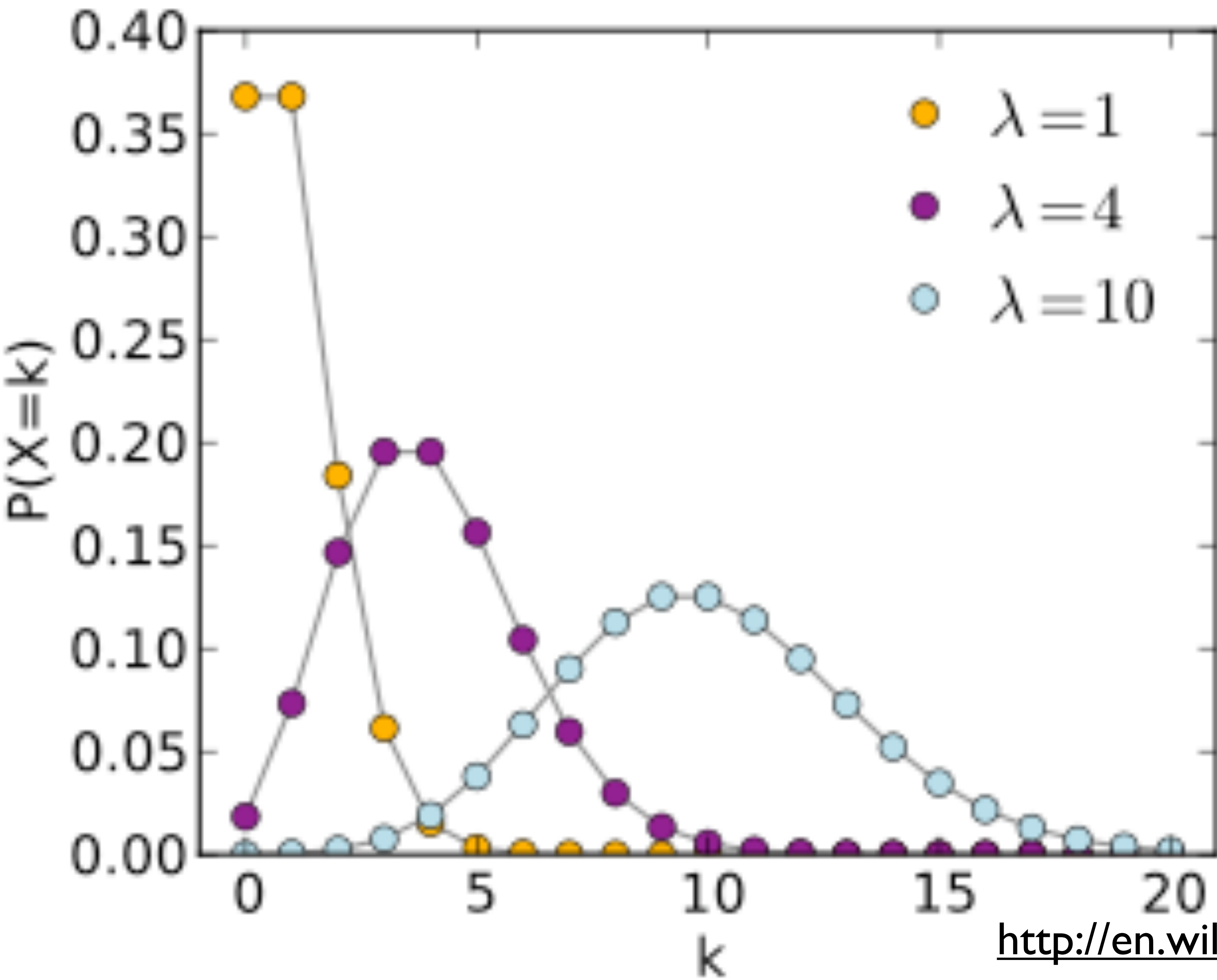
variance = σ^2

continous

http://en.wikipedia.org/wiki/Normal_distribution

The Poisson PDF

$$f(x|\mu) = \frac{\mu^x}{x!}\exp(-\mu)$$



mean = μ

variance = μ

x = 0, 1, 2, ... ; μ > 0

discrete

http://en.wikipedia.org/wiki/Poisson_distribution

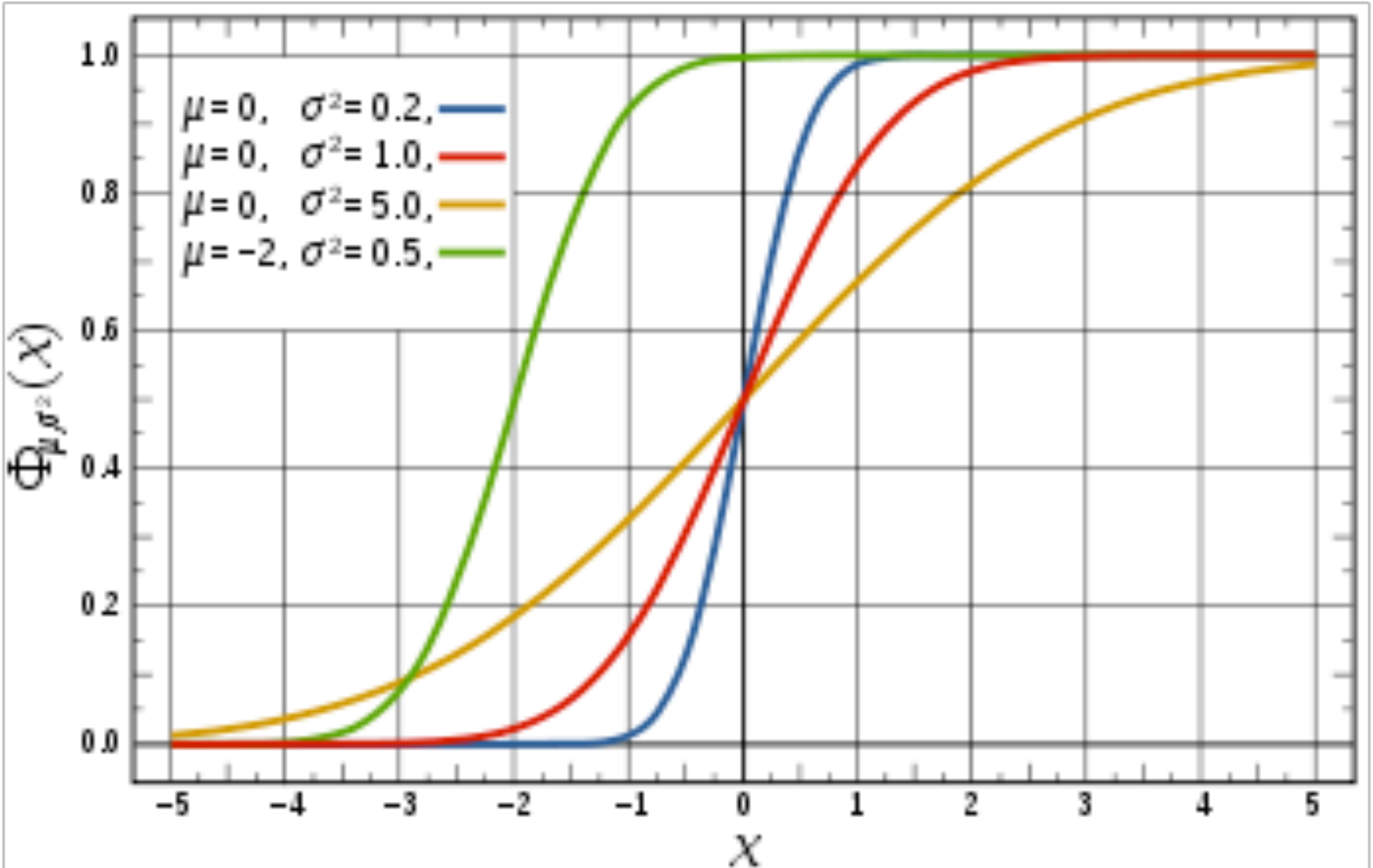
Cumulative Density Functions

german: "Verteilungsfunktion"

