Proposed exercise

The values of the parameter $\mu=\sigma/\sigma_{SM}$ for the Higgs boson for the three main decay channels measured in 2014 by ATLAS were:

$$\mu_{\gamma\gamma} = 1.55 \pm 0.30$$

$$\mu_{ZZ} = 1.43 \pm 0.37$$

$$\mu_{WW} = 0.99 \pm 0.29$$

Evaluate the compatibility among the three independent ATLAS results and calculate the best overall estimate of μ from ATLAS. Then evaluate the compatibility with the SM expectation (μ=1).

Proposed exercise

Consider the Higgs production (M_H =125 GeV) at a pp collider at $\sqrt{s}$ = 14 TeV. Evaluate the interval in rapidity y and the minimum value of x for direct Higgs production.

Bayesian vs frequentist intervals (revisited)

Bayesian intervals

posterior → $$p(\theta_{true}/x_0) = \frac{L(x_0/\theta_{true})\pi(\theta_{true})}{\int d\theta_{true} L(x_0/\theta_{true})\pi(\theta_{true})}$$ ← prior

Bayesian interval $$\int_{\theta_1}^{\theta_2} p(\theta_{true}/x_0) d\theta_{true} = \beta$$

The interval $[\theta_1, \theta_2]$ is called **credible interval**.

The edges θ_1 , θ_2 of the Bayesian intervals are not uniquely defined

$$\int_{\theta_1}^{\theta_2} p(\theta_{true}/x_0)d\theta_{true} = \beta$$

Central intervals: the pdf integral is the same above and below the interval:

$$\int_{-\infty}^{\theta_1} p(\theta_{true}/x_0)d\theta_{true} = \frac{1-\beta}{2}$$

$$\int_{\theta_2}^{+\infty} p(\theta_{true}/x_0)d\theta_{true} = \frac{1-\beta}{2}$$

Upper limits: θ_{true} is below a certain value. In this case the interval is between 0 (if θ is a non-negative quantity) and θ_{up}:

$$\int_{0}^{\theta_{up}} p(\theta_{true}/x_0)d\theta_{true} = \beta$$

Lower limits: θ_{true} is above a certain value θ_{low}:

$$\int_{\theta_{low}}^{+\infty} p(\theta_{true}/x_0)d\theta_{true} = \beta$$

Comments:

Bayes:
- Non informative prior (does it exist?)
- Recursive Bayes estimation => Bayes filter

$$\text{posterior} \propto \text{prior} \times \text{likelihood}$$

$$\downarrow$$

$$\textit{revised} \propto \textit{current} \times \textit{new likelihood}$$

$$\pi_{n+1}(\theta) \propto \pi_n(\theta) \times L_{n+1}(\theta) = \pi_n(\theta) f(x_{n+1} \mid \mathbf{x_n}, \theta).$$

In this dynamic perspective we notice that at time n we only need to keep a representation of π_n and otherwise can ignore the past.

The current π_n contains all information needed to revise knowledge when confronted with new information $L_{n+1}(\theta)$.

We sometimes refer to this way of updating as *recursive*.

Comments:

Bayes:
- Non informative prior (does it exist?)
- Recursive Bayes estimation => Bayes filter

Applications

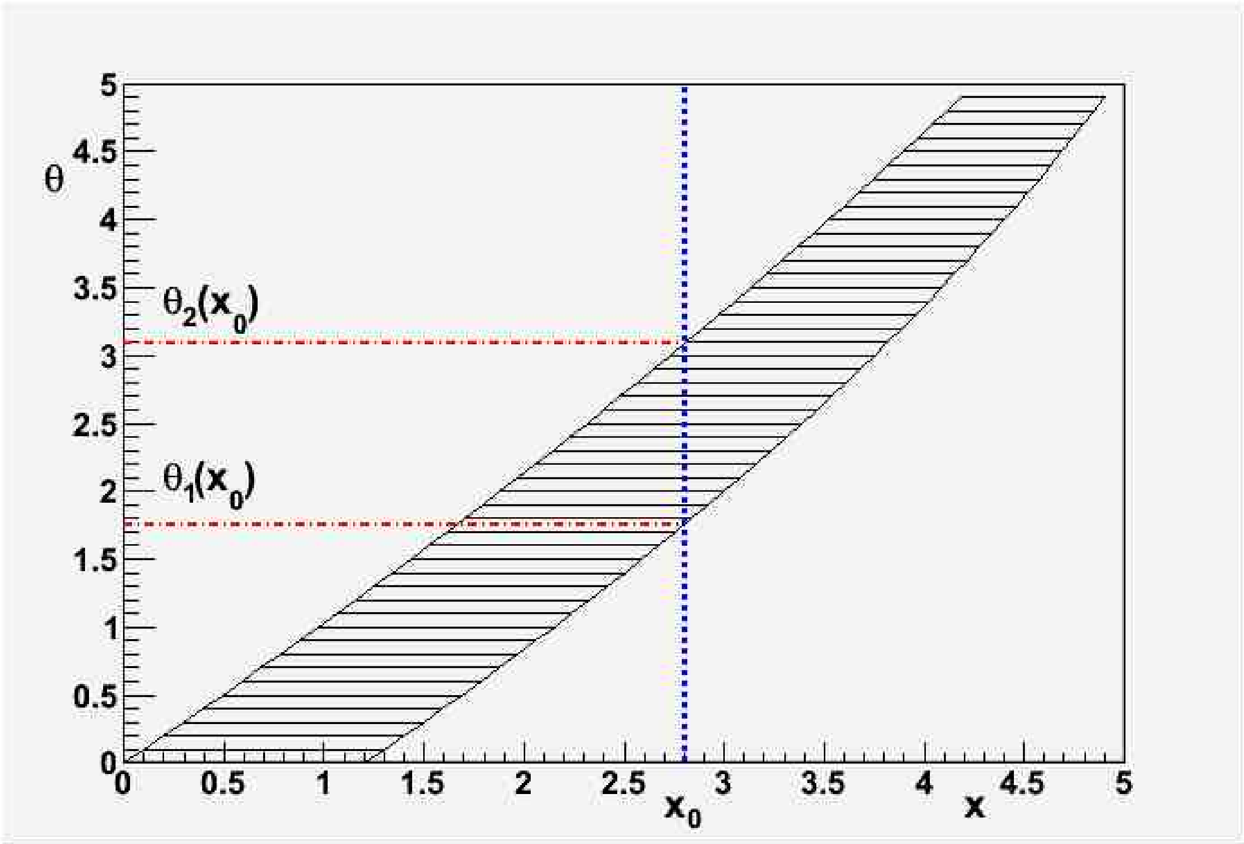
- Ballistics
- Robotics
  - Robot localization
- Tracking hands/cars/…
- Econometrics
  - Stock prediction
- Navigation

- Many more…

Frequentist intervals

Neynman construction of the confidence intervals

$$\int_{x_1(\theta)}^{x_2(\theta)} L(x/\theta)dx = \beta$$



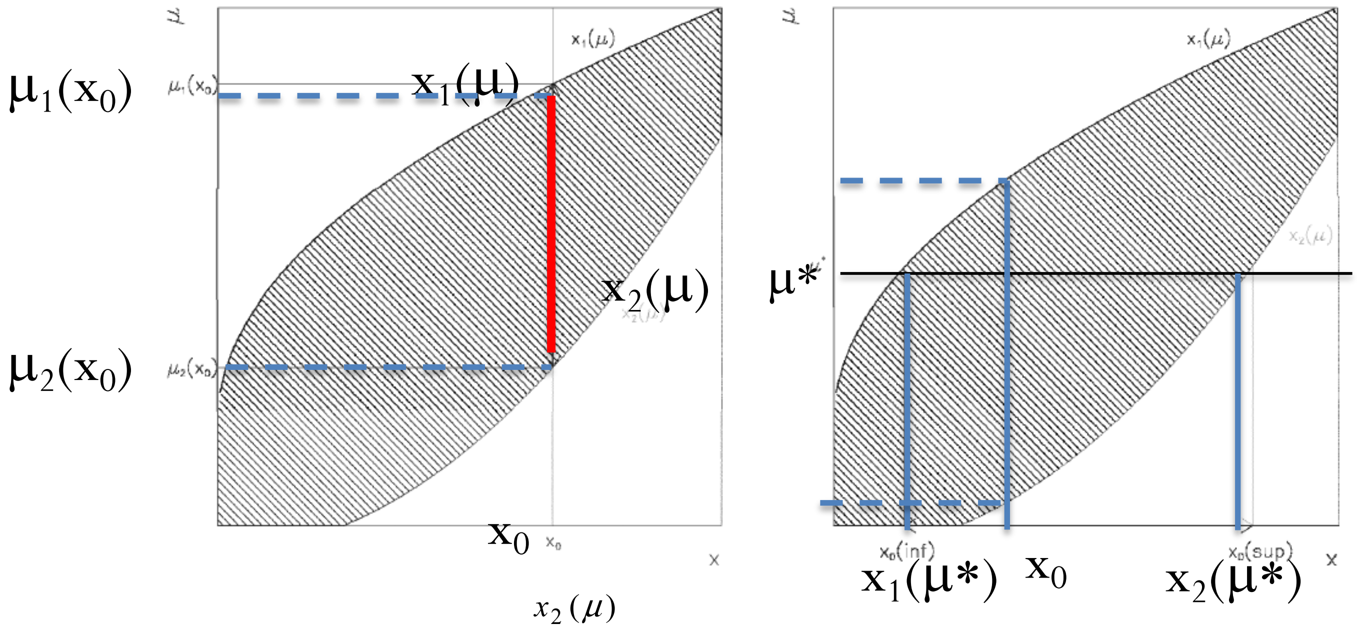
Coverage: $$p(\theta_1(x_0) < \theta_{true} < \theta_2(x_0)) = \beta$$

Neyman's construction

Pay attention, in the following β=1-α:

$$\int_{x_1(\mu)}^{x_2(\mu)} f(x/\mu)dx = 1\text{-}\alpha$$

Neyman's construction



$$\int_{x_1(\mu)}^{x_2(\mu)} f(x/\mu)dx = 1\text{-}\alpha$$

By construction the probability to measure $x_0' < x_0$ if the true value $\mu=\mu_1(x_0)$ is $\alpha/2$
$x_0' > x_0$ if the true value $\mu=\mu_2(x_0)$ is $\alpha/2$

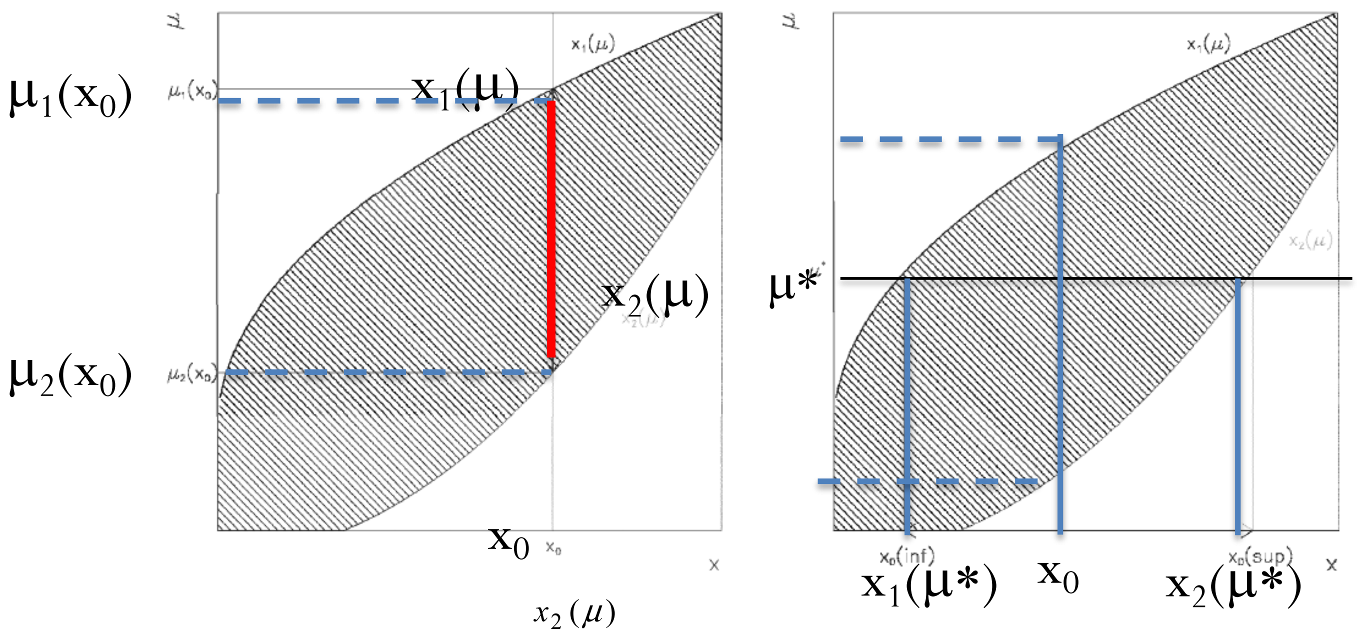
The determined C.I. is $[\mu_2(x_0), \mu_1(x_0)]$.

Check the correct coverage: suppose μ^* is the true value. I repeat N times the measurement and determine each time the C.I. By construction in a fraction 1-α of the cases x_0 is within $x_1(\mu^*)$ and $x_2(\mu^*)$ and the corresponding C.I. provides coverage of μ^*.

$$P(x_1(\mu^*) < x_0 < x_2(\mu^*)) = 1\text{-}\alpha$$

In α cases x_0 lies outside the interval $[x_1(\mu^*), x_2(\mu^*)]$ and the corresponding C.I. does not cover μ^*.

Neyman's construction



$$\int_{x_1(\mu)}^{x_2(\mu)} f(x/\mu)dx = 1\text{-}\alpha$$

By construction the probability to measure $x_0' < x_0$ if the true value $\mu = \mu_1(x_0)$ is $\alpha/2$
$x_0' > x_0$ if the true value $\mu = \mu_2(x_0)$ is $\alpha/2$

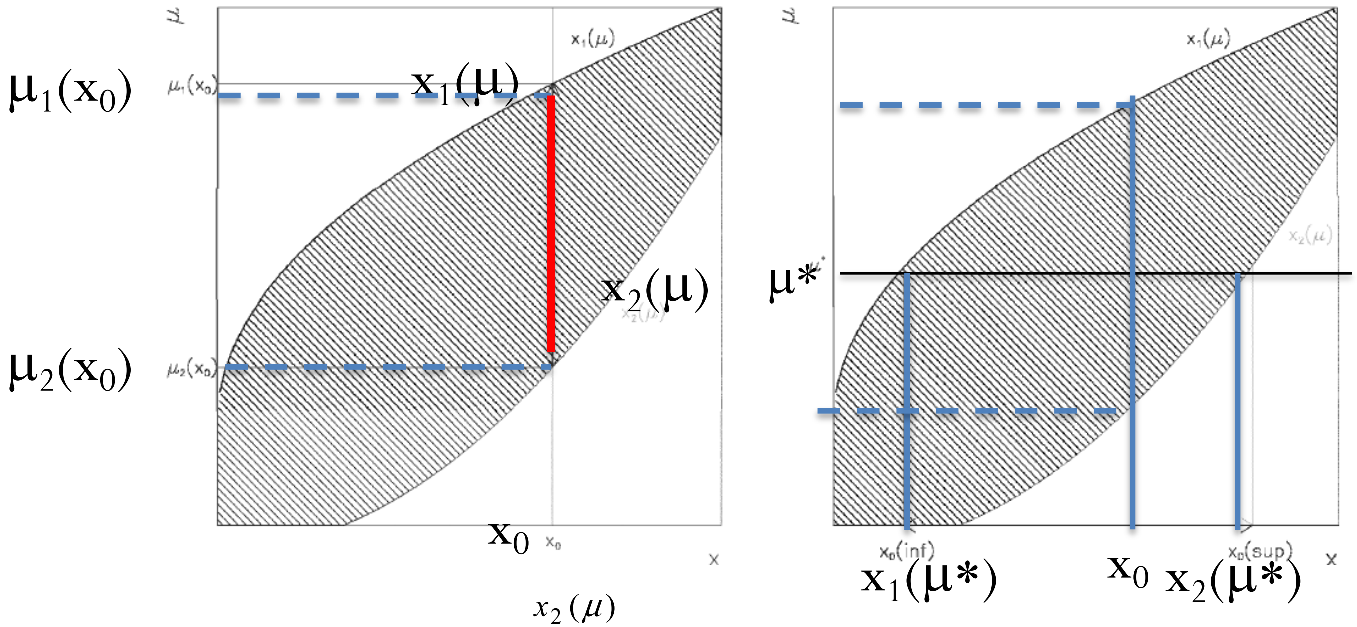
The determined C.I. is $[\mu_2(x_0), \mu_1(x_0)]$.

