Alberto Rotondi
Università di Pavia - Scuola di Alghero 2009

- Statistica 1: Neyman vs Bayes.
  frequenze in esp. di conteggio

- Statistica 2: likelihood ratio e test di ipotesi
  segnale su fondo

- Track fitting: tracking in GEANT3 &GEANT4
  filtri di Kalman e global fitting

Frequentist approach

Bayesian approach

What is Statistics ?

- ***a problem of probability calculus*: if $p = 1/2$ for having head in tossing a coin, what is the probability to have in 1000 coin tosses less than 450 heads?**

- ***the same problem in statistics*: if in 1000 coin tosses 450 heads have been obtained, what is the estimate of the true head probability?**

Statistical error: $s \approx \sigma$

$$\mu \pm \sigma = 500.0 \pm 15.8 \simeq 500 \pm 16 = [484, 516]$$

$$x \pm s = 450.0 \pm 15.7 \simeq 450 \pm 16 = [434, 466]$$

Statistics

We have 2 inferences

- **parameter estimation: to estimate p from 1000 coin tosses**
- **hypothesis testing: in in two experiments of 1000 coin tosses 450 and 600 tosses have been obtained, how much is probable that the two experiments use two consistent coins?**

Parametric Statistics: the probability depends on θ:

$$\mathcal{E}(\theta) \equiv (S, \mathcal{F}, P_\theta)$$

corresponding to a density

$$P\{X \in A\} = \int_A p(x;\theta)\,\mathrm{d}x$$

The hystorical path

	FREQUENTISTS	BAYESIANS
1763		Thomas Bayes writes a fundamental paper. Bayesian age
1900	Karl Pearson proposes the χ^2 test	
1910	Robert Fisher invents Maximum Likelihood	
1937	The J. Neyman frequentist interval estimate	
1940	The Hypothesis testing of Pearson. Frequentist age The Popper scheme Frequentist teaching	
1990		rediscovering of the bayesian works of Jeffreys, De Finetti and Jaynes
now	the debate is open: see on Confidence Limits	the CERN Workshop (Geneva 2000) neo-Bayesian age?

PHYSTAT 05 - Oxford 12th - 15th September 2005

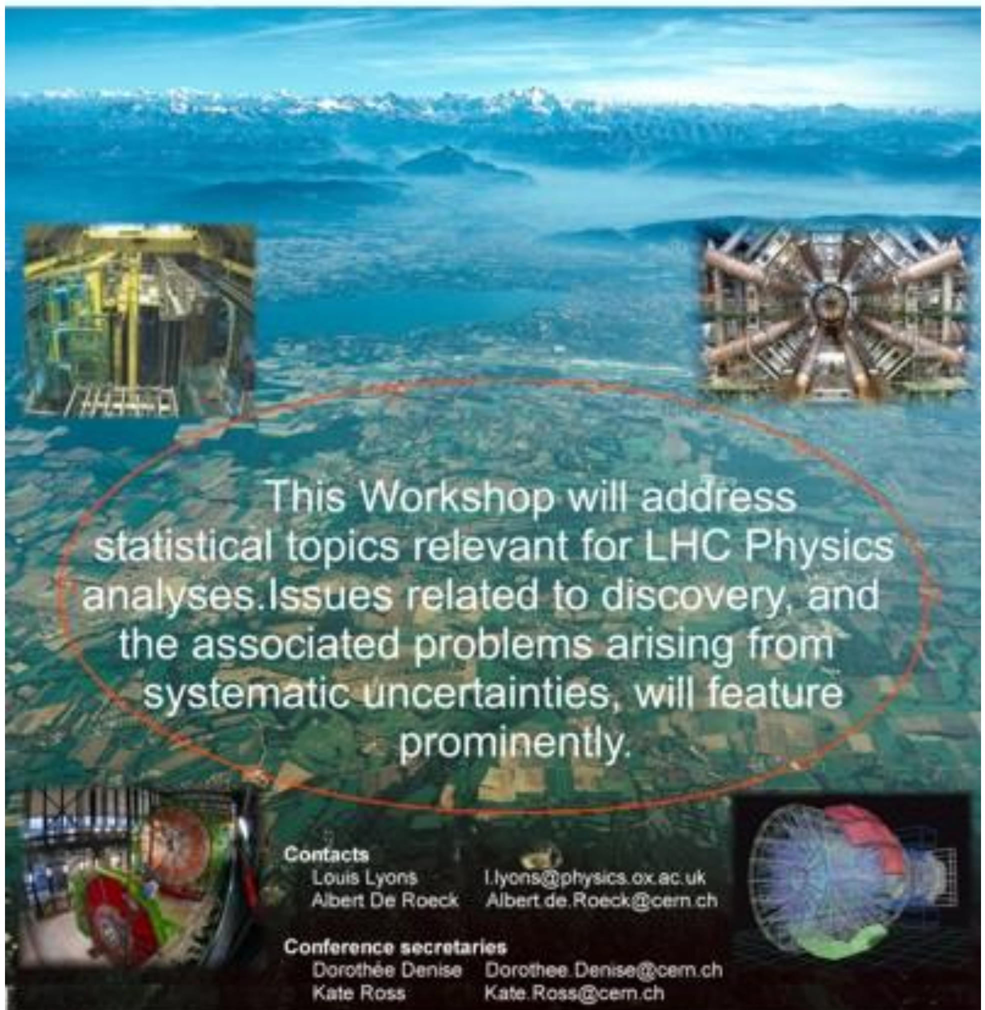
Statistical problems in Particle Physics, Astrophysics and Cosmology

PHYSTAT-LHC Workshop

on

Statistical Issues for LHC Physics

CERN Geneva June 27-29, 2007



Further information and registration at http://cern.ch/phystat-lhc

Physics and Statistics

- **Higgs mass (PDG 2000):**

$$m > 95.3\ GeV, CL = 95\%$$

- W **mass:**

$$m_W = 80.419 \pm 0.056\ GeV$$

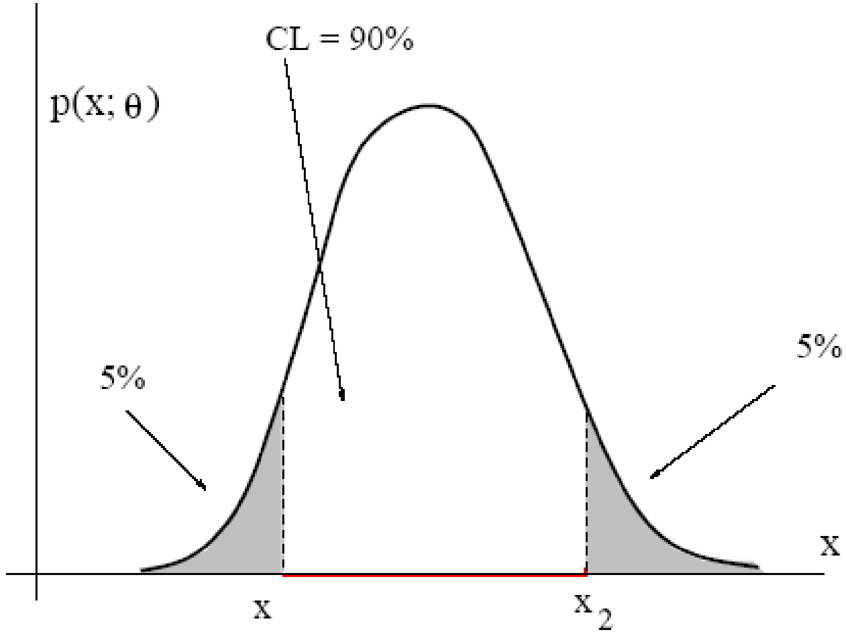
These are
confidence intervals

Frequentist Confidence Intervals

One (Neyman, 1937) starts from probability calculus

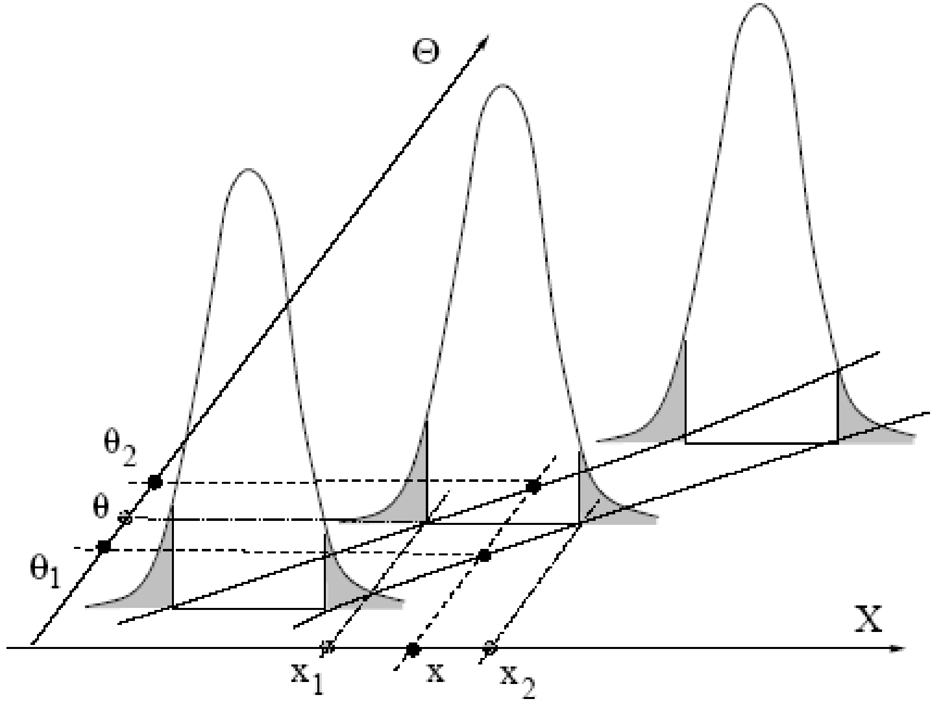
$$\int_{x_1}^{x_2} p(x;\theta)\,\mathrm{d}x = CL$$

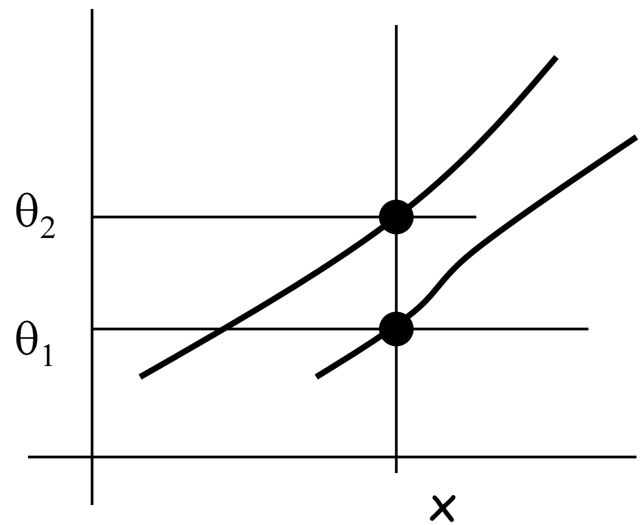
and the procedure is repeated



for all the possible θ values

Frequentist confidence intervals





It is possible to show that

$$X \in [x_1, x_2] \text{ iff } \Theta \in [\theta_1, \theta_2]$$

Since

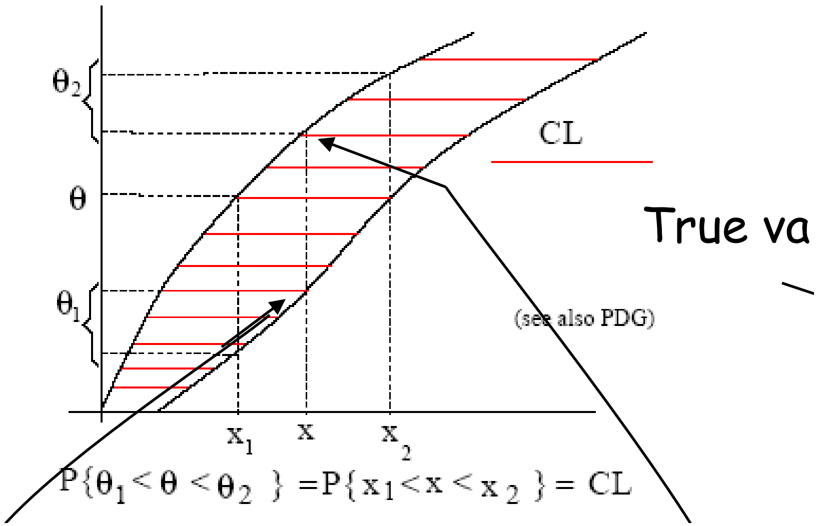
$$P\{X \in [x_1, x_2]\} = CL$$

then

$$P\{\Theta \in [\theta_1, \theta_2]\} = CL$$

Fundamental Neyman result (1937)

Cut and top views of the Neyman construction:

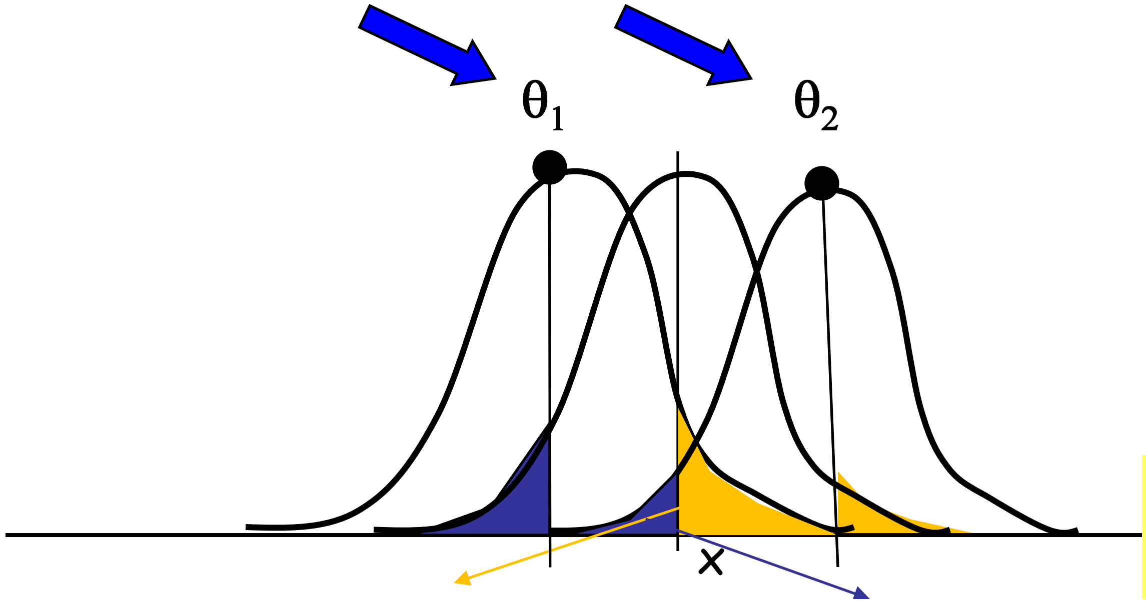


$x = x_1 \rightarrow \theta_1 < \theta < \theta$

$x = x_2 \rightarrow \theta < \theta < \theta_2$

$\theta_1 < \theta < \theta_2$ when $x_1 < x < x_2$

$P(\theta_1 < \theta < \theta_2) = P(x_1 < x < x_2) = CL$



NEYMAN INTEGRALS

Important:

$$\int_{\theta_1}^{\theta_2} p(\theta; x)\, \mathrm{d}\theta = CL$$

Elementary statistics may be WRONG!!

x

Pivot quantities

Avoid the calculation of the integrals

$$\int_A p(x;\theta)\,\mathrm{d}x = c_i$$

If $Q(x;\theta)$ is pivotal, $P\{Q \in A\}$ is independent of θ. Example:

$$Q = (X-\theta) \sim N(0,\sigma^2)$$

Method:

- **find** $P\{q_1 \le Q \le q_2\} = CL$ **;**
- **invert the equation:**

$$Q(x;\theta) = q \rightarrow \theta = T(x;q)$$

- **Then:**

Because P{Q} does not contain the parameter!

$$P\{q_1 \le Q \le q_2\} = P\{T_1 \le \theta \le T_2\} = CL$$

$$P\{\mu-\sigma \le X \le \mu+\sigma\} = P\{-\sigma \le X-\mu \le \sigma\}$$

$$= P\{X-\sigma \le \mu \le X+\sigma\}$$

Estimation of the sample mean

$$\mathrm{Var}[M] = \mathrm{Var}\left[\frac{1}{N}\sum_{i=1}^{N} X_i\right] = \frac{1}{N^2}\sum_{i=1}^{N}\mathrm{Var}[X_i] \ .$$

since $\mathrm{Var}[X_i] = \sigma^2 \quad \forall i,$

$$\mathrm{Var}[M] = \frac{1}{N^2}\sum_{i=1}^{N}\sigma^2 = \frac{1}{N^2}N\sigma^2 = \frac{\sigma^2}{N} \ .$$

Due to the Central Limit theorem we have a pivot quantity when N>>**1**

$$\frac{\mu - M}{\sigma/\sqrt{N}} \sim N(0,1)$$

Hence:

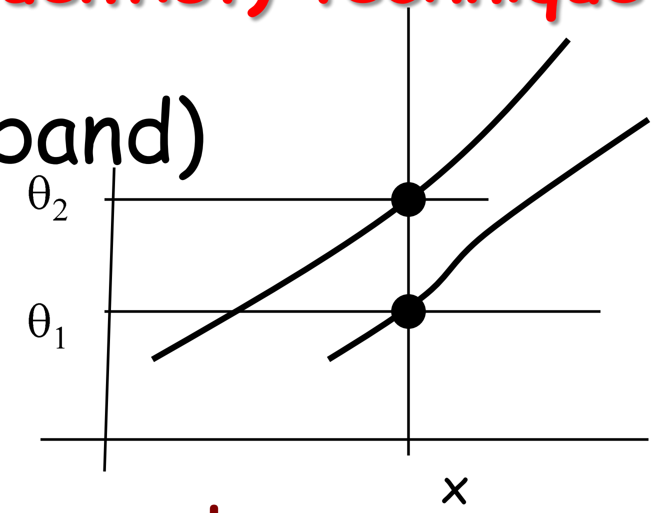
$$P\left\{\left|\frac{\mu - M}{\sigma/\sqrt{N}}\right| \leq 1\right\} = P\left\{M - \frac{\sigma}{\sqrt{N}} \leq \mu \leq M + \frac{\sigma}{\sqrt{N}}\right\}$$

$(N > 20-30)$:

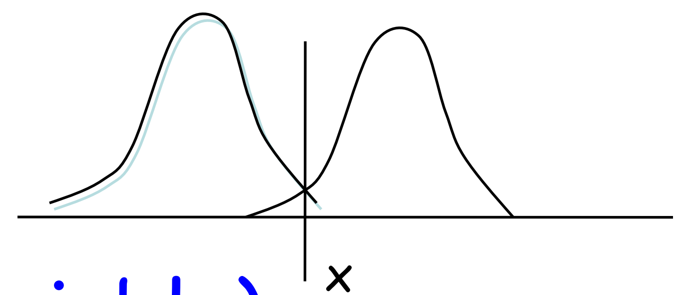
$$\mu = m \pm \frac{\sigma}{\sqrt{N}} \simeq \mu = m \pm \frac{s}{\sqrt{N}} \qquad CL \simeq 68\%$$

Hence, we have three methods to find confidence intervals with the Neyman (frequentist) technique:

- Graphical method (Neyman band)



- Analytic with the Neyman Integrals (Clopper Pearson method)



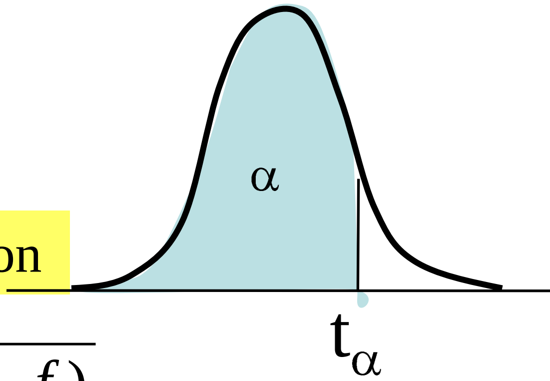
- Inversion method (pivot variable)

$$P\{\mu-\sigma \le X \le \mu+\sigma\} = P\{-\sigma \le X-\mu \le \sigma\}$$
$$= P\{X-\sigma \le \mu \le X+\sigma\}$$

Counting experiments: Binomial case

t=1, area 84%
Quantile α=0.84
P[|f-p|<t σ]= 68%

$$P\left\{\frac{|F-p|}{\sigma[p]} \le t_\alpha\right\} = P\left\{\frac{|F-p|}{\sqrt{\frac{p(1-p)}{n}}} \le t_\alpha\right\} = CL$$



t is the quantile of the normal distribution

$$\frac{|f-p|}{\sqrt{\frac{p(1-p)}{n}}} \le |t| \longrightarrow p = \frac{f+\frac{t^2}{2n}}{\frac{t^2}{n}+1} \pm \frac{t\sqrt{\frac{t^2}{4n^2}+\frac{f(1-f)}{n}}}{\frac{t^2}{n}+1}$$

