Statistical Methods in Particle Physics

7. Confidence Limits and Intervals

Prof. Dr. Klaus Reygers (lectures)
Dr. Sebastian Neubert (tutorials)

Heidelberg University
WS 2017/18

Point Estimates and Limits

http://pdg.lbl.gov/2017/reviews/rpp2016-rev-statistics.pdf

One often reports a point estimate and its standard deviation: $\hat{\theta},\ \hat{\sigma}_{\hat{\theta}}$

In some situation one rather reports an interval instead, e.g. when

- the p.d.f. of the estimator is not Gaussian
- one has physical boundaries on the possible values of the parameter

Goals

- communicate as objectively as possible the result of the experiment;
- provide an interval that is constructed to cover the true value of the parameter with a specified probability;
- provide the information needed by the consumer of the result to draw conclusions about the parameter or to make a particular decision;
- draw conclusions about the parameter that incorporate stated prior beliefs.

With a sufficiently large data sample, the point estimate and standard deviation (or for the multiparameter case, the parameter estimates and covariance matrix) satisfy essentially all of these goals.

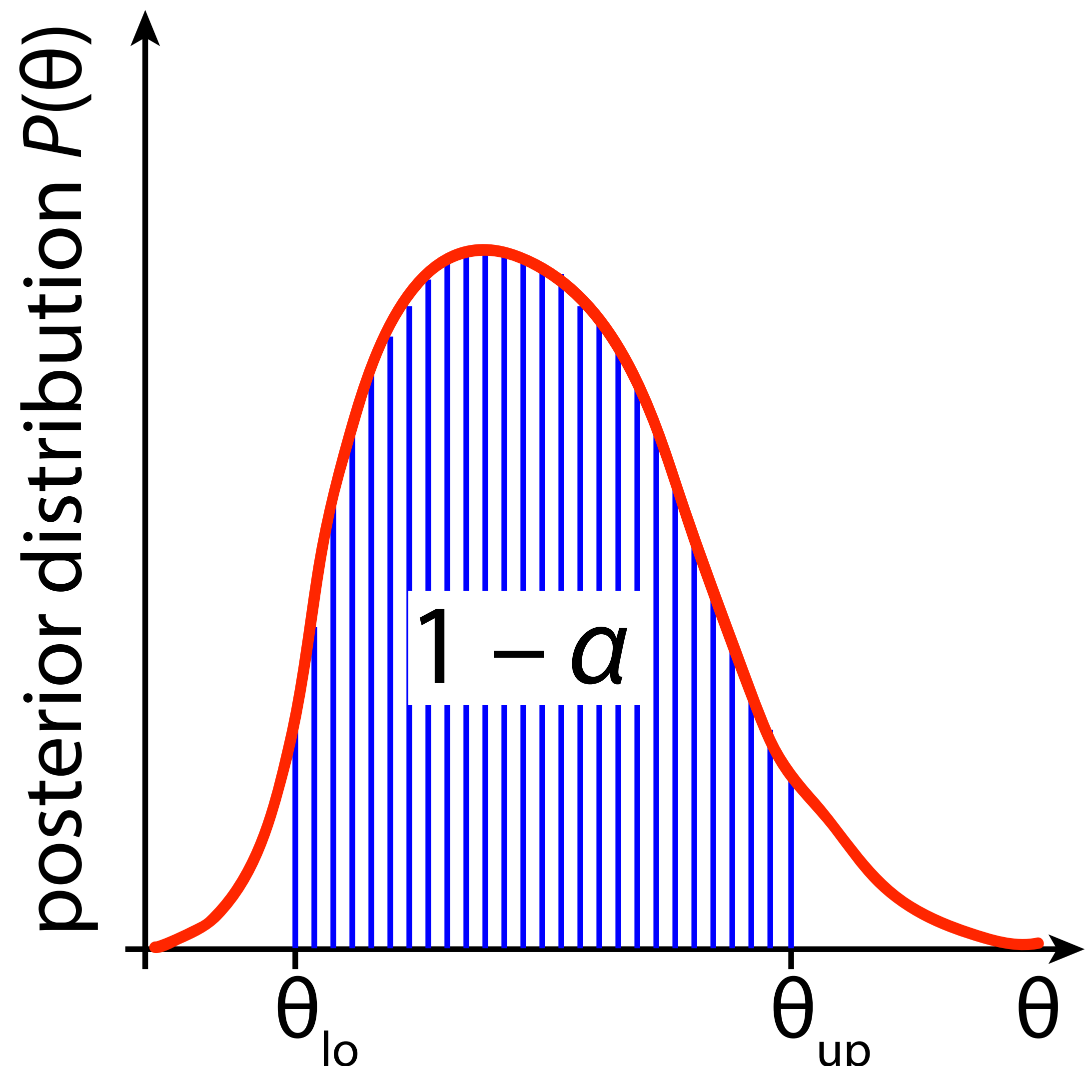
Bayesian Credible Intervals

Bayesian approach:
report full posterior p.d.f.

In case a range is desired:
integrate posterior p.d.f.

$$1 - \alpha = \int_{\theta_{\rm lo}}^{\theta_{\rm up}} p(\theta|x)\,\mathrm{d}\theta$$

cf. LIGO paper: $1 - \alpha = 0.9$
("90% credible interval")



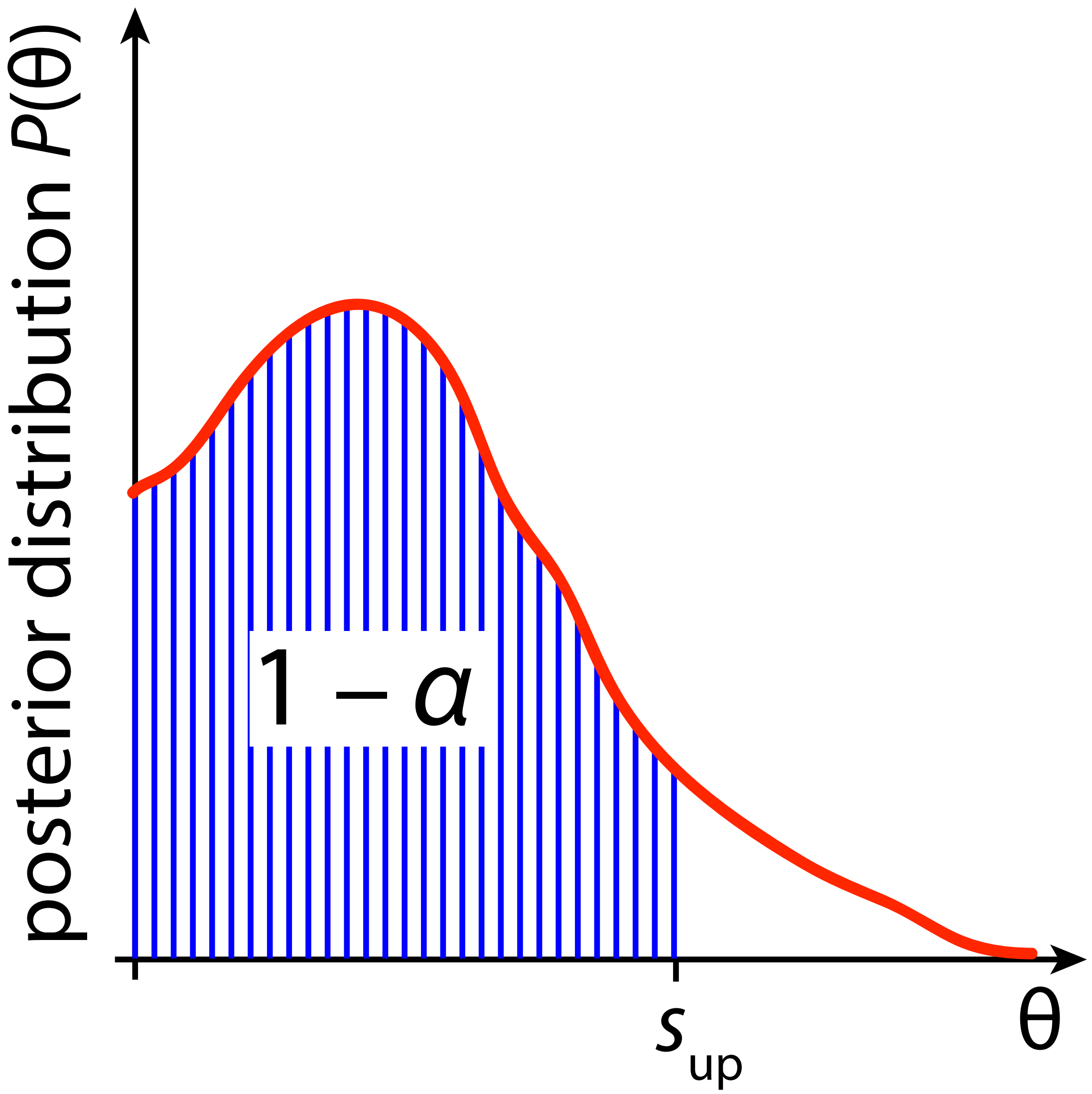
Different options to construct the interval $[\theta_{lo}, \theta_{up}]$:

- $[-\infty, \theta_{lo}]$ and $[\theta_{up}, \infty]$ both correspond to a probability $\alpha/2$
- Antisymmetric intervals, e.g. $[-\infty, \theta_{up}]$ corresponding to probability $1 - \alpha$
- Symmetric interval around maximum value corresponding to probability $1 - \alpha$
- $p(\theta|x)$ higher than for any θ not belonging to the set (could give disjoint intervals)
- …

Bayesian Upper Limits

$$1 - \alpha = \int_{-\infty}^{s_{\mathrm{up}}} p(\theta|x)\, \mathrm{d}\theta$$

In case of a physical lower bound, lower integration limit is replaces by physics bound (e.g., mass of a particle $m > 0$)



Example:
Bayesian Upper Limits for a Poisson Variable *n* (I)

In a counting experiment one would like to measure a signal *s*. Suppose the average number of background counts *b* is known:

Likelihood to find *n* counts:

$$P(n|s) = \frac{(s+b)^n}{n!} e^{-(s+b)}$$

Let's take the following prior for *s*:

$$\pi(s) = \begin{cases} 0, & < 0 \\ 1, & s \geq 0 \end{cases}$$

Upper limit:

$$1 - \alpha = \int_{-\infty}^{s_{\mathrm{up}}} p(s|n)\,\mathrm{d}s = \frac{\int_{-\infty}^{s_{\mathrm{up}}} P(n|s)\pi(s)\,\mathrm{d}s}{\int_{-\infty}^{\infty} P(n|s)\pi(s)\,\mathrm{d}s}$$

Solution (without going into details here):

$$s_{\mathrm{up}} = \frac{1}{2} F_{\chi^2}^{-1}[p, 2(n+1)] - b, \qquad p = 1 - \alpha\left(1 - F_{\chi^2}[2b, 2(n+1)]\right)$$

$F_{\chi^2}^{-1}$: quantile of the χ^2 distribution (inverse of the cumulative distribution)

Example:
Bayesian Upper Limits for a Poisson Variable *n* (II)

Special case: $b = 0$

$$s_{\rm up} = \frac{1}{2} F^{-1}_{\chi^2}[1-\alpha, 2(n+1)]$$

In root:

```
void bayesian_poisson_upper_limits() {

    // parameters: b = expected background counts,
    // alpha = confidence level
    Double_t b = 0, alpha = 0.1;

    // loop over observed counts n
    for (Int_t n = 0; n <= 10; ++n) {
        Double_t p = 1 - alpha * TMath::Prob(2 * b, 2 * (n + 1));
        Double_t s_up = 0.5 * TMath::ChisquareQuantile(p, 2 * (n + 1)) - b;
        cout << n << "  " << s_up << endl;
    }
}
```

n	s_up
0	2.30259
1	3.88972
2	5.32232
3	6.68078
4	7.99359
5	9.27467
6	10.5321
7	11.7709
8	12.9947
9	14.206
10	15.4066

