Statistical analysis of experimental data
Parameter Inference (2)

Aleksander Filip Żarnecki

FACULTY OF PHYSICS
UNIVERSITY OF WARSAW

Lecture 07

November 14, 2024

Parameter Inference (2)

1. Frequentist confidence intervals
2. Bayesian limits
3. Unified approach
4. Homework

Parameter covariance matrix

For the considered case of multivariate normal distribution, best parameter estimates $\hat{\boldsymbol{\lambda}}$ are given by the measured variable values $\mathbf{x}$.

Unlike parameters $\boldsymbol{\lambda}$, parameter estimates $\hat{\boldsymbol{\lambda}}$ are random variables (functions of $\mathbf{x}$) and so we can consider covariance matrix for $\hat{\boldsymbol{\lambda}}$:

$$\mathbb{C}_{\mathbf{x}} = \mathbb{C}_{\hat{\boldsymbol{\lambda}}} = \left(- \frac{\partial^2 \ell}{\partial \lambda_i \, \partial \lambda_j} \right)^{-1}$$

Knowing the likelihood function, we can not only estimate parameter values, but also extract uncertainties and correlations of these estimates!

For the uncorrelated parameters (diagonal covariance matrix):

$$\sigma_{\hat{\boldsymbol{\lambda}}_i} = \left(- \frac{\partial^2 \ell}{\partial \lambda_i^2} \right)^{-1/2}$$

Parameter covariance matrix

Considered example was based on the Gaussian distribution.
Standard deviation is one of the parameters of the p.d.f., can be easily extracted from log-likelihood:

$$\sigma_i = \sqrt{\mathbb{C}_{ii}}$$

However, this procedure works only in the Gaussian approximation.
How to define parameter uncertainty in the general case?

Recipe for a parameter uncertainty

G. Bohm and G. Zech

Standard error intervals of the extracted parameter are defined by the decrease of the log-likelihood function by 0.5 for one, by 2 for two and by 4.5 for three standard deviations.

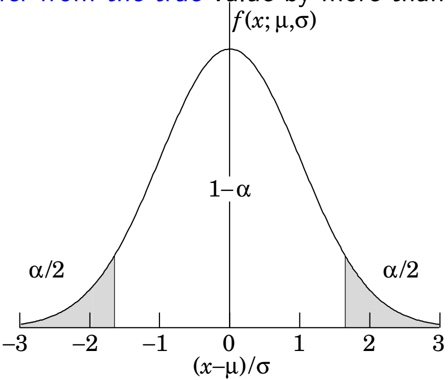
This definition works for arbitrary p.d.f. shape, also for multiple parameters

Normal distribution

Meaning of σ is well defined for Gaussian distribution.

Probability for the experimental result to differ from the true value by more than $N\sigma$:

		α	
$\pm 1\ \sigma$	$\Rightarrow$	31.73	%
$\pm 2\ \sigma$	$\Rightarrow$	4.55	%
$\pm 3\ \sigma$	$\Rightarrow$	0.27	%
$\pm 4\ \sigma$	$\Rightarrow$	0.0063	%
$\pm 5\ \sigma$	$\Rightarrow$	0.000057	%

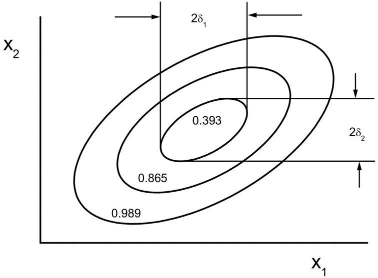


Fluctuations up and down are observed with equal probability...

Normal distribution in N-D

It is also important to notice that the fractions presented previously (eg. 68% within $\pm 1\sigma$) refer to one-dimensional normal distribution only!

If we consider 2-D distribution



Less than 40% is contained inside 1σ contour...

Fractions within $N\sigma$ contours:

Deviation	Dimension			
	1	2	3	4
1 σ	0.683	0.393	0.199	0.090
2 σ	0.954	0.865	0.739	0.594
3 σ	0.997	0.989	0.971	0.939
4 σ	1.	1.	0.999	0.997

1σ fraction above 50% only for N=1 !

G. Bohm and G. Zech

Interpreting results

So far we have only considered distribution of experimental results for given probability distribution, $f(\mathbf{x}; \boldsymbol{\lambda})$, when the parameter values $\boldsymbol{\lambda}$ are known.

Interpreting results

So far we have only considered distribution of experimental results for given probability distribution, $f(\mathbf{x};\boldsymbol{\lambda})$, when the parameter values $\boldsymbol{\lambda}$ are known.

The actual situation is usually different: for given set of measurements $\mathbf{x}$ we extract estimates of the parameter values $\hat{\boldsymbol{\lambda}}$.

Uncertainties estimated from log-likelihood variation indicate the expected level of agreement (in Gaussian approximation) between our estimate $\hat{\boldsymbol{\lambda}}$ and the true parameter values $\boldsymbol{\lambda}$.

Can we present measurement results in a way which gives us more precise information about the possible fluctuations in the estimate $\hat{\boldsymbol{\lambda}}$?

Interpreting results

So far we have only considered distribution of experimental results for given probability distribution, $f(\mathbf{x};\boldsymbol{\lambda})$, when the parameter values $\boldsymbol{\lambda}$ are known.

The actual situation is usually different: for given set of measurements $\mathbf{x}$ we extract estimates of the parameter values $\hat{\boldsymbol{\lambda}}$.

Uncertainties estimated from log-likelihood variation indicate the expected level of agreement (in Gaussian approximation) between our estimate $\hat{\boldsymbol{\lambda}}$ and the true parameter values $\boldsymbol{\lambda}$.

Can we present measurement results in a way which gives us more precise information about the possible fluctuations in the estimate $\hat{\boldsymbol{\lambda}}$?

Yes, but we need to define the problem differently... We should not consider probability of $\hat{\boldsymbol{\lambda}}$...

Frequentist confidence intervals

Classical (frequentist) definition of the confidence interval refers directly to the probability distribution of the experimental results, $f(\mathbf{x}; \boldsymbol{\lambda})$.

We do not try to make any prediction (nor guess) about the "probability" (degree of belief) of given parameter value $\boldsymbol{\lambda}$. This is the Bayesian concept we will discuss later...

In the frequentist approach we consider individually each $\boldsymbol{\lambda}$ value.
Given value of $\boldsymbol{\lambda}$ is allowed (on given confidence level, C.L.), if the actual outcome of our experiment, x_m, is within the corresponding **probability interval** for **variable x** for this value of $\boldsymbol{\lambda}$.

Frequentist confidence intervals

Classical (frequentist) definition of the confidence interval refers directly to the probability distribution of the experimental results, $f(\mathbf{x};\boldsymbol{\lambda})$.

We do not try to make any prediction (nor guess) about the "probability" (degree of belief) of given parameter value $\boldsymbol{\lambda}$. This is the Bayesian concept we will discuss later...

In the frequentist approach we consider individually each $\boldsymbol{\lambda}$ value.
Given value of λ is allowed (on given confidence level, C.L.), if the actual outcome of our experiment, x_m, is within the corresponding **probability interval** for **variable x** for this value of $\boldsymbol{\lambda}$.

This definition clearly depends on the way we define probability intervals for $\mathbf{x}$.
So this is rather a general concept, more assumptions are needed.
We always refer to probability distribution $f(\mathbf{x};\boldsymbol{\lambda})$ for random variable $\mathbf{x}$!

Frequentist confidence intervals

As mentioned above, to define confidence interval for our parameter, we need to define how the **probability interval** for our measurement is defined.

There are three "natural" choices: CL = 1 - α

- We constrain the measurement from above, define upper limit x_{ul}: $\alpha \ll 1$

$$\int_{-\infty}^{x_{ul}} dx\ f(x;\lambda) = CL \quad \text{or} \quad \int_{x_{ul}}^{+\infty} dx\ f(x;\lambda) = \alpha$$

- We constrain the measurement from below, define lower limit x_{ll}:

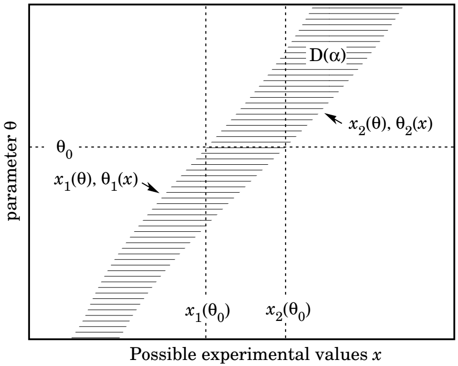
$$\int_{-\infty}^{x_{ll}} dx\ f(x;\lambda) = \alpha \quad \text{or} \quad \int_{x_{ll}}^{+\infty} dx\ f(x;\lambda) = CL$$

- We define central probability interval $[x_1, x_2]$: as presented for Gaussian pdf

$$\int_{-\infty}^{x_1} dx\ f(x;\lambda) \ = \ \alpha/2 \quad \text{and} \quad \int_{x_2}^{+\infty} dx\ f(x;\lambda) \ = \ \alpha/2$$

Frequentist confidence intervals

General procedure



- calculate limits of probability intervals for x, $x_1(\theta)$ and $x_2(\theta)$, for different values of θ
- calculated intervals define the "accepted region" in (θ, x)
- confidence interval for θ is defined by drawing line $x = x_m$ in the accepted region

$\Rightarrow$ limit on θ for given x_m, $\theta_1(x_m)$, corresponds to limit on x for given θ: $x_m = x_1(\theta_1)$.