Wilson interval (1934)

$$\xrightarrow{n>>1} p = f \pm t_\alpha\sqrt{\frac{f(1-f)}{n}}$$

Wald (1950)
Standard in Physics

Counting experiments: Poisson case

Wilson interval (1934)

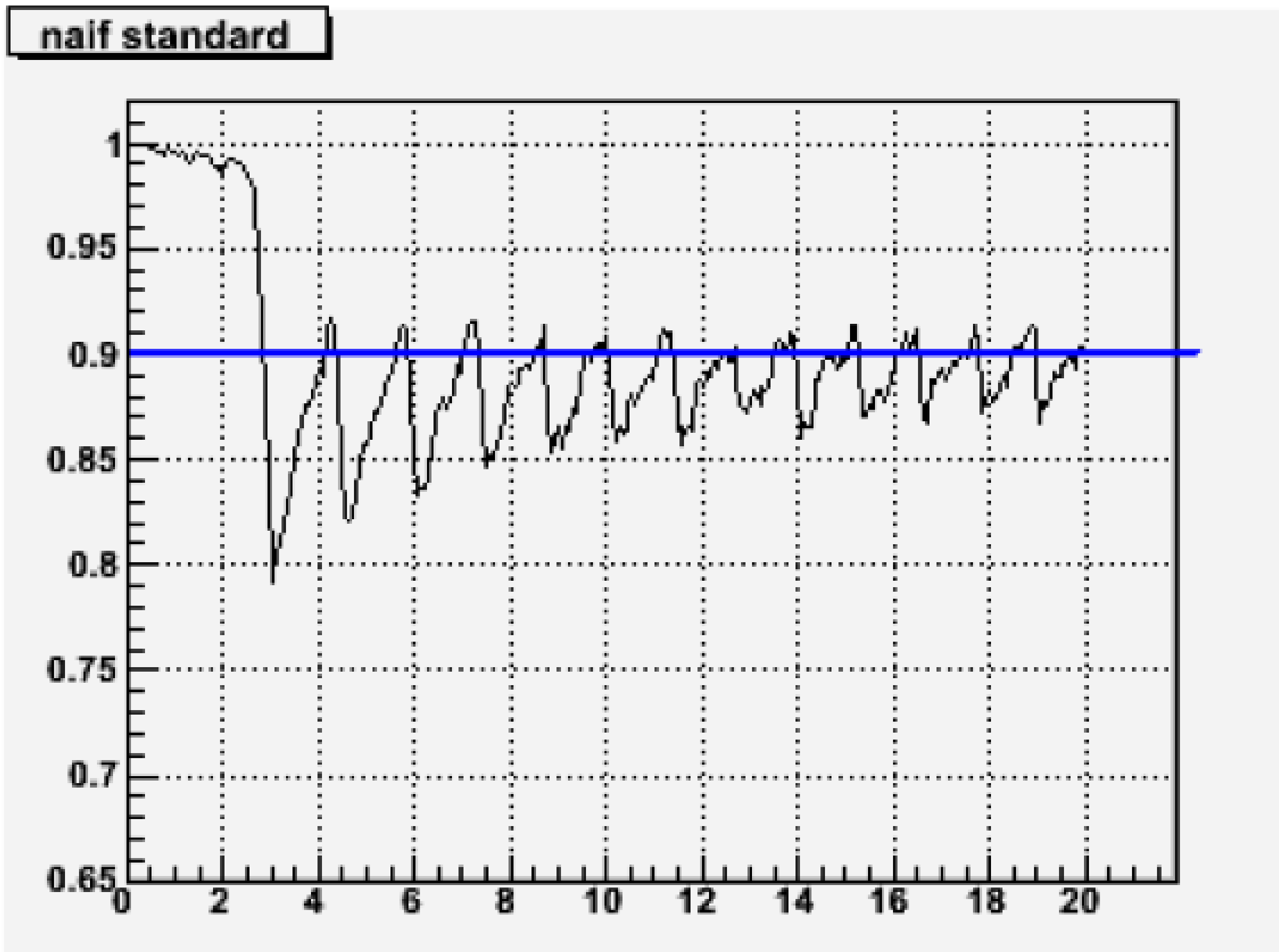
$$P\left\{\frac{|x-\mu|}{\sqrt{\mu}} \le t_\alpha\right\} = CL \rightarrow \quad \mu = x + \frac{t_\alpha^2}{2} \pm t_\alpha \sqrt{x + \frac{t_\alpha^2}{4}}$$

Wald (1950)
Standard in Physics

$$\xrightarrow{\mu \approx x} \mu = x \pm t_\alpha \sqrt{x}$$

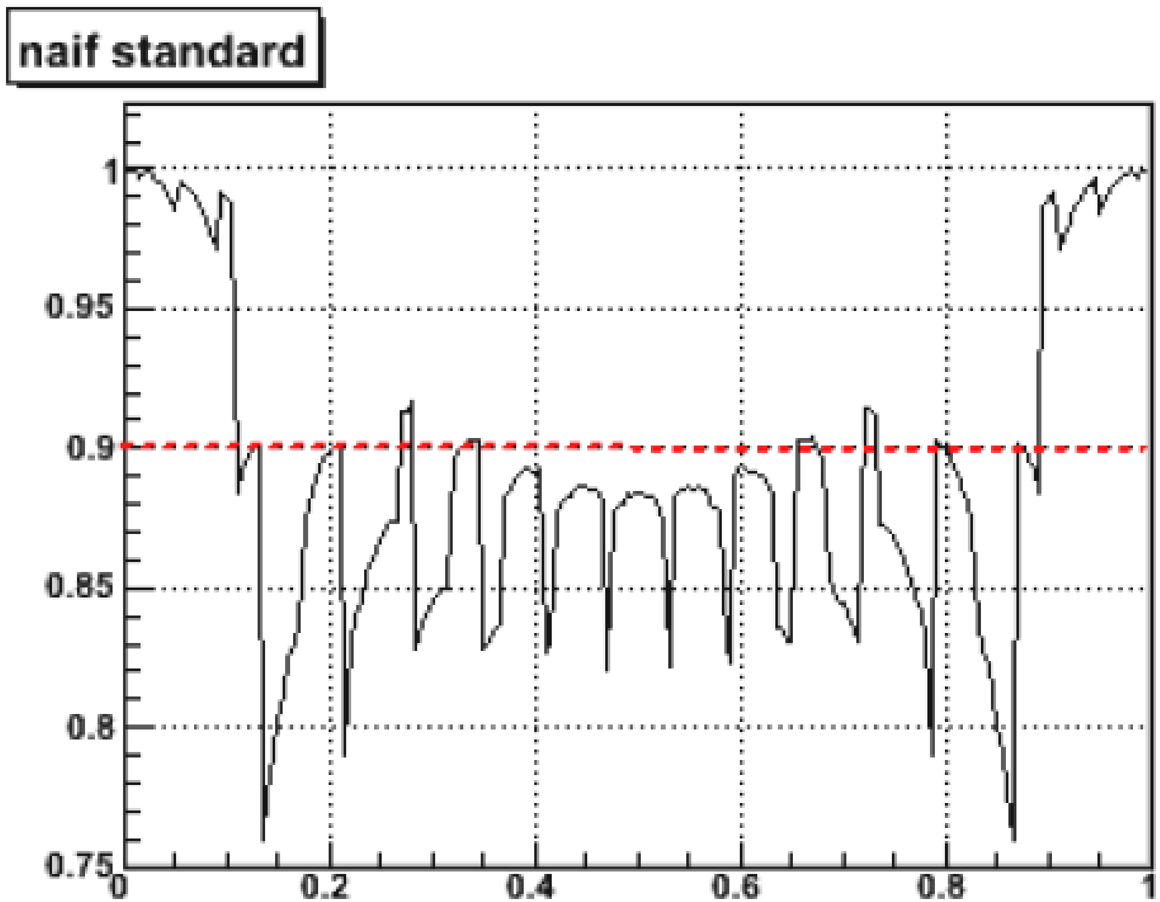
Why to complicate all this?

$$\cancel{\mu = x \pm t_\alpha \sqrt{x}}$$



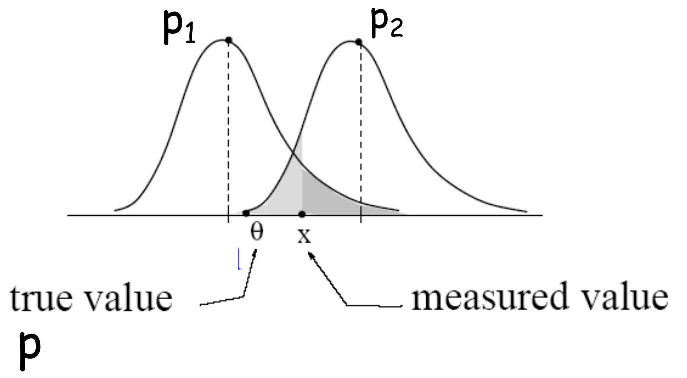
Why to complicate all this?

$$\xrightarrow{n>>1} \xcancel{\varepsilon = f \pm t_\alpha \sqrt{\frac{f(1-f)}{n}}}$$



n=20

small samples



... first difficulties ....
there are no pivot quantities:

$$\sum_{k=x}^{n} \binom{n}{k} p_1^k (1-p_1)^{n-k} = c_1 \ ,$$

$$\sum_{k=0}^{x} \binom{n}{k} p_2^k (1-p_2)^{n-k} = c_2 \ .$$

Symmetric case: $c_1 = c_2 = (1 - CL)/2 = \alpha/2$.
When $x = 0$, $x = n$, $c_1 = c_2 = 1 - CL$:

$$x = n \implies \quad p_1^n = 1 - CL \ ,$$
$$x = 0 \implies \quad (1-p_2)^n = 1 - CL \ .$$

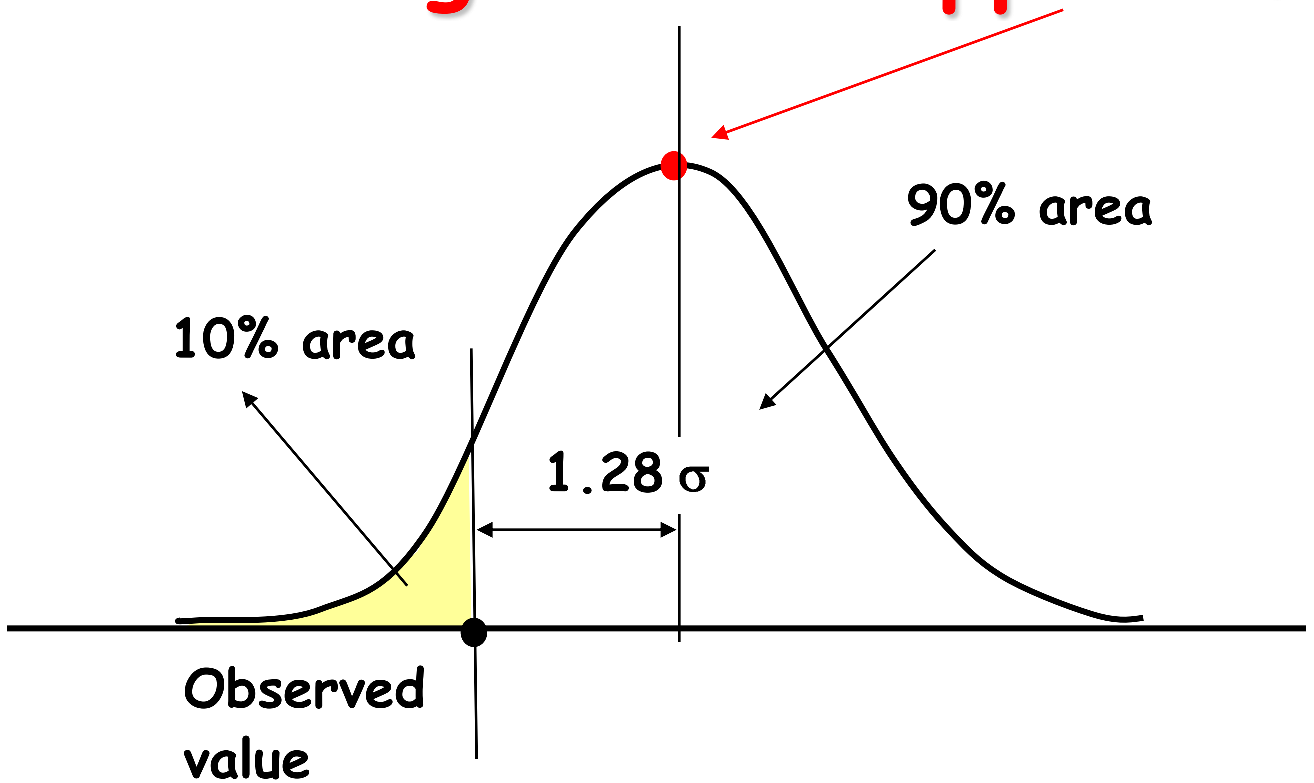
all the attempts had success:

$$p_1 = \sqrt[n]{1 - CL} \qquad p \in [p_1, 1]$$

no success:

$$p_2 = 1 - \sqrt[n]{1 - CL} \qquad p \in [0, p_2]$$

The 90% CL gaussian upper limit



Meaning I: with this upper limit, values less than the observed one are possible with a probability <10%

Meaning II: a larger upper limit should give values less than the observed one in less than 10% of the experiments

Meaning III: the probability to be wrong is 10%

Poisson Limits

$$\sum_{k=x}^{\infty} \frac{\mu_1^k}{k!} \exp(-\mu_1) = c_1 \ , \quad \sum_{k=0}^{x} \frac{\mu_2^k}{k!} \exp(-\mu_2) = c_2 \ ,$$

symmetric case: $c_1 = c_2 = (1 - CL)/2$.

Upper Limits to the mean number of events having obtained x events:

$$\sum_{k=0}^{x} \frac{\mu_2^k}{k!} \exp(-\mu_2) = 1 - CL \ .$$

For $x = 0, 1, 2$, where $\mu_2 \equiv \mu$

$$\begin{aligned} e^{-\mu} &= 1 - CL \ , \\ e^{-\mu} + \mu e^{-\mu} &= 1 - CL \ , \\ e^{-\mu} + \mu e^{-\mu} + \frac{\mu^2}{2} e^{-\mu} &= 1 - CL \end{aligned}$$

x	90%	95%	x	90%	95%
0	2.30	3.00	6	10.53	11.84
1	3.89	4.74	7	11.77	13.15
2	5.32	6.30	8	13.00	14.44
3	6.68	7.75	9	14.21	15.71
4	7.99	9.15	10	15.41	16.96
5	9.27	10.51	11	16.61	18.21

When $\mu > 2.3$, *one con observe no events* but in a number of experiments $< 10\%$.

The Bayes formula

$$P(B_k|A)P(A) = P(A|B_k)P(B_k)$$

if B_k **are disjoint and cover the set S,**

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$$

then $P(B_k|A)$ **can be written as:**

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)}\,, \quad P(A) > 0\ .$$

The trigger problem

$P(T \mid \mu) = 0.95$ prob. for a muon to give a trigger

$P(T \mid \pi) = 0.05$ prob. for a pion to give a trigger

$P(\mu) = 0.10$ prob. to be a muon

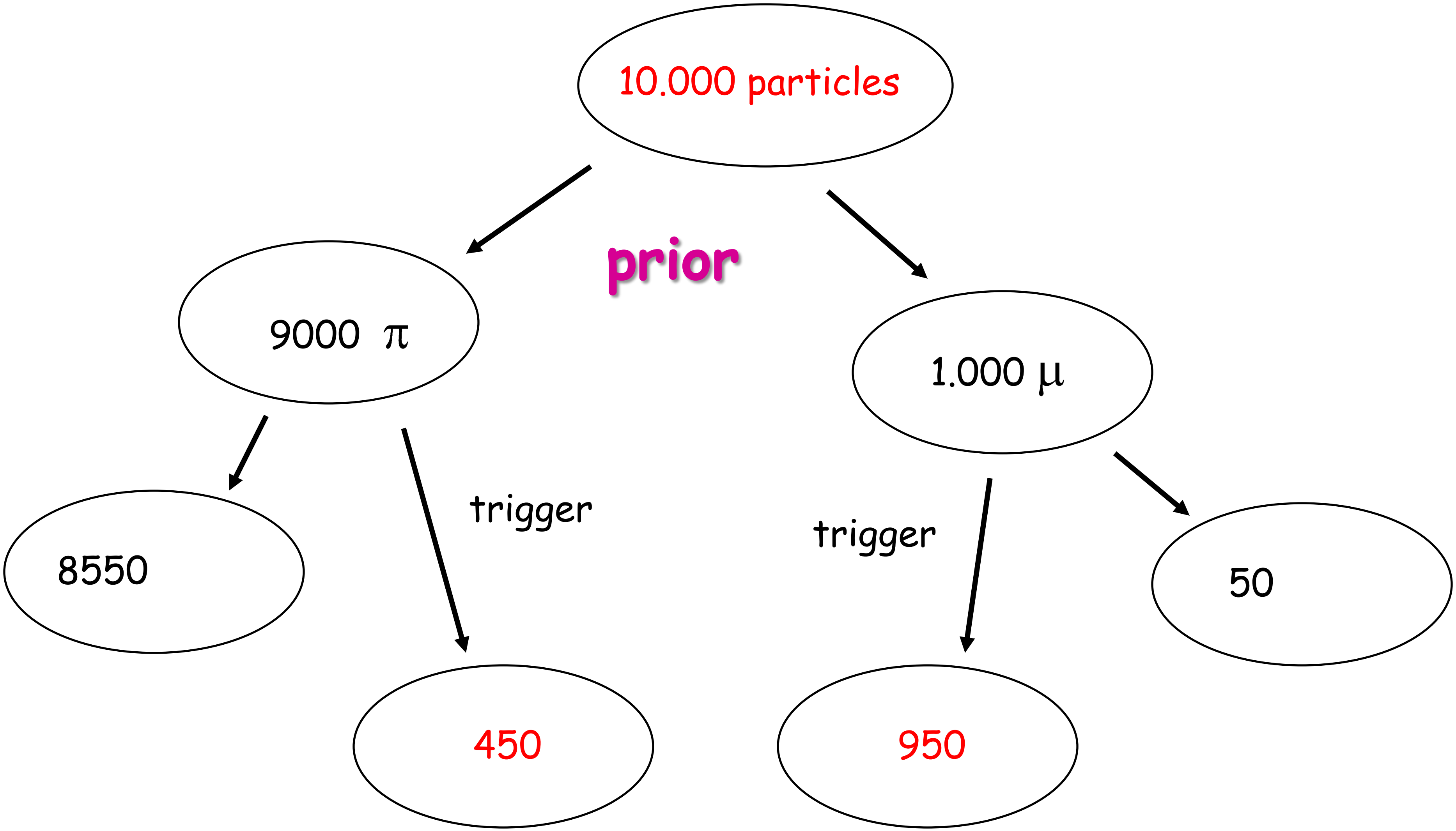
$P(\pi) = 0.90$ prob. to be a pion

$P(\pi \mid T)$ prob. that the trigger selects a pion

$P(\mu \mid T)$ prob. that the trigger selects a muon

The probability to be a muon after the trigger **P(μ|T):**

$$P(\mu \mid T) = \frac{P(T \mid \mu)P(\mu)}{P(T \mid \mu)P(\mu) + P(T \mid \pi)P(\pi)} = \frac{0.95 \times 0.10}{0.95 \times 0.10 + 0.05 \times 0.90} = 0.678$$



enrichment 950/(950+450) = 68%

Efficiency (950+450)/10.000 = 14%

Bayesian use of Bayes formula

The Bayes formula is employed starting from
subjective probabilities

$$P(H_k|\text{data}) = \frac{P(\text{data}|H_k)P(H_k)}{\sum_{i=1}^{n} P(\text{data}|H_i)P(H_i)} \ .$$

An important step,

$$P(H_k|\text{data}) \to P_{n-1}(H_k)$$

iteration:

$$P_n(H_k|E) = \frac{P(E_n|H_k)P_{n-1}(H_k)}{\sum_{i=1}^{n} P(E_n|H_i)P_{n-1}(H_i)} \ ,$$

Bayesian

The gambler problem
Bayesian approach

$$P(\mathrm{Win}|C) = 1 \qquad P(\mathrm{Win}|H) = 0.5$$

Problem: to find the probability that the gambler is cheat, as a function of the number of consecutive wins $\{W_n\}$

$$P(H) \equiv P(H|W_0)\,, \quad P(C) \equiv P(C|W_0) \qquad P(H) = 1 - P(C)$$

Iteration:

$$P(C|W_n) = \frac{P(W_n|C)\,P(C|W_{n-1})}{P(W_n|C)\,P(C|W_{n-1}) + P(W_n|H)\,[1 - P(C|W_{n-1})]}$$

that is

$$P(C|W_n) = \frac{P(C|W_{n-1})}{P(C|W_{n-1}) + 0.5\,[1 - P(C|W_{n-1})]}$$

Bayes:

$P(C)/n$	5	10	15	20
1%	24	91	99.7	99.99
5%	63	98	99.94	99.998
50%	97	99.9	99.997	99.999

The gambler problem
Frequentist approach

Let us suppose 15 cosecutive wins

Hypothesis testing:
The null hypothesis H_0 (honest player) gives a significance level (p-value in this case)

$$0.5^{15} = 3.05\,10^{-5}$$

The probability to be wrong discarding the hypothesis is less then 0.003 %.
The player is cheat.