Frequentist Confidence Intervals

Construct interval that includes (covers) the true value of the parameter with a probability p

- p is called the coverage probability
- Constructed confidence interval depends on data and would fluctuate if we were to repeat the experiment many times
- coverage probability = fraction of intervals that would cover the true value in repeated experiments

Neyman Construction (I)

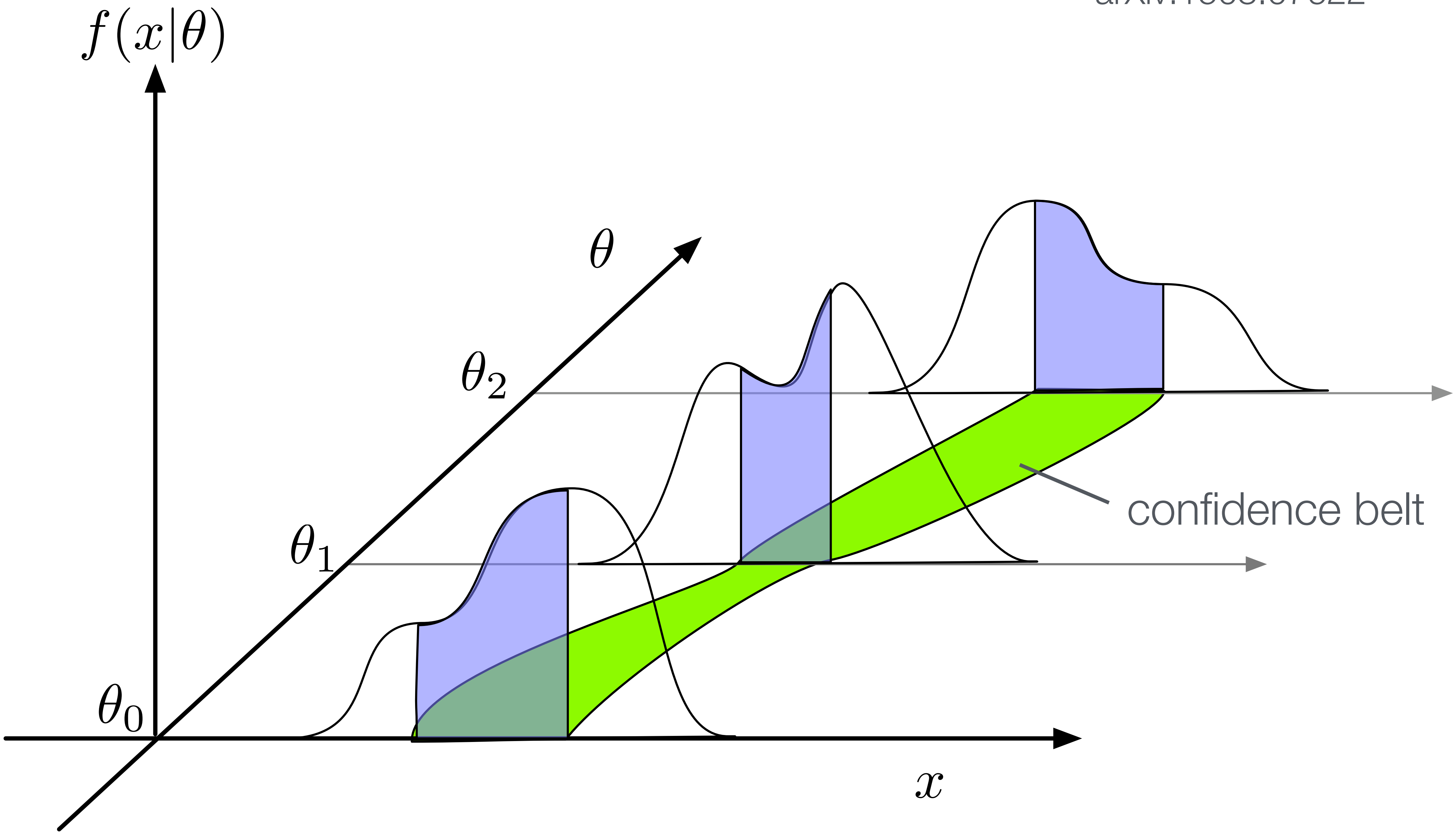
The Neyman construction for constructing frequentist confidence intervals involves the following steps:

1. Given a true value of the parameter θ, determine a p.d.f. f(*x*; θ) for the outcome of the experiment. Often *x* is an estimator for the θ.
2. Using some procedure, define an interval in *x* that has a specified probability (say, 90%) of occurring
3. Do this for all possible true values of θ, and build a confidence belt of these intervals.

In practice, the p.d.f. of step 1 might come from Monte Carlo simulation.

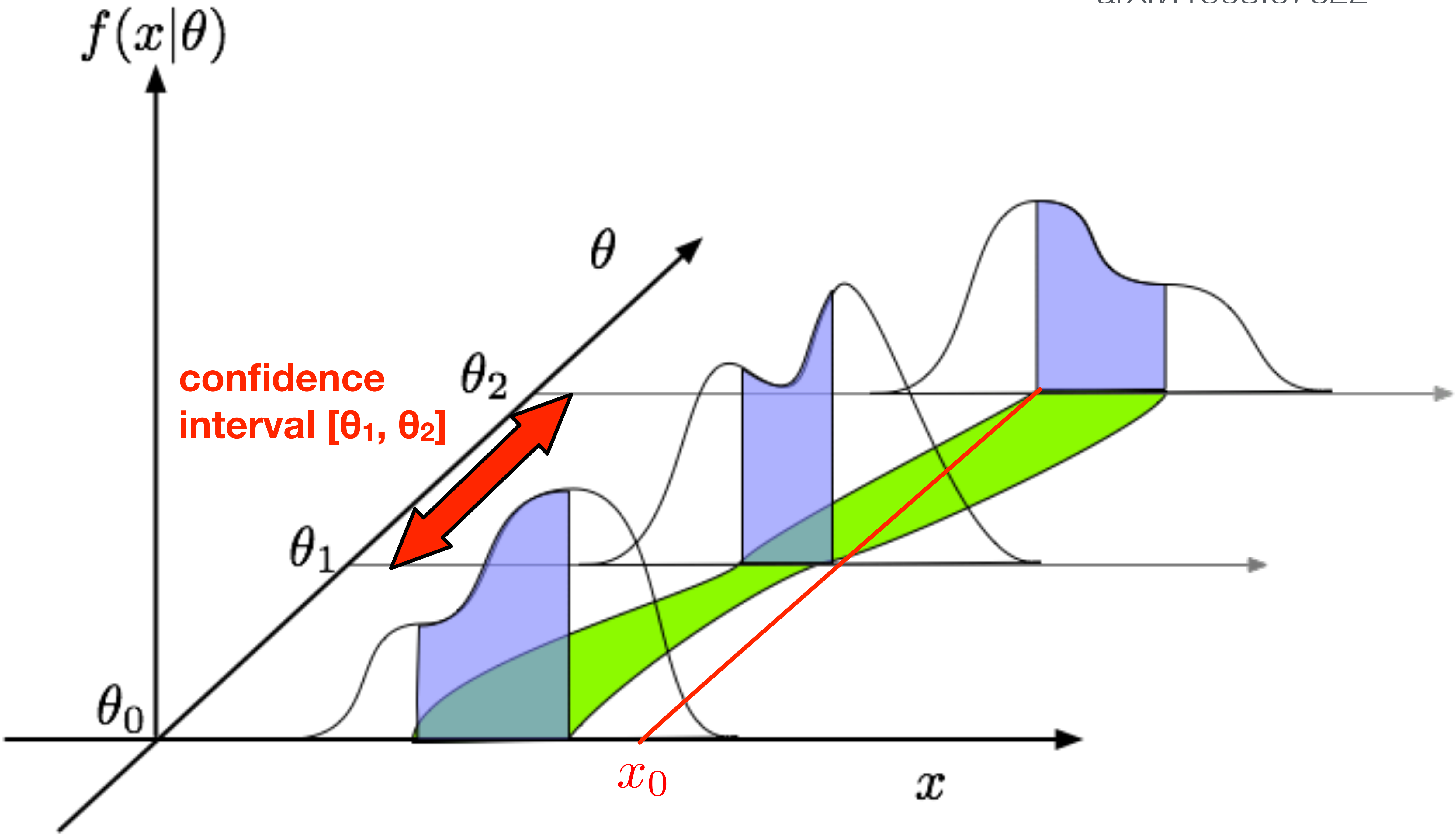
Neyman Construction (II)

arXiv:1503.07622



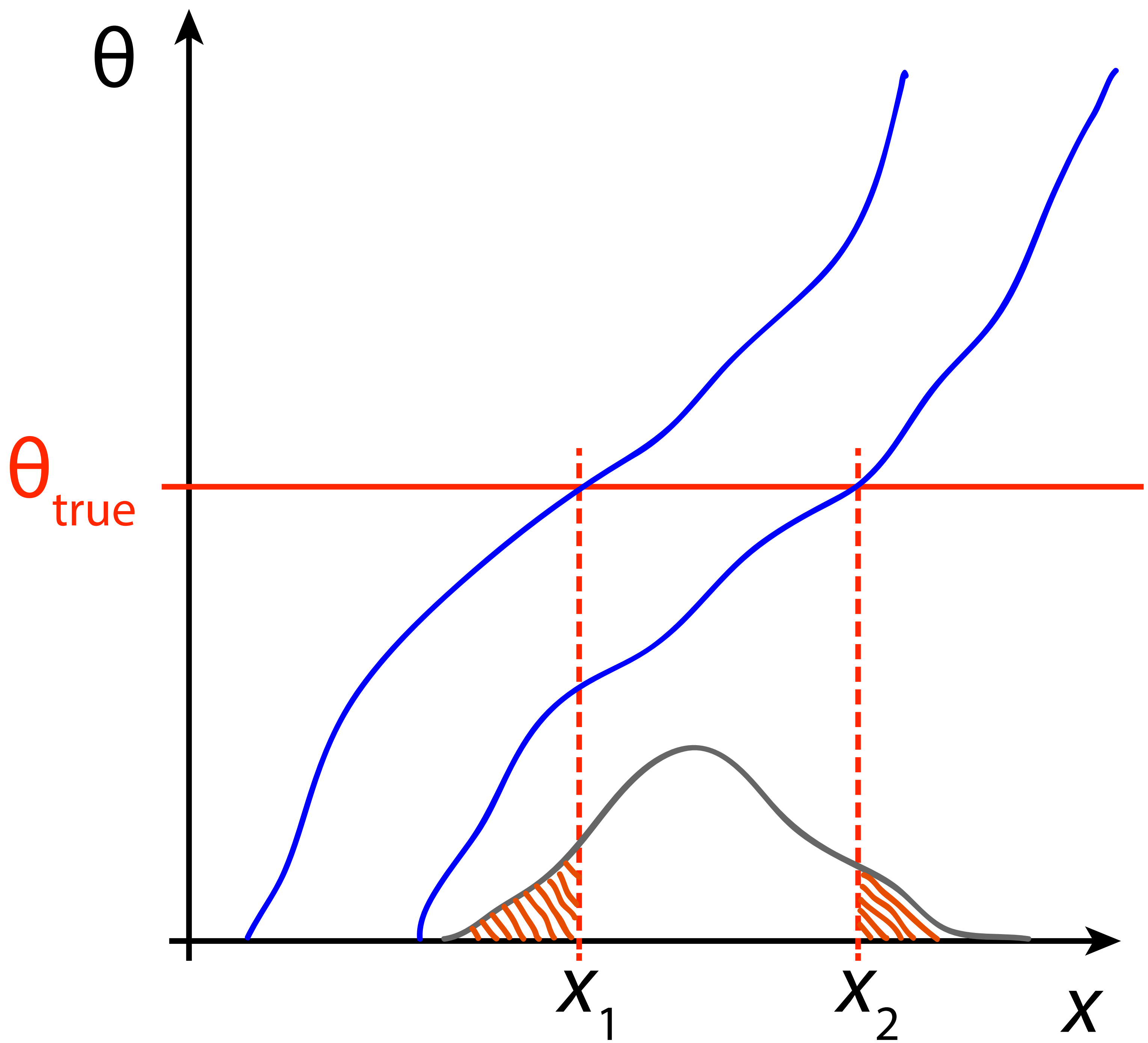
Neyman Construction (III)

arXiv:1503.07622



For a measured x_0 one draws a vertical line through x_0. The confidence interval for θ is the set of all values of θ for which the line is inside the confidence belt.