Check the correct coverage: suppose μ^* is the true value. I repeat N times the measurement and determine each time the C.I. By construction in a fraction 1-α of the cases x_0 is within $x_1(\mu^*)$ and $x_2(\mu^*)$ and the corresponding C.I. provides coverage of μ^*.

$$P(x_1(\mu^*) < x_0 < x_2(\mu^*)) = 1\text{-}\alpha$$

In α cases x_0 lies outside the interval $[x_1(\mu^*), x_2(\mu^*)]$ and the corresponding C.I. does not cover μ^*.

Neyman's construction



$$\int_{x_1(\mu)}^{x_2(\mu)} f(x/\mu)dx = 1\text{-}\alpha$$

By construction the probability to measure $x_0' < x_0$ if the true value $\mu = \mu_1(x_0)$ is $\alpha/2$

$x_0' > x_0$ if the true value $\mu = \mu_2(x_0)$ is $\alpha/2$

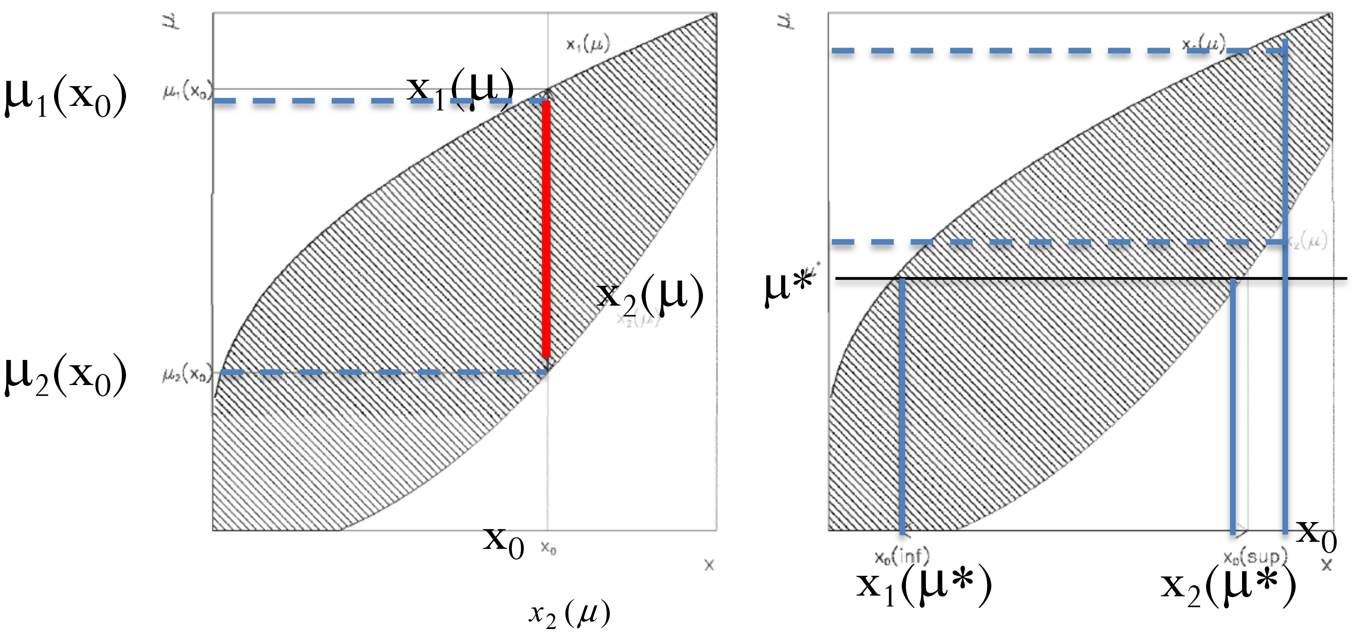
The determined C.I. is $[\mu_2(x_0), \mu_1(x_0)]$.

Check the correct coverage: suppose μ^* is the true value. I repeat N times the measurement and determine each time the C.I. By construction in a fraction 1-α of the cases x_0 is within $x_1(\mu^*)$ and $x_2(\mu^*)$ and the corresponding C.I. provides coverage of μ^*.

$$P(x_1(\mu^*) < x_0 < x_2(\mu^*)) = 1\text{-}\alpha$$

In α cases x_0 lies outside the interval $[x_1(\mu^*), x_2(\mu^*)]$ and the corresponding C.I. does not cover μ^*.

Neyman's construction



$$\int_{x_1(\mu)}^{x_2(\mu)} f(x/\mu)dx = 1\text{-}\alpha$$

By construction the probability to measure $x_0' < x_0$ if the true value $\mu = \mu_1(x_0)$ is $\alpha/2$
$x_0' > x_0$ if the true value $\mu = \mu_2(x_0)$ is $\alpha/2$

The determined C.I. is $[\mu_2(x_0), \mu_1(x_0)]$.

Check the correct coverage: suppose μ^* is the true value. I repeat N times the measurement and determine each time the C.I. By construction in a fraction 1-α of the cases x_0 is within $x_1(\mu^*)$ and $x_2(\mu^*)$ and the corresponding C.I. provides coverage of μ^*.

$$P(x_1(\mu^*) < x_0 < x_2(\mu^*)) = 1\text{-}\alpha$$

In α cases x_0 lies outside the interval $[x_1(\mu^*), x_2(\mu^*)]$ and the corresponding C.I. does not cover μ^*.

Confidence Interval & Coverage

- You claim, $CI_\mu=[\mu_1,\mu_2]$ at the 95% C.L
  i.e. In an ensemble of experiments C.L (95%) of the obtained confidence intervals will contain the true value of μ.

  - If your statement is accurate, you have full coverage
  - If the true C.L is>95%, your interval has an over coverage
  - If the true C.L is <95%, your interval has an undercoverage

Signal searches: upper and lower limits

(consider the simple example of counting experiment)

- **Discovery**: the Null Hypothesis H_0, based on the Standard Model is falsified by a goodness-of-fit test. This means that new physics should be included to account for the data. This is an important discovery.
- **Exclusion**: the Alternative Hypothesis H_1, based on an extension of the Standard Model (or on a new theory at all), doesn't pass the goodness-of-fit test. H_1 is excluded by data.

Exclusion means that the search has given a negative result. However a negative result is not a complete failure of the experiment, but it gives important informations that have to be expressed in a quantitative way so that theorists or other experimentalists can use them for further searches. These quantitative statements about negative results of a search for new phenomena are normally the "upper limits" or the "lower limits".

By **upper limit** we mean a statement like the following: such a particle, if it exists, is produced with a rate (or cross-section) below this quantity, with a certain probability. On the other hand, by **lower limit** statements like: this decay, if exists, takes place with a lifetime larger than this quantity, with a certain probability. Both statements concern an exclusion.

Bayes limits

$$L(n_0/s) = \frac{e^{-s} s^{n_0}}{n_0!}$$

Assume background b=0

If we count n_0=0

$$L(0/s) = e^{-s}$$

Let's consider Bayes theorem and assume uniform prior (π=cost for s>0 and π=0 for s<0)

$$p(s/0) = \frac{L(0/s)\pi(s)}{\int L(0/s)\pi(s)ds} = L(0/s) = e^{-s}$$

Given a probability content α (e.g. α=95%) the upper limit s_{up} will be such that:

$$\int_{s_{up}}^{\infty} p(s/0)ds = 1 - \alpha$$

$$\int_{s_{up}}^{\infty} e^{-s}ds = e^{-s_{up}} = 1 - \alpha$$

We easily find s_{up}=2.3 for α=90% and s_{up}=3 for α=95%.

Bayes limits

Assume background $b \neq 0$ with negligible uncertainty and same prior as before

If we count $n_0 \geq 0$

$$p(s/n_0) = \frac{e^{-(s+b)}(s+b)^{n_0}}{n_0!}$$

$$\int_{s_{up}}^{\infty} \frac{e^{-(s+b)}(s+b)^{n_0}}{n_0!} ds = 1 - \alpha$$

Bayes limits

$$\int_{s_{up}}^{\infty} \frac{e^{-(s+b)}(s+b)^{n_0}}{n_0!} ds = 1 - \alpha$$

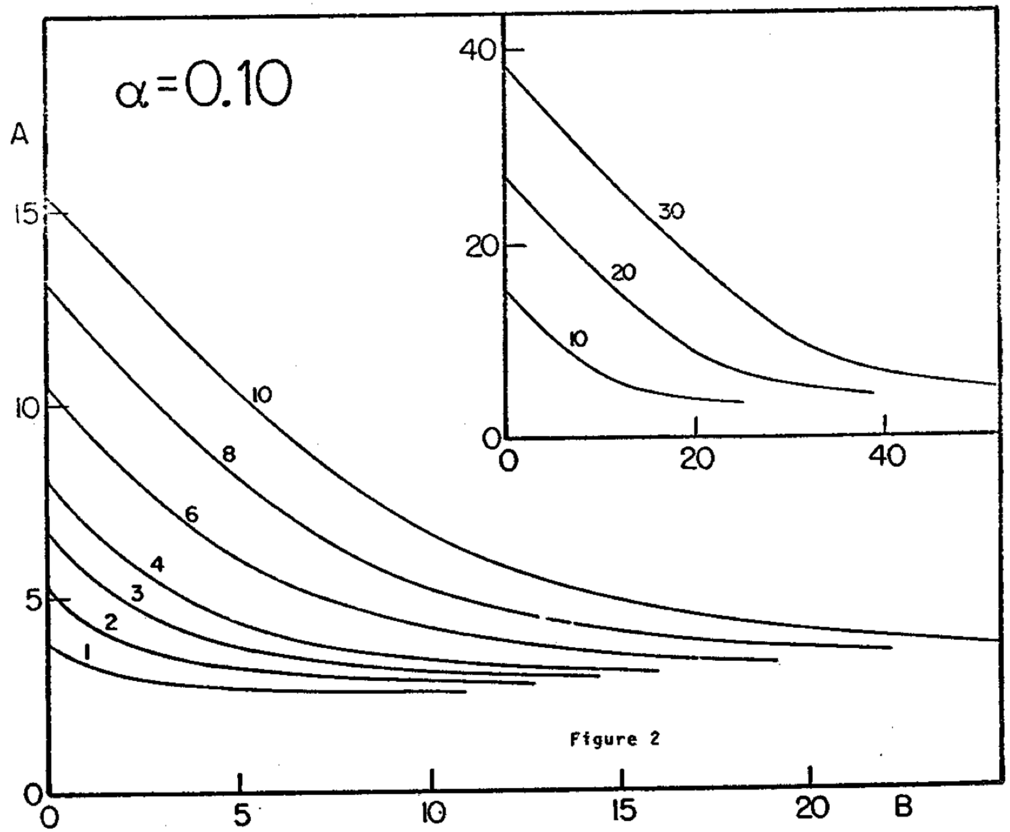


FIGURE 18. 90% limit s_{up} (A in the figure) vs. b (B in the figure) for different values of n_0. These are the upper limits resulting from a bayesian treatment with uniform prior. (taken from O.Helene, Nucl.Instr. and Meth. 212 (1983) 319)

Poisson

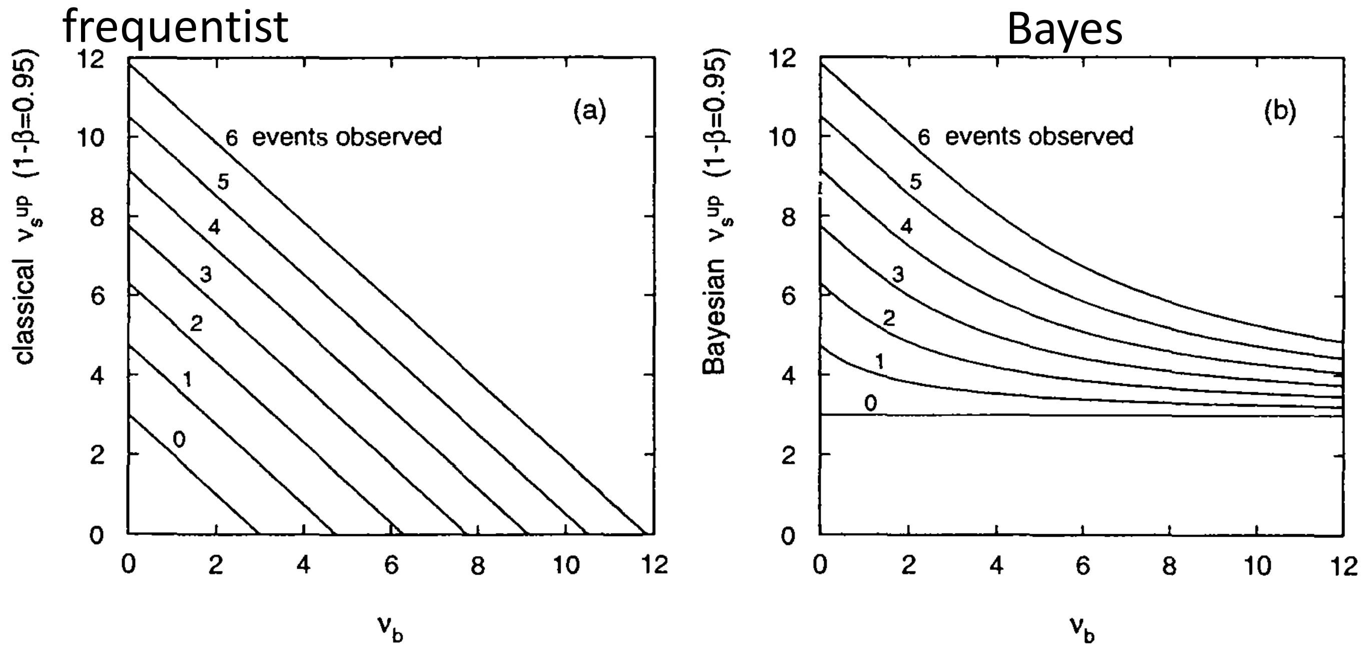


Fig. 9.9 Upper limits $\nu_s^{\rm up}$ at a confidence level of $1-\beta = 0.95$ for different numbers of events observed $n_{\rm obs}$ and as a function of the expected number of background events $\nu_{\rm b}$. (a) The classical limit. (b) The Bayesian limit based on a uniform prior density for ν_s.