Cheat probability estimation:
with $n = 15$ and $CL = 90\%$ the probability is

$$p = (0.1)^{1/15} \approx 0.86 \ .$$

With a "cheat probability" $p < 0.86$ it is possible to win for 15/15 times, but in a percentage of plays $< 10\%$

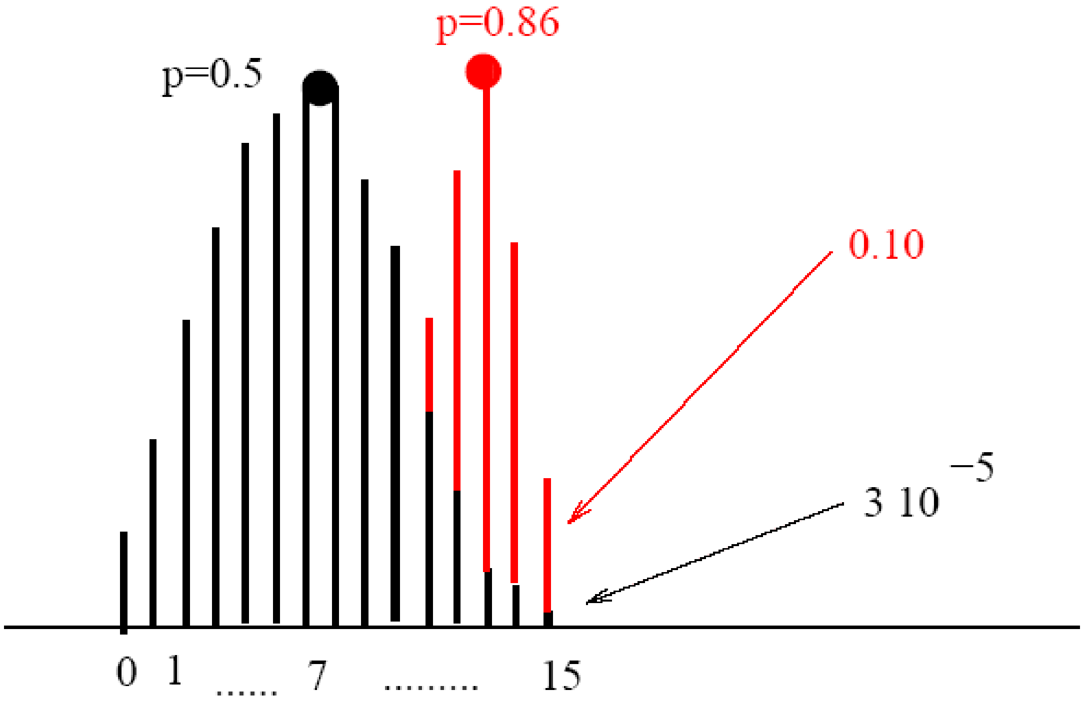
$$0.86 < p < 1 \qquad CL = 90\%$$

The gambler problem
Frequentist approach

Black: hypothesis testing

Red: probability estimation

These conclusions are independent of any a priori hypothesis!



Bayes for the continuum

$$p(x,y) = p_Y(y)\,p(x|y) = p_X(x)\,p(y|x)$$

hence

$$p(x|y) = \frac{p(y|x)\,p_X(x)}{p_Y(y)}$$

that is

$$p(x|y) = \frac{p(y|x)\,p_X(x)}{\int p(y|x)\,p_X(x)\,\mathrm{d}x}$$

Bayesian step:

$$\boxed{p(\mu;x) = \frac{p(x;\mu)\,p_\mu(\mu)}{\int p(x;\mu)\,p_\mu(\mu)\,\mathrm{d}x}}$$

that is

$$p(\mu;x) = \frac{\text{likelihood} \times \text{prior}}{\text{normalization}}$$

The prior

$$p_\mu(\mu)$$

that is the subjective probability assigned to μ, is **NEVER** used by frequentists

Bayesian Interval estimate

Degree of belief on μ for a measured x:

$$p(\mu;x) = \frac{L(x,\mu)\,p_\mu(\mu)}{\int L(x,\mu)\,p_\mu(\mu)\,\mathrm{d}p}$$

Estimate:

$$\mu \in [\mu_1, \mu_2]$$

Bayesian credible interval

with degree of belief

$$\int_{\mu_1}^{\mu_2} p(\mu;x)\,\mathrm{d}\mu = \text{degree of belief}$$

- **one integrates over μ considered as a random variable**
- **this coincides with the frequentist result if the prior $p_\mu(\mu)$ is uniform and the property**

  $$1 - F(\mu;x) = F(x;\mu)$$

  holds
- **but the interpretation is different!**

Bayesian coin tossing

$$p(p; n, x) = \frac{p^x (1-p)^{n-x}\, p_p(p)}{\int p^x (1-p)^{n-x}\, p_p(p)\, \mathrm{d}p}$$

With uniform prior,

$$p_p(p) = \text{const} \quad 0 < p < 1$$

Recalling the β function:

$$\int_0^1 p^x (1-p)^{n-x}\, \mathrm{d}p = \frac{x!(n-x)!}{(n+1)!}$$

one obtains the degree of belief of p

$$p(p; n, x) = \frac{(n+1)!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$$\langle p \rangle = \frac{x+1}{n+2}$$

$$\mathrm{Var}[p] = \frac{(x+1)(n-x+1)}{(n+3)(n+2)^2}$$

Elementary example

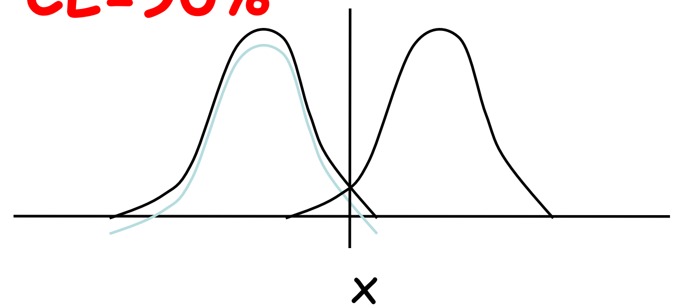
20 events have been generated and 5 passed the cut
What is the estimation of the efficiency with CL=90%?

Frequentist result:

x=5, n=20, CL=90%

$$\sum_{k=x}^{n} \binom{n}{k} \varepsilon_1^k (1-\varepsilon_1)^{n-k} = 0.05$$

$$\sum_{k=0}^{x} \binom{n}{k} \varepsilon_2^k (1-\varepsilon_2)^{n-k} = 0.05$$

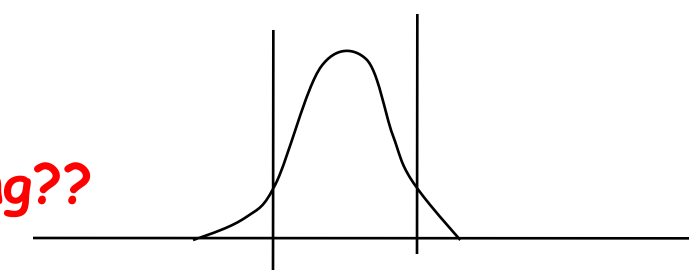


$\varepsilon_1, \varepsilon_2$ $\Rightarrow$ ε =[0.104, 0.455]

Bayesian result:

$$\frac{\int_{p_1}^{p_2} \varepsilon^x (1-\varepsilon)^{n-x} \mathbf{d}\varepsilon}{\int_{0}^{1} \varepsilon^x (1-\varepsilon)^{n-x} \mathbf{d}\varepsilon} = 0.90$$

What meaning??



$\varepsilon_1, \varepsilon_2$ $\Rightarrow$ ε =[0.122, 0.423]

Efficiency calculation: an OPEN PROBLEM!!

$$\frac{f+\frac{t^2}{2n}}{\frac{t^2}{n}+1} \pm \frac{t_\alpha \sqrt{\frac{t^2}{4n^2}+\frac{f(1-f)}{n}}}{\frac{t^2}{n}+1}$$

Wilson interval (1934)

$$\xrightarrow{n>>1} \varepsilon = f \pm t_\alpha \sqrt{\frac{f(1-f)}{n}}$$

Wald (1950)
Standard in Physics

$$\sum_{k=x}^{n} \binom{n}{k} \varepsilon_1^k (1-\varepsilon_1)^{n-k} = \alpha/2$$

$$\sum_{k=0}^{x} \binom{n}{k} \varepsilon_2^k (1-\varepsilon_2)^{n-k} = \alpha/2$$

Exact frequentist
Clopper Pearson (1934) (PDG)

$$\frac{\int_{p_1}^{p_2} \varepsilon^x (1-\varepsilon)^{n-x} \mathbf{d}\varepsilon}{\int_0^1 \varepsilon^x (1-\varepsilon)^{n-x} \mathbf{d}\varepsilon} = CL\,?$$

Bayes. This is not frequentist
but can be tested
in a frequentist way

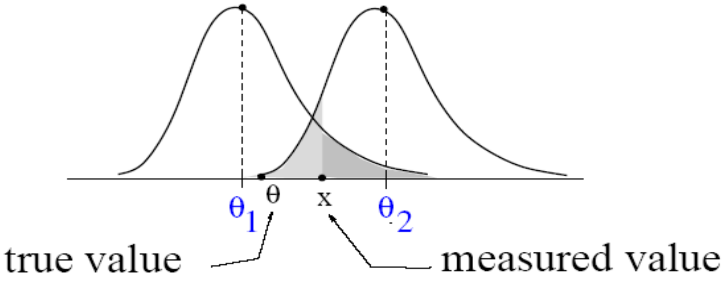
Coverage simulation

$$x = \text{gRandom} \rightarrow \text{Binomial}(p,N) \qquad \rightarrow \textcolor{red}{x}$$

$$1\text{-}CL = \alpha$$

$$\sum_{k=x}^{n} \binom{n}{k} p_1^k (1-p_1)^{n-k} = \alpha/2$$

$$\sum_{k=0}^{x} \binom{n}{k} p_2^k (1-p_2)^{n-k} = \alpha/2$$



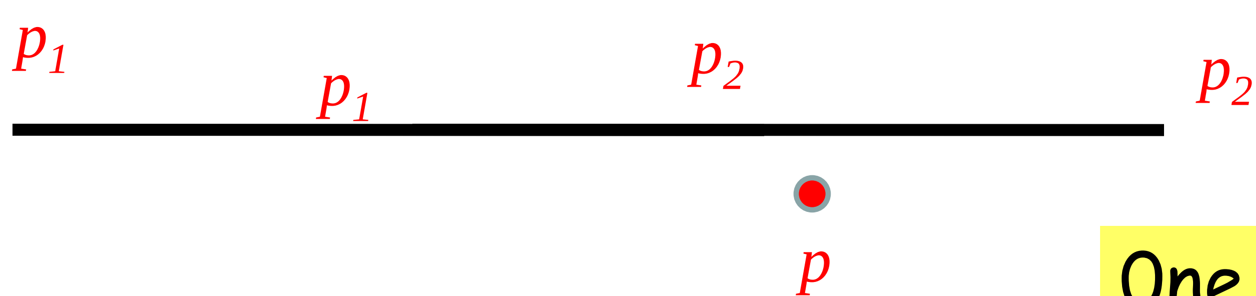
$\int_x^\infty p(x;\theta_1)\,\mathrm{d}x = c_1 \qquad \int_{-\infty}^x p(x;\theta_2)\,\mathrm{d}x = c_2$

where

$$\theta \in [\theta_1, \theta_2]\ , \quad 1-(c_1+c_2) = CL$$

MC techniques can be used: grid over θ to find the values θ_1 and θ_2 satisfying these integrals

Tmath:: BinomialI(p,N,x)



$\varepsilon = k/n$

One expects $\varepsilon \sim CL$

Statistical Science
2001, Vol. 16, No. 2, 101–133

Interval Estimation for a Binomial Proportion

Lawrence D. Brown, T. Tony Cai and Anirban DasGupta

Simulate many x with a true p and check when the intervals contain the true value p. Compare this frequency with the stated CL

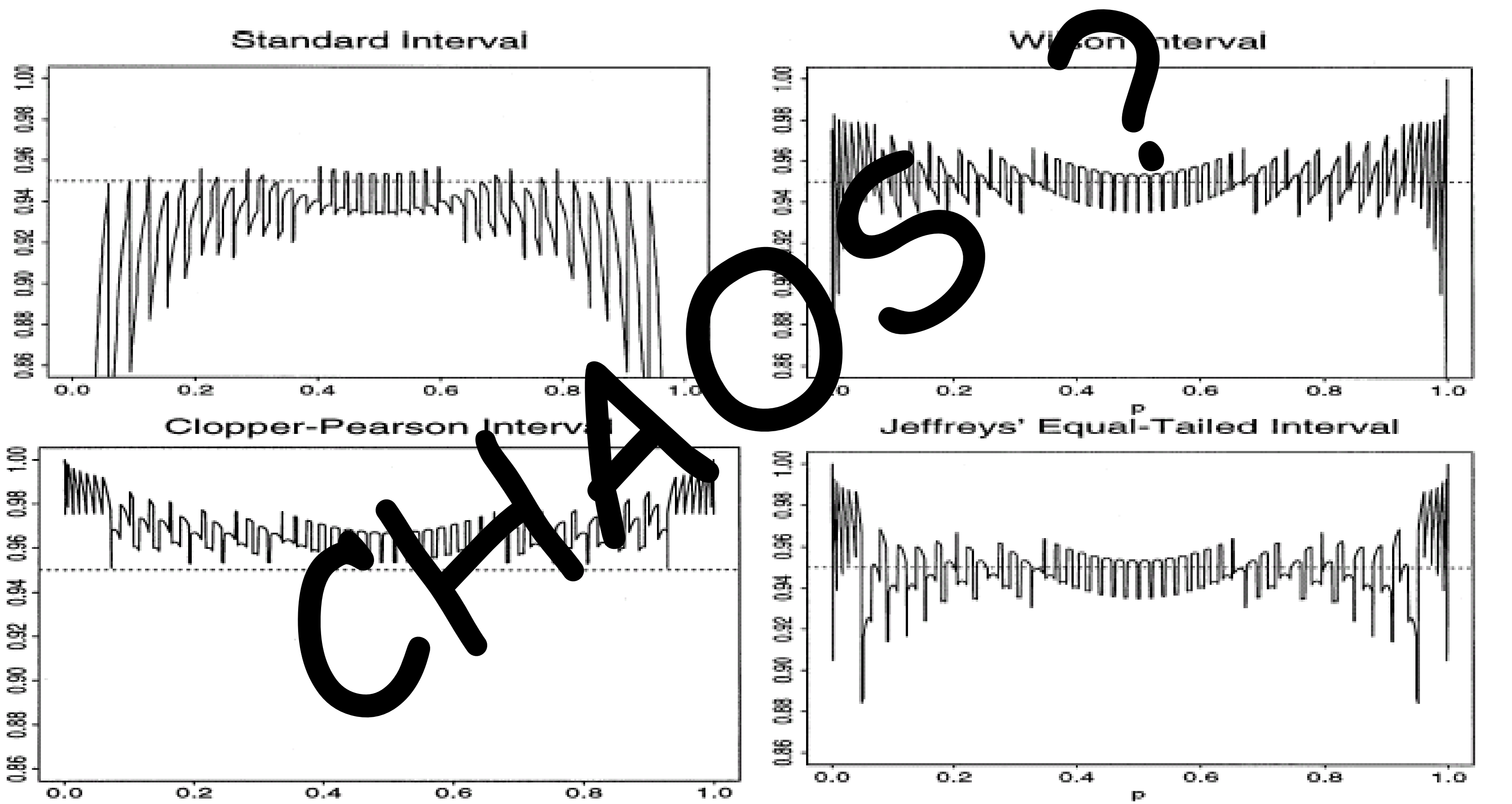
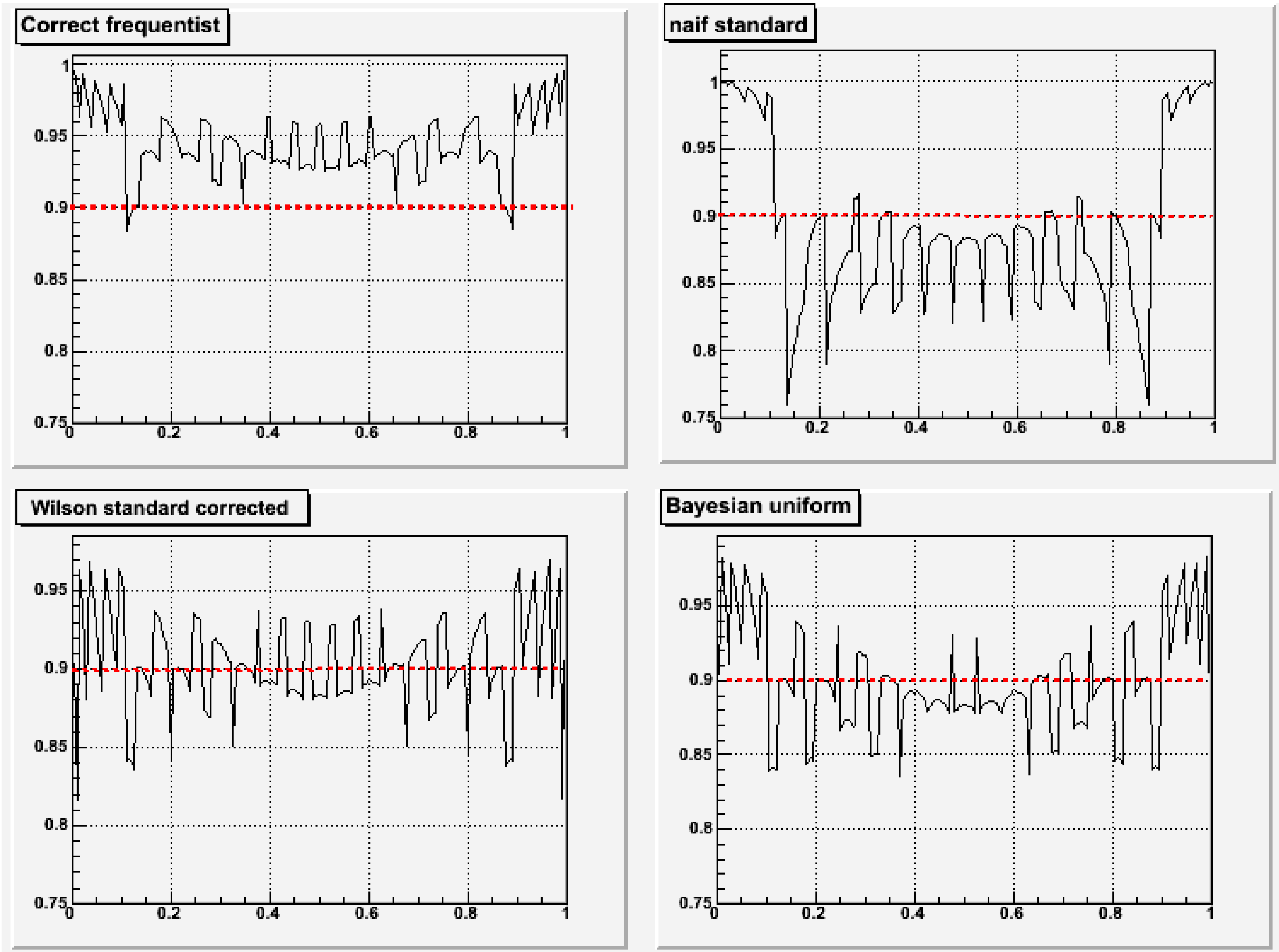


FIG. 5. *Coverage probability for $n = 50$.*

CL=0.95, n=50

Simulate many x with a true p and check when the intervals contain the true value p. Compare this frequency with the stated CL



CL=0.90, n=20

In the estimation of the efficiency (probability) the coverage is "chaotic"

The new standard (not yet for physicists) is to use the exact frequentist or the formula

$$\varepsilon = \frac{f + \dfrac{t_{\alpha/2}^{\,2}}{2n}}{\dfrac{t_{\alpha/2}^{\,2}}{n} + 1} \pm \frac{t_{\alpha/2}\sqrt{\dfrac{t_{\alpha/2}^{\,2}}{4n^2} + \dfrac{f(1-f)}{n}}}{\dfrac{t_{\alpha/2}^{\,2}}{n} + 1}, \qquad \begin{array}{l} x = n,\ [p_1, 1],\ p_1 = (1-CL)^{1/n} \\ x = 0,\ [0, p_2],\ p_2 = 1-(1-CL)^{1/n} \end{array}$$

$f = x/n$, $t_{\alpha/2}$ **gaussian,** $1\text{-}CL = \alpha$, $t = 1$ **is** 1σ

The standard formula

$$\cancel{\varepsilon = f \pm t_\alpha \sqrt{\frac{f(1-f)}{n}}}$$

BYE-BYE

should be abandoned

A further improvement:

The continuity correction is equivalent to The Clopper-Pearson formula

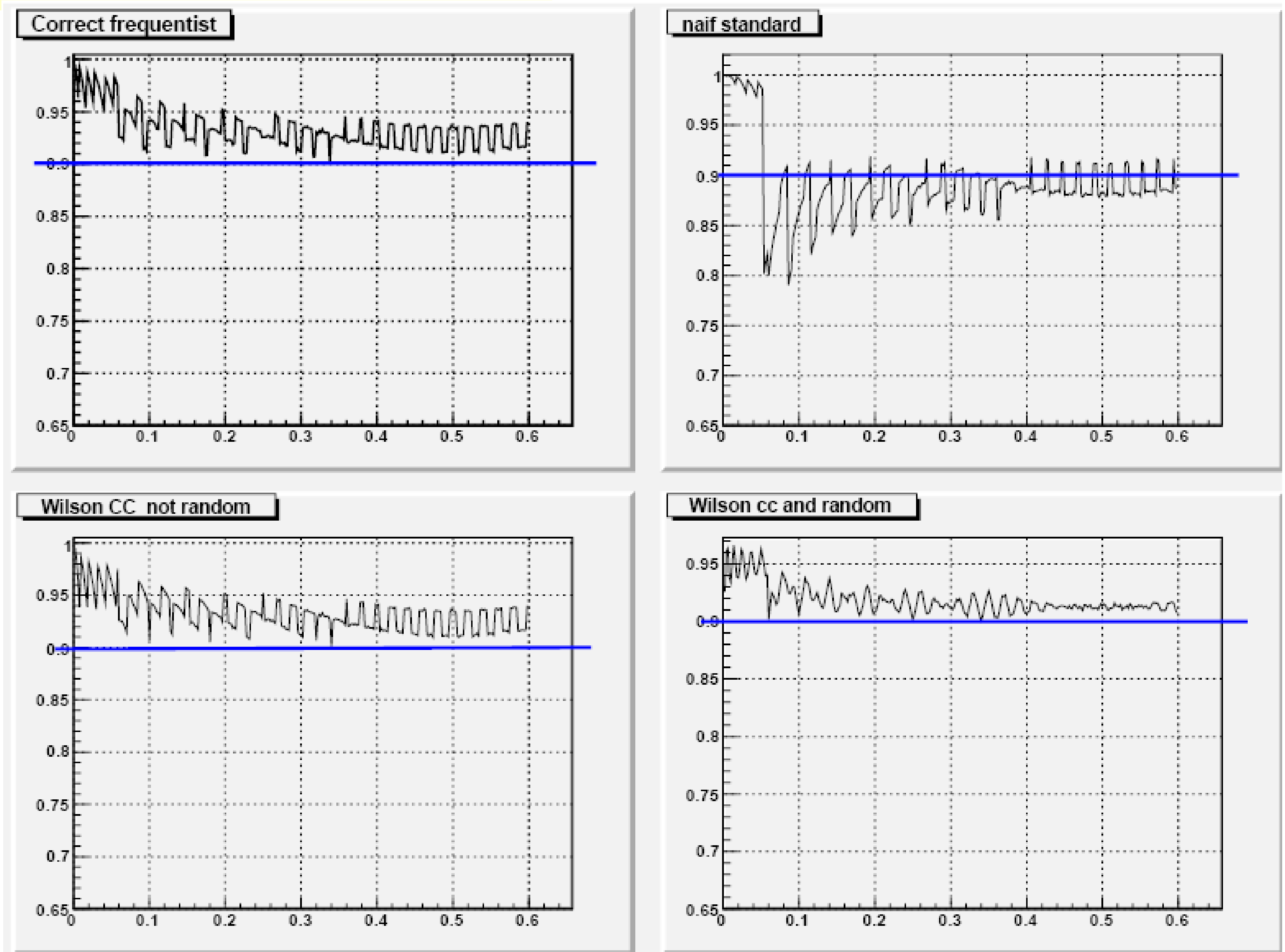
$$\varepsilon = \frac{f_{\pm} + \dfrac{t_{\alpha/2}^{\;2}}{2n}}{\dfrac{t_{\alpha/2}^{\;2}}{n} + 1} \pm \frac{t_{\alpha/2}\sqrt{\dfrac{t_{\alpha/2}^{\;2}}{4n^2} + \dfrac{f_{\pm}(1 - f_{\pm})}{n}}}{\dfrac{t_{\alpha/2}^{\;2}}{n} + 1}, \qquad \begin{array}{l} x = n,\ [p_1, 1],\ p_1 = (1 - CL)^{1/n} \\ x = 0,\ [0, p_2],\ p_2 = 1 - (1 - CL)^{1/n} \end{array}$$

$$f_{+} = (x + 0.5)/n,\ \ f_{-} = (x - 0.5)/n,$$

$$t_{\alpha/2}\ \text{gaussian}, 1\text{-}CL = \alpha,\ \ t = 1 \text{ is } 1\sigma$$

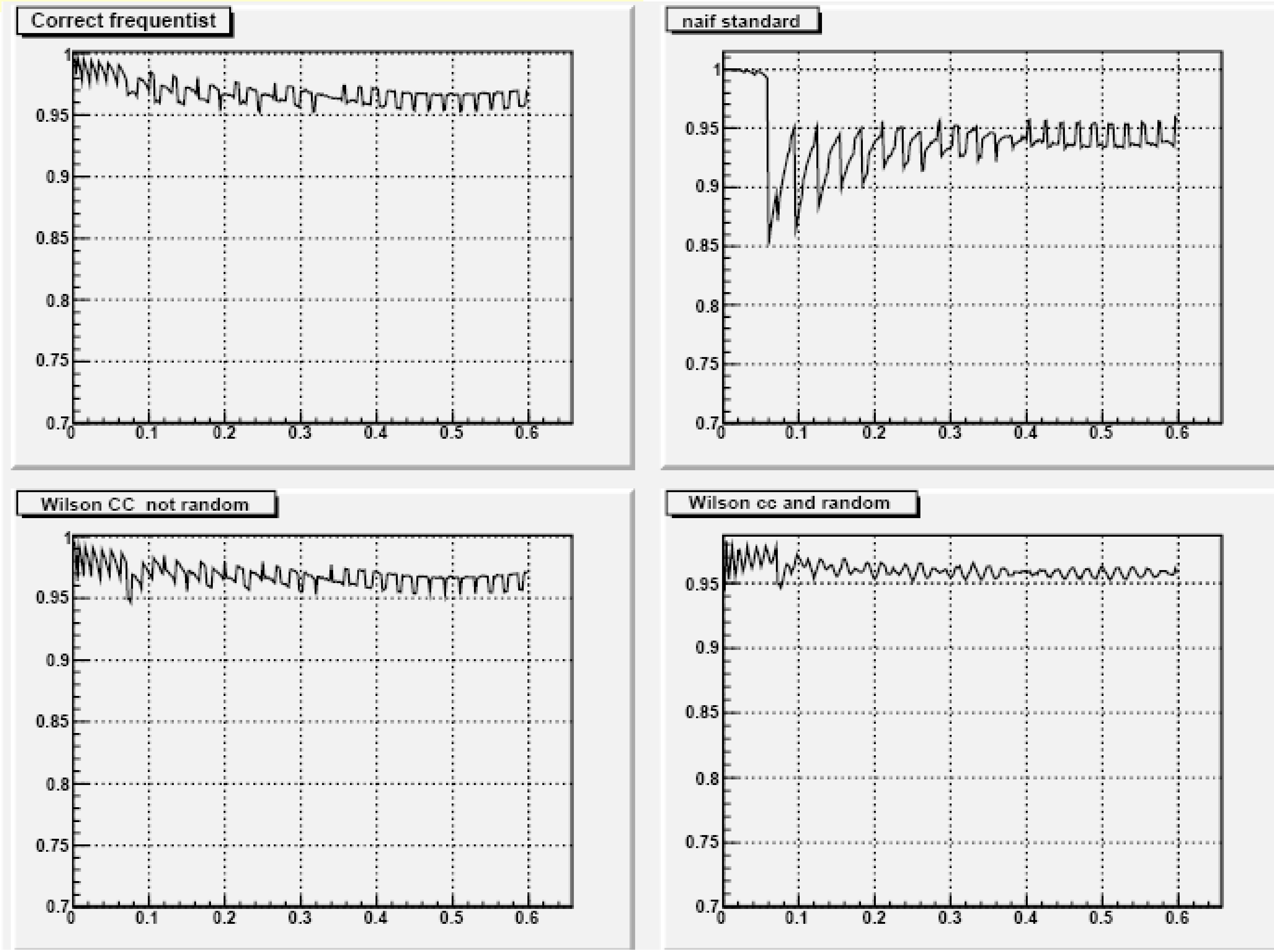
This should become the standard formula also for physicists

N=50 CL=0.90



Correct frequentist
naif standard
Wilson CC not random
Wilson cc and random

N=50 CL=0.95



Correct frequentist
naif standard
Wilson CC not random
Wilson cc and random
1
0.95
0.9
0.85
0.8
0.75
0.7
0
0.1
0.2
0.3
0.4
0.5
0.6

The likelihood ratio method

$$\lambda(p,x) = \frac{L(p,x)}{L(p_{best},x)}$$ Maximize

$$-2\ln\lambda(p,x) = 2\ln\frac{L(p_{best},x)}{L(p,x)}$$ Minimize

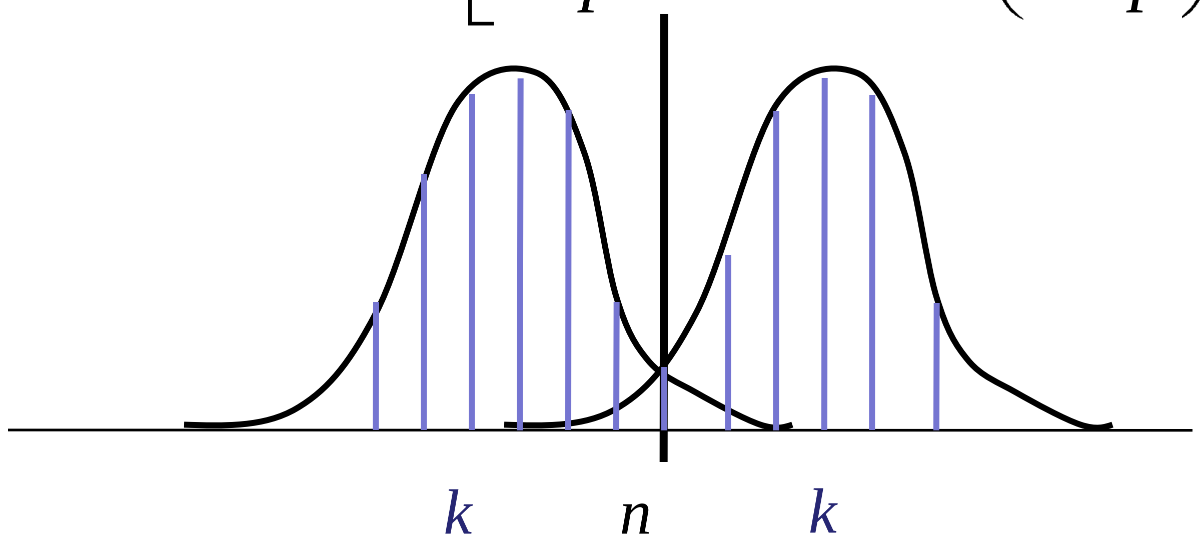
Binomial Coverage simulation max likelihood constraint

Feldman & Cousins, Phys. Rev. D 57(1998)3873
UNIFIED method

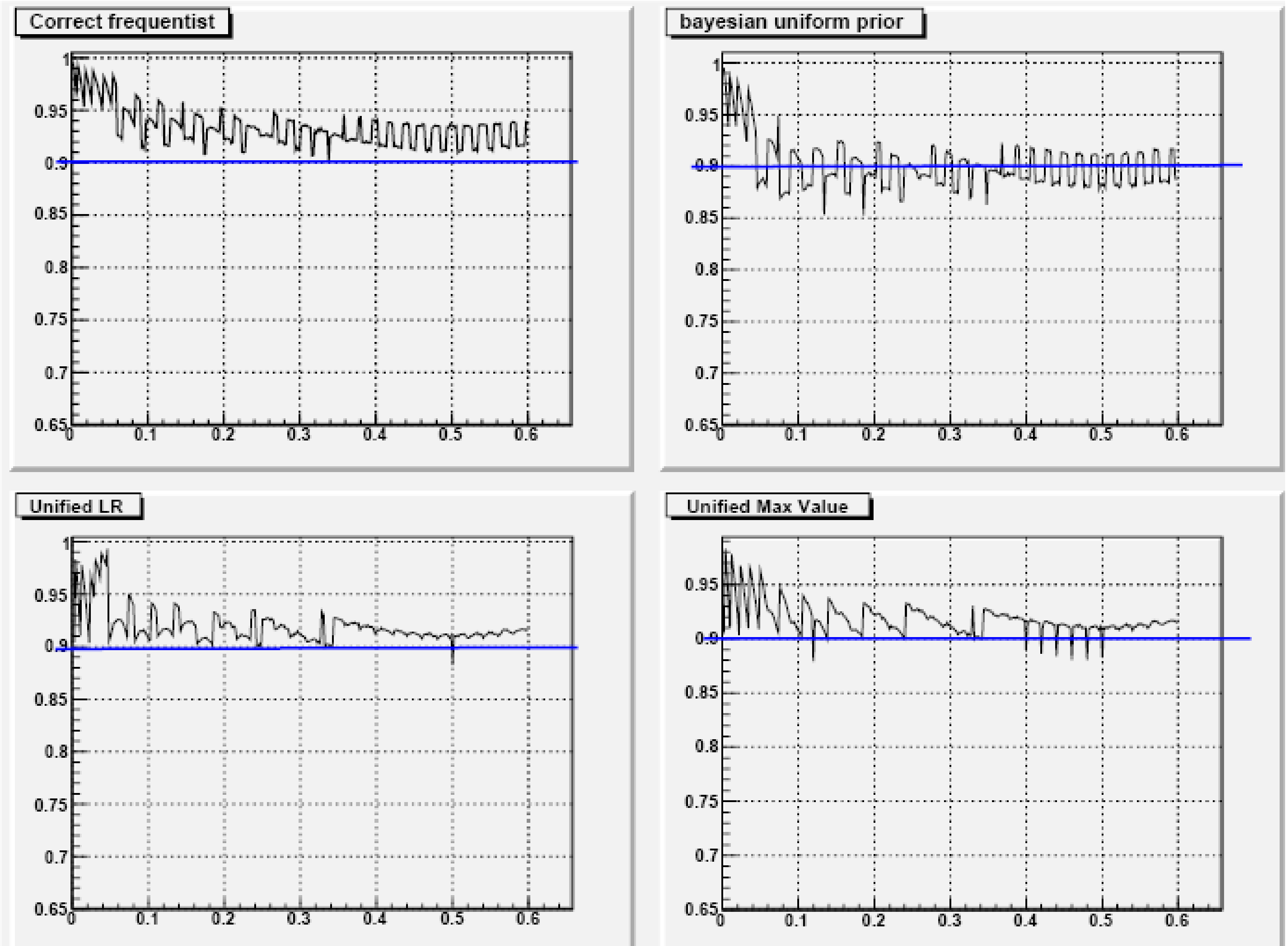
$$\sum_{k\in A}\frac{N!}{k!(N-k)!}\,p^{k}\,(1-p)^{N-k} < CL\,,$$

$$A = \{k \mid k \geq 0, \text{and} -2\ln\lambda(p,N,k) < -2\ln\lambda(p,N,x)\}$$

$$-2\ln\lambda(p,N,x) = 2\left[\ln\frac{f}{p} + (N-x)\,\ln\left(\frac{1-f}{1-p}\right)\right]$$

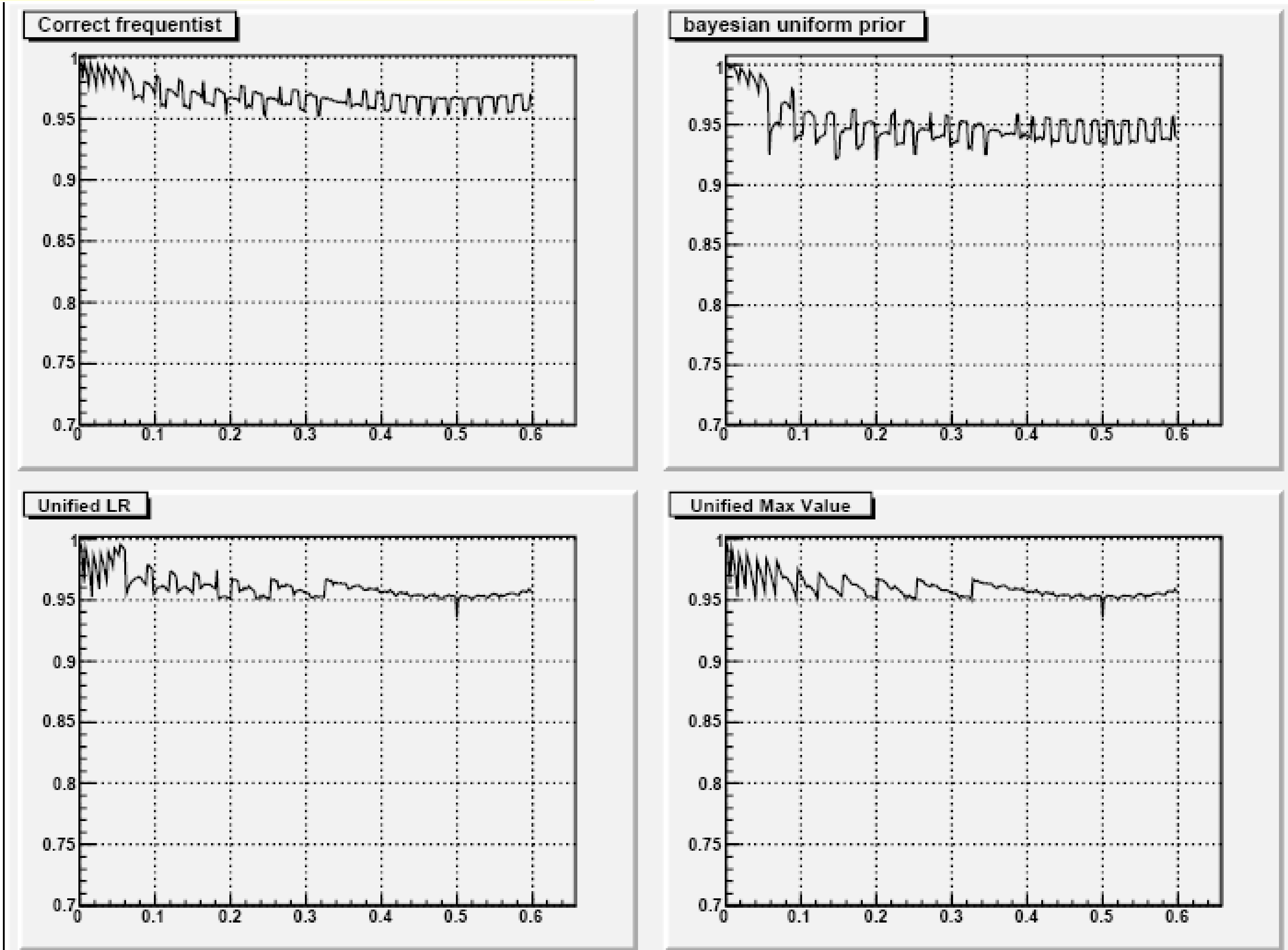


N=50 CL=0.90

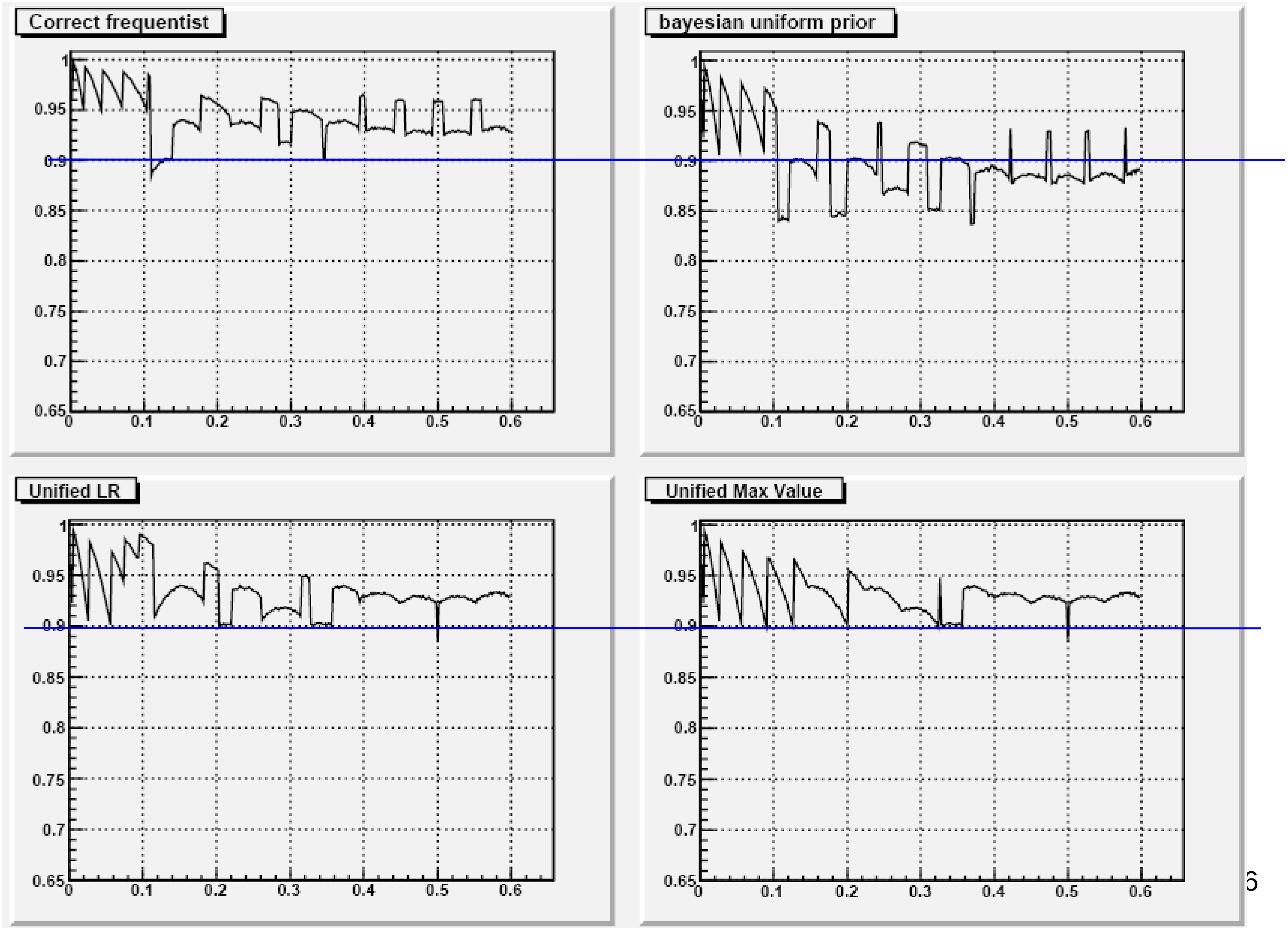


Correct frequentist
bayesian uniform prior
Unified LR
Unified Max Value
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0
0.1
0.2
0.3
0.4
0.5
0.6

N=50 CL=0.95


Correct frequentist
bayesian uniform prior
Unified LR
Unified Max Value
1
0.95
0.9
0.85
0.8
0.75
0.7
0
0.1
0.2
0.3
0.4
0.5
0.6

N=20 CL=090



Correct frequentist
bayesian uniform prior
Unified LR
Unified Max Value
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0
0.1
0.2
0.3
0.4
0.5
0.6

The problem persists also with large samples!