R.L. Workman et al. (Particle Data Group), Prog. Theor. Exp. Phys. 2022, 083C01 (2022), PDG web page

Parameter Inference (2)

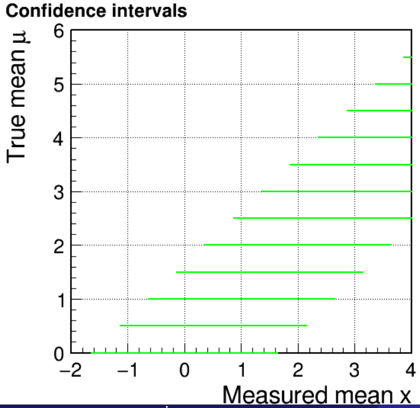
1. Frequentist confidence intervals
2. Bayesian limits
3. Unified approach
4. Homework

Procedure

07_gauss_interval.ipynb Open in Colab

Let us consider the 90% CL interval (or 95% CL limits) for Gaussian pdf: width fixed $\sigma \equiv 1$

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



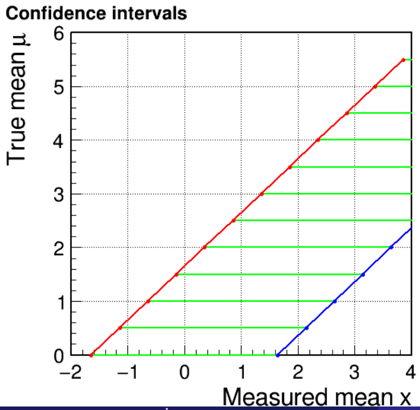
- calculate limits of probability intervals for x: $x_1(\mu)$ and $x_2(\mu)$, for different values of μ

Procedure

07_gauss_interval.ipynb Open in Colab

Let us consider the 90% CL interval (or 95% CL limits) for Gaussian pdf: width fixed $\sigma \equiv 1$

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



- calculate limits of probability intervals for x: $x_1(\mu)$ and $x_2(\mu)$, for different values of μ
- calculated intervals define the "accepted region" in the (μ, x) plane

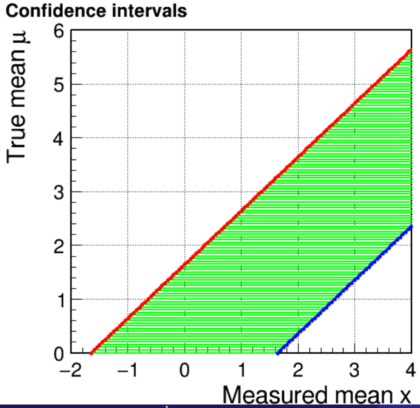
Frequentist confidence intervals

Procedure

07_gauss_interval.ipynb Open in Colab

Let us consider the 90% CL interval (or 95% CL limits) for Gaussian pdf: width fixed $\sigma \equiv 1$

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



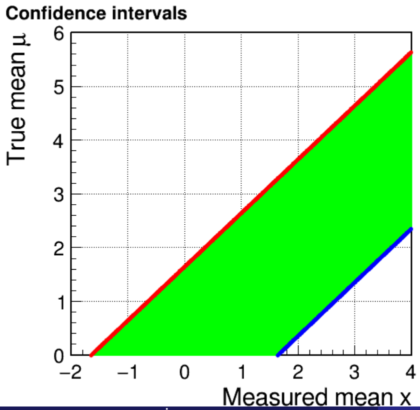
- calculate limits of probability intervals for x: $x_1(\mu)$ and $x_2(\mu)$, for different values of μ
- calculated intervals define the “accepted region” in the (μ, x) plane

Procedure

07_gauss_interval.ipynb Open in Colab

Let us consider the 90% CL interval (or 95% CL limits) for Gaussian pdf: width fixed $\sigma \equiv 1$

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



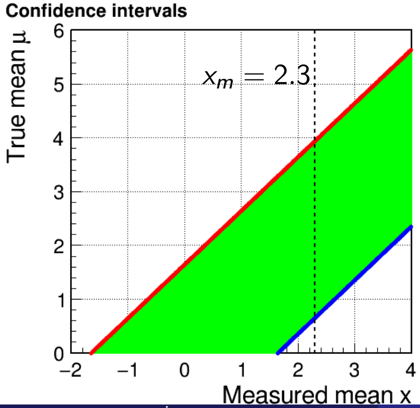
- calculate limits of probability intervals for x: $x_1(\mu)$ and $x_2(\mu)$, for different values of μ
- calculated intervals define the "accepted region" in the (μ, x) plane

Procedure

07_gauss_interval.ipynb Open in Colab

Let us consider the 90% CL interval (or 95% CL limits) for Gaussian pdf: width fixed $\sigma \equiv 1$

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



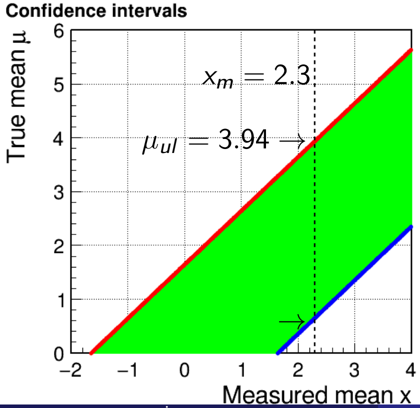
- calculate limits of probability intervals for x: $x_1(\mu)$ and $x_2(\mu)$, for different values of μ
- calculated intervals define the "accepted region" in the (μ, x) plane
- confidence interval for μ is found by drawing line $x = x_m$ in the accepted region

Procedure

07_gauss_interval.ipynb Open in Colab

Let us consider the 90% CL interval (or 95% CL limits) for Gaussian pdf: width fixed $\sigma \equiv 1$

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



- calculate limits of probability intervals for x: $x_1(\mu)$ and $x_2(\mu)$, for different values of μ
- calculated intervals define the “accepted region” in the (μ, x) plane
- confidence interval for μ is found by drawing line $x = x_m$ in the accepted region

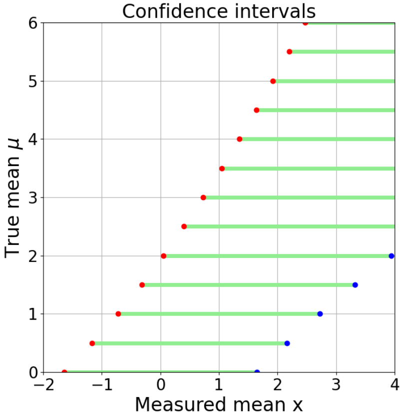
$\Rightarrow$ limit on μ for given x_m, $\mu_1(x_m)$, corresponds to the probability limit on x for given μ: $x_m = x_1(\mu_1)$.

Procedure

07_gauss_interval2.ipynb Open in Colab

The procedure can be easily used also for Gauss with variable σ: $CL = 90\%$

$$\sigma^2(\mu) \quad = \quad 1 + 0.1 \cdot \mu^2$$



- calculate limits of probability intervals for x: $x_1(\mu)$ and $x_2(\mu)$, for different values of μ

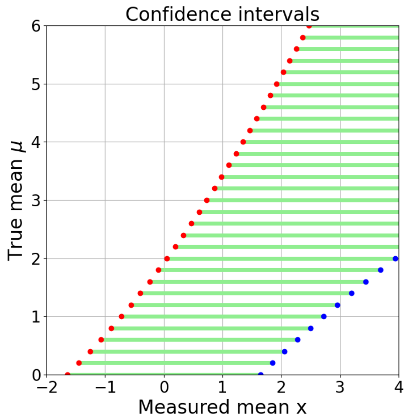
Procedure

07_gauss_interval2.ipynb Open in Colab

The procedure can be easily used also for Gauss with variable σ:

CL = 90%

$$\sigma^2(\mu) \quad = \quad 1 + 0.1 \cdot \mu^2$$



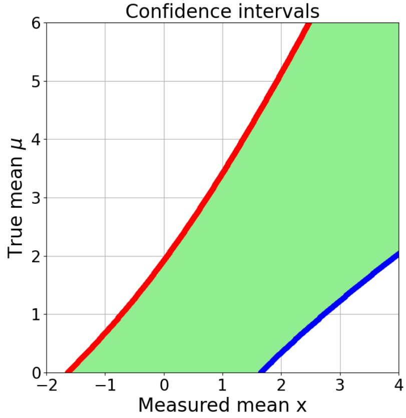
- calculate limits of probability intervals for x: $x_1(\mu)$ and $x_2(\mu)$, for different values of μ

Procedure

07_gauss_interval2.ipynb Open in Colab

The procedure can be easily used also for Gauss with variable σ: CL = 90%

$$\sigma^2(\mu) = 1 + 0.1 \cdot \mu^2$$



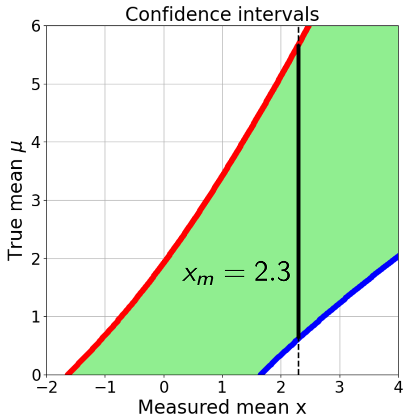
- calculate limits of probability intervals for x: $x_1(\mu)$ and $x_2(\mu)$, for different values of μ
- calculated intervals define the "accepted region" in the (μ, x) plane