Bayes limits

Assume background b ≠ 0 with uncertainty described by a pdf f(b) within interval bmin, bmax

$$p(s/n_0) = \frac{e^{-(s+b)}(s+b)^{n_0}}{n_0!}$$

↓ Convolution with the resolution f(b)

$$p(s/n_0) = \int_{b_{min}}^{b_{max}} \frac{e^{-(s+b')}(s+b')^{n_0}}{n_0!} f(b-b')db'$$

In general the width of f(b) affects the limit, large uncertainty on b => increase of S_{up}
The result in general depends on the prior (π(s)= cost, 1/s, 1/√s) (not in the case n_0=b=0)

Bayes limits

The General result for any n_0, is the pdf

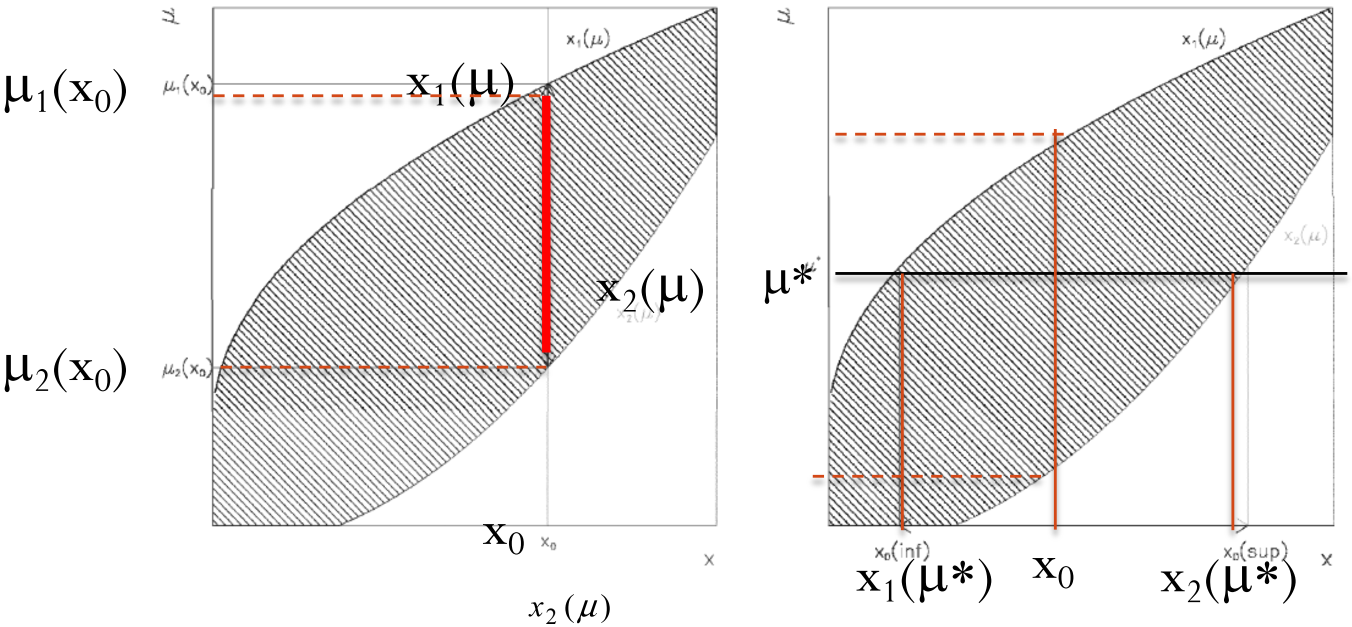
$$p(s/n_0)$$

If n0 significantly larger than b => observation of the signal
=> transition from upper limit to central interval:

$$\hat{s} = n_0 - b \pm \sqrt{n_0 + \sigma^2(b)}$$

Depending on the observed value and somewhat arbitrary => flip-flop problem (see next)

Neyman's construction



$$\int_{x_1(\mu)}^{x_2(\mu)} f(x/\mu)dx = \beta$$

By construction the probability to measure $x_0' < x_0$ if the true value $\mu = \mu_1(x_0)$ is $(1-\beta)/2$

$x_0' > x_0$ if the true value $\mu = \mu_2(x_0)$ is $(1-\beta)/2$

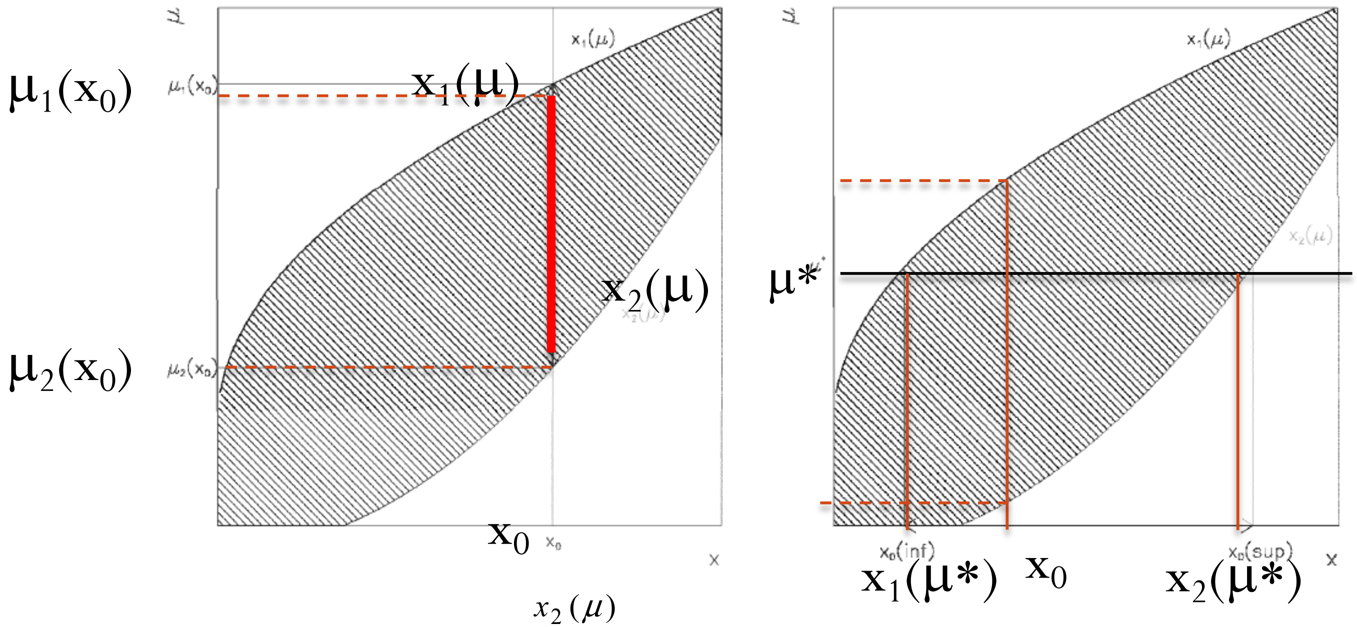
The determined C.I. is $[\mu_2(x_0)\ ,\ \mu_1(x_0)\ ]$.

Check the correct coverage: suppose μ^* is the true value. I repeat N times the measurement and determine each time the C.I. By construction in a fraction β of the cases x_0 is within $x_1(\mu^*)$ and $x_2(\mu^*)$ and the corresponding C.I. provides coverage of μ^*.

$$P(x_1(\mu^*) < x_0 < x_2(\mu^*)) = \beta$$

In $1-\beta$ cases x_0 lies outside the interval $[x_1(\mu^*), x_2(\mu^*)]$ and the corresponding C.I. does not cover μ^*.

Neyman's construction



$$\int_{x_1(\mu)}^{x_2(\mu)} f(x/\mu)dx = \beta$$

By construction the probability to measure $x_0' < x_0$ if the true value $\mu=\mu_1(x_0)$ is $(1-\beta)/2$

$x_0' > x_0$ if the true value $\mu=\mu_2(x_0)$ is $(1-\beta)/2$

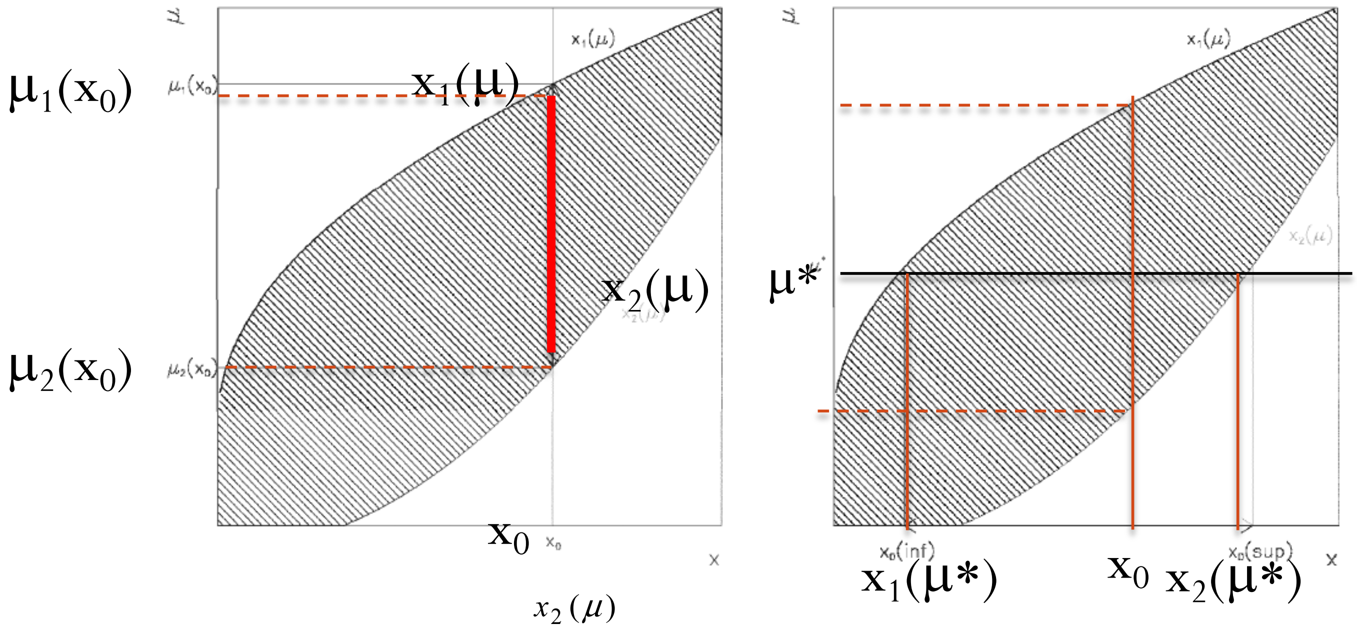
The determined C.I. is $[\mu_2(x_0), \mu_1(x_0)]$.

Check the correct coverage: suppose μ^* is the true value. I repeat N times the measurement and determine each time the C.I. By construction in a fraction β of the cases x_0 is within $x_1(\mu^*)$ and $x_2(\mu^*)$ and the corresponding C.I. provides coverage of μ^*.

$$P(x_1(\mu^*) < x_0 < x_2(\mu^*)) = \beta$$

In $1-\beta$ cases x_0 lies outside the interval $[x_1(\mu^*), x_2(\mu^*)]$ and the corresponding C.I. does not cover μ^*.

Neyman's construction



$$\int_{x_1(\mu)}^{x_2(\mu)} f(x/\mu)dx = \beta$$

By construction the probability to measure $x_0' < x_0$ if the true value $\mu = \mu_1(x_0)$ is $(1-\beta)/2$

$x_0' > x_0$ if the true value $\mu = \mu_2(x_0)$ is $(1-\beta)/2$

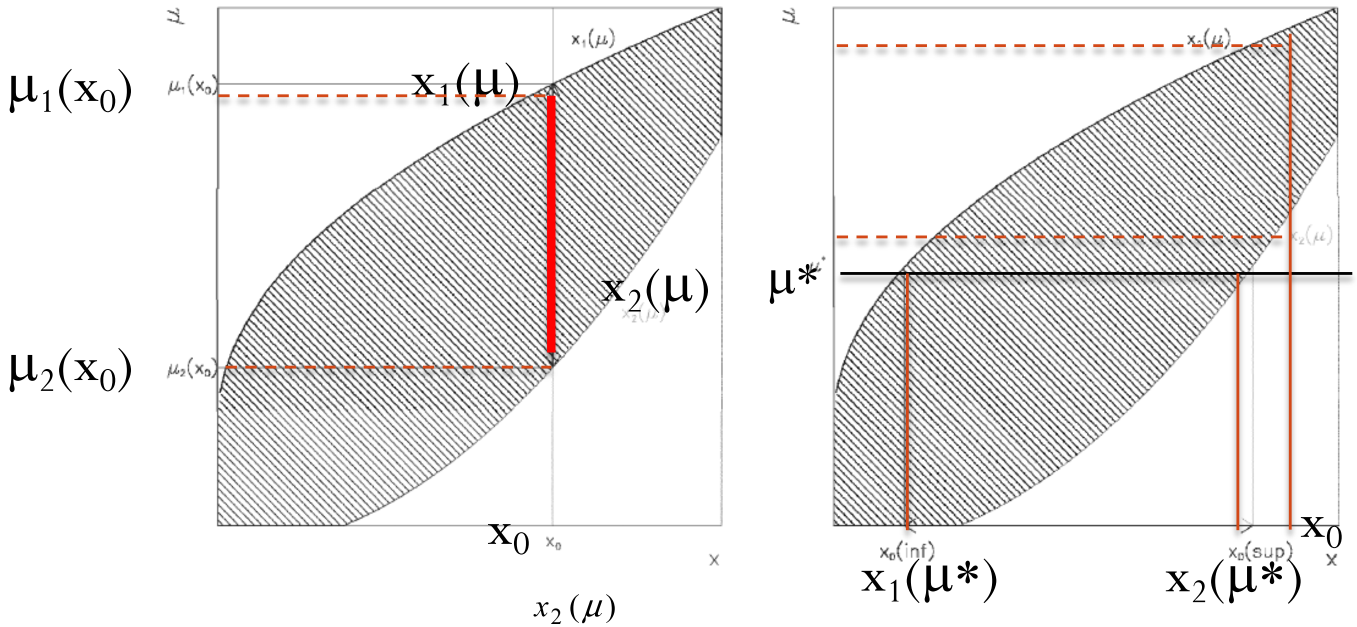
The determined C.I. is $[\mu_2(x_0)\ ,\ \mu_1(x_0)\ ]$.

Check the correct coverage: suppose μ^* is the true value. I repeat N times the measurement and determine each time the C.I. By construction in a fraction β of the cases x_0 is within $x_1(\mu^*)$ and $x_2(\mu^*)$ and the corresponding C.I. provides coverage of μ^*.

$$P(x_1(\mu^*) < x_0 < x_2(\mu^*)) = \beta$$

In $1-\beta$ cases x_0 lies outside the interval $[x_1(\mu^*), x_2(\mu^*)]$ and the corresponding C.I. does not cover μ^*.

Neyman's construction



$$\int_{x_1(\mu)}^{x_2(\mu)} f(x/\mu)dx = \beta$$

By construction the probability to measure $x_0' < x_0$ if the true value $\mu=\mu_1(x_0)$ is $(1-\beta)/2$

$x_0' > x_0$ if the true value $\mu=\mu_2(x_0)$ is $(1-\beta)/2$

The determined C.I. is $[\mu_2(x_0), \mu_1(x_0)]$.