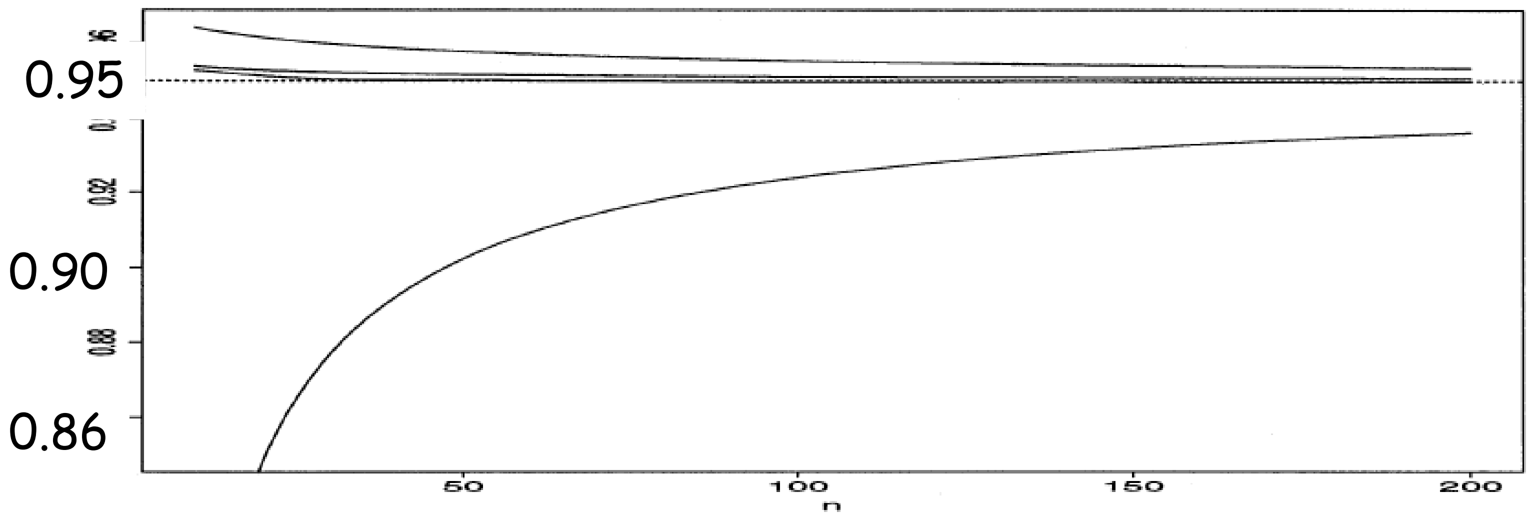
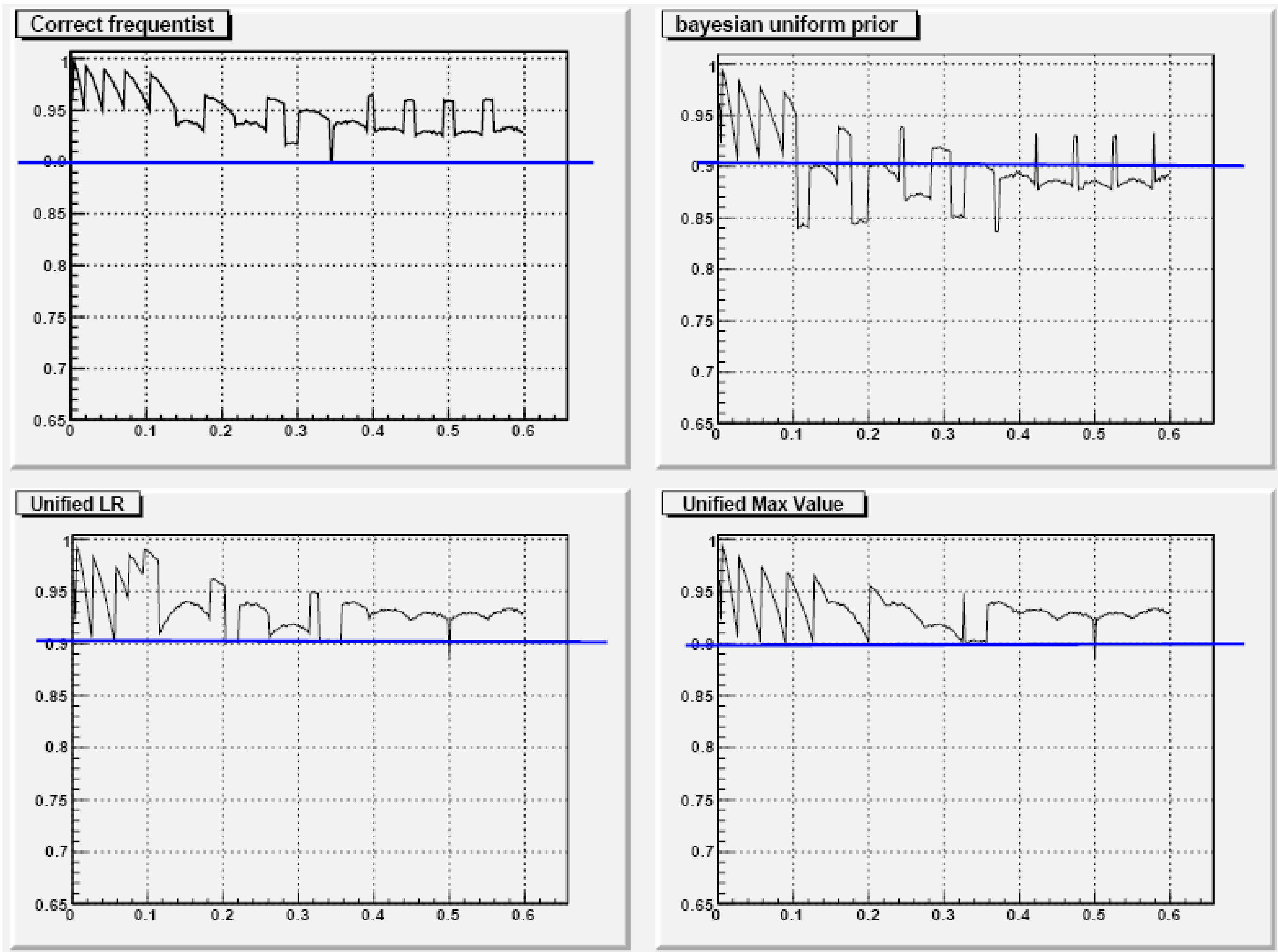


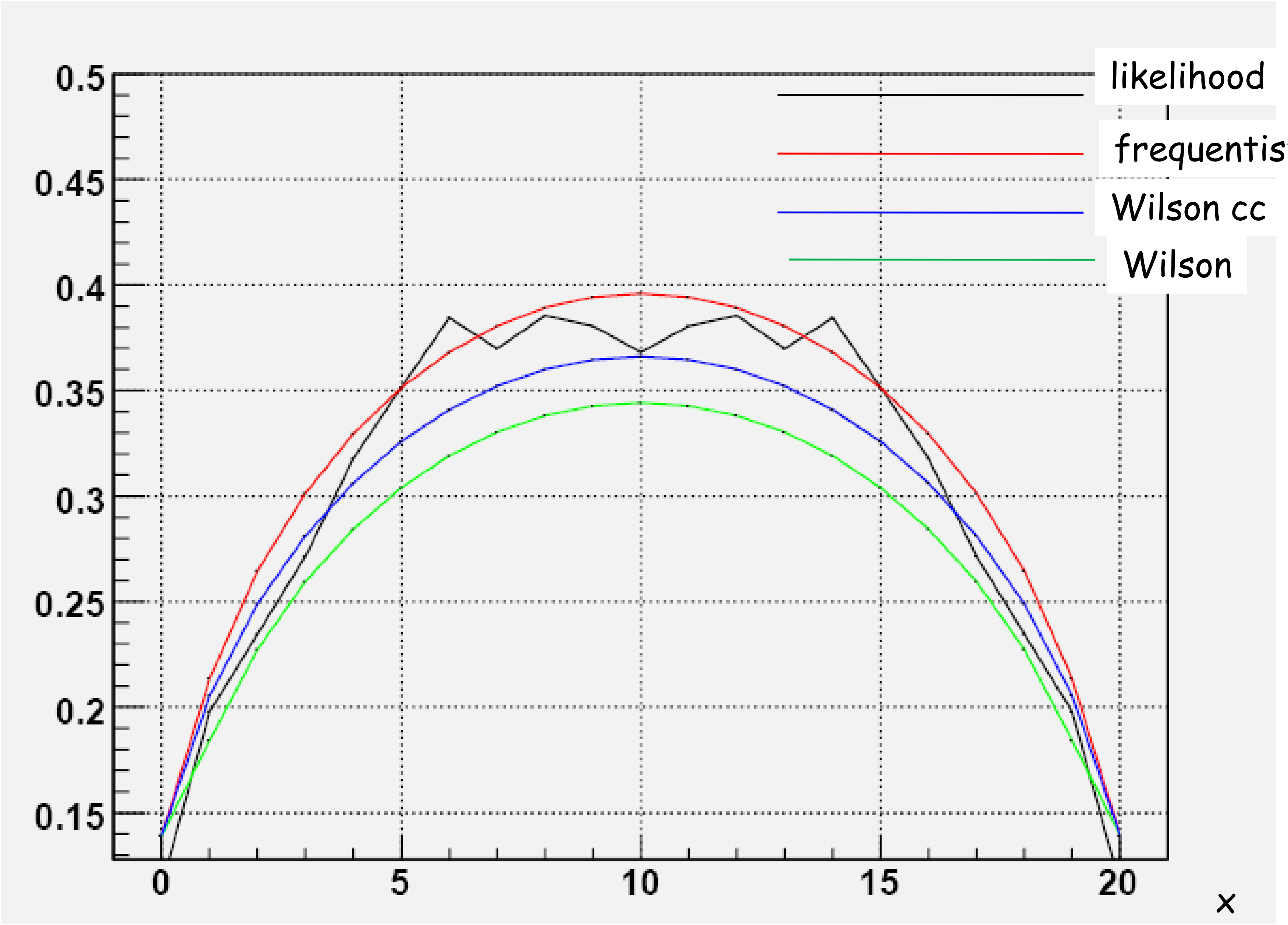
FIG. 6. *Comparison of the average coverage probabilities. From top to bottom: the Agresti–Coull interval CI_{AC}, the Wilson interval CI_W, the Jeffreys prior interval CI_J and the standard interval CI_s. The nominal confidence level is 0.95.*

N=20 CL=0.90

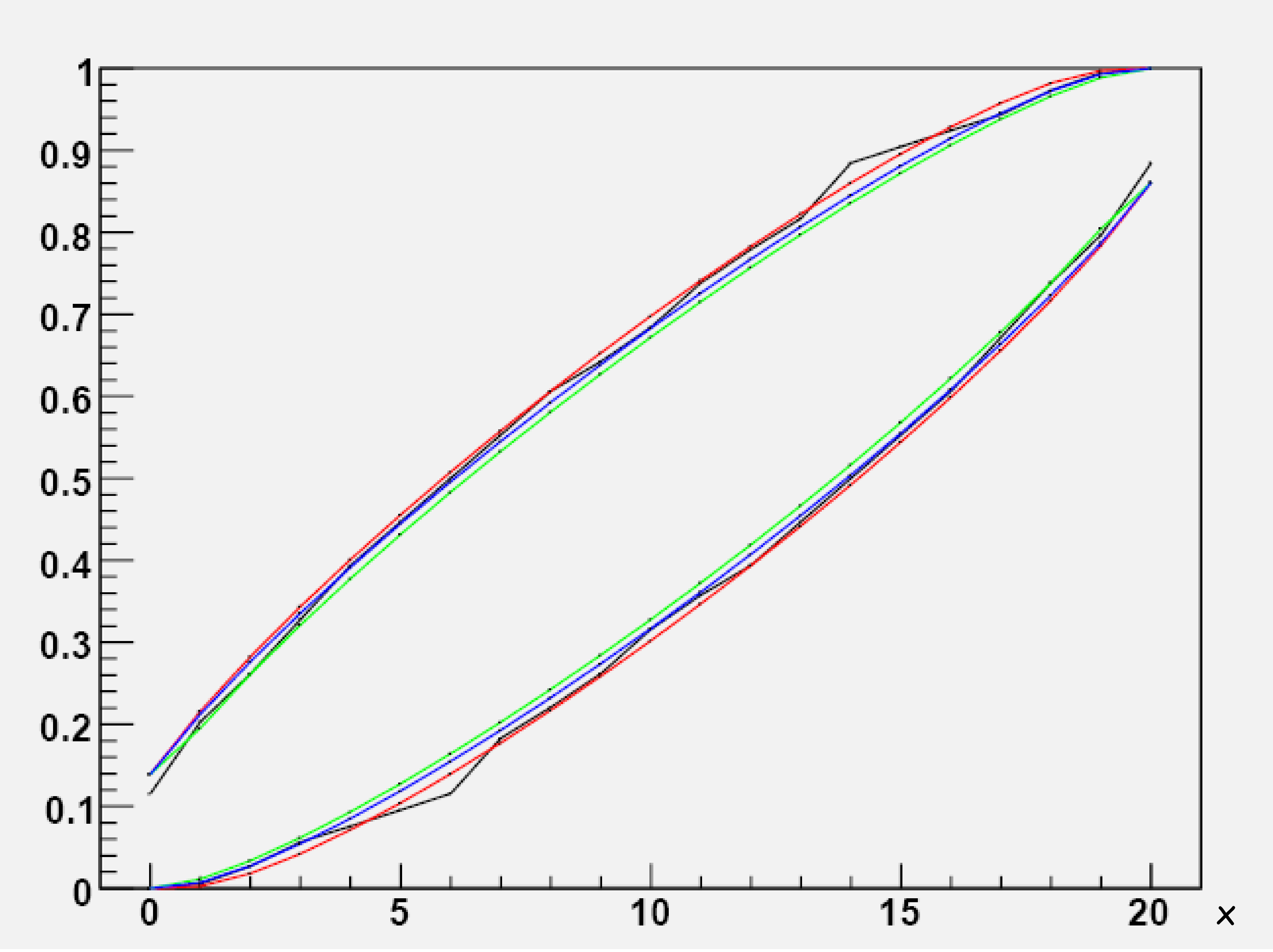


Correct frequentist
bayesian uniform prior
Unified LR
Unified Max Value
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0
0.1
0.2
0.3
0.4
0.5
0.6

N=20 CL=0.90   Interval amplitude



N=20 CL=0.90 Interval limits



Comment

George Casella

(2001)

$$\xrightarrow{n>>1} \cancel{\varepsilon = f \pm t_\alpha \sqrt{\frac{f(1-f)}{n}}}$$

1. INTRODUCTION

Professors Brown, Cai and DasGupta (BCD) are to be congratulated for their clear and imaginative look at a seemingly timeless problem. The chaotic behavior of coverage probabilities of discrete confidence sets has always been an annoyance, resulting in intervals whose coverage probability can be vastly different from their nominal confidence level. What we now see is that for the Wald interval, an approximate interval, the chaotic behavior is relentless, as this interval will not maintain $1-\alpha$ coverage for any value of n. Although fixes relying on ad hoc rules abound, they do not solve this fundamental defect of the Wald interval and, surprisingly, the usual safety net of asymptotics is also shown not to exist. So, as the song goes, “Bye-bye, so long, farewell” to the Wald interval.

Now that the Wald interval is out, what is in? There are probably two answers here, depending on whether one is in the classroom or the consulting room.