Procedure

07_gauss_interval2.ipynb Open in Colab

The procedure can be easily used also for Gauss with variable σ: $CL = 90\%$

$$\sigma^2(\mu) = 1 + 0.1 \cdot \mu^2$$



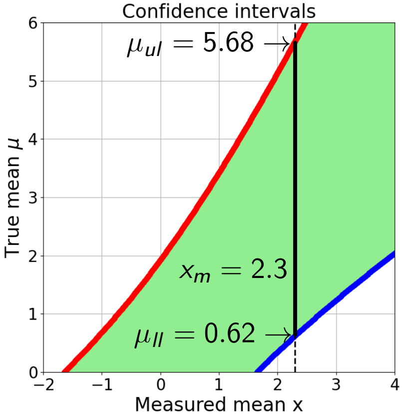
- calculate limits of probability intervals for x: $x_1(\mu)$ and $x_2(\mu)$, for different values of μ
- calculated intervals define the "accepted region" in the (μ, x) plane
- confidence interval for μ is found by drawing line $x = x_m$ in the accepted region

Procedure

07_gauss_interval2.ipynb Open in Colab

The procedure can be easily used also for Gauss with variable σ: $\mathrm{CL} = 90\%$

$$\sigma^2(\mu) \quad = \quad 1 + 0.1 \cdot \mu^2$$



- calculate limits of probability intervals for x: $x_1(\mu)$ and $x_2(\mu)$, for different values of μ
- calculated intervals define the "accepted region" in the (μ, x) plane
- confidence interval for μ is found by drawing line $x = x_m$ in the accepted region

$\Rightarrow$ limit on μ for given x_m, $\mu_1(x_m)$, corresponds to the probability limit on x for given μ: $x_m = x_1(\mu_1)$.

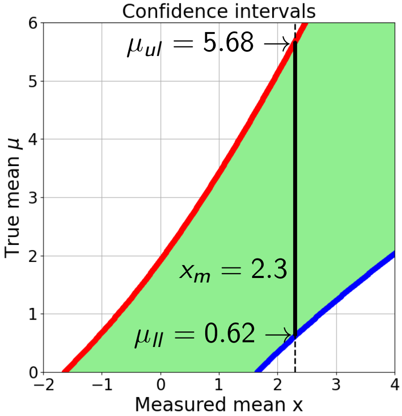
Procedure

07_gauss_interval2.ipynb Open in Colab

The procedure can be easily used also for Gauss with variable σ:

CL = 90%

$$\sigma^2(\mu) \quad = \quad 1 + 0.1 \cdot \mu^2$$



- calculate limits of probability intervals for x: $x_1(\mu)$ and $x_2(\mu)$, for different values of μ
- calculated intervals define the "accepted region" in the (μ, x) plane
- confidence interval for μ is found by drawing line $x = x_m$ in the accepted region

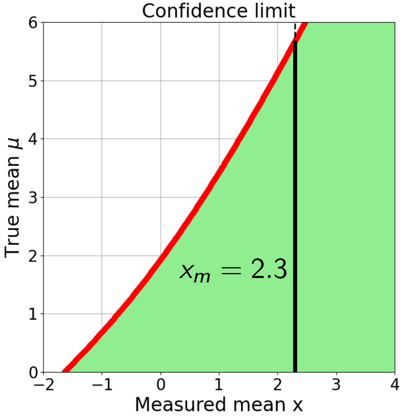
$\Rightarrow$ limit on μ for given x_m, $\mu_1(x_m)$, corresponds to the probability limit on x for given μ: $x_m = x_1(\mu_1)$.

Naive estimate, just taking x_m and $\sigma(x_m)$: [0.27,4.33]

Procedure

07_gauss_interval2_ul.ipynb Open in Colab

When considering one side (upper or lower) parameter limits (quite a common case) the procedure can be simplified. For upper limit (95% CL): F - cumulative distribution function



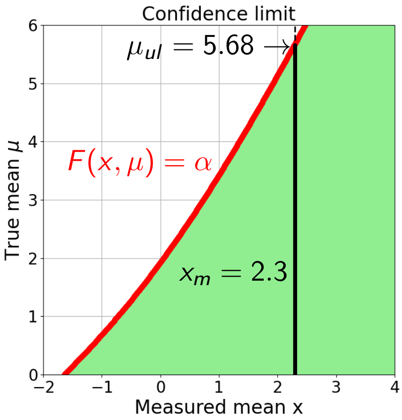
- for different values of μ, consider the probability of the experimental result $x < x_m$ (consistent with the measurement): $P(x < x_m; \mu) = F(x_m, \mu)$

Procedure

07_gauss_interval2_ul.ipynb Open in Colab

When considering one side (upper or lower) parameter limits (quite a common case) the procedure can be simplified. For upper limit (95% CL):

F - cumulative distribution function



- for different values of μ, consider the probability of the experimental result $x < x_m$ (consistent with the measurement): $P(x < x_m; \mu) = F(x_m, \mu)$
- scan parameter μ to find the value corresponding to:

$$P(x < x_m; \mu_{ul}) = \alpha$$

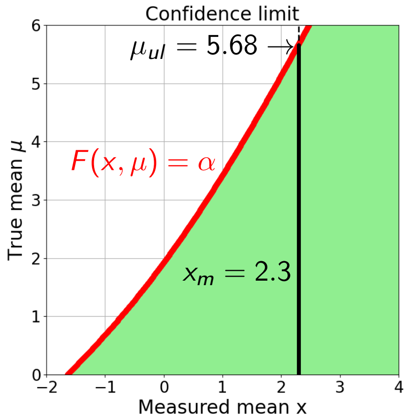
crossing of $F(x, \mu) = \alpha$ curve with $x = x_m$ one

Procedure

07_gauss_interval2_ul.ipynb Open in Colab

When considering one side (upper or lower) parameter limits (quite a common case) the procedure can be simplified. For upper limit (95% CL):

F - cumulative distribution function



- for different values of μ, consider the probability of the experimental result $x < x_m$ (consistent with the measurement): $P(x < x_m; \mu) = F(x_m, \mu)$
- scan parameter μ to find the value corresponding to:

$$P(x < x_m; \mu_{ul}) = \alpha$$

$\Rightarrow$ For higher parameter values, $\mu' > \mu_{ul}$, probability of reproducing experimental result:

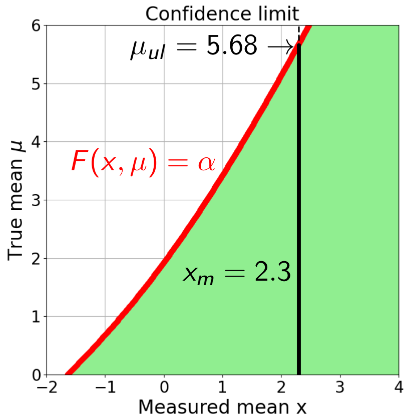
$$P(x < x_m; \mu') < \alpha$$

Procedure

07_gauss_interval2_ul.ipynb Open in Colab

When considering one side (upper or lower) parameter limits (quite a common case) the procedure can be simplified. For upper limit (95% CL):

F - cumulative distribution function



- for different values of μ, consider the probability of the experimental result $x < x_m$ (consistent with the measurement): $P(x < x_m; \mu) = F(x_m, \mu)$
- scan parameter μ to find the value corresponding to:
$$P(x < x_m; \mu_{ul}) = \alpha$$

$\Rightarrow$ For higher parameter values, $\mu' > \mu_{ul}$, probability of reproducing experimental result:
$$P(x < x_m; \mu') < \alpha$$

Check: $x_m = \mu_{ul} - 1.64 \cdot \sigma(\mu_{ul})$

Limit setting

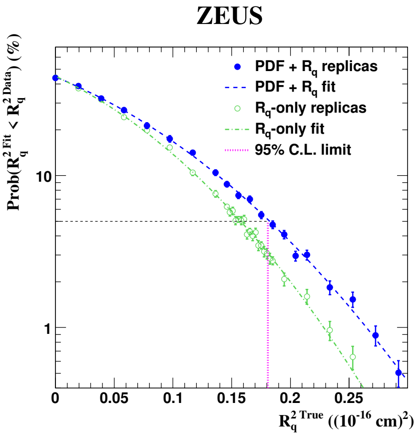
The probability of obtaining a $R_q^{2\,\mathrm{Fit}}$ value smaller than that obtained for the actual data

$$\mathrm{Prob}(R_q^{2\,\mathrm{Fit}} < R_q^{2\,\mathrm{Data}})$$

is studied as a function of $R_q^{2\,\mathrm{True}}$

$R_q^{2\,\mathrm{True}}$ values corresponding to the probability smaller than 5% are excluded at the 95% C.L.