Check the correct coverage: suppose μ^* is the true value. I repeat N times the measurement and determine each time the C.I. By construction in a fraction β of the cases x_0 is within $x_1(\mu^*)$ and $x_2(\mu^*)$ and the corresponding C.I. provides coverage of μ^*.

$$P(x_1(\mu^*) < x_0 < x_2(\mu^*)) = \beta$$

In $1-\beta$ cases x_0 lies outside the interval $[x_1(\mu^*), x_2(\mu^*)]$ and the corresponding C.I. does not cover μ^*.

frequentist limits

The belt is limited on one side only, and for any result of a measurement n_0 we identify s_{up} in such a way that if $s_{true} = s_{up}$, the probability to get a counting smaller than n_0 is $1-\beta$[31]. By considering the Poisson statistics without background (b=0) we get:

$$\sum_{n=0}^{n_0} \frac{e^{-s_{up}} s_{up}^n}{n!} = 1 - \beta$$

If $n_0 = 0$ we have

$$e^{-s_{up}} = 1 - \beta$$

$$s_{up} = \ln \frac{1}{1-\beta}$$

from which we get the same numbers for s_{up} obtained in the bayesian case.

frequentist limits

By construction the probability to measure $n_0' < n_0$ if the true value $s = s_{up}(n_0)$ is $(1-\beta)$ (only one limit) or the probability to measure $n_0' > n_0$ if the true value $s = s_{up}(n_0)$ is β

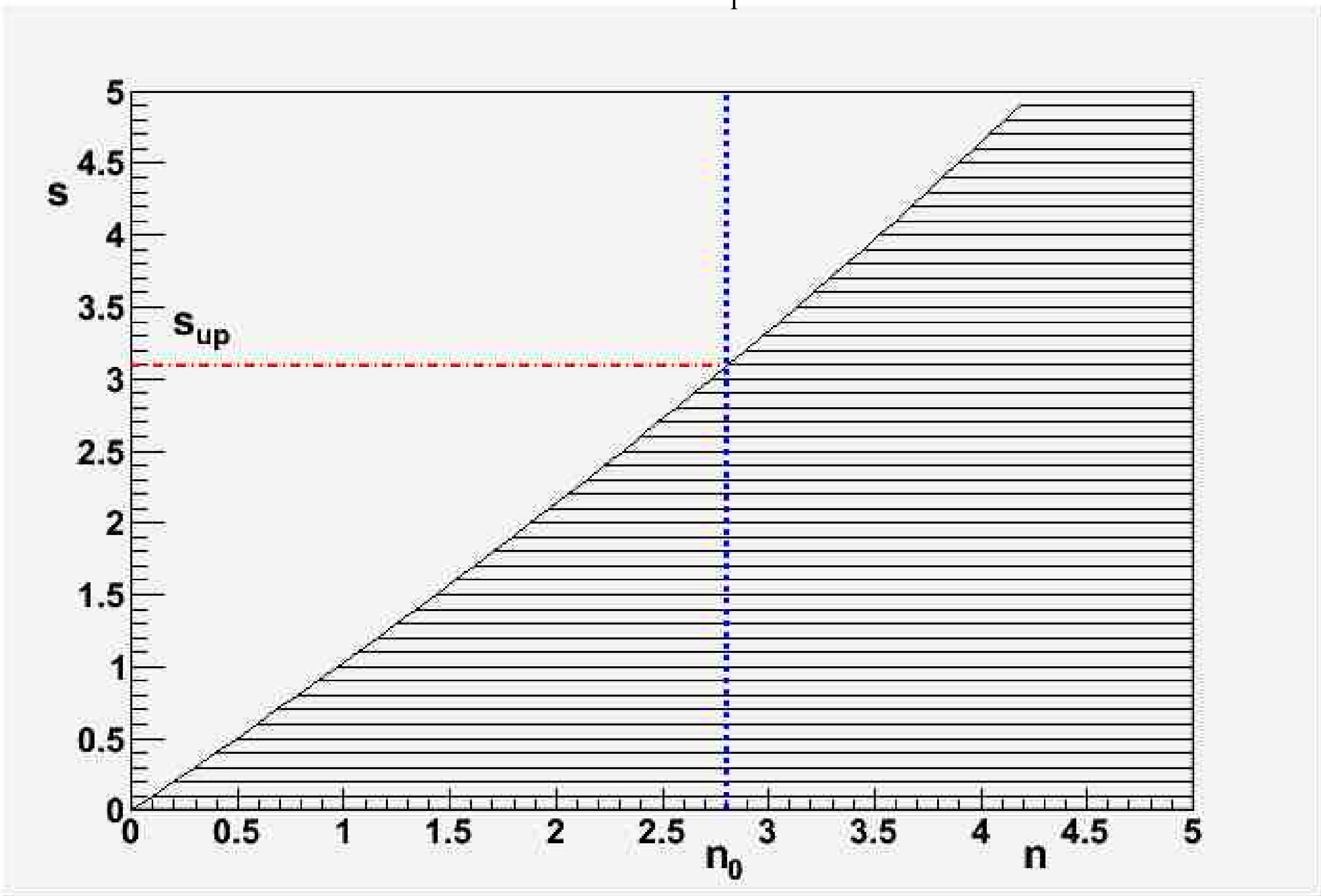


FIGURE 19. Neyman construction for the case of an upper limit. In this case a segment between $n_1(\theta)$ and ∞ is drawn for each value of the parameter θ. The segments define the confidence region. Once a value of n, n_0 is obtained, the upper limit s_{up} is found. (For simplicity the discrete variable n is considered as a real number here).

If b is not equal to 0 but is known,

$$\sum_{n=0}^{n_0} \frac{e^{-(s_{up}+b)}(s_{up}+b)^n}{n!} = 1 - \beta \tag{201}$$

and from this equation upper limits can be evaluated for the different situations.

It has been pointed out that the use of eq.201 gives rise to some problems. In particular negative values of s_{up} can be obtained using directly the formula[32]. This doesn't happen in the bayesian context where the condition $s > 0$ is put directly by using the proper prior.

[32]A rate is a positive-definite quantity. However, if a rate is 0 or very small with respect to the experimental sensitivity, the probability that n_0 is larger than b is exactly equal to the probability that n_0 is lower than b. This implies that a negative rate naturally comes out from an experimental analysis based on a difference between two counts. The acceptance of such results is a sort of "philosophical" question and is controversial.

Poisson

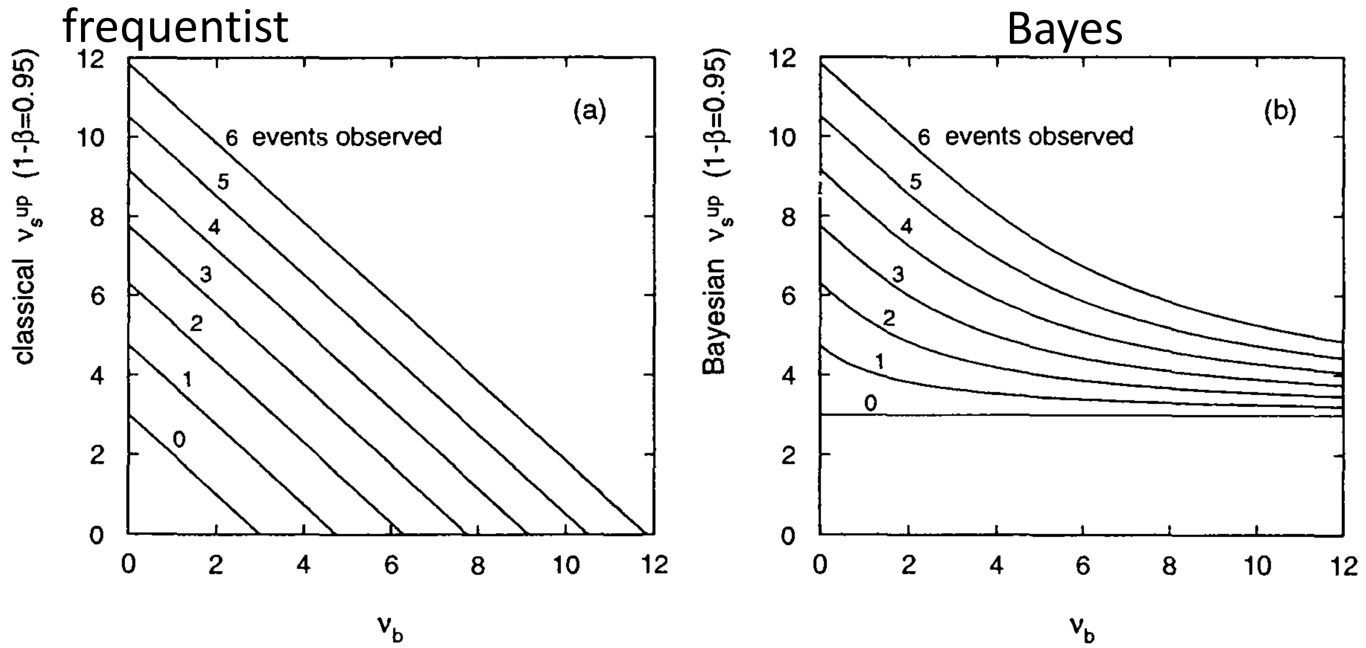


Fig. 9.9 Upper limits $\nu_s^{\rm up}$ at a confidence level of $1-\beta = 0.95$ for different numbers of events observed $n_{\rm obs}$ and as a function of the expected number of background events $\nu_{\rm b}$. (a) The classical limit. (b) The Bayesian limit based on a uniform prior density for ν_s.

Poisson

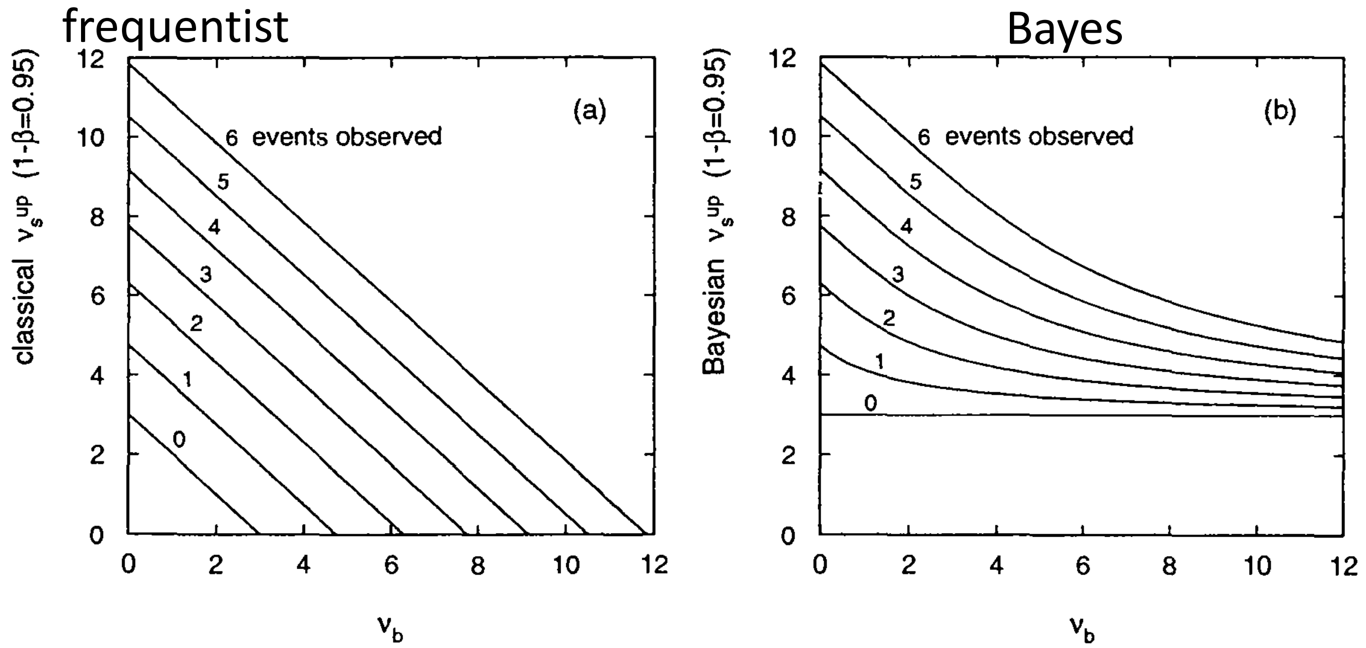


Fig. 9.9 Upper limits ν_s^{up} at a confidence level of $1-\beta = 0.95$ for different numbers of events observed n_{obs} and as a function of the expected number of background events ν_{b}. (a) The classical limit. (b) The Bayesian limit based on a uniform prior density for ν_s.

What if νb=4 and 0 events observed?

Flip-flop problem

Another general problem affecting both bayesian and frequentist approach is the so called **flip-flop** problem. When n_0 is larger than b, at a given point the experimentalist decides to present the result as a number $\pm$ an uncertainty rather than an upper limit. Such a decision is somehow arbitrary. A method to avoid this problem is the so called **unified approach** due to Feldman and Cousins, developed in the frequentist context.

(see next)

Example of discrepancy between frequentist and Bayesian approaches. (data available in ‘90s)

Results from fits;
PDG weighted average: $\overline{m}^2 = -54 \pm 30\ \mathrm{eV}^2$

How can this result be converted into an upper limit for the neutrino mass?

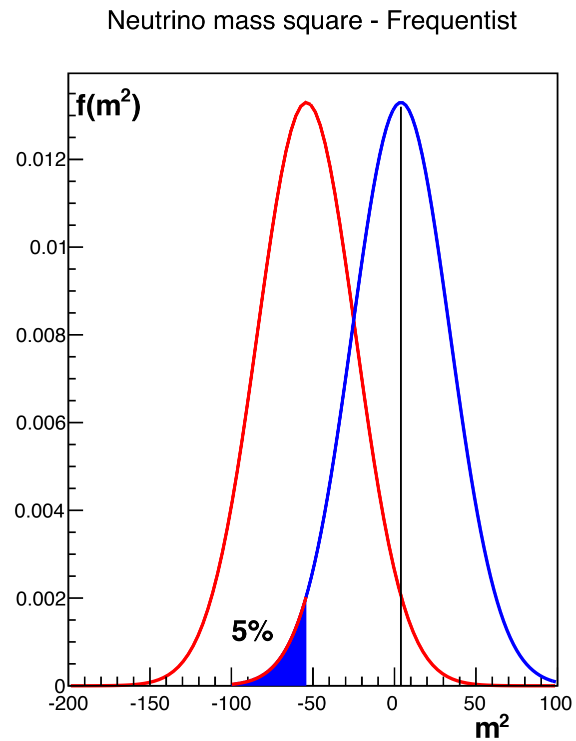
Example of discrepancy between frequentist and Bayesian approaches. (data available in '90s)

Results from fits;
PDG weighted average: $\overline{m}^2 = -54 \pm 30 \ \text{eV}^2$

How can this result be converted into an upper limit for the neutrino mass?

In the frequentist approach, Neyman's construction

At 97.5% CL => $m^2 < 4.6 \ \text{eV}^2$



Example of discrepancy between frequentist and Bayesian approaches. (data available in '90s)

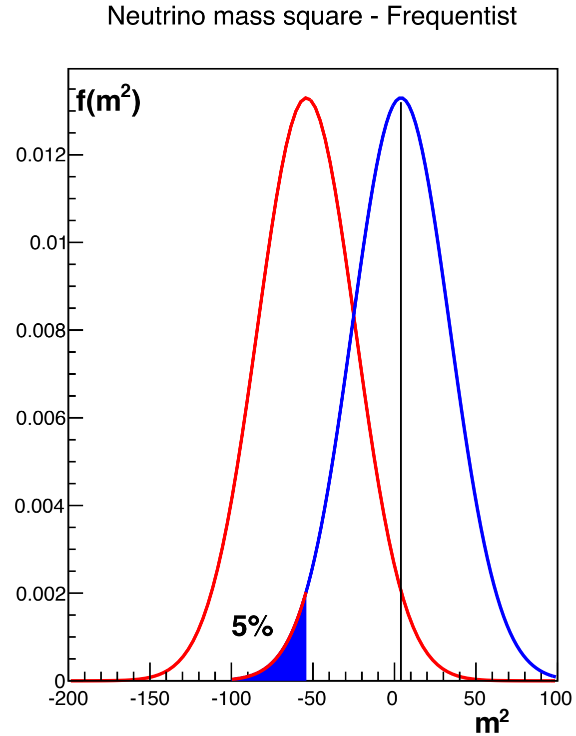
Results from fits;
PDG weighted average: $$\overline{m}^2 = -54 \pm 30 \ \mathrm{eV}^2$$

How can this result be converted into an upper limit for the neutrino mass?

In the frequentist approach, Neyman's construction

At 97.5% CL => $m^2 < 4.6 \ \mathrm{eV}^2$

At 90% CL => $m^2 < -16 \ \mathrm{eV}^2$???