George Casella is Arun Varma Commemorative Term Professor and Chair, Department of Statistics, University of Florida, Gainesville, Florida 32611-8545 (e-mail: casella@stat.ufl.edu).

$$\frac{f_\pm + \dfrac{t^2}{2n}}{\dfrac{t^2}{n}+1} \pm \frac{t_\alpha \sqrt{\dfrac{t^2}{4n^2} + \dfrac{f_\pm(1-f_\pm)}{n}}}{\dfrac{t^2}{n}+1}$$

$$\sum_{k=0}^{x} \binom{n}{k} \varepsilon_2^k (1-\varepsilon_2)^{n-k} = \alpha/2$$

$$\sum_{k=0}^{x} \binom{n}{k} \varepsilon_2^k (1-\varepsilon_2)^{n-k} = \alpha/2$$

Counting experiments: Poisson case

$$\frac{(x-\mu)}{\sqrt{\mu}} = t_\alpha \rightarrow \mu = x + \frac{t_\alpha^2}{2} \pm t_\alpha \sqrt{x + \frac{t_\alpha^2}{4}}$$

Wilson interval (1934)

$$\xrightarrow{\mu \approx x} \mu = x \pm t_\alpha \sqrt{x}$$

Wald (1950)
Standard in Physics

$$\sum_{k=0}^{x} \frac{\mu_2^{\,k}}{k!} e^{-\mu_2} = \alpha / 2$$

$$\sum_{k=x}^{\infty} \frac{\mu_1^{\,k}}{k!} e^{-\mu_1} = \alpha / 2$$

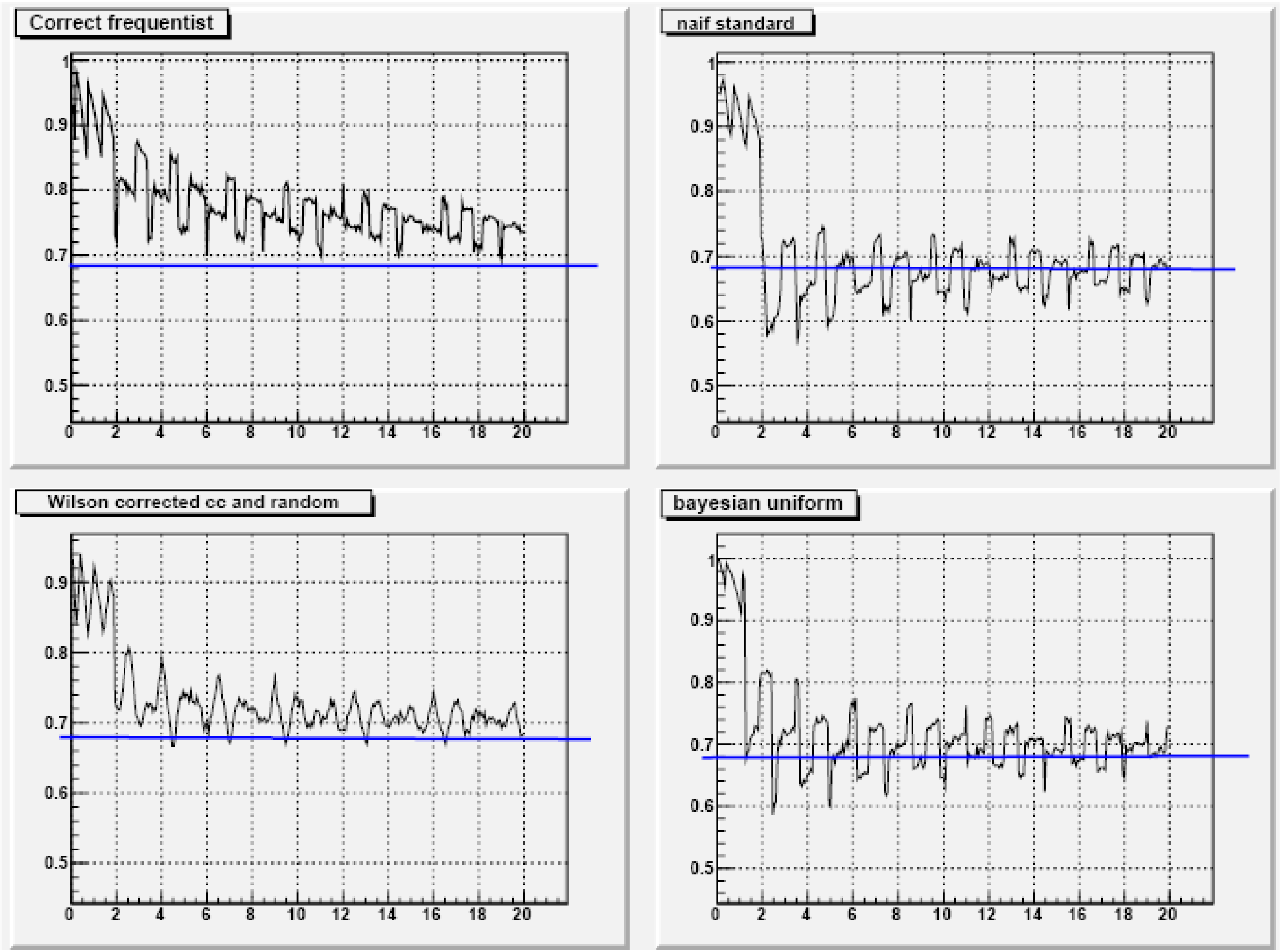
Exact frequentist
Clopper Pearson (1934) (PDG)

$$\frac{\int_{\mu_1}^{\mu_2} \frac{\mu^x e^{-\mu}}{x!} \mathrm{d}\mu}{\int_{0}^{\infty} \frac{\mu^x e^{-\mu}}{x!} \mathrm{d}\mu} = CL\,?$$

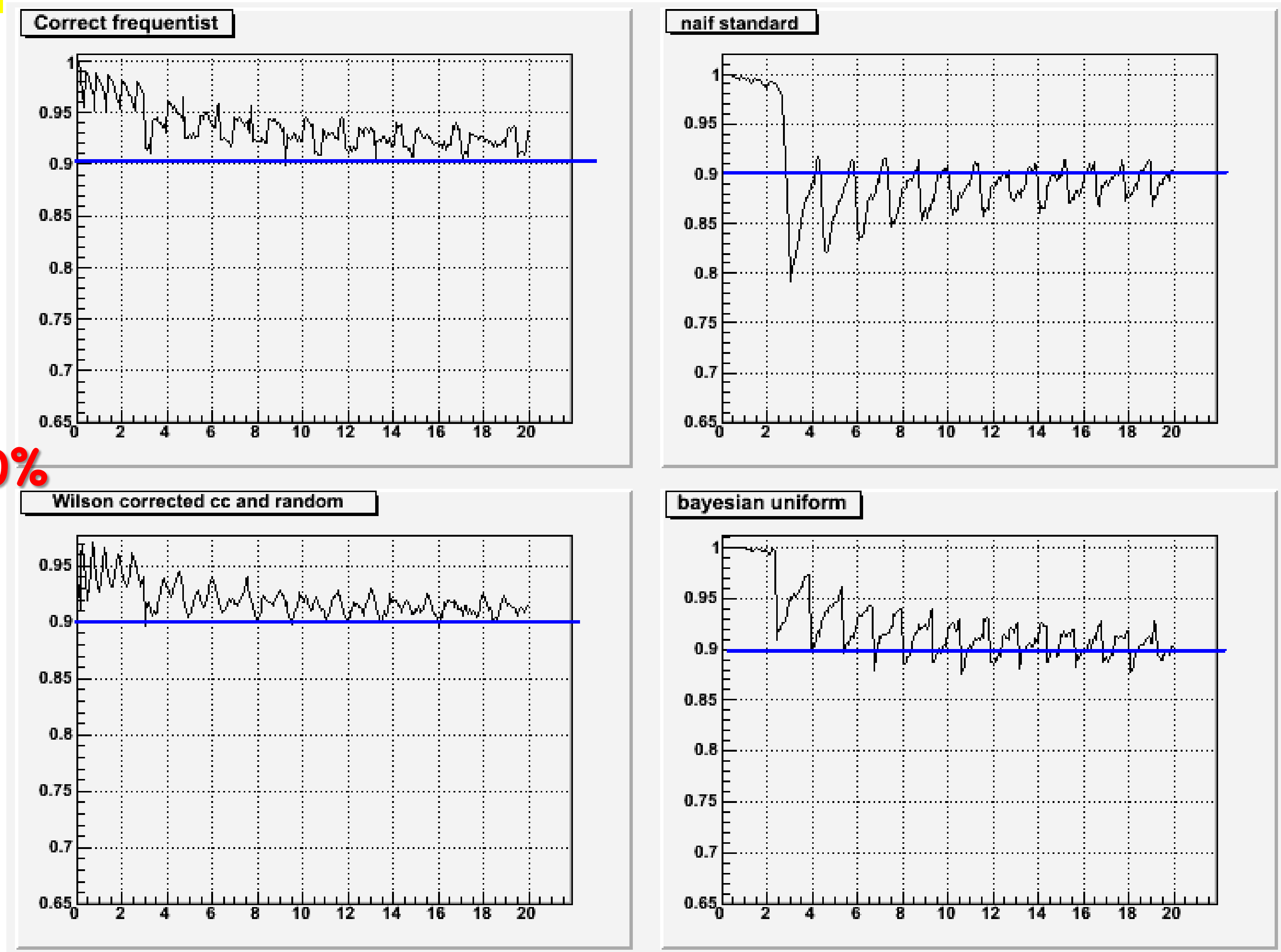
Bayes. This is not frequentist but can be tested in a frequentist way

Poissonian Coverage simulation

CL=68%

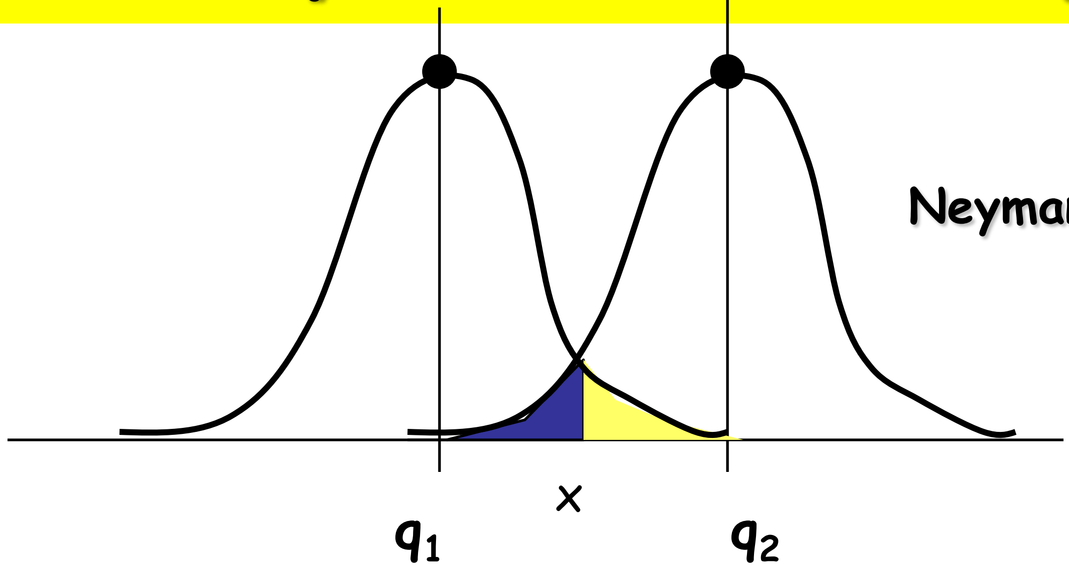


Poissonian Coverage simulation



CL=90%

The Neyman-FC integrals



Neyman integrals

$$\sum_{k \in A} p(x;\theta) < CL\,,$$

$$A = \{ k \mid k \geq 0, \text{and} - 2\ln\lambda(\theta,k) < -2\ln\lambda(\theta,x) \}$$

$$-2\ln\lambda(\theta,x) = 2\ln\frac{L(\theta_{best},x)}{L(\theta,x)}$$

minimize

Poissonian Coverage simulation max likelihood constraint

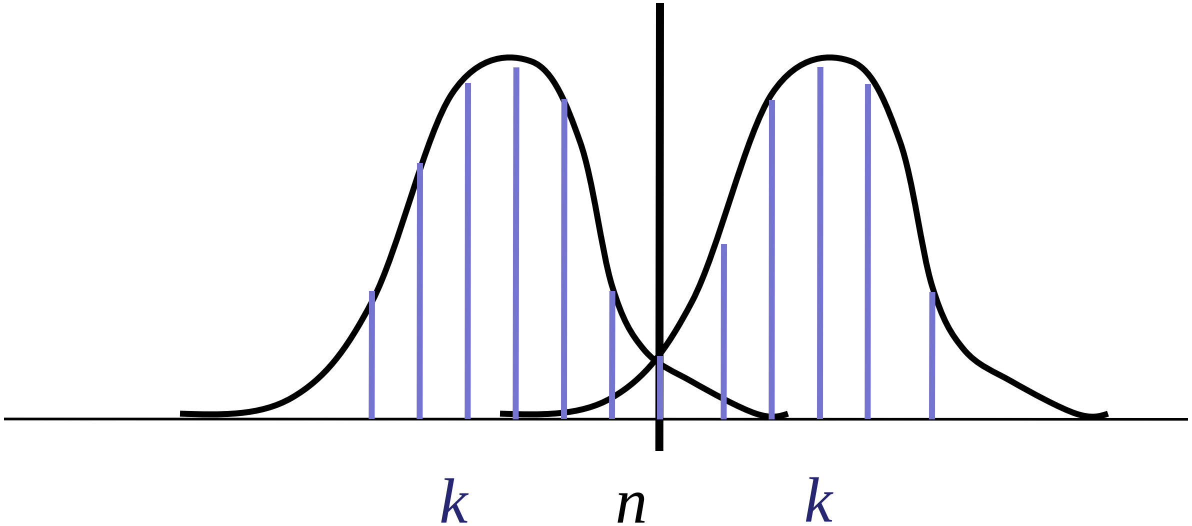
Feldman & Cousins, Phys. Rev. D 57(1998)3873

$$\sum_{k\in A}\frac{\mu^k}{k!}e^{-\mu} < CL$$

$$\mathcal{A}(\mu,n)=\{\,k\mid\ k\in\mathbf{Z},\,k\geq 0,\text{ and }-2\ln\lambda(\mu,k)<-2\ln\lambda(\mu,n)\ \}$$

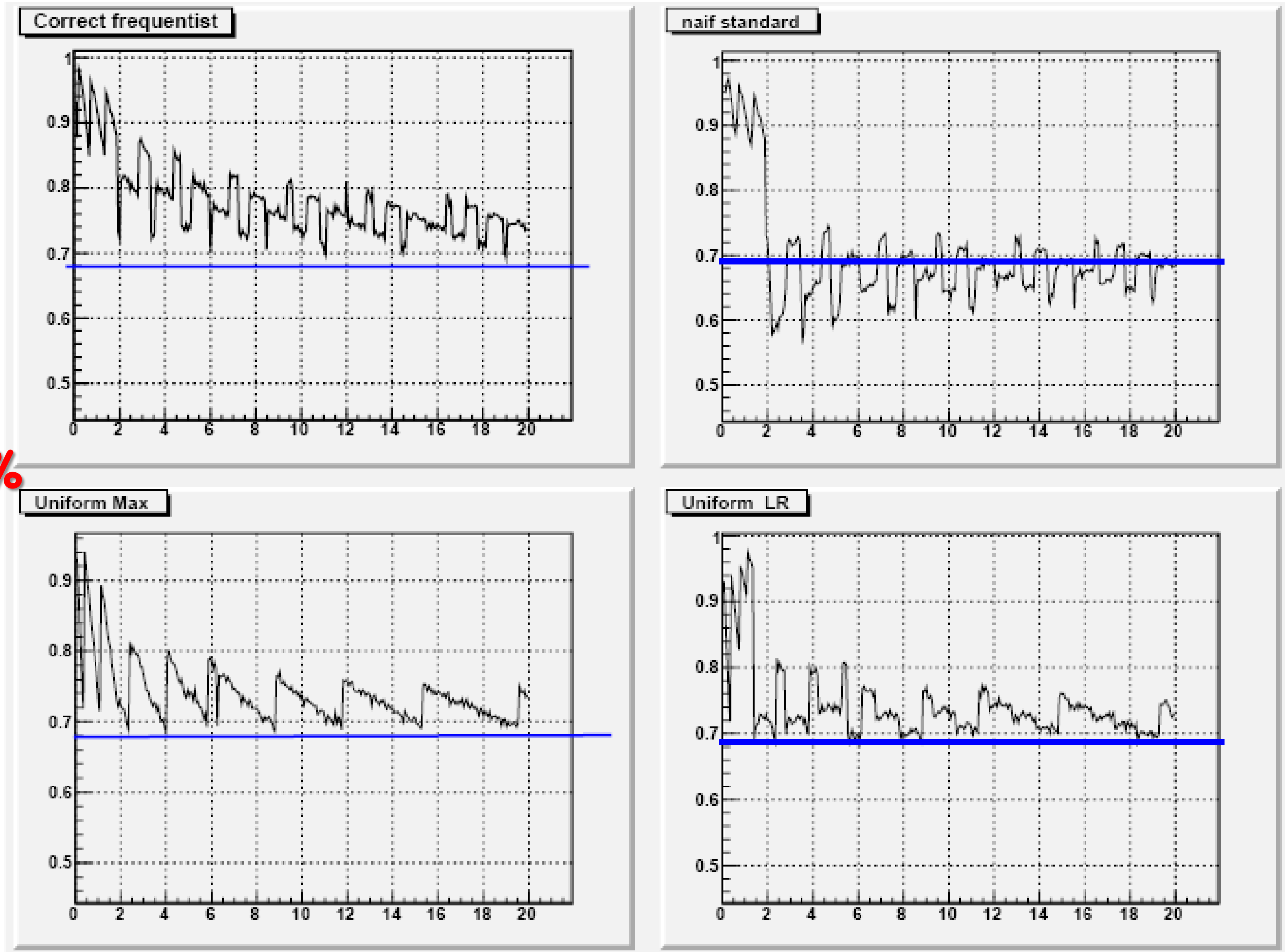
(crudely) describe $\mathcal{A}(\mu,n)$ as the set of all integers that give a "better fit" to μ than n does, where "better fit" is defined in terms of the likelihood ratio. Note that $n\notin\mathcal{A}(\mu,n)$.

$$-2\ln\lambda(\mu,n)=-2\ln\frac{\mu^n e^{-\mu}/n!}{n^n e^{-n}/n!}=(\mu-n)+n\ln\left(\frac{n}{\mu}\right)$$

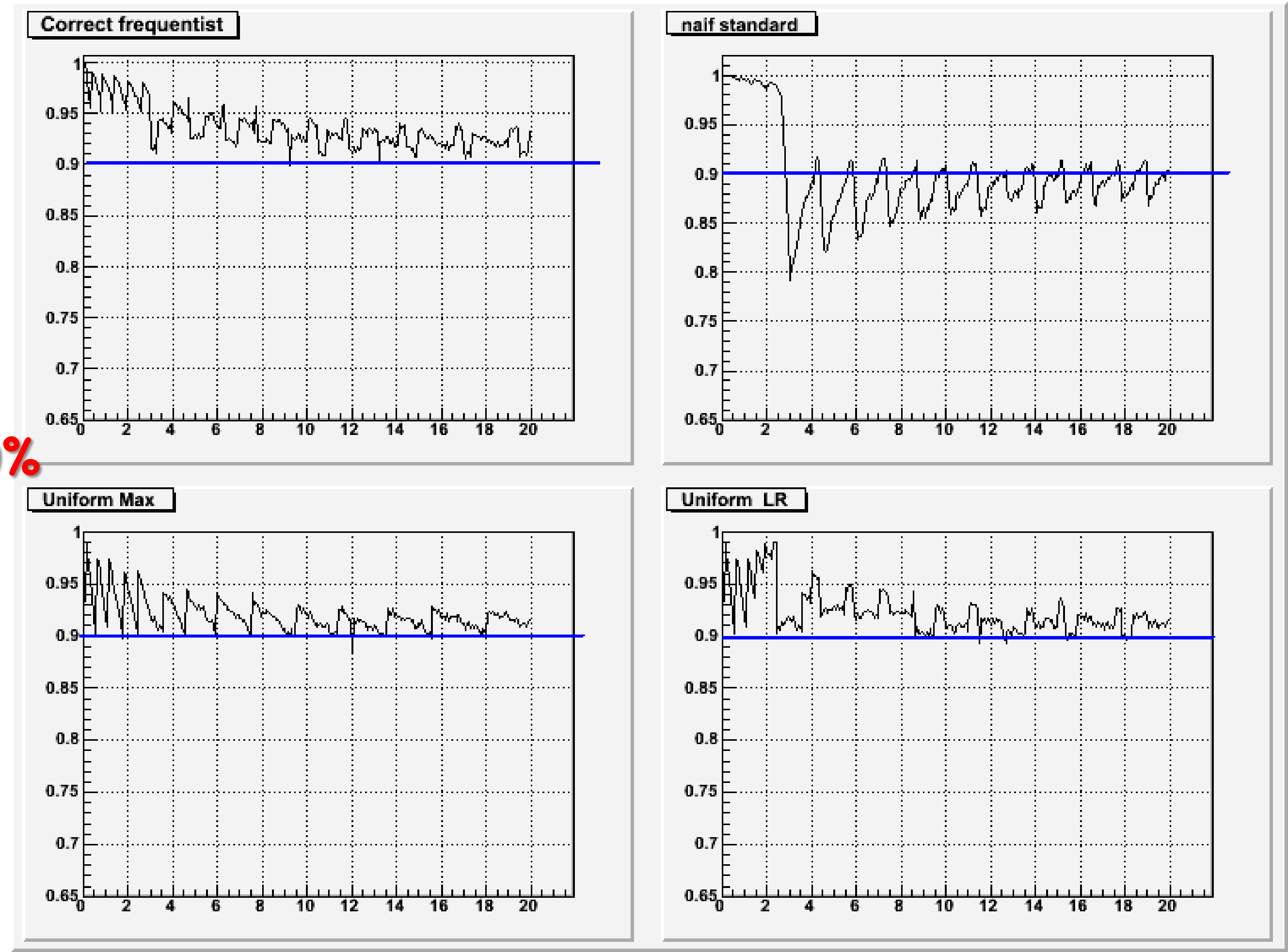


Poissonian Coverage simulation

CL=68%



Poissonian Coverage simulation



CL=90%

Counting experiments: new formula for the Poisson case

$$\frac{(x-\mu)}{\sqrt{\mu}} = t_\alpha \rightarrow \mu = x_\pm + \frac{t_\alpha^2}{2} \pm t_\alpha \sqrt{x_\pm + \frac{t_\alpha^2}{4}} \qquad x_\pm = x \pm 0.5$$

Wilson interval with Continuity correction gives the same results as …

$$\sum_{k=0}^{x} \frac{\mu_2^{\ x}}{x!} e^{-\mu_2} = \alpha/2$$

Exact frequentist
Clopper Pearson (1934) (PDG)

$$\sum_{k=x}^{\infty} \frac{\mu_1^{\ x}}{x!} e^{-\mu_1} = \alpha/2$$

The neutrino mass
...here Bayes helps!

An experiment with a Gaussian resolution of

$$\sigma = 3.3 \text{ eV}/c^2$$

measures the ν_e mass as:

$$m = -5.41 \text{ eV}/c^2$$

make the Bayesian estimate of m_ν.