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

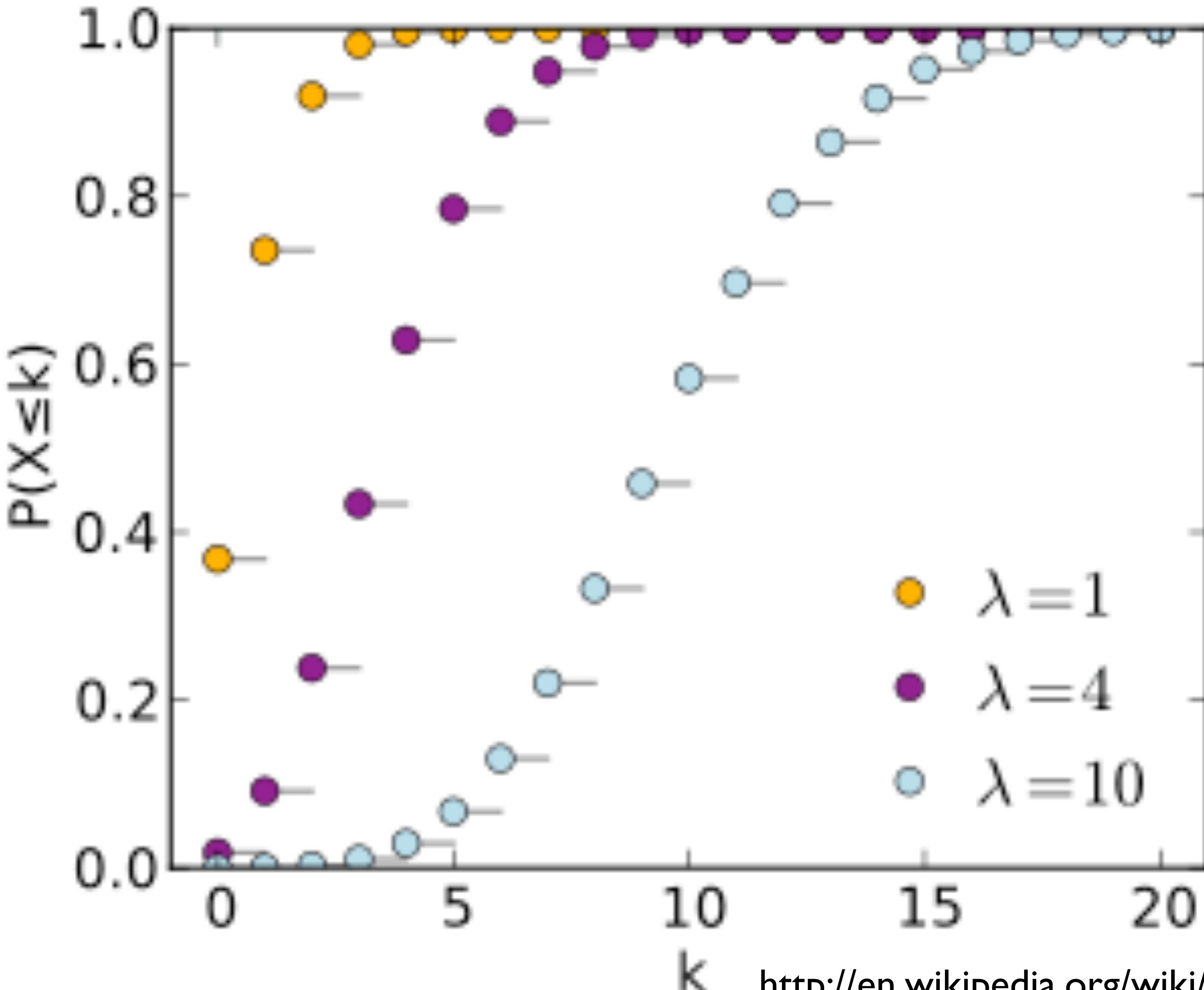
for the continous case

The Gaussian CDF



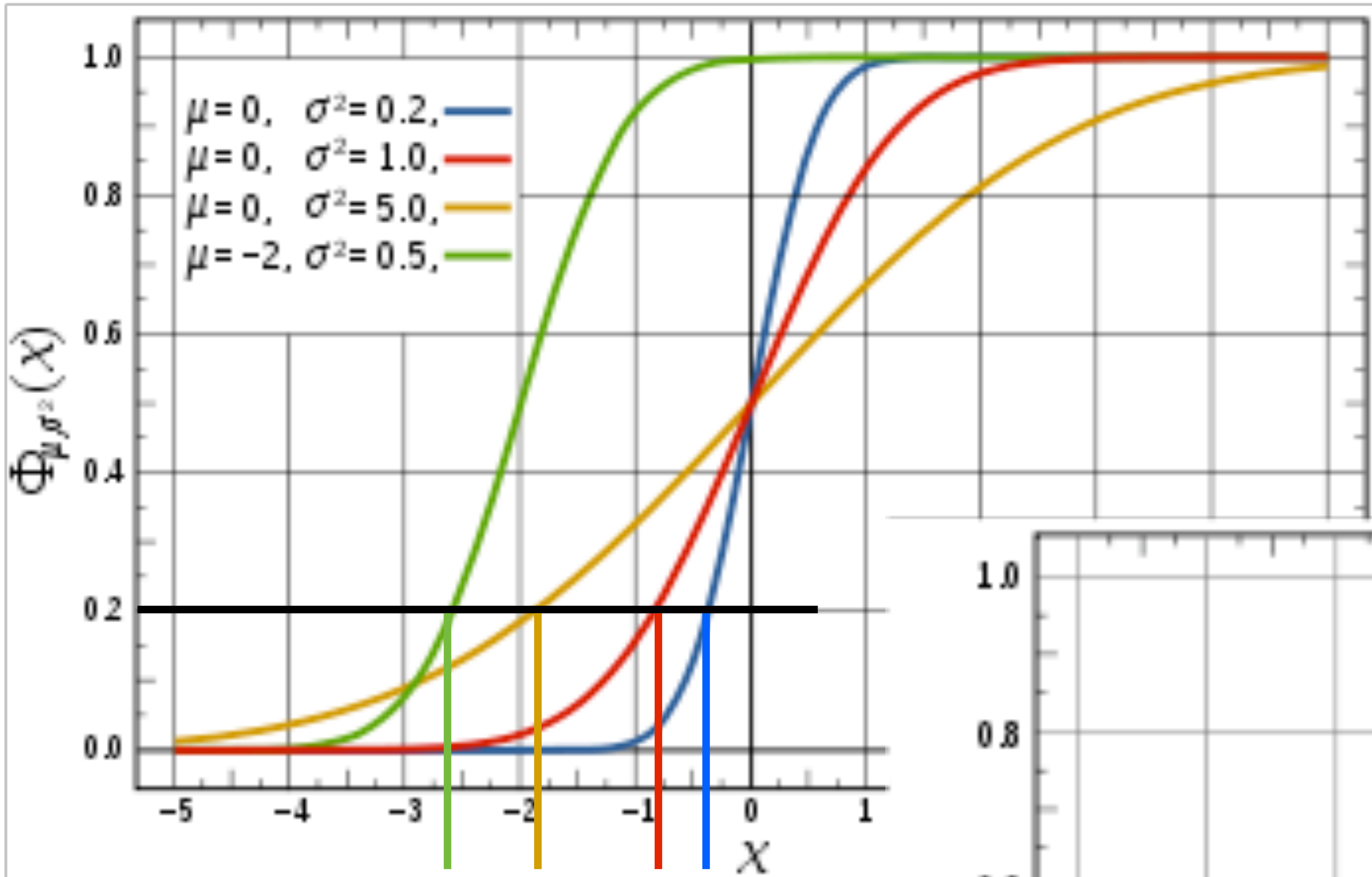
http://en.wikipedia.org/wiki/Normal_distribution