Example of discrepancy between frequentist and Bayesian approaches. (data available in '90s)

Results from fits;
PDG weighted average: $\overline{m}^2 = -54 \pm 30 \ \text{eV}^2$

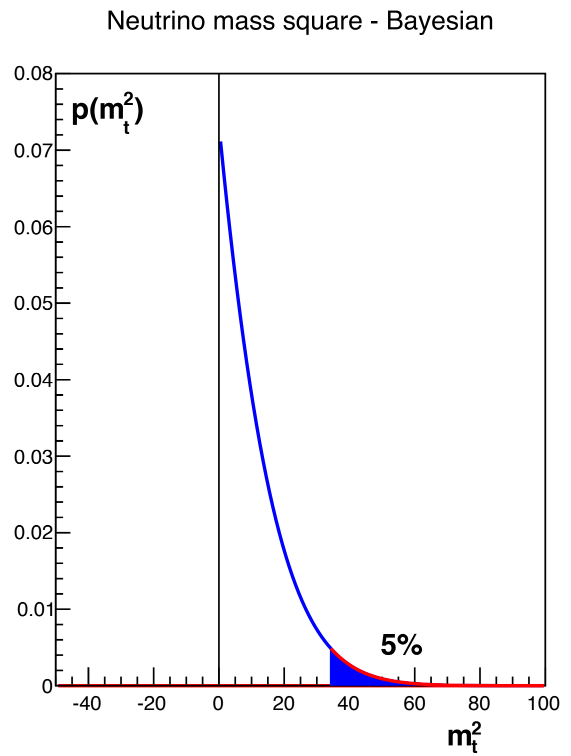
How can this result be converted into an upper limit for the neutrino mass?

In the Bayesian approach, using a prior forcing m_t^2 to be positive
(π=cost for $m_t^2>0$ and $\pi=0$ for $m_t^2<0$)

$$p(m_t^2/\overline{m}^2) = \frac{L(\overline{m}^2/m_t^2)\pi(m_t^2)}{\int dm_t^2 L(\overline{m}^2/m_t^2)}$$

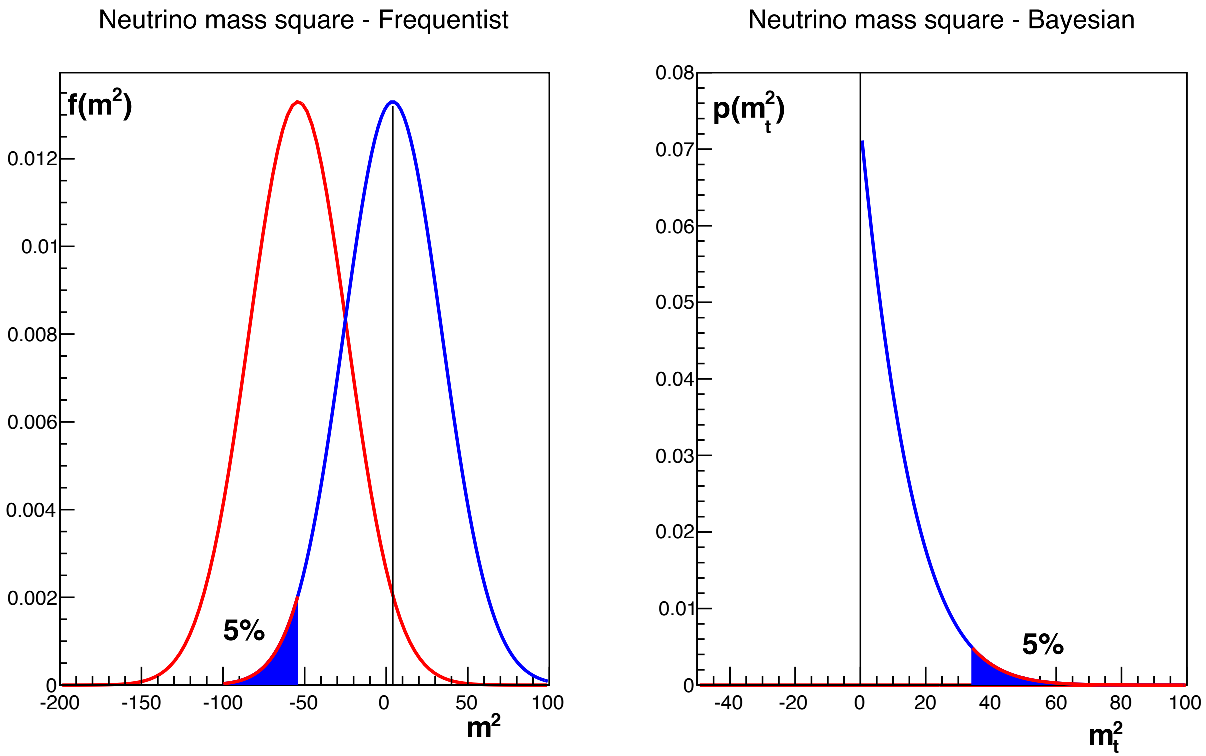
At 95% CL => $m_t^2 < 34 \ \text{eV}^2$

At 90% CL => $m_t^2 < 27 \ \text{eV}^2$



Example of discrepancy between frequentist and Bayesian approaches. (data available in '90s)

$m^2 < 4.6\ \mathrm{eV}^2$

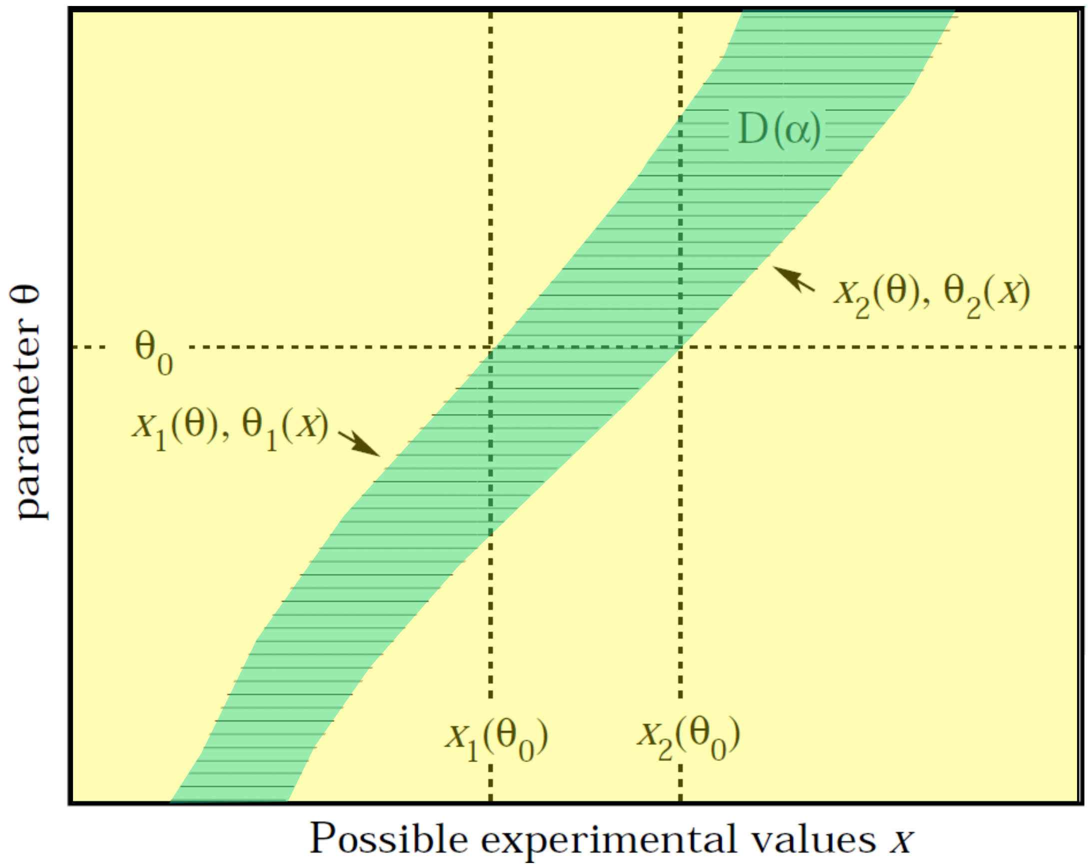


$m_t^2 < 34\ \mathrm{eV}^2$

FIGURE 21. Example of the square neutrino mass. Construction of the upper limit in the frequentist approach (left plot) and in the bayesian approach (right plot). (left) The red gaussian is the experimental likelihood, the blue gaussian corresponds to the 95% CL upper limit that leaves 5% of possible the experiment outcomes below the present experimental average. (right) The blue curve is the result of the Bayes theorem when a prior forcing to positive values is applied (eq.202).

The frequentist unified approach
(Feldman and Cousins PRD 57 3873 (1998))

- Scan an unknown parameter θ over its range
- Given θ, compute the interval $[x_1, x_2]$ that contain x with a probability $CL = 1-\alpha$
- Ordering rule is needed!
  - Central interval? Asymmetric? Other?
- Invert the confidence belt, and find the interval $[\theta_1, \theta_2]$ for a given experimental outcome of x
- A fraction $1-\alpha$ of the experiments will produce x such that the corresponding interval $[\theta_1, \theta_2]$ contains the true value of μ (coverage probability)
- Note that the random variables are $[\theta_1, \theta_2]$, not θ



From PDG statistics review

`RooStats::NeymanConstruction`

The frequentist unified approach
(Feldman and Cousins PRD 57 3873 (1998))

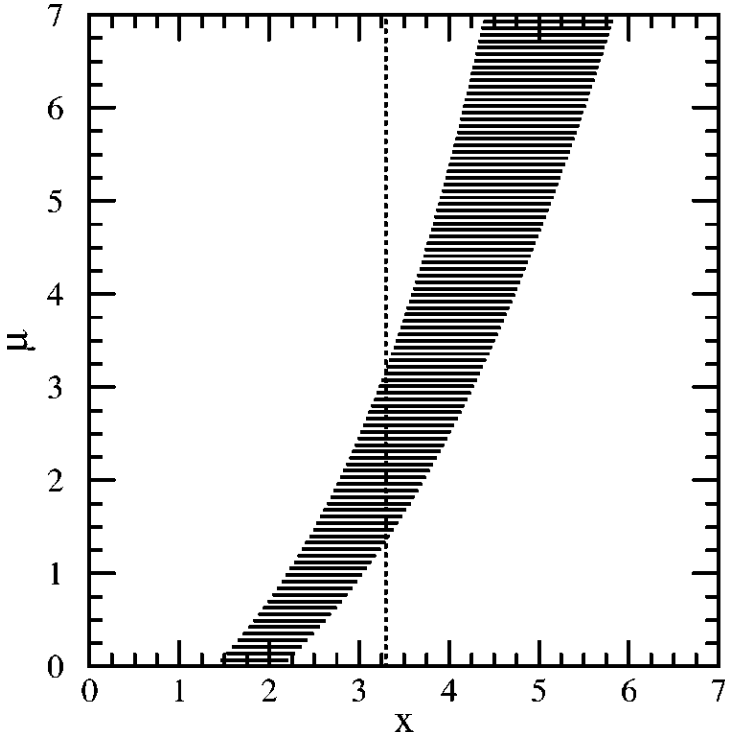


FIG. 1. A generic confidence belt construction and its use. For each value of μ, one draws a horizontal acceptance interval $[x_1, x_2]$ such that $P(x \in [x_1, x_2] | \mu) = \alpha$. Upon performing an experiment to measure x and obtaining the value x_0, one draws the dashed vertical line through x_0. The confidence interval $[\mu_1, \mu_2]$ is the union of all values of μ for which the corresponding acceptance interval is intercepted by the vertical line.

$$P(x \in [x_1, x_2] | \mu) = \alpha.$$

$$P(x \in [x_1, x_2] | \mu) = \alpha. \tag{2.4}$$

Such intervals are drawn as horizontal line segments in Fig. 1, at representative values of μ. We refer to the interval $[x_1, x_2]$ as the ‘‘acceptance region’’ or the ‘‘acceptance interval’’ for that μ. In order to specify uniquely the acceptance region, one must *choose* auxiliary criteria. One has total freedom to make this choice, *if the choice is not influenced by the data* x_0. The most common choices are

$$P(x < x_1 | \mu) = 1 - \alpha, \tag{2.5}$$

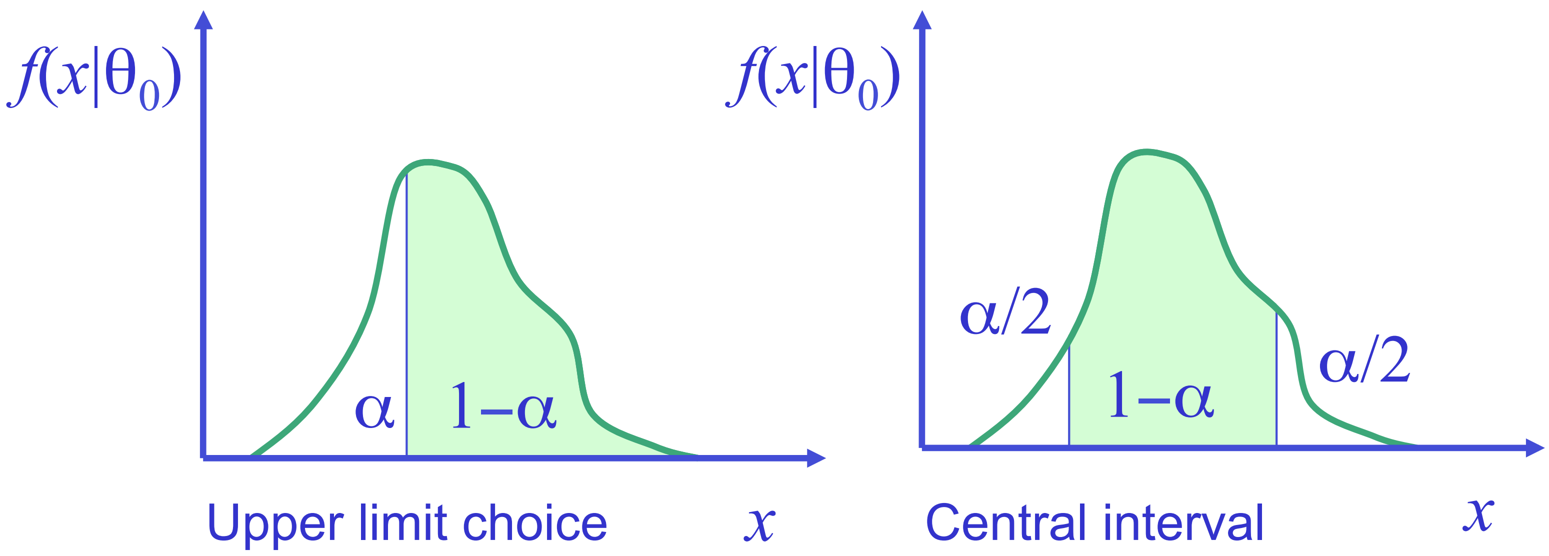
which leads to ‘‘upper confidence limits’’ (which satisfy $P(\mu > \mu_2) = 1 - \alpha$), and

$$P(x < x_1 | \mu) = P(x > x_2 | \mu) = (1 - \alpha)/2, \tag{2.6}$$

which leads to ‘‘central confidence intervals’’ [which satisfy $P(\mu < \mu_1) = P(\mu > \mu_2) = (1 - \alpha)/2$]. For these choices, the

Ordering rule

- For a fixed $\theta = \theta_0$ we can have different possible choices of intervals giving the same probability $1-\alpha$ are possible



The frequentist unified approach
(Feldman and Cousins PRD 57 3873 (1998))

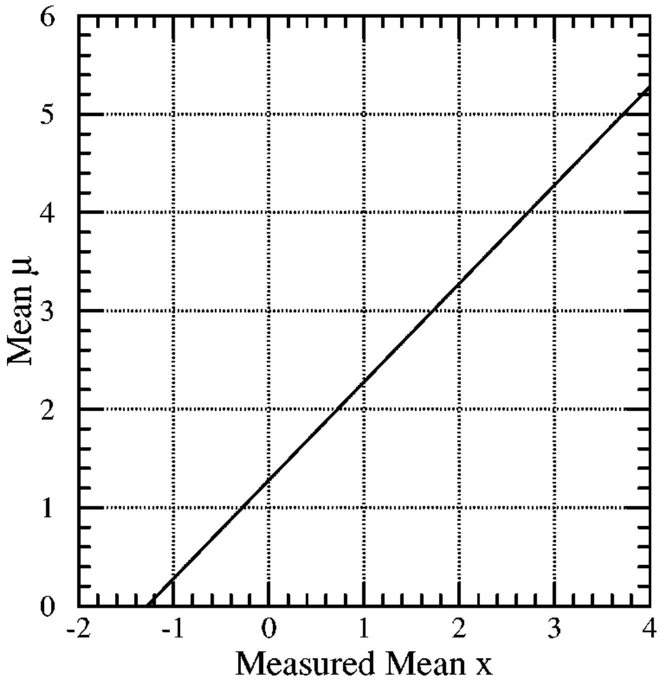


FIG. 2. Standard confidence belt for 90% C.L. upper limits for the mean of a Gaussian, in units of the rms deviation. The second line in the belt is at $x = +\infty$.