Coverage of the Neyman Interval



Suppose θ_{true} is the true value of θ.

With a probability α the measured value x is in $[x_1, x_2]$.

Constructed interval: $[\theta_1, \theta_2]$

x in $[x_1, x_2] \Rightarrow \theta_{true}$ in $[\theta_1, \theta_2]$

x not in $[x_1, x_2]$
$\Rightarrow \theta_{true}$ not in $[\theta_1, \theta_2]$

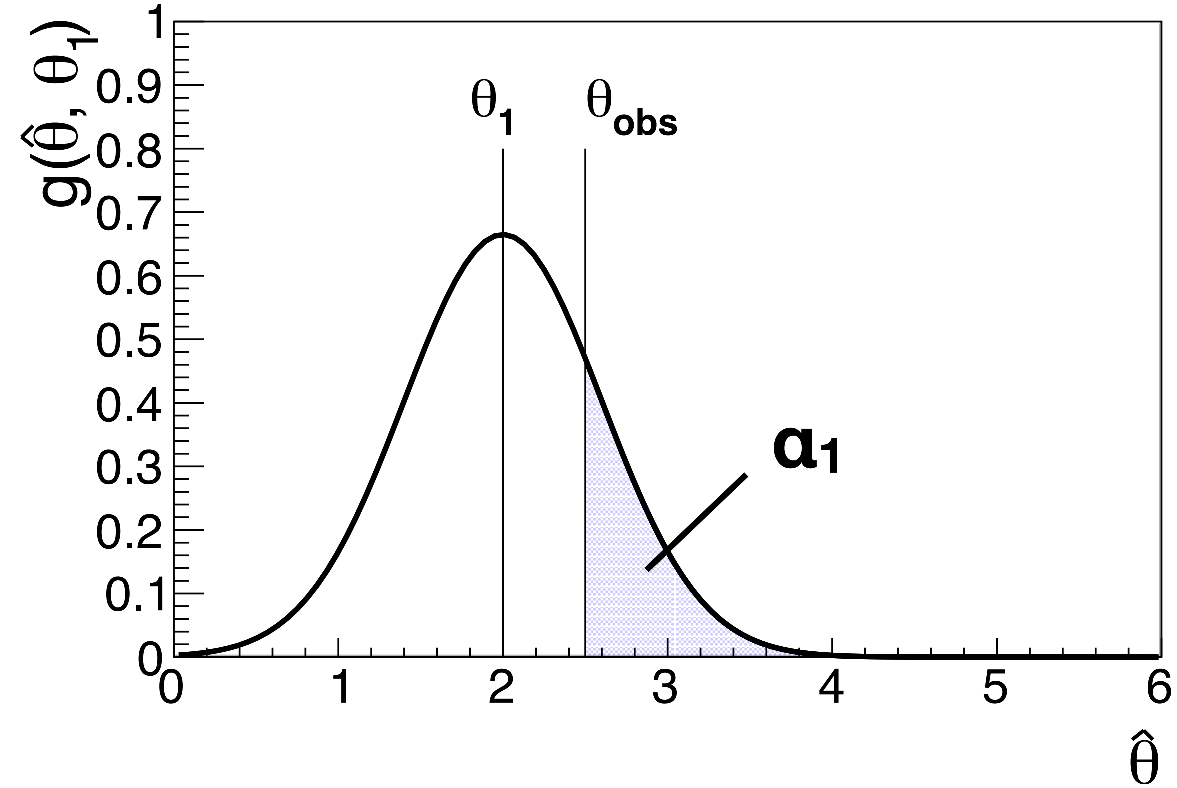
Coverage of $[\theta_1, \theta_2]$ is $1 - \alpha$

Confidence Interval for a Gaussian Distributed Estimator

Consider a parameter θ whose estimator is distributed as

$$g(\hat{\theta}; \theta) = \frac{1}{\sqrt{2\pi}\sigma_{\hat{\theta}}} \exp\left(-\frac{1}{2}\frac{(\hat{\theta}-\theta)^2}{\sigma_{\hat{\theta}}^2}\right)$$

"sampling distribution"



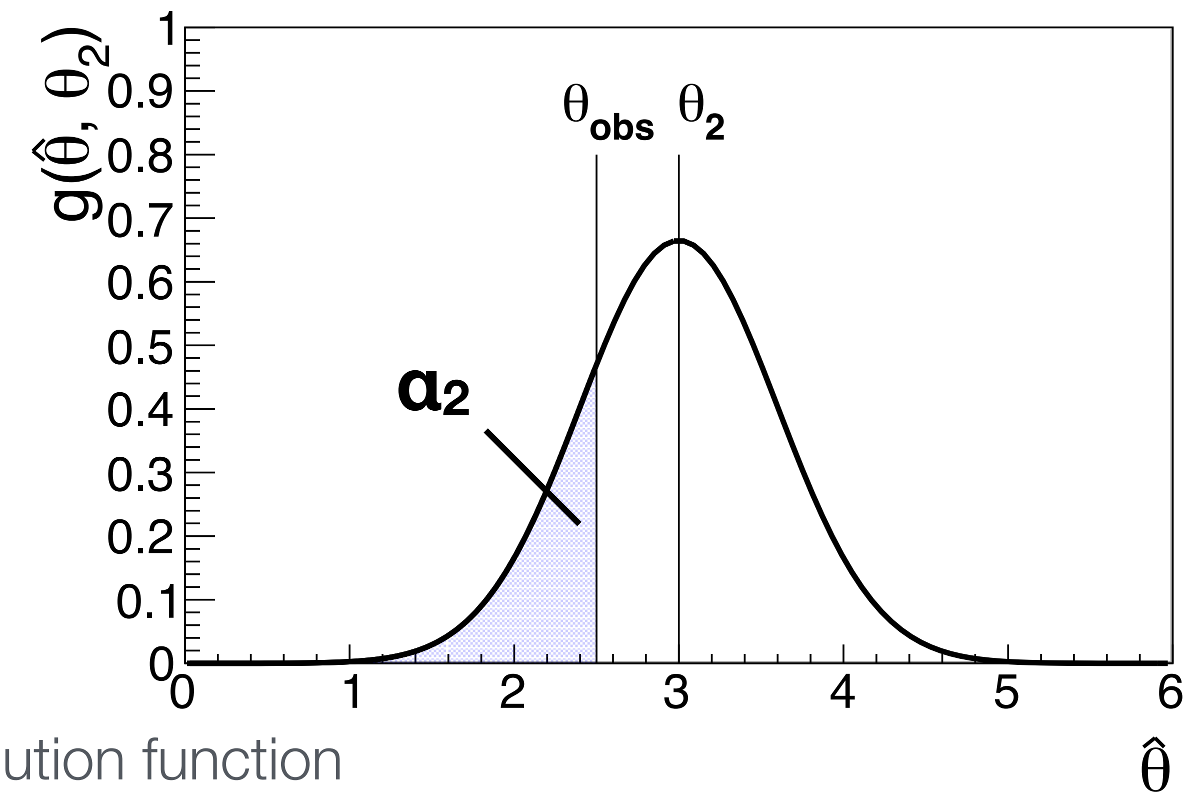
Determine lower bound θ_1 of the confidence interval for θ by solving

$$\alpha_1 = \int_{\hat{\theta}_{obs}}^{\infty} g(\hat{\theta}; \theta_1)\, d\hat{\theta} \equiv 1 - G(\hat{\theta}_{obs}, \theta_1)$$

Analogously for the upper bound θ_2:

$$\alpha_2 = \int_{-\infty}^{\hat{\theta}_{obs}} g(\hat{\theta}; \theta_2)\, d\hat{\theta} \equiv G(\hat{\theta}_{obs}, \theta_2)$$

cumulative distribution function



Confidence Interval for a Gaussian Distributed Estimator

With the aid of the CDF of the standard Gaussian Φ we can write this as:

$$\alpha_1 = 1 - G(\hat{\theta}_{\text{obs}}, \theta_1) = 1 - \Phi\left(\frac{\hat{\theta}_{\text{obs}} - \theta_1}{\sigma_{\hat{\theta}}}\right)$$

$$\alpha_2 = G(\hat{\theta}_{\text{obs}}, \theta_2) = \Phi\left(\frac{\theta_{\text{obs}} - \theta_2}{\sigma_{\hat{\theta}}}\right)$$

This gives:

$$\theta_1 = \theta_{\text{obs}} - \sigma_{\hat{\theta}} \Phi^{-1}(1 - \alpha_1)$$

$$\theta_2 = \theta_{\text{obs}} + \sigma_{\hat{\theta}} \Phi^{-1}(1 - \alpha_2)$$

$-\Phi^{-1}(y) = \Phi^{-1}(1 - y)$

Here Φ^{-1} is the inverse function of Φ, i.e., the quantile function of the standard Gaussian.

Classical Confidence Intervals for the Mean of the Poisson Distribution (I)

$$f(n;\nu) = \frac{\nu^n}{n!} e^{-\nu}$$

Equations for the confidence interval limits θ_1 and θ_2:

$$\alpha_1 = P(\hat{\nu} \geq \hat{\nu}_{\mathrm{obs}}; \theta_1)$$
$$\alpha_2 = P(\hat{\nu} \leq \hat{\nu}_{\mathrm{obs}}; \theta_2)$$

This gives:

$$\alpha_1 = \sum_{n=n_{\mathrm{obs}}}^{\infty} f(n;\theta_1) = 1 - \sum_{n=0}^{n_{\mathrm{obs}}-1} f(n;\theta_1) = 1 - \sum_{n=0}^{n_{\mathrm{obs}}-1} \frac{\theta_1^n}{n!} e^{-\theta_1}$$
$$\alpha_2 = \sum_{n=0}^{n_{\mathrm{obs}}} f(n;\theta_2) = \sum_{n=0}^{n_{\mathrm{obs}}} \frac{\theta_2^n}{n!} e^{-\theta_2}$$

Classical Confidence Intervals for the Mean of the Poisson Distribution (II)