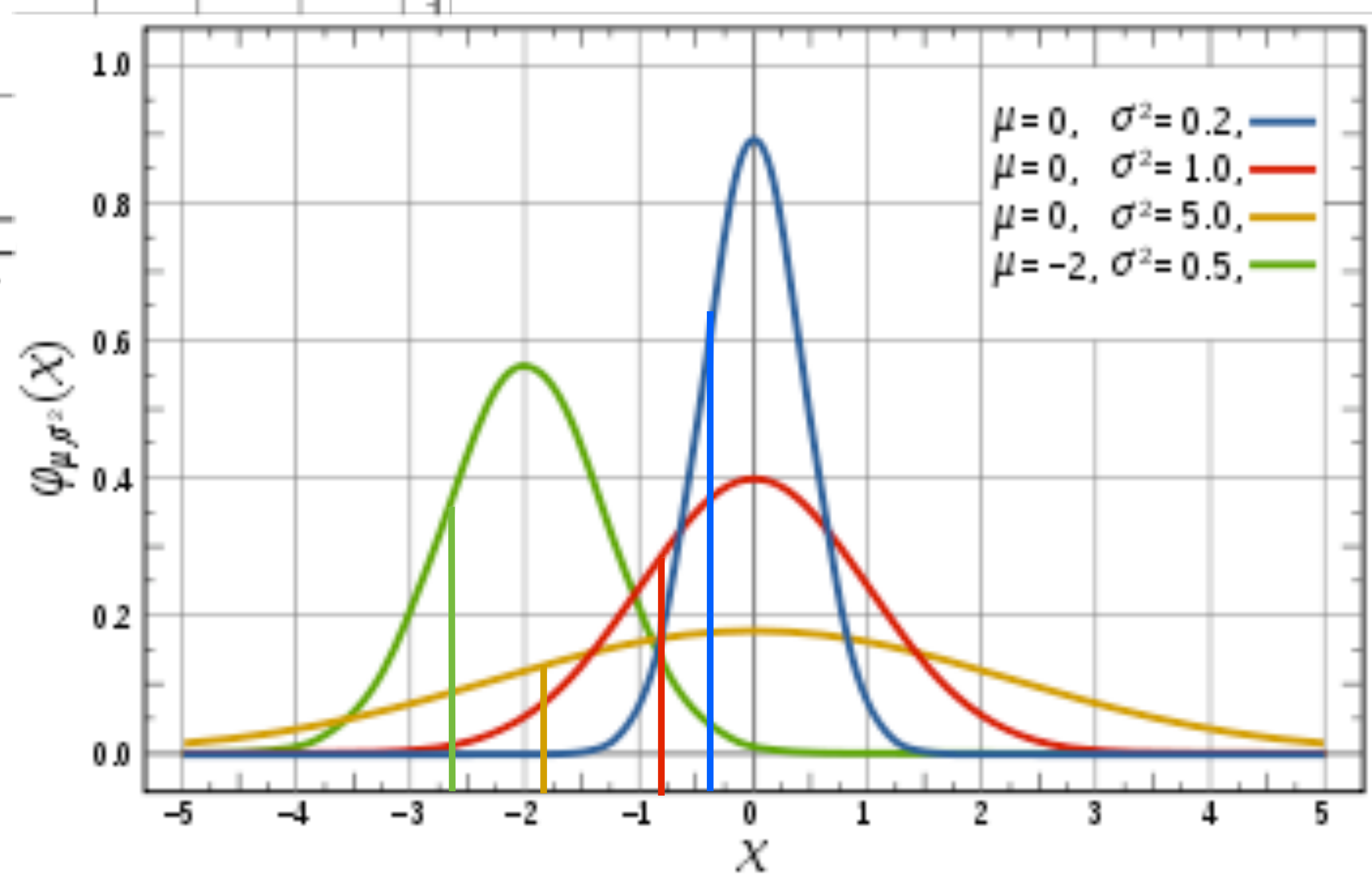
The Poisson CDF



http://en.wikipedia.org/wiki/Poisson_distribution

Example: Gaussian PDF + CDF





let's see where 20% quantiles are

Confidence Level - one sided

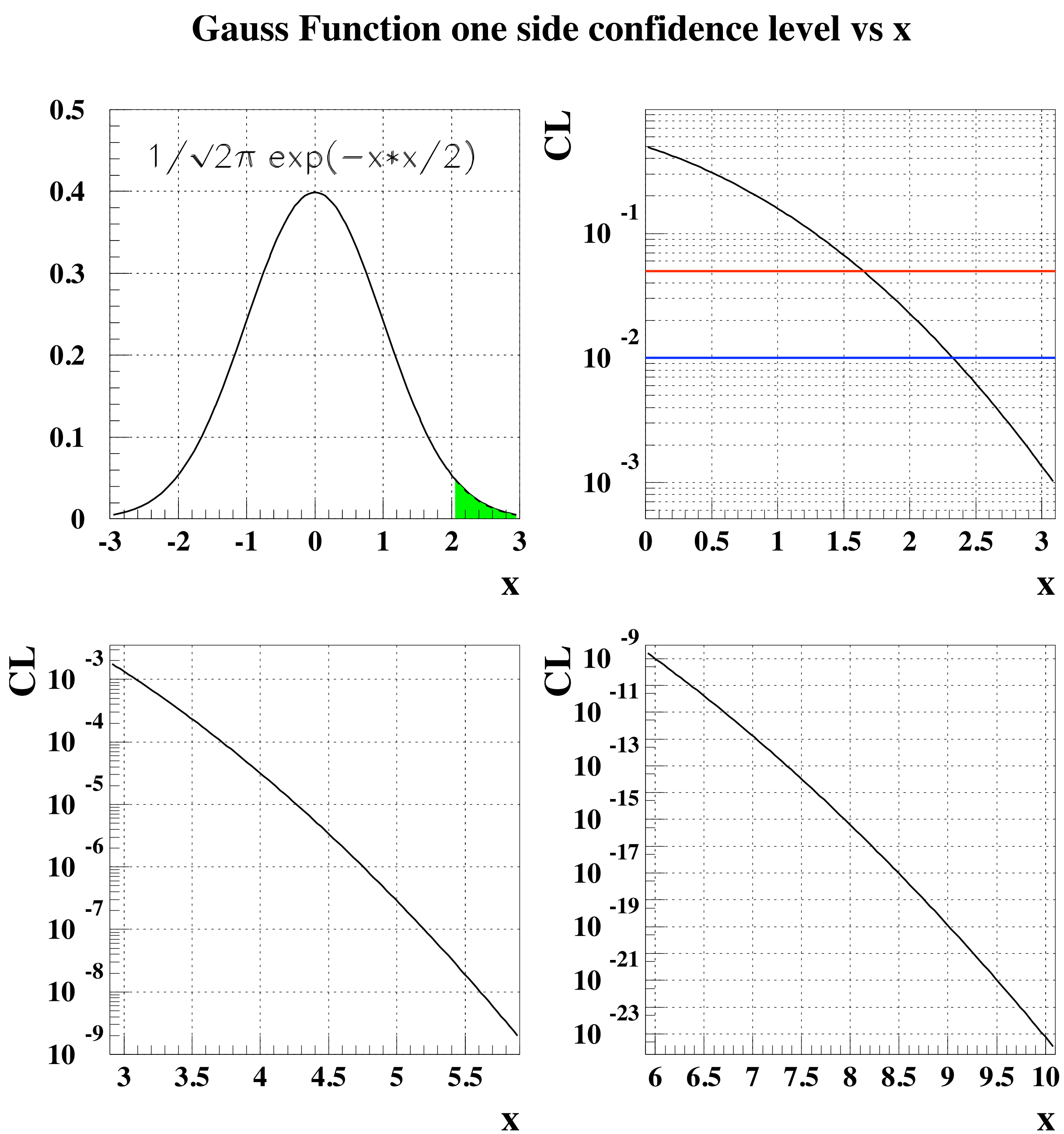
$$CL(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-x'^2/2) dx'$$