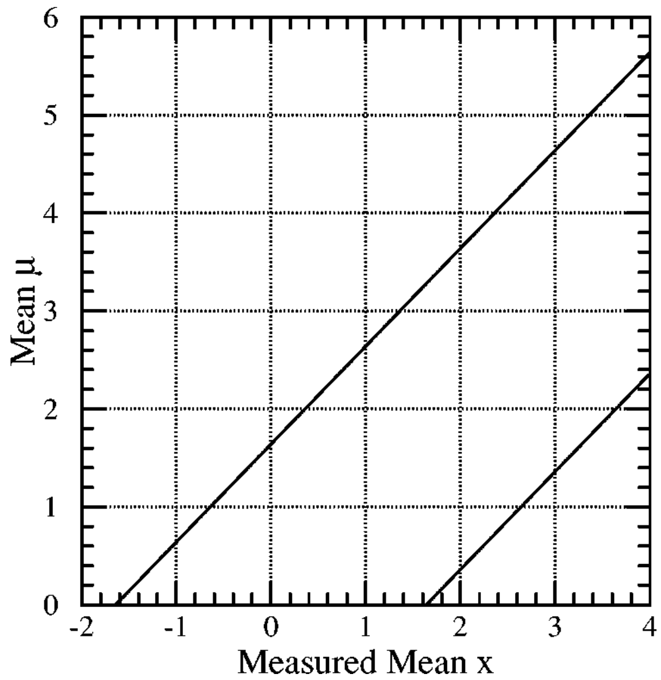


FIG. 3. Standard confidence belt for 90% C.L. central confidence intervals for the mean of a Gaussian, in units of the rms deviation.

The frequentist unified approach
(Feldman and Cousins PRD 57 3873 (1998))

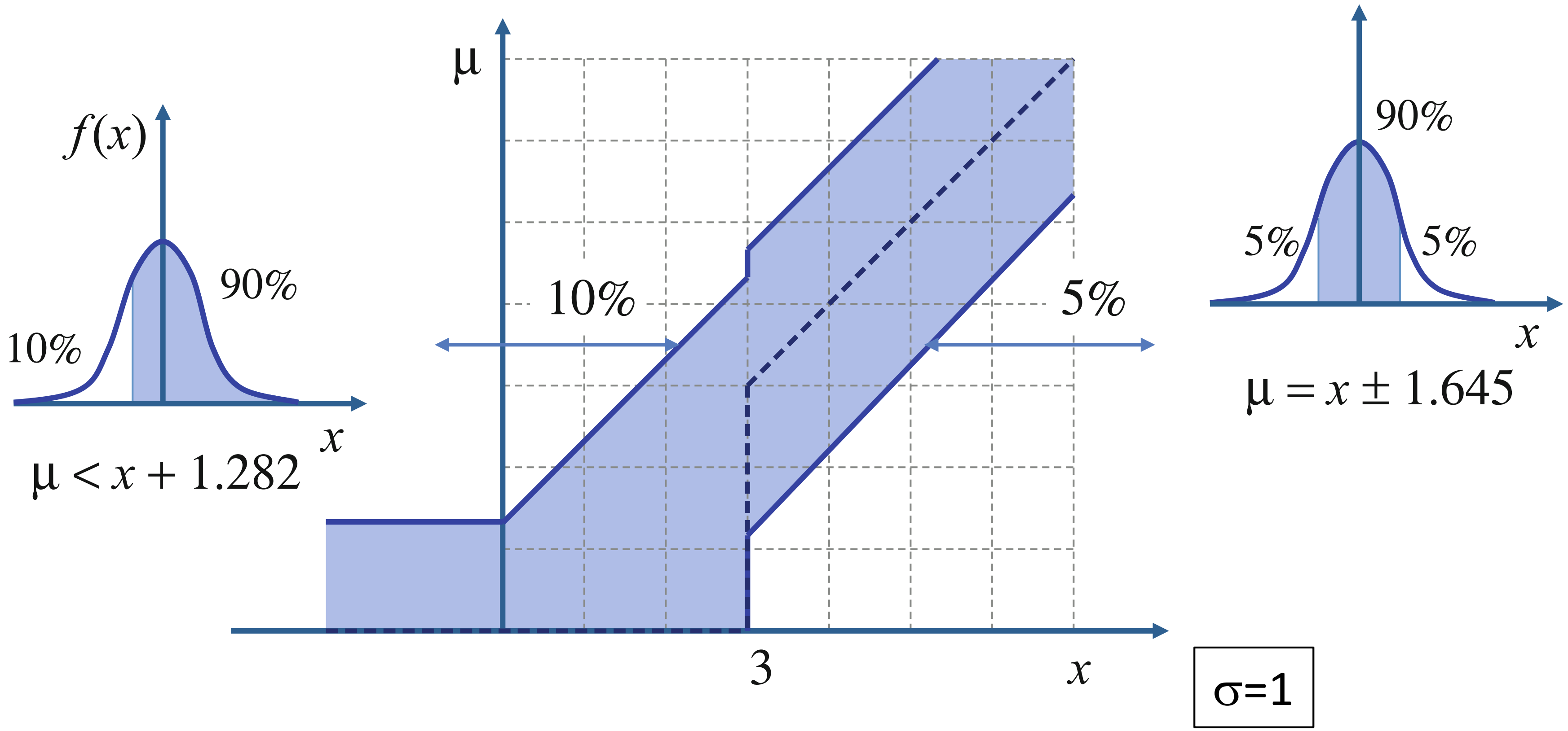


Fig. 7.6 Illustration of the *flip-flopping* problem. The plot shows the quoted central value of μ as a function of the measured x (*dashed line*), and the 90% confidence interval corresponding to the choice of quoting a central interval for $x/\sigma \geq 3$ and an upper limit for $x/\sigma < 3$. The coverage decreases from 90 to 85% for a value of μ corresponding to the horizontal lines with arrows

The frequentist unified approach
(Feldman and Cousins PRD 57 3873 (1998))

In order to avoid the flip-flopping problem and to ensure the correct coverage, the ordering rule proposed by Feldman and Cousins [3] provides a Neyman confidence belt, following the procedure described in Sect. 7.2, that smoothly changes from a central or quasi-central interval to an upper limit, in the case of low observed signal yield.

The proposed ordering rule is based on a likelihood ratio whose properties will be further discussed in Sect. 9.5. Given a value θ_0 of the unknown parameter θ, the chosen interval of the variable x used for the Neyman belt construction is defined by the ratio of two PDFs of x, one under the hypothesis that θ is equal to the considered fixed value θ_0, the other under the hypothesis that θ is equal to the maximum likelihood estimate value $\hat{\theta}(x)$, corresponding to the given measurement

The frequentist unified approach
(Feldman and Cousins PRD 57 3873 (1998))

The likelihood ratio must be greater than a constant k_α whose value depends on the chosen confidence level $1-\alpha$. In a formula:

$$\lambda(x \,|\, \theta_0) = \frac{f(x \,|\, \theta_0)}{f(x \,|\, \hat{\theta}(x))} > k_\alpha \,. \tag{7.11}$$

The constant k_α should be taken such that the integral of the PDF in the confidence interval R_α is equal to $1-\alpha$:

$$\int_{R_\alpha} f(x \,|\, \theta_0)\,\mathrm{d}x = 1-\alpha \,. \tag{7.12}$$

The confidence interval R_α for a given value $\theta = \theta_0$ is defined by Eq. (7.11):

$$R_\alpha(\theta_0) = \{x : \lambda(x \,|\, \theta_0) > k_\alpha\} \,. \tag{7.13}$$

The frequentist unified approach
(Feldman and Cousins PRD 57 3873 (1998))

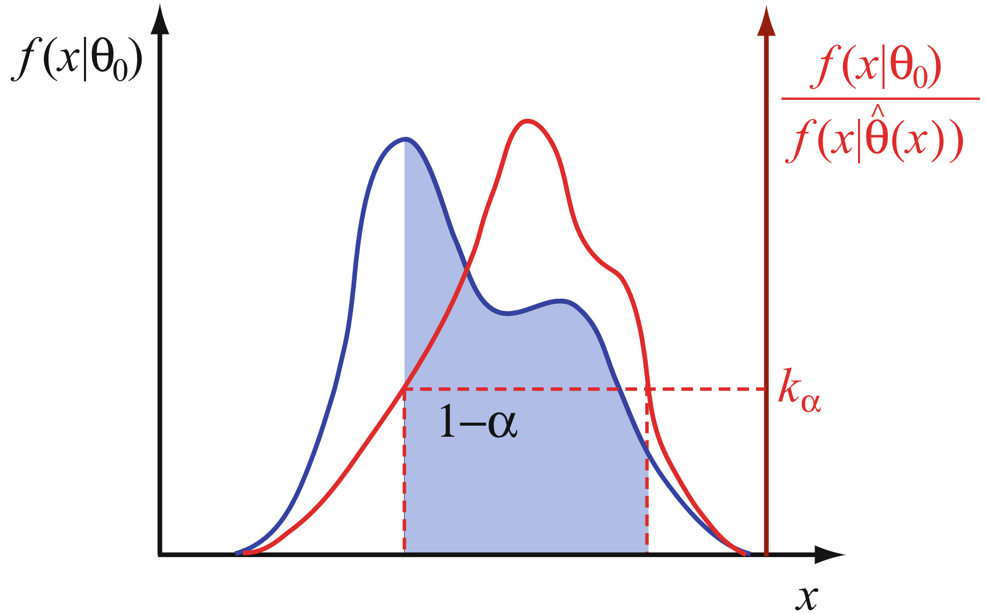


Fig. 7.7 Ordering rule in the Feldman–Cousins approach, based on the likelihood ratio $\lambda(x \mid \theta_0) = f(x \mid \theta_0)/f(x \mid \hat{\theta}(x))$

The frequentist unified approach
(Feldman and Cousins PRD 57 3873 (1998))

Two examples
1) Gaussian errors with a bounded physical region
2) Poisson processes with background

In contrast with the usual classical construction for upper limits, the unified construction “naturally” avoids the flip-flop problem and unphysical confidence intervals

The frequentist unified approach
(Feldman and Cousins PRD 57 3873 (1998))

$$P(x \in [x_1, x_2] | \mu) = \alpha.$$

Rank x in the acceptance interval $[x_1, x_2]$ by the ratio

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

where μ_{best} is the physically allowed value of μ for which $P(x|\mu)$ is maximum.

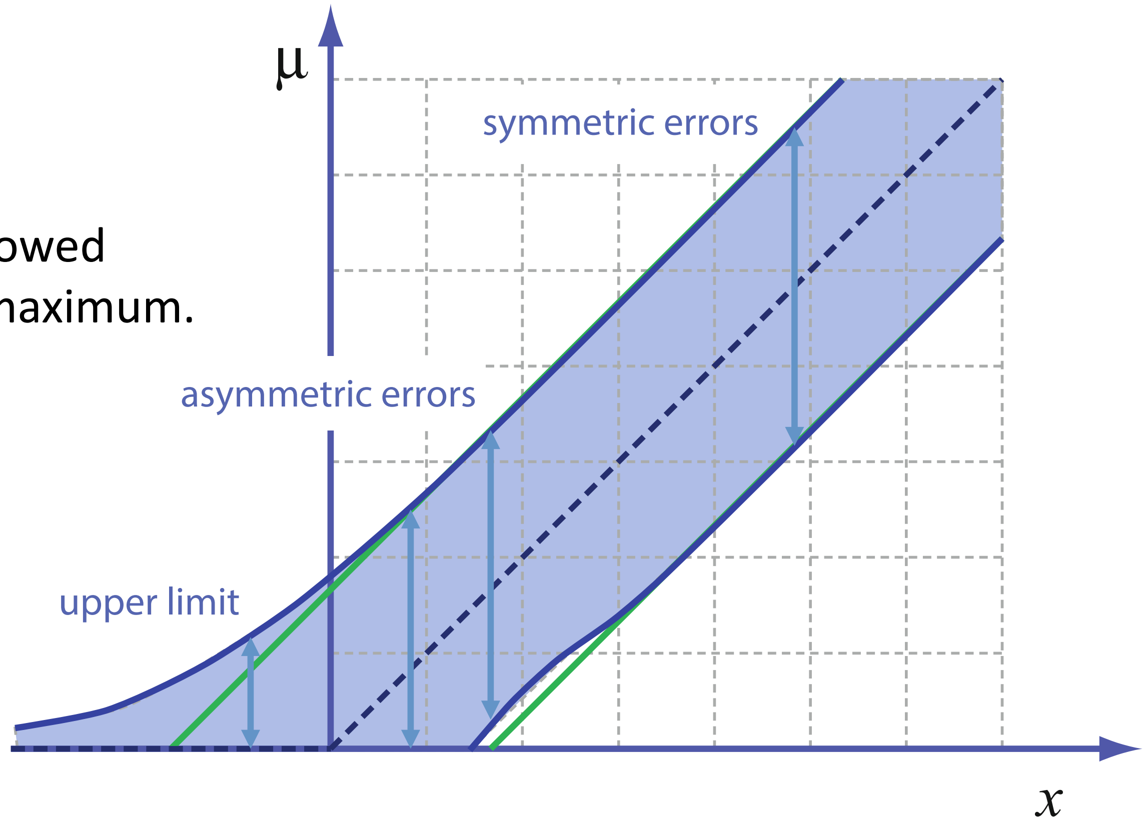


Fig. 7.8 Neyman confidence belt constructed using the Feldman–Cousins ordering

Flip-flop problem: the frequentist unified approach
(Feldman and Cousins PRD 57 3873 (1998))

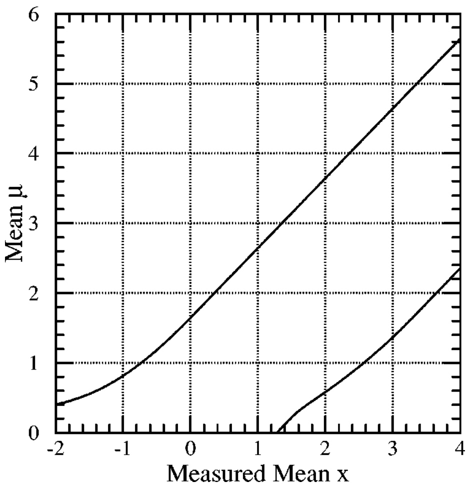


FIG. 10. Plot of our 90% confidence intervals for the mean of a Gaussian, constrained to be non-negative, described in the text.