Using the the following relation between the Poisson distribution and the χ^2 distribution

$$\sum_{n=0}^{n_\text{obs}} \frac{\nu^n}{n!} e^{-\nu} = \int_{2\nu}^{\infty} f_{\chi^2}(z; n_\text{df} = 2(n_\text{obs}+1)) \, \mathrm{d}z$$

$$= 1 - F_{\chi^2}(2\nu; 2(n_\text{obs}+1)))$$

F_{χ^2} : CFD of the χ^2 distribution

we obtain

$$\theta_1 = \frac{1}{2} F_{\chi^2}^{-1}[\alpha_1, 2n_\text{obs}]$$

$$\theta_2 = \frac{1}{2} F_{\chi^2}^{-1}[1-\alpha_2, 2(n_\text{obs}+1)]$$

Classical Confidence Intervals for the Mean of the Poisson Distribution (III)

n_{obs}	lower limit θ_1			upper limit θ_2		
	$\alpha_1 = 0.1$	$\alpha_1 = 0.05$	$\alpha_1 = 0.01$	$\alpha_2 = 0.1$	$\alpha_2 = 0.05$	$\alpha_2 = 0.01$
0	–	–	–	2.30	3.00	4.61
1	0.105	0.051	0.010	3.89	4.74	6.64
2	0.532	0.355	0.149	5.32	6.30	8.41
3	1.10	0.818	0.436	6.68	7.75	10.04
4	1.74	1.37	0.823	7.99	9.15	11.60
5	2.43	1.97	1.28	9.27	10.51	13.11

Classical Gaussian Upper Limits with Physical Limit

Suppose the estimator of a parameter θ follows a Gaussian with known standard deviation σ = 1:

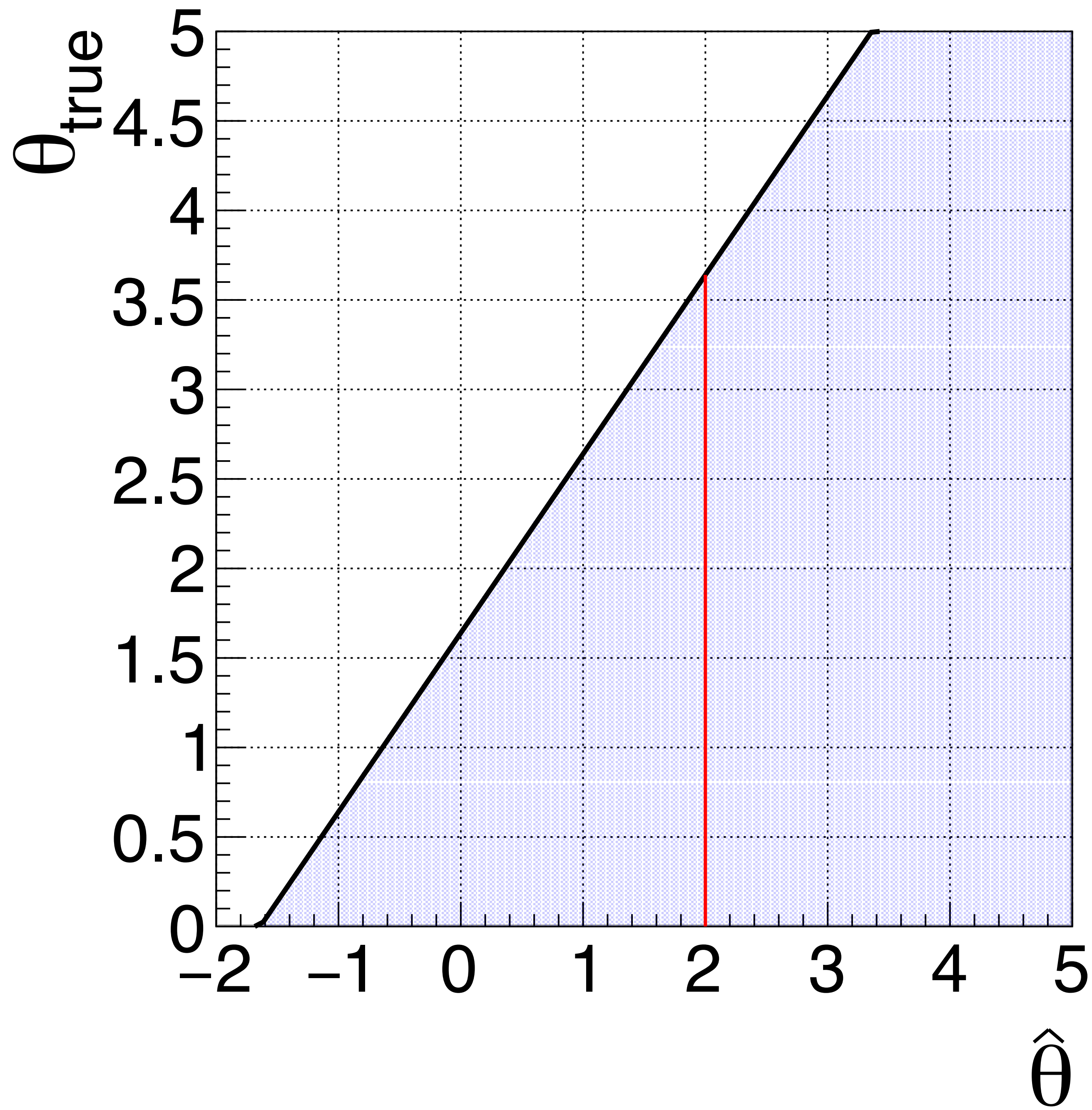
$$g(\hat{\theta};\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-(\hat{\theta}-\theta)^2/2\right)$$

Physically allowed region: $\theta \geq 0$

An example would be the measurement of the neutrino mass: $m \geq 0$

Let's construct the 95% CL upper limit confidence belt (1.64σ)

$$\hat{\theta} = 2 \rightsquigarrow s_{\mathrm{up}} = 3.64 \text{ @ } 95\%\mathrm{CL}$$



But what if we measured –2?: $\hat{\theta} = -2 \rightsquigarrow s_{\mathrm{up}} = -0.36 \text{ @ } 95\%\mathrm{CL}$

A negative upper limit? Has anything gone wrong?

Classical Gaussian Upper Limits with Physical Limit

$$\hat{\theta} = -2 \rightsquigarrow s_{\mathrm{up}} = -0.36 \text{ @ } 95\%\mathrm{CL}$$

We stipulated θ ≥ 0, i.e. the confidence interval is an empty set …

If we measured –1.63 the confidence interval would be [0, 0.01]. Does this really mean that in this case there is a 95% chance that the true value of is between 0 and 0.01?

No, it just means that we have observed a downward fluctuation

- Suppose the true value is zero (θ = 0) → acceptance region @ 95% CL is [–∞, 1.64]
- We expect a negative result in 50% of the cases
- We expect a measurement less than –1.64 in 5% of the cases
- We expect a measurement less than –2 in 2.3% of the cases

Sometimes a negative result is shifted to zero, i.e., 0 + 1.64 σ is reported as upper limit.

That's not helpful. Always report the observed value even if it is in the unphysical regime. Otherwise the result cannot be combined with other results in meta analyses.

Interpretation of Frequentist Confidence Intervals

So has anything gone wrong with the construction of the confidence interval?

Actually no, nothing has gone wrong.

- Even though one should not, there is a tendency to interpret frequentist confidence intervals as Bayesian objects. That is, if one constructs the confidence interval in our example one tends to think that the true value lies in this interval with 95% probability
- But that's not right. We have to think in terms of repeated experiments. The obtained interval covers the true value in 95% of the experiments.
- This does not mean that the interval obtained in a single experiment contains the true value with 95% probability.

The "flip-flop" Problem

Feldman, Cousins, physics/9711021v2

Let us suppose that Physicist X takes the following attitude in an experiment designed to measure a small quantity:

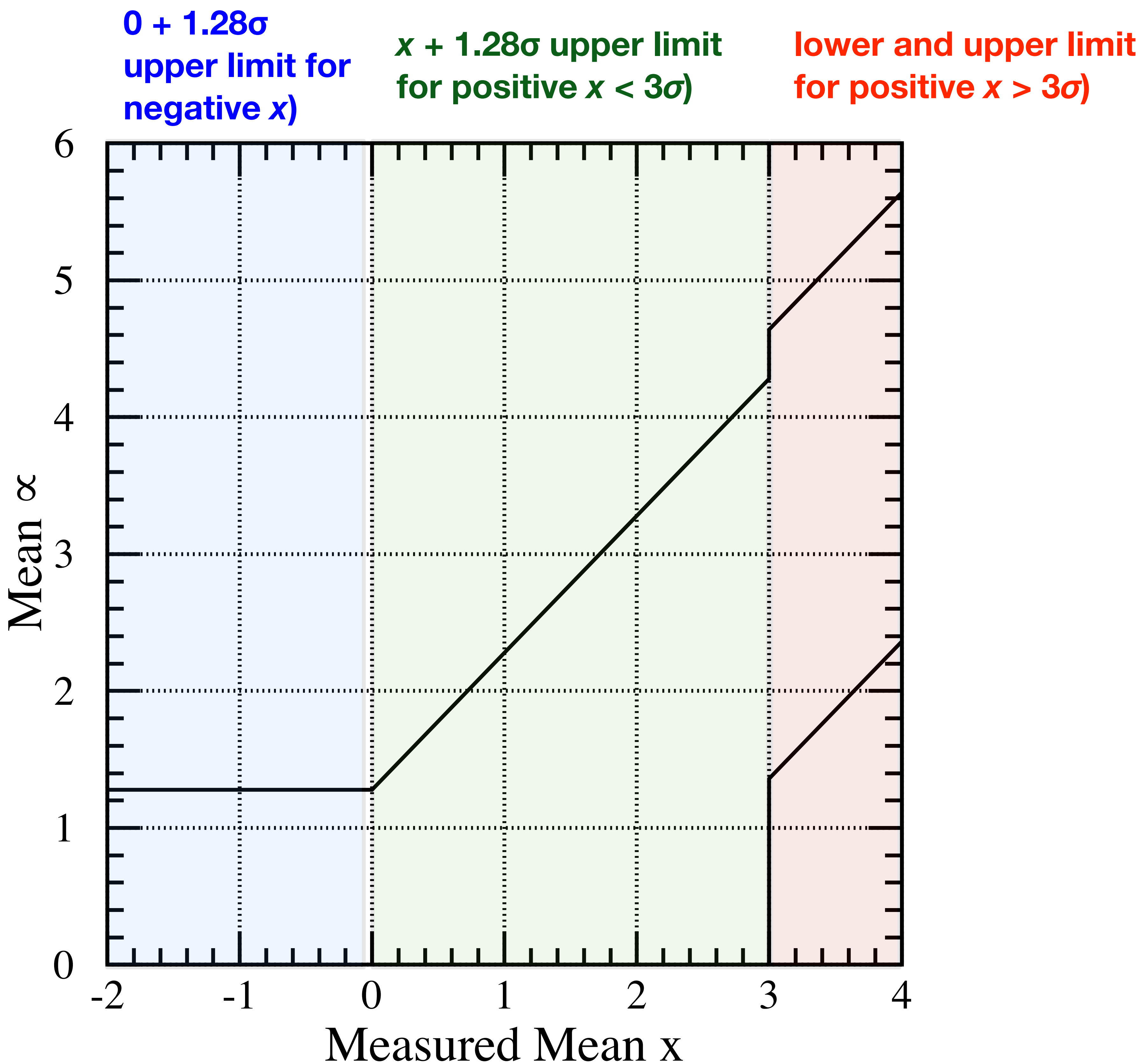
- If the result *x* is less then 3σ, I will state an upper limit
- If the result is greater than 3σ, I will state a central confidence interval from the standard tables