x is sigma deviation

CL(x) is p-value

5-sigma $\Leftrightarrow$ p = $2.9 \cdot 10^{-7}$

(1 - p) is Confidence Level



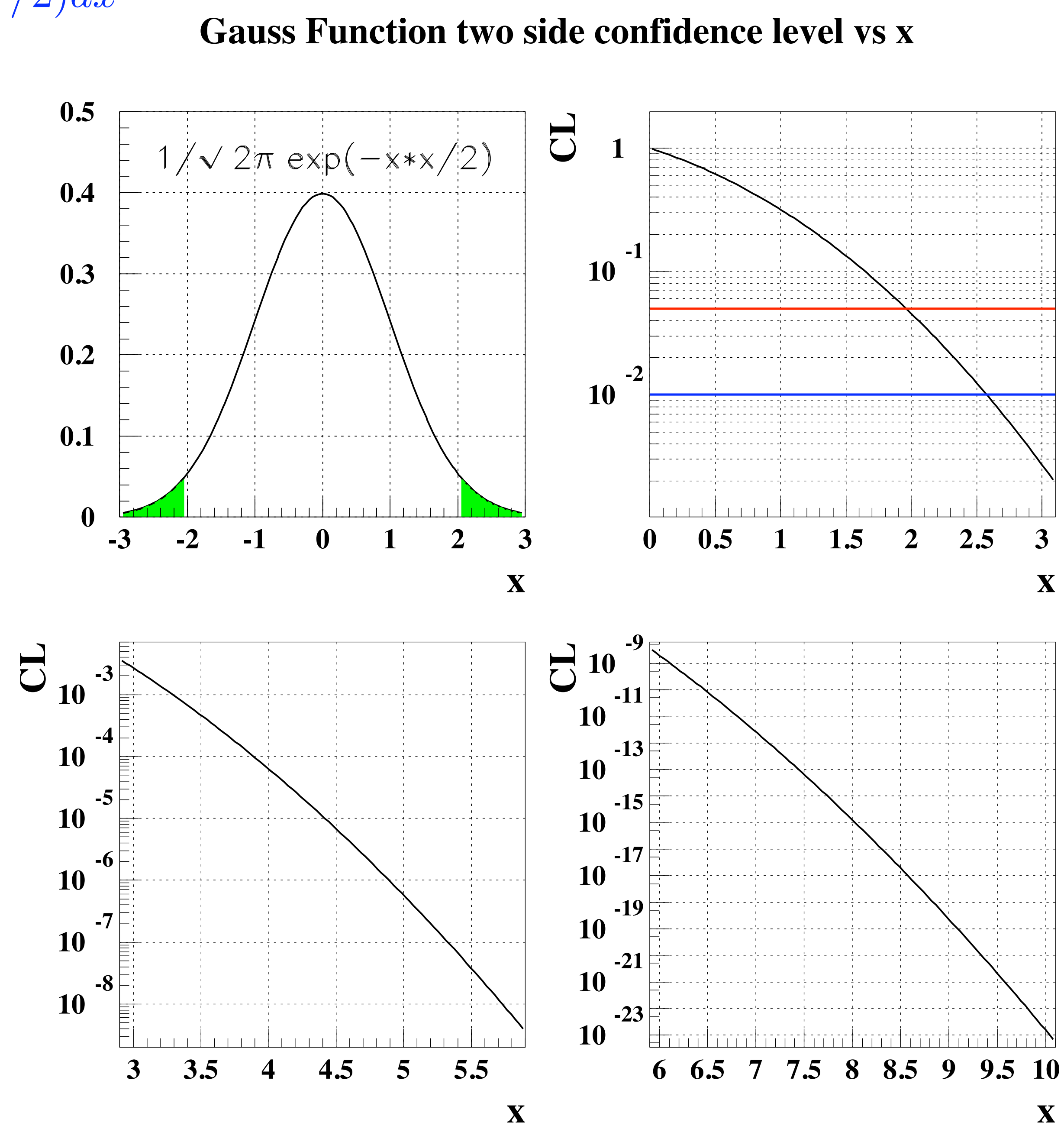
thanks to

O. Behnke, C. Kleinwort, S. Schmitt (DESY), from Terascale Statistics School 2008 exercises

Confidence Level - two sided

$$CL(x) = \int_{-\infty}^{-x} \frac{1}{\sqrt{2\pi}} \exp(-x'^2/2)dx' + \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-x'^2/2)dx'$$

p-value is 2·p(one sided)



thanks to

O. Behnke, C. Kleinwort, S. Schmitt (DESY), from Terascale Statistics School 2008 exercises

Confidence Interval Construction

"inverted" Hypothesis test (only short here, more tomorrow)

use PDFs for H_0 background-only

H_1 signal+background

accept or reject H_0 with measurement n events!