Limits obtained for fixed SM parameters are too strong by about 10%



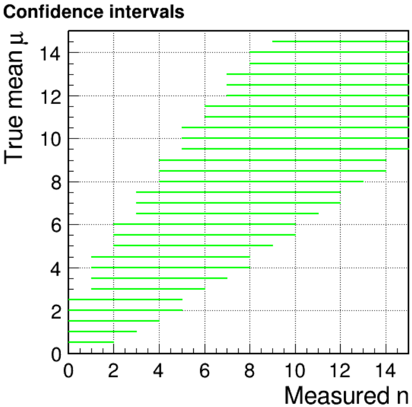
$$R_q < 0.43 \cdot 10^{-16}\mathrm{cm}$$

Procedure

07_poisson_interval.ipynb Open in Colab

The procedure can be also adapted for the counting experiment, Poisson distribution:

$$P(n;\mu) \quad = \quad \frac{\mu^n \; e^{-\mu}}{n!} \qquad \text{for } n = 0, 1, 2, \ldots$$



- calculate probability intervals for n for different values of μ

! As n is discrete random variable, we can not guarantee exact "coverage". The requirement is:

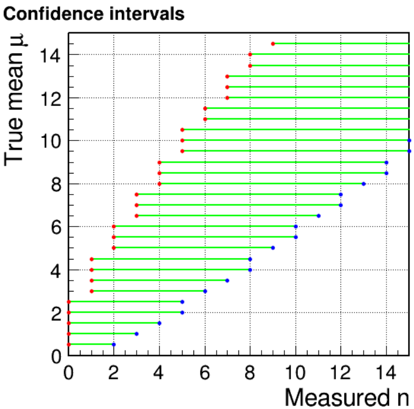
$$P(n_1(\mu) \leq n \leq n_2(\mu)) \geq 1 - \alpha$$

Procedure

07_poisson_interval.ipynb Open in Colab

The procedure can be also adapted for the counting experiment, Poisson distribution:

$$P(n;\mu) \quad = \quad \frac{\mu^n \, e^{-\mu}}{n!} \qquad \text{for } n = 0, 1, 2, \ldots$$



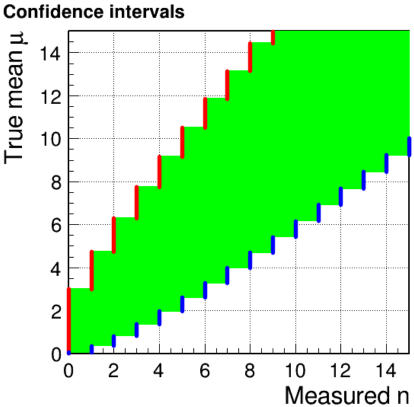
- calculate probability intervals for n for different values of μ
- calculated intervals define the "accepted region" in the (μ, n) plane

Procedure

07_poisson_interval.ipynb Open in Colab

The procedure can be also adapted for the counting experiment, Poisson distribution:

$$P(n;\mu) = \frac{\mu^n \, e^{-\mu}}{n!} \quad \text{for } n = 0, 1, 2, \ldots$$



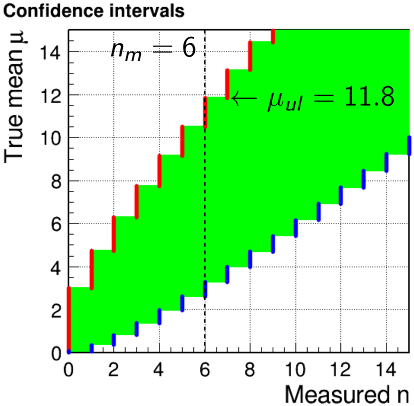
- calculate probability intervals for n for different values of μ
- calculated intervals define the "accepted region" in the (μ, n) plane

Procedure

07_poisson_interval.ipynb Open in Colab

The procedure can be also adapted for the counting experiment, Poisson distribution:

$$P(n;\mu) \quad = \quad \frac{\mu^n \, e^{-\mu}}{n!} \qquad \text{for } n = 0, 1, 2, \ldots$$



- calculate probability intervals for n for different values of μ
- calculated intervals define the "accepted region" in the (μ, n) plane
- confidence interval for μ is defined by drawing line $n = n_m$ in the accepted region (and taking maximal range)

Results

For the case of Poisson variable, calculation of the **upper limit** for the expected number of events μ, when observing n_m events, can be reduced to solving the equation:

$$P(n \leq n_m; \mu_{ul}) \;=\; \sum_{n=0}^{n_m} \frac{\mu_{ul}^{n}\, e^{-\mu_{ul}}}{n!} \;=\; \alpha$$

Results

For the case of Poisson variable, calculation of the **upper limit** for the expected number of events μ, when observing n_m events, can be reduced to solving the equation:

$$P(n \leq n_m; \mu_{ul}) \;=\; \sum_{n=0}^{n_m} \frac{\mu_{ul}^n \; e^{-\mu_{ul}}}{n!} \;=\; \alpha$$

For higher numbers of expected events $\mu' > \mu_{ul}$, probability that the repeated experiment will result in the measurement consistent with actual observation

$$P(n \leq n_m; \mu') \;<\; \alpha$$

$\Rightarrow$ these values are excluded on the assumed confidence level $(\mathrm{CL} = 1 - \alpha)$

Results

Lower and upper (one-sided) limits for the mean μ of a Poisson variable given n observed events in the absence of background, for confidence levels of 90% and 95%.

R.L. Workman et al. (Particle Data Group), Prog. Theor. Exp. Phys. 2022, 083C01 (2022)
PDG web page

	$1-\alpha=90\%$		$1-\alpha=95\%$	
n	$\mu_{\rm lo}$	$\mu_{\rm up}$	$\mu_{\rm lo}$	$\mu_{\rm up}$
0	–	2.30	–	3.00
1	0.105	3.89	0.051	4.74
2	0.532	5.32	0.355	6.30
3	1.10	6.68	0.818	7.75
4	1.74	7.99	1.37	9.15
5	2.43	9.27	1.97	10.51
6	3.15	10.53	2.61	11.84
7	3.89	11.77	3.29	13.15
8	4.66	12.99	3.98	14.43
9	5.43	14.21	4.70	15.71
10	6.22	15.41	5.43	16.96

Parameter Inference (2)

Bayesian approach

Bayes theorem can be used to generalize the concept of probability.

In particular, one can consider "probability" of given hypothesis H

(theoretical model or model parameter, eg. Hubble constant)

when taking into known outcome D (data) of the experiment

$$P(H|D) = \frac{P(D|H)}{P(D)} \cdot P(H)$$

There are two problems with this approach:

- H can not be considered an event, sampling space can not be defined (no experiment to repeat)
- we need to make a subjective assumption about the "prior" $P(H)$ describing our initial belief in hypothesis H

For these reasons I rather use term "degree of belief" for the result of the Bayesian procedure applied to non random events

Procedure

Bayes theorem can be applied to the case of counting experiment:

$$\mathcal{P}(\mu; n_m) = \frac{P(n_m; \mu)}{\int d\mu' \, P(n_m; \mu')} \cdot \mathcal{P}(\mu)$$

Procedure

Bayes theorem can be applied to the case of counting experiment:

$$\mathcal{P}(\mu; n_m) = \frac{P(n_m;\mu)}{\int d\mu' \; P(n_m;\mu')} \cdot \mathcal{P}(\mu)$$

Integral in the denominator is equal to 1 (Gamma distribution).
Assuming flat "prior distribution" for μ (no earlier constraints) we get:

$$\mathcal{P}(\mu; n) = \frac{\mu^n \; e^{-\mu}}{n!}$$

Procedure

Bayes theorem can be applied to the case of counting experiment:

$$\mathcal{P}(\mu; n_m) \quad = \quad \frac{P(n_m; \mu)}{\int d\mu' \; P(n_m; \mu')} \cdot \mathcal{P}(\mu)$$

Integral in the denominator is equal to 1 (Gamma distribution).

Assuming flat "prior distribution" for μ (no earlier constraints) we get:

$$\mathcal{P}(\mu; n) \quad = \quad \frac{\mu^n \; e^{-\mu}}{n!}$$

Upper limit on the expected number of events can be then calculated as:

$$\int_0^{\mu_{ul}} d\mu \;\; \mathcal{P}(\mu; n_m) \quad = \quad 1 - \alpha$$

Procedure

Bayes theorem can be applied to the case of counting experiment:

$$\mathcal{P}(\mu; n_m) \;=\; \frac{P(n_m;\mu)}{\int d\mu' \; P(n_m;\mu')} \cdot \mathcal{P}(\mu)$$

Integral in the denominator is equal to 1 (Gamma distribution).

Assuming flat "prior distribution" for μ (no earlier constraints) we get:

$$\mathcal{P}(\mu; n) \;=\; \frac{\mu^n \; e^{-\mu}}{n!}$$

Upper limit on the expected number of events can be then calculated as:

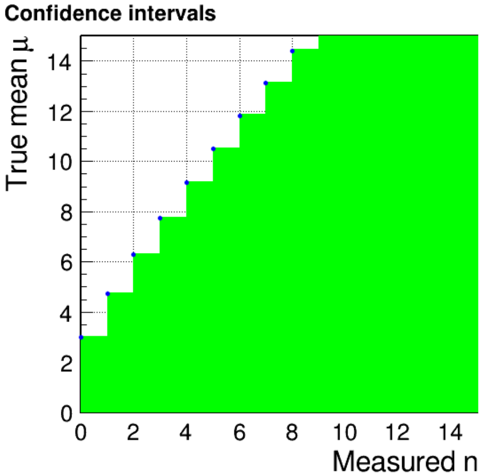
$$\int_0^{\mu_{ul}} d\mu \;\; \mathcal{P}(\mu; n_m) \;=\; 1-\alpha$$

Surprisingly, the numerical result is the same as for the Frequentist approach...

Numerical check

07_poisson_bayes.ipynb Open in Colab

Comparison of 95% C.L. upper limits from Frequentist approach (green) with corresponding limits obtained from Bayesian approach (blue).



Bayesian limits

Procedure

Bayes theorem can be applied to the Gaussian measurement as well:

$$\mathcal{P}(\mu; x_m) \quad = \quad \frac{G(x_m; \mu, \sigma)}{\int d\mu' \; G(x_m; \mu', \sigma)} \cdot \mathcal{P}(\mu)$$

Procedure

Bayes theorem can be applied to the Gaussian measurement as well:

$$\mathcal{P}(\mu; x_m) = \frac{G(x_m; \mu, \sigma)}{\int d\mu' \; G(x_m; \mu', \sigma)} \cdot \mathcal{P}(\mu)$$

Integral in the denominator is equal to 1 only if σ is fixed (!).