→ So what is reported in this case is decided *after* the measurement

Let's take a look at the confidence band

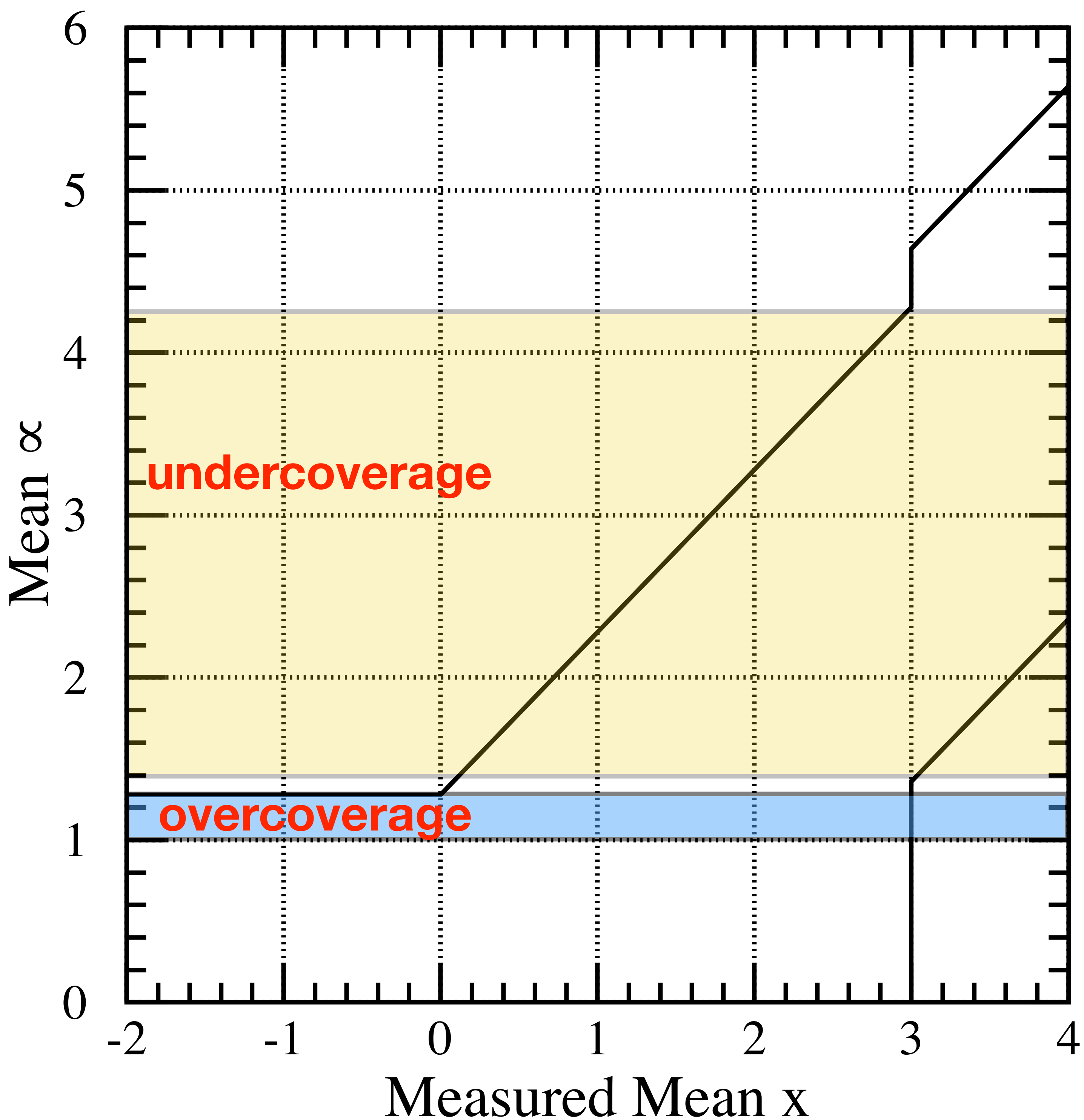
[Variables in the paper by Feldman and Cousins: $x \equiv \hat{\theta},\ \mu \equiv \theta$. Confidence band for 90% CL. Otherwise same situation: Gaussian sampling distribution with σ = 1 and physical regime μ ≥ 0. In the following we'll use *x* and μ.]

The "flip-flop" Problem: Confidence Band



Feldman, Cousins, physics/9711021v2

The "flip-flop" Problem: Coverage



The coverage of the intervals is wrong

- E.g., for 1.36 < μ < 4.28: chance of finding measured value *x* in acceptance region is only 85%, not the desired 90%
- Small μ: overcoverage

This is a serious problem of the flip-flopping approach

Feldman, Cousins, physics/9711021v2

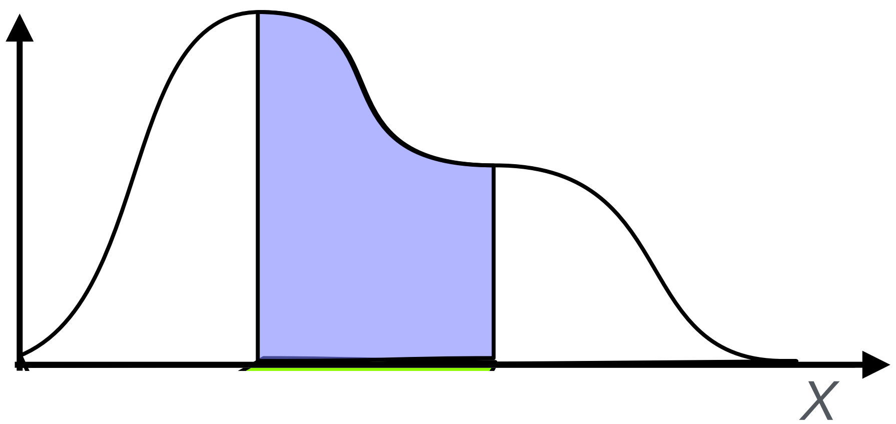
Problems with Classical Confidence Intervals

- in some situations the confidence interval can be an empty set
- they do not elegantly handle unphysical cases
- they do not continuously vary between
  a) giving upper limits in case of a very small signal and
  b) giving upper and lower limits in case of a more significant signal

Feldman & Cousins proposed a solution in their paper
→ Feldman-Cousins confidence intervals

Feldman Cousins Ordering Principle for the Construction of Confidence Intervals

The Neyman construction does not specify how, for a fixed true value μ, to define the interval that covers a fraction 1 – α (e.g. 95%) of the observed outcomes *x*.



Feldman & Cousins introduced an ordering principle based on the likelihood ratio:

$$R = \frac{P(x|\mu)}{P(x|\mu_{\text{best}})}$$

μ_{best} is the best fit obtained from data (maximum likelihood), taking the physically allowed region into account.

Order procedure for fixed μ: add values of *x* to the interval from highest *R* to lower *R* until the desired value 1 – α is reached.

Application of Feldman-Cousins to Gaussian Upper Limits with Physical Limit (I)

Sampling distribution in our example with physical limit μ ≥ 0 ($\sigma_x \equiv 1$):

$$g(x;\mu) = \frac{1}{\sqrt{2\pi}} \exp\left(-(x-\mu)^2/2\right)$$

In this case the best estimate is given by

$$\mu_{\text{best}} = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

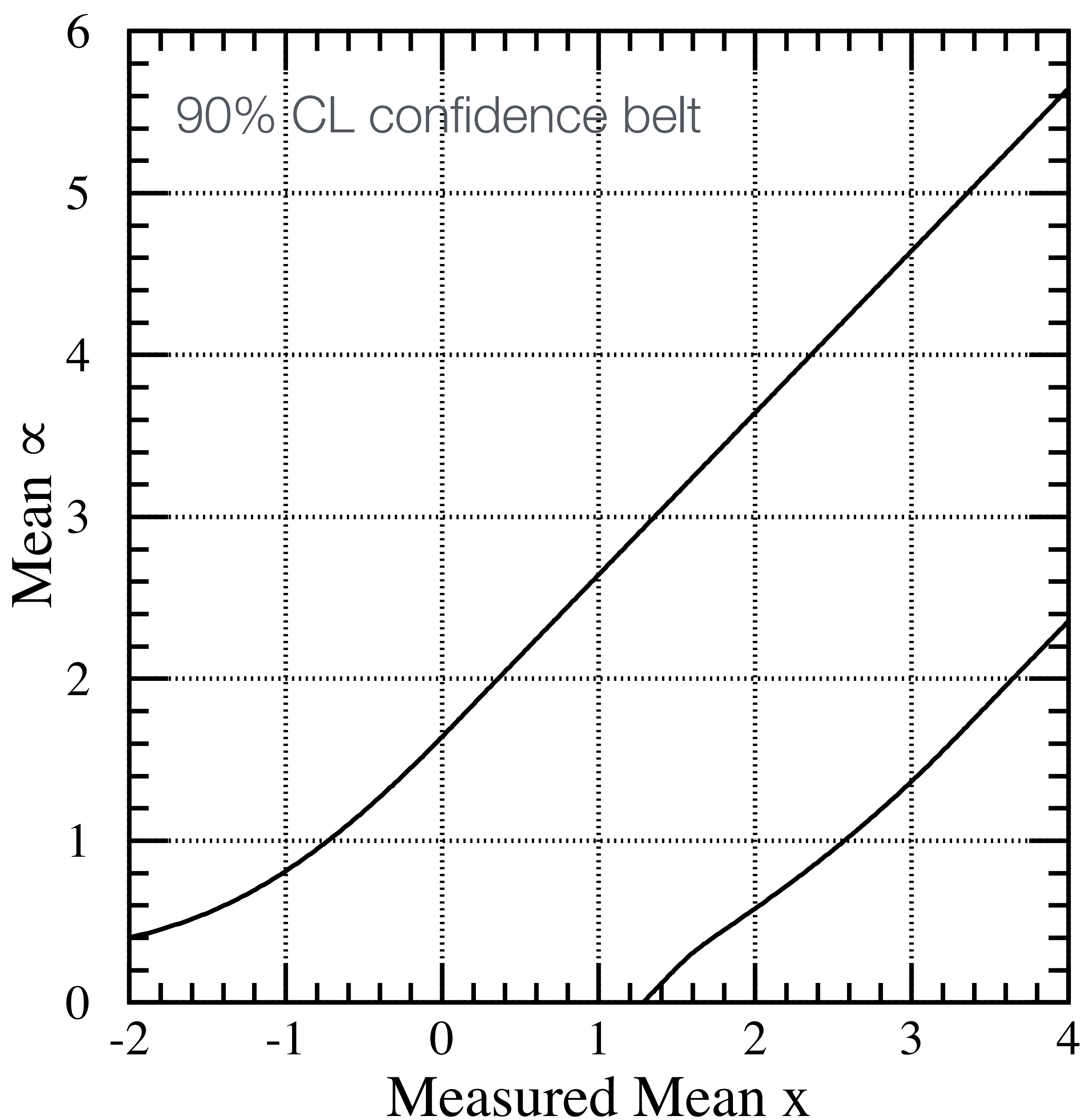
So *R* is given by

$$R = \frac{P(x|\mu)}{P(x|\mu_{\text{best}})} = \begin{cases} \dfrac{\exp\left(-\frac{1}{2}(x-\mu)^2\right)}{\exp\left(-\frac{1}{2}x^2\right)}, & x < 0 \\[2ex] \dfrac{\exp\left(-\frac{1}{2}(x-\mu)^2\right)}{1}, & x \geq 0 \end{cases}$$

In practice, for each μ find interval limits x_1 and x_2 by solving numerically:

$$R(x_1) = R(x_2) \qquad \text{and} \qquad \int_{x_1}^{x_2} g(x|\mu)\,\mathrm{d}x = 1-\alpha$$

Application of Feldman-Cousins to Gaussian Upper Limits with Physical Limit (II)



Some nice features:

- Confidence interval is never empty
- Smooth transition from giving upper limit to two-sided interval
- Tells you when to quote upper limit and when to quote an interval
- Correct coverage

Feldman-Cousins Confidence Intervals for the Mean of the Poisson Distribution (I)

Let's go back to the counting experiment with signal s and known average number of background counts b:

$$P(n|s) = \frac{(s+b)^n}{n!} e^{-(s+b)}$$

Classical method sometimes gives negative upper limit when $n_{\text{obs}} < b$.