$P(n \mid H_0) = P(n \mid b)$

$P(n \mid H_1) = P(n \mid s + b)$

$$P(n \mid \mu) = P(n \mid \mu s + b)$$

Parameter of Interest μ

Likelihood function for μ

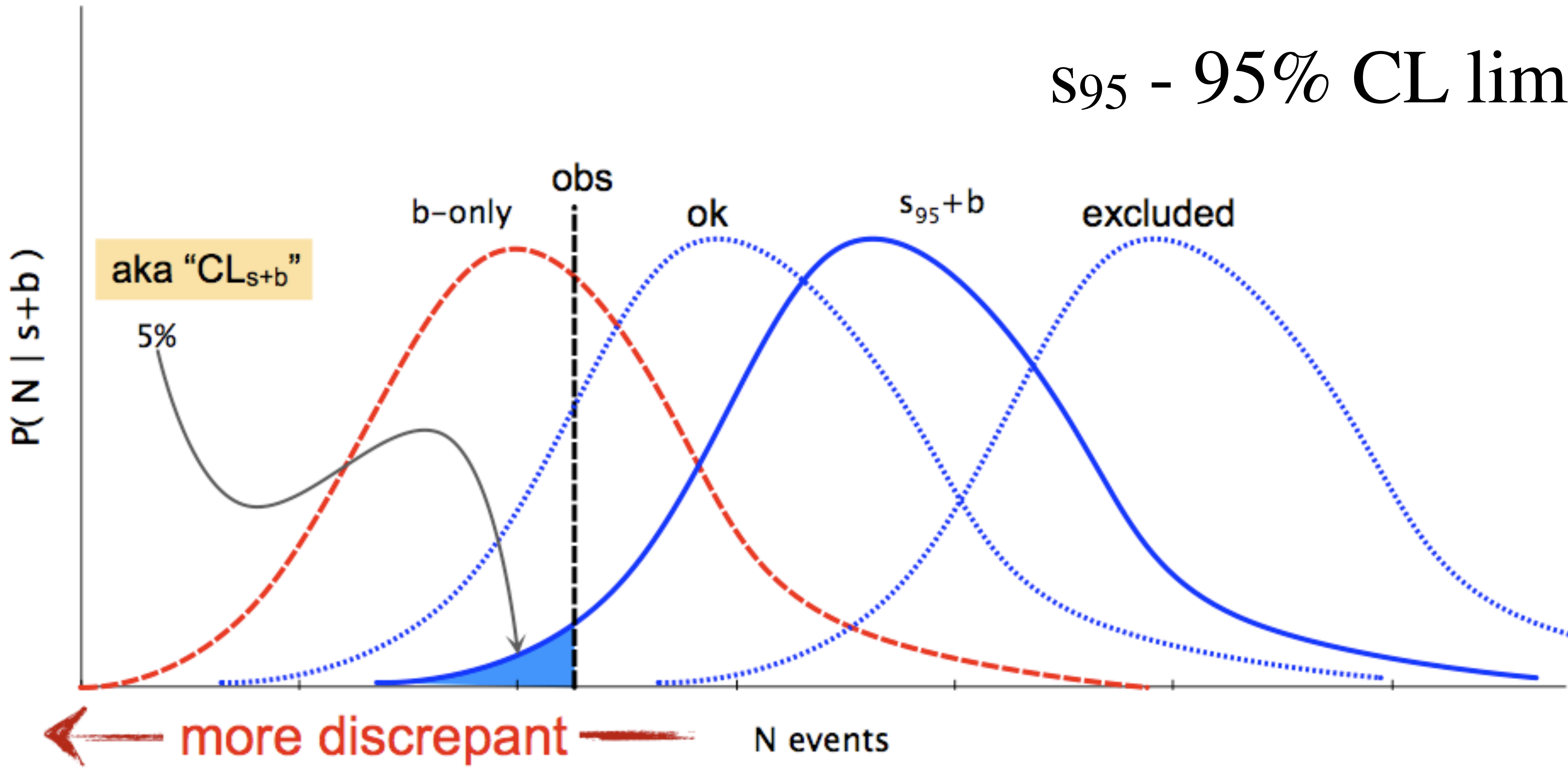
$\mu = 0 : H_0$

$\mu \neq 0 : H_1$

Confidence Intervals - CL_{s+b}

determine s+b consistent with observation (frequentist)
assume b is known! do toy monte carlo sufficiently often,
intervall $(0, s_{95})$ covers obs. value $(1 - p) = 95\%$ of the time
CL_{s+b} - confidence level for signal + background

s_{95} - 95% CL limit

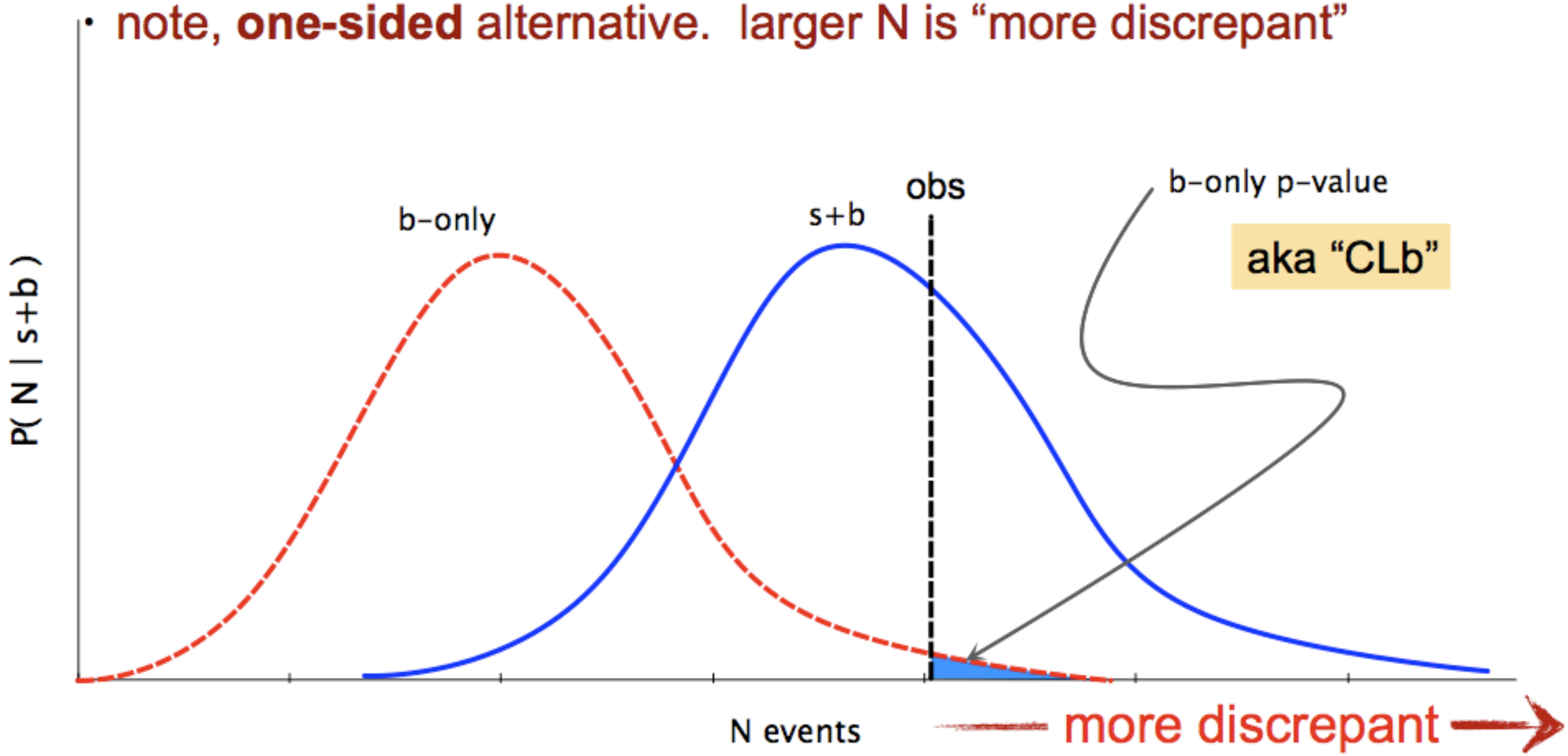


Confidence Intervals - CL_b

determine background fluctuation probability (frequentist)
assume b-only is known!
intervall $(0, s_{95})$ covers obs. background $(1 - p) = 95\%$ of the time
CL_b - confidence level for background only

Discovery: test b-only (null: s=0 vs. alt: s>0)