With flat "prior distribution" for μ (no earlier constraints) and fixed σ:

$$\mathcal{P}(\mu; x) = G(x; \mu, \sigma)$$

Procedure

Bayes theorem can be applied to the Gaussian measurement as well:

$$\mathcal{P}(\mu; x_m) = \frac{G(x_m; \mu, \sigma)}{\int d\mu' \; G(x_m; \mu', \sigma)} \cdot \mathcal{P}(\mu)$$

Integral in the denominator is equal to 1 only if σ is fixed (!).
With flat "prior distribution" for μ (no earlier constraints) and fixed σ:

$$\mathcal{P}(\mu; x) = G(x; \mu, \sigma)$$

Upper limit on the expected number of events can be then calculated as:

$$\int_0^{\mu_{ul}} d\mu \;\; \mathcal{P}(\mu; x_m) = 1 - \alpha$$

and the numerical result is (again) the same as for Frequentist approach...

General comments

For the two simplest cases, which one could consider, limits obtained from the Bayesian approach are exactly the same as the Frequentist limits.

General comments

For the two simplest cases, which one could consider, limits obtained from the Bayesian approach are exactly the same as the Frequentist limits.

However, this is not the case in the general!
Bayesian limits do not have well defined “confidence levels”,
probability of experimental result being consistent with considered measurement is not defined!

General comments

For the two simplest cases, which one could consider, limits obtained from the Bayesian approach are exactly the same as the Frequentist limits.

However, this is not the case in the general!
Bayesian limits do not have well defined "confidence levels",
probability of experimental result being consistent with considered measurement is not defined!

For complicated measurements (eg. in High Energy Physics) Bayesian approach is much easier to use, as it does not require generation of multiple experiment (MC samples assuming different parameter values) - only the measured distribution is compared with different models.

General comments

For the two simplest cases, which one could consider, limits obtained from the Bayesian approach are exactly the same as the Frequentist limits.

However, this is not the case in the general!
Bayesian limits do not have well defined "confidence levels",
probability of experimental result being consistent with considered measurement is not defined!

For complicated measurements (eg. in High Energy Physics) Bayesian approach is much easier to use, as it does not require generation of multiple experiment (MC samples assuming different parameter values) - only the measured distribution is compared with different models.

Resulting limits are only approximate, they should not be labeled with C.L.

Bayesian limits tend to correspond to higher C.L. than the assumed one...

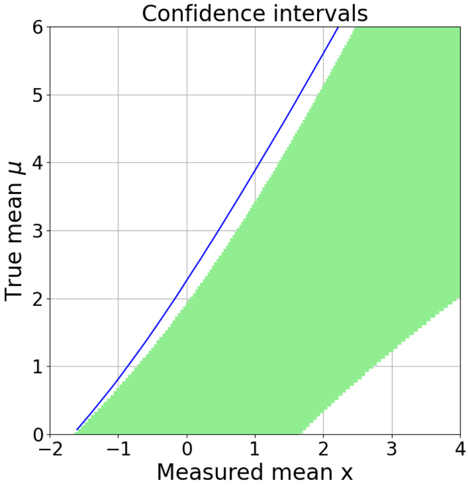
Comparison

07_gauss_bayes.ipynb Open in Colab

Comparison of 95% C.L. upper limits from Frequentist approach (green) with corresponding limits obtained from Bayesian approach (blue) for the example of Gaussian distribution with variable sigma:

$$\sigma^2(\mu) = 1 + 0.1 \cdot \mu^2$$

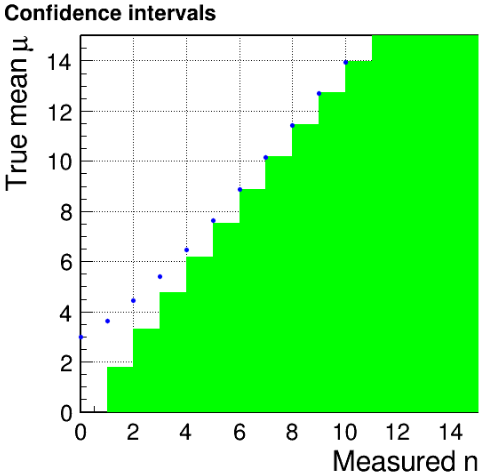
"Coverage" (corresponding measurement interval probability) for the Bayesian limit is higher than the assumed CL !



Comparison

07_poisson_bayes2.ipynb Open in Colab

Comparison of 95% C.L. upper limits from Frequentist approach (green) with corresponding limits obtained from Bayesian approach (blue) for the example of Poisson distribution with background ($\mu_{bg} = 3$).



General comments

One should also stress again that assumption made on prior distribution of the parameter is always arbitrary. Common approach is to use "flat prior", but extracted limits are then sensitive to the parameter choice.

Example: we want to set limits on the leptoquark production, based on the number of observed events. Signal expectation can be written as:

$$\mu_{sig} = \mathcal{L} \cdot A \cdot \sigma_{LQ}$$

where σ_{LQ} is the signal cross section, or as

$\mathcal{L}$ - integrated luminosity

A - acceptance

$$\mu_{sig} = \mathcal{L} \cdot A \cdot k\, \lambda_{LQ}^2$$

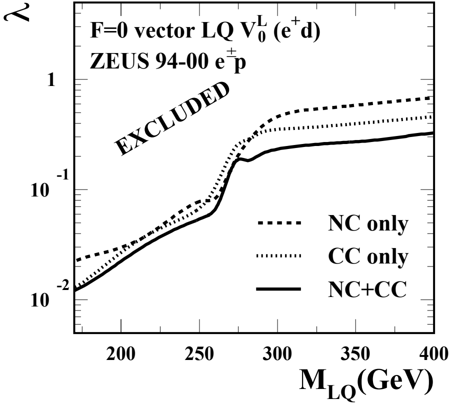
where λ_{LQ} is the leptoquark coupling. We can use Bayesian approach with flat prior to set limits on σ_{LQ} and λ_{LQ}, but they will not be consistent !!!

General comments

There is also arbitrariness in defining limits in multi-parameter space.

Consider leptoquark limits again.

ZEUS collaboration used Bayesian approach to set limits on coupling λ as a function of LQ mass M_{LQ}.
Assuming uniform λ^2 distribution.



ZEUS Collaboration, arXiv:hep-ex/0304008

General comments

There is also arbitrariness in defining limits in multi-parameter space.

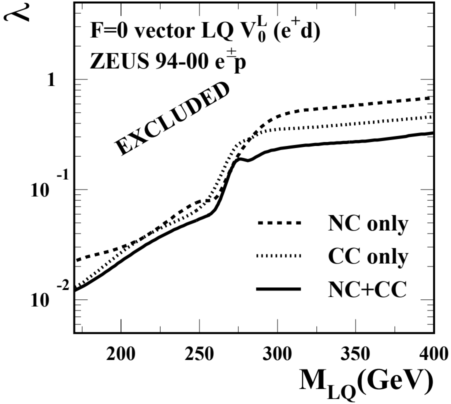
Consider leptoquark limits again.

ZEUS collaboration used Bayesian approach to set limits on coupling λ as a function of LQ mass M_{LQ}.
Assuming uniform λ^2 distribution.

But one could also consider setting limit on M_{LQ} as a function of λ, or limits on effective coupling $\eta = \left(\frac{\lambda}{M}\right)^2$

Limit curves in (M, λ) plane would be different!

Parameter choice is not relevant in frequentist approach!
Each point in parameter space is tested by itself...

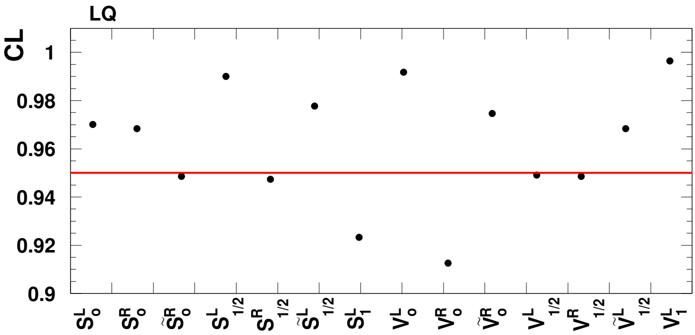


ZEUS Collaboration, arXiv:hep-ex/0304008

General comments

Limits presented in the ZEUS leptoquark publication were obtained with Bayesian approach.
We did not use "confidence level" term in our paper...

Confidence level of the obtained limits was verified for $M_{LQ} \gg \sqrt{s}$ case:



Most of the limits correspond to 95% or higher confidence level.

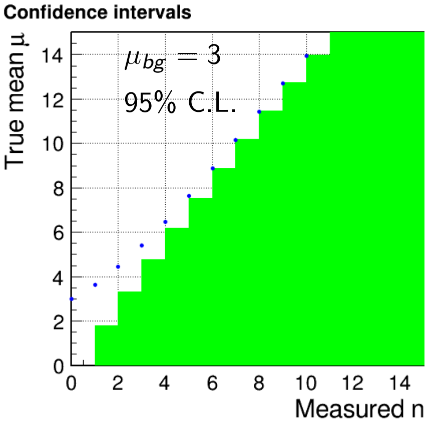
However, two of them are clearly too week...