This problem is addressed by the Feldman-Cousins method.

The paper contains look-up tables for upper limits and confidence intervals.

Feldman-Cousins Confidence Intervals for the Mean of the Poisson Distribution (II)

TABLE IV. 90% C.L. intervals for the Poisson signal mean μ, for total events observed n_0, for known mean background b ranging from 0 to 5.

$n_0 \backslash b$	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	5.0
0	0.00, 2.44	0.00, 1.94	0.00, 1.61	0.00, 1.33	0.00, 1.26	0.00, 1.18	0.00, 1.08	0.00, 1.06	0.00, 1.01	*0.00, 0.98*
1	0.11, 4.36	0.00, 3.86	0.00, 3.36	0.00, 2.91	0.00, 2.53	0.00, 2.19	0.00, 1.88	0.00, 1.59	0.00, 1.39	0.00, 1.22
2	0.53, 5.91	0.03, 5.41	0.00, 4.91	0.00, 4.41	0.00, 3.91	0.00, 3.45	0.00, 3.04	0.00, 2.67	0.00, 2.33	0.00, 1.73
3	1.10, 7.42	0.60, 6.92	0.10, 6.42	0.00, 5.92	0.00, 5.42	0.00, 4.92	0.00, 4.42	0.00, 3.95	0.00, 3.53	0.00, 2.78
4	1.47, 8.60	1.17, 8.10	0.74, 7.60	0.24, 7.10	0.00, 6.60	0.00, 6.10	0.00, 5.60	0.00, 5.10	0.00, 4.60	0.00, 3.60
5	1.84, 9.99	1.53, 9.49	1.25, 8.99	0.93, 8.49	0.43, 7.99	0.00, 7.49	0.00, 6.99	0.00, 6.49	0.00, 5.99	0.00, 4.99
6	2.21,11.47	1.90,10.97	1.61,10.47	1.33, 9.97	1.08, 9.47	0.65, 8.97	0.15, 8.47	0.00, 7.97	0.00, 7.47	0.00, 6.47
7	3.56,12.53	3.06,12.03	2.56,11.53	2.09,11.03	1.59,10.53	1.18,10.03	0.89, 9.53	0.39, 9.03	0.00, 8.53	0.00, 7.53
8	3.96,13.99	3.46,13.49	2.96,12.99	2.51,12.49	2.14,11.99	1.81,11.49	1.51,10.99	1.06,10.49	0.66, 9.99	0.00, 8.99
9	4.36,15.30	3.86,14.80	3.36,14.30	2.91,13.80	2.53,13.30	2.19,12.80	1.88,12.30	1.59,11.80	1.33,11.30	0.43,10.30
10	5.50,16.50	5.00,16.00	4.50,15.50	4.00,15.00	3.50,14.50	3.04,14.00	2.63,13.50	2.27,13.00	1.94,12.50	1.19,11.50
11	5.91,17.81	5.41,17.31	4.91,16.81	4.41,16.31	3.91,15.81	3.45,15.31	3.04,14.81	2.67,14.31	2.33,13.81	1.73,12.81
12	7.01,19.00	6.51,18.50	6.01,18.00	5.51,17.50	5.01,17.00	4.51,16.50	4.01,16.00	3.54,15.50	3.12,15.00	2.38,14.00
13	7.42,20.05	6.92,19.55	6.42,19.05	5.92,18.55	5.42,18.05	4.92,17.55	4.42,17.05	3.95,16.55	3.53,16.05	2.78,15.05
14	8.50,21.50	8.00,21.00	7.50,20.50	7.00,20.00	6.50,19.50	6.00,19.00	5.50,18.50	5.00,18.00	4.50,17.50	3.59,16.50
15	9.48,22.52	8.98,22.02	8.48,21.52	7.98,21.02	7.48,20.52	6.98,20.02	6.48,19.52	5.98,19.02	5.48,18.52	4.48,17.52
16	9.99,23.99	9.49,23.49	8.99,22.99	8.49,22.49	7.99,21.99	7.49,21.49	6.99,20.99	6.49,20.49	5.99,19.99	4.99,18.99
17	11.04,25.02	10.54,24.52	10.04,24.02	9.54,23.52	9.04,23.02	8.54,22.52	8.04,22.02	7.54,21.52	7.04,21.02	6.04,20.02
18	11.47,26.16	10.97,25.66	10.47,25.16	9.97,24.66	9.47,24.16	8.97,23.66	8.47,23.16	7.97,22.66	7.47,22.16	6.47,21.16
19	12.51,27.51	12.01,27.01	11.51,26.51	11.01,26.01	10.51,25.51	10.01,25.01	9.51,24.51	9.01,24.01	8.51,23.51	7.51,22.51
20	13.55,28.52	13.05,28.02	12.55,27.52	12.05,27.02	11.55,26.52	11.05,26.02	10.55,25.52	10.05,25.02	9.55,24.52	8.55,23.52

Feldman, Cousins, physics/9711021v2

Feldman-Cousins Method: Discussion

Nice features:

+ State-of-the art for frequentist confidence intervals
+ Avoids flip-flop problem, correct coverage
+ Handles interval estimates at physical boundaries

Drawbacks:

– Construction of F-C confidence intervals is complicated, usually has to be done numerically

– Systematic uncertainties not easily included

– Counter-intuitive result in case of counting experiments with different background (see next slide)

Feldman-Cousins Method: The Paradox of Fewer than Expected Background Events

Consider two counting experiments

- Experiment A: expects background $b = 0$ ("carefully designed experiment")
- Experiment B: expects background $b = 5$

Suppose now both experiments measure $n = 0$ counts.
Feldman-Cousins upper limits at 90% CL:

- Experiment A: $s_{\mathrm{up}} = 2.44$
- Experiment B: $s_{\mathrm{up}} = 0.98$

Weird: The FC method says that the experiment B in which a larger background is expected gives the better (more stringent) upper limit.

Experiment B must have observed a downward fluctuation of the background. How can a fluctuation result in a better upper limit?

Suggestion in the Feldman-Cousins Paper

"Our suggestion for doing this is that in cases in which the measurement is less than the estimated background, the experiment report both the upper limit and the “sensitivity” of the experiment, where the “sensitivity” is defined as the average upper limit that would be obtained by an ensemble of experiments with the expected background and no true signal. […]

Thus, an experiment that measures 2 events and has an expected background of 3.5 events would report a 90% C.L. upper limit of 2.7 events (from Tab. IV), but a sensitivity of 4.6 events (from Tab. XII)."

Feldman, Cousins, physics/9711021v2

CL_S Method: Motivation

Consider an experiment with low sensitivity ("background dominated experiment").

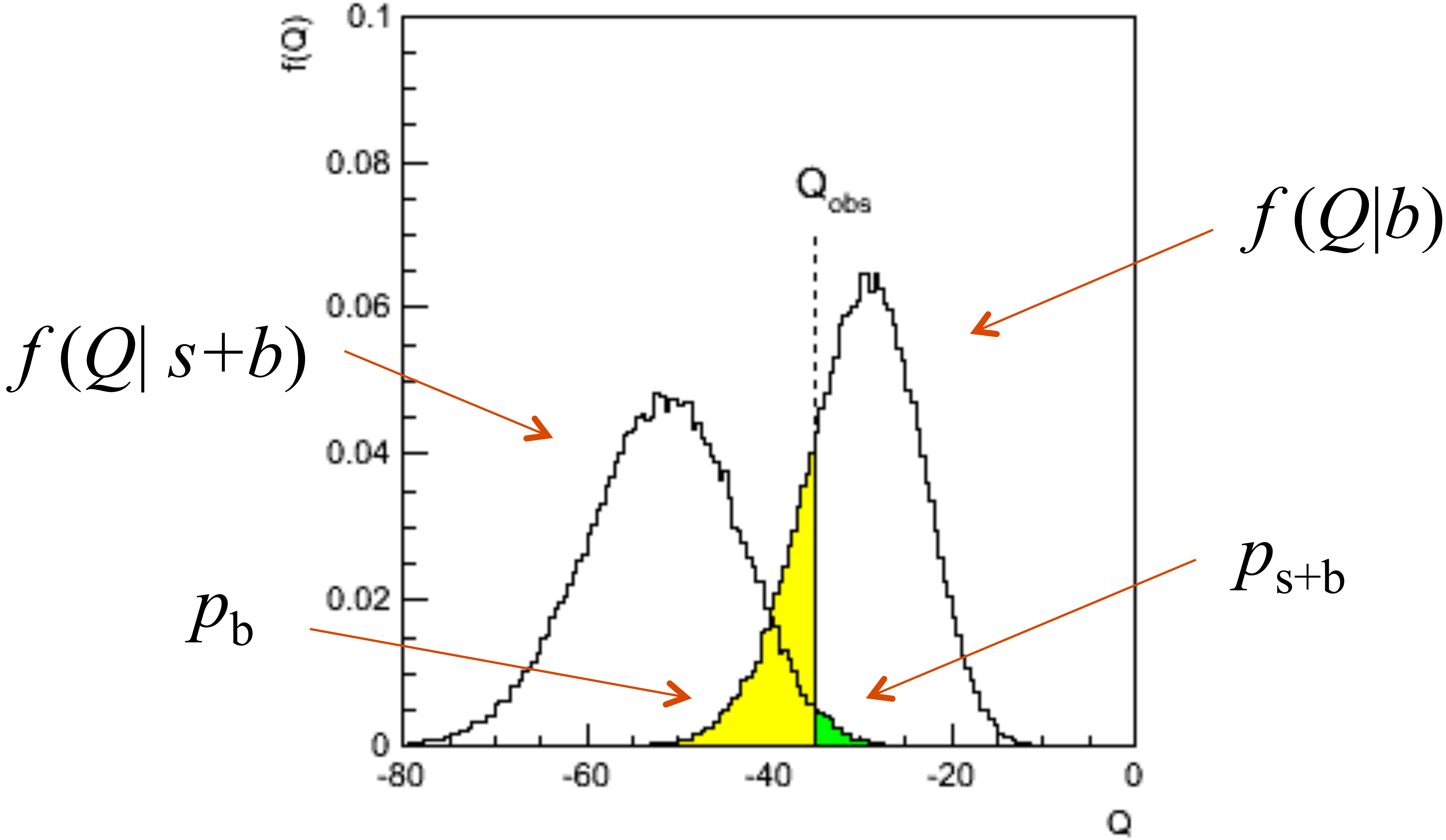
- By construction, one rejects a true null hypothesis with a certain probability (e.g. 5%)
- Problem: exclusion of parameter values to which one has no sensitivity
- Example Higgs search: m_H = 1000 TeV rejected with a chance of 5%
- "Spurious exclusion"

This problem was addressed for the LEP Higgs searches in the late 1990'ies and led to the CL_S method A. Read, J. Phys. G 28, 2693 (2002), T. Junk, NIM A, 434, 435 (1999)

- Explicitly consider experimental sensitivity in limit setting
- Reduce spurious exclusion by a particular choice of the critical region
- Frequentist-motivated approach, but NOT frequentist ("modified frequentist method")
- Name a bit misleading, as the CLs exclusion region is not a confidence interval
- Overcoverage by construction: conscious choice to give up frequentist coverage to take sensitivity into account
- "Despite its shaky foundations in statistical theory, it has been producing sensible results for over a decade" (http://cds.cern.ch/record/2203243)

CLs Procedure (I)