TABLE X. Our confidence intervals for the mean μ of a Gaussian, constrained to be non-negative, as a function of the measured mean x_0, for commonly used confidence levels. Italicized intervals correspond to cases where the goodness-of-fit probability (Sec. IV C) is less than 1%. All numbers are in units of σ.

x_0	68.27% C.L.	90% C.L.	95% C.L.	99% C.L.	x_0	68.27% C.L.	90% C.L.	95% C.L.	99% C.L.
−3.0	*0.00, 0.04*	*0.00, 0.26*	*0.00, 0.42*	*0.00, 0.80*	0.1	0.00, 1.10	0.00, 1.74	0.00, 2.06	0.00, 2.68
−2.9	*0.00, 0.04*	*0.00, 0.27*	*0.00, 0.44*	*0.00, 0.82*	0.2	0.00, 1.20	0.00, 1.84	0.00, 2.16	0.00, 2.78
−2.8	*0.00, 0.04*	*0.00, 0.28*	*0.00, 0.45*	*0.00, 0.84*	0.3	0.00, 1.30	0.00, 1.94	0.00, 2.26	0.00, 2.88
−2.7	*0.00, 0.04*	*0.00, 0.29*	*0.00, 0.47*	*0.00, 0.87*	0.4	0.00, 1.40	0.00, 2.04	0.00, 2.36	0.00, 2.98
−2.6	*0.00, 0.05*	*0.00, 0.30*	*0.00, 0.48*	*0.00, 0.89*	0.5	0.02, 1.50	0.00, 2.14	0.00, 2.46	0.00, 3.08
−2.5	*0.00, 0.05*	*0.00, 0.32*	*0.00, 0.50*	*0.00, 0.92*	0.6	0.07, 1.60	0.00, 2.24	0.00, 2.56	0.00, 3.18
−2.4	*0.00, 0.05*	*0.00, 0.33*	*0.00, 0.52*	*0.00, 0.95*	0.7	0.11, 1.70	0.00, 2.34	0.00, 2.66	0.00, 3.28
−2.3	0.00, 0.05	0.00, 0.34	0.00, 0.54	0.00, 0.99	0.8	0.15, 1.80	0.00, 2.44	0.00, 2.76	0.00, 3.38
−2.2	0.00, 0.06	0.00, 0.36	0.00, 0.56	0.00, 1.02	0.9	0.19, 1.90	0.00, 2.54	0.00, 2.86	0.00, 3.48
−2.1	0.00, 0.06	0.00, 0.38	0.00, 0.59	0.00, 1.06	1.0	0.24, 2.00	0.00, 2.64	0.00, 2.96	0.00, 3.58
−2.0	0.00, 0.07	0.00, 0.40	0.00, 0.62	0.00, 1.10	1.1	0.30, 2.10	0.00, 2.74	0.00, 3.06	0.00, 3.68
−1.9	0.00, 0.08	0.00, 0.43	0.00, 0.65	0.00, 1.14	1.2	0.35, 2.20	0.00, 2.84	0.00, 3.16	0.00, 3.78
−1.8	0.00, 0.09	0.00, 0.45	0.00, 0.68	0.00, 1.19	1.3	0.42, 2.30	0.02, 2.94	0.00, 3.26	0.00, 3.88
−1.7	0.00, 0.10	0.00, 0.48	0.00, 0.72	0.00, 1.24	1.4	0.49, 2.40	0.12, 3.04	0.00, 3.36	0.00, 3.98
−1.6	0.00, 0.11	0.00, 0.52	0.00, 0.76	0.00, 1.29	1.5	0.56, 2.50	0.22, 3.14	0.00, 3.46	0.00, 4.08
−1.5	0.00, 0.13	0.00, 0.56	0.00, 0.81	0.00, 1.35	1.6	0.64, 2.60	0.31, 3.24	0.00, 3.56	0.00, 4.18
−1.4	0.00, 0.15	0.00, 0.60	0.00, 0.86	0.00, 1.41	1.7	0.72, 2.70	0.38, 3.34	0.06, 3.66	0.00, 4.28
−1.3	0.00, 0.17	0.00, 0.64	0.00, 0.91	0.00, 1.47	1.8	0.81, 2.80	0.45, 3.44	0.16, 3.76	0.00, 4.38
−1.2	0.00, 0.20	0.00, 0.70	0.00, 0.97	0.00, 1.54	1.9	0.90, 2.90	0.51, 3.54	0.26, 3.86	0.00, 4.48
−1.1	0.00, 0.23	0.00, 0.75	0.00, 1.04	0.00, 1.61	2.0	1.00, 3.00	0.58, 3.64	0.35, 3.96	0.00, 4.58
−1.0	0.00, 0.27	0.00, 0.81	0.00, 1.10	0.00, 1.68	2.1	1.10, 3.10	0.65, 3.74	0.45, 4.06	0.00, 4.68
−0.9	0.00, 0.32	0.00, 0.88	0.00, 1.17	0.00, 1.76	2.2	1.20, 3.20	0.72, 3.84	0.53, 4.16	0.00, 4.78
−0.8	0.00, 0.37	0.00, 0.95	0.00, 1.25	0.00, 1.84	2.3	1.30, 3.30	0.79, 3.94	0.61, 4.26	0.00, 4.88
−0.7	0.00, 0.43	0.00, 1.02	0.00, 1.33	0.00, 1.93	2.4	1.40, 3.40	0.87, 4.04	0.69, 4.36	0.07, 4.98
−0.6	0.00, 0.49	0.00, 1.10	0.00, 1.41	0.00, 2.01	2.5	1.50, 3.50	0.95, 4.14	0.76, 4.46	0.17, 5.08
−0.5	0.00, 0.56	0.00, 1.18	0.00, 1.49	0.00, 2.10	2.6	1.60, 3.60	1.02, 4.24	0.84, 4.56	0.27, 5.18
−0.4	0.00, 0.64	0.00, 1.27	0.00, 1.58	0.00, 2.19	2.7	1.70, 3.70	1.11, 4.34	0.91, 4.66	0.37, 5.28
−0.3	0.00, 0.72	0.00, 1.36	0.00, 1.67	0.00, 2.28	2.8	1.80, 3.80	1.19, 4.44	0.99, 4.76	0.47, 5.38
−0.2	0.00, 0.81	0.00, 1.45	0.00, 1.77	0.00, 2.38	2.9	1.90, 3.90	1.28, 4.54	1.06, 4.86	0.57, 5.48
−0.1	0.00, 0.90	0.00, 1.55	0.00, 1.86	0.00, 2.48	3.0	2.00, 4.00	1.37, 4.64	1.14, 4.96	0.67, 5.58
0.0	0.00, 1.00	0.00, 1.64	0.00, 1.96	0.00, 2.58	3.1	2.10, 4.10	1.46, 4.74	1.22, 5.06	0.77, 5.68

In case of a Poisson variable n_0
in presence of background

The frequentist unified approach
(Feldman and Cousins PRD 57 3873 (1998))

“standard” confidence belt
b=3 case

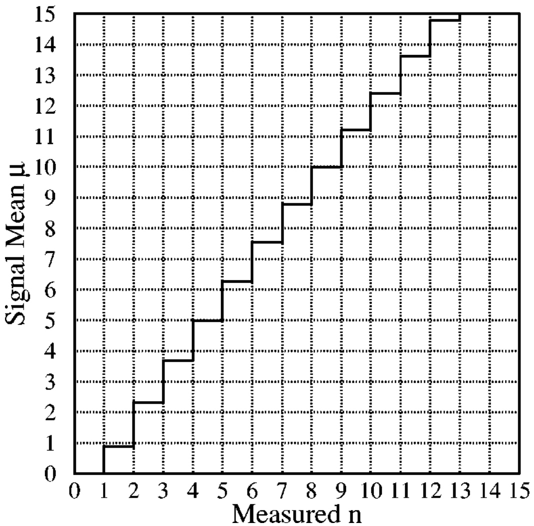


FIG. 5. Standard confidence belt for 90% C.L. upper limits, for unknown Poisson signal mean μ in the presence of a Poisson background with known mean $b=3.0$. The second line in the belt is at $n=+\infty$.

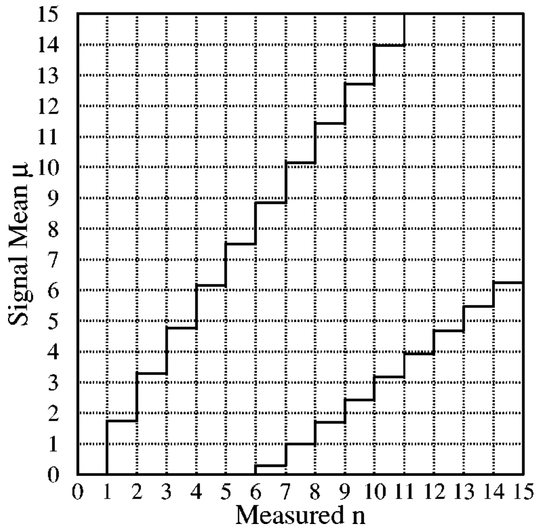


FIG. 6. Standard confidence belt for 90% C.L. central confidence intervals, for unknown Poisson signal mean μ in the presence of a Poisson background with known mean $b=3.0$.

In case of a Poisson variable n_0
in presence of background

The frequentist unified approach
(Feldman and Cousins PRD 57 3873 (1998))

“standard” confidence belt
b=3 case

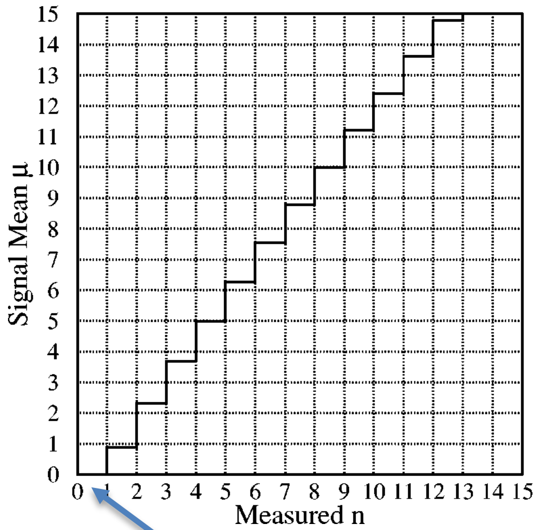


FIG. 5. Standard confidence belt for 90% C.L. upper limits, for unknown Poisson signal mean μ in the presence of a Poisson background with known mean $b=3.0$. The second line in the belt is at $n=+\infty$.

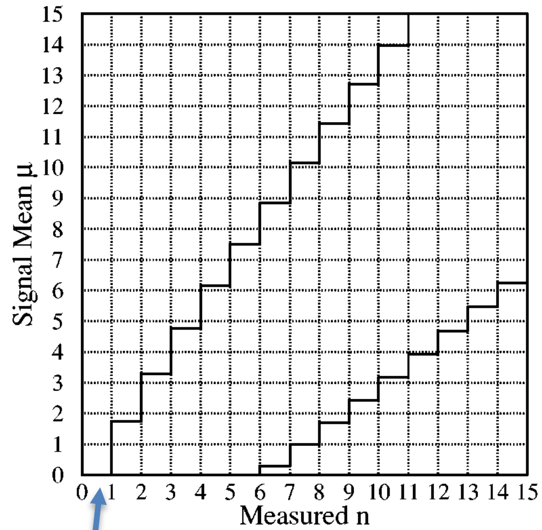


FIG. 6. Standard confidence belt for 90% C.L. central confidence intervals for unknown Poisson signal mean μ in the presence of a Poisson background with known mean $b=3.0$.

What for measured n=0?

In case of a Poisson variable n_0 in presence of background

The frequentist unified approach (Feldman and Cousins PRD 57 3873 (1998))

Feldman Cousins approach (ordering criteria) for confidence belt b=3 case

Let the known mean background be $b=3.0$, and consider the construction of the horizontal acceptance interval at signal mean $\mu=0.5$. Then $P(n|\mu)$ is given by Eq. (3.2), and is given in the second column of Table I.

Now consider, for example, $n=0$. For the assumed $b=3.$, the probability of obtaining 0 events is 0.03 if $\mu=0.5$, which is quite low on an absolute scale. However, it is not so low when compared to the probability (0.05) of obtaining 0 events with $b=3.0$ and $\mu=0.0$, which is the alternate hypothesis with the greatest likelihood. A *ratio* of likelihoods, in this case 0.03/0.05, is what we use as our ordering principle when selecting those values of n to place in the acceptance interval.

That is, for each n, we let $\mu_{\rm best}$ be that value of the mean signal μ which maximizes $P(n|\mu)$; we require $\mu_{\rm best}$ to be physically allowed, i.e., non-negative in this case. Then $\mu_{\rm best}=\max(0,n-b)$, and is given in the third column of Table I. We then compute $P(n|\mu_{\rm best})$, which is given in the fourth column. The fifth column contains the ratio

$$R=P(n|\mu)/P(n|\mu_{\rm best}), \tag{4.1}$$

and is the quantity on which our ordering principle is based. R is a ratio of two likelihoods: the likelihood of obtaining n given the actual mean μ, and the likelihood of obtaining n given the best-fit physically allowed mean. Values of n are added to the acceptance region for a given μ in decreasing order of R, until the sum of $P(n|\mu)$ meets or exceeds the desired C.L. This ordering, for values of n necessary to obtain total probability of 90%, is shown in the column labeled ''rank.'' Thus, the acceptance region for $\mu=0.5$ (analogous to a horizontal line segment in Fig. 1) is the interval $n=[0,6]$. Because of the discreteness of n, the acceptance region contains a summed probability greater than 90%; this is unavoidable no matter what the ordering principle, and leads to confidence intervals which are conservative.

For comparison, in the column of Table I labeled ''U.L.,'' we place check marks at the values of n which are in the acceptance region of standard 90% C.L. upper limits for this example, and in the column labeled ''central,'' we place check marks at the values of n which are in the acceptance region of standard 90% C.L central confidence intervals.

The construction proceeds by finding the acceptance region for all values of μ, for the given value of b. With a computer, we perform the construction on a grid of discrete values of μ, in the interval [0, 50] in steps of 0.005. This suffices for the precision desired (0.01) in the end points of confidence intervals. We find that a mild pathology arises as a result of the fact that the observable n is discrete. When the vertical dashed line is drawn at some n_0 (in analogy with in Fig. 1), it can happen that the set of intersected horizontal line segments is not simply connected. When this occurs we naturally take the confidence interval to have μ_1 corresponding to the bottommost segment intersected, and to have μ_2 corresponding to the topmost segment intersected.

We then repeat the construction for a selection of fixed values of b. We find an additional mild pathology, again caused by the discreteness in n: when we compare the results for different values of b for fixed n_0, the upper end point μ_2 is not always a decreasing function of b, as would be expected. When this happens, we force the function to be nonincreasing, by lengthening selected confidence intervals as necessary. We have investigated this behavior, and compensated for it, over a fine grid of b in the range [0, 25] in increments of 0.001 (with some additional searching to even finer precision).

Our compensation for the two pathologies mentioned in the previous paragraphs adds slightly to our intervals' conservatism, which however remains dominated by the unavoidable effects due to the discreteness in n.