Parameter Inference (2)

1. Frequentist confidence intervals
2. Bayesian limits
3. Unified approach
4. Homework

Problems

For counting experiment with background, results of both Frequentist and Bayesian approach are not very useful, when **no events are observed**.



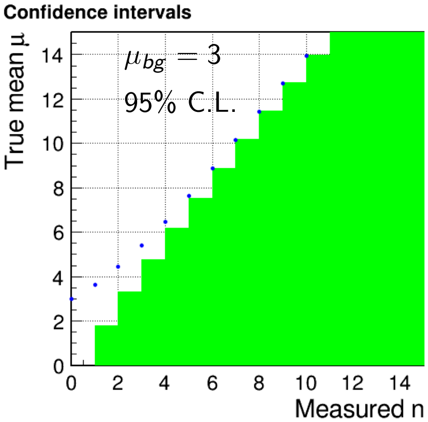
In the Frequentist approach, all values of $\mu > 0$ can be excluded, if background level is high and number of events observed is significantly lower than expected.

Probability of such background fluctuation is small, but finite.

We should not exclude small signals just because background has fluctuated...

Problems

For counting experiment with background, results of both Frequentist and Bayesian approach are not very useful, when **no events are observed**.



In the Bayesian approach, limits for $n_m = 0$ are almost the same as without background, while we would expect them to be stronger.

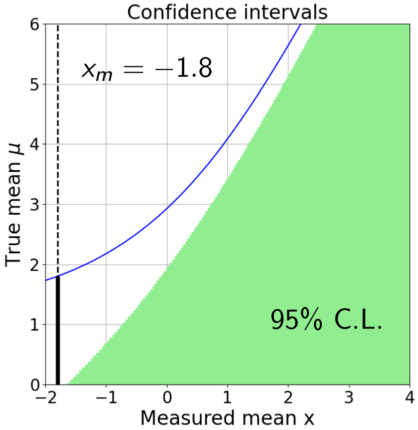
These limits correspond to much higher C.L. than the one assumed

As expected, the two approaches agree for $n_m \gg \mu_{bg}$

Problems

07_gauss_bayes2.ipynb Open in Colab

Similar problem is observed for our example Gaussian distribution, if we assume that true mean is constrained to positive values, $\mu > 0$.



If measured value x_m is below -1.64 then probability of $\mu = 0$ scenario is below 5%.

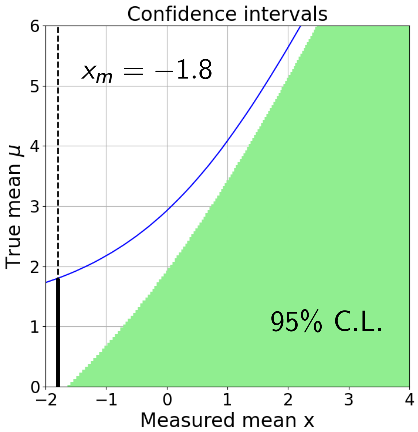
$\Rightarrow$ all values of μ are excluded in Frequentist approach

But we know this has to be fluctuation...

Problems

07_gauss_bayes2.ipynb Open in Colab

Similar problem is observed for our example Gaussian distribution, if we assume that true mean is constrained to positive values, $\mu > 0$.



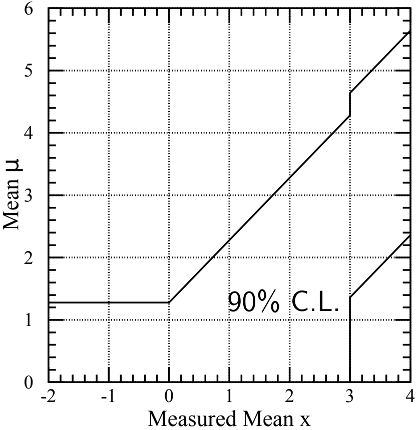
Bayesian limits, on the other hand, seem to be too week again.

Also limits for small positive x_m are significantly weaker...

Problems

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

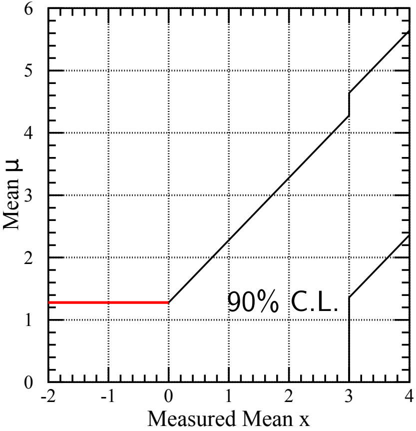
Another problem concerns the way we interpret the results of the Gaussian measurement, if true mean is constrained to positive values, $\mu > 0$.



Problems

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

Another problem concerns the way we interpret the results of the Gaussian measurement, if true mean is constrained to positive values, $\mu > 0$.



Following procedure could be applied:

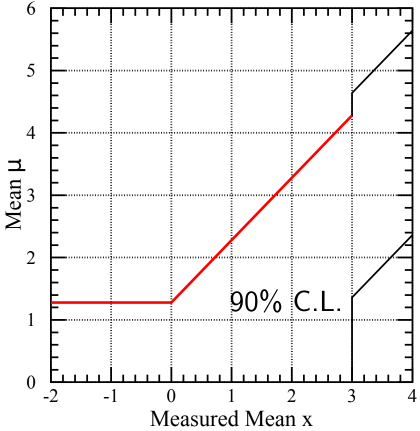
- If measured value x_m is below 0 then we assume it is fluctuation

  $\Rightarrow$ we quote limit for 0.

Problems

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

Another problem concerns the way we interpret the results of the Gaussian measurement, if true mean is constrained to positive values, $\mu > 0$.



Following procedure could be applied:

- If measured value x_m is below 0 then we assume it is fluctuation

  $\Rightarrow$ we quote limit for 0.

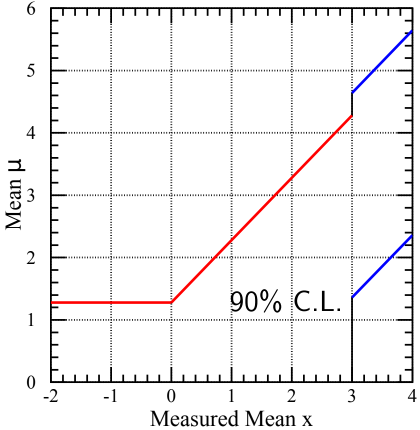
- If measured value is below 3σ

  $\Rightarrow$ we quote 90% CL upper limit

Problems

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

Another problem concerns the way we interpret the results of the Gaussian measurement, if true mean is constrained to positive values, $\mu > 0$.



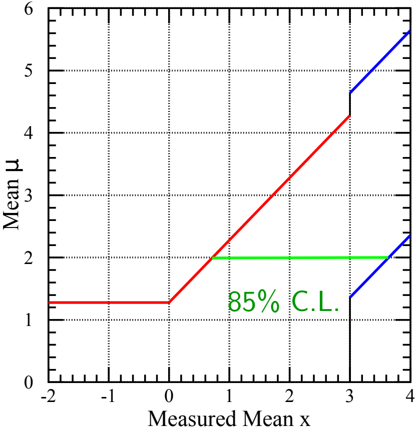
Following procedure could be applied:

- If measured value x_m is below 0 then we assume it is fluctuation
  $\Rightarrow$ we quote limit for 0.
- If measured value is below 3σ
  $\Rightarrow$ we quote 90% CL upper limit
- If measured value is above 3σ
  $\Rightarrow$ we quote 90% CL interval

Problems

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

Another problem concerns the way we interpret the results of the Gaussian measurement, if true mean is constrained to positive values, $\mu > 0$.



Following procedure could be applied:

- If measured value x_m is below 0 then we assume it is fluctuation
  $\Rightarrow$ we quote limit for 0.
- If measured value is below 3σ
  $\Rightarrow$ we quote 90% CL upper limit
- If measured value is above 3σ
  $\Rightarrow$ we quote 90% CL interval

This procedure seems "natural" but results in significant undercoverage! It is only 85% for $1.28 < \mu < 4.28$

Solution

Solution to these problem was proposed in

G.J.Feldman and R.D.Cousins,
A Unified Approach to the Classical Statistical Analysis of Small Signals,
Phys.Rev.D57:3873-3889,1998; arXiv:physics/9711021

New procedure gives proper confidence interval for all possible cases.

Solution

Solution to these problem was proposed in

G.J.Feldman and R.D.Cousins,
A Unified Approach to the Classical Statistical Analysis of Small Signals,
Phys.Rev.D57:3873-3889,1998; arXiv:physics/9711021

New procedure gives proper confidence interval for all possible cases.

We should not use central probability intervals to define limits!

Feldman and Cousin concluded that we should rather select our interval based on the likelihood of given hypothesis for the considered result.

"Best" probability interval for given hypothesis should be defined as the one covering experimental results most consistent with it (with highest likelihood).

Such definition also gives smooth transition between "limit setting" and "interval setting"...

Solution

We still want to start from constructing the probability intervals in random variable x (or n) for given hypothesis μ.

Let $\mu_{best}(x)$ be the parameter value best describing measurement x (maximum likelihood).

How consistent is the considered parameter value μ with our measurement (described by μ_{best}) can be described by likelihood ratio:

$$R(x;\mu) \;=\; \frac{P(x;\mu)}{P(x;\mu_{best}(x))} \;\leq 1$$

Solution

We still want to start from constructing the probability intervals in random variable x (or n) for given hypothesis μ.