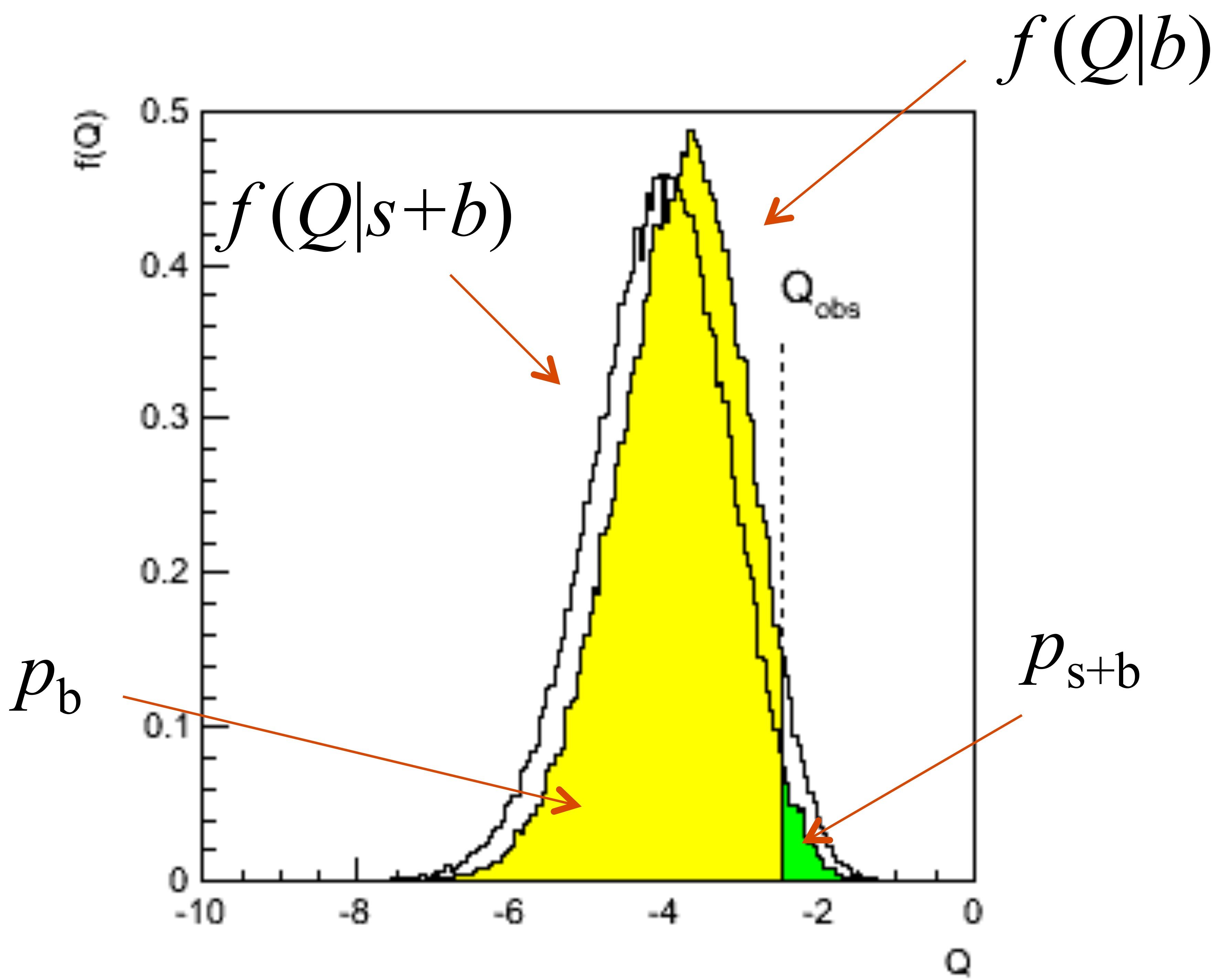
Test statistic: $Q = -2 \ln \frac{L(x|s+b)}{L(x|b)}$



CLs Procedure (II)

https://www.pp.rhul.ac.uk/~cowan/stat_course.html

Low sensitivity: the distributions under s and s+b are very close



CL_s Procedure (III)

https://www.pp.rhul.ac.uk/~cowan/stat_course.html

Standard p-value test:

Reject s+b hypothesis if

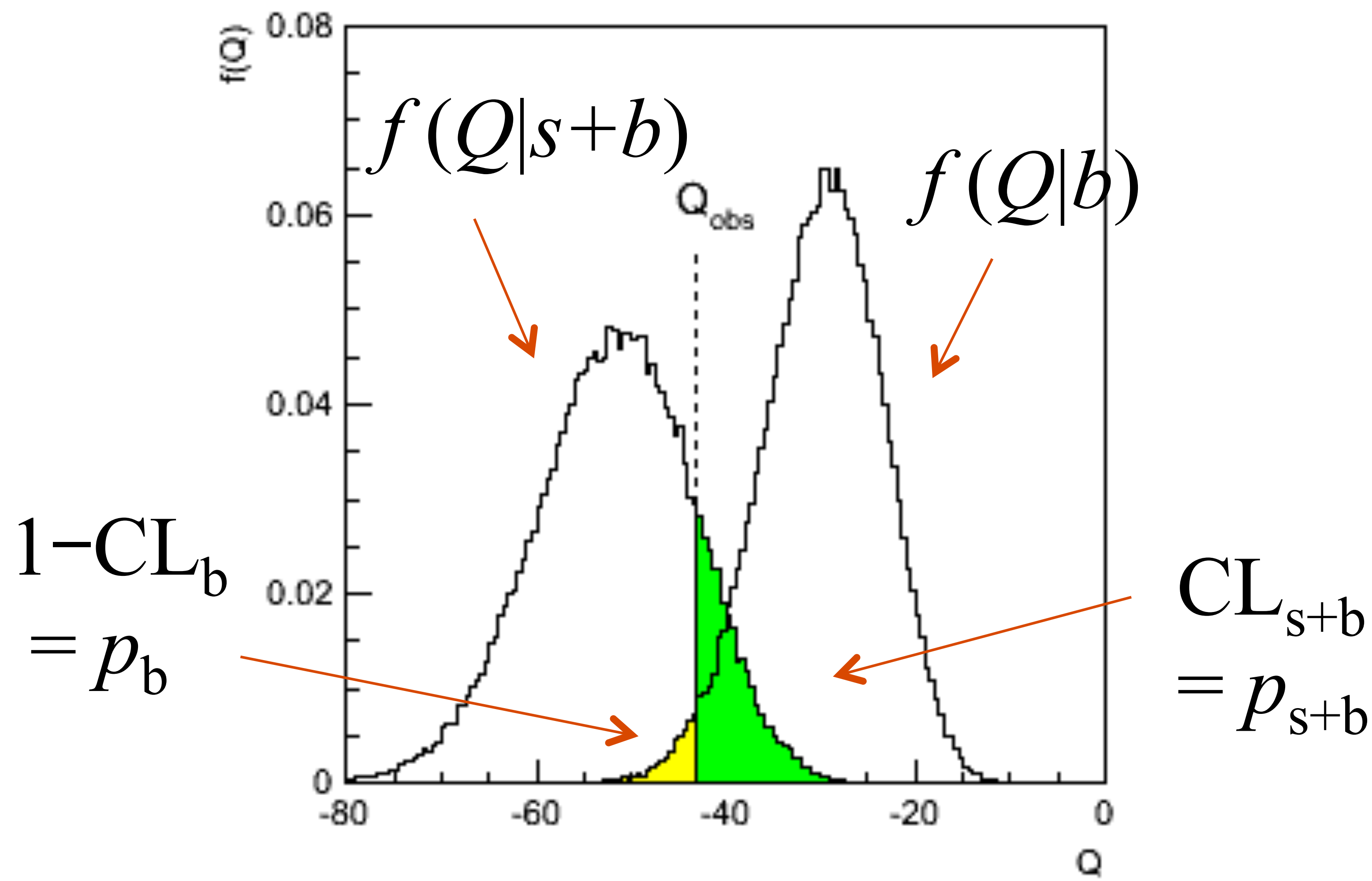
$$p_{s+b} \leq \alpha$$

CL_s method:

Reject s+b hypothesis if

$$CL_s := \frac{p_{s+b}}{1 - p_b} \leq \alpha$$

more stringent than standard p-value test as $1 - p_b \leq 1$



Increases “effective” p-value when the two distributions become close (prevents exclusion if sensitivity is low)

Upper Limits on $\mu = \sigma/\sigma_{SM}$ in Higgs Searches

Signal for Higgs hypothesis: $s(m_H) = L_{\text{int}} \cdot \sigma_{\text{SM}}$

Signal strength μ: $n = \mu \cdot s(m_H) + b, \quad \mu = \frac{L_{\text{int}} \cdot \sigma(m_H)}{L_{\text{int}} \cdot \sigma_{\text{SM}}(m_H)} = \frac{\sigma(m_H)}{\sigma_{\text{SM}}(m_H)}$

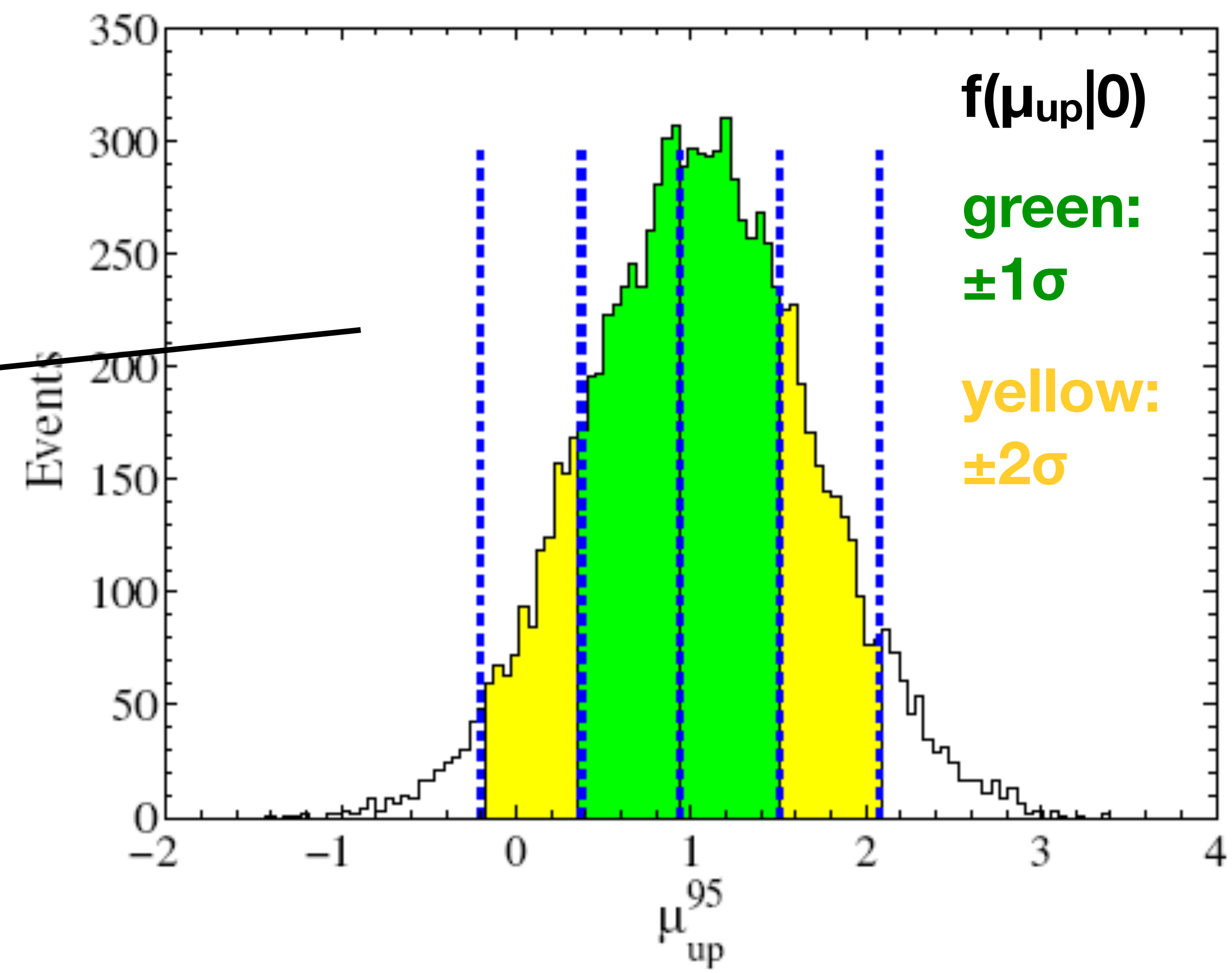
μ = 1: SM w/ Higgs, μ = 0: SM w/o Higgs

Carry out CL_s procedure for $\mu = \sigma/\sigma_{SM}$, i.e., reject μ if $\text{CL}_s := \frac{p_\mu}{1 - p_b} \leq 0.05$