Bayes formula

$$p(m_\nu; m, \sigma) = \frac{p(m; m_\nu, \sigma)\, p_\nu(m_\nu)}{\int p(m; m_\nu, \sigma)\, p_\nu(m_\nu)\, \mathrm{d}m_\nu}$$

Choosing the prior:

- **define** $0 \le m_\nu \le 20 - 30 \text{ eV}/c^2$ **;**
- **define** $\sigma_\nu = 10 \text{ eV}/c^2$
- **test three functional forms:**

1. **uniform:** $p_\nu = p_u(m_\nu) = 1/30\ ,\ \ 0 \le m_\nu \le 30$
2. **Gaussian:**

$$p_\nu = p_g(m_\nu) = \frac{2}{2\pi\sigma_\nu} \exp[-m_\nu^2/(2\sigma_\nu^2)]$$

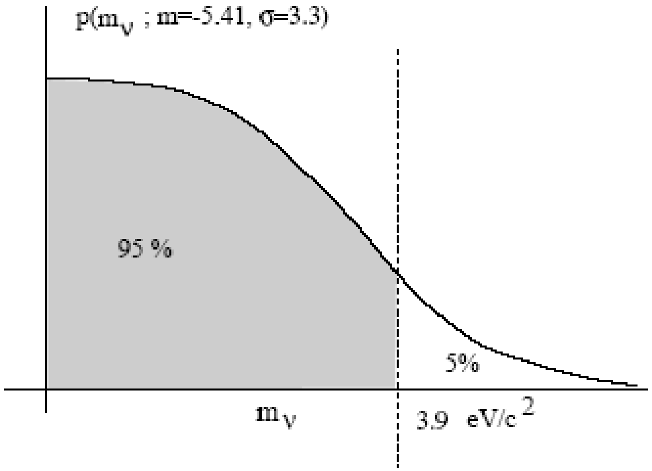
3. **triangular:** $p_\nu = p_t(m_\nu) = \frac{1}{450}\,(30 - m_\nu)$**,**
   $0 \le m_\nu \le 30 \text{ eV}/c^2$

The neutrino mass II

For example, using the uniform $p_u(m_\nu)$ and $\sigma = 3.3$, $m = -5.41$ ev$/c^2$:

$$p(m_\nu; m, \sigma) = \frac{\exp\left[-\frac{(m-m_\nu)^2}{2\sigma^2}\right]\frac{1}{30}}{\int_0^{30}\exp\left[-\frac{(m-m_\nu)^2}{2\sigma^2}\right]\frac{1}{30}\,\mathrm{d}m_\nu}$$

one obtains, at 95% probability:



- uniform: $0 \le m_\nu \le 3.9$ eV$/c^2$;
- Gaussian: $0 \le m_\nu \le 3.7$ eV$/c^2$;
- triangular: $0 \le m_\nu \le 3.7$ eV$/c^2$.

result "independent" of the prior!
Here the prior represent the knowledge, not the ignorance!!!

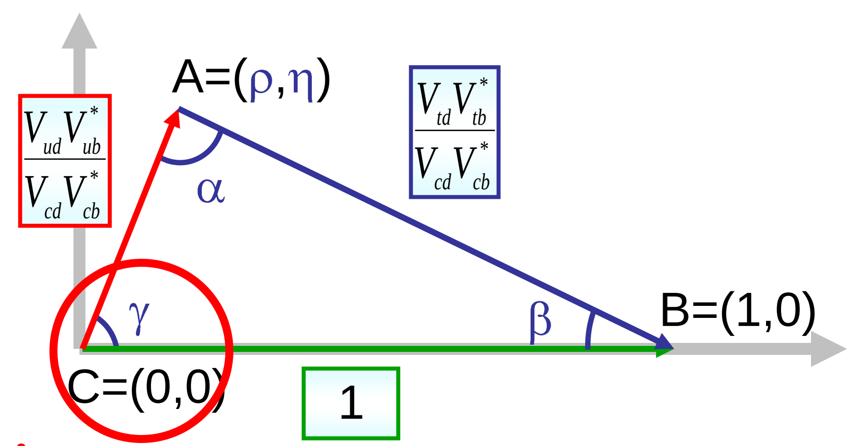
The Unitarity Triangle

$$\begin{array}{c} \\ u \\ c \\ t \end{array} \begin{array}{c} \begin{array}{ccc} d & s & b \end{array} \\ \begin{pmatrix} \boxed{V_{ud}} & V_{us} & \boxed{V_{ub}} \\ \boxed{V_{cd}} & V_{cs} & \boxed{V_{cb}} \\ \boxed{V_{td}} & V_{ts} & \boxed{V_{tb}} \end{pmatrix} \end{array}$$

- Quark mixing is described by the CKM matrix
- Unitarity relations on matrix elements lead to a triangle in the complex plane

$$V^{\dagger}V = I$$

$$\boxed{V_{ud}V_{ub}^{*}} + \boxed{V_{cd}V_{cb}^{*}} + \boxed{V_{td}V_{tb}^{*}} = 0$$



CP Violation

Constraints, Parameters	Value	Gauss Error	Flat Error	Comments
λ	0.2258	0.0014	-	
$\lvert V_{cb}\rvert$ (10^{-3})	39.2	1.1	-	Average of exclusive
$\lvert V_{cb}\rvert$ (10^{-3})	41.7	0.7	-	Average of inclusive
$\lvert V_{ub}\rvert$ 10^{-4} (excl.)	35.0	4.0	-	HFAG BR + Lattice QCD
$\lvert V_{ub}\rvert$ 10^{-4}(incl. HFAG)	39.9	1.5	4.0	HFAG average
m_b (GeV/c^2)	4.21	0.08	-	
m_c (GeV/c^2)	1.3	0.1	-	
$\Delta(m_d)$ (ps^{-1})	0.507	0.005	-	WA (CDF/CLEO/LEP/Babar/Belle)
$\Delta(m_s)$ (ps^{-1})	17.77	0.12	-	CDF Likelihood is used.
m_t (GeV/c^2)	161.2	1.7	-	(CDF/D0)
$f_{Bs}\sqrt{B_{Bs}}$ (MeV)	270	30	-	Lattice QCD
ξ	1.21	0.04	-	Lattice QCD
$\lvert\varepsilon_K\rvert 10^{-3}$	2.280	0.013	-	
B_K	0.75	0.07	-	Lattice QCD
f_K (GeV)	0.160	-	-	
$\Delta(m_K)$ (10^{-2} ps^{-1})	0.5301	-	-	
$\alpha_s(M_Z)$	0.119	0.003	-	
G_F (10^{-5} GeV^{-2})	1.16639	-	-	
m_W (GeV/c^2)	80.425	-	-	
m_{Bd} (GeV/c^2)	5.279	-	-	
m_{Bs} (GeV/c^2)	5.375	-	-	
m_{K^0} (GeV/c^2)	0.497648	-	-	

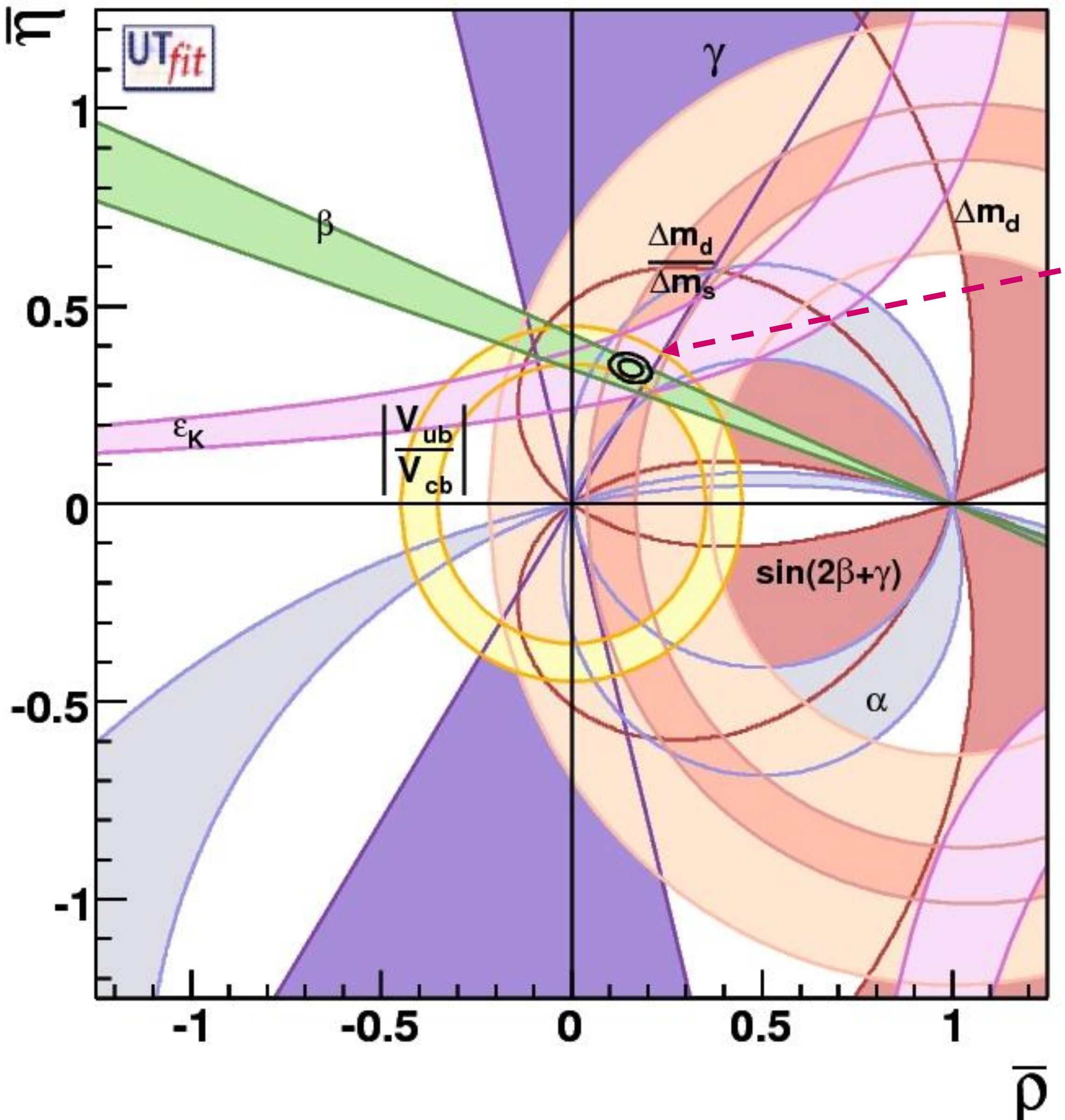
A Bayesian application: UTFit

- UTFit: Bayesian determination of the CKM unitarity triangle
  - Many experimental and theoretical inputs combined as product of PDF
  - Resulting likelihood interpreted as Bayesian PDF in the UT plane
- Inputs:
  - Standard Model experimental measurements and parameters
  - Experimental constraints

Combine the constraints

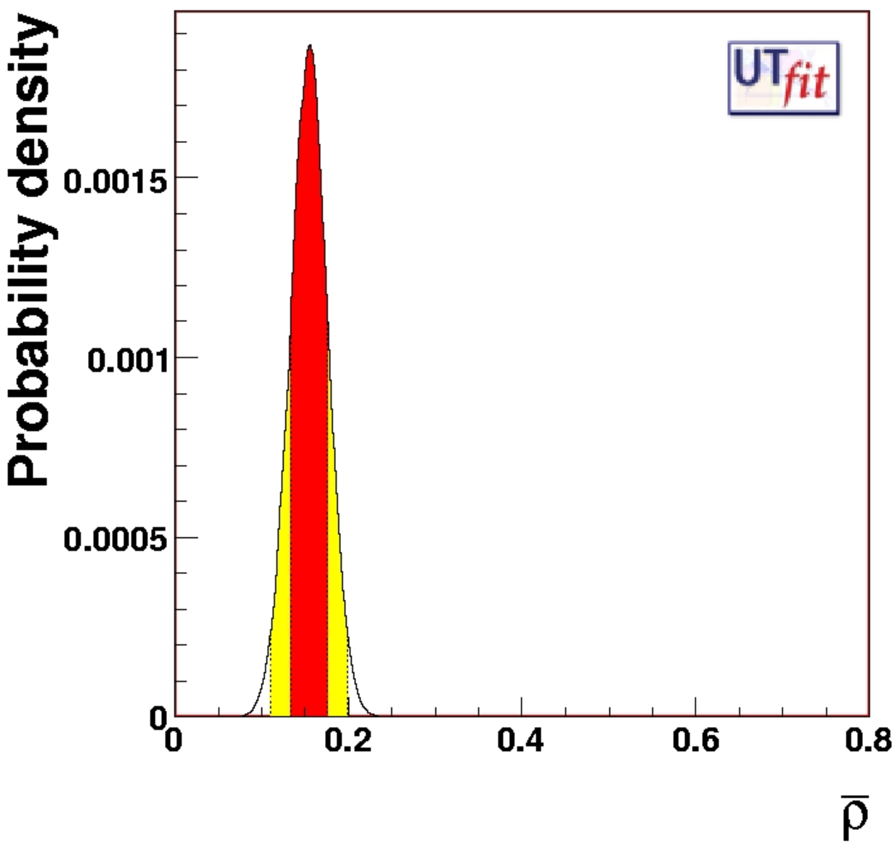
- Given $\{x_i\}$ parameters and $\{c_i\}$ constraints
- Define the combined PDF
  - $f(\rho, \eta, x_1, x_2, \ldots, x_N \mid c_1, c_2, \ldots, c_M) \propto$
    $\prod_{j=1,M} f_j(c_j \mid \rho, \eta, x_1, x_2, \ldots, x_N) \cdot$
    $\prod_{i=1,N} f_i(x_i) \cdot f_o(\rho, \eta)$ ← A priori PDF
  - PDF taken from experiments, wherever it is possible
- Determine the PDF of (ρ, η) integrating over the remaining parameters
  - $f(\rho, \eta) \propto$
    $\int \prod_{j=1,M} f_j(c_j \mid \rho, \eta, x_1, x_2, \ldots, x_N) \cdot$
    $\prod_{i=1,N} f_i(x_i) \cdot f_o(\rho, \eta)$

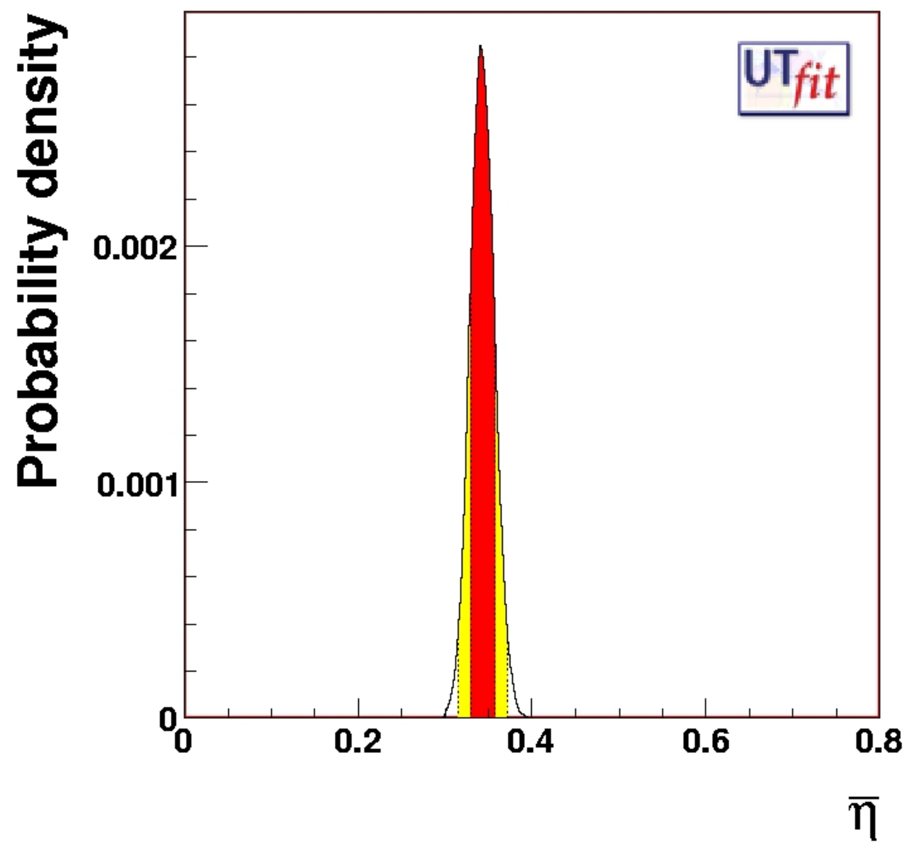
Unitarity Triangle fit



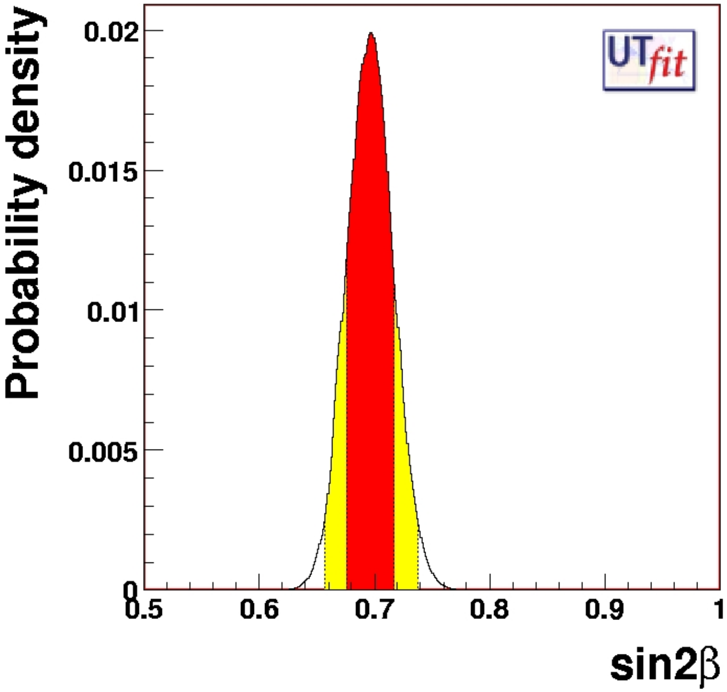
68%, 95% contours

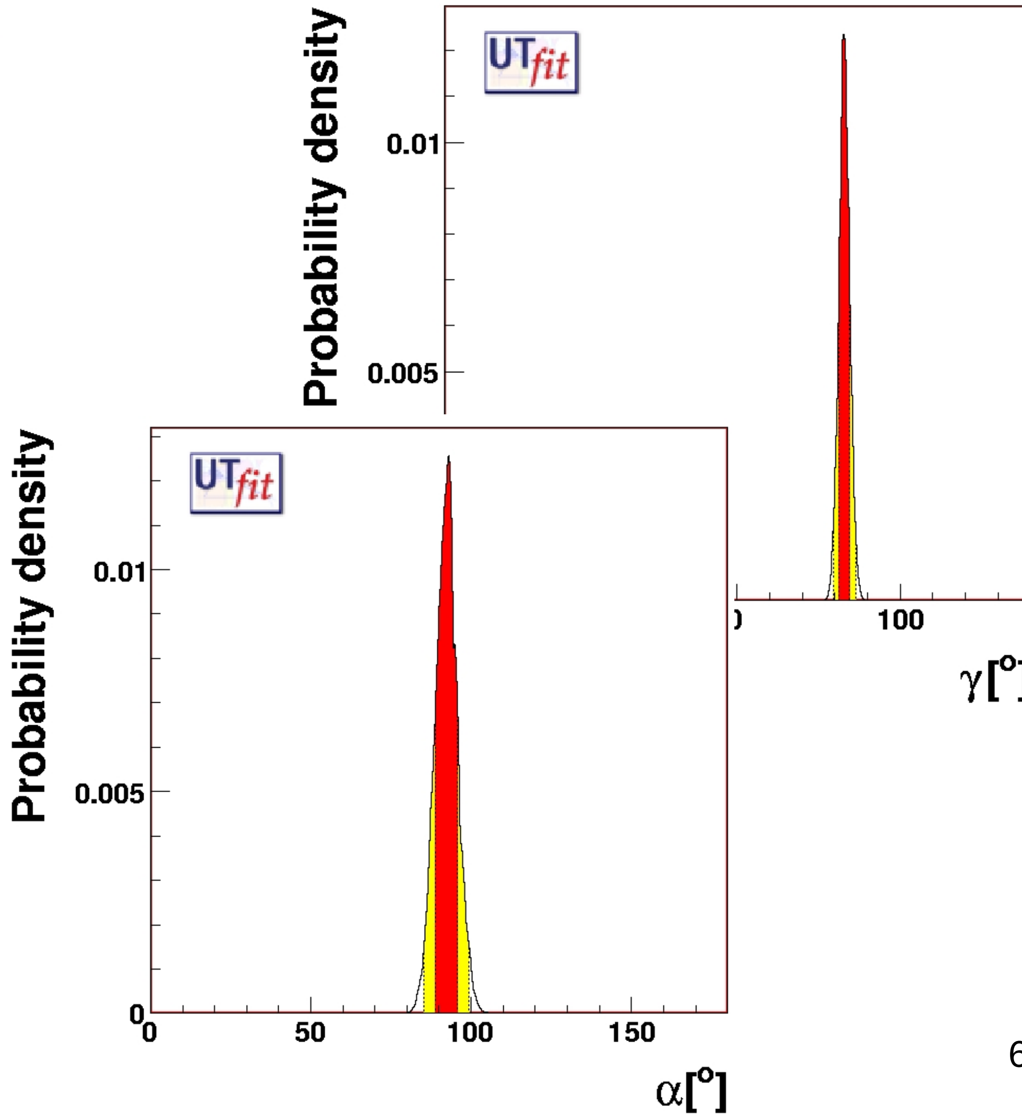
PDFs for ρ and η





Projections on other observables





A Frequentist application: RFit

- RFit: to choose a point in the $\rho-\eta$ plane, and ask for the best set of the parameters for this points. The χ^2 values give the requested confidence region.