Let $\mu_{best}(x)$ be the parameter value best describing measurement x (maximum likelihood).

How consistent is the considered parameter value μ with our measurement (described by μ_{best}) can be described by likelihood ratio:

$$R(x;\mu) \;=\; \frac{P(x;\mu)}{P(x;\mu_{best}(x))} \;\leq 1$$

We can now create the probability interval for x, $[x_1, x_2]$, by selecting values with highest R, up to given CL:

$$\int_{x_1}^{x_2} dx\, P(x;\mu) \;=\; 1-\alpha \quad \text{and} \quad \forall_{x\notin[x_1,x_2]} \; R(x) < R(x_1) = R(x_2)$$

Solution

We still want to start from constructing the probability intervals in random variable x (or n) for given hypothesis μ.

Let $\mu_{best}(x)$ be the parameter value best describing measurement x (maximum likelihood).

How consistent is the considered parameter value μ with our measurement (described by μ_{best}) can be described by likelihood ratio:

$$R(n;\mu) \quad = \quad \frac{P(n;\mu)}{P(n;\mu_{best}(x))} \quad \leq 1$$

We can now create the probability interval for n, $[n_1, n_2]$, by selecting values with highest R, up to given CL:

$$\sum_{n=n_1}^{n_2} P(n;\mu) \;\geq\; 1-\alpha \quad \text{and} \quad \forall_{n\notin[n_1,n_2]} \; R(n) < R(n_1) \;\cap\; R(n) < R(n_2)$$

Example

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

Calculations of 90% CL interval for $\mu = 0.5$, for counting experiment (Poisson variable) in the presence of known mean background $\mu_{bg} = 3.0$

n	$P(n\|\mu)$
0	0.030
1	0.106
2	0.185
3	0.216
4	0.189
5	0.132
6	0.077
7	0.039
8	0.017
9	0.007
10	0.002
11	0.001

Example

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

Calculations of 90% CL interval for $\mu = 0.5$, for counting experiment (Poisson variable) in the presence of known mean background $\mu_{bg} = 3.0$

n	$P(n\|\mu)$
0	0.030
1	0.106
2	0.185
3	0.216
4	0.189
5	0.132
6	0.077
7	0.039
8	0.017
9	0.007
10	0.002
11	0.001

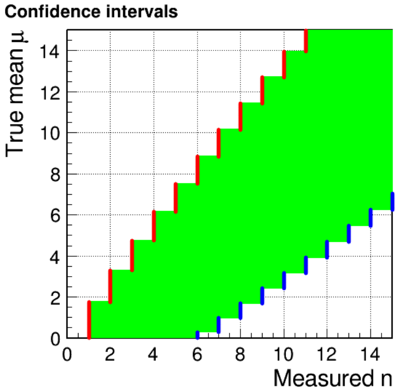
Central probability interval

Example

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

Calculations of 90% CL interval for $\mu = 0.5$, for counting experiment (Poisson variable) in the presence of known mean background $\mu_{bg} = 3.0$

n	$P(n\|\mu)$
0	0.030
1	0.106
2	0.185
3	0.216
4	0.189
5	0.132
6	0.077
7	0.039
8	0.017
9	0.007
10	0.002
11	0.001



Example

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

Calculations of 90% CL interval for $\mu = 0.5$, for counting experiment (Poisson variable) in the presence of known mean background $\mu_{bg} = 3.0$

n	$P(n\|\mu)$	μ_{best}
0	0.030	0.
1	0.106	0.
2	0.185	0.
3	0.216	0.
4	0.189	1.
5	0.132	2.
6	0.077	3.
7	0.039	4.
8	0.017	5.
9	0.007	6.
10	0.002	7.
11	0.001	8.

$$\mu_{best}(n) = \max(n - \mu_{bg},\ 0)$$

Example

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

Calculations of 90% CL interval for $\mu = 0.5$, for counting experiment (Poisson variable) in the presence of known mean background $\mu_{bg} = 3.0$

n	$P(n\|\mu)$	μ_{best}	$P(n\|\mu_{\mathrm{best}})$
0	0.030	0.	0.050
1	0.106	0.	0.149
2	0.185	0.	0.224
3	0.216	0.	0.224
4	0.189	1.	0.195
5	0.132	2.	0.175
6	0.077	3.	0.161
7	0.039	4.	0.149
8	0.017	5.	0.140
9	0.007	6.	0.132
10	0.002	7.	0.125
11	0.001	8.	0.119

Example

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

Calculations of 90% CL interval for $\mu = 0.5$, for counting experiment (Poisson variable) in the presence of known mean background $\mu_{bg} = 3.0$

n	$P(n\|\mu)$	μ_{best}	$P(n\|\mu_{\text{best}})$	R
0	0.030	0.	0.050	0.607
1	0.106	0.	0.149	0.708
2	0.185	0.	0.224	0.826
3	0.216	0.	0.224	0.963
4	0.189	1.	0.195	0.966
5	0.132	2.	0.175	0.753
6	0.077	3.	0.161	0.480
7	0.039	4.	0.149	0.259
8	0.017	5.	0.140	0.121
9	0.007	6.	0.132	0.050
10	0.002	7.	0.125	0.018
11	0.001	8.	0.119	0.006

Example

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

Calculations of 90% CL interval for $\mu = 0.5$, for counting experiment (Poisson variable) in the presence of known mean background $\mu_{bg} = 3.0$

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

Example

G.J.Feldman, R.D.Cousins, arXiv:physics/9711021

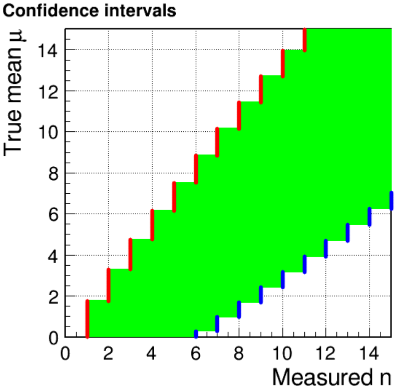
Calculations of 90% CL interval for $\mu = 0.5$, for counting experiment (Poisson variable) in the presence of known mean background $\mu_{bg} = 3.0$

n	$P(n\|\mu)$	μ_{best}	$P(n\|\mu_{\mathrm{best}})$	R	rank
0	0.030	0.	0.050	0.607	6 $\Leftarrow$ included! unlike in central int.
1	0.106	0.	0.149	0.708	5
2	0.185	0.	0.224	0.826	3
3	0.216	0.	0.224	0.963	2
4	0.189	1.	0.195	0.966	1
5	0.132	2.	0.175	0.753	4
6	0.077	3.	0.161	0.480	7
7	0.039	4.	0.149	0.259	
8	0.017	5.	0.140	0.121	
9	0.007	6.	0.132	0.050	
10	0.002	7.	0.125	0.018	
11	0.001	8.	0.119	0.006	

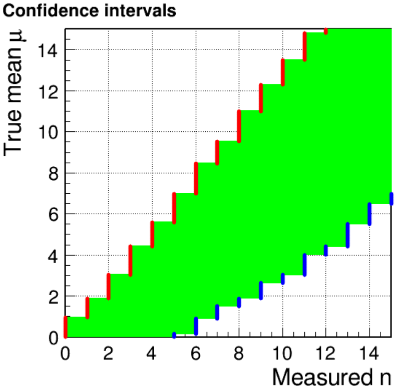
Example 07_poisson_interval2.ipynb 07_poisson_interval3.ipynb Open in Colab

Calculations of 90% CL interval for counting experiment (Poisson variable)
in the presence of known mean background $\mu_{bg} = 3.0$

Central 90% CL intervals



Unified 90% CL intervals

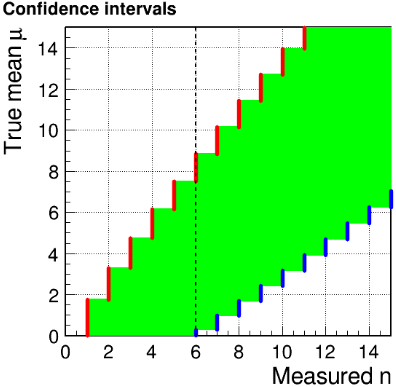


Example

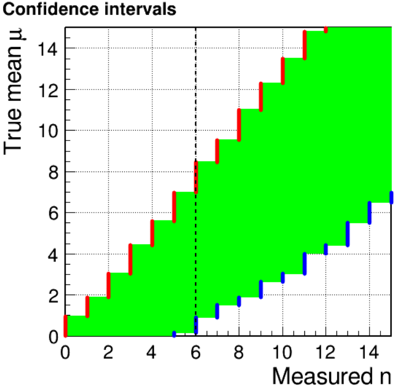
07_poisson_interval2.ipynb 07_poisson_interval3.ipynb Open in Colab

Calculations of 90% CL interval for counting experiment (Poisson variable) in the presence of known mean background $\mu_{bg} = 3.0$

Central 90% CL intervals



Unified 90% CL intervals

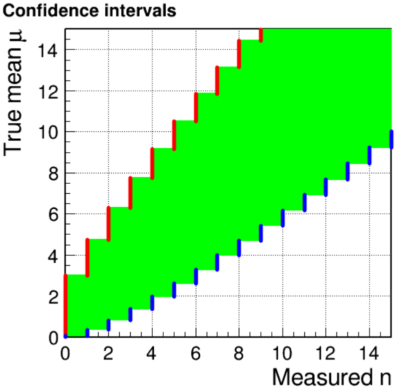


Example

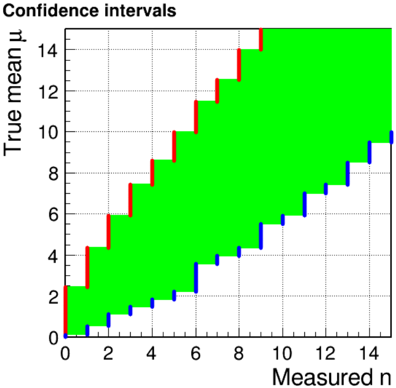
07_poisson_interval.ipynb 07_poisson_interval4.ipynb Open in Colab