→ upper limit μ_{up} at 95% CL

At a given value of m_H, we have an observed value of μ_{up}, and we can also find the distribution $f(\mu_{up}|0)$

background only hypothesis

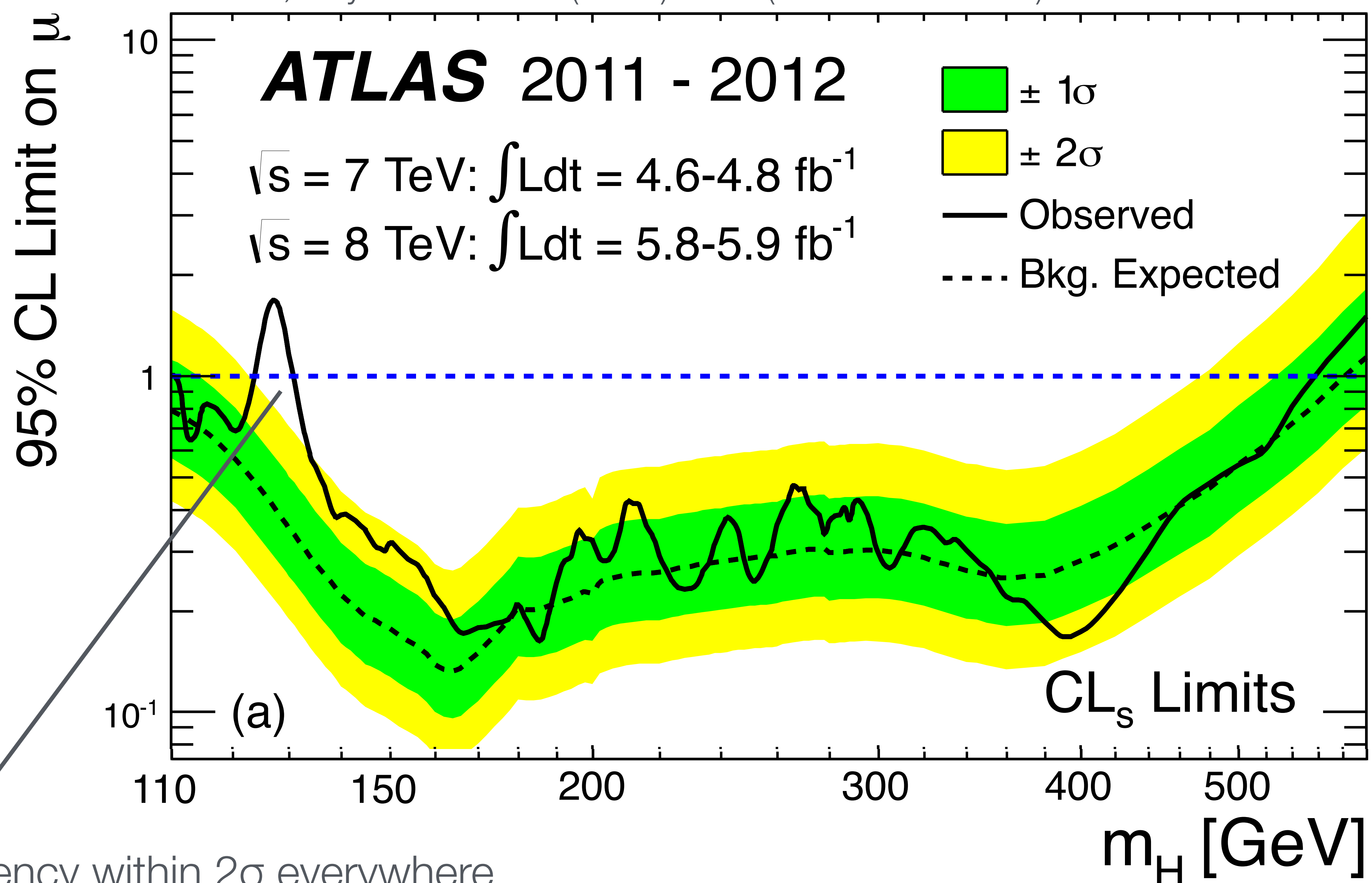


https://www.pp.rhul.ac.uk/~cowan/stat_course.htm

Upper Limits on $\mu = \sigma/\sigma_{SM}$ in Higgs Searches

95% CL limit on $\mu < 1 \Rightarrow$ Standard model with m_H rejected

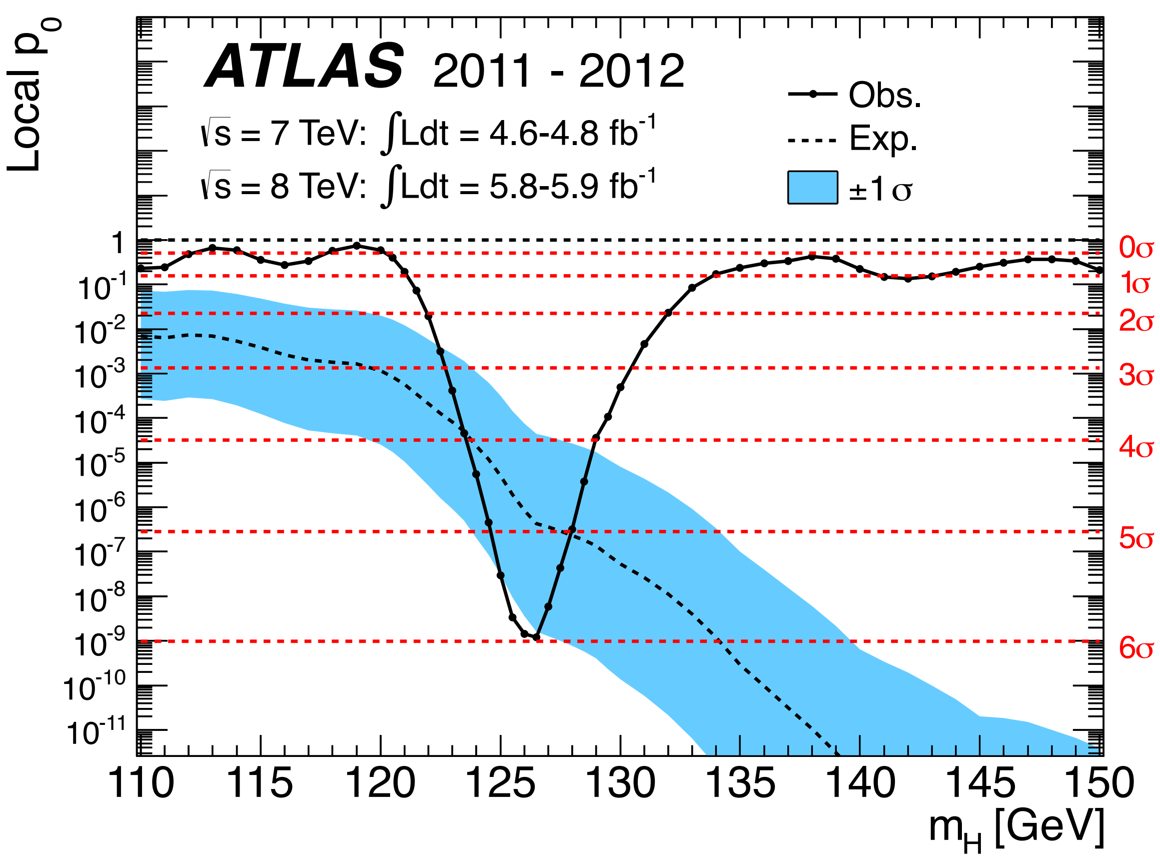
ATLAS, Phys.Lett. B716 (2012) 1-29 (arXiv:1207.7214)



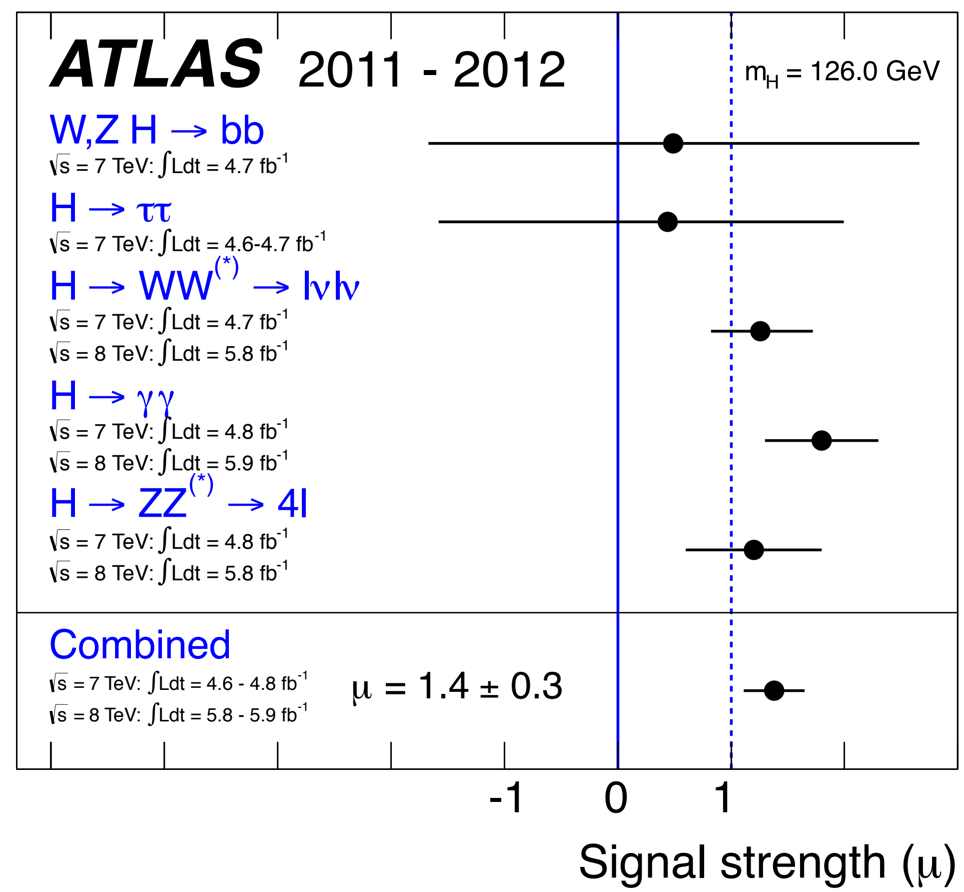
Consistency within 2σ everywhere except for m_H = 125 GeV

Higgs Discovery (this time from ATLAS Paper)

local p-value



signal strength



ATLAS, Phys.Lett. B716 (2012) 1-29 (arXiv:1207.7214)

Higgs Discovery

CERN Seminar on 4. July 2012

Higgs Discovery

“I think we have it!”

(Rolf Heuer,
CERN director general in 2012