In case of a Poisson variable n_0 in presence of background

The frequentist unified approach (Feldman and Cousins PRD 57 3873 (1998))

Feldman Cousins approach (ordering criteria) for confidence belt b=3 case

TABLE I. Illustrative calculations in the confidence belt construction for signal mean μ in the presence of known mean background $b=3.0$. Here we find the acceptance interval for $\mu=0.5$.

n	$P(n\|\mu)$	μ_{best}	$P(n\|\mu_{best})$	R	rank	U.L.	central
0	0.030	0.0	0.050	0.607	6		
1	0.106	0.0	0.149	0.708	5	√	√
2	0.185	0.0	0.224	0.826	3	√	√
3	0.216	0.0	0.224	0.963	2	√	√
4	0.189	1.0	0.195	0.966	1	√	√
5	0.132	2.0	0.175	0.753	4	√	√
6	0.077	3.0	0.161	0.480	7	√	√
7	0.039	4.0	0.149	0.259		√	√
8	0.017	5.0	0.140	0.121		√	
9	0.007	6.0	0.132	0.050		√	
10	0.002	7.0	0.125	0.018		√	
11	0.001	8.0	0.119	0.006		√	

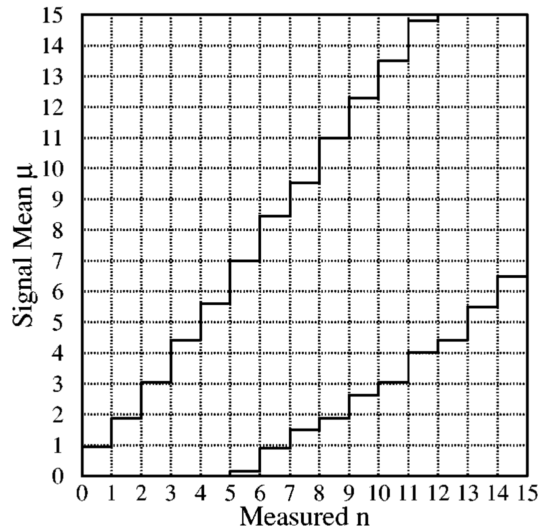


FIG. 7. Confidence belt based on our ordering principle, for 90% C.L. confidence intervals for unknown Poisson signal mean μ in the presence of a Poisson background with known mean $b=3.0$.

In case of a Poisson variable n_0 in presence of background

The frequentist unified approach (Feldman and Cousins PRD 57 3873 (1998))

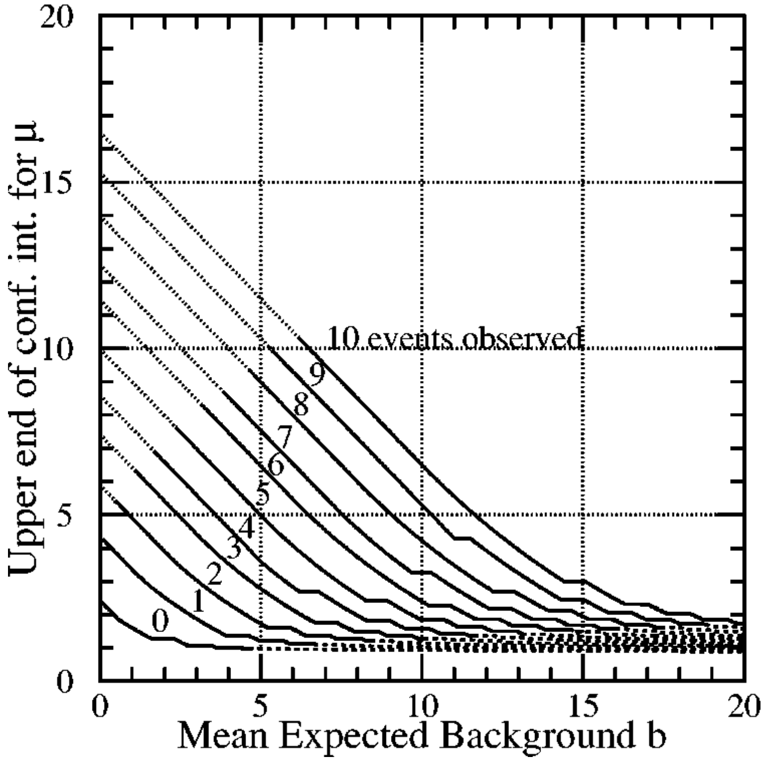


FIG. 8. Upper end μ_2 of our 90% C.L. confidence intervals $[\mu_1, \mu_2]$, for unknown Poisson signal mean μ in the presence of an expected Poisson background with known mean b. The curves for the cases n_0 from 0 through 10 are plotted. Dotted portions on the upper left indicate regions where μ_1 is non-zero (and shown in the following figure). Dashed portions in the lower right indicate regions where the probability of obtaining the number of events observed or fewer is less than 1%, even if $\mu = 0$.

In case of a Poisson variable n_0 in presence of background

The frequentist unified approach (Feldman and Cousins PRD 57 3873 (1998))

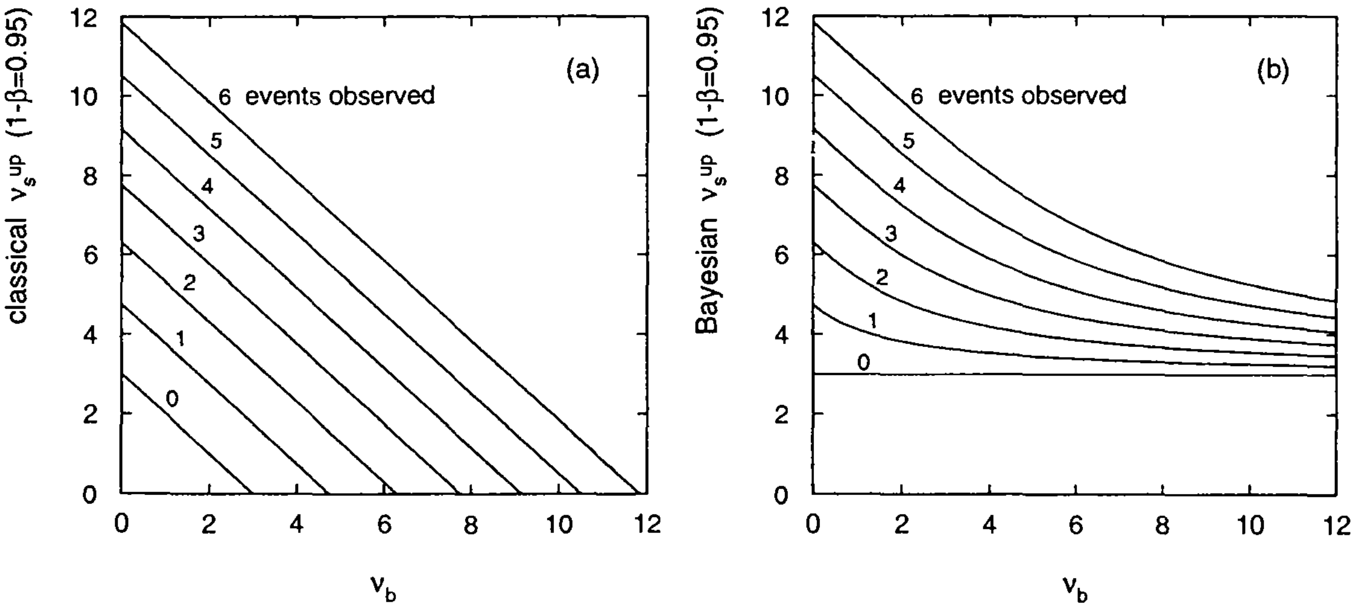


Fig. 9.9 Upper limits ν_s^{up} at a confidence level of $1 - \beta = 0.95$ for different numbers of events observed n_{obs} and as a function of the expected number of background events ν_b. (a) The classical limit. (b) The Bayesian limit based on a uniform prior density for ν_s.

In case of a Poisson variable n_0 in presence of background

The frequentist unified approach (Feldman and Cousins PRD 57 3873 (1998))

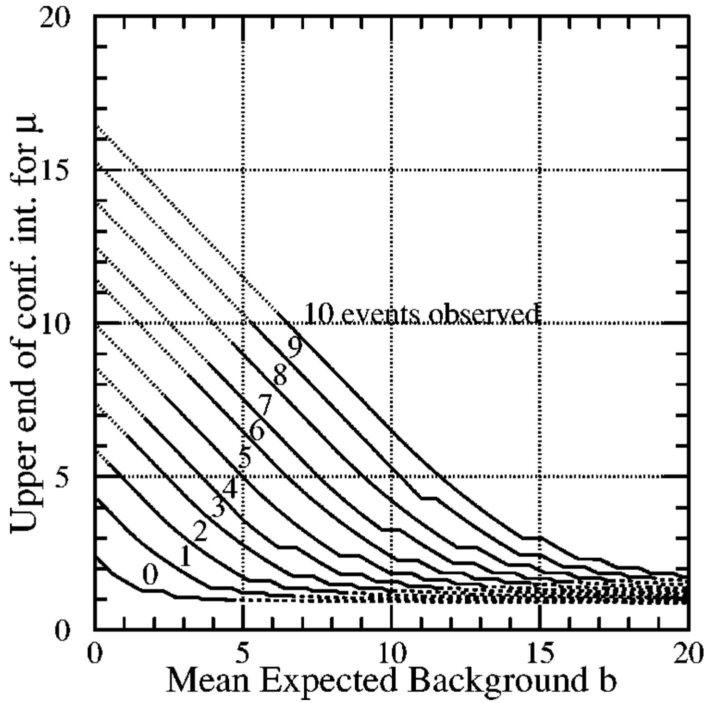


FIG. 8. Upper end μ_2 of our 90% C.L. confidence intervals $[\mu_1, \mu_2]$, for unknown Poisson signal mean μ in the presence of an expected Poisson background with known mean b. The curves for the cases n_0 from 0 through 10 are plotted. Dotted portions on the upper left indicate regions where μ_1 is non-zero (and shown in the following figure). Dashed portions in the lower right indicate regions where the probability of obtaining the number of events observed or fewer is less than 1%, even if $\mu=0$.

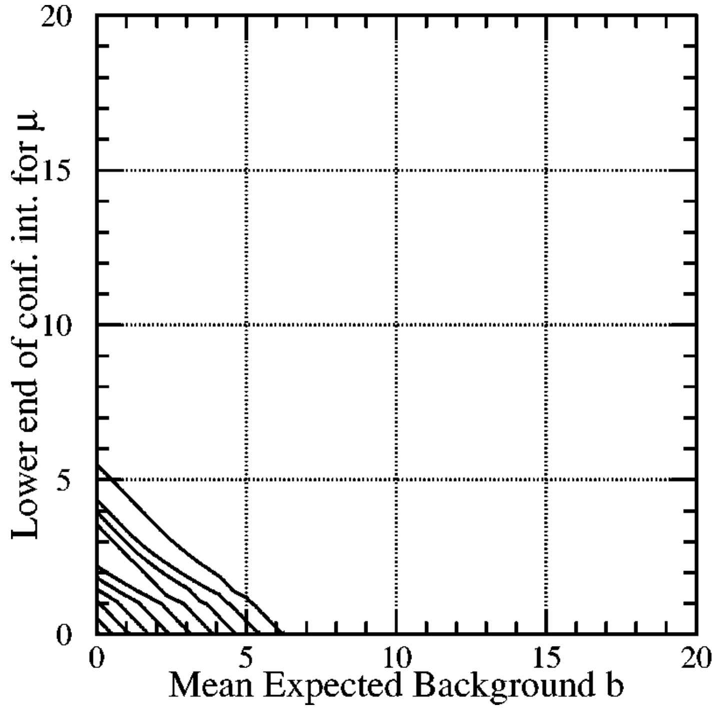


FIG. 9. Lower end μ_1 of our 90% C.L. confidence intervals $[\mu_1, \mu_2]$, for unknown Poisson signal mean μ in the presence of an expected Poisson background with known mean b. The curves correspond to the dotted regions in the plots of μ_2 of the previous figure, with again $n_0=10$ for the upper right curve, etc.

In case of a Poisson variable n_0 in presence of background

The frequentist unified approach (Feldman and Cousins PRD 57 3873 (1998))

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

In case of a Poisson variable n_0 in presence of background

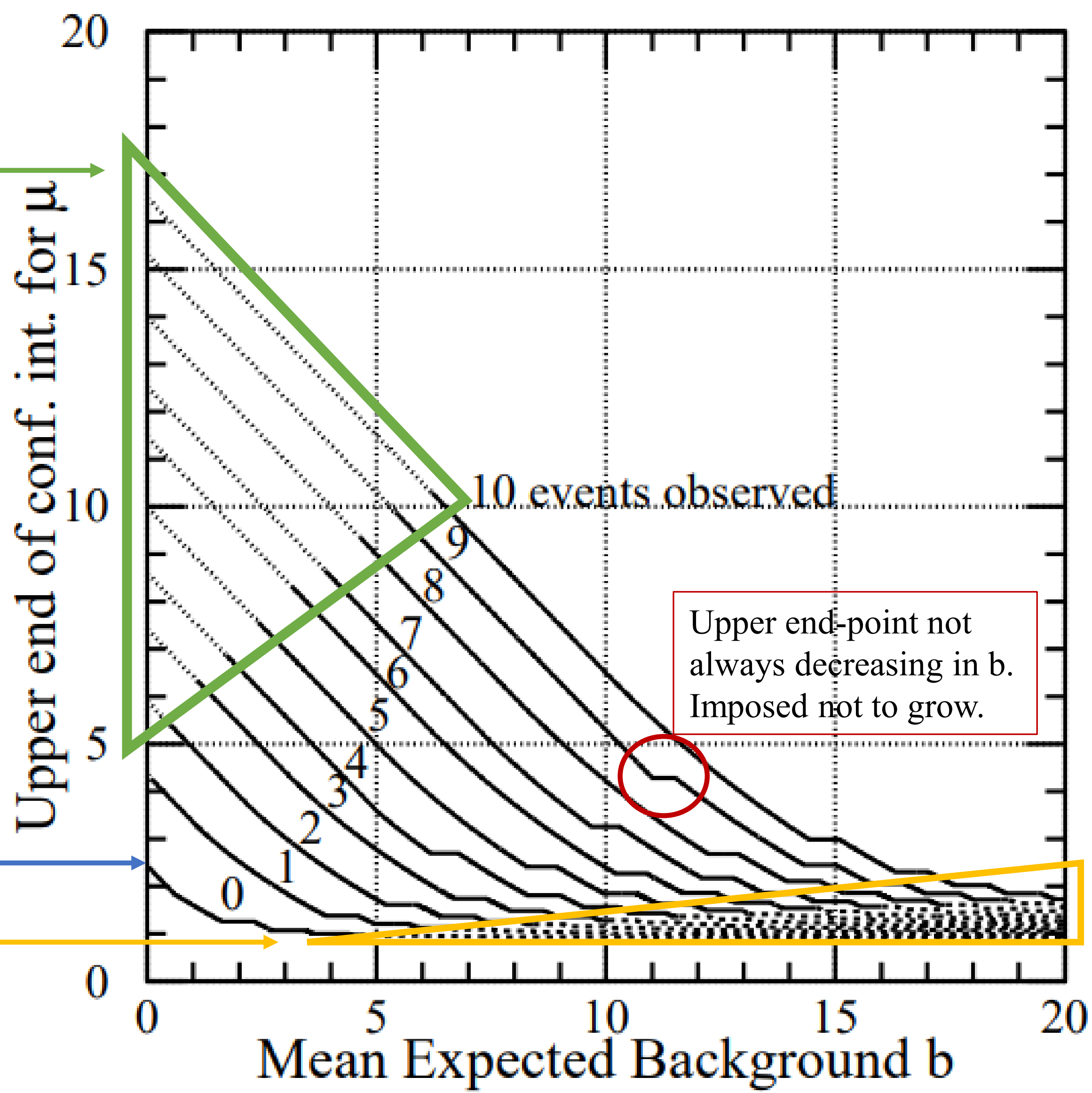
The frequentist unified approach (Feldman and Cousins PRD 57 3873 (1998))

The dotted lines means there is also a lower limit, not only an upper one

In the classical case, the upper limit on zero-counting without any background is 2.3, with 90% C.L.

The dashed portions indicate regions where the probability of obtaining the number of events observed or fewer is less than 1%

(very unlikely configuration, small n_0 with large b)



Homework n.6

The squared energy and momentum of a particle are independently measured:

$E^2 = 1010 \pm 17$ eV^2
$P^2 = 1064 \pm 25$ eV^2

Put an upper limit on the squared mass
$m^2 = E^2 - P^2$
of the particle using:

- The classical frequentist approach
- The unified approach (Feldman Cousins) with the mean of the Gaussian constrained to be non-negative
  (see Feldman and Cousins Phys.Rev.D 57 3873 (1998))
- The Bayesian approach (briefly comment the choice of the prior)