Calculations of 90% CL interval for counting experiment (Poisson variable) without background ($\mu_{bg} = 0$)

Central 90% CL intervals



Unified 90% CL intervals

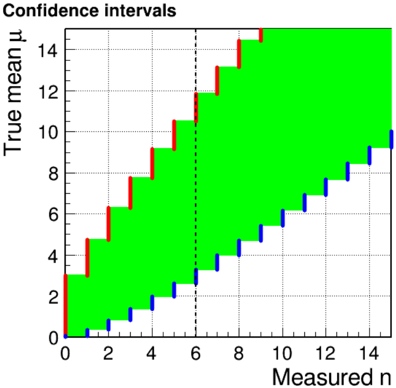


Example

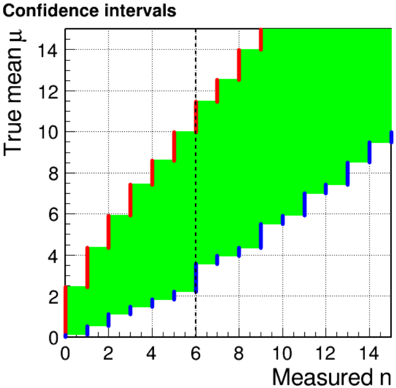
07_poisson_interval.ipynb 07_poisson_interval4.ipynb Open in Colab

Calculations of 90% CL interval for counting experiment (Poisson variable) without background ($\mu_{bg} = 0$)

Central 90% CL intervals



Unified 90% CL intervals



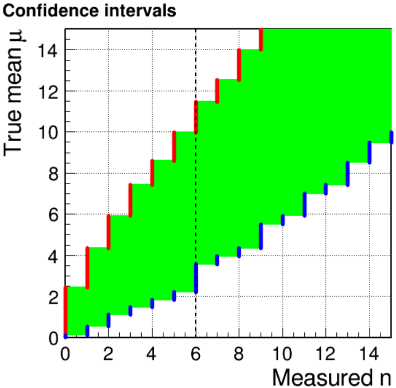
Example

Calculations of 90% CL interval for counting experiment (Poisson variable) without background ($\mu_{bg} = 0$)

	$1-\alpha$ =90%		$1-\alpha$ =95%	
n	μ_1	μ_2	μ_1	μ_2
0	0.00	2.44	0.00	3.09
1	0.11	4.36	0.05	5.14
2	0.53	5.91	0.36	6.72
3	1.10	7.42	0.82	8.25
4	1.47	8.60	1.37	9.76
5	1.84	9.99	1.84	11.26
6	2.21	11.47	2.21	12.75
7	3.56	12.53	2.58	13.81
8	3.96	13.99	2.94	15.29
9	4.36	15.30	4.36	16.77
10	5.50	16.50	4.75	17.82

RPP

Unified 90% CL intervals

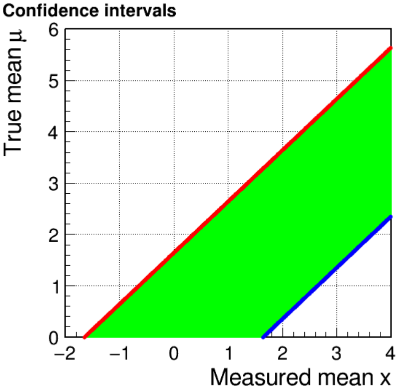


Example

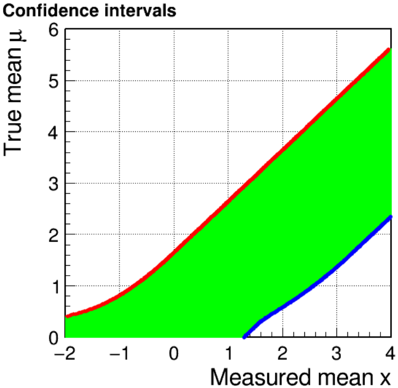
07_gauss_interval.ipynb 07_gauss_interval3.ipynb Open in Colab

Calculations of 90% CL interval for random variable with Gaussian pdf,
true mean constrained to be non-negative, $\mu \geq 0$. $\sigma \equiv 1$

Central 90% CL intervals



Unified 90% CL intervals



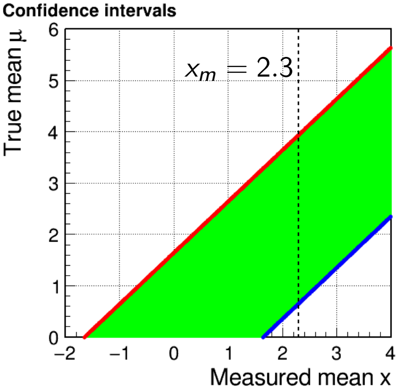
Unified approach

Example

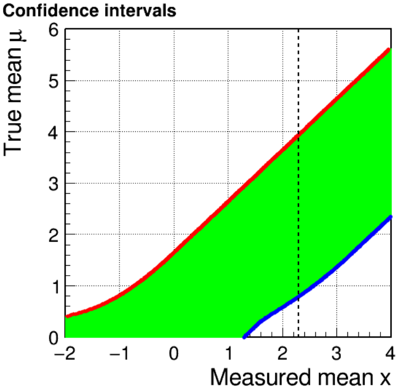
07_gauss_interval.ipynb 07_gauss_interval3.ipynb Open in Colab

Calculations of 90% CL interval for random variable with Gaussian pdf,
true mean constrained to be non-negative, $\mu \geq 0$. $\sigma \equiv 1$

Central 90% CL intervals



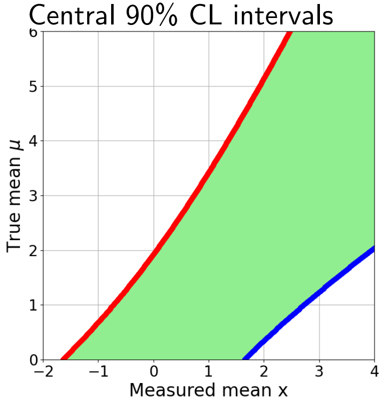
Unified 90% CL intervals

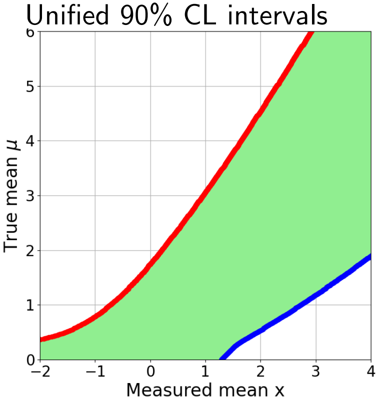


Example

07_gauss_interval2.ipynb 07_gauss_interval4.ipynb Open in Colab

Calculations of 90% CL interval for random variable with Gaussian pdf,
true mean constrained to be non-negative, $\mu \geq 0$. variable σ

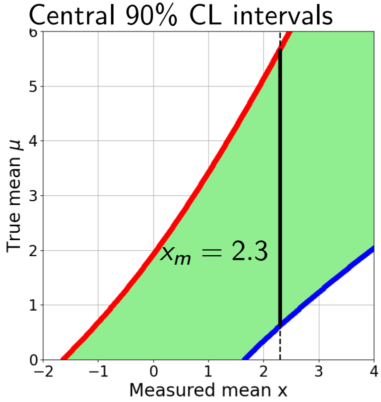


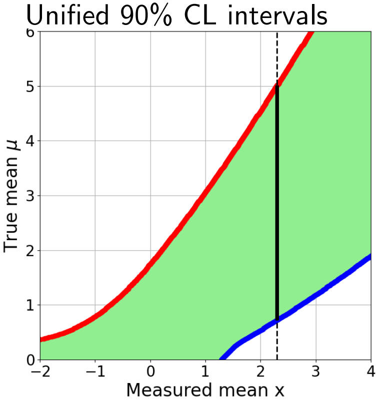


Example

07_gauss_interval2.ipynb 07_gauss_interval4.ipynb Open in Colab

Calculations of 90% CL interval for random variable with Gaussian pdf,
true mean constrained to be non-negative, $\mu \geq 0$. variable σ





Parameter Inference (2)

Homework

Solutions to be uploaded by November 27.

Calorimeter response to particle of given energy E [GeV] can be described by Gamma distribution (see lecture 3) with mean and variance given by:

$$\bar{x} = E + B$$
$$\sigma^2 = 0.25\ \text{GeV} \cdot (E + B)$$

where B is a known background level, $B = 1$ GeV.

Homework

Solutions to be uploaded by November 27.

Calorimeter response to particle of given energy E [GeV] can be described by Gamma distribution (see lecture 3) with mean and variance given by:

$$\begin{aligned} \bar{x} &= E + B \\ \sigma^2 &= 0.25\ \mathrm{GeV} \cdot (E + B) \end{aligned}$$

where B is a known background level, $B = 1$ GeV.

Calculate the **95% CL frequentist upper limit** for the particle energy E, if the measured value $x_m = 3\,\mathrm{GeV}$.