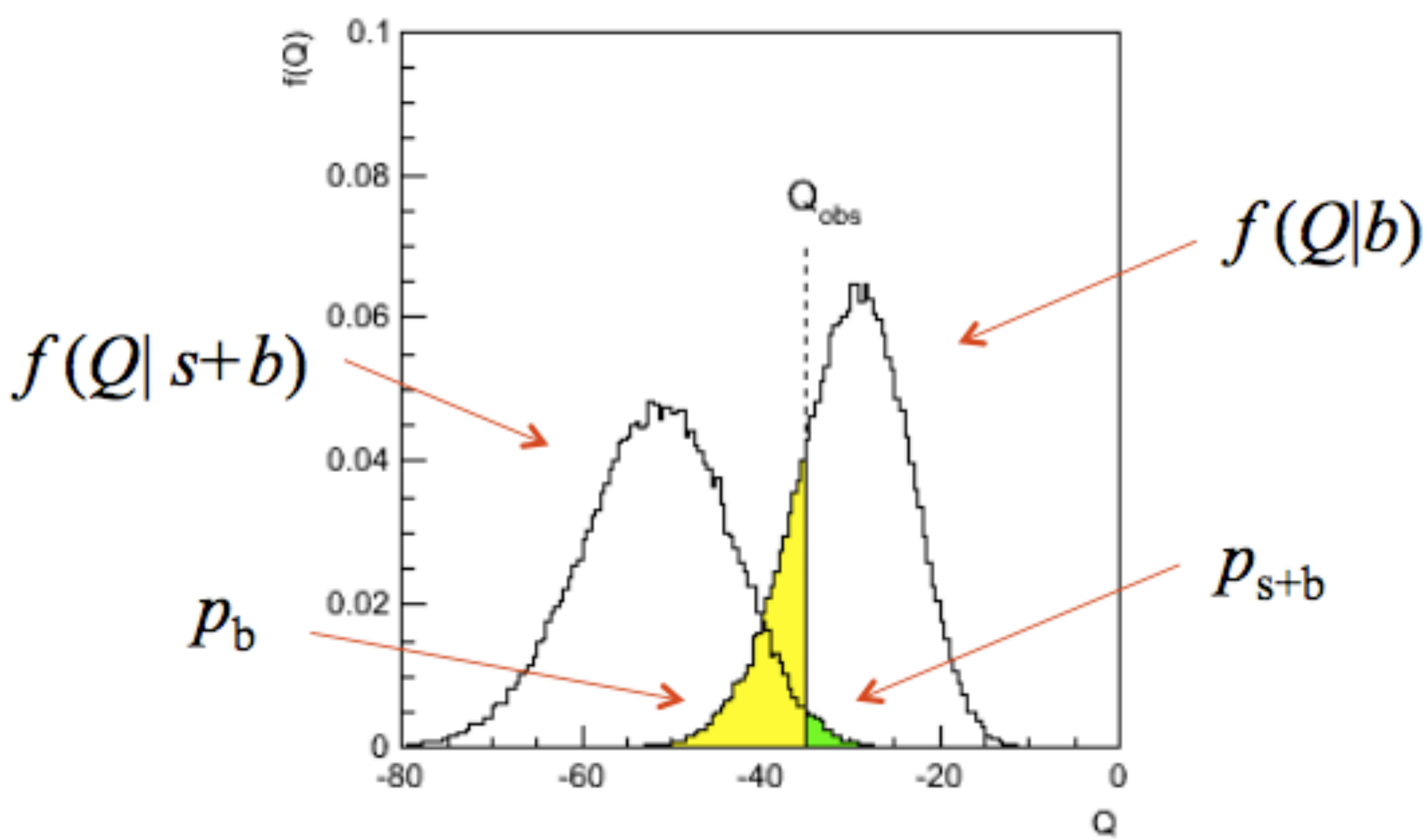
- note, **one-sided** alternative. larger N is "more discrepant"



Modified Frequentist Method - CL_s

The CL_s procedure

In the usual formulation of CL_s, one tests both the $\mu = 0$ (b) and $\mu = 1$ ($s+b$) hypotheses with the same statistic $Q = -2\ln L_{s+b}/L_b$:



use ratio

$CL_s = CL_{s+b}/ CL_b$

intervall $(0, s_{95})$ covers obs. value 95% of the time

"save" wrt background flucltutations

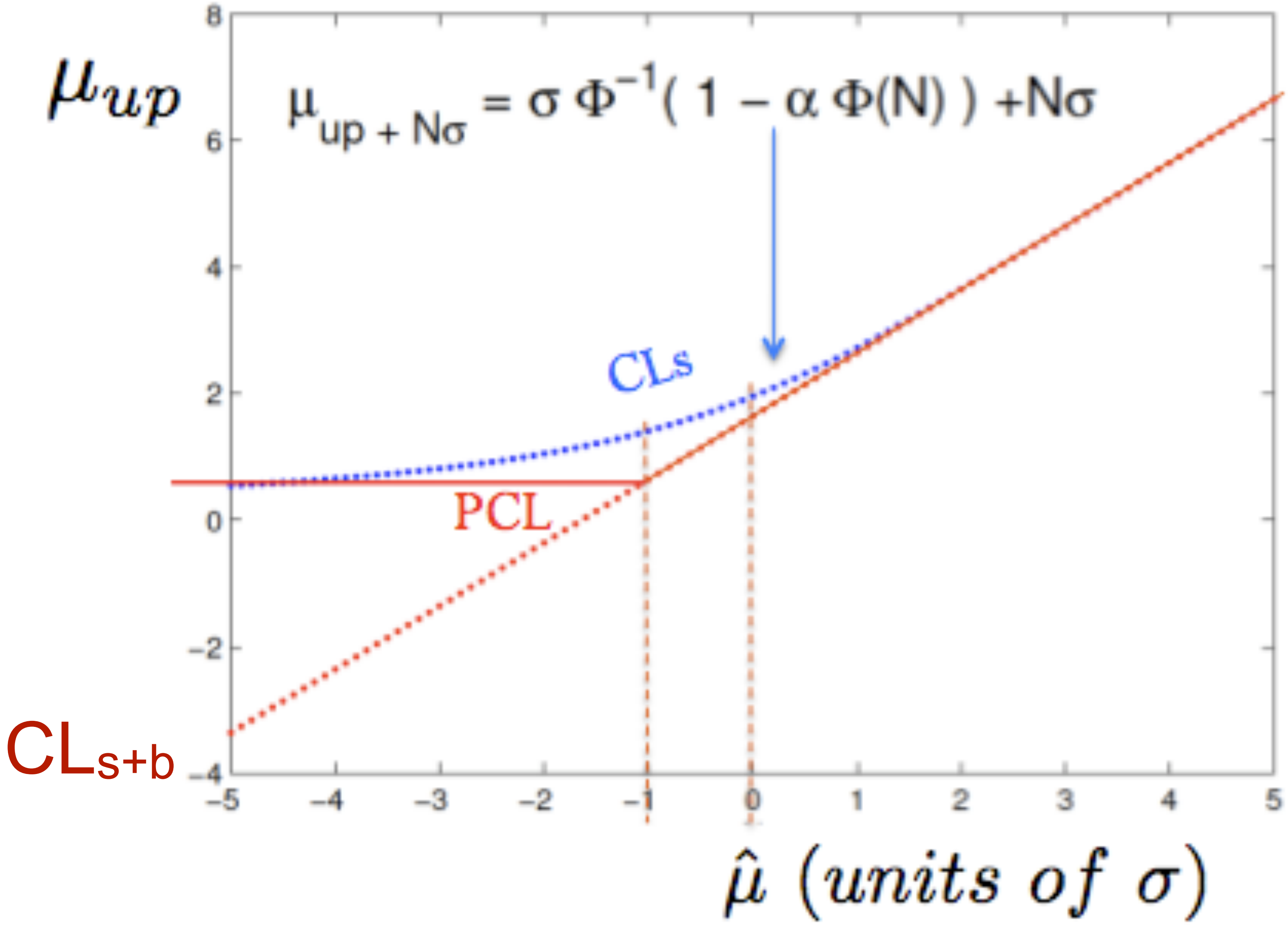
G. Cowan ATLAS Limits Workshop / PCL 11

CL_s vs. CL_{s+b} vs. PCL limits

PCL Power Constraint Limits

when background fluctuates down, CL_{s+b} gets negative

PCL more conservative wrt CL_S



from Ellam Gross Statistics Workshop ATLAS 14.April2011

Further Reading

A Unified Approach to the Classical Statistical Analysis of Small Signals
Gary J. Feldman, Robert D. Cousins
http://arxiv.org/abs/physics/9711021

WIKIPEDIA ARTICLES (frequentist vs. bayesian)

http://en.wikipedia.org/wiki/Statistical_hypothesis_testing

http://en.wikipedia.org/wiki/Bayesian_inference

CDSWEB VIDEO LECTURES (G. COWAN, K. CRANMER, B. COUSINS, ...)