- No a priori distribution of parameters is requested

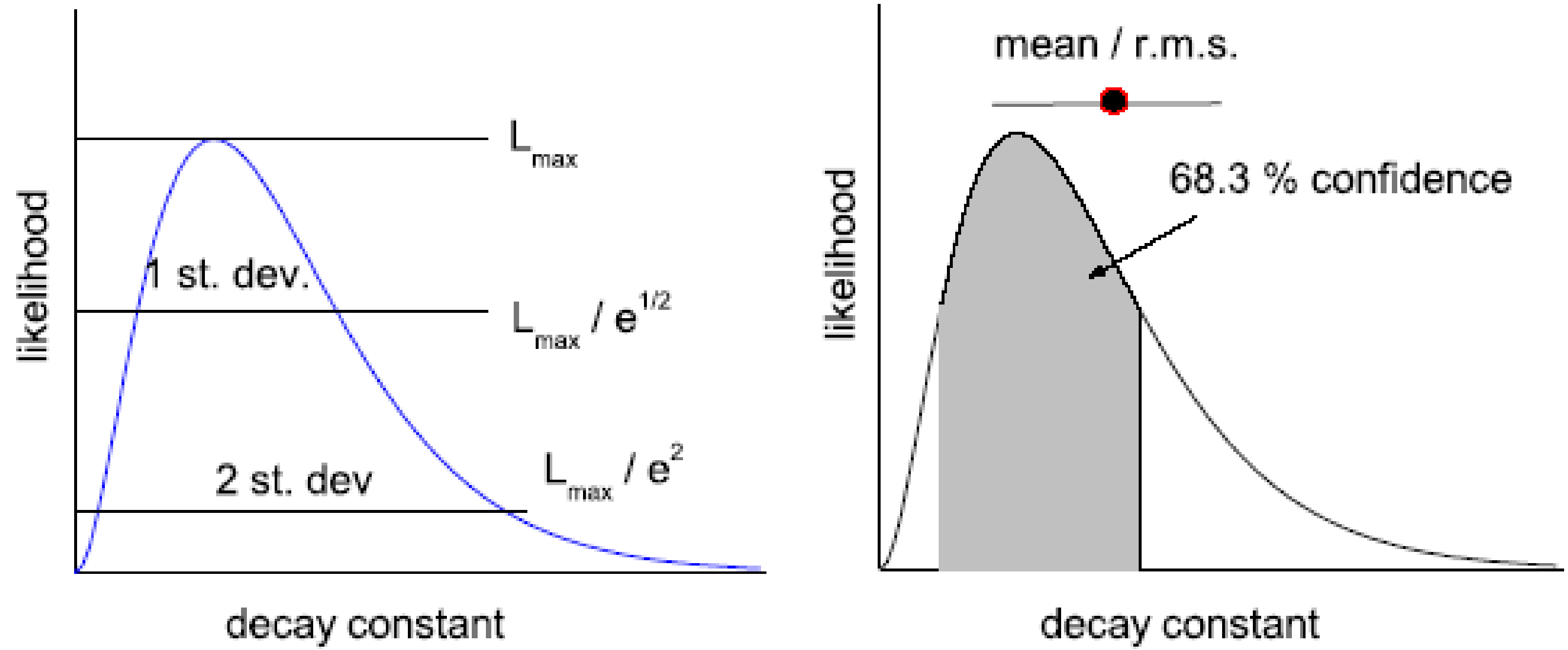
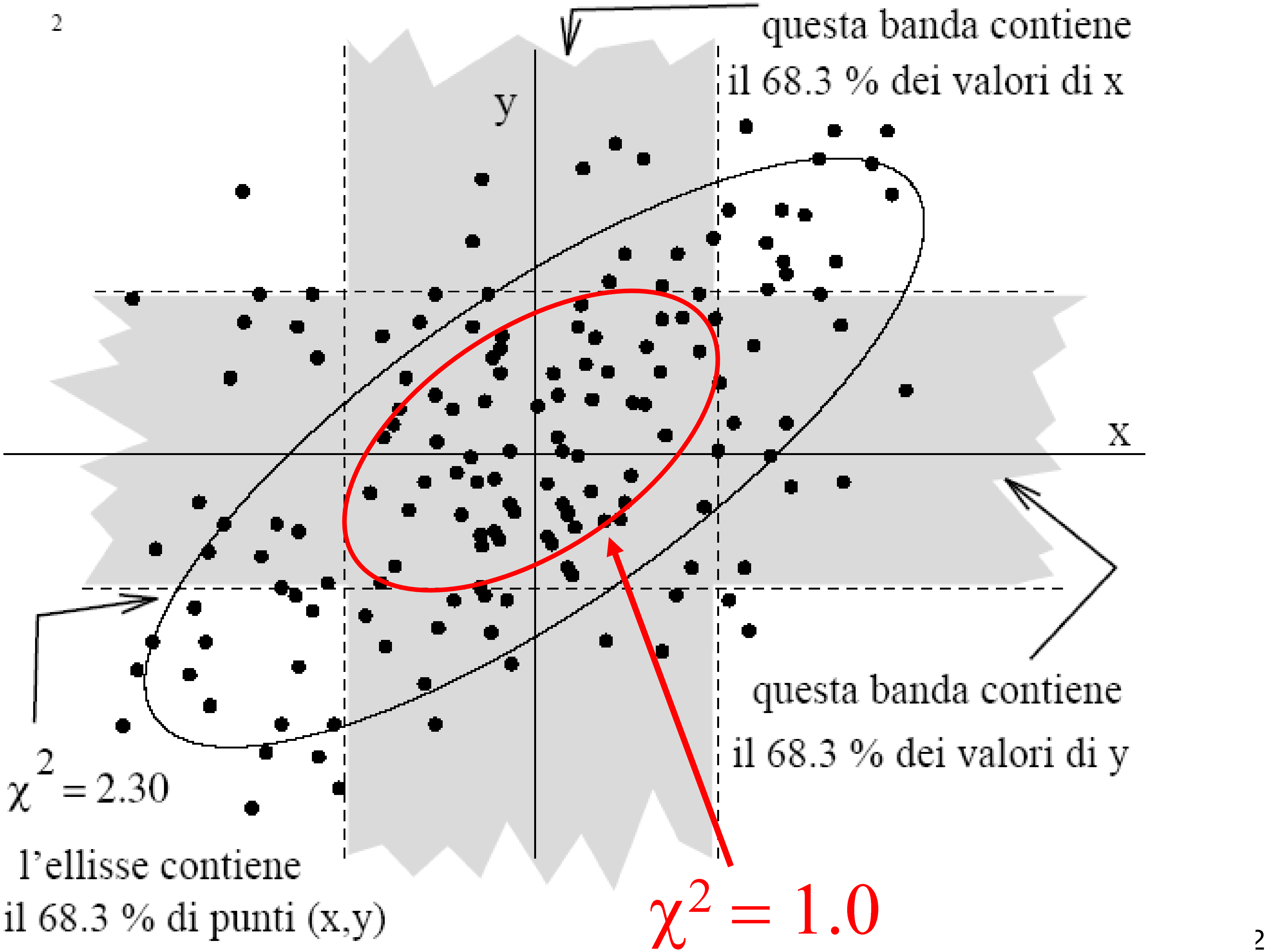


Fig. 18. Likelihood ratio limits (left) and Bayesian limits (right)

$$\ln e^{-\frac{1}{2}\frac{(x-\theta)^2}{\sigma}} = -\frac{1}{2}\frac{(x-\theta)^2}{\sigma^2} \Rightarrow -2\ln L(x;\theta) \approx \chi^2(\theta)$$


2
questa banda contiene
il 68.3 % dei valori di x
y
x
questa banda contiene
il 68.3 % dei valori di y
$\chi^2 = 2.30$
l'ellisse contiene
il 68.3 % di punti (x,y)
$\chi^2 = 1.0$
2

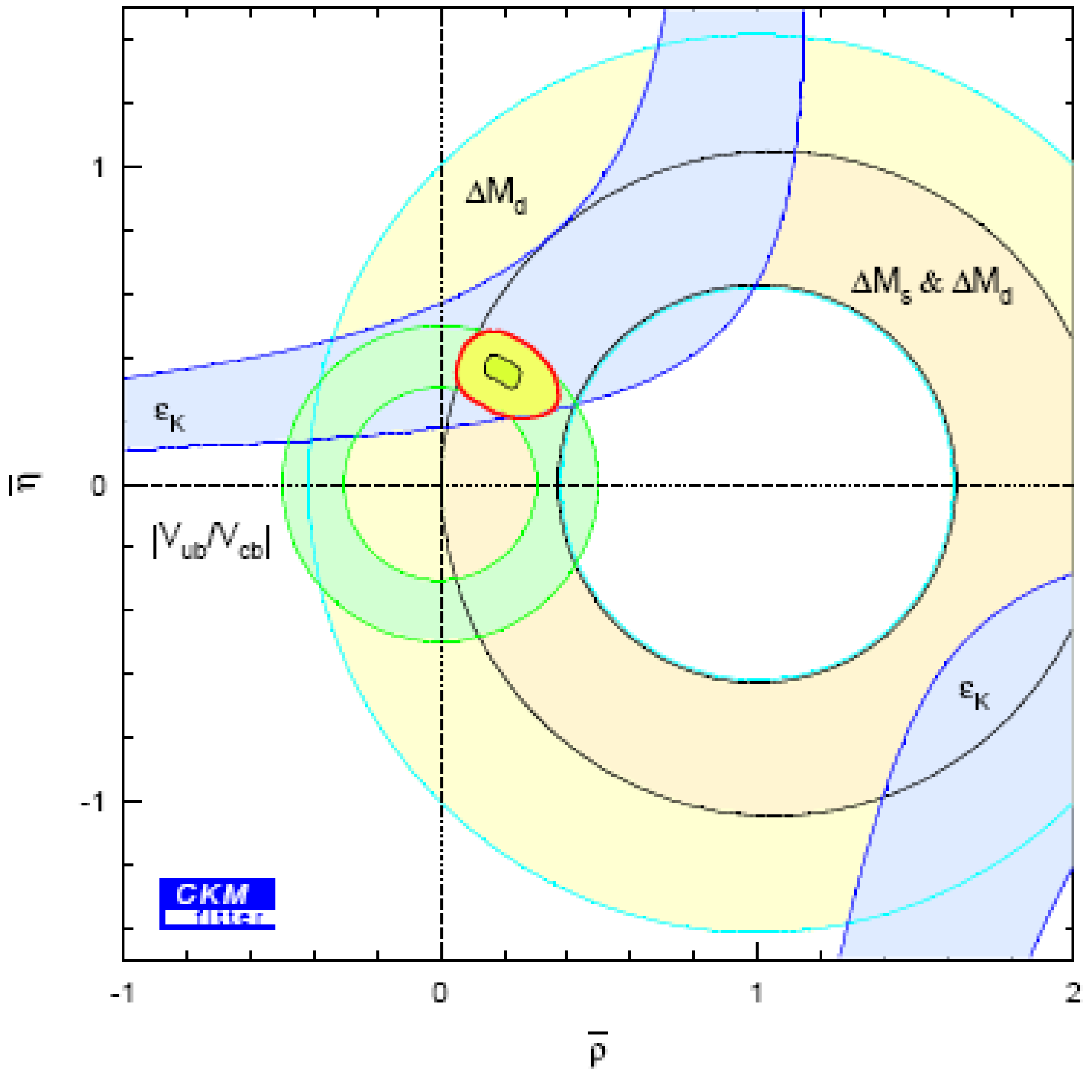


Fig. 5.4: *Different single constraints in the $\bar{\rho} - \bar{\eta}$ plane shown at 95 % CL contours. The 95 % and 10 % CL contours for the combined fit are also shown.*

5.3. Results and Comparison

*R*fit Method			
Parameter	$\leq 5\%$ CL	$\leq 1\%$ CL	$\leq 0.1\%$ CL
$\bar{\rho}$	0.091 - 0.317	0.071 - 0.351	0.042 - 0.379
$\bar{\eta}$	0.273 - 0.408	0.257 - 0.423	0.242 - 0.442
$\sin 2\beta$	0.632 - 0.813	0.594 - 0.834	0.554 - 0.855
γ°	42.1 - 75.7	38.6 - 78.7	36.0 - 83.5

Bayesian Method			
Parameter	5% CL	1% CL	0.1% CL
$\bar{\rho}$	0.137 - 0.295	0.108 - 0.317	0.045 - 0.347
$\bar{\eta}$	0.295 - 0.409	0.278 - 0.427	0.259 - 0.449
$\sin 2\beta$	0.665 - 0.820	0.637 - 0.841	0.604 - 0.863
γ°	47.0 - 70.0	44.0 - 74.4	40.0 - 83.6

Ratio *R*fit/Bayesian Method			
Parameter	5% CL	1% CL	0.1% CL
$\bar{\rho}$	1.43	1.34	1.12
$\bar{\eta}$	1.18	1.12	1.05
$\sin 2\beta$	1.17	1.18	1.16
γ°	1.46	1.31	1.09

Table 5.3: *Ranges at difference C.L for $\bar{\rho}$, $\bar{\eta}$, $\sin 2\beta$ and γ. The measurements of $|V_{ub}|$ / $|V_{cb}|$ and ΔM_d, the amplitude spectrum for including the information from the $B_s^0 - \overline{B}_s^0$ oscillations, $|\varepsilon_K|$ and the measurement of $\sin 2\beta$ have been used.*

Conclusions

- The usual formulae used by physicists in counting experimets should be abandoned

- By adopting a practical attitude, also bayesian formulae can be tested in a frequentist way

- frequentism is the best way to give the results of an experiment in the form

$$x \pm \sigma$$

  but other forms are also possible

- physicists should use Bayes formulae to parametrize the previous (th or exp)

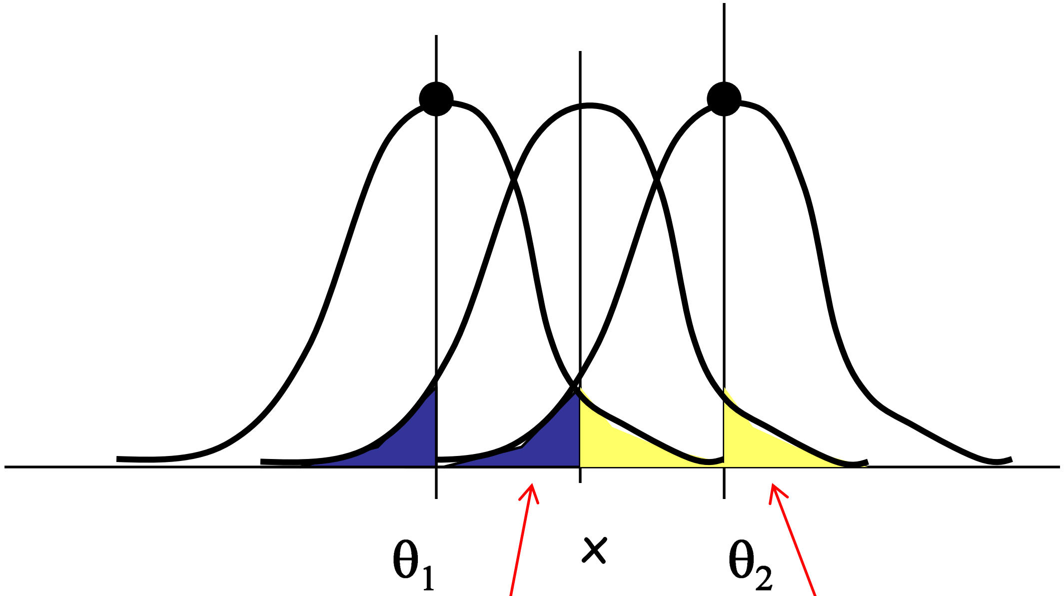
  knowledge, not the ignorance

Quantum Mechanics:
frequentist or bayesian?
Born or Bohr?

$$\int |\psi|^2 dx$$

The standard interpretation is frequentist

END



Neyman integrals

Bootstrap

$$F(x;\theta) = 1 - F(\theta;x)$$

Search for pivotal variables

This method avoids the graphic procedure and the resolution of the Neyman integrals

From coin tossing to physics: the efficiency measurement

$$
\begin{aligned}
P(\varepsilon;k,n) &= (n+1)\binom{n}{k}\varepsilon^k(1-\varepsilon)^{n-k} \\
&= \frac{(n+1)!}{k!\,(n-k)!}\,\varepsilon^k(1-\varepsilon)^{n-k}
\end{aligned}
$$

$$
\begin{aligned}
\bar{\varepsilon} &= \int_0^1 \varepsilon\; P(\varepsilon;k,n)\; d\varepsilon \\
&= \frac{(n+1)!}{k!\,(n-k)!}\int_0^1 \varepsilon^{k+1}(1-\varepsilon)^{n-k} d\varepsilon \\
&= \frac{k+1}{n+2}
\end{aligned}
$$

Valid also for *k=0* and *k=n*

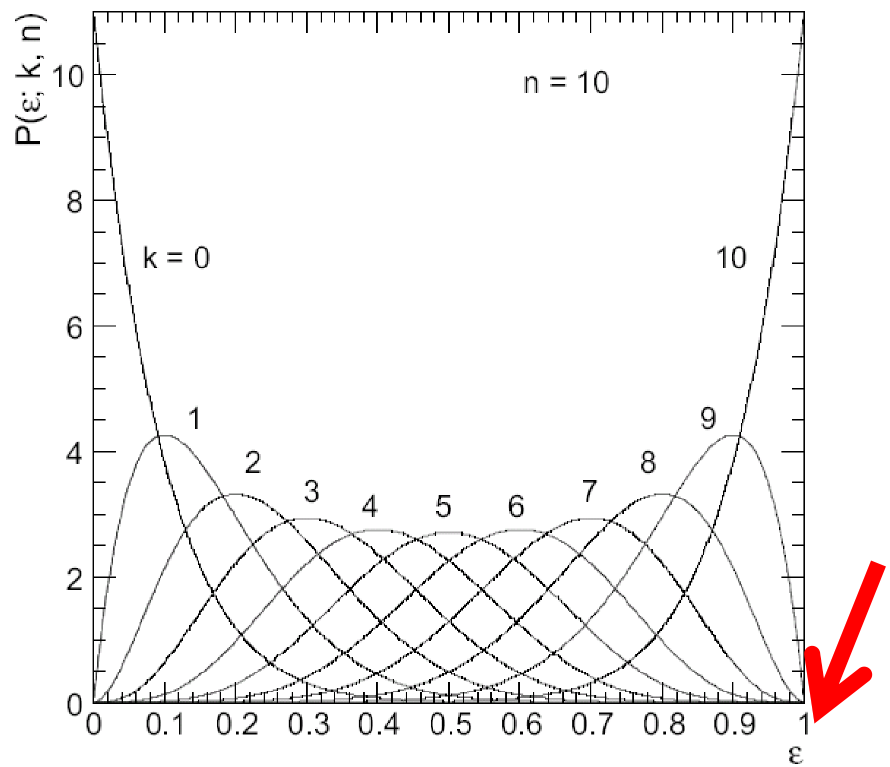
$$
\begin{aligned}
V(\varepsilon) &= \overline{\varepsilon^2} - \bar{\varepsilon}^2 \\
&= \int_0^1 \varepsilon^2\; P(\varepsilon;k,n)\; d\varepsilon - \bar{\varepsilon}^2 \\
&= \frac{(k+1)\,(k+2)}{(n+2)\,(n+3)} - \frac{(k+1)^2}{(n+2)^2}.
\end{aligned}
$$

ArXiv:physics/0701199v1

Treatment of Errors in Efficiency Calculations

T. Ullrich and Z. Xu
Brookhaven National Laboratory

February 2, 2008



re 1: The probability density function $P(\varepsilon;k,n)$ for $n=10$ and $k=0,1,\ldots,10$.

Frequentist C.I.
right and wrong definitions

RIGHT quotations:

- CL is the probability that the random interval $[T_1, T_2]$ covers the true value θ;
- in an infinite set of repeated identical experiments, a fraction equal to CL will succeed in assigning $\theta \in [\theta_1, \theta_2]$;
- if $\theta \notin [\theta_1, \theta_2]$, one can obtain $\{I = [\theta_1, \theta_2]\}$ in a fraction of experiments $\leq 1 - CL$
- if $H_0 : \theta \notin [\theta_1, \theta_2]$ the probability to reject a true H_0 is $1 - CL$ (falsification).
  see upper and lower limits estimates.

WRONG quotations

- CL is the degree of belief that the true value is in $[\theta_1, \theta_2]$
- $P\{\theta \in [\theta_1, \theta_2]\} = CL$
  (θ is not a random variable!)