http://cdsweb.cern.ch/search?cc=Video+Lectures&ln=en&jrec=1&p=statistics

ATLAS INFORMATION/ RECOMMENDATIONS

https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ATLASStatisticsFAQ

Summary

Statistics/ Probability

Probability Distributions

p-Value / Confidence Level

Frequentist/ Bayesian

Confidence Intervals   *Bayes' Law*

CL_s

"Bayesians address questions everyone is interested in, by using assumptions no-one believes"

"Frequentists use impeccable logic to deal with an issue of no interest to anyone."

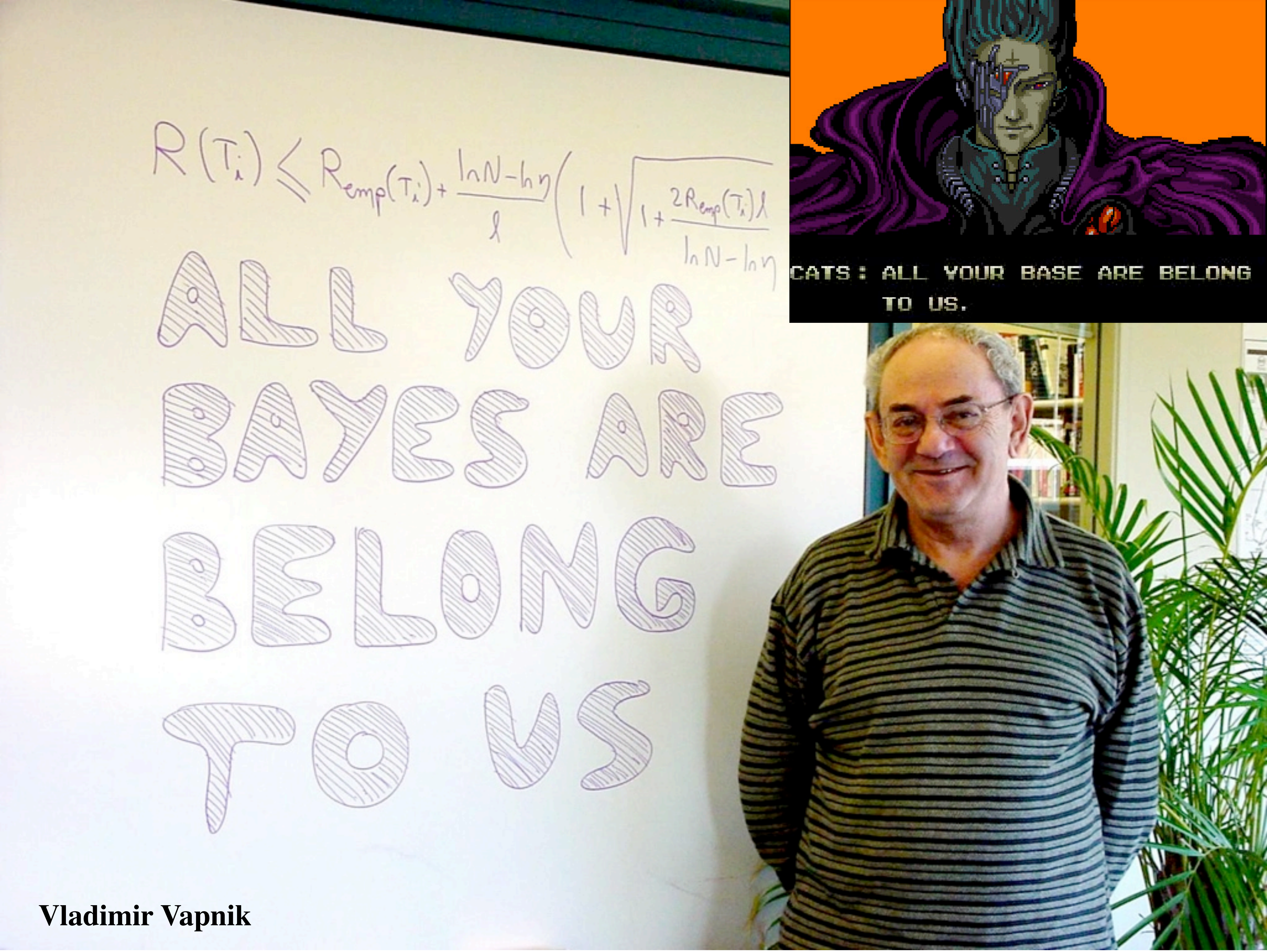
Louis Lions

Terascale Statistics School 2008
29. Sep. - 2. Oct. 2008
DESY Hamburg

Practical work

paper exercises

Authors: O. Behnke (DESY), C. Kleinwort (DESY), S. Schmitt (DESY)



Vladimir Vapnik